CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE (ČZU)

Multi-Source Quantification of Global Water Cycle Components

by

Mijael Rodrigo Vargas Godoy

A DISSERTATION
SUBMITTED TO THE FACULTY OF ENVIRONMENTAL SCIENCES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

WATER RESOURCES AND ENVIRONMENTAL MODELLING

PRAGUE, CZECHIA
MARCH, 2024

RESEARCH SUPERVISORY COMMITTEE MEMBERS

Dr. Vítězslav Moudrý, Chair, _____
Czech University of Life Sciences Prague, Czechia

Dr. Christoforos Pappas, Opponent, _____
University of Patras, Greece

Dr. Michal Kuráž, Opponent, _____
Czech University of Life Sciences Prague, Czechia

Dr. Peter Greve, Opponent, _____
Helmholtz-Centre Hereon, Germany

Dr. Marek Vach, _____
Czech University of Life Sciences Prague, Czechia

Dr. Michal Jeníček, _____
Charles University, Czechia

Dr. Petr Máca, _____
Czech University of Life Sciences Prague, Czechia

Dr. Ioannis Markonis, Advisor, _____
Czech University of Life Sciences Prague, Czechia

Dr. Simon M. Papalexiou, Co-Advisor, _____
University of Calgary, Canada

Abstract

Evolving in tandem with civilization and technology, our understanding of the water cycle has spurred continuous refinement in quantification methods. However, the uncertainty revealed in the first attempts to unify different data products has hindered this unique opportunity to obtain a robust quantification of the water cycle. Consequently, it remains a challenge to constrain the variability stemming from measurements by various sources, which will help us understand the effect of global warming on water resources in general.

Therefore, this dissertation aimed to improve the understanding of the global water cycle and advance multi-source quantification by: (a) providing an in-depth recapitulation of global water cycle research advancements regarding data sources and quantification methods; (b) enhancing the evaluation of climate reanalyses by proposing a comprehensive framework to study changes in the water cycle; (c) developing a powerful tool that eases the download, exploration, processing, and visualization of an all-around global precipitation data collection; and (d) comprehensively integrating the previous findings and tools to introduce a method for benchmarking multi-source hydroclimatic data fusion based on water cycle budget closure.

The results demonstrated that: (a) the consistency of global water cycle quantification has improved despite significant uncertainty challenges associated with spatiotemporal variability; (b) the signal-to-noise ratio of traditional metrics versus the natural variability is low, and as a consequence, the fluctuations in the water cycle are harder to detect; (c) there is significant heterogeneity among data distribution and formats, deterring robust quantification of global water cycle climatology; and (d) the proposed framework based on water cycle budget closure easily and quickly filters out data combinations that provide implausible results. Therefore, this dissertation improved the understanding of multi-source quantification methods and advanced frameworks for evaluating global and regional water cycle dynamics.

Abstrakt

Evoluujíc spolu s civilizací a technologií, naše chápání vodního cyklu vyvolalo neustálé zdokonalování kvantifikačních metod. Nejistota a nepřesnosti, které se projevily při prvních pokusech o sjednocení různých datových produktů, však bránily získání robustní kvantifikace koloběhu vody. Omezení variability vyplývající z různých zdrojů měření tak zůstává výzvou, stojící v cestě k pochopení vlivu globálního oteplování na vodní zdroje.

Tato disertační práce si kladla za cíl zlepšit porozumění globálnímu koloběhu vody a pokročit v kvantifikaci z více různých zdrojů prostřednictvím: (a) poskytnutí podrobné rekapitulace současného stavu poznání ve výzkumu globálního koloběhu vody, pokud jde o zdroje dat a kvantifikační metody; (b) zlepšení hodnocení druhotných analýz klimatu navržením komplexního rámce pro studium změn v koloběhu vody; (c) vývoje účinného nástroje, který usnadní stahování, průzkum, zpracování a vizualizaci univerzálního globálního sběru údajů o srážkách; (d) komplexní integrace předchozích zjištění a nástrojů pro zavedení metody pro srovnávání propojení hydroklimatických dat z více zdrojů na základě uzavření rozpočtu koloběhu vody.

Výsledky prokázaly, že: (a) navzdory nejistotě spojené s časoprostorovou variabilitou dat, konzistentnost kvantifikace globálního koloběhu vody se zlepšila; (b) poměr signálu k šumu tradičních metrik oproti přirozené variabilitě je nízký, a proto je obtížnější detekovat kolísání koloběhu vody; (c) existuje značná heterogenita mezi distribucí a formáty dat, která odrazuje od důkladné kvantifikace klimatologie globálního koloběhu vody; a za (d) navrhovaný přístup založený na uzavření rozpočtu vodního cyklu snadno a rychle odfiltruje kombinace údajů, které poskytují nevěrohodné výsledky. Tímto tato disertační práce zlepšila porozumění vícezdrojovým kvantifikačním metodám a pokročilým přístupům pro hodnocení globální a regionální dynamiky koloběhu vody.

Preface

This thesis entitled "Multi-source Quantification of Global Water Cycle Components" submitted to the Faculty of Environmental Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Water Resources and Environmental Modelling is an original work by the author.

- Chapter 2 of this thesis has been published as Vargas Godoy MR, Markonis Y, Hanel M, Kyselý J, Papalexiou SM (2021) The Global Water Cycle Budget: A Chronological Review. Surveys in Geophysics 42(5):1075–1107. https://doi.org/10.1007/s10712-021-09652-6

- Chapter 3 of this thesis has been published as Vargas Godoy MR, Markonis Y (2023) Water cycle changes in reanalyses: a complementary framework. Scientific Reports 13(1):1–12. https://doi.org/10.1038/s41598-023-31873-5

- Chapter 4 of this thesis has been published as Vargas Godoy MR, Markonis Y (2023) pRecipe: A global precipitation climatology toolbox and database. Environmental Modelling & Software 165:105711. https://doi.org/10.1016/j.envsoft.2023.105711

- Chapter 5 of this thesis has been published as Vargas Godoy MR, Markonis Y, Rakovec O, Jenicek M, Dutta R, Pradhan RK, Beštáková Z, Kyselý J, Juras R, Papalexiou SM, Hanel M (2024) Water cycle changes in Czechia: a multi-source water budget perspective. Hydrology and Earth System Sciences 28(1):1–19. https://doi.org/10.5194/hess-28-1-2024

The author's scholarly contributions encompass additional investigations, as listed below. Distinct from the chapters of the Ph.D. dissertation, these involve collaborative efforts as co-author participation. Full texts for these studies can be found in Appendix A):

- Shayeghi A, Rahmati Ziveh A, Bakhtar A, Teymoori J, Hanel M, Vargas Godoy MR, Markonis Y, AghaKouchak A. Assessing drought impacts on groundwater and agriculture in Iran using high-resolution precipitation and evapotranspiration products (2024). Journal of Hydrology. In press

- Beštáková Z, Strnad F, Vargas Godoy MR, Singh U, Markonis Y, Hanel M, Máca P, Kyselý J (2023) Changes of the aridity index in Europe from 1950 to 2019. Theoretical and Applied Climatology 151(1):587–601. https://doi.org/10.1007/s00704-022-04266-3

- Nasreen S, Součková M, Vargas Godoy MR, Singh U, Markonis Y, Kumar R, Rakovec O, Hanel M (2022) A 500-year annual runoff reconstruction for 14 selected European catchments. Earth System Science Data 14(9):4035–4056. https://doi.org/10.5194/essd-14-4035-2022

- Pradhan RK, Markonis Y, Vargas Godoy MR, Villalba-Pradas A, Andreadis KM, Nikolopoulos EI, Papalexiou SM, Rahim A, Tapiador FJ, Hanel M (2022) Review of GPM IMERG performance: A global perspective. Remote Sensing of Environment 268:112754. https://doi.org/10.1016/j.rse.2021.112754

M.Sc. Lic. Mijael Rodrigo Vargas Godoy

Acknowledgements

I immensely thank my Ph.D. supervisors, Yannis Markonis and Simon M. Papalexiou. Support and guidance from both of you throughout the project have been invaluable. Yannis, you have been a fantastic primary supervisor. Your doors were always open whenever I ran into a trouble spot or had a question about my research or the meaning of life. Simon, although our interaction was brief, you opened new horizons and perspectives for me. Thank you, Martin Hanel, head of the department, who trusted in me and green-lit many of my ventures.

It has been a fantastic experience working with two research groups across different universities; thanks to everyone from the Czech University of Life Sciences Prague and the University of Calgary for the stimulating discussions and the fun we had. Thank you, Bishal, Fernando, Remy, Roberto, and Santiago, for the beers and helping me stay sane through this journey.

I want to express my eternal gratitude to my family. I would be remiss not to thank my family because they have been a constant source of love, concern, support, and strength since I decided to pursue my dreams abroad. My extended family, especially my grandparents, have aided and encouraged me throughout this endeavor.

Amore, words cannot express how much your unwavering belief in me has meant. Thank you for always being my rock, accompanying me through the highs and lows of this academic journey. Your support and belief in me have been a source of strength and motivation.

Last but not least, I want to thank me. I want to thank me for believing in me. I want to thank me for doing all this hard work. I want to thank me for having no days off. I want to thank me for never quitting.

I dedicate this thesis to Cristobal Vargas Velasquez, my grandfather.

Who influenced countless people in his time here on Earth.

Gracias.

Table of Contents

Abstract iii

Abstrakt iv

Preface v

Acknowledgements vii

Dedication viii

Table of Contents x

List of Figures xiv

List of Tables xvi

Glossary xvii

Epigraph xxii

1 Introduction 1
- 1.1 Research Background . . . 1
  - 1.1.1 An Overview of the Water Cycle . . . 1
  - 1.1.2 Four Data Source Types . . . 6
  - 1.1.3 Multi-Source Quantification . . . 16
  - 1.1.4 Narrowing the Focus . . . 23
  - 1.1.5 Nexus Repercussions . . . 26
- 1.2 Research Objectives . . . 27
- 1.3 Dissertation Layout . . . 29

2 The global water cycle budget: A chronological review 31
- 2.1 Introduction . . . 31
- 2.2 Chronicle . . . 33
  - 2.2.1 Early Days of Hydrology . . . 33
  - 2.2.2 Model Simulations Period . . . 36
  - 2.2.3 Satellite Era . . . 39
  - 2.2.4 Age of Big Data . . . 43
- 2.3 Status Quo et Verisimile Futurum . . . 46
- 2.4 Discussion and Conclusions . . . 54

3 Water Cycle Changes in Reanalyses: A Complementary Framework 58
- 3.1 Introduction . . . 58
- 3.2 The Physical Basis . . . 60
  - 3.2.1 Water Cycle Budget . . . 61

3.2.2 Water Cycle Kinematics . . . 62
3.3 The Precipitation-Evaporation Space . . . 62
3.4 Results . . . 63
3.4.1 Climate Reanalyses . . . 63
3.4.2 Water Cycle Changes . . . 67
3.5 Discussion . . . 70
3.6 Methods . . . 74
3.6.1 Data . . . 74
3.6.2 Benchmarking Reanalyses . . . 74
3.6.3 Thermodynamics of Atmospheric Fluxes . . . 76

4 pRecipe: A global precipitation climatology toolbox and database 78
4.1 Introduction . . . 78
4.2 Methods . . . 80
4.2.1 Data . . . 80
4.2.2 Package Design . . . 82
4.3 Case Study . . . 87
4.4 Conclusions . . . 92

5 Water Cycle Changes in Czechia: A Multi-Source Water Budget Perspective 96
5.1 Introduction . . . 96
5.2 Data and Methods . . . 103
5.2.1 Study Area . . . 103
5.2.2 Data . . . 104
5.2.3 Data Evaluation . . . 108
5.2.4 Data Set Ranking . . . 108
5.2.5 Water Cycle Changes . . . 109
5.3 Results . . . 110
5.3.1 Benchmarking water cycle components . . . 110
5.3.2 Temporal changes in the water cycle . . . 115
5.3.3 Spatial patterns of water cycle changes . . . 116
5.4 Discussion . . . 123
5.5 Conclusions . . . 127

6 Summary & Conclusions 129
6.1 Novel contributions . . . 132
6.2 Recommendations for future research . . . 132

Bibliography 136

A Other Peer Review Publications During the Ph.D. 189

List of Figures

1.1 Water cycle diagram created by the USGS VizLab, in collaboration with the USGS Water Resources Mission Area Web Communications Branch, for the USGS Water Science School. 2
1.2 (a) Location of active stations since as of 2020 with at least 20 year long records. (b) Number of worldwide active stations during the years as per the Global Historical Climatology Network (GHCN) records. 7
1.3 NASA Earth Science Mission Profiles 1997 – 2021. Modified from `https://eospso.nasa.gov/`. 9
1.4 (a) Climate model conceptualization. Each grid cell can be represented by mathematical equations that describe the materials in it and the way energy moves through it. Credit: National Oceanic and Atmospheric Administration. (b) Development of the complexity of the computer models used to simulate Earth's climate. Credit: National Aeronautics and Space Administration. 11
1.5 (a) The principle of data assimilation in climate reanalysis. (b) Representation of four basic strategies for data assimilation, as a function of time. 13
1.6 Schematics of bilinear and conservative regridding. Original cell are in dashed lines and target cell in solid lines. The bilinear interpolation weight corresponding to the node on the lower left is the ratio of the shaded area over the source cell area. The conservative interpolation weight associated with the target cell is the ratio of the shaded area over the source cell area. Modified from Pletzer and Hayek (2019) . . . 21
1.7 Map of the 50 longest rivers of the Czech Republic. Credit: Pavel Hrdlička 24
1.8 Sankey diagram depicting the links between impacts of water cycle changes and Sustainable Development Goals 28

2.1 Schematic structure of a General Circulation Model modified from Bralower and Bice (2012). 36
2.2 Schematic of the Budyko bucket model implemented by Manabe and Bryan (1969). The model represents a single layer soil reservoir with a defined maximum field water capacity of 15 cm from which soil water evaporates at a rate proportional to the remaining water content. 38
2.3 The Water Cycle. Credit: Howard Perlman, United States Geological Survey (USGS). 48
2.4 Probability density distribution of global water cycle fluxes from tables 2.1 and 2.6. The dashed line represents the mean value of each flux. (a) Overland fluxes where P is precipitation, ET is evapotranspiration, and Q is runoff. (b) Over ocean fluxes where P is precipitation and E is evaporation. (c) Same as (a) but only for table 2.6. (d) same as (b) but only for table 2.6. 50
2.5 Chronological estimates of global water cycle fluxes over land in 10^3 km^3/year. P is precipitation, ET is evapotranspiration, and Q is runoff. The years listed correspond to the publication date and do not necessarily reflect the data sets' reference period used by the authors. 50
2.6 Chronological estimates of global water cycle fluxes over oceans in 10^3 km^3/year. P is precipitation and E is evaporation. The years listed correspond to the publication date and do not necessarily reflect the data sets' reference period used by the authors. 51

2.7 Chronological estimates of the evaporative index. Which is defined as the ratio between evapotranspiration and precipitation overland (ET/P). The years listed correspond to the publication date and do not necessarily reflect the data sets' reference period used by the authors. 52

3.1 The global water cycle regime in the precipitation-evaporation space. Vectors represent water cycle changes, where P is precipitation, and E is evaporation. Contours of equal $P - E$ (no change in water cycle storage) are shown as blue dashed lines, and movement across these lines (blue vector) describe changes in water cycle storage. Contours of equal $P + E$ (no change in water cycle intensity) are shown as green dashed lines, and movement across these lines (green vector) describe changes in water cycle intensity. 64
3.2 Benchmarking global spatial weighted average values of reanalysis precipitation compared to GPCP v2.3 as the observation-based reference. (A) Precipitation anomalies annual time series between 1950-2010 (common period between all reanalyses), spread of reanalysis estimates is shown in gray and their mean in white, GPCP v2.3 is shown in turquoise. (B) The 30-year average for the data sets compared, reanalysis estimates are shown in violet and GPCP v2.3 in turquoise. (C) Summary statistics of linear correlation between reanalysis products and GPCP v2.3 annual time series. 65
3.3 Benchmarking global spatial weighted average values of reanalysis temperature compared to HadCRUT5 as the observation-based reference. (A) Temperature anomalies annual time series between 1950-2010 (common period between all reanalyses), spread of reanalysis estimates is shown in gray and their mean in white, HadCRUT5 is shown in red. (B) The 30-year average for the data sets compared, reanalysis estimates are shown in violet (HadCRUT5 is not available). (C) Summary statistics of linear correlation between reanalysis products and HadCRUT5 annual time series. 66
3.4 Global spatial weighted average of annual total precipitation minus evaporation in [mm] as depicted in four reanalysis data sets for their respective available record. Annual values are shown in gray. 30-year moving average values are shown in blue. . 66
3.5 Global spatial weighted average of annual total precipitation plus evaporation and annual mean temperature in four reanalysis data sets for their respective available record. Precipitation plus evaporation in [mm] is shown in green. Temperature in [$^{\circ}C$] is shown in red. 68
3.6 The precipitation-evaporation space graphical framework for the assessment of global water cycle changes. P and E are global total precipitation and evaporation in [$mm/year$]. Contour of $P = E$ is shown as a blue dashed line (stable atmosphere). Contours of equal $P+E$ are shown as green dashed lines (equal water cycle intensity). Changes in $P - E$ and $P + E$ are shown as blue and green vectors correspondingly. Light gray points show the 30-year moving average trajectory, black points mark the labeled 30-year period of interest, and stars mark the position of the average for the full record of each reanalyses. I.e., 1836-2015 average for 20CR v3, 1900-2010 average for ERA-20C, 1950-2020 average for ERA5, and 1948-2020 average for NCEP1. (A) Relative position of reanalyses with respect to each other in the precipitation-evaporation space. (B) Zoomed in panel on the 20CRv3. (C) Zoomed in panel on ERA20C. (D) Zoomed in panel on ERA5. (E) Zoomed in panel on NCEP1. 70

4.1 Different applications of the *plot_line()* function illustrating area-weighted average precipitation over Czechia for all available data sets in *pRecipe*'s database between 1981 and 2020. (A) Monthly precipitation. (B) Annual total precipitation. (C) Annual minimum precipitation. (D) Annual maximum precipitation. (E) Annual median precipitation. (F) Annual average precipitation. 88

4.2 Illustration of the *plot_taylor()* function comparing the entire *pRecipe* database to the observational reference of the CHMI. Upper left, gauge-based data sets as listed in Table 4.1. Upper right, hydrological model forcing data sets as listed in Table 4.4. Bottom left, reanalysis data sets as listed in Table 4.3. Bottom right, satellite-based data sets as listed in Table 4.2. 90
4.3 Illustration of the *plot_taylor()* function comparing selected data sets by season. . . . 91
4.4 Illustration of the *plot_density()* function comparing the empirical distribution of monthly precipitation between the 1981-2000 and 2001-2020 periods for selected data sets. 93
4.5 Illustration of the *plot_map()* function comparing the spatial distribution of the 20-year median precipitation between the 1981-2000 and 2001-2020 periods for selected data sets. 94

5.1 The three drainage basins within Czechia's administrative boundaries (red line). Elbe (light gray shade), Danube (black stripes), and Oder (dark gray points). 104
5.2 Benchmarking spatial weighted average annual water fluxes over Czechia between 1961 and 2020. For consistency and comparability between different water fluxes, annual anomalies were computed using the 1981-2010 average as a reference, the common period among all data sets. The 1981-2010 average and standard deviation are listed at the bottom left of each panel. Linear correlation summary statistics are displayed at the bottom right of each panel. The spread of the estimates being evaluated is shown in gray, and their mean is in white. (a) Precipitation evaluation. CHMI data is shown in blue. (b) Evapotranspiration evaluation. (c) Runoff evaluation. GRDC data is shown in purple. 111
5.3 Spatial weighted average annual water fluxes over Czechia (first row), Labe River (Elbe basin inside Czechia; second row), Morava River (Danube basin inside Czechia; third row), and Odra River (Oder basin inside Czechia; fourth row). Where P is precipitation in blue, E is evapotranspiration in green, Q is runoff in purple, ξ is the residual ($P-E-Q$) in black, and c(ξ) is the cumulative sum of the residual in orange. Left column: TerraClimate (P), TerraClimate (E), and TerraClimate (Q). Middle column: mHM(E-OBS) (P), mHM (E), and mHM (Q). Right column: ERA5-Land (P), ERA5-Land (E), and ERA5-Land (Q). 112
5.4 Empirical distribution of the data set combinations listed on Table 5.7 colored based on their ranking as determined by Equation 5.3. The color gradient goes from higher ranked combinations colored in shades green to lower ranked combinations colored in shades of brown. 113
5.5 Box plots of spatial weighted average annual water fluxes over Czechia, where P is precipitation, E is evapotranspiration, Q is runoff, and $P-E$ is precipitation minus evapotranspiration. Data are divided into two 30-year periods: 1961-1990 (blue) and 1991-2020 (yellow). Note that outliers are present only in the latter period (i.e., 1991-2020) as expected from the recorded severe drought of 2003. 116
5.6 Box plot of spatial weighted average monthly water fluxes over Czechia, where P is precipitation, E is evapotranspiration, Q is runoff, and $P-E$ is precipitation minus evapotranspiration. Data are divided into two 30-year periods: 1961-1990 (blue) and 1991-2020 (yellow). Left column: TerraClimate (P), TerraClimate (E), and TerraClimate (Q). Middle column: mHM(E-OBS) (P), mHM (E), and mHM (Q). Right column: ERA5-Land (P), ERA5-Land (E), and ERA5-Land (Q). 117
5.7 Spatial pattern of changes in median water fluxes over Czechia between two 30-year periods: 1961-1990 and 1991-2020. I.e., the value of each grid cell is equal to the median value of 1991-2020 minus the median value of 1961-1990. P is precipitation, E is evapotranspiration, Q is runoff, $P-E$ is precipitation minus evapotranspiration, and ξ is the residual ($P-E-Q$). Left column: TerraClimate (P), TerraClimate (E), and TerraClimate (Q). Middle column: mHM(E-OBS) (P), mHM (E), and mHM (Q). Right column: ERA5-Land (P), ERA5-Land (E), and ERA5-Land (Q). 119

5.8 TerraClimate spatial pattern of changes in seasonal median water fluxes over Czechia between two 30-year periods: 1961-1990 and 1991-2020. I.e., the value of each grid cell is equal to the seasonal median value of 1991-2020 minus the seasonal median value of 1961-1990. P is precipitation, E is evapotranspiration, and Q is runoff. The seasons are defined as follows: winter as December, January, and February; spring as March, April, and May; summer as June, July, and August; autumn as September, October, and November. 120
5.9 mHM spatial pattern of changes in seasonal median water fluxes over Czechia between two 30-year periods: 1961-1990 and 1991-2020. I.e., the value of each grid cell is equal to the seasonal median value of 1991-2020 minus the seasonal median value of 1961-1990. P is precipitation, E is evapotranspiration, and Q is runoff. The seasons are defined as follows: winter as December, January, and February; spring as March, April, and May; summer as June, July, and August; autumn as September, October, and November. 121
5.10 ERA5-Land spatial pattern of changes in seasonal median water fluxes over Czechia between two 30-year periods: 1961-1990 and 1991-2020. I.e., the value of each grid cell is equal to the seasonal median value of 1991-2020 minus the seasonal median value of 1961-1990. P is precipitation, E is evapotranspiration, and Q is runoff. The seasons are defined as follows: winter as December, January, and February; spring as March, April, and May; summer as June, July, and August; autumn as September, October, and November. 122

List of Tables

1.1 Documentation of 20 most recent precipitation data products. The T62 Gaussian grid has 192 longitude equally spaced and 94 latitude unequally spaced grid points. Data type classification is based on the nomenclature by Sun et al. (2018). 20

2.1 Modified from Baumgartner and Reichel (1972) to exclude incomplete rows. All the fluxes are in 10^3 km^3/year. P_{L} is precipitation overland, ET is evapotranspiration overland, Q is runoff, P_{O} is precipitation over oceans, E is evaporation over oceans, P_{TOT} is total global precipitation, and E_{TOT} is total global evaporation. 35
2.2 Modified from Haddeland et al. (2011). LW is downward longwave radiation flux, $\mathrm{LW}_{\mathrm{net}}$ is net longwave radiation flux, P is precipitation (rain or snow distinguished in the model), q is specific humidity, RR is rainfall rate, S is snowfall rate, SW is downward shortwave radiation flux, SP is surface pressure, T is air temperature, $T_{\max}$ is maximum daily air temperature, $T_{\min}$ is minimum daily air temperature, and W is wind speed. Bulk formula: Bulk transfer coefficients are used when calculating the turbulent heat fluxes. 40
2.3 Compiled from Rodell et al. (2015). P is precipitation, ET is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage. 44
2.4 Modified from Zhang et al. (2016). P is precipitation, ET is evapotranspiration, Q is runoff, and TWS is total water storage. 45
2.5 Modified from Munier and Aires (2018). P is precipitation, ET is evapotranspiration, Q is runoff, and Δ TWS is total water storage change. 46
2.6 All the fluxes are in 10^3 km^3/year. P_{L} is precipitation overland, ET is evapotranspiration overland, Q is runoff, P_{O} is precipitation over oceans, E is evaporation over oceans, P_{TOT} is total global precipitation, and E_{TOT} is total global evaporation. . . 49

3.1 Linear relationship between global spatial weighted average of total atmospheric water fluxes and mean temperature, where P is precipitation, E is evaporation, and T is temperature. Long-term columns report the correlation between the annual values (i.e., $(P \pm E)$ vs. T). Year-to-year columns report the correlation between the annual consecutive differences (i.e., $\delta(P \pm E)$ vs. δT). 67
3.2 Linear relationship between global spatial weighted average of total atmospheric water fluxes and mean temperature, where P is precipitation, E is evaporation, and T is temperature. $P+E$ columns report the correlation ($(P+E)$ vs. T). P columns report the correlation (P vs. T). E columns report the correlation (E vs. T). Slopes are in [%/°C], where the reference for atmospheric flux percentage change and temperature anomaly is their 1981-2010 average. RSE is Residual Standard Error. 69
3.3 Data Set Overview 75

4.1 Gauge-based Products 83
4.2 Satellite-based Products 84
4.3 Reanalysis Products 85
4.4 Hydrological Model Forcing 86

5.1 Compiled from Sahoo et al. (2011). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage. 98
5.2 Compiled from Pan et al. (2012). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage. 99
5.3 Compiled from Rodell et al. (2015). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage. 100
5.4 Compiled from Zhang et al. (2016). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage. 101
5.5 Compiled from Munier and Aires (2018). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage. 102
5.6 Data set description. P is precipitation, E is evapotranspiration, and Q is runoff. . . 106
5.7 Data set ranking as determined by Equation 5.3. P is precipitation, E is evapotranspiration, Q is runoff, $\bar{\xi}$ is the mean residual over 60 years, σ_ξ is the standard deviation of the residual over 60 years, $cor(P-E,Q)$ is the correlation between $P-E$ and Q for the i-th ranked combination, $cor(P,P_o)$ is the correlation between P of the i-th ranked combination and CHMI, and $cor(Q,Q_o)$ is the correlation between Q of the i-th ranked combination and GRDC. 114

Glossary

20C	20th Century
20CR	20th Century Reanalysis
ASR	Arctic System Reanalysis
CAMS	Climate Anomaly Monitoring System
CDAT	Climate Data Analysis Tools
CDO	Climate Data Operators
CHIRPS	Climate Hazards Group Infrared Precipitation with Station Data
CFSR	Climate Forecast System Reanalysis
CHMI	Czech Hydrometeorological Institute
CLM3	Community Land Model version 3
CLSM	Catchment Land Surface Model
CMAP	Climate Prediction Center Merged Analysis of Precipitation
CMIP	Coupled Model Intercomparison Project
CMORPH	Climate Prediction Center Morphing Method
CPC	Climate Prediction Center
CRU TS	University of East Anglia Climatic Research Unit Time-Series
CSI	Critical Success Index
CSR	Center for Space Research at University of Texas
CSRC	Complex Systems Research Center at the University of New Hampshire
CSU	Colorado State University
DMSP	Defense Meteorological Satellite Program
DOE	Department of Energy
ECMWF	European Centre for Medium-Range Weather Forecasts

EM-EARTH	Ensemble Meteorological Dataset for Planet Earth
ERA	European Centre for Medium-Range Weather Forecasts Re-Analysis
ESM	Earth System Model
FAR	False Alarm Rate
FEWS-NET	Famine Early Warning Systems Network
FLDAS	Famine Early Warning Systems Network Land Data Assimilation System
GADM	Database of Global Administrative Areas
GCM	General Circulation Model
GEOS	Goddard Earth Observing System
GEWEX	Global Energy and Water Exchanges
GDAL	Geospatial Data Abstraction Library
GFZ	Deutschen GeoForschungsZentrum
GHCN	Global Historical Climatology Network
GHP	Global Energy and Water Exchanges Hydrometeorology Panel
GLDAS	Global Land Data Assimilation System
GLEAM	Global Land Evaporation Amsterdam Model
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GPM	Global Precipitation Measurement
GRACE	Gravity Recovery and Climate Experiment
GRDC	Global Runoff Data Centre
GRGS	Groupe de Recherche de Géodésie Spatiale
GRIB	General Regularly-distributed Information in Binary form
GRUN	Global Runoff Reconstruction
GSFC	Goddard Space Flight Center
GSMaP	Global Satellite Mapping of Precipitation
GWAVA	Global Water Availability Assessment
H08	Hanasaki 2008
HDF	Hierarchical Data Format

H-TESSEL	Land Surface Hydrology Tiled European Centre for Medium-Range Weather Forecasts Scheme for Surface Exchanges Over Land
IMERG	Integrated Multi-Satellite Retrievals for Global Precipitation Measurement
IPCC	Intergovernmental Panel on Climate Change
ISIMIP	The Inter-Sectoral Impact Model Intercomparison Project
JMA	Japan Meteorological Agency
JPL	Jet Propulsion Laboratories
JRA	Japanese Global Atmospheric Reanalysis Project
JULES	Joint UK Land Environment Simulator
LPJmL	Lund-Potsdam-Jena Managed Land
KGE	Kling-Gupta Efficiency
MacPDM	Macro-Scale Probability-Distributed Moisture
MATSIRO	Minimal Advanced Treatments of Surface Interaction and Runoff
MERRA	Modern-Era Retrospective Analysis for Research and Applications
mHM	Mesoscale Hydrologic Model
MPI-HM	Max Planck Institute - Hydrology Model
MOD16	Moderate Resolution Imaging Spectroradiometer Global Evapotranspiration Project
MODIS	Moderate Resolution Imaging Spectroradiometer
MSWEP	Multi-Source Weighted-Ensemble Precipitation
MSWX	Multi-Source Weather
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NetCDF	Network Common Data Form
NOAA	National Oceanic and Atmospheric Administration
NOAH	National Centers for Environmental Prediction; Oregon State University; Air Force; Hydrology Lab
NRCS	Natural Resources Conservation Service
NRL	Naval Research Laboratory

NSE	Nash–Sutcliffe Model Efficiency
NTSG	Numerical Terradynamic Simulation Group
Orchidee	Organising Carbon and Hydrology in Dynamic Ecosystems
PDSI	Palmer Drought Severity Index
PERSIANN	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
PGF	Princeton Global Forcing
POD	Probability of Detection
PREC/L	Precipitation Reconstruction Over Land
RMSE	Root Mean Square Error
SDG	Sustainable Development Goal
SPEI	Standardized Precipitation Evapotranspiration Index
SPI	Standard Precipitation Index
SRB-CFSR-SEBS	Surface Radiation Budget - Climate Forecast System Reanalysis - Surface Energy Balance System
SRB-CFSR-PM	Surface Radiation Budget - Climate Forecast System Reanalysis - Penman-Monteith
SRB-CFSR-PT	Surface Radiation Budget - Climate Forecast System Reanalysis - Priestly-Taylor
SRB-PGF-PM	Surface Radiation Budget - Princeton Global Forcing - Penman-Monteith
SSM/I	Special Sensor Microwave Imager
SSMIS	Special Sensor Microwave Imager Sounder
SWOT	Surface Water and Ocean Topography
TIROS	Television InfraRed Observation Satellite
TMPA	Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis
TRMM	Tropical Rainfall Measuring Mission
UDel	University of Delaware
USGS	United States Geological Survey
VIC	Variable Infiltration Capacity
WaterGAP	Water Global Assessment and Prognosis

WCRP	World Climate Research Program
WSAG	Water Systems Analysis Group
WSL	Windows Subsystem for Linux

Epigraph

El que lee mucho y anda mucho, ve mucho y sabe mucho.

- Miguel de Cervantes, *Don Quixote*

Chapter 1

Introduction

1.1 Research Background

1.1.1 An Overview of the Water Cycle

The water cycle is an intricate and dynamic system involving the perpetual movement of water through various physical, chemical, and biological processes, where water transitions between different states and traverses diverse landscapes (Figure 1.1). These processes include, among others, condensation, evaporation, groundwater flow, infiltration, percolation, plant uptake, precipitation, runoff, sublimation, transpiration, and water vapor transport (Allan et al., 2020). The water cycle is driven by energy from the sun, which causes water evaporation from surface-water bodies (e.g., lakes, rivers, and oceans), soils, and plants (i.e., transpiration). Evaporation is one of the most significant components of the water cycle, transporting water from the Earth's surface to the atmosphere and contributing to about 90% of atmospheric moisture. Condensation is the opposite physical process, where water vapor cools and condenses back into the liquid phase, forming clouds. When the water droplets in the clouds become large enough, under the appropriate cloud microphysics conditions, they are released as precipitation in the form of rain, freezing rain, sleet, snow, or hail. Precipitation is another critical component of the water cycle, complementary to evaporation, as it delivers atmospheric water to the Earth's surface. On land, precipitation can infiltrate the soil, replenish groundwater reservoirs, or run off into streams, rivers, and lakes. The water that runs off can eventually return to the oceans, where the cycle begins anew. In addition to total water storage, the above components represent the water cycle's major inputs, outputs, and storage. To some extent, the rest of the water cycle processes are encompassed in these four components (Bengtsson, 2010).

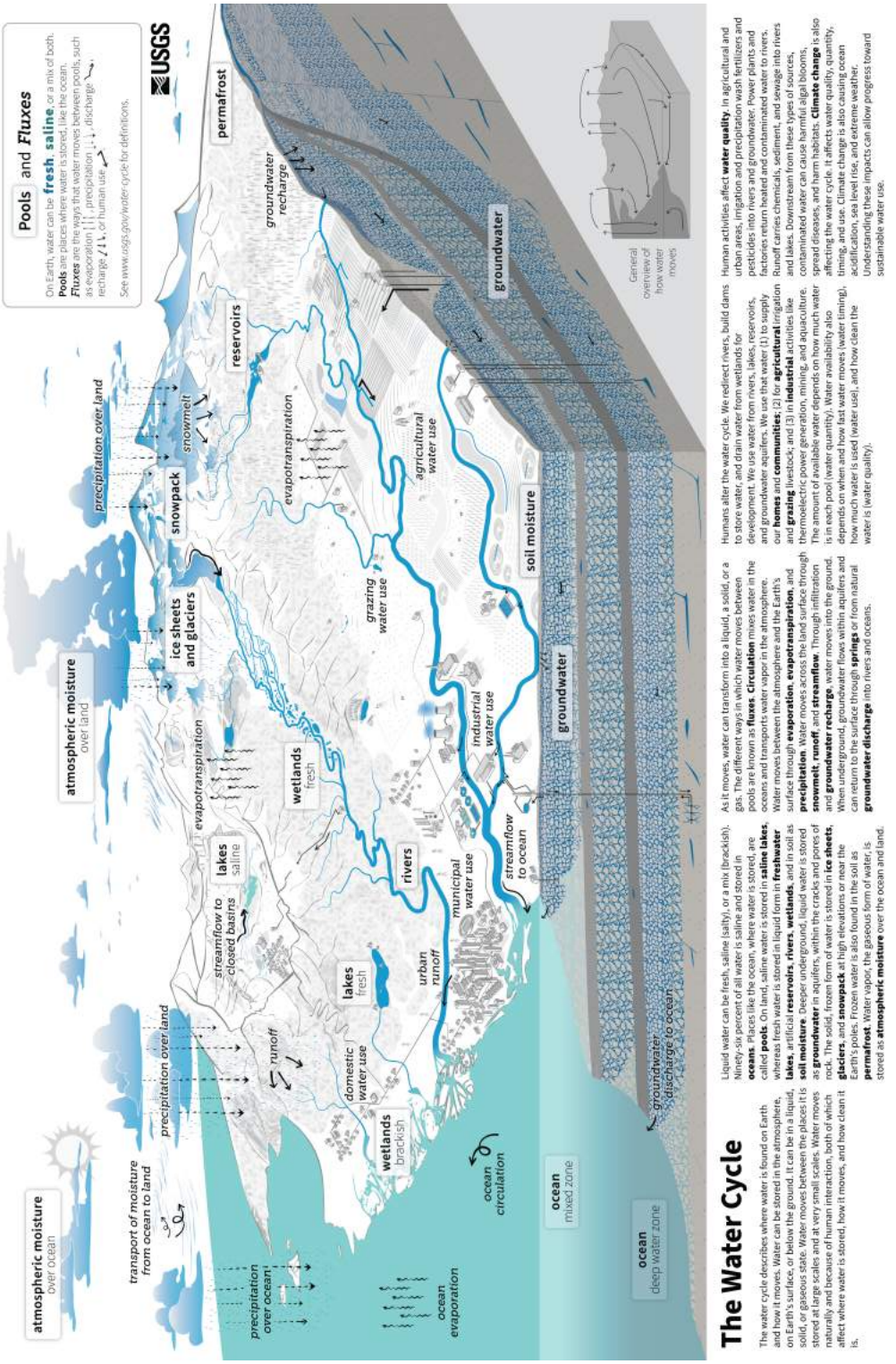


Figure 1.1: Water cycle diagram created by the USGS VizLab, in collaboration with the USGS Water Resources Mission Area Web Communications Branch, for the USGS Water Science School.

The water cycle budget, which keeps track of the inflows and outflows in a given area, is formulated as follows in its expanded form (Scanlon et al., 2002):

$$P + Q_{\text{in}}^{\text{sw}} + Q_{\text{in}}^{\text{gw}} = E^{\text{sw}} + E^{\text{gw}} + E^{\text{uz}} + \Delta S^{\text{sw}} + \Delta S^{\text{snow}} + \Delta S^{\text{uz}} + \Delta S^{\text{gw}} + Q_{\text{out}}^{\text{gw}} + \text{RO} + Q^{\text{bf}} \quad (1.1)$$

where P is precipitation, $Q_{\text{in}}^{\text{sw}}$ is surface water inflow, $Q_{\text{in}}^{\text{gw}}$ is groundwater inflow, E^{sw} is evaporation from surface water, E^{gw} is evaporation from groundwater, E^{uz} is evaporation from the unsaturated zone, ΔS^{sw} is change in surface water storage, ΔS^{snow} is change in snow water storage, ΔS^{uz} is change in unsaturated zone water storage, ΔS^{gw} is change in groundwater storage, $Q_{\text{out}}^{\text{gw}}$ is groundwater outflow and withdrawal by pumping, RO is surface runoff, and Q^{bf} is the base flow. However, for most practical applications, even in small watersheds, it can be expressed in terms of its four major components:

$$\Delta\text{TWS} = P - E - Q \quad (1.2)$$

where ΔTWS is the change in total water storage, E is total evaporation (from soils, surface-water bodies, and plants), and Q is runoff. Inadvertently, aggregating global water cycle components to the most dominant ones, also aggregate their underlying uncertainties.

In addition to the intricate natural processes of the water cycle, anthropogenic processes affect the availability and quality of water resources (Abbott et al., 2019). These processes fall under one of the following categories: water use, land cover change, and climate change. Water use could be divided into green (soil moisture use), blue (water withdrawals), and grey (pollution assimilation) (Rockström et al., 2012; Hoekstra and Mekonnen, 2012). Land cover changes (e.g., agriculture, deforestation, and wetland destruction) alter evapotranspiration, groundwater recharge, river discharge, and precipitation at continental scales (Ellis et al., 2010; Falkenmark et al., 2019). Lastly, climate change disrupts water flow and storage patterns locally and globally (Durack et al., 2012; Haddeland et al., 2014; Huang et al., 2016). This last alteration of the water cycle is the most submerged under a mist of uncertainty as we have yet to fully decouple the natural and anthropogenic forced responses of the water cycle. Thermodynamics, Clausius–Clapeyron scaling, in particular, determines the dependence of vapor pressure at a discontinuous phase transition between two phases of matter of a single constituent (water), i.e., the relationship between atmospheric water vapor and temperature (Clapeyron, 1834). However, the Earth's energy balance governs the water cycle's atmospheric fluxes and constrains the hydrological sensitivity, defined by the increase in mean precipitation (or evaporation) for a given change in mean temperature (Allan et al., 2020).

The lack of a comprehensive understanding of water cycle changes resulted in two main hypotheses: the "changing character of precipitation" (Trenberth et al., 2003) and the "dry gets drier, wet gets wetter" (Held and Soden, 2006). The former addresses precipitation intensity, duration, frequency, and phase instead of total amounts only. Extreme precipitation events, leading to floods and droughts, exert significant environmental and societal impacts, necessitating an enhanced understanding of precipitation characteristics (Gimeno et al., 2022). The increase in moisture content in the atmosphere regulated by the Clausius–Clapeyron equation is expected to rise much faster than the total precipitation amount. In addition, the precipitation rate far surpasses the rate at which moisture is replenished through surface evaporation, i.e., most precipitation originates from the moisture preexisting in the atmosphere when the event initiates (Trenberth, 1998). Therefore, heavy or extreme rainfall will become more frequent, while light or moderate precipitation will decline. In other words, the increase in global mean precipitation will be unevenly distributed in precipitation events. The latter presents a mathematical derivation that starts with the Clausius–Clapeyron equation:

$$\frac{d \ln e_s}{dT} = \frac{L}{RT^2} = \alpha\left(T\right) \tag{1.3}$$

where e_s is the saturation vapor pressure, T is temperature, L is the latent heat of vaporization, and R is the gas constant. Through a series of boundary conditions, assumptions, and mathematical approximations (for details, see Held and Soden, 2006), the equation ends in the following form:

$$\delta\left(P - E\right) = \alpha \delta T \left(P - E\right) \tag{1.4}$$

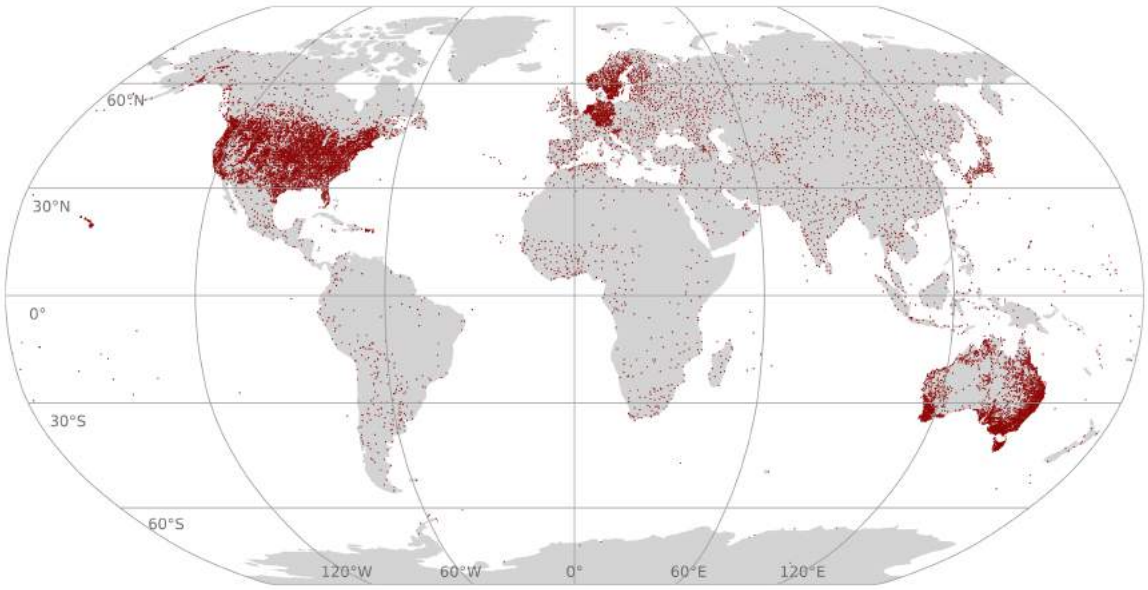
where $\alpha \approx 0.07[K^{-1}]$. The mathematical framework presented implies a reduction in the mass exchange between the boundary layer and mid-troposphere. Given that a significant portion of this exchange takes place during moist convection in the Tropics, there must be a decrease in convective mass flux. Under the assumption that the lower tropospheric relative humidity and the flow remain unchanged, the poleward vapor transport and the pattern of precipitation minus evaporation will increase proportionally to the lower tropospheric water vapor. It is worth mentioning that this hypothesis is probably the most known and the most commonly misunderstood because it is applied or evaluated without consideration for all the assumptions and boundary conditions that constrain its validity. Nevertheless, both hypotheses are presently under vigorous debate (Seager et al., 2010; O'Gorman and Muller, 2010; Greve et al., 2014; Roderick et al., 2014; Byrne and O'Gorman, 2015; Kumar et al., 2015; Salzmann, 2016; Skliris et al., 2016; Wang et al., 2017; Markonis et al., 2019).

Undeniably, our understanding of the global water cycle has evolved over the years thanks to humankind's relentless pursuit of technological advancements and innovative solutions. Needless to say, throughout history, our methods for quantifying hydro-meteorological variables have transformed in tandem with the exponential growth of technology from solely human observation to the integration of cutting-edge machine learning techniques. Pursuing precise quantification of the global water cycle led to the establishment of the Global Energy and Water Exchanges (GEWEX) project. Originally named the Global Energy and Water Cycle Experiment, this initiative commenced in 1990 with the primary focus of investigating Earth's water and energy cycles (Chahine, 1992a). GEWEX created a platform for global collaboration in research, facilitating engagement through various panels, meetings, and projects. GEWEX oversees eight continental-scale experiments concerning the enhancements of data sets and modeling: GEWEX Americas Prediction Project (GAPP; Lawford 1999), Baltic Sea Experiment (BALTEX; Raschke et al. 1998, 2001), GEWEX Asian Monsoon Experiment (GAME; Yasunari 1994), Large Scale Biosphere Atmosphere Experiment in Amazonia (LBA; Marengo 2005), Mackenzie GEWEX Study (MAGS; Stewart et al. 1998), La Plata Basin (LPB; Cavalcanti et al. 2015), The African Monsoon Multidisciplinary Analysis (AMMA; Redelsperger et al. 2006), and Murray-Darling Basin (MDB; Evans and McCabe 2010). Beyond logistical and political considerations, the selection of these sites aimed to gather data from diverse climate regimes, providing a representative assessment of the global water cycle. The collaborative endeavors of international teams contributed to an enhanced understanding of regional water balance and feedback processes. The data generated from these continental-scale experiments are accessible to the public, indirectly contributing to establishing a scientific framework for quantifying the global water cycle and achieving a budget closure with a 10% non-closure tolerance. Notwithstanding these momentous advancements, data sources' diverse performance levels and associated uncertainties restrict their suitability for global-scale analyses (Hegerl et al., 2015). Even though all data estimates, regardless of their source, inherently possess a degree of uncertainty (Steen, 1990), the quest to minimize this uncertainty to the greatest extent possible and acquire higher quality and more accurate data for water cycle assessment has mutually driven technological advancements (e.g., satellite instruments; Hildebrand et al., 2003; Levizzani and Cattani, 2019). At present, we could say that the available data products fall under one of four main categories: ground station observations, satellite remote sensing, hydrological model simulations, and reanalysis outputs. Note that while some products could be categorized easily as they exclusively rely on a unique data source type, usually the categorization is based on the primary source or bulk input, as in reality, it is typical for a data product to have multiple inputs from different source types.

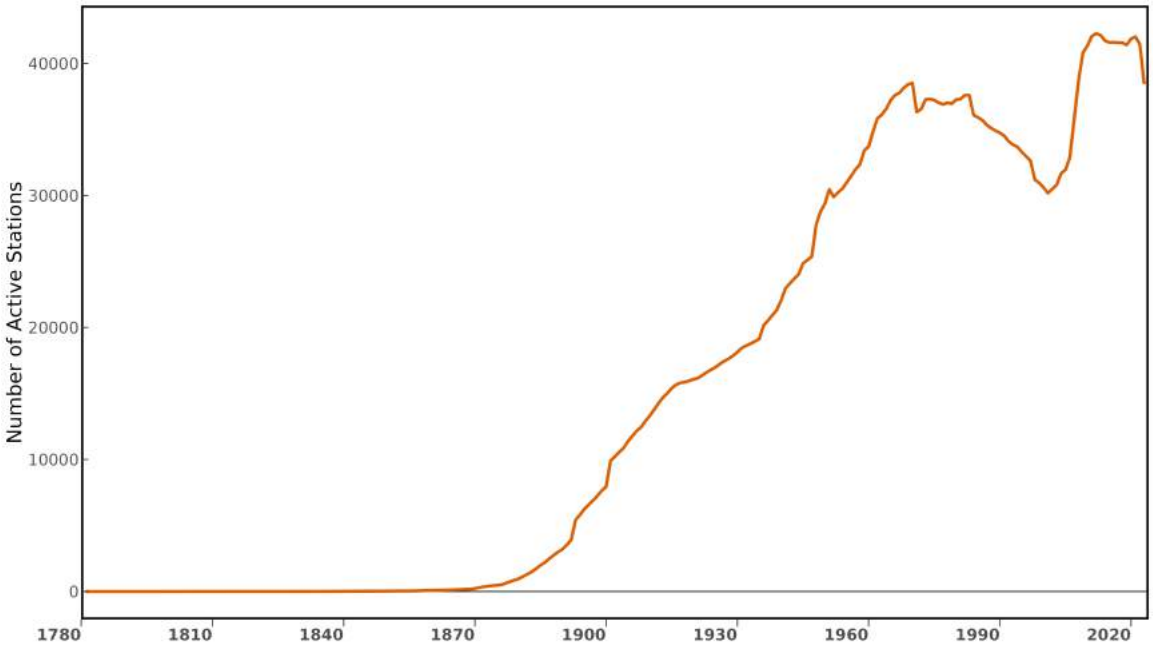
1.1.2 Four Data Source Types

As better and more accurate instruments became available, ground-based observations improved their detection capabilities. The development of different instruments like rain gauges (standard, tipping bucket, and weighing), lysimeters (percolation, weighing, and wick), atmometers, flumes, and sounders led to different data products. Even nowadays, despite well-known spatial heterogeneity and variability, ground-based observations are regarded as the closest measurements to the actual values and are operationally used for calibration, evaluation, and assimilation. Nevertheless, the distribution of active stations is quite heterogeneous worldwide (Figure 1.2(a)), extremely geopolitically dependent (Kibler et al., 2014), and their number has been declining for the past decades (Figure 1.2(b)). Deploying dense monitoring networks to observe hydroclimate spatial heterogeneity accurately implies high operational and maintenance costs (Saltikoff et al., 2017). Consequently, it is unsurprising that in many developing countries, ground observational records, if available, tend to have multiple temporal discontinuities or non-standardized data quality check protocols (Walker et al., 2016). In other words, we still lack a comprehensive global network.

For example, the Global Precipitation Climatology Centre (GPCC) offers data from 1891 to the present based on quality-controlled data from 67 200 stations worldwide that feature record durations of 10 years or longer on a regular grid with a spatial resolution of 0.25 degrees (Schneider et al., 2011). GPCC is derived from the most extensive gauge network currently available, but it represents only about 1% of the Earth's surface (assuming a 5 km non-overlapping radius per gauge; Kidd and Huffman, 2011). As evidenced above, the distribution of stations worldwide is heterogeneous, which denotes that interpolation methods are relied on to generate regular global grids. Moreover, a variable number of stations per grid over time can be a significant source of inhomogeneity, inconsistency, and uncertainty (Herrera et al., 2019). While a kindred initiative for evaporation exists (FLUXNET; Pastorello et al., 2020), the network's instrumentation does not directly measure evaporation, but rather, it employs the eddy covariance technique to measure carbon, water, and energy cycling between the biosphere and atmosphere. As a result, evaporation is more commonly derived from atmospheric moisture and precipitation measurements. Runoff has an additional layer of complexity because the Global Runoff Data Centre (GRDC; Fekete et al., 2002), provides ground-based data for river discharge only, which is only a fraction of total water outflow from a catchment. To our knowledge, total water storage, the last major water cycle component, has no comparable database. GPCC, FLUXNET, and GRDC constitute the three most extensive collections of ground-based observations for water cycle components.



(a)



(b)

Figure 1.2: (a) Location of active stations since as of 2020 with at least 20 year long records. (b) Number of worldwide active stations during the years as per the Global Historical Climatology Network (GHCN) records.

Satellite remote sensing data complemented the traditional gauge-based measurements and offered unprecedented coverage over previously inaccessible or ungauged regions. Notwithstanding, ground observations remain crucial for satellite data calibration, evaluation, and, in some cases, post-processing refinement. Several satellite missions have been launched into orbit since 1960, starting with the Television Infrared Observation Satellite (TIROS-1 or TIROS-A). Among the most notable are the National Aeronautics and Space Administration's (NASA) Earth Observing System (EOS) missions (Figure 1.3). Satellites are characterized by their sensor type and orbit. Satellite sensors could be active or passive, and it is no longer uncommon for both types to be onboard simultaneously. Active sensors provide their energy source to illuminate the objects they observe. In contrast, passive sensors detect energy emitted or reflected from the environment. Satellite orbits could either be geosynchronous (GEO) or polar, yet many of the satellites in the EOS missions have a nearly polar orbit. Polar-orbit satellites move around the Earth in a Sun-synchronous orbit, so the overpass occurs at the same local time every day, taking around 100 minutes to complete an orbit. However, it soon became apparent that satellite instruments have heterogeneous performances across the globe (Maggioni et al., 2016), limiting their applicability. Moreover, satellite remote sensing data records are, to date, too short to assess long-term water cycle changes.

Understanding that satellite design (orbit and sensor type) limits remote sensing observations is fundamental. The satellite's orbit delimits its spatiotemporal resolution and coverage. Generally, a satellite with high spatial resolution comes with coarse temporal resolution and vice-versa, and high spatiotemporal resolution comes with limited coverage. Satellite remote sensing estimates have been shown to vary considerably depending on the sensor type (i.e., active or passive; Fekete et al., 2004). Aware of such systematic limitations, recent missions rely on a constellation of multiple satellites rather than a single satellite, even if it carries both passive and active sensors. Similarly to ground-based station data, satellite-based flagship products exist for each water cycle component: the Global Precipitation Measurement (GPM) for precipitation (Huffman et al., 2015), the Moderate Resolution Imaging Spectroradiometer (MODIS) for evapotranspiration (Mu et al., 2011), the Surface Water and Ocean Topography (SWOT) for runoff (Durand et al., 2010), and the Gravity Recovery and Climate Experiment (GRACE) for total water storage (Tapley et al., 2004). Because the SWOT mission was launched in 2022, its data record is currently too short for usage. However, runoff could be derived from other satellite remote sensing data sets, for instance, Tropical Rainfall Measuring Mission (TRMM) precipitation (Huffman et al., 2007), and MODIS land cover (Friedl et al., 2002) using the Natural Resources Conservation Service (NRCS) runoff curve number method (Cronshey, 1986; Burges et al., 1998).

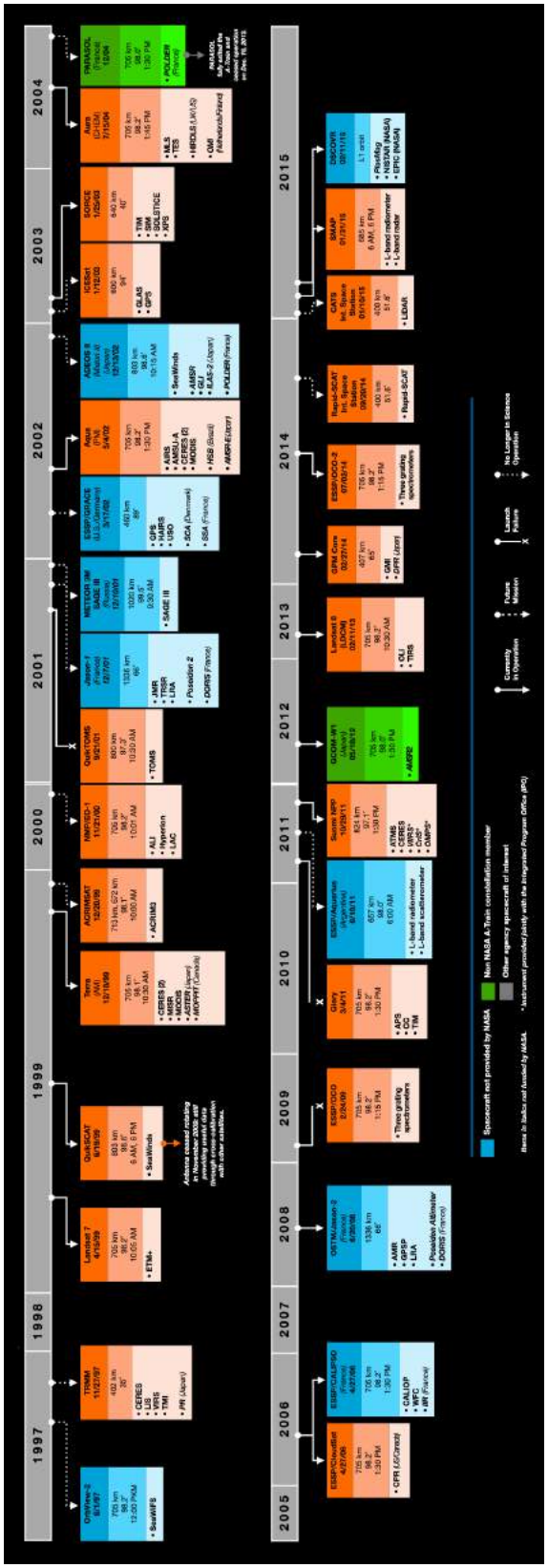
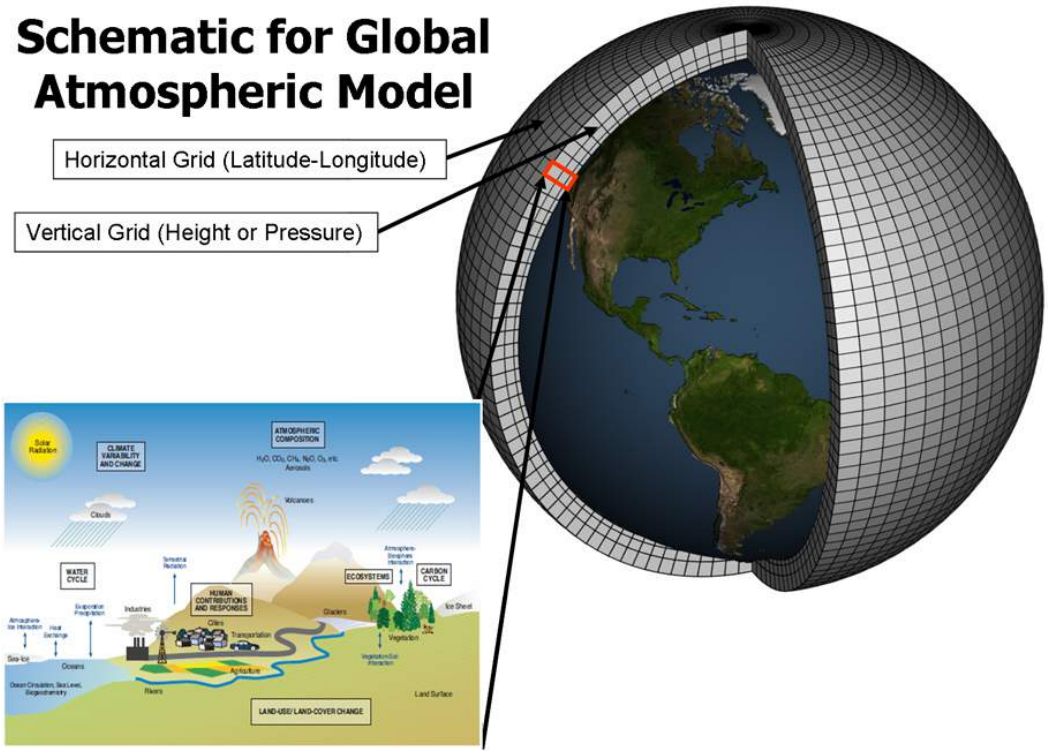


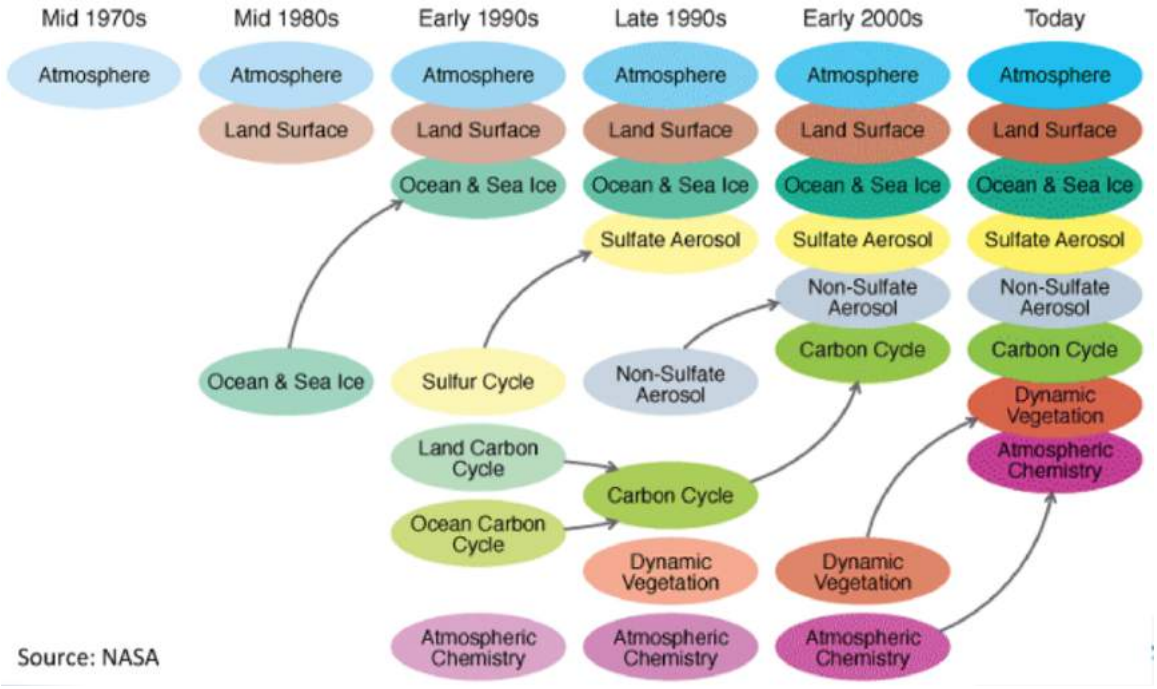
Figure 1.3: NASA Earth Science Mission Profiles 1997 – 2021. Modified from `https://eospso.nasa.gov/`.

Almost contemporary to satellite remote sensing flourishment, general circulation models were developed as powerful tools that comprehensively explore the intricate processes of the global water cycle (Figure 1.4(a)). These models, rooted in fundamental principles of atmospheric physics, fluid dynamics, and thermodynamics, opened the doors to simulate and study the complex interactions between the atmospheric, oceanic, and terrestrial components that shape the global water cycle (McGuffie and Henderson-Sellers, 2001). Models differ in the details of the description of processes, parameter estimation approaches, time scales, and spatial resolution of input data and simulations (Haddeland et al., 2011). The stand-alone models are usually applied at the basin or catchment scale and tend to have many parameters that need to be calibrated or estimated regionally (e.g., Santhi et al., 2006). On the other hand, global hydrological models usually have few parameters and are calibrated at eco-region, climatic-region, or large river basin scales (Vörösmarty et al., 1989). Global model simulations are becoming more complex and resolute as more functionality is added and the availability of finer global spatial data sets increases (Sood and Smakhtin, 2015). A shared concern for model simulations across different scales is that they are calibrated to existing observational data sets and retain some residual error. In practice, said error is typically unknown and often ignored, implicitly trusting simulated responses as deterministic quantities that might not even represent the observed ones (Farmer and Vogel, 2016).

The importance of model simulations is not only in their ability to reproduce present climate conditions but also in their capacity to project past and future scenarios, aiding in anticipating potential shifts in water availability and distribution. As previously mentioned, even though the water cycle's response to global warming is under vigorous debate, it is non-debatable that the intensification theory draws its foundations from model experiments involving the doubling of carbon dioxide (CO_2) concentrations (Manabe and Wetherald, 1975). While the Earth's energy balance influences alterations in global mean precipitation, variations at regional to local scales arise from the complex interactions among factors such as CO_2 levels, aerosols, land use changes, and human water consumption (Allan et al., 2020). Despite the exponential growth in computing power efficiency and growing complexity of models to try to capture local dynamical processes (Figure 1.4(b)), many of them, like radiative transfer, convection initiation, hydrometeor phase change, and cloud microphysics that occur between the sub-kilometer scale and the microscale (i.e., nine orders of magnitude less than current model resolutions) are parameterized, as they cannot be resolved at the model resolution. As a result, we observe artifacts like a correlation between an increase in precipitation extremes and an increase in model resolution and anti-correlation between precipitation extremes and changes in light-moderate precipitation (Thackeray et al., 2018).



(a)



(b)

Figure 1.4: (a) Climate model conceptualization. Each grid cell can be represented by mathematical equations that describe the materials in it and the way energy moves through it. Credit: National Oceanic and Atmospheric Administration. (b) Development of the complexity of the computer models used to simulate Earth's climate. Credit: National Aeronautics and Space Administration.

In a two-fold endeavor to handle the water cycle's spatial heterogeneity and further improve the spatiotemporal resolution of the measurements, reanalysis data products rose to the avant-garde (Parker, 2016). Reanalysis products preponderantly assimilate data from observations and model simulations (e.g., general circulation or earth system models). These models undergo continuous updates and enhancements, leading to significant alterations in the fundamental climatological characteristics of the model over time. As a result, relying on a long time series of operational analyses becomes impractical for studying extended trends or variations in climate. Reanalyses generate a dynamically consistent global analysis of the atmospheric state over an extended duration, spanning many years or decades, without spatial or temporal gaps. This process entails utilizing a "frozen" iteration of the analysis model and conducting a retrospective analysis using historical observation records (Betts et al., 2006). The benefit of this approach lies in utilizing a broader range of observational data, tapping into high-quality observations that might not be accessible to operational models in real-time (Figure 1.5(a)). Assimilation algorithms recursively combine observational and model simulation data within a Bayesian statistical framework (Figure 1.5(b)), which usually falls under one of four types of algorithms:

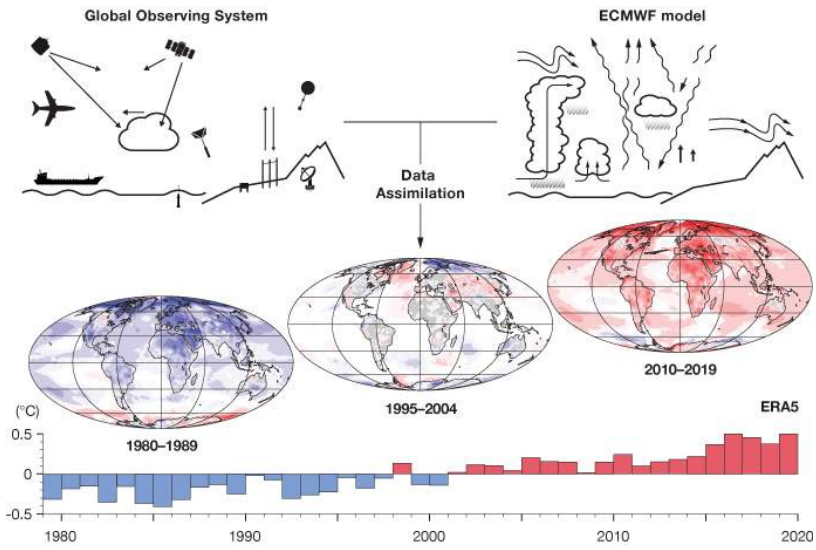
- Sequential, intermittent assimilation. This method involves intermittent assimilation, where available observations over a range of time are introduced at regular intervals into the model, e.g., the Arctic System Reanalysis (ASR; Bromwich et al., 2018).

- Sequential, continuous assimilation. In this method every observation is introduced into the model at the time it was registered, allowing for a continuous update of the model state, e.g., the Modern-Era Retrospective analysis for Research and Applications (MERRA; Rienecker et al., 2011).

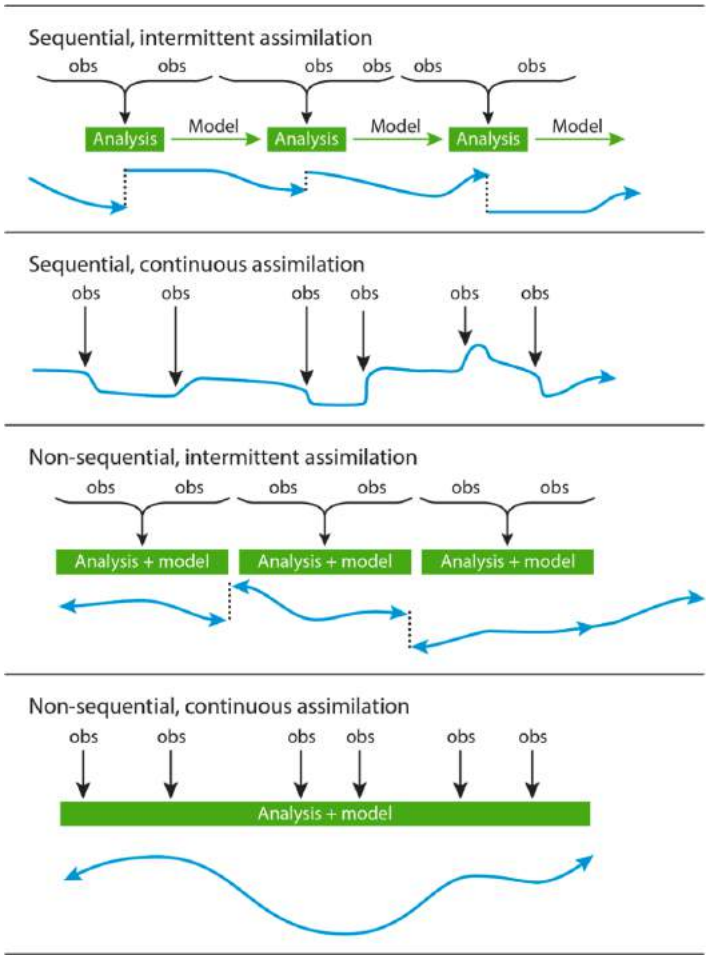
- Non-sequential, intermittent assimilation. This method incorporates various observations into a numeric weather prediction model at regular intervals, e.g., the European Centre for Medium Range Weather Forecasts (ECMWF) Reanalysis product v5 (ERA5; Hersbach et al., 2020).

- Non-sequential, continuous assimilation. This method allows for the assimilation of a broader range of observations in a more continuous and real-time manner, e.g., MERRA version 2 (MERRA-2; Gelaro et al., 2017)

Regretfully, since assimilation algorithms are statistically grounded, physical conservation principles might be overstepped, reflected in substantial variability compared to other data sources (Prein and Pendergrass, 2019).



(a)



(b)

Figure 1.5: (a) The principle of data assimilation in climate reanalysis. (b) Representation of four basic strategies for data assimilation, as a function of time.

Successive generations of reanalyses showcase continuous enhancements regarding better observational data availability, newer models used for simulations, different assimilation algorithms, fixing previous errors, or extending the record length. These improvements aim to refine the accuracy and reliability of the generated data sets over time. Examples include:

a.i National Centers for Environmental Prediction – National Center for Atmospheric Research Reanalysis 1 (NCEP/NCAR R1), an ongoing product with records starting in 1948, used the T62/28-level NCEP global spectral model and a three-dimensional variational (3DVAR) assimilation scheme (Kalnay et al., 1996).

a.ii NCEP – Department of Energy Reanalysis 2 (NCEP/DOE R2), an ongoing product with records starting in 1979, used the T62/28-level NCEP global spectral model with updated parameterizations and a four-dimensional variational (4D-Var) assimilation scheme (Kanamitsu et al., 2002).

b.i MERRA, a discontinued product with 1979 to 2016 records, used the Goddard Earth Observing System Model Version 5 (GEOS-5) and an incremental analysis update (IAU) assimilation scheme (Rienecker et al., 2011).

b.ii MERRA-2, an ongoing product with records starting in 1980, used the Goddard Earth Observing System Model Version 5.12.4 (GEOS-5.12.4) and a 3DVAR assimilation scheme (Gelaro et al., 2017).

c.i Japanese 25-year Reanalysis project (JRA-25), a discontinued produc with 1979 to 2004 records, used the Japan Meteorological Agency (JMA) Global Spectrum Model and a 3DVAR assimilation scheme (Onogi et al., 2007).

c.ii Japanese 55-year Reanalysis (JRA-55), an ongoing product with records starting in 1958, used the JMA Global Spectrum Model and a 4D-Var assimilation scheme (Kobayashi et al., 2015).

d.i ERA5, an ongoing product with records starting in 1940, used the Integrated Forecasting System (IFS) Cycle 41r2 model and a 4D-Var assimilation scheme (Hersbach et al., 2020).

d.ii ERA5-Land is produced using the tiled ECMWF Scheme for Surface Exchanges over Land incorporating land surface hydrology (H-TESSEL) IFS Cycle 45r1 model forced by meteorological fields from ERA5 without any further assimilation scheme, delivering hourly estimates from 1950 with a spatial resolution of 0.1 degrees (Muñoz-Sabater et al., 2021).

Unfortunately, reanalysis data has yet to overcome several issues already identified in their first generation of products. Trenberth and Guillemot (1995, 1998) evaluated the water cycle and the global atmospheric moisture budget in NCEP/NCAR R1 data between 1987 and 1993. They found that the reanalysis data sets could represent evaporation, precipitation, and moisture transport reasonably well. However, limitations in reanalysis data were identified, including marked biases in precipitable water as well as discrepancies in tropics' moisture structures and precipitation patterns, including a misrepresented intertropical convergence zone in the central Pacific. They noted that the reanalysis data quality depends on the input observations' availability and quality, the model's vertical resolution, computation methods, and analysis initialization. A decade later, Bosilovich et al. (2008) evaluated global precipitation across JRA-25, ERA-40 (Uppala et al., 2005), NCEP/NCAR R1, and NCEP/DOE R2 between 1979 and 2005. While there was a general sense that the precipitation reanalysis data are improving in recent systems, the study highlighted that this is not always the case, particularly in certain ocean regions. The authors emphasized the importance of using climate records of observed precipitation through a merged satellite and gauge data set as a reference for comparison, underscoring the ongoing need for careful assessment and evaluation when utilizing reanalysis precipitation data for various applications. Trenberth et al. (2011) assessed the global energy and water cycles, focusing on atmospheric moisture transports from the ocean to land, using eight current atmospheric reanalyses. These were the NCEP/NCAR R1, NCEP/DOE R2, Climate Forecast System Reanalysis (CFSR; Saha et al., 2010), Twentieth Century Reanalysis (20CR; Compo et al., 2011), ERA-40, ERA-Interim (Dee et al., 2011), JRA-55, and MERRA. The authors reported three main issues: first, moisture transport from ocean to land is not similar in most reanalyses; second, land discharge into the ocean differs significantly from observational estimates; and third, moisture recycling is too large and its lifetime too short. Once again, discrepancies among water cycle components are identified to arise with the changes in the observing system, model treatment, and analysis increments.

In short, the nature of atmospheric reanalyses, geared toward generating time series of the best available analyses considering the observing system through a statistical rather than a physical framework, introduces challenges. These reanalyses do not consistently conserve quantities dictated by conservation laws, leading to spurious changes with evolving observing systems over time. While natural variations in shorter time frames often overshadow these factors, their impact becomes more pronounced in decadal climate change or trends. Notably, the reanalyses tend to yield satisfactory results for land-based precipitation, but atmospheric fluxes over the oceans exhibit instability, disagreement, unreliability, and violations of fundamental physical constraints.

1.1.3 Multi-Source Quantification

While advancements in technology and research have provided various data sources for quantifying the global water cycle, it becomes evident that no single data source is without its flaws. Although valuable, climate model simulations and reanalysis have limitations (Valmassoi et al., 2023), and the often short and heterogeneous observational data records propagate uncertainties into our understanding (Schneider et al., 2017). However, one data source's strengths can compensate for another's weaknesses, emphasizing the importance of a comprehensive and integrated strategy. As we stand on the brink of a new era in water cycle research, the paradigm has shifted from the quest for the single best data source for each component to a focus on developing optimal methods for integrating information from diverse sources. Various integration methodologies have emerged, the most widely used ones being Bayesian Model Averaging (BMA; Hoeting et al., 1999), Constrained Linear Regression (CLR; Clemen, 1986), Modified Triple Collocation (MTC; Pan et al., 2015), Neural Networks (NN; Bishop, 1996), Optimal Interpolation (OI; Daley, 1999), and Simple Weighting (Rodgers, 2000). Additionally, post-processing closure methodologies, which distributed the budget residual among the components based on each component's uncertainties, explored Closure Correction Models (Munier and Aires, 2018) and Kalman Filter variations (Pan and Wood, 2006).

BMA is a statistical method that addresses model uncertainty by integrating information from multiple models to improve predictive performance. In BMA, the plausibility of each model is described by the posterior model probability, which is determined using Bayesian principles. This method provides a principled way to define model weights as posterior model probabilities, which is universal to all data-generating processes. BMA can account for model uncertainty when estimating model parameters and is particularly useful in applications with several plausible models where there is no definitive reason to choose a particular model over the others. CLR is a statistical method used when there is prior information available about a linear relation that the coefficients of the linear model should satisfy. This method is beneficial when there is a need to enforce specific constraints on the model parameters, such as ensuring that specific coefficients have a predefined relationship or that the model satisfies known physical laws or theoretical expectations. MTC is a mathematical technique to evaluate product error statistics without requiring the ground truth. MTC was introduced to model the error associated with wind speed measurements but can be applied to measuring any geophysical variable using three or more collocated target variables measured at the same time and place. The method simultaneously calculates linear calibration coefficients and measurement error variances. However, implementing TC is tedious and requires at least three

independent measurement systems, which are often hard to establish. NN is a computational model inspired by the structure and function of the human brain. It consists of interconnected nodes, or neurons, organized in layers. Each connection between nodes has an associated weight, and the network uses training data to learn and improve its accuracy over time. NN integrates and processes complex, multidimensional data to identify patterns, make predictions, or classify and cluster data. OI is a statistical method based on multidimensional analysis equations that combine observations with model simulations to produce more accurate output than the individual input data sets. The method involves merging background data, such as model outputs, with observation data, such as point measurements, to perform interpolation. OI assigns weights to the background and observation data sets based on their relative accuracy, which are determined from the background error variance, observation error variance, and background error correlation. Simple Weighting refers to the adjustment of integrating data sets to represent the target population accurately. This process involves assigning weights to individual data sets, calculated by dividing a given data set by the sample mean. This adjustment ensures that the integrated data set reflects the target mean and corrects any bias in the sample.

Aires (2014)compared the performance of the aforementioned integration methods and reported that Simple Weighting is the most suitable. Simple Weighting offers a straightforward formulation, and more elaborate methods do not offer enough improvement in results to justify the increased complexity they carry along. The Simple Weighting method assumes that the errors associated with the different products are Gaussian (zero-mean) and independent. However, there might be cases where this assumption may not hold, especially for gauge-based data products, and the dependence among products will cause an underestimation of the error associated with the integrated data set. The combined data set for a given component of the water cycle (P, ET, Q, or Δ TWS) is equal to:

$$x = \sum_{i=1}^{n} w_i x_i \tag{1.5}$$

where x is the combined data set for the single component of the water cycle being integrated, $x_1,\ x_2,\ x_3,\ ...,\ x_n$ are the different products considered, w_i is the associated weight of product x_i and is defined as:

$$w_i = \frac{(\bar{x} - x_i)^{-2}}{\sum_{j=1}^{n} (\bar{x} - x_j)^{-2}} \tag{1.6}$$

where $\bar{x}$ is the arithmetic mean of the n data products considered, and $(\bar{x} - x_i)^2$ is defined as the

error variance. That is to say, the weight associated to each product is proportional to the inverse of its error variance. Finally, the error associated to the combined data set x is:

$$e_x = \frac{1}{\sum_{i=1}^{n} (\bar{x} - x_i)^{-2}} \tag{1.7}$$

This perspective shift with multi-source data integration as a new north is crucial for achieving a more holistic and accurate understanding of the global water cycle. Therefore, the twenty-first century demands advancements in data collection because the ensemble is as strong as the weakest link and refinement of the analytical tools used in water cycle research.

Robust statistical methods for uncertainty quantification, sophisticated downscaling and disaggregation techniques, and the ability to analyze data across multiple scales are essential to this evolving paradigm. However, amidst this progress and a "clear" path to follow, a formidable challenge emerges—the lack of a unified standard for data distribution (Table 1.1). The various data sources often come in different grid reference systems, spatial resolutions, temporal resolutions, and file formats like the Hierarchical Data Format (HDF), General Regularly-distributed Information in Binary form (GRIB), and network Common Data Form (NetCDF) among the most common. This lack of uniformity poses a significant obstacle to gathering them prior to homogenizing data from different sources. The latter, data homogenization, feels like dealing with an irreversible process because spatiotemporal aggregation is easily achievable, but disaggregation is not. Regarding precipitation, for example, to aggregate in time, averaging precipitation rates or summing total precipitation will do the trick, and in space, the spatially weighted average will suffice.

There are different types of temporal rainfall disaggregation methods available. These methods can be broadly categorized into two broad methods, i.e., stochastic simulation and random cascade models. When it comes to sub-daily precipitation, however, stochastic simulations become the one reliable option because zeros, a frequent precipitation measurement at such scales, are fundamentally incompatible with the notion of self-similarity and multiplicative random cascades (Gupta and Waymire, 1993), yet they can be introduced artificially (e.g., Pathirana et al., 2003; Gires et al., 2013; Lombardo et al., 2017). Regarding stochastic simulations, it would be amiss not to discuss the framework proposed by Papalexiou (2018). The framework aims to preserve the processes' marginal distributions, correlation structures, and intermittency grounded on the assumption that any process can emerge by transforming a parent Gaussian process with a specific correlation structure. This approach unifies, extends, and improves a general-purpose modeling strategy, providing a consistent and fully general description that supersedes previous specific parameterizations and is applicable

for simulating a variety of hydroclimatic variables, such as precipitation, river discharge, wind speed, and humidity, as well as for multivariate applications.

Spatial disaggregation, or downscaling, techniques are also grouped under two primary umbrellas: statistical and dynamical downscaling. Statistical methods rely on empirical statistical relationships to associate local-scale variables with large-scale variables (e.g., univariate or multivariate regression), while dynamical downscaling is based on mathematical representations of complex physical processes (e.g., high-resolution regional climate models). Even though dynamical methods do not always produce significantly better results and are often considered too computationally demanding (Hellström et al., 2001), they are less demanding regarding data ingestion than statistical methods.

In addition to downscaling, there is regridding, also known as remapping or interpolation, in the spatial domain. Multiple grids, such as regular, rectilinear, curvilinear, and unstructured, are used in climate research. Therefore, even if two data sets have similar spatial resolution, regridding might be necessary if the data are on different or shifted grids. Among the most used regridding methods we have first-order conservative, distance-weighted approaches (e.g., nearest neighbour or bilinear interpolation) (Figure 1.6). First-order conservative remapping is primarily employed when working with latitude-longitude rectangular grids and ensures the preservation of the integral of the source field during the regridding process. This method calculates weights based on the ratio of the source cell area overlapping with the corresponding target cell (Jones, 1999). Distance-weighted regridding approaches assume that the variation of the interpolated quantity is linear between the grid points and assign weights to the source grid points based on their distance from the target grid point. For example, bilinear regridding considers the four nearest cells in a 2D grid to determine the value of the new cell (e.g., Accadia et al., 2003). Rajulapati et al. (2021) assessed the effects of regridding, and reported a substantial impact on the statistical properties of precipitation, with marked differences between the original and regridded datasets. These differences were most notable at high and low quantiles, particularly in tropical land regions and polar regions, respectively. The authors emphasized that the impacts of regridding vary spatially and at different quantiles, indicating that regridding should be approached with caution.

The data homogenization challenge not only demands technical solutions but also necessitates a broader conversation within the scientific community about establishing standards and protocols for data sharing in hydrology. Without a concerted effort to address this issue, the full potential of multi-source quantification may remain unrealized, hindering our comprehensive understanding of the complexities within the global water cycle. In the context of this technical need, specialized software is available to efficiently deal with processing and analyzing large amounts of data,

Table 1.1: Documentation of 20 most recent precipitation data products. The T62 Gaussian grid has 192 longitude equally spaced and 94 latitude unequally spaced grid points. Data type classification is based on the nomenclature by Sun et al. (2018).

Name	Domain	Spatial Resolution	Record Length	Temporal Resolution	Data Type	Format	Reference(s)
CMAP	Global	2.5° x 2.5°	1979–present	Pentad	Gauge-based	NetCDF	Xie and Arkin (1997)
CMORPH	60°SN	8 km x 8 km	1998–present	30 min	Satellite-based	NetCDF	Xie et al. (2017)
CPC-Global	Land	0.5° x 0.5°	1979–present	Daily	Gauge-based	NetCDF	Xie et al. (2010b)
CRU TS v4.07	Land	0.5° x 0.5°	1901–2022	Monthly	Gauge-based	NetCDF	Harris et al. (2020)
EM-Earth	Land	0.1° x 0.1°	1950–2019	Hourly	Gauge-based	NetCDF	Tang et al. (2022)
ERA5	Global	0.25° x 0.25°	1940–present	Hourly	Reanalysis	GRIB	Hersbach et al. (2020)
ERA5-Land	Land	0.1° x 0.1°	1950–present	Hourly	Reanalysis	GRIB	Muñoz-Sabater et al. (2021)
FLDAS	Land	0.1° x 0.1°	1982–present	Monthly	Model	NetCDF	McNally et al. (2017)
GPCC v2022	Land	0.25° x 0.25°	1891–2020	Monthly	Gauge-based	NetCDF	Schneider et al. (2017)
GPCP v3.2	Global	0.5° x 0.5°	2000–present	Daily	Satellite-based	NetCDF	Huffman et al. (2023)
GPM IMERG v07	Global	0.1° x 0.1°	2000–present	30 min	Satellite-based	HDF	Huffman et al. (2015)
GSMaP V05/v8	Global	0.1° x 0.1°	2000-present	Hourly	Satellite-based	NetCDF	Kubota et al. (2007)
JRA-55	Global	1.25° x 1.25°	1958–present	3-Hourly	Satellite-based	GRIB	Kobayashi et al. (2015)
MERRA-2	Global	0.5° x 0.625°	1980–present	Hourly	Reanalysis	NetCDF	Gelaro et al. (2017)
MSWEP v2.8	Global	0.1° x 0.1°	1979–present	3-Hourly	Satellite-based	NetCDF	Beck et al. (2019)
NCEP/DOE R2	Global	T62 Gaussian	1979–present	Daily	Reanalysis	NetCDF	Kanamitsu et al. (2002)
NCEP/NCAR R1	Global	T62 Gaussian	1948–present	4xDaily	Reanalysis	NetCDF	Kalnay et al. (1996)
PERSIANN-CDR	60°SN	0.25° x 0.25°	1983–present	Daily	Satellite-based	NetCDF	Ashouri et al. (2015)
PREC/L	Land	0.5° x 0.5°	1948–present	Monthly	Gauge-based	NetCDF	Chen et al. (2002)
TerraClimate	Land	1/24° x 1/24°	1958–present	Monthly	Model	NetCDF	Abatzoglou et al. (2018)

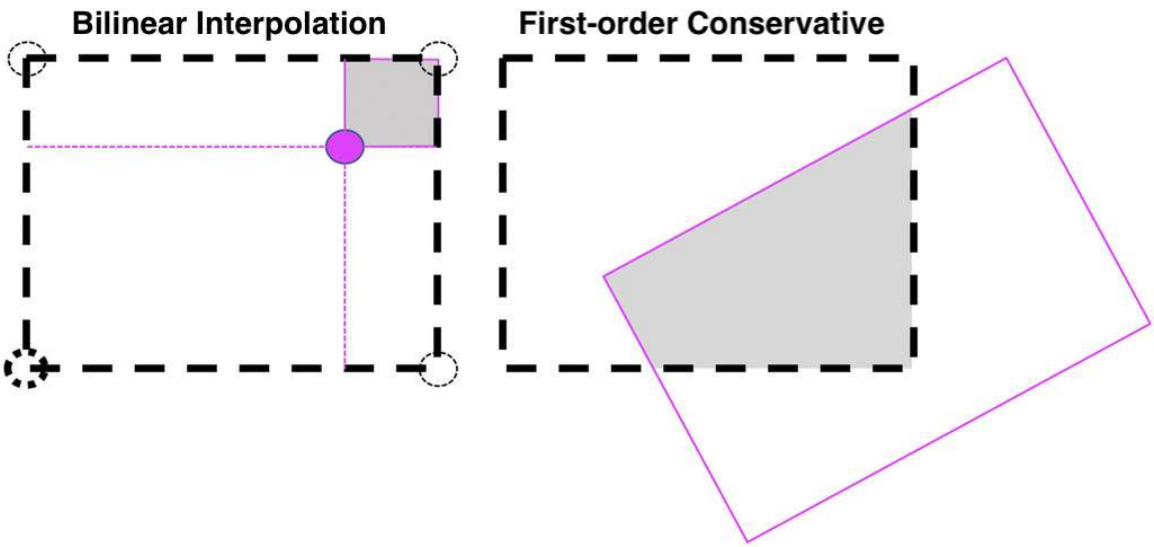


Figure 1.6: Schematics of bilinear and conservative regridding. Original cell are in dashed lines and target cell in solid lines. The bilinear interpolation weight corresponding to the node on the lower left is the ratio of the shaded area over the source cell area. The conservative interpolation weight associated with the target cell is the ratio of the shaded area over the source cell area. Modified from Pletzer and Hayek (2019)

namely the Climate Data Operators (CDO; Schulzweida, 2022) and Climate Data Analysis Tools (CDAT; Williams et al., 2009). These software provide data cleaning, merging, and analysis tools. While working with big data can be challenging and time-consuming, these or similar alternative software allow researchers to automate and streamline the data analysis process for reproducibility. However, while tools like CDO and CDAT provide valuable capabilities, they encounter a notable limitation—compatibility issues with Windows, the predominant desktop operating system globally. Although installing CDO and CDAT on Windows is technically feasible, the process involves utilizing the Windows Subsystem for Linux (WSL), essentially introducing a GNU/Linux environment to Windows (Singh, 2020). Moreover, it is crucial to note that CDAT is anticipated to undergo deprecation and cease support around the conclusion of the 2023 calendar year, adding a layer of consideration for researchers seeking sustained tools for their work.

More multipurpose software available include Fortran (Backus et al., 1957), MATLAB (Moler et al., 1982), Python (Van Rossum and Drake Jr, 1995), and R (Ihaka and Gentleman, 1996), which are the most used hydrology programming languages. Fortran continues to be a prevalent choice in hydrological modeling, leveraging the performance advantages inherent in low-level programming languages, particularly for computationally intensive tasks. While low-level languages offer computational efficiency, they are often criticized for being less readable and requiring more intricate programming compared to high-level languages such as MATLAB, Python, and R. MATLAB of-

fers comprehensive toolboxes for data analysis, statistical modeling, and visualization. Regrettably, these toolboxes are not widely adopted as they are, for many users, out of reach behind a paywall. Python and R have gained traction due to their flexibility, open-source nature, and ease of learning. Python, in particular, is valued for its extensive machine-learning libraries and general-purpose programming. Nonetheless, because of Python's Global Interpreter Lock (GIL), the threads within each process cannot truly run in parallel. Unlike Python, R is not limited to serial programming or convoluted workarounds because it supports Open Multi-Processing (OpenMP). Therefore, R is better suited for a broad spectrum of functions that range from data acquisition and manipulation to analysis, modeling, statistics, and visualization in High-Performance Computing (HPC) setups.

Furthermore, R has well-developed capabilities in geospatial and geographic information systems (GIS) applications (Gokceoglu and Pourghasemi, 2019), a particularly noteworthy facet of hydrological research. One of the remarkable strengths of particular interest lies in the thriving and actively engaged computational hydrology community that has flourished over the past five to ten years and the availability of documentation, tutorials, and online discussion platforms (Slater et al., 2019). The R hydrological community has significantly grown until it acquired a pivotal role in hydrological research and the operational practice of hydrology. This evolution is marked by the development of packages designed for various hydrological tasks, encompassing data retrieval and pre-processing from hydrological and meteorological sources, hydrograph and spatial analysis functions, and a spectrum of process-based and stochastic modeling tools.

The community-driven development of packages has led to a substantial expansion in functionality, catering to diverse needs within the hydrological domain. Nevertheless, more often than not, these packages are still developed around specific data sets or providers. For instance, packages like *easyclimate* facilitate access to high-resolution daily climate data for Europe (Cruz-Alonso et al., 2023), while *dataRetrieval* is tailored for the US Geological Survey (USGS) National Water Information System (DeCicco et al., 2022). Tools-centered packages offer comprehensive functionalities and often require more generic inputs. Examples include *envoutliers* adept at identifying outliers in environmental time series data (Čampulová et al., 2022), and *CoSMoS* a tool for generating univariate/multivariate non-Gaussian time series and random fields for environmental and hydroclimatic processes (Papalexiou et al., 2021). These latter kinds of packages give the users more flexibility on the account they are to deal with data gathering and pre-processing on their own.

These advancements position the R language seamlessly within production-ready ecosystems, leveraging cutting-edge technologies and tools to enhance reproducibility, testing, and continuous integration. These tools are imperative for addressing the inherent challenges when relying on diverse

data sources, particularly in the absence of ground observations. These tools are better exploited at regional scales, where ground observations for one or more water cycle components are available because these measurements constitute clear benchmarking or evaluation targets. Furthermore, localized studies at regional scales are essential in understanding the intricate interactions between climate, geography, and human activities within a specific area. By focusing on specific geographical areas, these studies provide insights into the unique characteristics, challenges, and variations within the water cycle of a particular region, such as Local Resource Management, Climate Change Impact, Extreme Events, Ecosystem Preservation, Water Quality, Infrastructure Planning, and Policy Development.

1.1.4 Narrowing the Focus

One region of particular interest is Czechia, a small Central European country with diverse landscapes and a growing population (United Nations, 2022). Czechia is a landlocked country (surrounded by Germany, Austria, Slovakia, and Poland) that covers an area of 78 864 km^2. The country experiences a temperate climate in the transition zone between the oceanic and continental climate types, with warm summers and cold, cloudy, and snowy winters (Tolasz et al., 2007). Czechia, an essential headwaters region of the European continent, is marked by various topographical features, including mountain ranges, plateaus, and lowlands, influencing its hydrological landscape. The country is home to several large rivers, including the Vltava, the Labe, the Morava, and the Oder, all of which have their sources within it (Figure 1.7). The Vltava River, in particular, originates in the Šumava Mountains and flows through Prague, the capital, shaping the central part of the country. Czechia is situated at the intersection of three sea drainage basins: the North Sea, the Baltic Sea, and the Black Sea, which, in return, divide Czechia into three main hydrological catchment areas: the Elbe, Oder, and Danube basins. All of these major watercourses drain water into neighboring states. The water sources of Czechia are thus almost exclusively dependent on precipitation. Sporadic rainfall is throughout the year, but concentrated rainfall is more frequent in the summer (Řehoř et al., 2021). Snowfall in the winter, especially in the mountainous regions, adds to the water storage and gradually feeds into rivers during the melting period (Jenicek and Ledvinka, 2020).

In recent times, Czechia has undergone notable transformations in its water cycle, impacting diverse facets of the water balance within the region. These alterations encompass shifts in river flow regimes, modifications in water quality, the decline of wetlands, and variations in the occurrence and intensity of extreme events. Since Czechia experiences a temperate climate with sufficient rainfall,

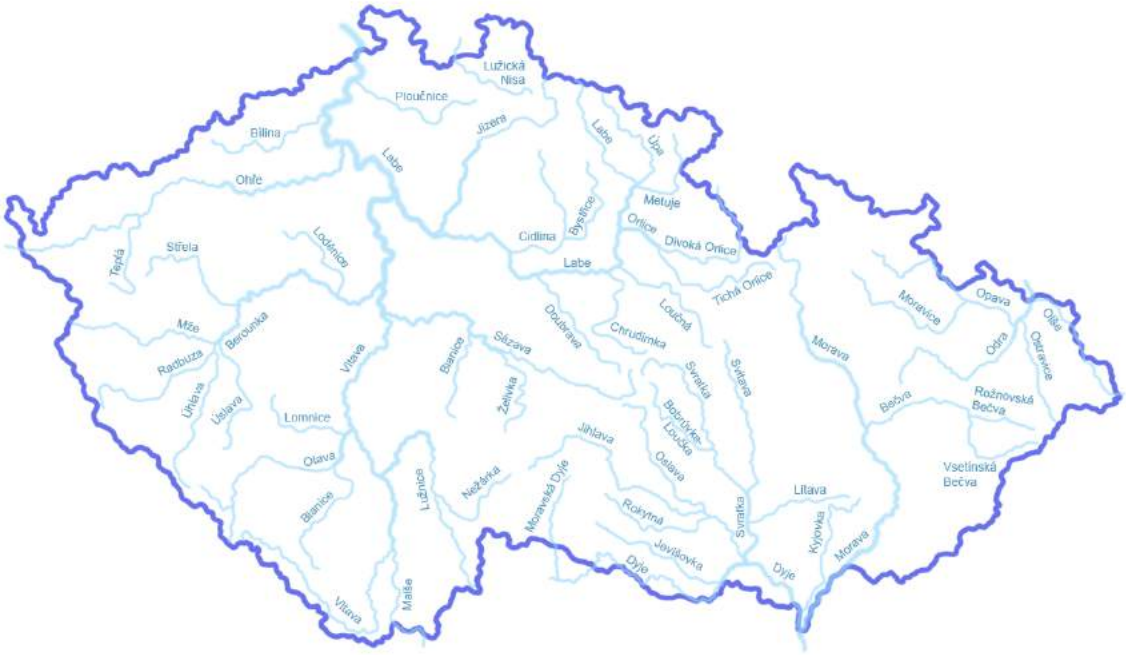


Figure 1.7: Map of the 50 longest rivers of the Czech Republic. Credit: Pavel Hrdlička

the water cycle is more likely to be limited by energy availability, such as solar radiation, rather than water availability. Case in point: an 18% increase in mean evaporation between 2001 and 2008 compared to the 1971–2000 average associated with rising global radiation and vapor pressure deficit. In addition, regional climate models estimate significant increases between 27–54%, with potential implications for water resources and ecosystems in the Czech Republic (Mozny et al., 2020).

Along the same lines, Nedelcev and Jenicek (2021) investigated trends in snowpack for 1965–2014 in 40 catchments in five mountain regions in Czechia. Therein, the snowpack was reported to be more sensitive to changes in air temperature at elevations below 900 meters above mean sea level (MAMSL), while precipitation had a more significant effect at elevations above 1200 MAMSL. However, snowpack sensitivity to air temperature increased at all elevations in the last few decades, resulting in changes in the rainfall-snowfall partition, which have decreased snow cover and premature snowmelt. Besides, snow water equivalent projections showed a decrease in annual maximum by 30–70%, occurring on average 3–4 weeks prematurely until the end of the 21st century (Jenicek et al., 2021). Regarding precipitation, an ensemble of regional climate models predicts that heavy precipitation events are likely to increase in severity, particularly in winter, with less agreement among models for the summer season (Kyselý and Beranová, 2009; Kyselý et al., 2011). At the same time, there is a projected decrease in the number of heavy rainfall events (Svoboda et al.,

2016), i.e., future projections estimate less yet more intense heavy precipitation events. In addition, increased human activities, such as urbanization and agriculture, have led to changes in land use and land cover, which in turn has contributed to the occurrence of floods and droughts. Consequently, the water cycle in Czechia and human activity find themselves on a causal feedback loop of sorts. Droughts, have had disastrous consequences for agriculture, forestry, water management, and other human activities (Brázdil et al., 2009). The abovementioned observed changes and future projections of the water cycle forecast the overall drying conditions over Czechia to extend well into the end of the century. Despite the identified trends in Czechia's water cycle changes, it is binding to acknowledge the substantial uncertainty inherent in climate projections. Complex climate systems, coupled with the limitations of current models, make it challenging to assert definitive outcomes. While there are observed shifts in precipitation patterns, snowpack dynamics, and evaporation rates, the range of potential future scenarios is wide (N–2N%) and necessitates caution in drawing absolute conclusions.

Understanding the regional scale water cycle's intricate dynamics prompts consideration of its broader implications for the global water cycle. On scales circa 4000 km and less, alterations in the water cycle are primarily influenced by the transport of moisture, contingent upon a combination of thermodynamic and dynamical processes (Dagan and Stier, 2020). The constraints imposed by energy budgets on a global scale and moisture budgets on a regional scale lead to changes in fundamental water cycle features, including changes in precipitation intensity, duration, and frequency, as the climate undergoes warming (Döll et al., 2018). Future water availability is driven by changes in evaporation, a process shaped by the overall rise in atmospheric evaporative demand and subject to modulation through vegetation's regulatory role in controlling evaporative losses (Vicente-Serrano et al., 2020). Regional water cycle changes result from the interplay between multiple potential drivers, including CO_2, aerosols, land use change, and human water use (IPCC, 2023). These changes can contribute to alterations in larger-scale water circulation patterns, potentially influencing the global water cycle. Recognizing the interconnectedness of regional and global water systems is essential in addressing the uncertainties and variability inherent in climate change impacts on water resources. Research efforts and collaborative initiatives on both regional and global scales are crucial for advancing the understanding of these complex interactions and refining data estimates to anticipate future water cycle dynamics better and address multiple Sustainable Development Goals (SDGs).

1.1.5 Nexus Repercussions

SDGs are a set of 17 global goals established by the United Nations in 2015 as part of the 2030 Agenda for Sustainable Development. These goals encompass a broad range of social, economic, and environmental objectives, aiming to address the world's various challenges. There are intricate connections between water dynamics and broader societal and environmental objectives (Figure 1.8), and the quantification of changes in the water cycle holds significant relevance to several of these goals:

- SDG 1 (No Poverty): While the direct connection might not be immediately apparent, understanding water availability, access, and management is crucial for communities that can harness water for various purposes, such as irrigation, aquaculture, or small-scale industries, have a better chance of breaking the cycle of dependency on a single source of income (Borgomeo et al., 2018).

- SDG 2 (Zero Hunger): Changes in the water cycle can impact agriculture, a sector highly dependent on water availability. Accurate quantification helps understand and address potential shifts in atmospheric water flux patterns and freshwater availability, ensuring food security (Alcamo, 2019).

- SDG 3 (Good Health and Well-being): Reliable water cycle data is vital for assessing and managing water-related health risks. Changes in precipitation patterns and consequent floods can influence the proliferation of waterborne diseases, and accurate quantification supports efforts to safeguard public health (Ternes et al., 2015).

- SDG 6 (Clean Water and Sanitation): Quantifying changes in the water cycle directly aligns with this goal, ensuring sustainable water management and sanitation practices. It aids in developing strategies for efficient water use and pollution prevention (Ho et al., 2020a).

- SDG 9 (Industry, Innovation, and Infrastructure): Understanding the evolving water cycle and the shifting characteristics of extreme events is crucial for sustainable infrastructure development. Accurate data helps design resilient infrastructure, especially in regions susceptible to changing water availability (Di Baldassarre et al., 2013).

- SDG 11 (Sustainable Cities and Communities): Cities are vulnerable to water-related challenges, from flooding to water scarcity. Quantifying water cycle changes supports urban

planning, helping communities build resilience and sustainable water management systems (Bhaduri et al., 2016).

- SDG 13 (Climate Action): The water cycle and its characteristics fundamentally define climate patterns. Accurate quantification contributes to climate models and predictions, supporting effective climate action and mitigation strategies (Mortimer et al., 2023).

- SDG 14 (Life Below Water): Changes in the water cycle impact aquatic ecosystems (e.g., salinity increases). Accurate quantification is crucial for monitoring and conserving marine environments, preserving biodiversity, and ensuring sustainable fisheries (Singh et al., 2019).

- SDG 15 (Life on Land): Terrestrial ecosystems are sensitive to water availability. Quantifying changes in the water cycle aids in understanding and mitigating the impact on land ecosystems and abnormal migration of species, promoting biodiversity conservation (Salleh, 2016).

In summary, quantifying water cycle changes is fundamental for achieving various SDGs, ensuring sustainable development, resilience, and the well-being of both human and natural systems.

1.2 Research Objectives

Considering the aforementioned research challenges and hindrances for water cycle quantification, this dissertation's overarching objective was to better understand the effects of different data sources, their integration, and the metrics/methods used to characterize water cycle changes on both global and regional scales. In addition, through a multidimensional approach, advance the implementation of open-source publicly available tools for data acquisition and processing, uncertainty quantification, and evaluation frameworks. In particular, the specific objectives of this dissertation were:

1. To chronologically trace the evolution of global water cycle quantification methods, emphasizing the challenges posed by spatiotemporal variability. Chapter 2 aims to comprehensively review historical attempts, data sources, and methods while critically assessing their contribution to improving spatiotemporal monitoring.

   Embarking on a chronological review was a foundational choice driven by the need to comprehensively understand the existing state of the art in water cycle quantification. If the power of the first computers, which occupied 167 square meters, now fits in our pockets, why has the uncertainty of our estimates not been reduced as drastically?

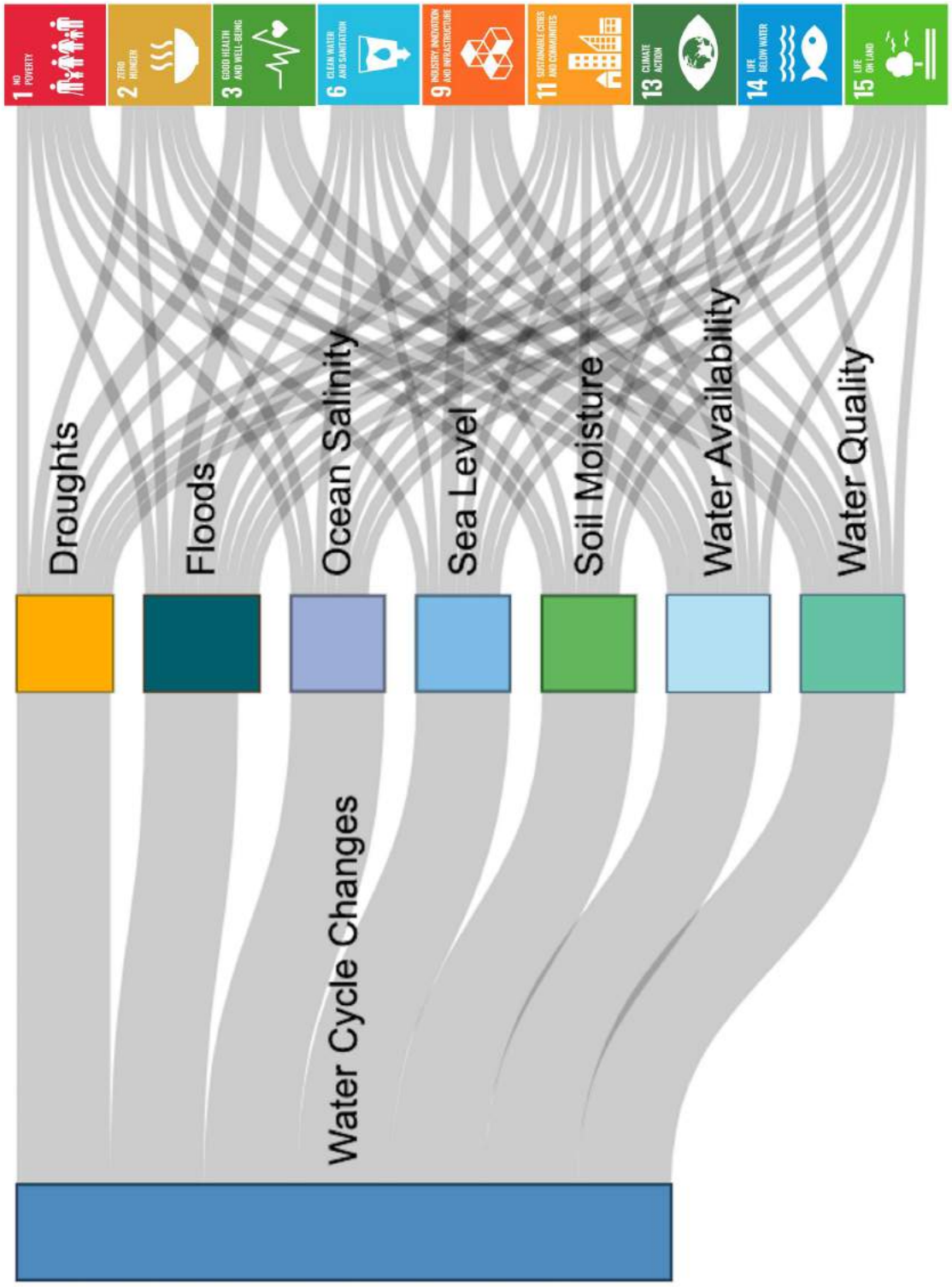


Figure 1.8: Sankey diagram depicting the links between impacts of water cycle changes and Sustainable Development Goals

2. To enhance the applicability of climate reanalyses in quantifying the global water cycle by proposing a framework that considers precipitation, evaporation, their difference, and their sum. Chapter 3 addresses the poor representation of the water cycle in reanalysis products despite being the most recent and advanced data sets available.

   The decision to delve into reanalyses emerged organically from their identified inconsistencies. If reanalyses, in theory, have the best of both worlds by exploiting observational data and model simulations in conjunction, why do their estimates describe a water cycle so different to the observed one?

3. To improve the robust quantification of climatologic properties of global precipitation by integrating multiple data sources addressing the heterogeneity among existing data products. Chapter 4 tackles data inaccessibility due to different spatiotemporal scales and distribution formats limiting the available assets to understand water cycle changes.

   The growing volume of data, both in terms of quantity and diversity, became increasingly apparent. This realization underscored the urgent need to develop an open-access, innovative, and user-friendly tool to confront the challenges posed by this burgeoning data landscape.

4. To investigate and understand recent regional water cycle changes via a novel method for benchmarking hydroclimatic data fusion based on water cycle budget closure. Chapter 5 demonstrates a case study to exploit the gained knowledge and developed tools to depict a multi-source water cycle budget perspective.

   Recognizing that regional changes in the water cycle pose unique challenges that demand tailored solutions, creating a robust and flexible framework that makes use of the attained outcomes for studying these intricate processes was the natural next step.

1.3 Dissertation Layout

This dissertation is composed of 6 chapters. The present Chapter 1 provides the overview and background of the conducted research and specifies the research objectives. The proceeding four chapters (Chapters 2 to 5) are structured as journal articles but without Abstract and Keywords. As each of these five chapters contains an Introduction section, which provides the state-of-the-art literature review on its topic, this dissertation does not include an individual chapter entitled “Literature

Review". In particular, Chapter 2 delves into the historical evolution of global water cycle quantification, highlighting the challenges posed by spatiotemporal variability (Vargas Godoy et al., 2021). Chapter 3 introduces the role of climate reanalyses in complementing traditional measurements, emphasizing the need for a comprehensive framework to address uncertainties (Vargas Godoy and Markonis, 2023b). Chapter 4 focuses on the integration of remote sensing data and model simulations to enhance precipitation quantification, presenting the *pRecipe* package as a tool for analysis (Vargas Godoy and Markonis, 2023a). Chapter 5 narrows the focus to the water cycle in Czechia, addressing recent changes in precipitation and evapotranspiration rates (Vargas Godoy et al., 2024).

Chapter 2

The global water cycle budget: A chronological review

2.1 Introduction

Water and the continuous circulation through its global cycle have played a fundamental role in sustaining life on Earth since its formation. The global water cycle is a complex phenomenon composed of several physiochemical processes such as condensation, evaporation, groundwater flow, infiltration, percolation, plant uptake, precipitation, runoff, sublimation, transpiration, and water vapor transport (Allan et al., 2020), coupled with anthropogenic interactions like water withdrawals and soil moisture use for livestock, crop irrigation, and forestry (Abbott et al., 2019). The longstanding representation of the global water cycle's conceptual model has been limited to three variables, namely precipitation, evaporation and runoff. Recently, this coarse representation has been partitioned to include the aforementioned sub-processes and their feedbacks. Our understanding of the global water cycle has been evolving over the years, and the methods we use to quantify hydrometeorological variables have adapted to exploit new technologies. Furthermore, the need to better estimate the components of the global water cycle has driven tailor-made technological developments as well (e.g., satellite instruments; Hildebrand et al. 2003; Levizzani and Cattani 2019).

Remote sensing data and model simulations complemented the traditional surface-based measurements and offered unprecedented coverage over previously inaccessible or unmonitored regions. Even though these advances provided vast data sources, and aided to quantify water cycle components at multiple scales, their varying performances and uncertainties limit their applicability

to global scale analyses (Brocca et al., 2019). Thus, the number of primary components used to quantify the global water cycle has not changed much. The most substantial differences that arose with the inclusion of satellite data are the decomposition of total evaporation into evaporation over oceans and evapotranspiration over land (Dickinson, 1984), and the addition of total water storage (Lvovitch, 1973). The above components represent the major inputs, outputs, and storage of the global water cycle. Hence, if we apply the mass conservation principle, we may write the water budget equation, which relates to these four components as follows.

$$\Delta\text{TWS} = P - \text{ET} - Q \tag{2.1}$$

where Δ TWS is the change in total water storage (as the sum of groundwater, soil moisture, and surface water such as river water, snow water, and water in lakes), P is precipitation, ET is evapotranspiration, and Q is the net water transport. The rest of the global water cycle processes are, to some extent, encompassed in these four components (Bengtsson, 2010). Inadvertently, aggregating global water cycle components to the most dominant ones also aggregates the underlying uncertainties of the minor components, which are overshadowed by the uncertainties of the major components with the available accuracy at the moment. Global water cycle quantification accuracy is further hindered by the inherent biases revealed in the first attempts to unify multiple data sources for a single component due to the vast heterogeneity of algorithms and data used (Hegerl et al., 2015).

Uncertainties in the quantification of global water cycle components are indispensable when attempting to close the water budget. We can express equation 2.1 as:

$$P - \text{ET} - Q - \Delta\text{TWS} = \xi \tag{2.2}$$

where ξ is the budget residual, which in a closed budget equals to zero. Through the years, there have been various attempts to close the budget (Starr and Peixoto, 1958; Willmott et al., 1985; Sheffield et al., 2009; Sahoo et al., 2011). They have used different data sources and methods to minimize the residual, but non-closure of the water budget still prevails. Alternatively, rather than using budget closure as the performance metric, some researchers prefer to look at runoff as a diagnostic flux to assess their results (Sheffield et al., 2009). Closing the water budget not only will improve our understanding of the global water cycle, but will necessarily lead to improvement of the accuracy of the data involved. Enhancing data accuracy is of critical importance for applications in climatology, hydrology, meteorology, and water resource management, to name a few.

To keep moving forward towards closure of the global water cycle, ergo more accurate data, it

would be beneficial to assess previous achievements. Herein, we present a review of the chronological evolution of the paradigms regarding the global water cycle budget. We provide an in-depth recapitulation of the advancements in global water cycle quantification. In addition, we present a comparison between budgets reported in the literature, with highlights on the methods and data sources used. Using significant technological improvements as timeline reference milestones, we considered four epochs, namely Early Days of Hydrology, Model Simulations Period, Satellite Era, and Age of Big Data. Each epoch is characterized by its own accomplishments and challenges. Some of the latter were overcome in succeeding epochs and some prevailed up to the present. Despite data reaching unprecedented availability, detail, and coverage, the quest for robust quantification of the global water cycle remains.

2.2 Chronicle

2.2.1 Early Days of Hydrology

Studies of the global water cycle are as old as hydrology. In classical Greece, Plato and Aristotle philosophized that groundwater might be the component responsible for circulating water resources by connecting rivers and lakes. However, Marcus Vitruvius is most commonly credited to be the first one to conceptualize the water cycle. In the first century BCE, Vitruvius proposed a philosophical description of the water cycle that placed precipitation instead of groundwater as a critical component of water transport (Pollio, 1648). Vitruvius planted a seed that would later lead both, yet independently, during the sixteenth century, Leonardo da Vinci and Bernard Palissy into describing a water cycle with three principal components: precipitation, evaporation, and runoff (Palissy, 1580; Pfister et al., 2009). Therefore, equation 2.1 was originally formulated as:

$$P - E = Q \tag{2.3}$$

where P is precipitation, E is evaporation, and Q is the runoff or exceeding precipitation. With this theoretical formulation, the scientific community ventured into quantifying the above components during the seventeenth century. Pierre Perrault and Edmund Halley were among the pioneers that supplemented experimental science to hydrology with their research on catchment precipitation and evaporation, respectively (Brutsaert, 2023). John Dalton was the first to quantify all three above-listed components for England and Wales, providing a comprehensive quantification of a water cycle and not just a single component of it (Dalton, 1799).

With catchment scale quantification achieved, the next step was to aim for global-scale quantification. During the next years and up to the end of the 1960s, numerous studies, mainly coming from Germany and Russia, attempted to quantify the global water cycle. Baumgartner and Reichel (1972) surveyed the literature on the global water cycle quantification during the 1900s and added their findings to the previous compilation by Reichel (1952), accounting for over 40 studies (Table 2.1). Over land, precipitation, range between (99 to 122) $\times 10^3$ km^3/year, evapotranspiration range between (52 to 97) $\times 10^3$ km^3/year, and and runoff range between (25 to 48) $\times 10^3$ km^3/year. Over oceans, precipitation and evaporation range between (242 to 412) $\times 10^3$ km^3/year and (273 to 458) $\times 10^3$ km^3/year, respectively. Note that evaporation and evapotranspiration have the most extensive ranges, presumably, because these values were derived from other measurements since, at the time, it was not possible to obtain direct observations. Even so, several reported fluxes are similar, if not identical, which may be caused by the fact that despite using different approximations or formulations, the initial data set used was the same. Over land precipitation estimates were derived from gauge and chart data, runoff estimates were derived from the river measurements by Marcinek (1964), and evaporation estimates were computed as the difference between precipitation and runoff. Over oceans, heat balance maps, and climatological data for fixed locations constituted evaporation estimates, runoff is the same as overland because of atmospheric water balance (Rasmussen, 1970), and precipitation estimates were the difference between evaporation and runoff.

Due to the high variability in time and space of global water cycle components, ground station reports were not representative of the surrounding areas. Besides, it has been typical for developing countries not to possess a ground station network dense enough to monitor global water cycle components in those regions (Willmott et al., 1994). Aware of the above, Baumgartner and Reichel (1972) introduced very strong yet somewhat arbitrary correction assumptions, and estimated the errors based on the biggest difference between the values compiled on their survey. Considering that the precipitation measured by rain gauges is smaller than the amount reaching the surface and there are different zonal climatic conditions overland, the authors suggest three different options to correct precipitation underestimation. They pointed out that the scenario selected is the most probable, yet no explanation is provided towards why that is. Correcting precipitation overland has a ripple effect because it is used to compute runoff, which is then used to compute precipitation over the oceans. Based on their assumptions, they report the quantification of the global water cycle had been achieved within a margin of ten percent relative error.

A decade later, Willmott et al. (1985) presented the first study with sufficient spatial coverage. Their study was based on temperature and precipitation observational data records from 13,332

globally distributed stations, and estimated terrestrial snow-cover, soil moisture, and evapotranspiration. Their work extended on previous regional studies over Africa (Mather, 1962), Asia excluding U.S.S.R. (Mather, 1963a), U.S.S.R. (Mather, 1963b), Australia, New Zealand, and Oceania (Mather, 1963c), Europe (Mather, 1964a), North America excluding U.S.A. (Mather, 1964b), U.S.A. (Mather, 1964c), and South America (Mather, 1965). The above cumulatively used only 8,565 stations from the same network Willmott et al. (1985) used on their study. Still, they had to use empirical equations and a revised version of the potential evapotranspiration method of Thornthwaite (1948) in order to derive snow-cover, soil moisture, and evapotranspiration from the temperature and precipitation observational data available. Willmott et al. (1985) did not report single values as annual averages, but presented their results in maps where it could be seen that annual mean evapotranspiration is approximately 173×10^3 km^3/year over continental regions near the equator, 43×10^3 km^3/year towards the poles, and below 43×10^3 km^3/year across the Sahara, Arabia and Central Asia. Nonetheless, we know now, technological limitations and the lack of data sources place the findings of the above discussed studies in a best-guess scenario only.

Table 2.1: Modified from Baumgartner and Reichel (1972) to exclude incomplete rows. All the fluxes are in 10^3 km^3/year. P_{L} is precipitation overland, ET is evapotranspiration overland, Q is runoff, P_{O} is precipitation over oceans, E is evaporation over oceans, P_{TOT} is total global precipitation, and E_{TOT} is total global evaporation.

Author	P_{L}	ET	Q	P_{O}	E	P_{TOT}	E_{TOT}
Brückner (1905)	122	97	25	359	384	481	481
Fritzsche (1906)	112	81	31	353	384	465	465
Schmidt (1915)	112	81	31	242	273	354	354
Wüst (1922)	112	75	37	267	304	379	379
Cherubim (1931)	112	75	37	334	371	446	446
Meinardus (1934)	99	62	37	412	449	511	511
Halbfaß (1934)	100	52	48	410	458	510	510
Wüst and Defant (1936)	99	62	37	297	334	396	396
Wundt (1938)	99	62	37	346	383	445	445
L'vovitch (1945)	107	71	36	412	448	519	519
Möller (1951)	99	62	37	$\leq$324	$\leq$361	$\leq$423	$\leq$423
Reichel (1952)	100	70	30	315	345	415	415
Wüst et al. (1954)	100	73	27	324	351	424	424
Budyko (1955)	100	66	34-38	370	408	470	474
Albrecht (1960)	100	67	33	378	411	478	478
Budyko (1963)	107	61	46-48	404	452	512	513
Mira (1964)	108	72	36	412	448	520	520
Nace (1968)	100	69	31	319	350	419	419
Kessler (1968)	100	60	40	410	450	510	510
Mather (1969)	106	69	37	382	419	488	488
L'vovitch (1970)	109	72	37	411	448	520	520
Budyko (1970)	107	64	43	412	455	519	519

2.2.2 Model Simulations Period

In simple terms, General Circulation Models (GCMs) are a set of theoretical and empirical mathematical expressions that attempt to simulate climate's physical processes. They could be an atmospheric GCM, an oceanic GCM, or a coupled GCM. The first atmospheric GCM was introduced by Norman Phillips (1956), and it opened the door to new opportunities for global water cycle quantification (McGuffie and Henderson-Sellers, 2001). Not long after, towards the end of the 1960s, the National Oceanic and Atmospheric Administration Geophysical Fluid Dynamics Laboratory developed the first coupled GCM (Manabe and Bryan, 1969). The basic structure of a GCM can be seen in figure 2.1. The GCM spatial domain is composed of 3D cells, whose horizontal grid is typically formed by latitude and longitude, and pressure levels determine the cell height. The number of physical processes considered and the complexity to which they are represented have continuously improved since the introduction of GCMs. Today's models further account for terrestrial vegetation and the carbon cycle with an explicit representation of biogeochemical processes - such models are referred to as Earth System Models or ESMs (Flato, 2011; Collins et al., 2013; Hurrell et al., 2013; Flato et al., 2013; Otto-Bliesner et al., 2016).

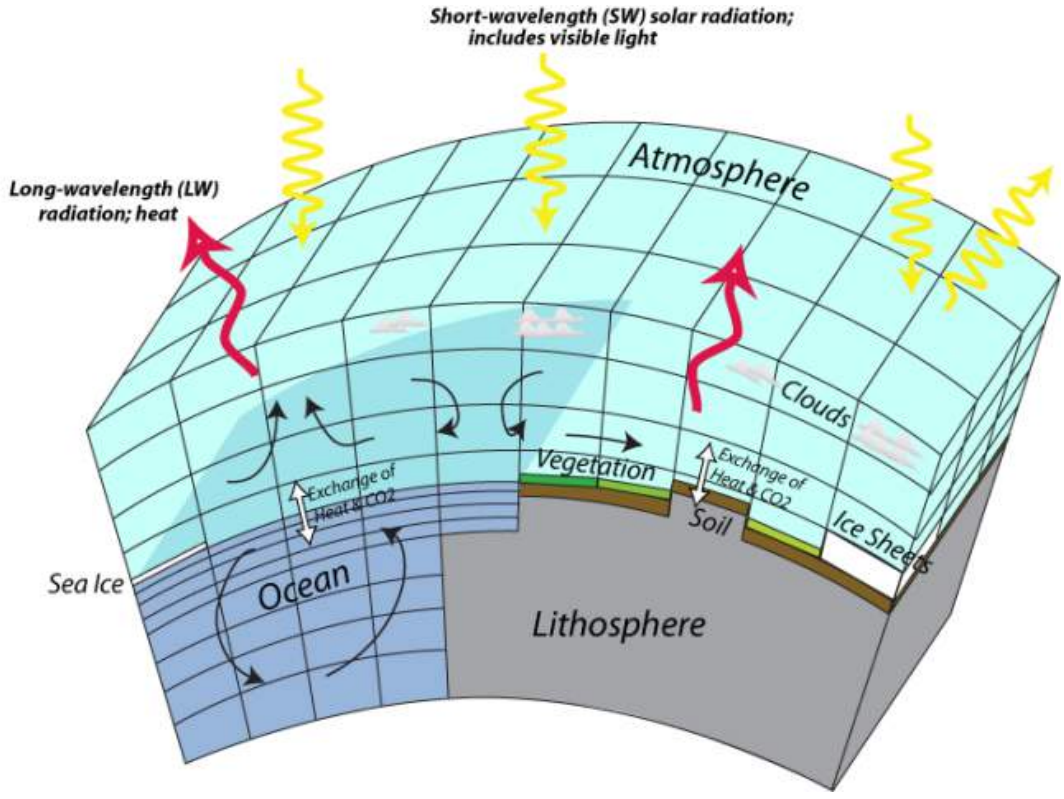


Figure 2.1: Schematic structure of a General Circulation Model modified from Bralower and Bice (2012).

Model simulations were initially driven exclusively by ground observations. Later on, satellite remote sensing, model reanalysis data sets, or different combinations of them were assimilated. Hydrological models revolutionized the quantification of the global water cycle by providing regular gridded data with global coverage as well as constant time steps. On top of that, both statistical and dynamical downscaling of GCMs and ESMs have evolved over the past decades to enable more reliable estimates (Tapiador et al., 2020). For example, the most recent release of the European Centre for Medium Range Weather Forecasts Reanalysis product (ERA-5), which is a reanalysis based on the European Centre for Medium-Range Weather Forecasts' Integrated Forecasting System (ECMWF's IFS) weather model, provides a 30 km global coverage with 137 atmospheric pressure levels capped at 80 km with uncertainty ranges reported at each level (Hersbach et al., 2020). Despite the exponential growth in computing power efficiency, many fundamental processes like radiative transfer, convection initiation, hydrometeor phase change, and cloud microphysics that occur between the sub-kilometer scale and the microscale (i.e., nine orders of magnitude less than current model resolutions) are parameterized, as they cannot be resolved at the model resolution. On that account, while GCMs and ESMs provide global coverage of water cycle components, their spatial and temporal resolution are still relatively coarse, hindering validation attempts.

Model simulations further changed global water cycle quantification by providing more robust formulations towards the estimation of evapotranspiration. The bucket model developed by Budyko (1961) was implemented for the evapotranspiration scheme used in the first coupled GCM (Manabe and Bryan, 1969). This scheme oversimplified the physical processes surrounding evapotranspiration (figure 2.2); nevertheless, its results were not significantly different from much more complex formulations attempted in contemporaneous GCMs (Carson, 1982). In the aforementioned scheme, evapotranspiration depends on potential evaporation, soil water content, field capacity (defined as the amount of soil moisture or water content held in the soil after excess water has drained away and the rate of downward movement has decreased), and water holding capacity (Carson, 1982). Federer et al. (1996) compared five surface-independent and four surface-dependent potential evapotranspiration approximation schemes in models, and their results suggest that, at that time, none of the methods significantly differ from each other for most surface types. Still, the authors point out that the Penman-Monteith (Monteith and Unsworth, 2013) and Shuttleworth & Wallace (Shuttleworth and Wallace, 1985) methods might pose as the most comprehensive for global-scale analysis, a hypothesis that was later confirmed for Penman-Monteith (Wang and Dickinson, 2012).

The coupled GCM introduced by Manabe and Bryan (1969) simulated average values of 93.4×10^3 km^3/year overland precipitation, 69.5×10^3 km^3/year evapotranspiration, 23.9×10^3 km^3/year runoff,

359.3×10^3 km^3/year over ocean precipitation, and 429×10^3 km^3/year evaporation. In recent years, Haddeland et al. (2011) compared 11 model simulations for the period 1985-1999 (Table 2.2). Observation-based data for global precipitation overland had an average value of 126×10^3 km^3/year, simulated evapotranspiration, and runoff mean values range between (60 to 85) $\times 10^3$ km^3/year and (42 to 66) $\times 10^3$ km^3/year, respectively. Note that Manabe's evapotranspiration estimate is the only flux within the values reported by Haddeland et al. (2011). Besides, the later estimates are within the range for annual averages reported by Baumgartner and Reichel (1972), hinting that despite the substantial uncertainties and approximations, the values reported in the previous period were not that far from the current ones.

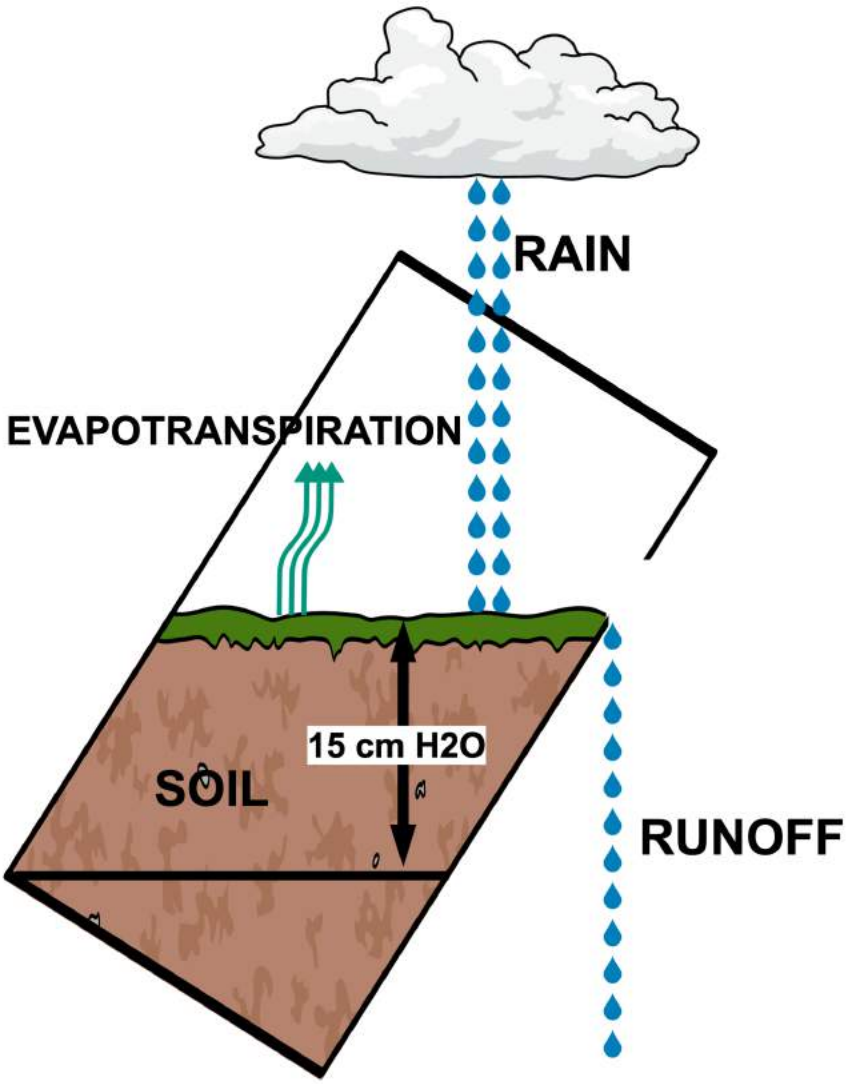


Figure 2.2: Schematic of the Budyko bucket model implemented by Manabe and Bryan (1969). The model represents a single layer soil reservoir with a defined maximum field water capacity of 15 cm from which soil water evaporates at a rate proportional to the remaining water content.

Model simulations did represent a new data source with seeming advantages over observations like the ability to generate global coverage data and perhaps more revolutionary to forecast, predict, and project. Nevertheless, once again, the scientific community relied heavily on observational data because it was crucial for model calibration and validation. Consequently, this novel opportunity to research global water cycle variability and its response to global warming further stressed the need for better observation-based measurements and more accurate quantification of the cycle components.

2.2.3 Satellite Era

Shortly after the introduction of climate models (Phillips, 1956), the Television Infrared Observation Satellite (TIROS-1 or TIROS-A) became the first weather satellite successfully launched in 1960, and so it began the satellite era (NOAA, 1987). Barnes and Bowley (1968) proved the effectiveness of satellite observations in hydrology when they published their findings on snow cover mapping over the Missouri and Upper Mississippi River basins. Thereafter, several satellite missions made it into orbit, among the most notable, we may mention the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) missions. Based on their orbits, satellites could be grouped into two major groups, either geosynchronous orbit (GEO) or polar orbit. Many of the satellites involved in the EOS missions have a nearly polar orbit. Polar-orbit satellites move around the Earth in a Sun-synchronous orbit so that the overpass occurs at the same local time every day, taking around 100 minutes to complete an orbit. These satellites overpass the equator at the same local solar time each day. Satellite sensors could be active or passive, and it is not uncommon for both to be onboard the same satellite. For example, the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), a passive sensor, and the Precipitation Radar (PR), an active sensor, were onboard the TRMM satellite. Regarding satellites and missions of particular interest for global water cycle quantification, we have the TRMM (Huffman et al., 2007) and the Global Precipitation Measurement (GPM) (Huffman et al., 2015) for precipitation, the Moderate Resolution Imaging Spectroradiometer (MODIS) for evapotranspiration (Mu et al., 2011), and the Gravity Recovery and Climate Experiment (GRACE) for total water storage (Tapley et al., 2004). There is no specific instrument nor mission dedicated solely to runoff yet (Hong et al., 2007). However, runoff could be derived from other satellite observations, for instance, TRMM precipitation (Huffman et al., 2007), and MODIS landcover (Friedl et al., 2002) using the Natural Resources Conservation Service (NRCS) runoff curve number method (Cronshey, 1986; Burges et al., 1998).

Satellite observations complemented the traditional surface measurements and offered unprece-

Table 2.2: Modified from Haddeland et al. (2011). LW is downward longwave radiation flux, LW_{net} is net longwave radiation flux, P is precipitation (rain or snow distinguished in the model), q is specific humidity, RR is rainfall rate, S is snowfall rate, SW is downward shortwave radiation flux, SP is surface pressure, T is air temperature, T_{max} is maximum daily air temperature, T_{min} is minimum daily air temperature, and W is wind speed. Bulk formula: Bulk transfer coefficients are used when calculating the turbulent heat fluxes.

Model name	Model time step	Meteorological forcing variables	ET scheme	Reference(s)
GWAVA	Daily	P, T, W, q, LW_{net}, SW, SP	Penman-Monteith	Meigh et al. (1999)
H08	6h	RR, S, T, W, q, LW, SW, SP	Bulk formula	Hanasaki et al. (2008)
HTESSEL	1h	RR, S, T, W, q, LW, SW, SP	Penman-Monteith	Balsamo et al. (2009)
JULES	1h	RR, S, T, W, q, LW, SW, SP	Penman-Monteith	Cox et al. (1999); Essery et al. (2003)
LPJmL	Daily	P, T, LW_{net}, SW	Priestley-Taylor	Bondeau et al. (2007); Rost et al. (2008)
MacPDM	Daily	P, T, W, q, LW_{net}, SW	Penman-Monteith	Arnell (1999); Gosling and Arnell (2011)
MATSIRO	1h	RR, S, T, W, q, LW, SW, SP	Bulk formula	Takata et al. (2003); Koirala (2010)
MPI-HM	Daily	P, T	Thornthwaite	Hagemann and Dümenil (1997); Hagemann and Gates (2003)
Orchidee	15 min	RR, S, T, W, q, SW, LW, SP	Bulk formula	de Rosnay and Polcher (1998)
VIC	Daily/3h	P, T_{max}, T_{min}, W, q, LW, SW, SP	Penman-Monteith	Liang et al. (1994)
WaterGAP	Daily	P, T, LW_{net}, SW	Priestley-Taylor	ALCAMO et al. (2003)

dented observational coverage on a global scale (McCabe et al., 2017). The Defense Meteorological Satellite Program (DMSP) near-polar orbiting satellites have been key providers of data over the oceans since 1987 (Dubach and Ng, 1988). Onboard their satellites, the most notable instruments are the Special Sensor Microwave Imager (SSM/I) (Hollinger, 1991) and its successor, the Special Sensor Microwave Imager Sounder (SSMIS) (Kunkee et al., 2008). These passive microwave radiometers provide measurements used to derive data on surface wind speed, atmospheric water vapor, cloud liquid water, and rain rate, which are critical to quantifying the global water cycle (Robertson et al., 2014). Furthermore, various present-day models and reanalysis products assimilate satellite observations (van Dijk and Renzullo, 2011). Nonetheless, like for GCMs, ground observations are crucial for satellite data validation. Notwithstanding, the number of ground stations worldwide has been declining since the 1970s (Walker et al., 2016). It was not before Trenberth et al. (2007) that the availability of observational and modeled data to quantify the global water cycle was exploited. A year prior, Oki and Kanae (2006) presented a quantitative synthesis of the global water cycle. Instead of estimating the budget, they made a compilation of individual studies to stress the importance of global water cycle quantification and further assessment to manage renewable freshwater resources properly. This concern has been in the minds of the scientific community for quite some time now (Falkenmark and Lindh, 1974). The budget assessments by Trenberth et al. and Oki & Kanae are held in high regard and are often used as a sort of validation reference (Rodell et al., 2015).

Oki and Kanae (2006) addressed the availability of renewable freshwater resources for human consumption within the global water cycle. The authors stressed that freshwater availability would be better assessed by fluxes than by storages because water is a circulating resource. Also, given the high variability of the water cycle in time and space, water stress is not a problem of how much water is available but a matter of when and where it is available (Postel et al., 1996). To better represent their research, they synthesized previous estimates of global water cycle fluxes and storages (Korzoun 1978; Shiklomanov 1998; Dirmeyer et al. 2006; Oki 2006). By doing so, they also presented a much more comprehensive mean state of the global water cycle. Their results showed overland precipitation of 111×10^3 km^3/year, evapotranspiration of 65.5×10^3 km^3/year, and runoff of 45.5×10^3 km^3/year. Moreover, precipitation is divided into rainfall and snowfall, plus the fluxes are allocated to different terrains or land uses. Over oceans, precipitation was 391×10^3 km^3/year and evaporation was 436.5×10^3 km^3/year.

Trenberth et al. (2007) used different data sources to quantify the global water cycle and its components. Three data sets were selected for precipitation, the Global Precipitation Climatology

Project (GPCP v2; Adler et al. 2003), the University of East Anglia Climatic Research Unit time-series (CRU TS 2.1; Mitchell and Jones 2005), and the PRECipitation REConstruction over Land (PREC/L; Chen et al. 2002). Evapotranspiration was simulated using the Community Land Model version 3 (CLM3; Bonan et al. 2002; Qian et al. 2006), which was forced using a combined PREC/L and GPCP precipitation data set. Surface plus subsurface runoff was derived from two climatic water balance estimates (evapotranspiration minus precipitation), the first from the European Centre for Medium Range Weather Forecasts Reanalysis 45 year product (ERA40; Uppala et al. 2005) using the methods described by Trenberth and Guillemot (1998), and the second using evapotranspiration from CLM3 and GPCP precipitation. Additionally, the authors relied on previous work for some components of the global water cycle like surface runoff (Dai and Trenberth, 2002), ice volumes (Houghton et al., 2001), soil moisture (Webb et al., 1993), and groundwater (Schlesinger, 2005). It was common for prior studies to cite values that, in return, cite another and so on. Unlike them, the authors documented, and traced back as far as possible, the origins of the values used. They reported 113×10^3 km^3/year overland precipitation, 73×10^3 km^3/year evapotranspiration, 40×10^3 km^3/year runoff, 373×10^3 km^3/year over ocean precipitation, and 413×10^3 km^3/year evaporation.

It is important to note that satellite data records are recently of sufficient time frame lengths and with methods "mature" enough to develop meaningful global water cycle climatology records that can provide information on its components mean state and variability (Schlosser and Houser, 2007; Robertson et al., 2014). Exploiting the increasing availability and maturity of satellite products, Sheffield et al. (2009) addressed the feasibility of closing the water budget, relying solely on satellite-based products. They combined the TRMM Multi-satellite Precipitation Analysis (TMPA; Huffman et al. 2007) and the Climate Prediction Center morphing method (CMORPH; Joyce et al. 2004) products for precipitation, the University of Colorado GRACE time series (CSR RL04; Wahr et al. 1998) for total water storage, and they derived evapotranspiration from Aqua satellite data using the Penman-Monteith revised formulation proposed by Mu et al. (2007). Then they evaluated their findings over the Mississippi River basin comparing their runoff estimates, computed as the budget residual, with ground observations. Their results indicate that the data products selected do not close the budget because the computed runoff is greatly overestimated compared to ground measurements. The authors suggest that further improvement of satellite-based products may reduce the residual, and suggest multi-source data merging as a complementary means to achieve budget closure.

2.2.4 Age of Big Data

In this day and age, we have transitioned from minimal data coverage and sources into a widely heterogeneous abundance. In contrast to the continuous decline in the number of ground stations, satellite-based and model-derived data products have proliferated. However, while some components of the global water cycle have multiple products to choose from (e.g., precipitation), others do not (e.g., total water storage). Some products assimilate or calibrate against ground station data to improve their performance (Rudolf and Schneider, 2005); others implemented machine learning processing to do so (Hong et al., 2004). It is not uncommon to find performance comparisons between products in the literature, evincing large differences in the magnitude and the variability of the estimates (e.g., as much as 300 mm/year difference between precipitation data sets; Sun et al. 2018). In their global comparison of 30 data sets at multiple spatiotemporal scales, Sun et al. (2018) found that, in general, variability from reanalysis data sets is more substantial than that from other data sources. Conversely, we can see that no single data set performs the best in all regions and at all scales. Aware of that fact, some studies did not look for the best individual data set, but the best combination of data sets towards budget closure of the water cycle over one (Azarderakhsh et al., 2011) or multiple basins (Lorenz et al., 2014). It should be pointed out that the above studies' success metric was not budget closure itself, but validation versus in situ runoff instead.

The paradigm of quantifying the global water cycle is steadily shifting from identifying the best data source per water cycle component into developing the best way to merge data from various sources to complement each other. Various integration methodologies have emerged, among the most widely used ones are: bayesian model averaging, constrained linear regression, neural networks, optimal interpolation, and simple weighting (Bishop, 1996; Hoeting et al., 1999; Rodgers, 2000; Aires et al., 2004). Also, post-processing closure methodologies, which distributed the budget residual R among the components based on each component's uncertainties, explored Monte Carlo applications and Kalman filter variations (Pan and Wood, 2006; Munier and Aires, 2018). Specifics vary from method to method, but, in general, combining different data sets consists of three steps. These steps are an initial assessment of the products to be combined, followed by the integration of the products, and finally, budget closure post-processing.

Data integration is not a new concept nor the methods mentioned above, but its implementation altogether with closure constraints into the quantification of the water cycle is. Sahoo et al. (2011) used 16 data sets (eight for precipitation, six for evapotranspiration, one for runoff, and one for total water storage) applying simple weighting integration over ten basins across the globe, determining

water cycle budget non-closure between $5-25\%$. Likewise, Pan et al. (2012) used eight data sets (four for precipitation, two for evapotranspiration, one for runoff, and one for total water storage) in 32 different basins. The authors focused on describing the uncertainty contribution of each component rather than focusing on budget closure, and found that, in general, most of the closure error comes from evapotranspiration.

To date, only a few studies have adopted multi-source data integration at the global scale (Rodell et al., 2015; Zhang et al., 2016; Munier and Aires, 2018). The differences between studies and their results reside either on the data sets selected or in the post-processing. Rodell et al. (2015), using six data sets (one for precipitation, three for evapotranspiration, one for runoff, and one for total water storage; table 2.3), reported a non-closure residual of less than 10%. The authors adopted the variational data assimilation algorithm of L'Ecuyer and Stephens (2002) and adjusted it to optimize the global water cycle budget closure at the annual scale. They reported $(116.5 \pm 5.1) \times 10^3$ km^3/year overland precipitation, $(70.6 \pm 5.0) \times 10^3$ km^3/year evapotranspiration, $(45.9 \pm 4.4) \times 10^3$ km^3/year runoff, $(403.5 \pm 22.2) \times 10^3$ km^3/year over ocean precipitation, and $(449.5 \pm 22.2) \times 10^3$ km^3/year evaporation. Note that the estimates reported by Oki and Kanae (2006) and Trenberth et al. (2011) lie within the above findings with the only two exceptions of overland precipitation from Oki and Kanae (2006) and runoff from Trenberth et al. (2011).

Table 2.3: Compiled from Rodell et al. (2015). P is precipitation, ET is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage.

Data source	Variable	Reference(s)
GPCP v2.2	P	Adler et al. (2003)
		Huffman et al. (2009)
Princeton ET	ET	Vinukollu et al. (2011b)
MERRA and MERRA-Land	ET	Rienecker et al. (2011)
		Bosilovich et al. (2011)
		Reichle (2012)
GLDAS	ET	Rodell et al. (2004)
University of Washington runoff	Q	Clark et al. (2015)
CSR RL05	Δ TWS	Chambers and Bonin (2012)
		Johnson and Chambers (2013)
		Tapley et al. (2004)

Zhang et al. (2016), using 14 data sets (five for precipitation, six for evapotranspiration, one for runoff, and two for total water storage; table 2.4), assessed the effect of different data sources in the estimation of the water cycle and its budget closure. By removing/replacing in situ observations, reanalysis products, model simulations, or satellite products before data integration, the authors observed that removing non-satellite sources worsens closure errors. Furthermore, as for satellite data sets, they indicate that budget closure error depends on the use of satellite-only data sets or

satellite-gauge combined data sets. Regardless of the combination of data sets, the budget could not be closed and, thus, a constrained Kalman filter was used, as developed by Sahoo et al. (2011). They reported a non-closure residual that ranges between $7.6 - 10.4\%$ when using satellite products that lack gauge-based corrections, which is reduced to $4.2 - 9.0\%$ when using gauge-corrected satellite products.

Table 2.4: Modified from Zhang et al. (2016). P is precipitation, ET is evapotranspiration, Q is runoff, and TWS is total water storage.

Data source	Variable	Reference(s)
CSU	P	Bytheway and Kummerow (2013)
PGF	P	Sheffield et al. (2006)
CHIRPS	P	Funk et al. (2014)
GPCC(v6)	P	Schneider et al. (2014)
TMPA-RT	P	Huffman et al. (2007, 2010)
SRB-PGF-PM	ET	Vinukollu et al. (2011a)
VIC	ET	Sheffield and Wood (2007)
ERA-interim	ET	Simmons (2006)
MERRA	ET	Rienecker et al. (2011)
GLEAM	ET	Miralles et al. (2011)
SRB-CFSR-SEBS	ET	Vinukollu et al. (2011a)
SRB-CFSR-PM	ET	Vinukollu et al. (2011a)
SRB-CFSR-PT	ET	Vinukollu et al. (2011a)
VIC	Q	Sheffield and Wood (2007)
VIC	TWS	Sheffield and Wood (2007)
GRACE	TWS	Landerer and Swenson (2012)

Munier and Aires (2018) integrated 12 data sets (four for precipitation, three for evapotranspiration, one for runoff, and four for total water storage; table 5) over 11 basins to test a budget closure correction model. The authors define the Calibration Index for Closure (CIC), which depends on the values of precipitation minus evapotranspiration (P − ET) and the Normalized Difference Vegetation Index (NDVI), and based on the CIC values, assigned the basins into one of four classes. Then the closure correction model is calibrated to each basin using the corresponding CIC class, and it optimizes budget closure for the fluxes one at the time. While no absolute values are reported, the authors describe how this novel method reduced non-closure residuals by 26% of the value it would have using constrained Kalman filter post-processing.

In the above-mentioned studies, there is a methodological consensus to use simple weighting when integrating data from various sources. This is in good agreement with Aires (2014) who compared the performance of different integration methods, and reported that simple weighting is the most suitable one. Simple weighting offers a straightforward formulation, and more elaborate methods do not offer enough improvement on results to justify the increased complexity they carry along. The assumption for the simple weighting method is that the errors associated with the different products

Table 2.5: Modified from Munier and Aires (2018). P is precipitation, ET is evapotranspiration, Q is runoff, and Δ TWS is total water storage change.

Data source	Variable	Reference(s)
TMPA	P	Huffman et al. (2007)
CMORPH	P	Joyce et al. (2004)
NRL	P	Turk et al. (2010)
GPCP	P	Adler et al. (2003)
GLEAM	ET	Miralles et al. (2011)
MOD16	ET	Mu et al. (2007)
NTSG	ET	Zhang et al. (2010)
GRDC	Q	`http://www.grdc.sr.unh.edu/`
CSR	Δ TWS	`http://www2.csr.utexas.edu/grace/`
GFZ	Δ TWS	`ftp://isdcftp.gfz-potsdam.de/grace/`
JPL	Δ TWS	`https://grace.jpl.nasa.gov/data/get-data/`
GRGS	Δ TWS	`https://grace.obs-mip.fr/`

are Gaussian (zero-mean) and independent. However, there might be cases that this assumption may not hold, especially for gauge-based data products, and the dependence among products will cause an underestimation of the error associated with the integrated data set. The combined data set for a given component of the global water cycle (P, ET, Q, or Δ TWS) is equal to:

$$x = \sum_{i=1}^{n} w_i x_i \tag{2.4}$$

where x is the combined data set for the single component of the global water cycle being integrated, $x_1,\ x_2,\ x_3,\ ...,\ x_n$ are the different products considered, w_i is the associated weight of product x_i and is defined as:

$$w_i = \frac{(\bar{x} - x_i)^{-2}}{\sum_{j=1}^{n} (\bar{x} - x_j)^{-2}} \tag{2.5}$$

where $\bar{x}$ is the arithmetic mean of the n data products considered, and $(\bar{x} - x_i)^2$ is defined as the error variance. That is to say, the weight associated to each product is proportional to the inverse of its error variance. Finally, the error associated to the combined data set x is:

$$e_x = \frac{1}{\sum_{i=1}^{n} (\bar{x} - x_i)^{-2}} \tag{2.6}$$

2.3 Status Quo et Verisimile Futurum

It might have been noticed that the chronology of global water cycle quantification does not follow a linear timeline. The epochs started at different points in time without replacing the one before. Each

epoch did not only continue to develop, but just like global water cycle components, they interacted with each other in a feedback loop. A convergence point is the fact that model simulations and satellite-based measurements depend upon ground observations either for validation or calibration. The latest epoch, the age of big data, does not intend to merge all the previous into one, but to exploit the various data sources stemming from them to generate the most accurate estimates possible. Therefore, we should keep working on the continuous improvement of ground measurements, model simulations, and satellite observations, which will inherently improve their integration.

Abbott et al. (2019) provided one of the most recent descriptions of the global water cycle. Analogously to Oki and Kanae (2006), the authors did not quantify the global water cycle components themselves but synthesized data from the literature. The authors did not aim to quantify the components of the global water cycle but to assess its correct representation. To do so, they compiled over 464 diagrams (e.g., figure 2.3) and estimates from over 80 studies. Human interaction was absent in approximately 85% of the diagrams, highlighting the omission of the non-negligible anthropogenic component of the water cycle. In addition, the authors stress the necessity to represent seasonal and interannual variability of the global water cycle fluxes and storages in diagrams because the general understanding of temporal variability of the global water cycle is absent in the collective consciousness (Cardak, 2009). Within the studies, not all of them reported estimates for all components of the global water cycle. The synthesis resulted in the following estimates: overland precipitation 110×10^3 km^3/year, evapotranspiration 69×10^3 km^3/year, and runoff 46×10^3 km^3/year; over oceans, precipitation 380×10^3 km^3/year and evaporation 420×10^3 km^3/year.

Herein, building upon the previous compendium done by Baumgartner and Reichel (1972), we surveyed the recent literature, and to the best of our knowledge, compiled all the different estimates of global water cycle components available in peer review journals that at least report the average annual fluxes for the terrestrial or oceanic water cycle (Table 2.6). Since 2010 it has become more common for studies to address only the terrestrial water cycle (e.g., van der Ent et al. 2010; Haddeland et al. 2011; Jasechko et al. 2013; Zhang et al. 2018). On the other hand, ocean salinity measurements are being exploited to study the oceanic branch of the water cycle (Durack, 2015), yet there are very few studies focusing solely on the oceanic water cycle (e.g., Syed et al. 2010; Robertson et al. 2014; Gutenstein et al. 2021). Inspecting the chronology of global water cycle flux annual average estimates over land and over oceans, it is safe to state that uncertainty estimates associated with fluxes over oceans is higher than that over land (figures 2.4(a) and 2.4(b)). Comparing the standard deviation and the interquartile range of the estimates from Oki (1999) onward with the ones from all the estimates (1905-2019), we can affirm that variability has diminished in recent

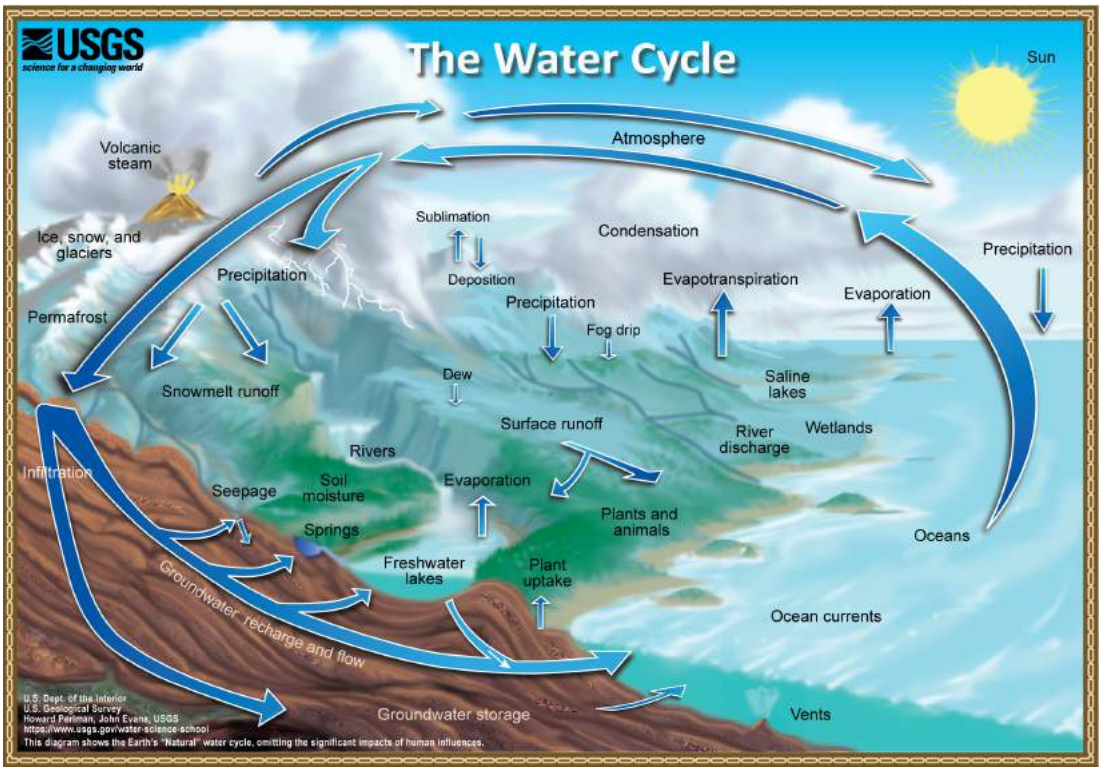


Figure 2.3: The Water Cycle. Credit: Howard Perlman, United States Geological Survey (USGS).

years (figures 2.4(c) and 2.4(d)). Moreover, the variability of ocean precipitation and evaporation was reduced by more than 70%. These findings advocate that the consistency of the estimates has been improved.

Despite our survey compiling estimates available in the literature rather than presenting a more "traditional" estimates' time series, we observe an increasing trend in the global water cycle fluxes annual average as the year of publication progresses (figure 2.5). We should remark that the years listed correspond to the publication date and do not necessarily reflect on the data sets' reference period used by the authors therein. Hence, our observations are of qualitative and not quantitative character. An increasing trend in global water cycle fluxes, commonly referred to as intensification, is often attributed to global warming; however, the processes that drive the global water cycle's response are yet to be fully understood (Allan et al., 2020). Take note that these estimates are global and do not describe changes in the water cycle at different smaller scales. On top of that, we should not assess these results conclusively because most studies used different data sources and different methods at different development stages, as discussed in the previous section. For example, if we were to look only at table 2.6 entries in figure 2.6 (from Baumgartner and Reichel (1972) onward), we would not be able to clearly discriminate a trend from the variability present in those estimates.

Table 2.6: All the fluxes are in 10^3 km^3/year. P_{L} is precipitation overland, ET is evapotranspiration overland, Q is runoff, P_{O} is precipitation over oceans, E is evaporation over oceans, P_{TOT} is total global precipitation, and E_{TOT} is total global evaporation.

Author	P_{L}	ET	Q	P_{O}	E	P_{TOT}	E_{TOT}
Manabe and Bryan (1969)	93.4	69.5	23.9	359.3	429	452.7	498.5
Baumgartner and Reichel (1972)	100	65	35	383	418	483	483
Falkenmark and Lindh (1974)	114	73	41	412	453	526	526
Speidel and Agnew (1982)	111	71	39.7	385	425	496	496
NRC (1986)	107	71	36	398	434	505	505
VanDerLeeden et al. (1991)	100	70	39.6	320	350	420	420
Gleick (1993)	119	72	47	458	505	577	577
Schmitt (1995)	110.4	69.4	41	384.7	425.7	495.1	495.1
Shiklomanov (1998)	119	74.2	42.7	458	502.8	577	577
Oki (1999)	115	75	40	391	431	506	506
Oki and Kanae (2006)	111	65.5	45.5	391	436.5	502	502
Schlosser and Houser (2007)	103.5	63	40.5	376	417	479.5	480
Trenberth et al. (2007)	113	73	40	373	413	486	486
Lim and Roderick (2009)	113	78.8	34.1	417.7	451.8	530.7	530.8
Syed et al. (2010)			36.1	374.2	409.2		
van der Ent et al. (2010)	117	82	35				
Chapin et al. (2011)	110	71	40	385	425	495	496
Haddeland et al. (2011)	126	72.5	54				
Trenberth et al. (2011)	114	74	40	386	426	500	500
Jasechko et al. (2013)	110	72.7	37.3				
Durack (2015)	110.4	85.1	39.4	384.7	410	495.1	495.1
Rodell et al. (2015)	116.5	70.6	45.9	403.5	449.5	520	520.1
Schneider et al. (2017)	117.6	71.8	45.8	386	431.8	503.6	503.6
Zhang et al. (2018)	114.7	68	46.6				
Abbott et al. (2019)	110	69	46	380	420	490	489

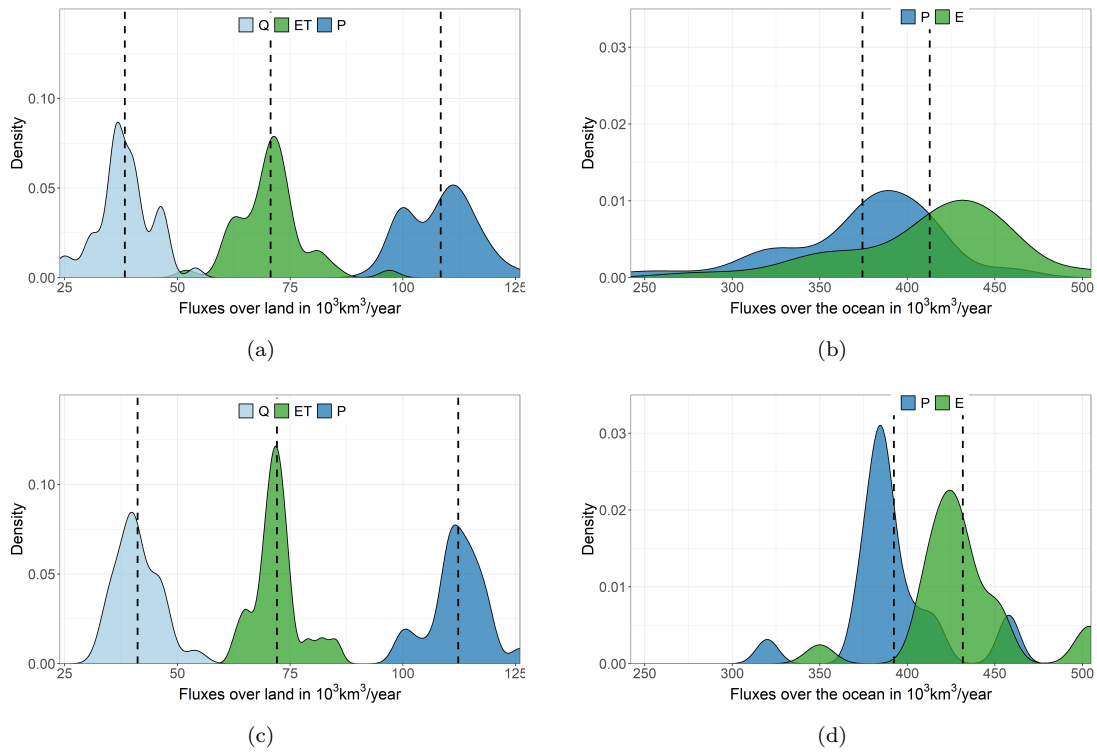


Figure 2.4: Probability density distribution of global water cycle fluxes from tables 2.1 and 2.6. The dashed line represents the mean value of each flux. (a) Overland fluxes where P is precipitation, ET is evapotranspiration, and Q is runoff. (b) Over ocean fluxes where P is precipitation and E is evaporation. (c) Same as (a) but only for table 2.6. (d) same as (b) but only for table 2.6.

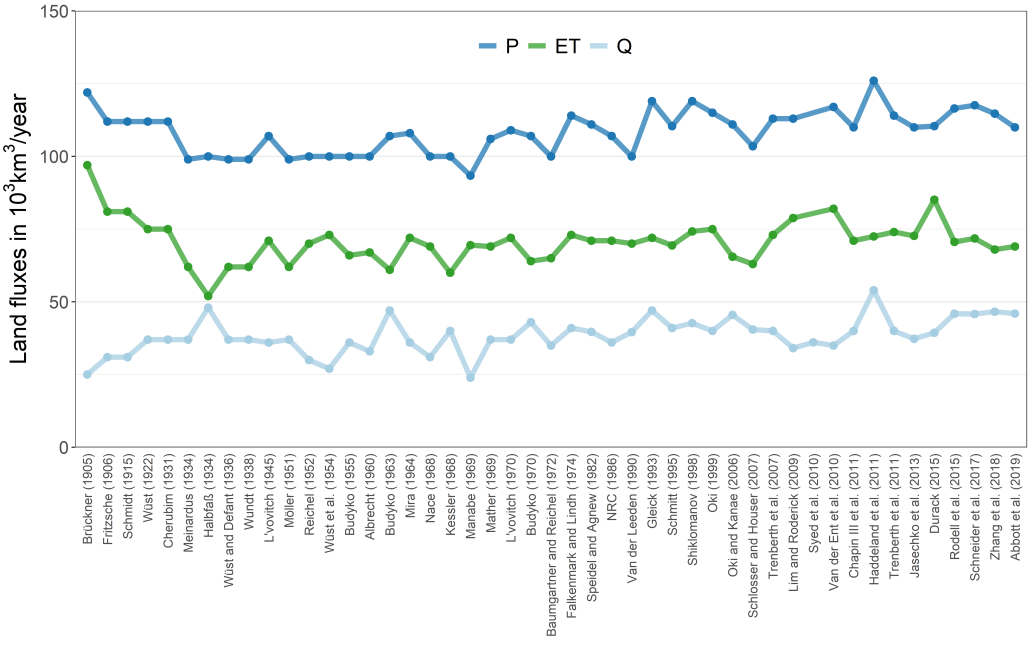


Figure 2.5: Chronological estimates of global water cycle fluxes over land in 10^3 km^3/year. P is precipitation, ET is evapotranspiration, and Q is runoff. The years listed correspond to the publication date and do not necessarily reflect the data sets' reference period used by the authors.

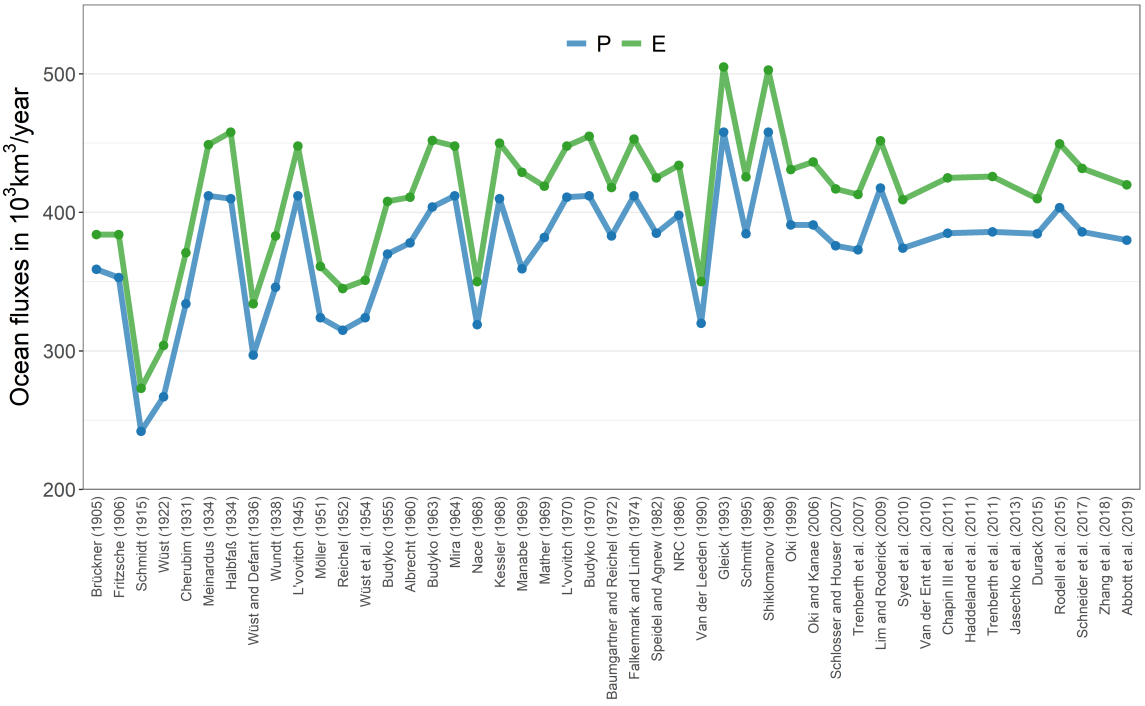


Figure 2.6: Chronological estimates of global water cycle fluxes over oceans in 10^3 km^3/year. P is precipitation and E is evaporation. The years listed correspond to the publication date and do not necessarily reflect the data sets' reference period used by the authors.

Moreover, suppose we were to omit the estimates reported between VanDerLeeden et al. (1991) and Shiklomanov (1998), there seem to be minor oscillations around an overall flat trend, attesting the narrative is dependent on the data being observed. Latch onto the ratio between evapotranspiration and precipitation over land, also known as the Evaporative Index (ET/P; figure 2.7), and it is interesting to see how, despite some clear multiannual oscillations, there seems to be no sharp trend. The Evaporative Index is the fraction of available water consumed by evapotranspiration (Budyko, 1974), and assuming no significant change in total water storage, its residual (1− ET/P) could be inferred as the fraction that turns into available freshwater. This, at least on paper, would suggest global freshwater availability has not diminished on average.

Through the previous sections, we have described how our understanding of the global water cycle has been evolving over the years as we exploit novel technologies and methods to quantify the components of the global water cycle more accurately. Accordingly, to assess future changes in the global water cycle and its response to global warming, we should study both past shifts documented in observational records and possible changes predicted by model simulations. While there are inherent fluctuations in the global water cycle, some of them are driven by natural phenomena like variations in the sun and volcanic eruptions (e.g., the year without a summer; Stommel and Stommel

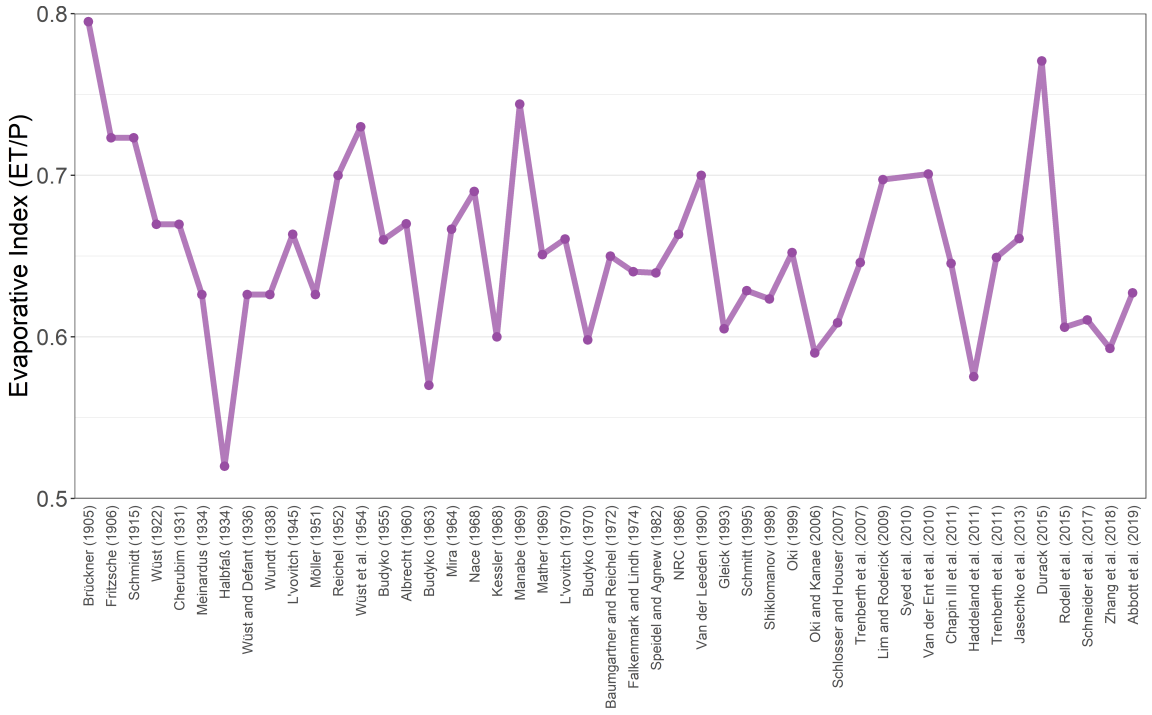


Figure 2.7: Chronological estimates of the evaporative index. Which is defined as the ratio between evapotranspiration and precipitation overland (ET/P). The years listed correspond to the publication date and do not necessarily reflect the data sets' reference period used by the authors.

1979), and anthropogenic activities. The latter exerts a continuously increasing influence directly via interference with land surface and water consumption, and indirectly via greenhouse gases and aerosols emissions (Abbott et al., 2019). The nature of the driver and the spatial scale they exercise domain over, alter key water cycle characteristics, e.g., precipitation frequency, intensity, or duration (Pendergrass and Hartmann, 2014).

Concurrently, model simulations predicted that global mean precipitation would rise in response to CO_2 doubling (Mitchell et al., 1987). The relationship between climate and water cycle caught the attention of both climatic and hydrological communities (Chahine, 1992b; Loaiciga et al., 1996). Models and the relationship between climate and water cycle are continuously evaluated in the Assessment Reports of the Intergovernmental Panel on Climate Change (IPCC; Collins et al. 2013; Flato et al. 2013). The Clausius-Clapeyron expression for the saturation vapor pressure establishes that at temperatures typical of the lower troposphere, the water holding capacity increases by about 7% for each 1K increase in temperature. It is safe to assume that an increase in lower-tropospheric water vapor will lead to a chain reaction affecting the entire global water cycle. The poorly understood response of the global water cycle resulted in two main hypotheses: the "changing character of precipitation" and the "dry gets drier, wet gets wetter". The former shows that the increase

in global mean precipitation will be unevenly distributed in precipitation events (Trenberth et al., 2003). Heavy or extreme rainfall will become more frequent, while light or moderate precipitation will decline. The latter suggests that the increased vertical gradient of atmospheric water vapor would offset atmospheric wind convergence in the tropics making wet regions wetter and dry regions drier (Roderick et al., 2014). Both hypotheses are today under vigorous debate (Held and Soden, 2006; Seager et al., 2010; O'Gorman and Muller, 2010; Greve et al., 2014; Roderick et al., 2014; Byrne and O'Gorman, 2015; Kumar et al., 2015; Salzmann, 2016; Skliris et al., 2016; Wang et al., 2017; Markonis et al., 2019; Allan et al., 2020).

Global precipitation and evapotranspiration, however, are further associated with Earth's energy budget rather than the Clausius-Clapeyron equation (O'Gorman et al., 2012; Roderick et al., 2014). Model simulations report that in response to global warming, global precipitation and evapotranspiration, independently of climate forcing, would increase constrained by Earth's energy budget to an expected rate between 2-3%/K (Samset et al., 2018). Precipitation's response to global warming, also known as apparent hydrological sensitivity, comprises a fast reaction proportional to radiative forcings and a slow temperature-dependent response to the radiative forcings (Bala et al., 2010). Across multiple model simulations, precipitation increases with global warming are generally suppressed over land compared to the global mean (0.8-2.4%/K vs. 2.3-2.7%/K), a behavior partly expected due to limitations on moisture convergence product of the more significant warming over land than oceans (Richardson et al., 2018). Considering that global precipitation's response to global warming is slower than the response of atmospheric water vapor, atmospheric water vapor lifetime must increase to reconcile these different response rates (Hodnebrog et al., 2019). By doing so, regional characteristics of precipitation such as seasonal duration, frequency, and intensity are altered (Pendergrass, 2018).

As atmospheric water vapor content increases and its lifetime prolongs, the increased horizontal moisture transport induces an intensification of precipitation minus evapotranspiration patterns. Over the continents, precipitation minus evapotranspiration is positive and accounts for the freshwater flux from the atmosphere to the surface, whereas over the ocean, precipitation minus evaporation is negative and represents the freshwater flux from the oceans to the atmosphere. In dry regions, where evapotranspiration is constrained by water availability, changes in precipitation minus evapotranspiration will be mainly credited to precipitation changes (Roderick et al., 2014). Precipitation minus evapotranspiration over land can be negative during dry seasons or extended drought periods (Kumar et al., 2015). Given that evapotranspiration is a compound flux of evaporation and transpiration, the response of vegetation to global warming and increased CO_2 concentrations in the

atmosphere will also determine the characteristics of regional precipitation minus evapotranspiration patterns. Besides, over land, we cannot neglect anthropogenic activities like irrigation, land-use change, deforestation, urbanization, and water withdrawals, among others that directly alter precipitation minus evapotranspiration regimes. On this account, we can expect several factors like topography, atmospheric circulation, anthropogenic tampering, and vegetation response to generate different and complex water cycle responses to global warming.

2.4 Discussion and Conclusions

Early attempts to quantify the global water cycle date back to the early 1900s (Brückner, 1905). To date, despite tremendous advances in terms of data and technology, accuracy regarding the components of the global water cycle has not increased accordingly. Ultimately, unquantified uncertainties on remote sensing satellite products (Sheffield et al., 2009), limitations of climate model simulations (Trenberth et al., 2011), short and heterogeneous observational data records (Schneider et al., 2017), and the natural fluctuations of water cycle components Markonis et al. (2018) keep the understanding of the global water cycle ambiguous and human contribution unattributed. Within the twenty-first century, the paradigm of quantifying the global water cycle has been shifting from identifying the best data source per water cycle component into developing the best way to integrate data from various sources (Aires, 2014). Therefore, proper statistical tools for uncertainty quantification (Papalexiou, 2018), robust downscaling/disaggregation (Papalexiou et al., 2018), along with analysis over multiple scales (Hanel et al., 2017; Markonis et al., 2021b) are required.

The quest for accurate global water cycle quantification gave birth to the Global Energy and Water Exchanges (GEWEX) project. The GEWEX project, formerly known as the Global Energy and Water Cycle Experiment, started in 1990 and is dedicated to studying the Earth's water and energy cycles (Chahine, 1992a). GEWEX established a channel for international research collaboration through different panels, meetings, and projects. Among the most renowned outcomes, we could mention the work of Trenberth et al. (2007), which we further discussed in section 2.2.3. Speaking of data sets and modeling improvements, GEWEX overlooks eight continental-scale experiments: GEWEX Americas Prediction Project (GAPP; Lawford 1999), Baltic Sea Experiment (BALTEX; Raschke et al. 1998, 2001), GEWEX Asian Monsoon Experiment (GAME; Yasunari 1994), Large Scale Biosphere Atmosphere Experiment in Amazonia (LBA; Marengo 2005), Mackenzie GEWEX Study (MAGS; Stewart et al. 1998), La Plata Basin (LPB; Cavalcanti et al. 2015), The African Monsoon Multidisciplinary Analysis (AMMA; Redelsperger et al. 2006), and Murray-Darling Basin

(MDB; Evans and McCabe 2010). Other than the logistic and political criteria, these sites were selected in order to collect data from different climate regimes to assess the global water cycle in a representative manner. The collaborative effort of the international teams involved improved the understanding of regional water balance and feedback processes. The data resulting from the continental-scale experiments are publicly available. Thus, they indirectly started to set up a scientific framework to quantify the global water cycle and close its budget; the latter was obtained within a 10% non-closure tolerance.

As a rule of thumb, ground observations are regarded as the closest measurements to the actual values. However, it is evident that ground observations suffer from systematic errors, mainly because of different environmental and meteorological conditions. For example, the precipitation phase, evaporation from the gauge, and wind drift induce precipitation undercatch on rain gauges (Fuchs et al., 2001). The scientific community is aware that good quality ground observations data represent a cornerstone to quantify the global water cycle, yet we are still unable to deploy a homogeneously distributed global network. Spatial coverage of the Global Precipitation Climatology Centre (GPCC), currently the most comprehensive gauge network available, represents only about 1% of the Earth's surface (assuming no overlap of a 5 km radius per gauge) (Kidd et al., 2017). One of the main reasons behind the struggle to deploy a comprehensive network is that ground stations, and ergo observational data records, are extremely geopolitically dependant (Kibler et al., 2014). In addition, deploying dense monitoring networks unavoidably imply high operational and maintenance costs and spatial requirements (Saltikoff et al., 2017). Consequently, in many developing countries, ground observational records, if available, tend to have multiple temporal discontinuities or non-standardized data quality check protocols (Walker et al., 2016). Different techniques have been used to fill spatiotemporal gaps in observational records. Reconstructing these time series could be achieved using several tools that could be grouped in the following, self-contained infilling (Kemp et al., 1983; Pappas et al., 2014), spatial interpolation (Shepard, 1968; Young, 1992; Eischeid et al., 1995, 2000), quantile mapping (Simolo et al., 2010; Newman et al., 2015, 2019; Devi et al., 2019), and machine learning methods (Dastorani et al., 2010; Wambua et al., 2016). On a different front, there is an opportunity to use data from amateur networks and the internet of things (i.e., big data with large uncertainty) to enhance spatial coverage and spatiotemporal resolution of traditional ground stations via crowdsourcing and the internet. Needless to say, appropriate validation and quality control procedures must be adopted and implemented to fully exploit the potential to provide a valuable source of high spatiotemporal resolution real-time data (Muller et al., 2015). As of now, however, the lack of adequate ground-based data and station networks still hampers our

ability to monitor the water cycle robustly.

Model simulations can generate past climate, current climate, and climate projections data. Moreover, they are capable to switch anthropogenic forcing on precipitation on and off, while the decoupling of natural and anthropogenic forcing remains a challenge on observational data (Allen and Ingram, 2002). However, compared to observational data, various characteristics of global water cycle fluxes, and precipitation, in particular, hold uncertainty (Prein and Pendergrass, 2019). The simulated projections' temporal length appears to influence precipitation trends, e.g., variability in precipitation estimates are indistinguishable from the noise of internal variability in 20-year or longer runs (Hawkins et al., 2016). Specifications differ from model to model, but in general, recycling of moisture is too large, and the lifetime of moisture is too short across most models, inducing premature precipitation (Trenberth et al., 2011). Also, inaccurate convective parameterizations evidenced that models overestimate precipitation frequency and underestimate its intensity (Trenberth et al., 2017). Analysis focusing on convective precipitation highlighted that its model representation is strongly dependant on the model depiction of cloud microphysics and cloud spatiotemporal variability (Zhao et al., 2016). There is a threefold spread in mean precipitation change with global temperature ($1 - 3\%\ K^{-1}$), and model simulations showed that there is a correlation between an increase in precipitation extremes and an increase in model resolution, precipitation extremes at the same time showed an anticorrelation with changes in light-moderate precipitation (Thackeray et al., 2018). Furthermore, both the spread and magnitude of change in extreme precipitation vastly exceed those of mean precipitation ($4-10\%\ K^{-1}$) (Kharin et al., 2013). Last but not least, despite the known link between the energy and water global cycles, solar dimming and brightening (the effect of aerosols) are not well represented or sometimes not even considered at all in models; thus, model simulations fail to reproduce variability in the global water cycle intensity (Wild and Liepert, 2010).

Satellite remote sensing observations, like models, are limited by their design. Both the orbit they follow and the instrument type (i.e., active or passive) influence global water cycle components' monitoring. The satellite's orbit would delimit its spatiotemporal resolution or coverage. In general, a satellite with high spatial resolution comes with coarse temporal resolution and vice-versa, and high spatiotemporal resolution comes with limited coverage. It has been shown that estimates from active sensors can considerably vary from passive sensor ones, yet they complement each other (Petković and Kummerow, 2017). In addition, similarly to ground observations, satellite remote sensing has to deal with different meteorological conditions. For instance, satellite-based global water cycle estimates accuracy is affected by cloud-top reflectance and thermal radiance, making uncertainty larger during the winter or in dry climates (Kummerow et al., 2004). While satellites

can monitor the water cycle at the global scale and cover regions inaccessible by ground stations, they still have to tackle the problems involved in complex topography regions. In some cases, the relative biases reach as much as 300% for precipitation estimates (Fekete et al., 2004). Further complications arise from the unique spatiotemporal characteristics of different remotely sensed global water cycle components, making it impossible to assess the water budget without some sort of prior downscaling or integration (Sheffield et al., 2018). E.g., TMPA's precipitation at 25 km every three hours (Huffman et al., 2007), MODIS' evapotranspiration at 1 km daily (Mu et al., 2007), and GRACE's total water storage at $\sim$ 500 km every 30 days(Tapley et al., 2004). Despite all the issues mentioned above, satellite products continue to be the most widely used sources to monitor global water cycle components due to their comprehensive spatial coverage.

It is clear that no global water cycle data source is without fail, and in some cases, one data source strengths cover for other weaknesses. It is typical for satellite-based measurements and model simulations to use ground-based data for validation, calibration, and enhancement purposes. Along the same line, model simulations additionally assimilate satellite-based observations for the above plus for reanalysis. In contrast to the top-down estimation approach used in satellite remote sensing, a bottom-up approach, referred to as reverse hydrology, has been recently proposed (Ciabatta et al., 2020). A physically-based selection of surface explanatory variables, like soil moisture, vegetation cover, and topography, is expected to preserve process dynamics and interlinkages within data sets that remain unresolved in conventional statistical downscaling bias-correction methods (Wehbe et al., 2020). It is of utmost importance that the research community strives to improve ground observations, model simulations, and satellite remote sensing measurements individually because more accurate and robust individual data sources will subsequently refine the outcome of multi-source integration. Hence, a three-way integration of satellite remote sensing, model reanalysis, and ground-based measurements, as discussed in section 2.2.4, is widely acknowledged as the current best practice, particularly when leveraging machine learning tools to handle large data sets.

Chapter 3

Water Cycle Changes in Reanalyses: A Complementary Framework

3.1 Introduction

Understanding the global water cycle and its balance is crucial for Earth system science and climate change studies. To assess the water cycle at multiple spatiotemporal scales, we observe and measure the fluxes and storage that comprise its budget. The data sources we rely on for such research have continuously evolved, even though they remain thwarted by uncertainty (Vargas Godoy et al., 2021). Ground observations are regarded as the closest measurements to the actual values, but we still lack a comprehensive global network. E.g., the Global Precipitation Climatology Centre (GPCC) (Schneider et al., 2011), currently the most extensive gauge network available, represents only about 1% of the Earth's surface (assuming a 5[km] non-overlapping radius per gauge) (Kidd and Huffman, 2011). While a kindred initiative for evaporation exists (FLUXNET) (Pastorello et al., 2020), evaporation is more commonly derived from atmospheric moisture and precipitation measurements than directly observed. Satellite remote sensing data complement ground measurements by offering observational coverage on a global scale. Its record, nonetheless, is too short to assess long-term changes of water cycle fluxes.

Reanalysis products assimilate observational data into general circulation models or, most re-

cently, earth system models. Broadly speaking, assimilation algorithms recursively combine the model outputs and observations within a Bayesian statistical framework. As a result, the physical conservation principles are overstepped, which is reflected in substantial variability compared to that of other data sources (Prein and Pendergrass, 2019). Since their early implementation, concerns about the reliability of reanalyses to assess the global water cycle have been raised (Trenberth and Guillemot, 1995). Despite individual advancements in model simulations as well as assimilation algorithms, it has been reported that moisture recycling is too large and its lifetime too short (Trenberth et al., 2011). Regardless, reanalyses remain one of the most comprehensive data sources because of their high spatiotemporal resolution and capability to switch anthropogenic forcing on and off (Allen and Ingram, 2002).

Therefore, it is no surprise that they have been used in the estimation of water cycle fluxes and their changes. A prime example is the work of Trenberth et al. (2007), therein the authors describe the global water budget and its annual cycle. While the authors have reservations about reanalysis-based results, they acknowledge the potential for in-depth analysis using reanalyses. It is worth mentioning that the work of Trenberth et al. (2007) is in high regard by the scientific community, and their results are often used to benchmark more recent studies. During the last decade, global water cycle research has explored multi-source data integration, exploiting observational and reanalysis data availability. Specifics between methods vary, but, in general, data sets are merged in three steps: initial assessment of the data, integration of the products, and budget closure post-processing.

Some examples at the global scale are the works of Rodell et al. (2015) and Zhang et al. (2016), where the authors use multiple reanalysis evaporation/evapotranspiration products to assess the water cycle and budget closure. Rodell et al. (2015) relied on reanalyses at various other stages of their analysis, such as data sources for other variables (e.g., atmospheric convergence, wind, and surface pressure), to downscale observations, and to fill data gaps. The authors convey that independent reanalysis estimates enable assessing uncertainty with a higher degree of confidence. Zhang et al. (2016) studied the influence of data sources on water budget closure experiments and concluded that integrating reanalysis data reduces the non-closure errors significantly. Yet, further efforts are needed to understand the discrepancies among different data sources.

In this study, the representation of global water cycle changes is assessed in four reanalysis data sets for the first time. To achieve it, we physically define precipitation plus evaporation to unveil hidden details that have been overlooked due to the lack of a more exhaustive framework. We assessed the following data sets: 20CR v3 (Slivinski et al., 2019), ERA-20C (Poli et al., 2016), ERA5 (Hersbach et al., 2020), and NCEP1 (Kalnay et al., 1996). First, we compare the reanalyses using

ground-based data as a reference GPCP v2.3(Adler et al., 2018) and HadCRU5T5 (Morice et al., 2021) for precipitation and temperature, respectively. Then, we inspect $P - E$ to check for budget closure, and we evaluate the $P+E$ behavior in terms of hydrological sensitivity. Next, we present the application of $P + E$ in a framework that describes the changes in the water cycle. We achieve this by exploring the changes in atmospheric water fluxes and storage redistribution between land-ocean and the atmosphere. Finally, we discuss the possible connotations of the findings regarding $P + E$ and its application as a performance metric for reanalysis data.

3.2 The Physical Basis

Over land, the net water flux into the surface, a vital aspect of the water cycle for human society, is described by the difference between precipitation and evaporation ($P-E$). Thus, $P-E$ characterizes atmosphere-land surface interactions and represents the maximum available renewable freshwater (Oki and Kanae, 2006). Analogously, evaporation minus precipitation ($E - P$) determines the surface salinity of the ocean, which helps determine the stability of the water column (Cheng et al., 2020). There was a consensus that as precipitation increases overland, so does evaporation over the oceans to balance the global water cycle (Held and Soden, 2006). Nonetheless, it has recently become evident that there are contrasting responses between the terrestrial and oceanic water cycles (Sherwood and Fu, 2014; Byrne and O'Gorman, 2015). Furthermore, at the regional scale moisture convergence can increase precipitation (Espinoza et al., 2018). Assuming radiation is not limiting, evapotranspiration will be equally enhanced. On the one hand, $P - E$ would suggest no change in the hydrological cycle, while, on the other hand, the increase in $P + E$ would correctly indicate that the water cycle is indeed changing, with more water being circulated in total through the surface-atmosphere continuum.

Huntington et al. (2018) have already shown that the sum of precipitation and evapotranspiration can be adequately applied to quantify the changes in the terrestrial portion of the water cycle. We argue that this approach can be extended to the description of the whole water cycle because $P + E$ has a robust physical meaning; it describes the total flux of water exchanged between the atmosphere and the surface. Furthermore, like the human heart, the Earth cycles far more water through the atmosphere than its holding capacity. In this manner, it would make sense to also look into the addition of fluxes rather than only their difference when assessing the global water cycle intensification. The proposed framework is based on quantifying precipitation, evaporation, their difference, and their sum. The latter, precipitation plus evaporation, is mathematically complemen-

tary to the widely used $P - E$ metric. Nonetheless, math alone does not suffice to improve our understanding of the global water cycle. Thus, we will define $P + E$ from a mass balance and a kinematic perspective.

3.2.1 Water Cycle Budget

The global water cycle's mass balance is expressed with the water budget equation:

$$P + Q_{\text{in}} = E + Q_{\text{out}} + \Delta S \tag{3.1}$$

where P is precipitation, Q_{in} is water flow into the Earth, E is evaporation (since we are at the global scale we will refer to it simply as evaporation for brevity, but we acknowledge it encompasses evaporation from soils, surface-water bodies, and plants), ΔS is water storage change in the land-ocean continuum (biological water, fresh lakes, ice, nonrenewable groundwater, oceans, permafrost, reservoirs, renewable groundwater, rivers, saline lakes, seasonal snow, soil moisture, and wetlands), and Q_{out} is water flow out of the Earth. All terms are averaged globally over a fixed time period (e.g., $[mm/yr]$). At the global scale, due to Earth's gravity and temperature, water inflow or outflow leaking between the atmosphere and outer space is negligible compared with precipitation and evaporation and water storage change. Consequently, $Q_{in} \to 0$ and $Q_{out} \to 0$ leaving us with:

$$\Delta S = P - E \tag{3.2}$$

where ΔS represents a storage redistribution from the atmosphere towards the land-ocean continuum (positive), from the land-ocean continuum towards the atmosphere (negative), or steady state equilibrium (zero). Now, we define global water cycle intensity as:

$$\text{GWCI} = P + E \tag{3.3}$$

In this manner, intensity is defined as the total total flux of water exchanged between the atmosphere and the land-ocean continuum. This definition is in line with previous formulations in the literature (Huntington et al., 2018; Weiskel et al., 2007). Furthermore, different ways to integrate precipitation and evaporation to describe the hydroclimatic regime have been in use for over half a century now (e.g., Budyko curve; Budyko 1974).

3.2.2 Water Cycle Kinematics

As established above, precipitation plus evaporation describes the water cycle intensity from a mass balance perspective by quantifying the total flux of water exchanged between the atmosphere and the land-ocean continuum. If we describe these atmospheric water fluxes from a kinematic perspective, we have two velocity vectors:

$$\begin{aligned} \vec{P}_{lon,lat} &= \mathbf{P}(x, y, z) \\ \vec{E}_{lon,lat} &= \mathbf{E}(x, y, z) \end{aligned} \tag{3.4}$$

where, at any location on Earth's surface, $\vec{P}_{lon,lat}$ is the precipitation vector with magnitude $\mathbf{P}$ and $\vec{E}_{lon,lat}$ is the evaporation vector with magnitude $\mathbf{E}$. These velocities are parallel to each other but are oriented in opposite directions. We define the direction from the atmosphere to the surface as positive and the opposite (from the surface to the atmosphere) as negative, then:

$$\begin{aligned} \vec{P}_{lon,lat} &= \mathbf{P}(0\hat{i}, 0\hat{j}, 1\hat{k}) \\ \vec{E}_{lon,lat} &= \mathbf{E}(0\hat{i}, 0\hat{j}, -1\hat{k}) \end{aligned} \tag{3.5}$$

Precipitation and evaporation are heavily intertwined through moisture recycling. Therefore, we could characterize their interdependence relationship by defining the velocity of the global water cycle as the Newtonian relative velocity of precipitation with respect to evaporation:

$$\begin{aligned} \overrightarrow{GWC}_{lon,lat} &= \vec{P}_{lon,lat} - \vec{E}_{lon,lat} \\ &= \mathbf{P}(0\hat{i}, 0\hat{j}, 1\hat{k}) - \mathbf{E}(0\hat{i}, 0\hat{j}, -1\hat{k}) \\ &= (0-0)\hat{i} + (0-0)\hat{j} + (\mathbf{P} - (-\mathbf{E}))\hat{k} \\ &= 0\hat{i} + 0\hat{j} + (\mathbf{P} + \mathbf{E})\hat{k} \\ &= (\mathbf{P} + \mathbf{E})(0\hat{i}, 0\hat{j}, 1\hat{k}) \end{aligned} \tag{3.6}$$

where **(P + E)** is the magnitude of global water cycle velocity. Hence, we can safely ascertain that assessing changes in $P + E$ refers to acceleration or deceleration of the global water cycle.

3.3 The Precipitation-Evaporation Space

Including precipitation, evaporation, their difference, and their sum provides a synthesized visual of the overall response of the water cycle to global warming. The global water cycle regimes in

this framework would be described in the precipitation-evaporation space by their precipitation and evaporation coordinates, and vectors represent changes between two periods (Figure 3.1). By transforming the changes in the relationship of P and E to changes in $P - E$ and $P + E$, we can describe the water cycle dynamics in terms of atmospheric water storage and fluxes correspondingly. Precipitation and evaporation may increase, decrease, or remain constant. From equation (3.2), changes in atmospheric water storage ($P - E$) shown as blue contours are planes that increase from the bottom right (wetter) to the top left corner (drier). It is important to note that Huntington et al. (2018) focused on terrestrial water storage, as such, the directions for drier and wetter are reversed therein. From equation (3.3), water cycle acceleration ($P + E$) is a plane shown as green contours that increases from the bottom left (cooler) to the top right (warmer). $P - E$ is negative to the right of the identity diagonal, zero along this line, and positive to the left of the line. At the global scale, negative values describe an increase in atmospheric water storage (wetter), positive values describe an increase in land-ocean water storage (drier), and zero describes steady-state equilibrium. $P + E$ increases describe shifts from cooler regimes into warmer ones.

3.4 Results

3.4.1 Climate Reanalyses

Our analyses, taken together, show the potential of precipitation plus evaporation to assess reanalysis data and complement water cycle changes research. We start by exploring precipitation and temperature as portrayed in reanalyses with GPCP v2.3 and HadCRUT5 as the existing references. The variability from reanalysis precipitation becomes readily visible by the wide spread of values (Figure 3.2a). We observe an abrupt reduction in reanalysis precipitation variance after the mid-1960s (narrowing of the gray area; Figure 3.2a), coinciding with the satellite era's beginning. To a greater or lesser extent, all reanalyses products overestimate precipitation, with NCEP1 having its 30-year average closest to GPCP v2.3 (Figure 3.2b). Nevertheless, ECMWF reanalyses perform better than the 20CR v3 and NCEP1 (0.4 vs. 0.1 R-squared; Figure 3.2c). Regarding temperature, there is considerably less variability among reanalyses and no visible abrupt changes in said behavior (Figure 3.3a). Concurrently, temperature in reanalyses is centered around the 14[$^{\circ}C$] average (Figure 3.3b). Furthermore, all reanalysis products exhibit a strong and statistically significant correlation to HadCRUT5 (R-squared $\gtrsim$ 0.9; Figure 3.3c). Overall, ERA5, with the highest R-squared values, most comprehensively captures both precipitation (R-squared 0.43) and temperature (R-squared 0.97)

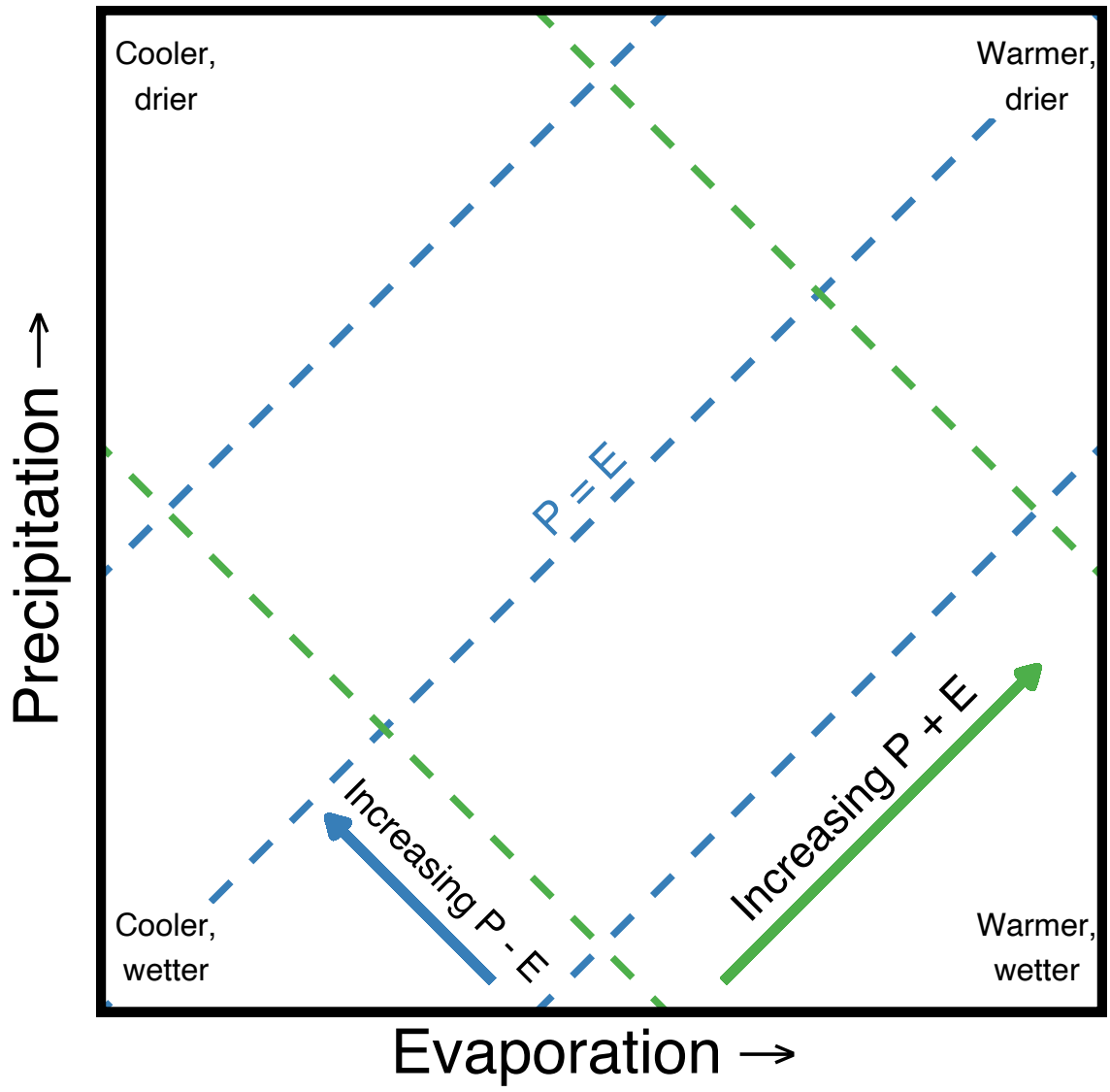
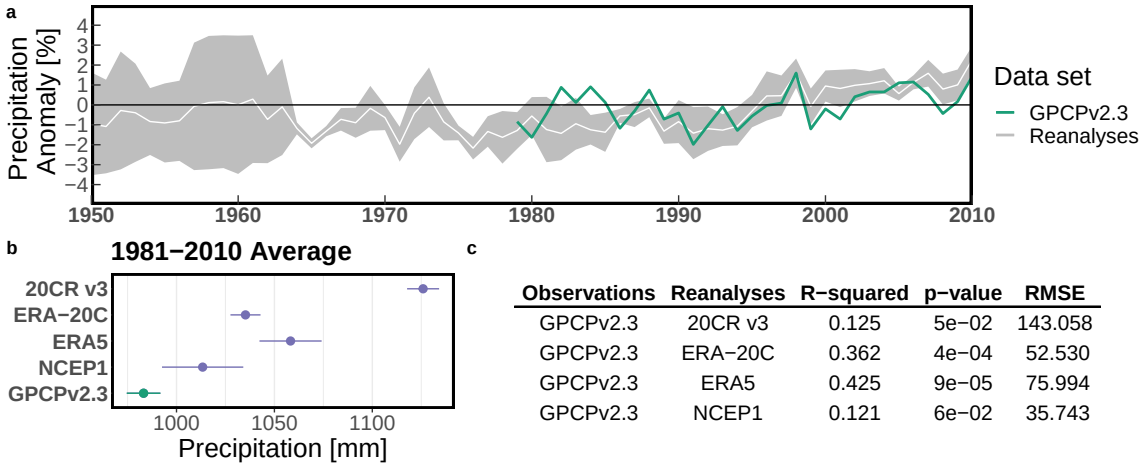


Figure 3.1: The global water cycle regime in the precipitation-evaporation space. Vectors represent water cycle changes, where P is precipitation, and E is evaporation. Contours of equal $P - E$ (no change in water cycle storage) are shown as blue dashed lines, and movement across these lines (blue vector) describe changes in water cycle storage. Contours of equal $P + E$ (no change in water cycle intensity) are shown as green dashed lines, and movement across these lines (green vector) describe changes in water cycle intensity.

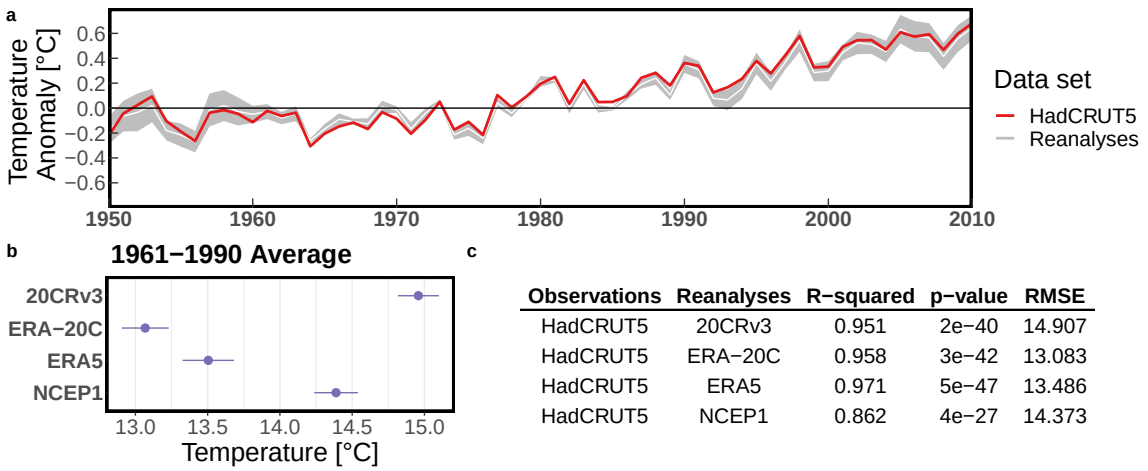
changes among the four reanalyses. While no direct assessment of reanalysis evaporation is possible due to the lack of observation-based reference data, and despite the general biases reported above, it is feasible to rely on reanalyses to assess global water cycle changes based on their performance versus precipitation and temperature observations.



Observations	Reanalyses	R−squared	p−value	RMSE
GPCPv2.3	20CR v3	0.125	5e−02	143.058
GPCPv2.3	ERA−20C	0.362	4e−04	52.530
GPCPv2.3	ERA5	0.425	9e−05	75.994
GPCPv2.3	NCEP1	0.121	6e−02	35.743

Figure 3.2: Benchmarking global spatial weighted average values of reanalysis precipitation compared to GPCP v2.3 as the observation-based reference. (A) Precipitation anomalies annual time series between 1950-2010 (common period between all reanalyses), spread of reanalysis estimates is shown in gray and their mean in white, GPCP v2.3 is shown in turquoise. (B) The 30-year average for the data sets compared, reanalysis estimates are shown in violet and GPCP v2.3 in turquoise. (C) Summary statistics of linear correlation between reanalysis products and GPCP v2.3 annual time series.

Atmospheric water residence time is circa nine days, and as previously stated, this lifetime is underestimated in reanalyses. Thus at annual or longer time steps, what goes into the atmosphere as evaporation has to equal what comes out as precipitation. Due to the assimilation algorithms and systematic uncertainty in reanalyses, we expected budget non-closure to some extent. However, it was surprising that even the 30-year moving average of $P - E$ in reanalyses is not steady (Figure 3.4). A gripping behavior in both ongoing reanalyses, i.e., ERA5 (Figure 3.4c) and NCEP1 (Figure 3.4d), is that the $P - E$ trend appears to be directed towards $0[mm/year]$ ($P = E$). In the case of the long-term reanalyses, we found opposing conducts. ERA-20C has the "flattest" $P - E$ mean at approximately $-5.5[mm/year]$ (Figure 3.4b). On the other hand, the 20CR v3 has considerably more variability and the highest $P - E$ absolute values (Figure 3.4a). A particular characteristic of 20CR v3 $P - E$ is that it seems to exhibit two regimes, one before 1900 centered around $-54[mm/year]$ and the second from 1900 onwards centered around $-69[mm/year]$.



Observations	Reanalyses	R−squared	p−value	RMSE
HadCRUT5	20CRv3	0.951	2e−40	14.907
HadCRUT5	ERA−20C	0.958	3e−42	13.083
HadCRUT5	ERA5	0.971	5e−47	13.486
HadCRUT5	NCEP1	0.862	4e−27	14.373

Figure 3.3: Benchmarking global spatial weighted average values of reanalysis temperature compared to HadCRUT5 as the observation-based reference. (A) Temperature anomalies annual time series between 1950-2010 (common period between all reanalyses), spread of reanalysis estimates is shown in gray and their mean in white, HadCRUT5 is shown in red. (B) The 30-year average for the data sets compared, reanalysis estimates are shown in violet (HadCRUT5 is not available). (C) Summary statistics of linear correlation between reanalysis products and HadCRUT5 annual time series.

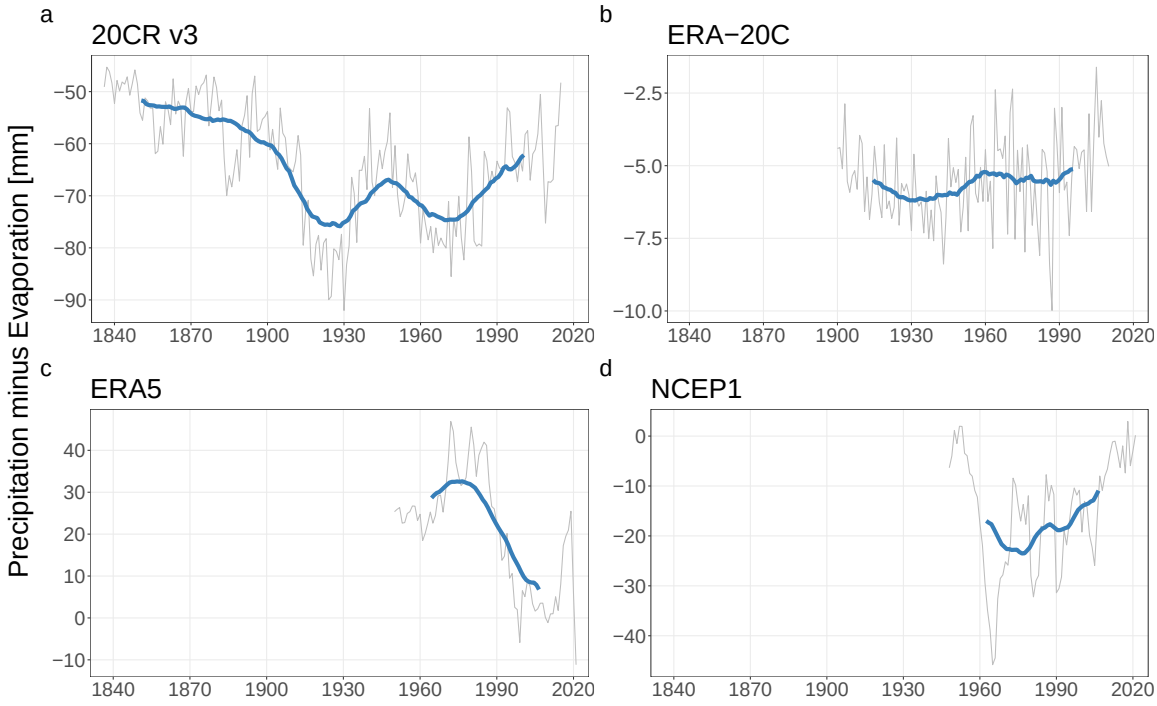


Figure 3.4: Global spatial weighted average of annual total precipitation minus evaporation in $[mm]$ as depicted in four reanalysis data sets for their respective available record. Annual values are shown in gray. 30-year moving average values are shown in blue.

3.4.2 Water Cycle Changes

The physical soundness of the $P + E$ metric becomes readily visible by the superimposition of the annual mean global temperature and the annual total global $P + E$ of the four reanalysis data sets (Figure 3.5). Their coupling is statistically supported by quantifying the linear relationship between these variables (Table 3.1). The dominant behavior in the long-term relationships reports two common markers: a strong $P+E$ correlation (R-squared ≈ 0.8; Figures 3.5a, 3.5b, and 3.5c), and an apparent decoupling between $P+E$ and temperature around the 1960s. We observe particular traits for ERA5 and NCEP1. ERA5 shows a moderate $P - E$ correlation (R-squared = 0.39). NCEP1, not resembling the other three data sets, has a higher correlation for the difference than the sum of precipitation and evaporation(0.18 vs 0.12 R-squared). Moreover, the coupling between $P + E$ and temperature occurs only after the mid-1970s (Figure 3.5d). The robust performance of $P + E$ as a metric to substantiate the relationship between atmospheric water fluxes and temperature carries from the long-term onto the year-to-year variability (Table 3.1). Estimating the annual differences, we now observe a homogeneous behavior in all the reanalyses data sets with moderate $\delta(P + E)$ correlation (R-squared between 0.2-0.4) and no $\delta(P - E)$ correlation (R-squared ≤ 0.02). This independence in $\delta(P - E)$ imply that the correlation observed between $P - E$ and temperature was due to the long-term trends, while $P + E$ correlates both to short-term and long-term temperature variability.

Table 3.1: Linear relationship between global spatial weighted average of total atmospheric water fluxes and mean temperature, where P is precipitation, E is evaporation, and T is temperature. Long-term columns report the correlation between the annual values (i.e., $(P \pm E)$ vs. T). Year-to-year columns report the correlation between the annual consecutive differences (i.e., $\delta(P \pm E)$ vs. δT).

	Long-term				Year-to-year			
	$P + E$		$P - E$		$\delta(P + E)$		$\delta(P - E)$	
Reanalysis	R^2	p-value	R^2	p-value	R^2	p-value	R^2	p-value
20CR v3	0.82	$< 2 \times 10^{-16}$	0.01	0.1	0.19	1×10^{-9}	2×10^{-4}	0.9
ERA-20C	0.80	$< 2 \times 10^{-16}$	0.06	1×10^{-2}	0.37	2×10^{-12}	0.02	0.2
ERA5	0.75	$< 2 \times 10^{-16}$	0.39	3×10^{-9}	0.35	3×10^{-8}	0.02	0.2
NCEP1	0.12	2×10^{-3}	0.18	2×10^{-4}	0.22	2×10^{-5}	4×10^{-3}	0.6

Thermodynamics, Clausius–Clapeyron scaling in particular, determine the relationship between atmospheric water vapor and temperature. However, it is the Earth's energy balance that governs global precipitation and evaporation, and constraining the hydrological sensitivity (Allan et al., 2020). The hydrological sensitivity, defined by the increase in global mean precipitation (or evaporation) for a given change in global mean temperature, has been estimated at 2.1–3.1 [$\%/^\circ C$] (Fläschner et al., 2016). Consequently, $P + E$ should also increase at approximately 2-3 [$\%/^\circ C$].

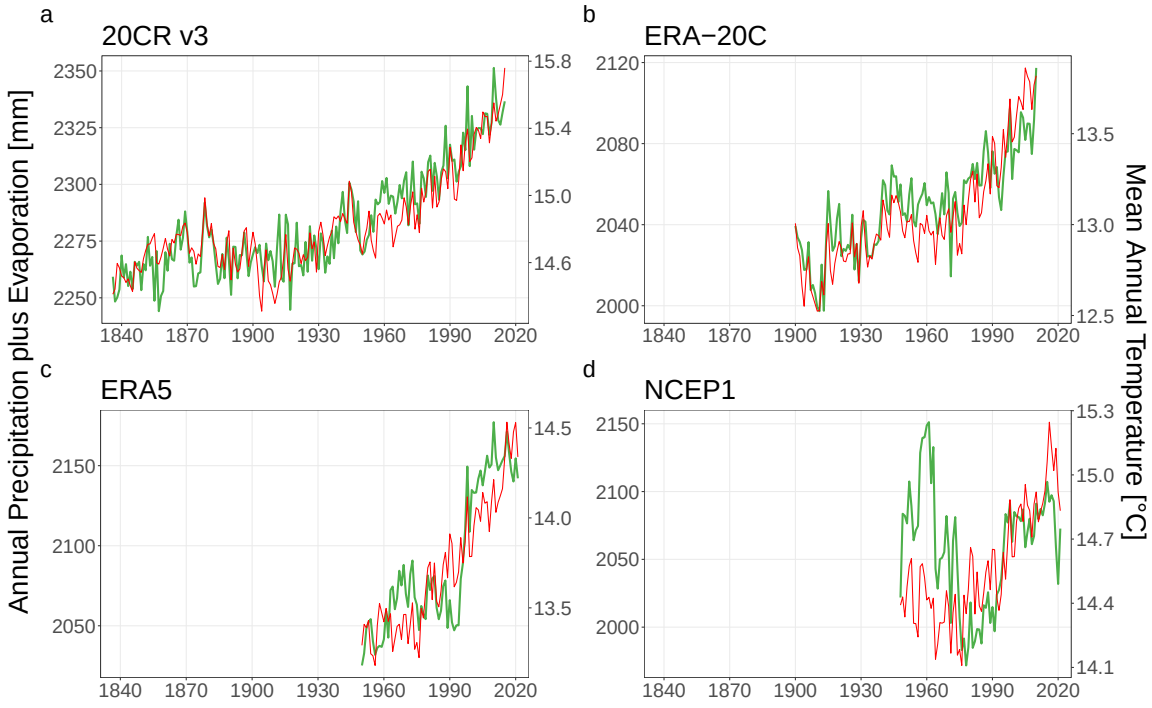


Figure 3.5: Global spatial weighted average of annual total precipitation plus evaporation and annual mean temperature in four reanalysis data sets for their respective available record. Precipitation plus evaporation in $[mm]$ is shown in green. Temperature in $[°C]$ is shown in red.

To validate our hypothesis, we looked into the slopes of linear regression fits between $P+E$, P, E, and temperature (Table 3.2). We validated the anticipated increases for $P+E$ except for ERA5, which had a rate of 5.0 ± 0.3 $[\%/°C]$, but also a rather high evaporation increase of 6.2 ± 0.4 $[\%/°C]$. R-squared offers some insight about the proportion of variance in $P+E$, P, and E that can be explained by temperature. Interestingly, evaporation has the lowest correlation to temperature across all reanalyses. 20CR v3 and ERA5 have higher R-squared values for $P+E$ than for P, with differences of 0.12 and 0.09, respectively. In contrast, ERA-20C and NCEP1 have higher R-squared values for precipitation (differences of 0.01 and 0.05). Note that while precipitation has a higher R-squared for ERA-20C and NCEP1, the difference is one order of magnitude smaller than those whose $P+E$ has a higher R-squared (20CR v3 and ERA5). These results demonstrate a good coupling between $P+E$ and hydrological sensitivity.

The above analysis establishes the usability of reanalysis data to assess changes in atmospheric water fluxes and temperature. It also highlights the different insight gained from $P-E$ and $P+E$. We will now unveil further details through a graphical framework that integrates precipitation, evaporation, their difference, and their sum (Huntington et al., 2018). By transforming the changes in the relationship of P and E to changes in $P-E$ and $P+E$, we can describe the water cycle

Table 3.2: Linear relationship between global spatial weighted average of total atmospheric water fluxes and mean temperature, where P is precipitation, E is evaporation, and T is temperature. $P + E$ columns report the correlation ($(P + E)$ vs. T). P columns report the correlation (P vs. T). E columns report the correlation (E vs. T). Slopes are in [%/°C], where the reference for atmospheric flux percentage change and temperature anomaly is their 1981-2010 average. RSE is Residual Standard Error.

	$P + E$			P			E		
Reanalysis	slope [%/°C]	RSE	R^2	slope [%/°C]	RSE	R^2	slope [%/°C]	RSE	R^2
20CR v3	3.2 ± 0.1	0.42	0.82	3.1 ± 0.2	0.58	0.7	3.3 ± 0.2	0.69	0.65
ERA-20C	3.3 ± 0.2	0.5	0.8	3.3 ± 0.2	0.49	0.81	3.2 ± 0.2	0.52	0.78
ERA5	5.0 ± 0.3	1.05	0.75	3.8 ± 0.3	1	0.66	6.2 ± 0.4	1.32	0.74
NCEP1	2.8 ± 0.9	2.01	0.12	4 ± 1	2.18	0.17	1.9 ± 0.9	1.95	0.06

dynamics in terms of atmospheric water storage and fluxes correspondingly. We apply this procedure to the four reanalyses to explore their representation of water cycle between two 30-year periods (1951-1980 and 1981-2010; Figure 3.6). It is easy to pinpoint some distinguishable features for each data set. The 20CR v3 appears to have substantially higher atmospheric water flux estimates than any other reanalysis. However, if we decompose it in $P - E$ and $P + E$ terms, we can see that in the two periods examined, the difference between precipitation and evaporation increased (blue vector), implying atmospheric water loss (Figure 3.6B). In ERA5, the exact opposite behavior emerges. The atmospheric water content has been increasing, but overall the average conditions suggest that the atmosphere has been getting drier since 1950 (Figure 3.6D). The remaining two reanalyses manifest a stationary relationship in the water storage with no changes in the $P - E$ component (Figures 3.6C and 3.6E). Surprisingly, the flux of atmospheric water is decreasing in NCEP1, suggesting a weakening of the water cycle (green arrow; Figure 3.6C).

It is evident that no two reanalyses are alike when it comes to the exchange of water between the land-ocean continuum and the atmosphere at the global scale. In terms of magnitude, ERA5 reports changes in $P + E$ accelerating almost twice as fast as in the 20CR v3 and ERA-20C (41.5 $[mm/yr]$ versus 23.69 $[mm/yr]$ and 25.3 $[mm/yr]$, respectively). The $P + E$ change in NCEP1 is similar to that observed in the 20CR v3 and ERA-20C. Although as already mentioned, in the opposite direction. Looking beyond 1950, in the reanalyses with longer records (20CR v3 and ERA-20C), we can see an agreement in the direction of change since 1921. Additionally, both reanalyses show a higher increase in $P + E$ between 1951-1980 and 1981-2010 than between 1921-1950 and 1951-1980. What is different, though, is the behavior of $P - E$, especially if analyzed over their 30-year average trajectory (Figures 3.6B and 3.6C light gray points). In ERA-20C, $P-E$ changes remain consistently stationary and very close to zero (0.15 $[mm/yr]$), while in the 20CR v3 oscillates substantially

following both increasing and decreasing patterns over the last 120 years. The trajectories of the other two reanalyses show behaviors somewhere in between, with more flexibility in $P-E$ compared to ERA-20C but not as much freedom as in the 20CR v3. Overall, the combination of $P-E$ and $P+E$ revealed a wealth of additional information about the reanalyses performance that is easily communicable and reproducible through the precipitation-evaporation space graphical framework, shaping the path for further investigations into the reasons behind these differences.

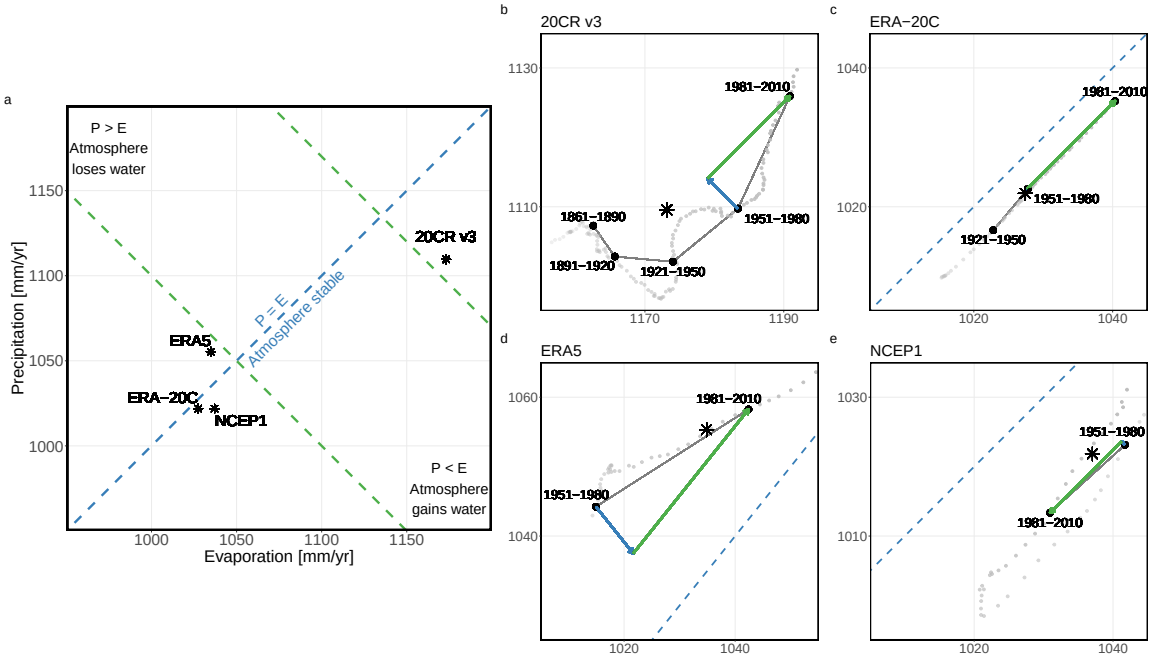


Figure 3.6: The precipitation-evaporation space graphical framework for the assessment of global water cycle changes. P and E are global total precipitation and evaporation in $[mm/year]$. Contour of $P = E$ is shown as a blue dashed line (stable atmosphere). Contours of equal $P + E$ are shown as green dashed lines (equal water cycle intensity). Changes in $P - E$ and $P + E$ are shown as blue and green vectors correspondingly. Light gray points show the 30-year moving average trajectory, black points mark the labeled 30-year period of interest, and stars mark the position of the average for the full record of each reanalyses. I.e., 1836-2015 average for 20CR v3, 1900-2010 average for ERA-20C, 1950-2020 average for ERA5, and 1948-2020 average for NCEP1. (A) Relative position of reanalyses with respect to each other in the precipitation-evaporation space. (B) Zoomed in panel on the 20CRv3. (C) Zoomed in panel on ERA20C. (D) Zoomed in panel on ERA5. (E) Zoomed in panel on NCEP1.

3.5 Discussion

Due to the lack of robust observational-based data for crucial water cycle components such as evaporation, reanalyses data is still one of our best tools for researching changes in the global water cycle. The results fall within persistent criticism toward reanalyses (e.g., substantial variability

Prein and Pendergrass (2019) and overestimations Trenberth et al. (2011)) but advocate for the framework proposed to acquire new insight and improve climate reanalysis. We displayed how while $P - E$, a key diagnostic, is not directly observable at the global scale, $P + E$ is not held back by scale limitations and complements global water cycle research. Most significantly, including $P + E$ revealed additional information about the water cycle changes characteristics in four reanalyses. Information that could be implemented to address non-physical trends and inhomogeneities due to changes in the observing system (e.g., Ho et al. (2020b)) and water budget non-closure (e.g., Trenberth and Fasullo (2013a,b)). The latter is an ongoing challenge in global water cycle research (Hegerl et al., 2015), and non-closure is present in all reanalyses. Unexpectedly, although, we found a spurious long-term correlation between $P - E$ and temperature, suggesting such an artifact might be rooted in model processes and not only due to assimilation schemes. Along that line, we were surprised to find that the correlation between evaporation and temperature is smaller than that between precipitation and temperature in reanalyses, except for ERA5.

Needless to say, a persistent challenge is the unconstrained uncertainty in quantifying water cycle fluxes. Of particular relevance herein is that global $P - E$ is small, and its uncertainty might easily be much larger than its value. Thus, the signal-to-noise ratio of changes in $P - E$ versus the natural variability will be low and as a consequence the fluctuations in water cycle harder to detect. This limitation can be overcome when using $P + E$, which is less prone to the reanalyses uncertainties. These uncertainties could be encapsulated by their assimilation scheme, considering the assimilation scheme includes, among others: the forecast model, boundary conditions, observations, observation operators, and covariance models (Dee, 2005). Put simply, differences in reanalysis assimilation schemes can significantly impact precipitation and evaporation inherent uncertainties. NCEP1 reanalysis uses a 3D-Var data assimilation system, which minimizes the difference between the model and observations by adjusting the atmospheric state variables (Kalnay et al., 1996). On the other hand, ERA-20C and ERA5 use a 4D-Var data assimilation system, which adjusts the atmospheric state variables over a series of time steps to minimize the difference between the model and observations (Poli et al., 2016; Hersbach et al., 2020). The 20CR v3 uses a hybrid 4D-Var/Ensemble Kalman Filter data assimilation system, which combines the strengths of both 3D-Var and 4D-Var to improve the accuracy of the precipitation and evaporation estimates (Slivinski et al., 2019). Although it is not the scope of this study to address the underlying uncertainties or the effects of different assimilation schemes, looking into discrepancies among reanalyses estimates offers a handy demonstration of what can be learned by utilizing the precipitation-evaporation space to assess water cycle changes.

Whilst some features were common for all or most reanalyses, like changes in $P - E$ being much smaller than in $P + E$ and an increase in $P + E$ between the two most recent 30-year periods, we observed various individual distinctions. Out of the four reanalyses, ERA5 had the most comprehensive representation of precipitation and temperature variability compared to observational-based references, and was found to represent better the acceleration dynamics between 1951-1980 and 1981-2010. At the same time, ERA5 has the most pronounced changes for $P - E$, showcasing improvements in its terrestrial water storage computations (Eicker et al., 2020). However, ERA5 has the steepest acceleration of $P + E$ and is the only reanalysis above the $P = E$ isoline for the entirety of its record, which could be an artifact attributed to precipitation overestimations identified across different regions (Hassler and Lauer, 2021). To the opposite end, NCEP1 shows a decline in atmospheric water fluxes over time with a slight decrease in atmospheric water storage. Forbye, the 30-year average trajectory exhibits an acute u-turn between the mid-1960s and the late 1970s. Around this trajectory inversion, the behavior is similar to ERA-20C with little to no variability along a $P - E$ isoline. A possible explanation for this abnormal behavior could be traced back to remote sensing data assimilation. Inconsistencies in its atmospheric data pre-1979 have previously been reported and associated with the lack of satellite observations before 1979, e.g., in the Southern Hemisphere (Tennant, 2004).

Using solely $P + E$ comes with its own limitations and could mask the true dynamics of global water cycle change. The reciprocal complementarity of $P + E$ and $P - E$ is better perceived on the long-record reanalyses. The overview clearly shows that the 20CR v3 portrays a warmer and wetter Earth relative to the rest of the reanalyses. This is consistent with a systematic bias in tropical precipitation (Slivinski et al., 2019), and biases in the vertical structure of mass and circulation determined throughout the atmosphere (Slivinski et al., 2021). Having said that, the magnitude of changes in $P + E$ are consistent with those of ERA-20C. The most recent increase is higher than the preceding ones and suggests that the global water cycle acceleration signal has further strengthened in the last three decades (Greve et al., 2014). The above would suggest that changes in the global water cycle are similarly represented on both data sets. In sooth, $P - E$ changes in the 20CR v3 oscillate substantially following both increasing and decreasing patterns, whereas ERA-20C shows little to no variability (no spurious jumps or trends) and steadily moves along a $P - E$ isoline. Said stability lies around water cycle budget non-closure because evaporation is higher than precipitation despite known systematic precipitation overestimation (Rustemeier et al., 2019). Reportedly, there are only subtle differences in the data assimilated and the data assimilation schemes between these two reanalyses (Poli et al., 2016), yet we can see contrasting behaviors exposed within the framework

proposed herein.

Our findings, including the good agreement with the range of hydrological sensitivity, advocate for the definitions of $P+E$ to be physically sound. It is important, nonetheless, to note that such an agreement is not a two-way relationship. As seen in our examination, the fact that all the reanalyses have similar hydrological sensitivities does not necessarily mean that they express a similar rate of water cycle changes. Assuming so could be misleading, whereas we can get more insight and avoid these pitfalls by decomposing the change into $P-E$ and $P+E$ (i.e., into water storage and fluxes). It could be argued that introducing a new metric for acceleration into the current broad spectrum of metrics may lead to inconsistent hydroclimatology analysis terminology, such as that recently reported for wetter and drier (Roth et al., 2021). Nevertheless, $P+E$ is not just an index because it is physically grounded and, as such, is better suited to describe climate models and reanalyses (McColl et al., 2022).

Along the same line, it could be argued that assessing changes in precipitation or evaporation alone can directly indicate changes in the water cycle. It is easy to imagine that altering the state of one component in the water cycle would affect the others. However, the global water cycle is a complex phenomenon composed of several processes that we are yet to understand fully. Hence, changes in one component might not be instantly observed in the others. The compound behavior of precipitation and evaporation provides a more comprehensive picture of the water balance because it considers both the supply and demand of water or, within the precipitation-evaporation space, both atmospheric water storage ($P-E$) and water cycle intensity ($P+E$). As evinced by our results, precipitation increases are evident in all reanalyses. Regardless, until we inspect these reanalyses in the precipitation-evaporation space, we cannot observe that, in reality, no reanalysis is alike as they all describe different water cycle dynamics.

The above applications highlight the potential of $P+E$ to complement water cycle research at the global scale. The proposed framework could advance our understanding of water cycle changes and improve climate modeling. We have already revealed some discrepancies between the reanalysis data sets. Still to properly address them, the observational limitations at global scale, especially in evaporation, need to be overcome (McCabe et al., 2016). Additionally, it is intriguing to see how the total water transfer between the land-ocean continuum and atmosphere appears in Earth System Models and whether it can be also applied as a metric for the model performance. Future research into global spatial patterns of $P+E$ could also shed more light on how they relate to regional changes and hydroclimatic extremes such as droughts. To this extent, quantifying the surface-atmosphere water exchange in the form of $P+E$ can enhance our insight into past, present,

and future hydroclimatic variability.

3.6 Methods

3.6.1 Data

We selected four reanalysis data products (Table 3.3). These are the Twentieth Century Reanalysis (20CR) v3 (Slivinski et al., 2019), European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalyses ERA-20C (Poli et al., 2016) and ERA5 (Hersbach et al., 2020), and the National Centers for Environmental Prediction & the National Center for Atmospheric Research NCEP/NCAR Reanalysis 1 (Kalnay et al., 1996). The 20CRv3 and the ERA-20C have two of the longest record among reanalyses, with 180 and 100 years, respectively. ERA5 and NCEP1 are two distinctive ongoing projects. ERA5 is a fifth-generation reanalysis (the most recent to date), and NCEP1 is a first-generation reanalysis. NCEP1 it is the longest-running reanalysis that uses rawindsonde data, but the model and data assimilation scheme are antiquated (Trenberth et al., 2011). Notwithstanding, the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP) data set (Xie and Arkin, 1997), which is highly regarded as an observational-based reference (Yin et al., 2004), blends NCEP1 to fill missing data.

Additionally, we used two observation-based products. For precipitation, the Global Precipitation Climatology Project (GPCP) v2.3 (Adler et al., 2018), which merges data from rain gauge stations, satellites, and sounding observations. For temperature, the HadCRUT5 (Morice et al., 2021) from the Met Office Hadley Centre and the Climatic Research Unit at the University of East Anglia, which blends data from meteorological stations, ships, and buoys. All of the above data sets are available for download on the dedicated websites of their providers. Through the pRecipe R package (`https://cran.r-project.org/package=pRecipe`), we computed the area-weighted average of gridded data and generated annual time series for total atmospheric water fluxes and global mean temperature.

3.6.2 Benchmarking Reanalyses

We examined some commonly used statistical metrics to benchmark the reanalysis data products. Their aptness to capture the temporal variability of the water cycle was quantified via:

- The square of the Pearson correlation coefficient (R-squared or R^2)

$$R^2 = 1 - \frac{\sum_i^n (y_i - \hat{y}_i)^2}{\sum_i^n (y_i - \overline{y})^2}$$

Table 3.3: Data Set Overview

Name	Source	Model Resolution	Record Length	Assimilation Schemes	Reference
20CR v3	NOAA	T254 ($\approx 75[km]$ at the equator)	1836 - 2015	Ensemble Kalman Filter and 4-dimensional incremental analysis update (EnKF-4DIAU)	(Slivinski et al., 2019)
ERA-20C	ECMWF	T159 ($\approx 125[km]$ at the equator)	1900 - 2010	4-dimensional variational assimilation (4D-Var)	(Poli et al., 2016)
ERA5	ECMWF	T639 ($\approx 31[km]$ at the equator)	1950 - now	4-dimensional variational assimilation (4D-Var)	(Hersbach et al., 2020)
NCEP1	NCEP NCAR	T62 ($\approx 210[km]$ at the equator)	1948 - now	3-dimensional variational assimilation (3D-Var)	(Kalnay et al., 1996)

where i starts on the first year of the available record, n is the last year of the available record, y_i is the observational estimate on year i, $\hat{y}_i$ is the reanalysis estimate on year i, and $\overline{y}$ is the mean observational estimate for the full available record.

- Root Mean Square Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{\sum_i^n (y_i - \hat{y}_i)^2}{N}}$$

where i starts on the first year of the available record, n is the last year of the available record, y_i is the observational estimate on year i, $\hat{y}_i$ is the reanalysis estimate on year i, and N is the total number of years in the full available record.

Only precipitation and temperature records were evaluated because there is no robust observation-based evaporation data set. Note that precipitation and temperature were compared using two different reference periods because GPCP v2.3 record starts in 1979, and HadCRU5T5 provides only temperature anomalies using the 1961-1990 average as a reference. Thus, we could not homogenize the reference period for both variables and selected 1981-2010 for precipitation and 1961-1990 for temperature. Subsequently, we inspected global water budget closure via the 30-year moving average of $P - E$.

3.6.3 Thermodynamics of Atmospheric Fluxes

For superimposing the temperature to the precipitation plus evaporation time series, without incurring in any kind of data tampering, we simply rescaled temperature to precipitation plus evaporation in the same way one would rescale degrees Fahrenheit to degrees Celsius. I.e.:

$$y_i' = \left((T_i - min(T)) * \frac{max(P+E) - min(P+E)}{max(T) - min(T)} \right) + min(P+E)$$

where y_i' denotes the value used to plot Temperature in the same scale of precipitation plus evaporation for any given year, T_i is the temperature reanalysis estimate on year i, $min(T)$ is the minimum temperature reanalysis estimate in the full available record, $max(P+E)$ is the maximum precipitation plus evaporation reanalysis estimate in the full available record, $min(P+E)$ is the minimum reanalysis estimate in the full available record, and $max(T)$ is the maximum temperature reanalysis estimate in the full available record.

As thermodynamics dictates, we expect a linear relationship between atmospheric water fluxes

and temperature. This correspondence was quantified via the square of the Pearson correlation coefficient (R-squared or R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i - f(T_i)\right)^2}{\sum_{i=1}^{n} \left(y_i - \overline{y}\right)^2}$$

where n is the total number of years in the full available record, y_i is the i-th reanalysis estimate for atmospheric water flux, $f(T_i)$ is the i-th predicted estimate by temperature, and $\overline{y}$ is the mean reanalysis estimate for the full available record. The same metrics were computed again for the annual differences of each time series (i.e., $\delta(y_i) = y_i - y_{i-1}$). To this extent, we can characterize the long-term and the year-to-year association between atmospheric water fluxes and temperature. While the correlation coefficient describes the presence or absence of a linear relationship, it does not quantify the rate of change of one variable relative to the other. Henceforth, we used linear regression to estimate the corresponding slopes and describe the rate of change at which atmospheric water fluxes respond to changes in temperature. To compare the slopes between data sets on a one-to-one basis, we estimated atmospheric water fluxes and temperature in terms of global anomalies with respect to the 1981-2010 period.

$$\text{slope} = \frac{n\left(\sum_{i=1}^{n} T_i y_i\right) - \left(\sum_{i=1}^{n} T_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\left(\sum_{i=1}^{n} T_i^2\right) - \left(\sum_{i=1}^{n} T_i\right)^2}$$

where n is the total number of years in the full available record, y_i is the i-th reanalysis estimate for atmospheric water flux anomaly, and T_i is the i-th reanalysis estimate for temperature anomaly. Lastly, we relied on the Residual Standard Error (RSE) to assess the goodness-of-fit of the slopes, i.e., how well these slopes represent the linear relationship between our variables.

$$\text{RSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(y_i - f(T_i)\right)^2}{n-1}}$$

where n is the total number of years in the full available record, y_i is the i-th reanalysis estimate for atmospheric water flux anomaly, and $f(T_i)$ is the i-th predicted estimate by temperature anomaly.

Author Contributions

M.R.V.G. and Y.M. designed the research. M.R.V.G. performed the analyses and wrote a draft of the paper. M.R.V.G. and Y.M. discussed the results and wrote the manuscript.

Chapter 4

pRecipe: A global precipitation climatology toolbox and database

4.1 Introduction

It is common practice to describe and quantify the water cycle focusing on its four major components evaporation, precipitation, runoff, and total water storage (Harding et al., 2011). Out of these four components, precipitation has been extensively researched because it is the primary factor determining water availability across several spatiotemporal scales (Trenberth and Zhang, 2018). Accurate estimates of precipitation climatologies are crucial for water resource management (Marques et al., 2022), water-related engineering design, and long-term agricultural policy making (Bezner Kerr et al., 2022). Notwithstanding, to date, a comprehensive network of ground stations remains elusive due to practical, economical, or political reasons (Vargas Godoy et al., 2021). When ground observations are unavailable, we may rely on data from different sources, such as satellite remote sensing, model simulations, and reanalyses. Regardless of the source, having a good grasp of the uncertainty of the estimates becomes imperative. Consequently, using different data products from various, ideally independent, sources is the most appropriate direction for current research and operational needs.

Although precipitation understanding has improved dramatically due to the vast amount of different data sources nowadays, their information has not been comprehensive enough due to substantial uncertainty between sources, with biases reaching as much as 300% (Fekete et al., 2004). Such uncertainty could be partially attributed to the intrinsic heterogeneity of multiple aspects,

from data distribution to end-product specifications, e.g., spatial resolution, time step, measuring units, file format, etc. Therefore, we find ourselves with a broad spectrum of data renditions, a research matter on its own (Sun et al., 2018), and a homogenization pre-processing hindrance. The latter, i.e., data preparation, is acknowledged to be an often unavoidable and rather time-consuming step of the analysis (Young et al., 2017). Some data distributors mitigate the above challenges by facilitating different tools for extraction (e.g., `https://www.earthdata.nasa.gov/`), exploration (e.g., `https://climexp.knmi.nl/`), and, in some cases, visualization (e.g., `https://giovanni.gsfc.nasa.gov/`). However, being online services, they are heavily oriented toward graphic user interfaces, are limited to elementary operations for exploratory data analysis, and allow for simultaneous analysis of at most two data sets at a time. Undeniably, a broader, more inclusive framework integrating multiple data sets is still missing.

Tailored software is available to deal with the processing and analyze large amounts of data efficiently, namely the Climate Data Operators (CDO; Schulzweida, 2022) and Climate Data Analysis Tools (CDAT; Williams et al., 2009). These packages provide data cleaning, analysis, and visualization tools. While working with precipitation data can be challenging and time-consuming, these or similar alternative software allow researchers to automate and streamline the data analysis process for reproducibility. Nevertheless, a significant limitation of tools like CDO and CDAT is their incompatibility with Windows, the dominant desktop operating system globally. It could be argued that installing both CDO and CDAT in Windows is possible. However, the installation is done through the Windows Subsystem for Linux (WSL), which provides a GNU/Linux environment, including command-line tools and utilities, on Windows (Singh, 2020). Moreover, CDAT is staged for deprecation and cease of support around the end of the calendar year 2023.

Over the last decade, R, an open-source programming language (R Core Team, 2023), has continuously increased its presence until it acquired a central role in hydrological research and the operational practice of hydrology (Slater et al., 2019). The R hydrological community has grown significantly in the last decade with applications, or packages, that involve data retrieval and pre-processing from hydrological and meteorological sources, hydrograph and spatial analysis functions, and tools for process-based and stochastic modeling. Nevertheless, more often than not, these packages are still developed around one data set or one data provider. E.g., *easyclimate* to access high-resolution daily climate data for Europe (Cruz-Alonso et al., 2023) or *dataRetrieval* for the US Geological Survey (USGS) National Water Information System (DeCicco et al., 2022). Tools-centered packages tend to be more comprehensive and require more generic inputs. E.g., *envoutliers* identifies outliers in environmental time series data (Čampulová et al., 2022), and *CoSMoS* gen-

erates univariate/multivariate non-Gaussian time series and random fields for environmental and hydroclimatic processes (Papalexiou et al., 2021). These latter kinds of packages give the users more flexibility on the account they are to deal with data gathering and pre-processing on their own.

Despite R flourishing in hydrology, the previously mentioned support supplied by data providers focuses more on different programming languages like MATLAB or Python. On that account, addressing both data preparation time consumption and the lack of a comprehensive R-based alternative, we introduce *pRecipe*. We acknowledge that various yet exclusive data retrieval packages are already available (see Albers et al. (2022)). Moreover, while claiming to provide all available data sets would be fraudulent, we can ascertain that *pRecipe* provides a ready-for-analysis homogenized database with products from various sources. The *pRecipe* package database consists of 24 data sets at monthly time step and 0.25° resolution. These are derived from gauge, satellite, reanalysis, and hydrological model forcing precipitation products. Furthermore, *pRecipe* offers additional processing tools to subset the record length and spatial coverage, crop data based using shapefiles, and various graphical aesthetics for visualization and exploratory data analysis. The package can be downloaded from the CRAN repository or from `https://github.com/MiRoVaGo/pRecipe`.

4.2 Methods

4.2.1 Data

The *pRecipe* package offers a database of 24 precipitation data sets homogenized to common a spatial (0.25°) and temporal (monthly) resolution. These include:

- Seven gauge-based products: CPC-Global (Xie et al., 2010a), CRU TS v4.06 (Harris et al., 2020), EM-EARTH (Tang et al., 2022), GHCN v2 (Peterson and Vose, 1997), GPCC v2020 (Schneider et al., 2011), PREC/L (Chen et al., 2002), and UDel v5.01 (Willmott and Matsuura, 2001).

- Eight satellite-based products: CHIRPS v2.0 (Funk et al., 2015), CMAP (Xie and Arkin, 1997), CMORPH (Joyce et al., 2004), GPCP v2.3 (Adler et al., 2018), GPM IMERGM v06 (Huffman et al., 2020), MSWEP v2.8 (Beck et al., 2019), PERSIANN-CDR (Ashouri et al., 2015), and TRMM 3B43 v7 (Huffman et al., 2010).

- Five reanalysis products: 20CR v3 (Slivinski et al., 2019), ERA-20C (Poli et al., 2016), ERA5 (Hersbach et al., 2020), NCEP/NCAR R1 (Kalnay et al., 1996), and NCEP/DOE R2 (Kana-

mitsu et al., 2002).

- Four hydrological model forcing products: GLDAS-CLSM v2.0 (Rodell et al., 2004), GLDAS-NOAH v2.0 (Rodell et al., 2004), GLDAS-VIC v2.0 (Rodell et al., 2004), and TerraClimate (Abatzoglou et al., 2018).

Their native specifications, as well as download links to their original providers, and their respective references, are detailed in Tables 4.1, 4.2, 4.3, and 4.4, respectively. If multiple distributions were available, the one closest to the target spatiotemporal resolution was chosen to minimize the pre-processing uncertainty. Remapping data from one spatial or temporal resolution to another can result in information loss when processing data from higher to lower resolution. Hence, these uncertainties' magnitudes depend on source data quality and are proportional to the times we manipulate data. Consequently, the less pre-processing we have to perform on the data, the less uncertainty we introduce.

Overall, the package focuses on three fronts: formatting, homogenization, and storage. To begin with, most providers either natively have data in the Network Common Data Form (NetCDF) format or offer the option to download in that format. In the same fashion, we chose the NetCDF format for our database. The GPM (Huffman et al., 2020) and TRMM (Huffman et al., 2010) data sets use the Hierarchical Data Format (HDF) instead. TRMM data is in HDF4 format and was reformatted into NetCDF using the conversion toolkit from the HDF group (`https://hdfeos.org/`). GPM is in HDF5 format, and no direct conversion tool was available. Thus, we extracted the values and stored them in NetCDF files using R. Note that no reprojection or manipulation of any kind took place at this stage. Once all data sets were in NetCDF files, if there were multiple files per data set (i.e., one file per day, month, or year), we merged them in time into a single NetCDF file using Climate Data Operators (CDO; Schulzweida (2022)).

Then, data homogenization addressed the variable type (total precipitation; tp), the measuring units (millimeters; mm), the temporal resolution (monthly), and the spatial resolution (0.25°). If the providers offered both precipitation rate and total precipitation, total precipitation files were downloaded to minimize data tampering. We converted the precipitation rate from $[mm/day]$ or $[kg/m2/s]$ into total precipitation $[mm]$. Else, we just converted the units of total precipitation where needed (e.g., $[m]$ into $[mm]$). Subsequently, daily data was aggregated into monthly. Thereafter, spatial remapping was performed using CDO. When regridding coarser than 0.25° resolution data, the 'remapnn' operator was used for nearest-neighbor interpolation. Otherwise, the 'gridboxmean' operator would be used for regridding via area-weighted averaging (accounting for the area of each

grid cell in proportion to the total area being averaged) plus 'remapnn' when 0.25° is not divisible by the original resolution. Arguably, nearest-neighbor interpolation potentially leads to abrupt changes in the values of the remapped data when used to fill in missing data. However, if used simply for regridding, we do not introduce any significant artifacts as evinced by differences of less than 0.01% in total precipitation volume between raw and remapped data. Finally, the database has been deposited in a public Zenodo repository under the following naming convention:

<data set>_<variable>_<units>_<coverage>_<start date>_<end date>_<resolution>_<time step>.nc

E.g., GPCP v2.3 (Adler et al., 2018) would be: *gpcp_tp_mm_global_197901_202205_025_monthly.nc*

4.2.2 Package Design

Designed with reproducible science in mind, the *pRecipe* package facilitates the download, exploration, visualization, and analysis of multiple precipitation data products across various spatiotemporal scales. The general workflow is as follows:

1. Direct download of a single, multiple, or all data sets available in the *pRecipe* database is done via the *download_data()* function, which has two arguments *data_name* and *destination*. The *data_name* argument is set to "all" by default, but the users can specify the name(s) of their interest: 20cr, chirps, cmap, cmorph, cpc, cru-ts, em-earth, era20c, era5, ghcn, gldas-clsm, gldas-noah, gldas-vic, gpcc, gpcp, gpm-imerg, mswep, ncep-doe, ncep-ncar, persiann, precl, terraclimate, trmm-3b43, and/or udel. The *destination* argument is set to "." by default. I.e., the current working directory. By replacing it for "your_project_folder", the downloaded files will be stored in "your_project_folder" instead.

2. Data processing functions are built upon the raster package (Hijmans et al., 2022), with the additional advantage that saving data will do so in a NetCDF format compatible with CDO. Currently, *pRecipe* offers spatial subsetting by either a bounding box or an irregular polygon via shapefile. Besides, temporal upscaling from monthly to yearly scale offers basic statistical options, such as maximum, minimum, median, average, and sum. Last but not least, the *make_ts()* function computes the area-weighted average of each time step, be it monthly or annual, to transform the raster into a time series comma-separated values (CSV) file.

3. Prompt and aesthetic visualization is available at any stage of analysis. The *pRecipe* graphical framework allows the user to explore and present analysis results of its data via maps, time series curves, boxplots, histograms, and heat maps. It is important to note that the above-

Table 4.1: Gauge-based Products

Data Set	Spatial Resolution	Spatial Coverage	Temporal Resolution	Record Lenght	Get Data	Reference
CPC-Global	0.5°	Land	Daily	1979-01 to 2022-08	Download	Xie et al. (2010a)
CRU TS v4.06	0.5°	Land	Monthly	1901-01 to 2021-12	Download	Harris et al. (2020)
EM-EARTH	0.1°	Land	Daily	1950-01 to 2019-12	Download	Tang et al. (2022)
GHCN v2	5°	Land	Monthly	1900-01 to 2015-05	Download	Peterson and Vose (1997)
GPCC v2020	0.25°	Land	Monthly	1891-01 to 2022-08	Download	Schneider et al. (2011)
PREC/L	0.5°	Land	Monthly	1948-01 to 2022-08	Download	Chen et al. (2002)
UDel v5.01	0.5°	Land	Monthly	1901-01 to 2017-12	Download	Willmott and Matsuura (2001)

Table 4.2: Satellite-based Products

Data Set	Spatial Resolution	Spatial Coverage	Temporal Resolution	Record Lenght	Get Data	Reference
CHIRPS v2.0	0.05°	50°S-N Land	Monthly	1981-01 to 2022-07	Download	Funk et al. (2015)
CMAP	2.5°	Global	Monthly	1979-01 to 2022-07	Download	Xie and Arkin (1997)
CMORPH	0.25°	60°S-N Global	Daily	1998-01 to 2021-12	Download	Joyce et al. (2004)
GPCP v2.3	0.5°	Global	Monthly	1979-01 to 2022-05	Download	Adler et al. (2018)
GPM IMERGM v06	0.1°	Global	Monthly	2000-06 to 2020-12	Download	Huffman et al. (2020)
MSWEP v2.8	0.1°	Global	Monthly	1979-02 to 2022-06	Download	Beck et al. (2019)
PERSIANN-CDR	0.25°	60°S-N Global	Monthly	1983-01 to 2022-06	Download	Ashouri et al. (2015)
TRMM 3B43 v7	0.25°	50°S-N Global	Monthly	1998-01 to 2019-12	Download	Huffman et al. (2010)

Table 4.3: Reanalysis Products

Data Set	Spatial Resolution	Spatial Coverage	Temporal Resolution	Record Lenght	Get Data	Reference
20CR v3	1°	Global	Monthly	1836-01 to 2015-12	Download	Slivinski et al. (2019)
ERA-20C	1.125°	Global	Monthly	1900-01 to 2010-12	Download	Poli et al. (2016)
ERA5	0.25°	Global	Monthly	1959-01 to 2021-12	Download	Hersbach et al. (2020)
NCEP/NCAR R1	1.875°	Global	Monthly	1948-01 to 2022-08	Download	Kalnay et al. (1996)
NCEP/DOE R2	1.875°	Global	Monthly	1979-01 to 2022-08	Download	Kanamitsu et al. (2002)

Table 4.4: Hydrological Model Forcing

Data Set	Spatial Resolution	Spatial Coverage	Temporal Resolution	Record Lenght	Get Data	Reference
GLDAS CLSM v2.0	0.25°	Land	Daily	1948-01 to 2014-12	Download	Rodell et al. (2004)
GLDAS NOAH v2.0	0.25°	Land	Monthly	1948-01 to 2014-12	Download	Rodell et al. (2004)
GLDAS VIC v2.0	1°	Land	Monthly	1948-01 to 2014-12	Download	Rodell et al. (2004)
TerraClimate	$4km$	Land	Monthly	1958-01 to 2021-12	Download	Abatzoglou et al. (2018)

mentioned graphical framework is based on the *ggplot2* package (Wickham et al., 2022). As such, the outputs are easily adjusted to suit the user's needs using the grammar of graphics.

The *pRecipe* package is publicly available in the Comprehensive R Archive Network (CRAN) at `https://CRAN.R-project.org/package=pRecipe`. More experienced users may find all the functions' source code at `https://github.com/MiRoVaGo/pRecipe` and can easily modify them to fit the user specific needs if needed.

4.3 Case Study

The user-friendly accessibility that *pRecipe* provides makes analysis reproducibility as simple as following a recipe. In this introductory recipe, we downloaded the entire *pRecipe* database (Section 4.2.1) using the *download_data()* function. We then subsetted the downloaded data to the 1981-2020 period using the *subset_time()* function, and cropped it within the administrative borders of Czechia via the *crop_data()* in conjunction with a shapefile provided by the Database of Global Administrative Areas (GADM). We then generated time series using the *make_ts()* function. The time series were generated by computing the area-weighted average of all the grid cells of interest, and the values were stored in data.table objects with four columns: date, value, and name, type. The last two are mainly used for graphical aesthetics. Note that storing the time series in data.table objects enables further calculations with ease. Herein we calculated the sum, min, max, median, and mean of our monthly data by year in order to visually assess the similarities and discrepancies between data sources using the *plot_line()* function (Figure 4.1). It is evident at first glance that even limiting the data record to just 40 years, a line plot is not the best graphical aesthetic to represent our data due to the high clustering and overlapping of lines (Figure 4.1a). Upscaling into annual time steps, it is easier to observe that while there is considerable variability between different products (Figures 4.1b, 4.1c, 4.1e, and 4.1f), there is higher agreement in measuring high precipitation (Figure 4.1d).

To validate data using local observations or one of the downloaded data sets as the reference, we can assess their correlation and variance through Taylor diagrams using the *plot_taylor()* function (Figure 4.2). We used data from the Czech Hydrometeorological Institute (CHMI) to validate the database in this case study. As expected, observational data from gauge-based and satellite-based products are highly correlated with the CHMI reference, with most of their correlation coefficients above 0.95 and 0.9, respectively (Figure 4.2). In terms of variance, we observe that the hydrological model forcing data exhibits almost identical locations on the diagram. In contrast, reanalysis data

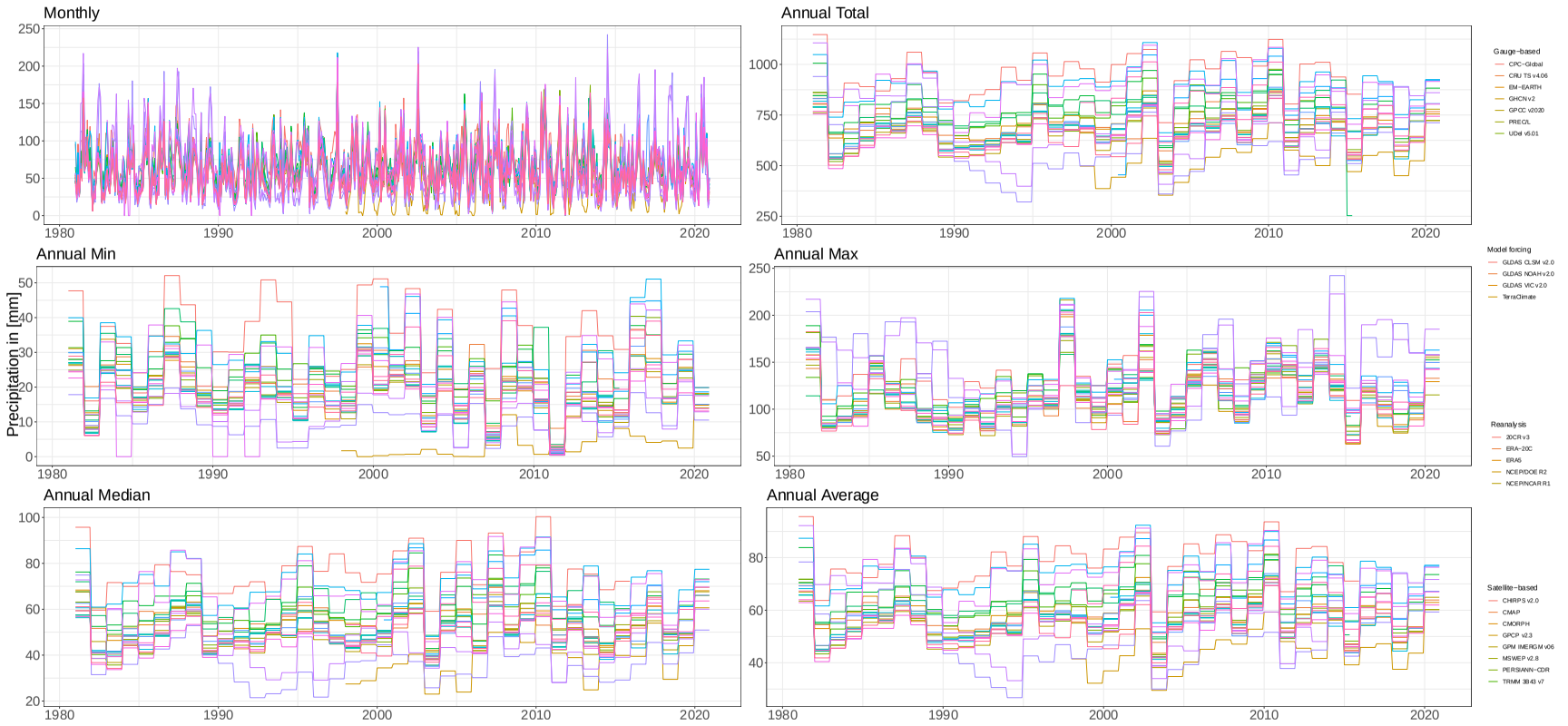


Figure 4.1: Different applications of the *plot_line()* function illustrating area-weighted average precipitation over Czechia for all available data sets in *pRecipe*'s database between 1981 and 2020. (A) Monthly precipitation. (B) Annual total precipitation. (C) Annual minimum precipitation. (D) Annual maximum precipitation. (E) Annual median precipitation. (F) Annual average precipitation.

are the most scattered of all four data sources. From this quick inspection, we can say GPCC v2020 data estimates are the closest to our validation reference, while NCEP/NCAR R1 and NCEP/DOE R2 are the most inconsistent (lowest correlation and highest variance). The former is presumably due to GPCC v2020 likely drawing data by the same network of stations CHMI oversees (Becker et al., 2013). The latter might be an artifact produced by the coarse native resolution of NCEP reanalyses compared to the area of Czechia. I.e., the area of two grid cells (approximately 86 632 [km^2]) fully covers Czechia (78 867 [km^2]). Moreover, NCEP/NCAR R1 is a first-generation reanalysis that uses antiquated data assimilation and model (Kalnay et al., 1996). NCEP/DOE R2 is a direct update that fixed some errors and updated the parameterizations of NCEP/NCAR R1 (Kanamitsu et al., 2002). Nonetheless, it did not address other limitations like higher horizontal and vertical resolution, direct assimilation of radiances, proper use of Special Sensor Microwave Imager (SSM/I) data, and assimilation of rainfall data.

Further insight into the validation of our data sets can be reckoned with by looking into their correlation and variance across different seasons (Figure 4.3). This time we looked only into eight data sets: CAMP, CPC-Global, ERA-20C, ERA5, GLDAS CLSM v2.0, GLDAS NOAH v2.0, GPCC v2020, and TRMM 3B43 v7. Using high correlation to CHMI as a preliminary filter, we selected the two best data sets from each source. Visualizing the data sets' correlation by season, we discover that the best agreement with the CHMI reference occurs during Fall, where most data sets have a correlation above 0.95, normalized standard deviation around 1, and centered root mean square error 0.5. Out of the selected data sets, ERA-20C has the lowest correlation regardless of the season, with its correlation further away from the rest, dropping to 0.8 in Summer. Observations assimilated by ERA-20C include surface pressure from the International Surface Pressure Databank (Compo et al., 2015) as well as from ICOADS (Woodruff et al., 2011) and surface winds over the oceans from ICOADS. Upper-air and satellite data are omitted (Poli et al., 2016). Due to the limited observations used, ERA-20C does not provide the best estimate since 1979, when major advancements in the observing system occurred with the dawn of the satellite era. Another point of interest is that TRMM 3B43 v7 correlation is consistent across seasons, but its variance visibly increases in Winter. TRMM precipitation radar algorithm has been reported to underestimate precipitation at higher latitudes (40°N; Chen and Li, 2016) in conjunction with winter precipitation characteristics (i.e., lighter rain events, snow, and mixed-phase precipitation) would explain larger biases in the Winter season (Maggioni et al., 2016).

The toolbox that *pRecipe* offers can also be applied to assess changes in precipitation regimes. We selected four data sets (GPCC v2020, ERA5, GLDAS NOAH v2.0, and MSWEP v2.8), we divided

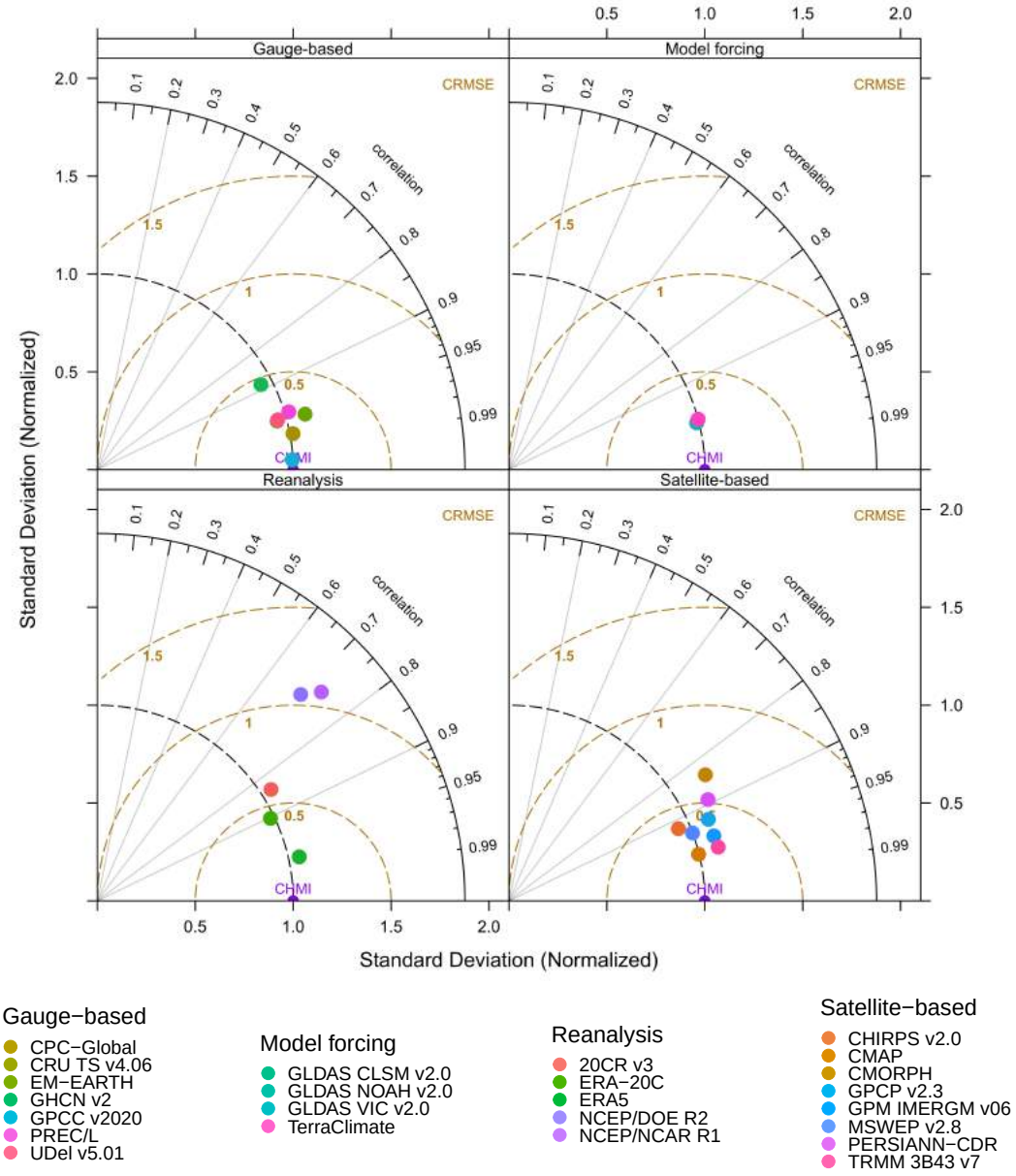


Figure 4.2: Illustration of the *plot_taylor()* function comparing the entire *pRecipe* database to the observational reference of the CHMI. Upper left, gauge-based data sets as listed in Table 4.1. Upper right, hydrological model forcing data sets as listed in Table 4.4. Bottom left, reanalysis data sets as listed in Table 4.3. Bottom right, satellite-based data sets as listed in Table 4.2.

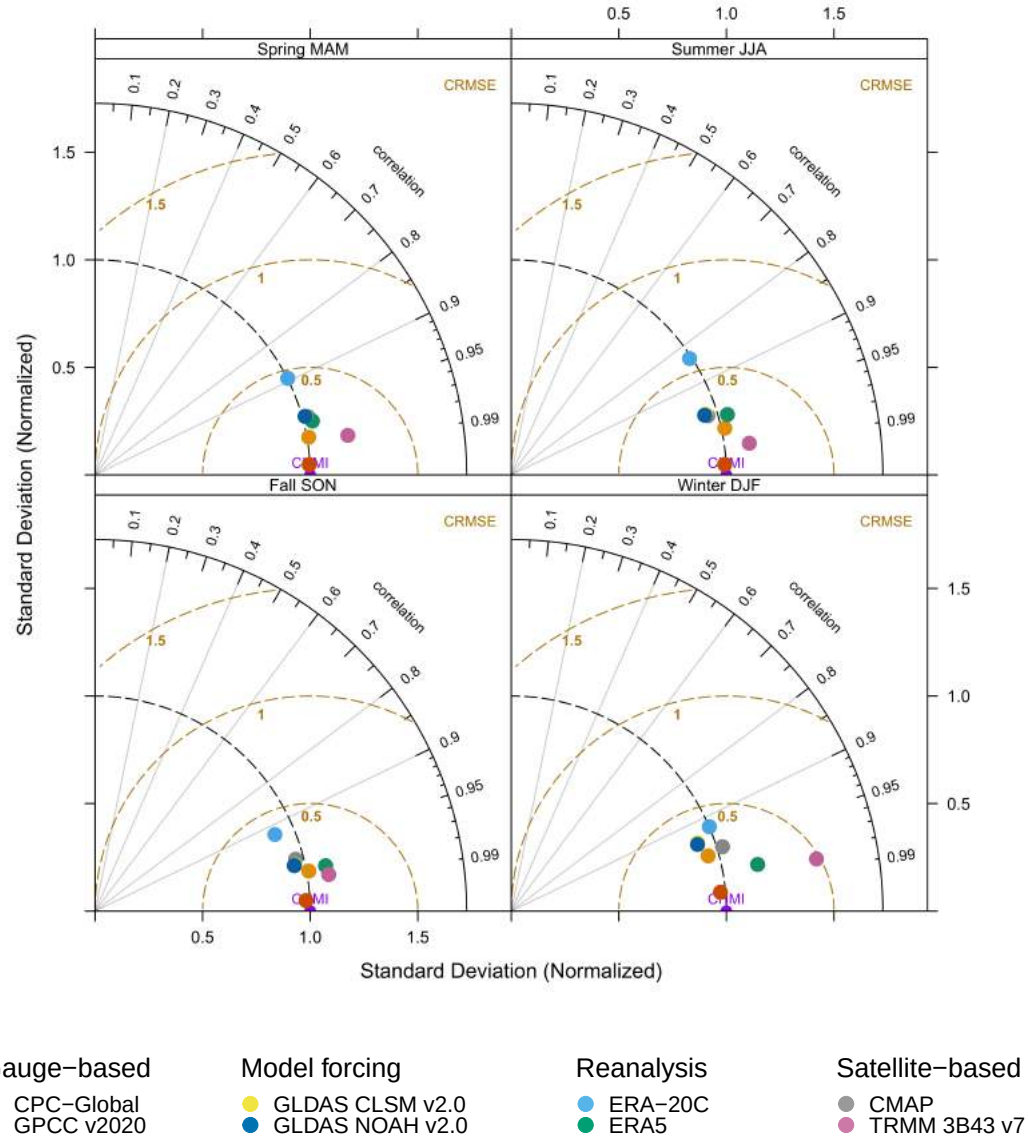


Figure 4.3: Illustration of the *plot_taylor()* function comparing selected data sets by season.

the time series in two 20-year periods (1981-2000 and 2001-2020), and examined their empirical distribution; to do so, we used the *plot_density()* function (Figure 4.4). We identify two common traits: a density peak around 50[mm] for the first 20-year period and a general widening of the density curve towards higher precipitation across the selected data sets in the last 20 years. A particular distinction in the distribution of ERA5 compared with the other data sets is that the density peak is shorter for both periods. A shorter density peak and a bigger area under the density curve to the right of said peak indicate that ERA5 precipitation estimates are higher than those of the other data sets. In line with our findings through the empirical distribution of ERA5, overestimation of precipitation has previously been identified across different regions (Hassler and Lauer, 2021).

A different approach to analyzing changes in precipitation regimes is to explore their spatial patterns. We computed the median monthly precipitation at each grid cell for two 20-year periods, and then portrayed them using the *plot_map()* function (Figure 4.5). The maps show that no drastic changes in spatial patterns took place between 1981-2000 and 2001-2020, except for a slight increase in precipitation in relatively uniform manner. Intercomparing data sets, we observe a common high precipitation center located around the Šumava Mountains on the southwestern border of Czechia. While the empirical distribution of precipitation estimates already pointed at ERA5 overestimating precipitation, it is now observable, and perhaps more easily conveyed, on the maps that ERA5 estimates are overall higher than the rest. Contrarily, we can see the lack of spatial contrast between high and low precipitation in GLDAS NOAH v2.0 estimates appearing as more homogeneous color maps. Conversely, this artifact is due to precipitation underestimation compared to the other data sets (e.g., Xue et al., 2013). We found explicitly higher estimates around the Sudetic, Šumava, and Ore Mountains, supporting previous reports that in ERA5, too much precipitation can occur on the leeward side of an orographic barrier (Lavers et al., 2022).

More details about further functions, including a simpler and fully reproducible example, can be found at `https://cran.r-project.org/web/packages/pRecipe/vignettes/pRecipe.html`.

4.4 Conclusions

The *pRecipe* package provides a common starting point for the hydrology scientific community through its homogenized database. By encompassing widely used products from multiple sources and establishing a common ground from which to start analysis, *pRecipe* guarantees a fully reproducible framework for precipitation research. Its versatility to export data at any processing stage in NetCDF (raster) or CSV (data.table) facilitates a seamless transition for the user into different R packages

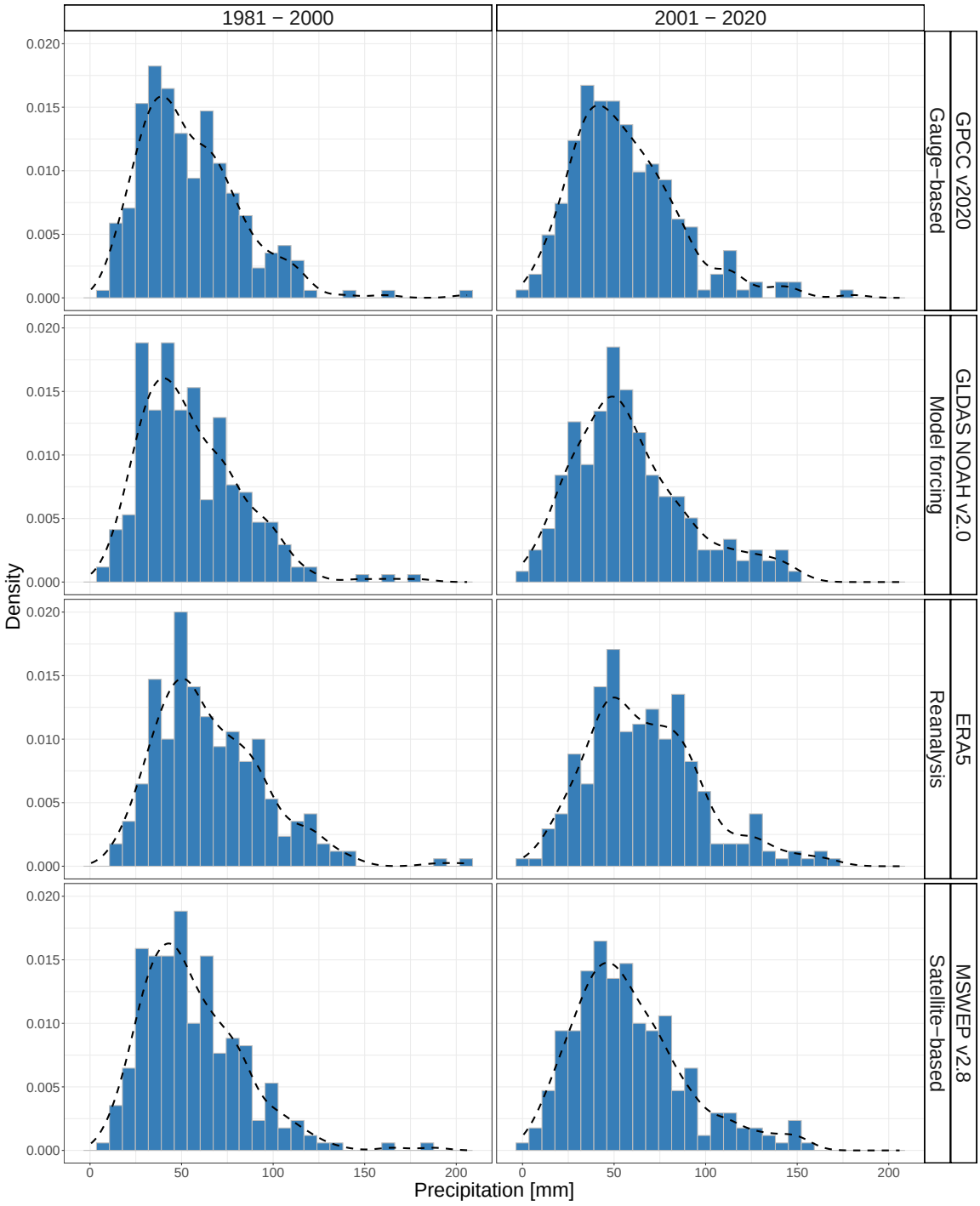


Figure 4.4: Illustration of the *plot_density()* function comparing the empirical distribution of monthly precipitation between the 1981-2000 and 2001-2020 periods for selected data sets.

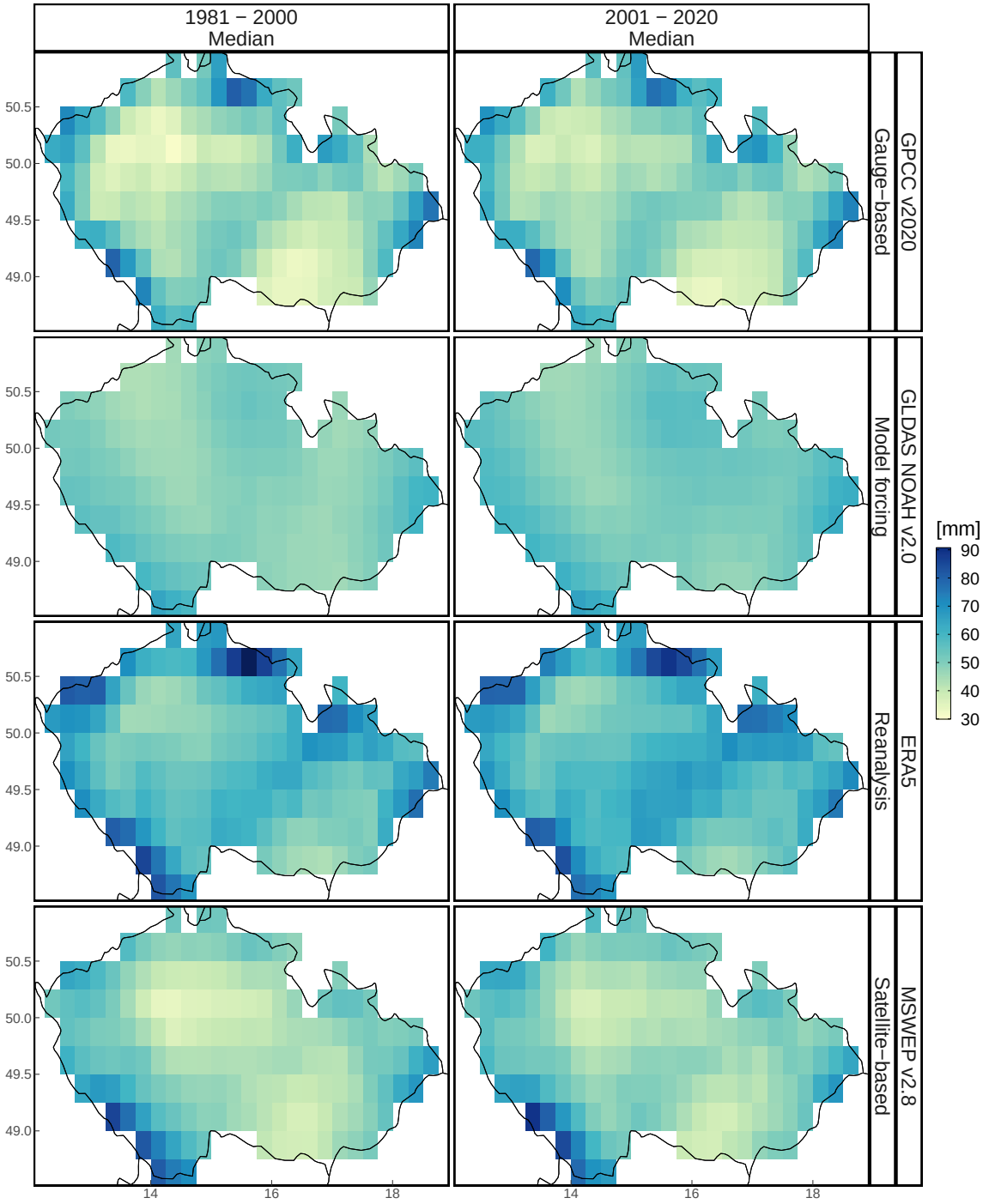


Figure 4.5: Illustration of the *plot_map()* function comparing the spatial distribution of the 20-year median precipitation between the 1981-2000 and 2001-2020 periods for selected data sets.

like *CoSMoS* (Papalexiou, 2018), *csa* (Markonis et al., 2021b), and *somspace* (Markonis and Strnad, 2020) for further analysis.

The *pRecipe* package constitutes a valuable resource for academics, government agencies, and private sector professionals because it provides a consistent and transparent approach to precipitation research. Through *pRecipe*, these users can easily access and analyze precipitation data from multiple sources, visualize various characteristics of precipitation climatology, and seamlessly transition into different R packages for further analysis. Overall, the *pRecipe* package is a powerful tool that can benefit anyone involved in precipitation research, from novice researchers to seasoned experts.

Herein, we have exemplified some of its key capabilities and showcased the ease of accessibility for the user to visualize various characteristics of precipitation climatology. We aim to provide, through *pRecipe*, an open-access database and toolbox that the hydrology community could adopt for a more consistent and reproducible science. The latter aspect will be strengthened in coming years under the following roadmap: implementation of probabilistic significance estimation for slopes/changes in 2023 and uncertainty quantification functions in 2024. At the same time, we plan to develop an evaporation twin package of *pRecipe*, with similar functionalities and integrate them together into a holistic framework for the study of terrestrial water cycle as suggested by Vargas Godoy and Markonis (2023b). To this end, we invite all scientists involved in precipitation hydroclimatology research to actively contribute with their suggestions, additions, and requests for future versions of *pRecipe*.

CRediT authorship contribution statement

Mijael Rodrigo Vargas Godoy: Conceptualization, Software, Data Curation, Writing - Original Draft, Visualization. Yannis Markonis: Conceptualization, Supervision, Writing - Review & Editing.

Chapter 5

Water Cycle Changes in Czechia: A Multi-Source Water Budget Perspective

5.1 Introduction

During the last decades, there have been significant advances in analyzing the water cycle and its response to global warming. While we expect alterations in the water cycle to respond to climate change and global warming, the actual extent and characteristics of this reaction are poorly understood (Zaitchik et al., 2023). It was hypothesized that an increased vertical gradient of atmospheric water vapor would offset atmospheric wind convergence in the tropics making wet regions wetter and dry regions drier (Held and Soden, 2006). Nevertheless, such claims lack conclusive support of observed measurements and have lit the fire of controversy in the field (Vecchi et al., 2006; Allan, 2012; Skliris et al., 2016).

Undoubtedly, the advances in remote sensing observations and process-based modeling have shaped current research the most. However, as the data sources increased, it soon became apparent that large discrepancies between the data sets still exist due to biases and uncertainties (Vargas Godoy et al., 2021). Observational data is hampered by short and heterogeneous ground-based records (Schneider et al., 2017), and unquantified uncertainties on satellite-based products (e.g., the impact of cloud filtering; Povey and Grainger, 2015). Therefore, reanalysis data providing global

coverage through models while assimilating observation-based data has attained an essential role in assessing water cycle changes (Lorenz and Kunstmann, 2012). Each data source has limitations and uncertainties; when multiple sources are combined, these can compound and result in conflicting or unclear results. Hence, in addition to uncertainty due to the complex water cycle system, which involves multiple feedback mechanisms and interactions between different components, we must account for data merge uncertainty. Accordingly, various methodologies for multi-source data integration have emerged. Among the most widely used ones are: Bayesian model averaging, constrained linear regression, neural networks, optimal interpolation, and simple weighting (Rodgers, 2000; Aires, 2014; Moazamnia et al., 2019; Pellet et al., 2019; Xiao et al., 2020). Subsequently, once merged data is generated, it is subject to post-processing for water cycle budget closure via Monte Carlo applications and Kalman filter variations (Pan and Wood, 2006).

Several studies have quantified the water cycle by implementing data integration methods and budget closure constraints, e.g.,: Sahoo et al. (2011) integrated 16 data sets over 10 globally distributed river basins (eight for precipitation, six for evapotranspiration, one for runoff, and one for total water storage; Table 5.1). Pan et al. (2012) integrated eight data sets over 32 globally distributed river basins (four for precipitation, two for evapotranspiration, one for runoff, and one for total water storage; Table 5.2). Rodell et al. (2015), integrated six data sets over continents and ocean basins (one for precipitation, three for evapotranspiration, one for runoff, and one for total water storage; Table 5.3). Zhang et al. (2016), integrated 14 data sets globally (five for precipitation, six for evapotranspiration, one for runoff, and two for total water storage; Table 5.4). Munier and Aires (2018) integrated 12 data sets at the global scale (four for precipitation, three for evapotranspiration, one for runoff, and four for total water storage; Table 5.5).

The studies mentioned above focus on merging multiple data sets to end up with a single data set per water cycle component at different spatial scales. It is evident that unconstrained uncertainty remains despite the plethora of data products derived from satellites, ground-based measurements, and climate models. This is true even for localized studies at regional scales where "ground-truth" measurements for one or more components of the water cycle are available. One region of particular interest is Czechia, a small country in Central Europe with diverse landscapes and a growing population (United Nations, 2022). The water cycle over Czechia has been experiencing significant changes in recent times, affecting various aspects of the water balance in the region, including changes in river flow regimes and water quality, loss of wetlands, and changes in the frequency and severity of extreme events (Mozny et al., 2020). Besides, changes in the rainfall-snowfall partition have given rise to a decrease in snow cover and premature snowmelt (Nedelcev and Jenicek, 2021). These

Table 5.1: Compiled from Sahoo et al. (2011). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage.

Name	Variable	Spatial Resolution	Temporal Resolution	Record Length	Data Type	Reference(s)
GPCP	P	1°	Daily	1997-2006	Satellite-based	Adler et al. (2003)
TMPA 3B42RT	P	0.25°	3h	1997-2019	Satellite-based	Huffman et al. (2007)
CMORPH	P	8km	30min	2003-2006	Satellite-based	Joyce et al. (2004)
PERSIANN	P	0.25°	3h	2000-2006	Satellite-based	Hong et al. (2004)
CPC PRECŁ	P	2.5°	Monthly	1950-Present	Gauge-based	Chen et al. (2002)
CRU TS3.0	P	0.5°	Monthly	1901-2006	Gauge-based	Mitchell and Jones (2005)
WM v2.01	P	0.5°	Monthly	1900-2008	Gauge-based	Willmott and Matsuura (2001)
GPCC	P	0.5°	Monthly	1900-2007	Gauge-based	Schneider et al. (2011)
PM (ISCCP)	E	2.5°	3h	1984-2005	Satellite-based	Sheffield et al. (2010)
PM (EOS)	E	5km	Daily	2003-2006	Satellite-based	Vinukollu et al. (2011b)
PT (EOS)	E	5km	Daily	2003-2006	Satellite-based	Vinukollu et al. (2011b)
SEBS (EOS)	E	5km	Daily	2003-2006	Satellite-based	Vinukollu et al. (2011b)
VIC	E	1.0°	3h	1948-2006	Model	Sheffield and Wood (2007)
ERA-interim	E	T255	12h	1989-2006	Reanalysis	Simmons (2006)
GRACE	Δ TWS	Basin	∼Monthly	2002-2006	Satellite-based	Swenson and Wahr (2006)
GRDC	Q	Basin	Monthly	1900-2006	Station	`www.bafg.de/GRDC`

Table 5.2: Compiled from Pan et al. (2012). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage.

Name	Variable	Spatial Resolution	Temporal Resolution	Record Length	Data Type	Reference(s)
GPCP v2.2	P	2.5°	Monthly	1950-Present	Gauge-based	Adler et al. (2003)
CRU TS3.0	P	0.5°	Monthly	1901-2006	Gauge-based	Mitchell and Jones (2005)
WM v2.01	P	0.5°	Monthly	1900-2008	Gauge-based	Willmott and Matsuura (2001)
GPCC	P	0.5°	Monthly	1900-2007	Gauge-based	Schneider et al. (2011)
MPI	E	0.5°	Monthly	1982–2008	Flux tower-based	Jung et al. (2010)
SEBS (EOS)	E	5km	Daily	2003-2006	Satellite-based	Vinukollu et al. (2011b)
GRACE	Δ TWS	Basin	∼Monthly	2002-2006	Satellite-based	Swenson and Wahr (2006)
GRDC	Q	Basin	Monthly	1900-2006	Station	`www.bafg.de/GRDC`

Table 5.3: Compiled from Rodell et al. (2015). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage.

Name	Variable	Spatial Resolution	Temporal Resolution	Record Length	Data Type	Reference(s)
GPCP v2.2	P	1°	Daily	1997-2006	Satellite-based	Adler et al. (2003)
						Huffman et al. (2009)
Princeton ET	E	5km	Daily	2003-2006	Satellite-based	Vinukollu et al. (2011b)
MERRA	E	0.5°x0.667°	Hourly	1980-2016	Reanalysis	Rienecker et al. (2011)
						Bosilovich et al. (2011)
						Reichle (2012)
GLDAS	E	0.25°	3h	1948-2014	Model	Roderick et al. (2014)
U of Washington	Q	2°	Monthly	1998–2008	Model	Jung et al. (2010)
GRACE	Δ TWS	Basin	∼Monthly	2002-2006	Satellite-based	Swenson and Wahr (2006)

Table 5.4: Compiled from Zhang et al. (2016). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage.

Name	Variable	Spatial Resolution	Temporal Resolution	Record Length	Data Type	Reference(s)
CSU	P	0.25°	3h	1998–2010	Satellite-based	Bytheway and Kummerow (2013)
PGF	P	0.25°	3h	1948–2010	Satellite-based	Sheffield et al. (2006)
CHIRPS	P	0.5°	Monthly	1981–present	Satellite-based	Funk et al. (2014)
GPCC(v6)	P	0.5°	Monthly	1901–2010	Gauge-based	Schneider et al. (2014)
TMPA-RT	P	0.25°	Monthly	2001-2019	Satellite-based	Huffman et al. (2007, 2010)
SRB–PGF–PM	E	0.5°	3h	1984–2007	Satellite-based	Vinukollu et al. (2011b)
VIC	E	0.25°	3h	1948-2010	Model	Sheffield and Wood (2007)
ERA-interim	E	T255	12h	1989-2006	Reanalysis	Simmons (2006)
MERRA	E	0.5°x0.667°	Hourly	1980-2016	Reanalysis	Rienecker et al. (2011)
GLEAM	E	0.5°	3h	1984-2017	Satellite-based	Miralles et al. (2011)
SRB-CFSR-SEBS	E	0.5°	Daily	1984–2007	Satellite-based	Vinukollu et al. (2011b)
SRB-CFSR-PM	E	0.5°	Daily	1984–2007	Satellite-based	Vinukollu et al. (2011b)
SRB-CFSR-PT	E	0.5°	Daily	1984–2007	Satellite-based	Vinukollu et al. (2011b)
VIC	Q	0.25°	3h	1948-2010	Model	Sheffield and Wood (2007)
VIC	Δ TWS	0.25°	3h	1948-2010	Model	Sheffield and Wood (2007)
GRACE	Δ TWS	1°	Monthly	2002-present	Satellite-based	Landerer and Swenson (2012)

Table 5.5: Compiled from Munier and Aires (2018). P is precipitation, E is evapotranspiration, Q is runoff, and Δ TWS is changes in total water storage.

Name	Variable	Spatial Resolution	Temporal Resolution	Record Length	Data Type	Reference(s)
TMPA	P	0.25°	Monthly	1998-2019	Satellite-based	Huffman et al. (2007)
CMORPH	P	0.25°	Daily	1998–present	Satellite-based	Sheffield et al. (2006)
NRL	P	0.25°	12h	2003–2010	Satellite-based	Turk et al. (2010)
GPCP	P	2.5°	Monthly	1979–present	Satellite-based	Schneider et al. (2014)
GLEAM	E	0.25°	3h	1980–2011	Satellite-based	Miralles et al. (2011)
MOD16	E	1km	8-day	2000–2012	Satellite-based	Mu et al. (2007)
NTSG	E	8km	Daily	1983–2006	Satellite-based	Zhang et al. (2010)
CSR	Δ TWS	Basin	Monthly	2002-present	Satellite-based	`http://grace.jpl.nasa.gov/data/`
GFZ	Δ TWS	Basin	Monthly	2002-present	Satellite-based	`http://grace.jpl.nasa.gov/data/`
JPL	Δ TWS	Basin	Monthly	2002-present	Satellite-based	`http://grace.jpl.nasa.gov/data/`
GRGS	Δ TWS	Basin	Monthly	2002-present	Satellite-based	`http://grgs.obs-mip.fr/grace/`
GRDC	Q	Basin	Monthly	1900-present	Station	`http://www.grdc.sr.unh.edu/`

changes in the water cycle are expected to continue in the near-future (Kyselý and Beranová, 2009; Jenicek et al., 2021). Precipitation, in particular, is expected to increase its mean mainly in winter and extreme rates throughout the year (Kyselý et al., 2011). In addition, increased human activities, such as urbanization and agriculture, have led to changes in land use and land cover, which in turn has contributed to the occurrence of floods and droughts (Svoboda et al., 2016). Droughts, have had disastrous consequences for agriculture, forestry, water management, and other human activities (Brázdil et al., 2009). Consequently, the water cycle in Czechia and human activity find themselves on a causal feedback loop.

In this study, we aim to estimate the water cycle changes over Czechia between the 1961-1990 and 1991-2020 periods, and determine the current trends and patterns in water cycle components. Our analysis includes various data sets at different spatiotemporal scales allowing us to assess 96 data combinations for budget closure. Rather than enforcing budget closure on a multi-source integrated data set or assessing different integration methods, we explored an empirical method to rank how multiple data set combinations close the water cycle budget while correlating to referential data estimates of individual water cycle components. In this manner, we are not generating yet another new data set but are identifying the best combination among the data sets available for our study domain. Only the data sets with the best rankings as determined by our proposed benchmarking were used in all subsequent computations. We found that hydroclimatic models, as expected, have better water budget closure. However, ERA5-Land is not far off despite known non-closure limitations associated with reanalyses. We identified an overall acceleration of atmospheric water fluxes. Simultaneously, we report a heterogeneous distribution of freshwater availability.

5.2 Data and Methods

5.2.1 Study Area

Czechia is a landlocked (surrounded by Germany, Austria, Slovakia, and Poland) European country that covers an area of 78 864 km^2. Czechia is an essential headwaters region of the European continent. The country is home to several large rivers, including the Vltava, the Elbe, the Morava, and the Oder, all of which have their sources within it. Czechia is situated at the intersection of three sea drainage basins: the North Sea, the Baltic Sea, and the Black Sea, which, in return, divide Czechia into three main hydrological catchment areas: the Elbe, Oder, and Danube basins (Figure 5.1). All of these major watercourses drain water into neighboring states. The water sources of

Czechia are thus almost exclusively dependent on precipitation.

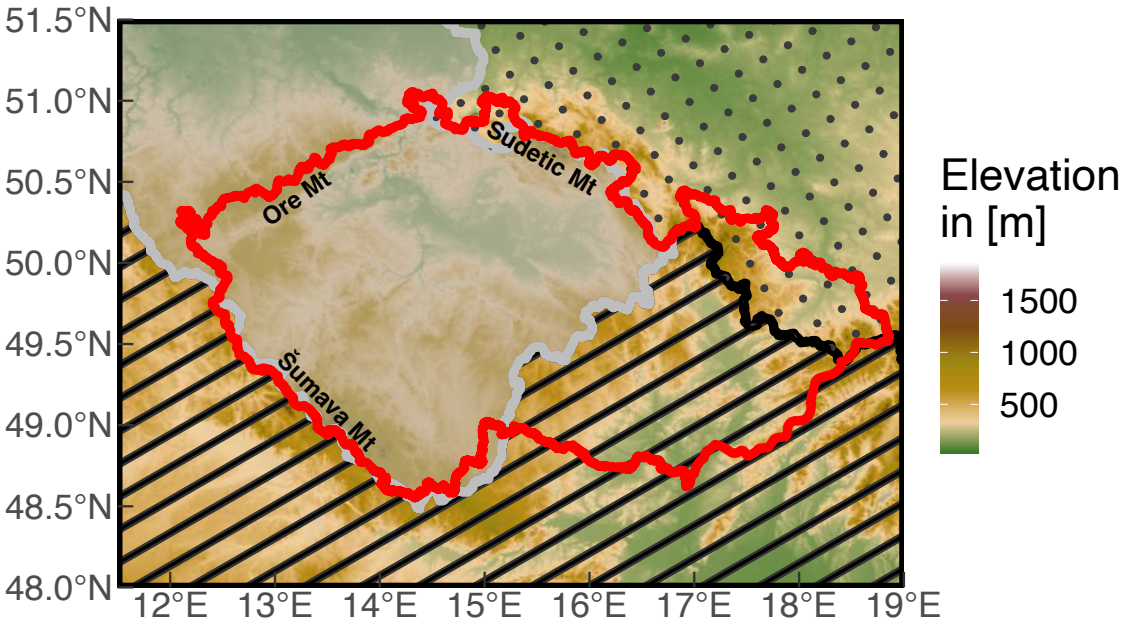


Figure 5.1: The three drainage basins within Czechia's administrative boundaries (red line). Elbe (light gray shade), Danube (black stripes), and Oder (dark gray points).

5.2.2 Data

To assess water cycle acceleration we gathered data sets with at least 60 years of record. This first filter reduced the plethora of publicly available data sets to nine data sets from multiple sources (observation-based, reanalysis, and hydrological model products) plus three evaluation references (Table 5.6). The evaluation data sets for precipitation and runoff are the Czech Hydrometeorological Institute (CHMI)and the Global Runoff Data Centre (GRDC), respectively. Six precipitation data sets: Climatic Research Unit at the University of East Anglia (CRU TS v4.06; Harris et al. (2020)), European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis (ERA5-Land; Muñoz-Sabater et al. (2021)), the E-OBS data set from the Copernicus Climate Change Service (Cornes et al., 2018), National Centers for Environmental Prediction & the National Center for Atmospheric Research Reanalysis One (NCEP/NCAR R1; Kalnay et al. (1996)), Precipitation Reconstruction Over Land (PREC/L; Chen et al. (2002)), and TerraClimate (Abatzoglou et al., 2018). Note that, E-OBS (hereinafter mHM(E-OBS)) was used as meteorologic input for the mesoscale Hydrologic Model (mHM; Samaniego et al. (2010); Kumar et al. (2013)). Four evapotranspiration data sets:

ERA5-Land, mHM, NCEP/NCAR R1, and TerraClimate. Four runoff data sets: ERA5-Land, mHM, NCEP/NCAR R1, and TerraClimate. Using the above listed data sets we assessed a total of 96 different combinations.

5.2.2.1 Evaluation References

As evaluation references, we relied solely on ground station data sets. A distinct advantage of station data over hydrological models or reanalyses is their capability to capture detailed and localized information. These in-situ measurements directly reflect the local climatic conditions, offering a more accurate representation of the water cycle.

The Czech Hydrometeorological Institute (CHMI) provides station derived precipitation data. The CHMI station network consists of approximately 700 stations distributed with a mean density of one station per each 100 km^2, adequately representing the distinct geographical features of Czechia (Kašpar et al., 2021). Although the data collection and related services for a specific station are generally managed by the regional branches of CHMI, the entire territory station data can be accessed from the Department of Climatology of CHMI at once. All the data sets are undergone robust quality control checks by CHMI before being added to the database. Herein, we gathered the country level estimates calculated by CHMI (one value per month) for a period of 60 years (1961-2020).

The Global Runoff Data Centre (GRDC) is a collection of river discharge data from more than 8000 stations in 157 countries. The GRDC operates under the the World Meteorological Organization (WMO) since 1988 to collect, manage, and distribute data related to river discharge and runoff from around the world. The data collected at GRDC undergoes quality control to check for errors, inconsistencies, and outliers in the data before its dissemination. While data is available at daily and monthly time step, the record length varies by location. We selected three stations from GRDC, namely the Bohumin (Oder), Decin (Elbe), and Moravsky Jan (Danube) stations, which are placed near the borders of the country and country level estimates were calculated by their wieighted average based the catchment area as registered by GRDC.

5.2.2.2 Observational-based Products

CRU TS is a popularly used gridded data set generated by the University of East Anglia's Climate Research Unit (Harris et al., 2020). It is known for its historical long-term coverage, which is available from 1901 to the near present. The data set comes with a 0.5° spatial resolution at the monthly scale. It compiles station data from multiple sources such as the Food and Agricultural Organisation (FAO),

Table 5.6: Data set description. P is precipitation, E is evapotranspiration, and Q is runoff.

Name	**Variable(s)**	**Spatial Resolution**	**Temporal Resolution**	**Record Length**	**Data Type**	**Reference**
CHMI	P	Point	Daily	1961-2020	Stations	`http://portal.chmi.cz`
CRU TS v4.06	P	1°	Monthly	1901-2020	Gauge	Harris et al. (2020)
E-OBS	P	0.125°	Daily	1950-2020	Gauge	Cornes et al. (2018)
ERA5-Land	P, E, Q	0.1°	Monthly	1950-2020	Reanalysis	Muñoz-Sabater et al. (2021)
GRDC	Q	Point	Daily	1921-2017	Stations	`www.bafg.de/GRDC`
mHM	E, Q	0.125°	Daily	1950-2020	Model	Samaniego et al. (2010)
NCEP/NCAR R1	P, E, Q	T62	Monthly	1948-2020	Reanalysis	Kalnay et al. (1996)
PREC/L	P	0.5°	Monthly	1948-2020	Gauge	Chen et al. (2002)
TerraClimate	P, E, Q	4 km	Monthly	1958-2020	Model	Abatzoglou et al. (2018)

the World Meteorological Organisation (WMO), and the National Meteorological Agencies (NMA's) (Sun et al., 2018). CRU TS v4, its latest version, implemented angular distance based interpolation to facilitate tracing back the stations upon which the gridded data set has been constructed.

PREC/L, created by the US Climate Prediction Center (CPC), is a gridded product entirely based on the station data set (Chen et al., 2002) with global coverage and monthly time step. PREC/L draws data from over 17 000 stations from the Global Historical Climatology Network version2 (GHCN v2; Peterson and Vose, 1997) and the Climate Anomaly Monitoring System (CAMS; Janowiak and Xie, 1999). Subsequently, the data is interpolated to construct the gridded product at three different resolutions (0.5°, 1°, and 2.5°). Herein, we used the 0.5° monthly precipitation, whose record extends from 1948 to the present.

5.2.2.3 Hydrological Models

The mesoscale Hydrologic Model (mHM; Samaniego et al., 2010; Kumar et al., 2013) is a conceptual grid-based model representing dominant hydrological fluxes and storage at the Earth's surface and subsurface through a system of ordinary differential equations. mHM represents processes such as interception, snow, soil moisture, evapotranspiration, and various runoff components like fast/slow interflow and baseflow. The model was established, parameterized and evaluated over the European continent (Rakovec et al., 2016b; Samaniego et al., 2019; Rakovec et al., 2022). The meteorological inputs were based on daily E-OBS data (Cornes et al., 2018) of precipitation in addition to minimum, maximum and average temperature. The potential evapotranspiration was derived using the method of (Hargreaves and Samani, 1982). The spatial resolution of the model grid corresponds to 0.125°.

Terraclimate is a high-resolution gridded global climate data set that provides the mean climate and mean water balance data covering a time span of 1958 to the present (Abatzoglou et al., 2018). The data set is commonly known for its high spatial resolution (4 km). It uses various global gridded climate data sets such as WorldClim v2 (Fick and Hijmans, 2017) and v1.4 (Hijmans et al., 2005), CRU TS v4 (Harris et al., 2020), Japanese 55-year Reanalysis (JRA55) (Kobayashi et al., 2015), and Root zone storage capacity (Wang-Erlandsson et al., 2016) in order to generate the high-resolution monthly climate variables time series at the global level. An additional advantage of the Terraclimate is that it produces monthly surface water balance based on a water balance model along with primary climatic variables such as temperature, precipitation, solar radiation, etc.

5.2.2.4 Reanalyses

ERA5-Land is the latest fifth-generation global atmospheric reanalysis product developed by the European Center for Medium-Range Weather Forecast (ECMWF) (Muñoz-Sabater et al., 2021). ERA5-Land, as the name implies, builds upon the terrestrial component of ERA5 and downscales the model spatial grid resolution from 31 km into 9 km. As a result, ERA5-Land delivers either hourly or monthly estimates with a spatial resolution of 0.1°. Given its high spatiotemporal resolution and long record, ERA5-Land provides valuable data for comprehensive analysis and diverse hydrological applications at the global scale.

The NCEP/NCAR Reanalysis project one is produced by the collaboration between the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) (Kalnay et al., 1996). It is the longest-running reanalysis that uses rawindsonde data, at the expense that the model and data assimilation scheme are antiquated Trenberth et al. (2011). The data set is distributed on a T62 Gaussian grid (approximately 1.875° at the equator) and its record start dates back to 1948.

5.2.3 Data Evaluation

We validated the gathered data sets to capture the temporal variability of water cycle components as described by the three observational references via the coefficient of determination (R-squared or R^2) and the Root Mean Square Error (RMSE). All data sets were spatial weighted averaged over Czechia and temporally aggregated to an annual scale over the calendar year. Note that only precipitation data sets could be evaluated over the entire 60-year period of 1961-2020. In contrast, runoff was evaluated over 1961-2017. In order to compare a 30-year mean among all water cycle components, the common period of 1981-2010 was selected.

5.2.4 Data Set Ranking

A success metric widely used among several studies is getting the budget closure residual (ξ) as close to zero as possible. Herein, we define the budget closure residual as follows:

$$\xi_n = P_n - E_n - Q_n \tag{5.1}$$

where P_n is precipitation, E_n is evapotranspiration, and Q_n is runoff for a given year n. Thus, we have 60 annual values for each of the 96 possible combinations. Note that, the water flux time

series used to compute the residuals are the spatial weighted average values. Under steady state conditions the mean of these residuals should tend to zero:

$$\overline{\xi_i} = \frac{\sum_{n=1}^{N} \xi_n}{N} \rightarrow 0 \tag{5.2}$$

where $\overline{\xi_i}$ is the mean of the $N = 60$ annual residuals for the i-th combination. The score to be used in the ranking of a given data set combination was determined via:

$$score = \frac{|\overline{\xi_i}|\sigma_{\xi_i}}{\left(cor(P_i - E_i, Q_i)cor(P_i, P_o)cor(E_i, E_o)cor(Q_i, Q_o)\right)^2} \tag{5.3}$$

where $|\overline{\xi_i}|$ is the absolute value of the mean of the 60 annual residuals for the i-th combination, σ_{ξ_i} is the standard deviation of the 60 annual residuals for the i-th combination, $cor(P_i - E_i, Q_i)$ is the correlation between $P - E$ and Q for the i-th combination, $cor(P_i, P_o)$ is the correlation between P of the i-th combination and the precipitation evaluation reference, $cor(E_i, E_o)$ is the correlation between E of the i-th combination and the evapotranspiration evaluation reference, and $cor(Q_i, Q_o)$ is the correlation between Q of the i-th combination and the runoff evaluation reference. The ranking method proposed herein can easily be applied to any other referential data set for evaluation. In data-limited areas or those with a poor observational network, the ranking method may still be applied using external data as an evaluation reference, or the corresponding term in the equation can be simply left out. E.g., if evapotranspiration data for evaluation is not available, Equation 5.3 becomes:

$$score = \frac{|\overline{\xi_i}|\sigma_{\xi_i}}{\left(cor(P_i - E_i, Q_i)cor(P_i, P_o)cor(Q_i, Q_o)\right)^2}$$

In the case of Czechia, we used this modified version due to the absence of access to observational evapotranspiration data.

5.2.5 Water Cycle Changes

We assessed the empirical distribution of spatial weighted average values (accounting for the area of each grid cell in proportion to the total area being averaged) of annual water cycle fluxes between 1961-1990 and 1991-2020 for three of the best data set combinations. To account for the influence of extreme value in the latter period due to the 100-year drought of 2003 (Brázdil et al., 2013), we compared the median values rather than their means (see Figure 5.5). To deepen our assessment of changes in the distribution of water cycle fluxes, we compared their monthly values between

1961-1990 and 1991-2020. To determine the statistical significance of the above-mentioned changes, we employed non-parametric bootstrapping of 10 000 iterations. Subsequently, we performed an analogous analysis in space. We computed the change in the median values between 1961-1990 and 1991-2020 over each grid cell. Note that each data set was assessed at its native resolution for this part of the analysis. Finally, we examined the change patterns of water cycles through the seasons. Herein, we considered: winter as December, January, and February; spring as March, April, and May; summer as June, July, and August; autumn as September, October, and November.

5.3 Results

5.3.1 Benchmarking water cycle components

Our analysis describes the most recent spatiotemporal changes on the water cycle in Czechia. For starters, we examined precipitation, evapotranspiration, and runoff estimates from the gathered data sets. Further, precipitation and runoff were compared to CHMI (Figure 5.2a) and GRDC (Figure 5.2c) as the respective evaluation references. The variability of estimates from precipitaion and runoff data sets (Figure 5.2a and c) visibly have a broader spread than those of evapotranspiration (Figure 5.2b). While one may suspect the spread in precipitation is due to the higher number of data sets available, they correlate better to their evaluation reference than runoff. The data set with the highest correlation values for precipitation is mHM(E-OBS) with R-squared of approximately 0.99 (Figure 5.2a). mHM has the highest correlation for runoff, with R-squared circa 0.93 (Figure 5.2c). In contrast, NCEP/NCAR R1 consistently reports the lowest correlation values regardless of the water flux of interest. Additionally it has substantially higher RMSE values than the rest of the data sets for precipitation. To some degree, ERA5-Land is the in-betweener data set because it has high correlation values and simultaneously has high RMSE for precipitation, yet for runoff, ERA5-Land exhibits moderate correlation and small RMSE.

The water cycle budget is meant to close over hydrological units. Accordingly, we examined the water fluxes of the data sets with the best evaluation over the subbasins enclosed by the Czech administrative borders (Figure 5.3). For simplicity, we will refer to them by their river names inside Czechia. I.e., Morava for the Danube basin, Labe for the Elbe basin, and the Odra for the Oder basin. It can be seen that within each data set, no extremely deviant behavior is exhibited between basins or at the country level. In other words, the precipitation time series depicted by TerraClimate for Czechia is similar to the one depicted for the Morava, Labe, and Odra Rivers. Comparing data sets,

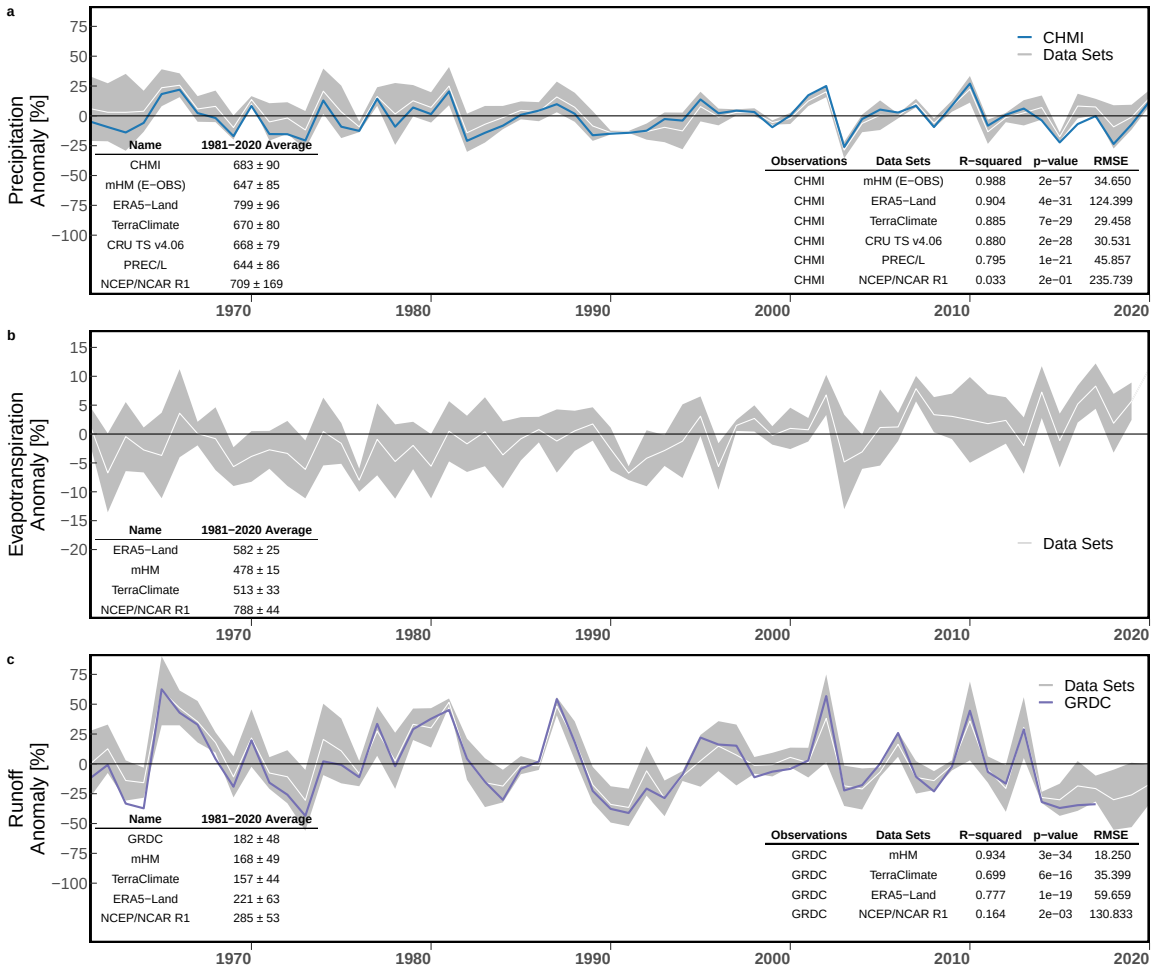


Figure 5.2: Benchmarking spatial weighted average annual water fluxes over Czechia between 1961 and 2020. For consistency and comparability between different water fluxes, annual anomalies were computed using the 1981-2010 average as a reference, the common period among all data sets. The 1981-2010 average and standard deviation are listed at the bottom left of each panel. Linear correlation summary statistics are displayed at the bottom right of each panel. The spread of the estimates being evaluated is shown in gray, and their mean is in white. (a) Precipitation evaluation. CHMI data is shown in blue. (b) Evapotranspiration evaluation. (c) Runoff evaluation. GRDC data is shown in purple.

however, it is evident that ERA5-Land is different. At first glance, we evince higher magnitudes for ERA5-Land precipitation and evapotranspiration, yet the residuals do not appear to be that far off from those of mHM or TerraClimate. It is not until we look at the cumulative sum of the residuals that we can distinguish ERA5-Land water budget residuals are nonstationary with a decreasing trend.

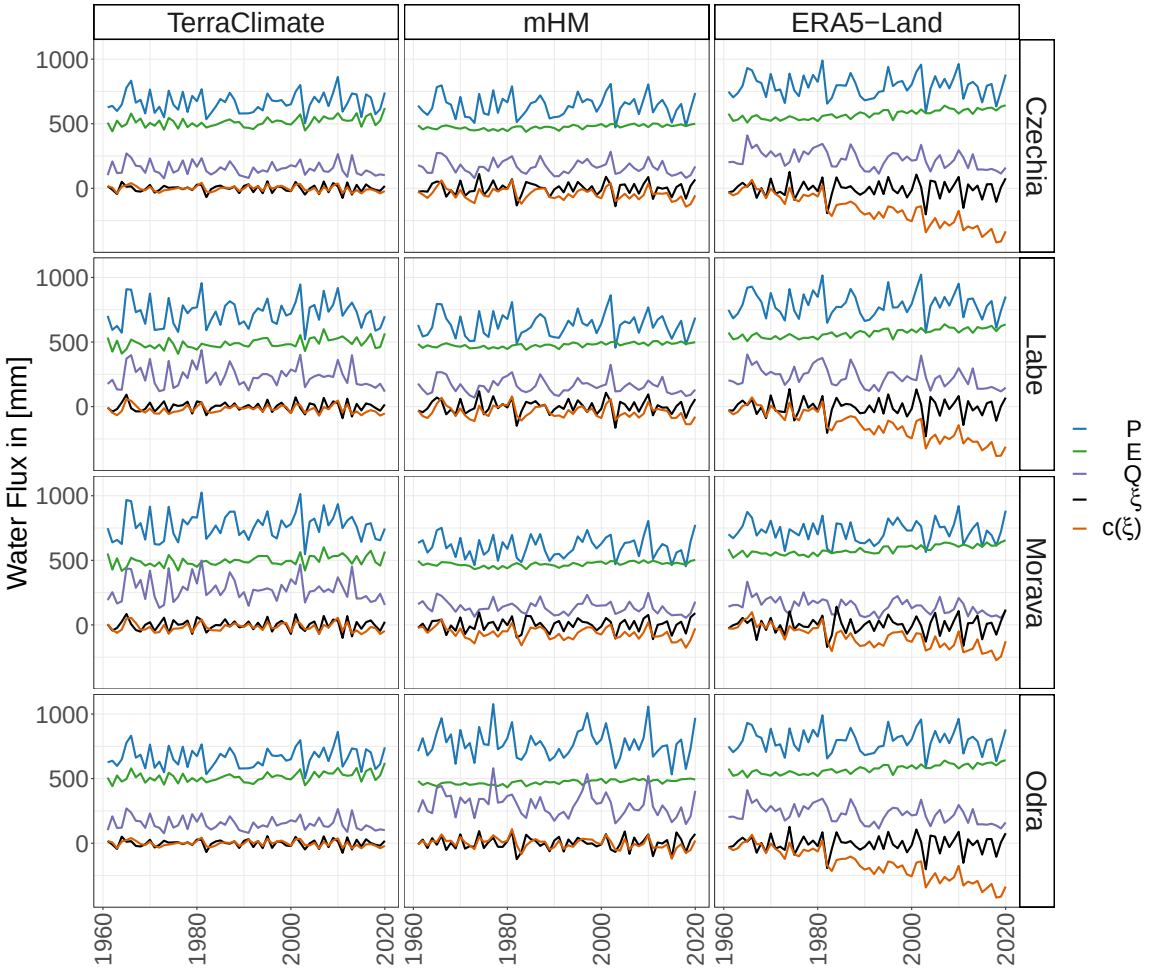


Figure 5.3: Spatial weighted average annual water fluxes over Czechia (first row), Labe River (Elbe basin inside Czechia; second row), Morava River (Danube basin inside Czechia; third row), and Odra River (Oder basin inside Czechia; fourth row). Where P is precipitation in blue, E is evapotranspiration in green, Q is runoff in purple, ξ is the residual $(P - E - Q)$ in black, and c(ξ) is the cumulative sum of the residual in orange. Left column: TerraClimate (P), TerraClimate (E), and TerraClimate (Q). Middle column: mHM(E-OBS) (P), mHM (E), and mHM (Q). Right column: ERA5-Land (P), ERA5-Land (E), and ERA5-Land (Q).

It would be sensible to use the best data set for each water flux to proceed with further analysis. However, we first verified if the best data sets individually would depict the best water cycle budget in conjunction. Conventional metrics like R-squared and RMSE cannot be directly applied to a

combination of data sets. We defined an empirical scoring metric, as described by Equation 5.3, where the smallest the value, the better the data set combination. While our ranking approach is empirical and simple, Equation 5.3 correctly identifies narrow distribution centered mean zero with higher ranked positions compared to wider distributions centered around positive or negative values (Figure 5.4). Upon ranking all 96 possible combinations (Table 5.7), we observe that even though mHM outperformed TerraClimate for individual water flux estimates, the TerraClimate exclusive combination offers the best water budget closure. We expected combinations with hydrological model data to be highly ranked and reanalyses to be poorly ranked due to the above-reported considerable biases of the latter. Notwithstanding, we were surprised to see the ERA5-Land exclusive combination (i.e., all flux estimates from the same data set) among the top six ranks despite non steady water budget residuals (Figure 5.3) as well as biases 1.7-3.3 and 3.8-4.2 times larger than those of models for runoff (Figure 2.4c) and precipitation (Figure 2.4a), respectively. The first combination that includes at least one estimate from NCEP/NCAR R1 is at the 38th rank, and the NCEP/NCAR R1 exclusive combination is at the 87th rank.

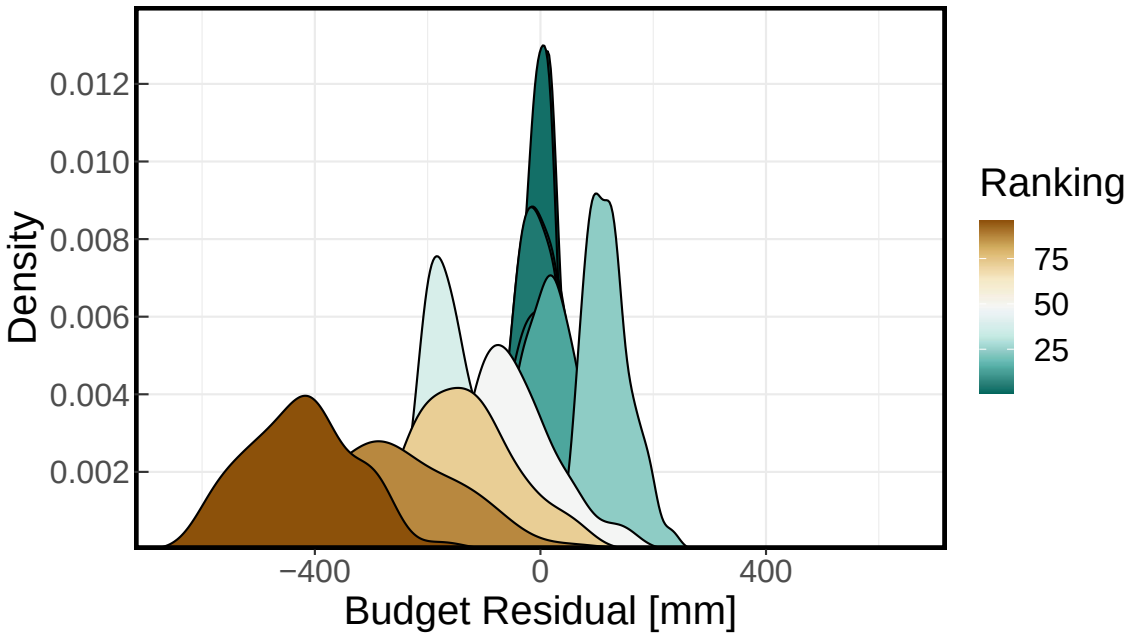


Figure 5.4: Empirical distribution of the data set combinations listed on Table 5.7 colored based on their ranking as determined by Equation 5.3. The color gradient goes from higher ranked combinations colored in shades green to lower ranked combinations colored in shades of brown.

Table 5.7: Data set ranking as determined by Equation 5.3. P is precipitation, E is evapotranspiration, Q is runoff, $\bar{\xi}$ is the mean residual over 60 years, σ_ξ is the standard deviation of the residual over 60 years, $cor(P-E,Q)$ is the correlation between $P-E$ and Q for the i-th ranked combination, $cor(P,P_\mathrm{o})$ is the correlation between P of the i-th ranked combination and CHMI, and $cor(Q,Q_\mathrm{o})$ is the correlation between Q of the i-th ranked combination and GRDC.

Ranking	P	E	Q	$\bar{\xi}$	σ_ξ	$cor(P-E,Q)$	$cor(P,\mathbf{P_o})$	$cor(Q,\mathbf{Q_o})$
1st	TerraClimate	TerraClimate	TerraClimate	-0.346	30.204	0.846	0.941	0.836
2nd	mHM(E-OBS)	mHM	mHM	-0.912	51.231	0.816	0.994	0.967
3rd	CRU TS v4.06	TerraClimate	TerraClimate	-1.749	29.944	0.843	0.938	0.836
4th	TerraClimate	TerraClimate	mHM	-8.861	39.847	0.730	0.941	0.967
5th	CRU TS v4.06	TerraClimate	mHM	-10.265	40.613	0.711	0.938	0.967
6th	ERA5-Land	ERA5-Land	ERA5-Land	-5.554	66.606	0.701	0.951	0.882
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
14th	PRECL/L	mHM	TerraClimate	17.013	60.281	0.658	0.891	0.836
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
24th	ERA5-Land	TerraClimate	mHM	114.628	44.721	0.763	0.951	0.967
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
38th	ERA5-Land	NCEP/NCAR R1	mHM	-166.746	60.420	0.714	0.951	0.967
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
48th	PREC/L	mHM	ERA5-Land	-52.549	82.751	0.382	0.891	0.882
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
72nd	mHM(E-OBS)	mHM	NCEP/NCAR R1	-134.044	87.923	0.237	0.994	0.405
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
87th	NCEP/NCAR R1	NCEP/NCAR R1	NCEP/NCAR R1	-292.024	137.297	0.675	0.181	0.405
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
96th	CRU TS v4.06	NCEP/NCAR R1	NCEP/NCAR R1	-424.772	93.962	-0.019	0.938	0.405

5.3.2 Temporal changes in the water cycle

Moving forward, we computed the change in water fluxes' annual distribution via shifts on their 30-year median (Figure 5.5). Also, we assessed the statistical significance of the observed change in the medians by non-parametric bootstrapping (10 000 iterations). Hereupon, we will report results only for the first- (TerraClimate exclusive), second- (mHM exclusive), and sixth-ranked (ERA5-Land exclusive) data combinations. Because the third- (CRU TS v4.06, TerraClimate, Terraclimate), fourth- (TerraClimate, TerraClimate, mHM), and fifth-ranked (CRU TS v4.06, TerraClimate, mHM) data combinations have a single data set different from the first- and second-ranked ones, as such, we would be showing the same plots and statistics multiple times. TerraClimate and mHM show similar increases in precipitation and evapotranspiration circa 20 mm, but only evapotranspiration manifests a statistically significant change (p ¡ 0.01). Evapotranspiration changes underwhelming those of precipitation stand further accentuated in ERA5-Land, whose magnitude of the change in evapotranspiration is almost 60 mm and in precipitation is less than -1 mm. Another peculiarity of ERA5-Land is that runoff, with a change of -56 mm at p = 0.01 statistical significance. Regarding the estimates for precipitation minus evapotranspiration, we observe three different behaviors: TerraClimate has a change in P-E in the opposite direction of runoff (1 mm vs. -5 mm); mHM has a change in P-E of smaller magnitude than runoff (-2 mm vs. -9 mm); ERA5-Land has similar changes for both P-E and runoff (-55 mm vs. -56 mm), but with values one order of magnitude higher than those of TerraClimate and mHM.

The above results, seemingly disagreeing with the expected increases reported in previous literature (Kyselý and Beranová, 2009; Svoboda et al., 2016; Kašpárek and Kožín, 2022), indicate that there have not been any statistically significant changes in median annual precipitation over Czechia between the last two 30-year periods. Thereafter, we proceeded to look into changes between 1961-1990 and 1991-2020 monthly water fluxes (Figure 5.6). Note that hereinafter we mention only months with statistically significant changes (p ¡ 0.01). Regarding precipitation, we observe a consistent increase of around 14 mm during October and circa 11 mm during July present in TerraClimate, mHM(E-OBS), and ERA5-Land. Besides, mHM(E-OBS) and ERA5-Land had decreasing changes in April of -6 mm and -9 mm, respectively. We also found a -5 mm decrease during November, present only in mHM(E-OBS). In terms of evapotranspiration, as expected from the statistically significant changes described for annual values, we report increases between 1-10 mm depending on the month. TerraClimate has the shortest period of continuous changes with gradually increasing magnitude from January (1 mm) to March(9 mm). mHM on top of said evapotranspiration behavior

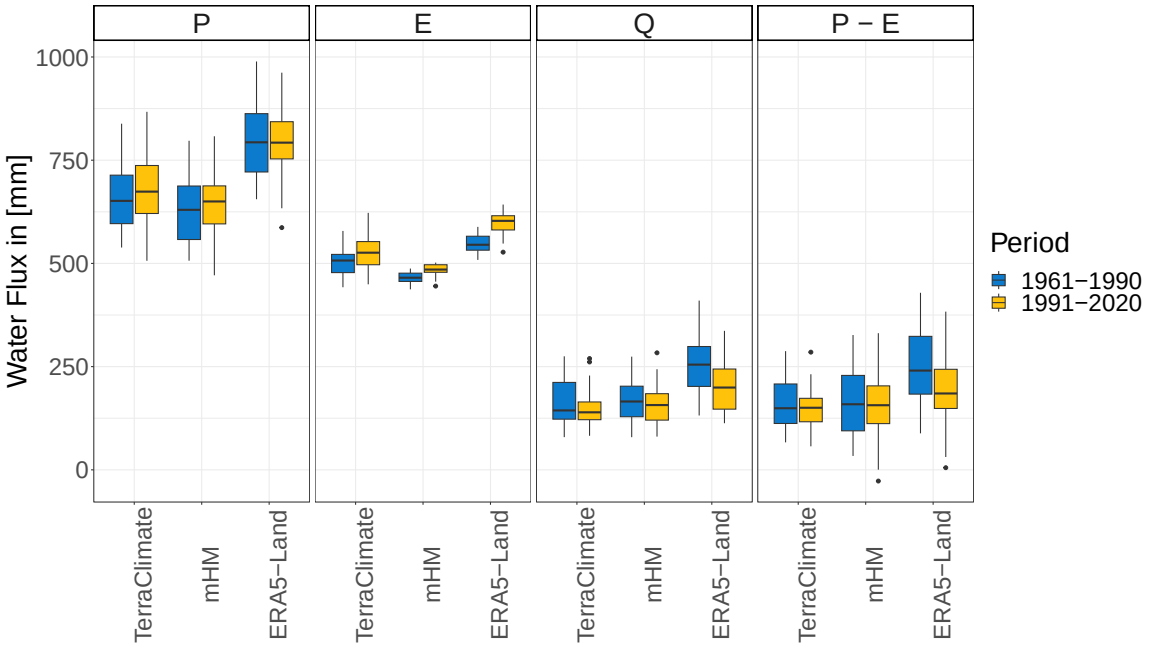


Figure 5.5: Box plots of spatial weighted average annual water fluxes over Czechia, where P is precipitation, E is evapotranspiration, Q is runoff, and $P - E$ is precipitation minus evapotranspiration. Data are divided into two 30-year periods: 1961-1990 (blue) and 1991-2020 (yellow). Note that outliers are present only in the latter period (i.e., 1991-2020) as expected from the recorded severe drought of 2003.

from January (1 mm) to April (4 mm) also shows the subsequent oscillating behavior: May (2 mm), June (2 mm), July(4 mm), and August (3 mm). ERA5-Land changes in evapotranspiration have a behavior similar to mHM but with overall higher magnitudes and two months longer. I.e., a consecutive increase from December (1 mm) to April (8 mm) and subsequent swings back and forth: May (7 mm), June (7 mm), July(10 mm), August (8 mm), and September (3 mm). Concerning runoff, there is a striking unique visual for TerraClimate, whose range of values from February to April is considerably larger than those of mHM or ERA5-Land. A runoff decrease is present in all data sets for April and May, with an added magnitude of -18 mm, -8 mm, and -12 mm for TerraClimate, mHM, and ERA5-Land, respectively. Interestingly, these runoff decreases are translated only into mHM and ERA5-Land through precipitation minus evapotranspiration decrease in April (-6 mm and -15 mm).

5.3.3 Spatial patterns of water cycle changes

The results shown so far provide insight into the temporal changes water cycle components have undergone in the past 60 years, considering spatial weighted averaged values across Czechia. To expand our analysis from the temporal into the spatial domain and provide insight into the spa-

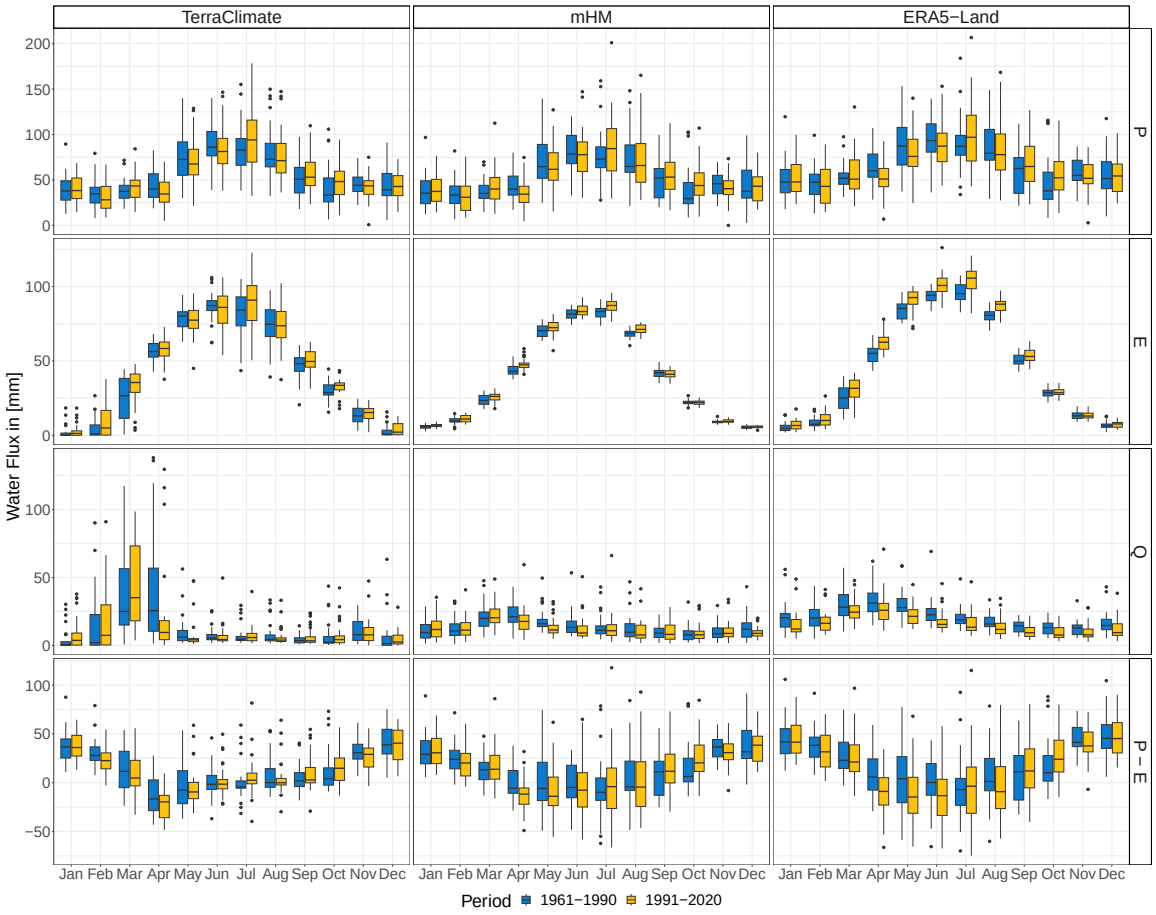


Figure 5.6: Box plot of spatial weighted average monthly water fluxes over Czechia, where P is precipitation, E is evapotranspiration, Q is runoff, and $P-E$ is precipitation minus evapotranspiration. Data are divided into two 30-year periods: 1961-1990 (blue) and 1991-2020 (yellow). Left column: TerraClimate (P), TerraClimate (E), and TerraClimate (Q). Middle column: mHM(E-OBS) (P), mHM (E), and mHM (Q). Right column: ERA5-Land (P), ERA5-Land (E), and ERA5-Land (Q).

tiotemporal features of the selected data sets, we mapped the difference between the 1991-2020 and the 1961-1990 medians for P, E, Q, and $P - E$ (Figure 5.7). Note that maps for each product were generated at their native resolutions, i.e., TerraClimate at 4 km, mHM at 0.125°, and ERA5-Land at 0.1°. At first glance, we observe overall agreement in spatial patterns between data sets for evapotranspiration and runoff, with slight discrepancies around the Sudetic (northeast), Šumava (southwest), and Ore (northwest) Mountains. In particular, ERA5-Land exhibits changes of higher magnitude in evapotranspiration (increase) and runoff (decrease) than TerraClimate and mHM.

Contrary to the above-described agreement, there is no consensus on spatial precipitation patterns among data sets. We discern three different patterns: TerraClimate shows a homogeneous increase across the country with a particular contour of higher increase that starts at the Šumava Mountains and diminishes toward the Ore Mountains and a slight decrease around the Sudetes; ERA5-Land portrays a somewhat zonal pattern with increasing bands north of 50.5°N and south of 49.5°N of the country and a decreasing band in the middle; mHM pattern is in between those of TerraClimate and ERA5-Land, with the band of precipitation decrease being smaller than that of ERA5-Land confined west of 15°E. While some of these heterogeneities are echoed in $P - E$ spatial patterns, there is a general decrease across data sets over Czechia. Therefore, evapotranspiration changes appear to dominate the spatial distribution of water availability.

Based on the results observed in Figure 5.6, we have previously identified that monthly patterns of increase or decrease in water fluxes are, to some extent, aligned with their seasonal variability. Thus this time around, we aggregated the data seasonally rather than looking at the monthly spatial distribution of changes in the median between the two 30-year periods. While individual characteristics for each data set are further emphasized by looking into seasonal spatial patterns, we identify some common traits. A dominant pattern of precipitation decrease is localized to the Westernmost part of Czechia during winter and expands to the rest of the country during spring. Evapotranspiration increases of the highest magnitude take place during spring and summer. As a result of this opposing direction, during spring, we see the most substantial decrease in runoff and $P - E$ therein. Furthermore, it is safe to state that if evapotranspiration generally increases despite decreasing patches of precipitation (present to a greater or lesser extent across all seasons), the water cycle in Czechia is dominated by changes in energy rather than water availability.

TerraClimate, with a resolution of 4 km, offers far more detail on spatial patterns than other data sets (Figure 5.8). It has a semester split for precipitation, with a decreasing pattern dominating winter and spring and an increasing pattern dominating summer and autumn. Evapotranspiration decreases during spring and summer but does not cover nearly as much area of Czechia as precip-

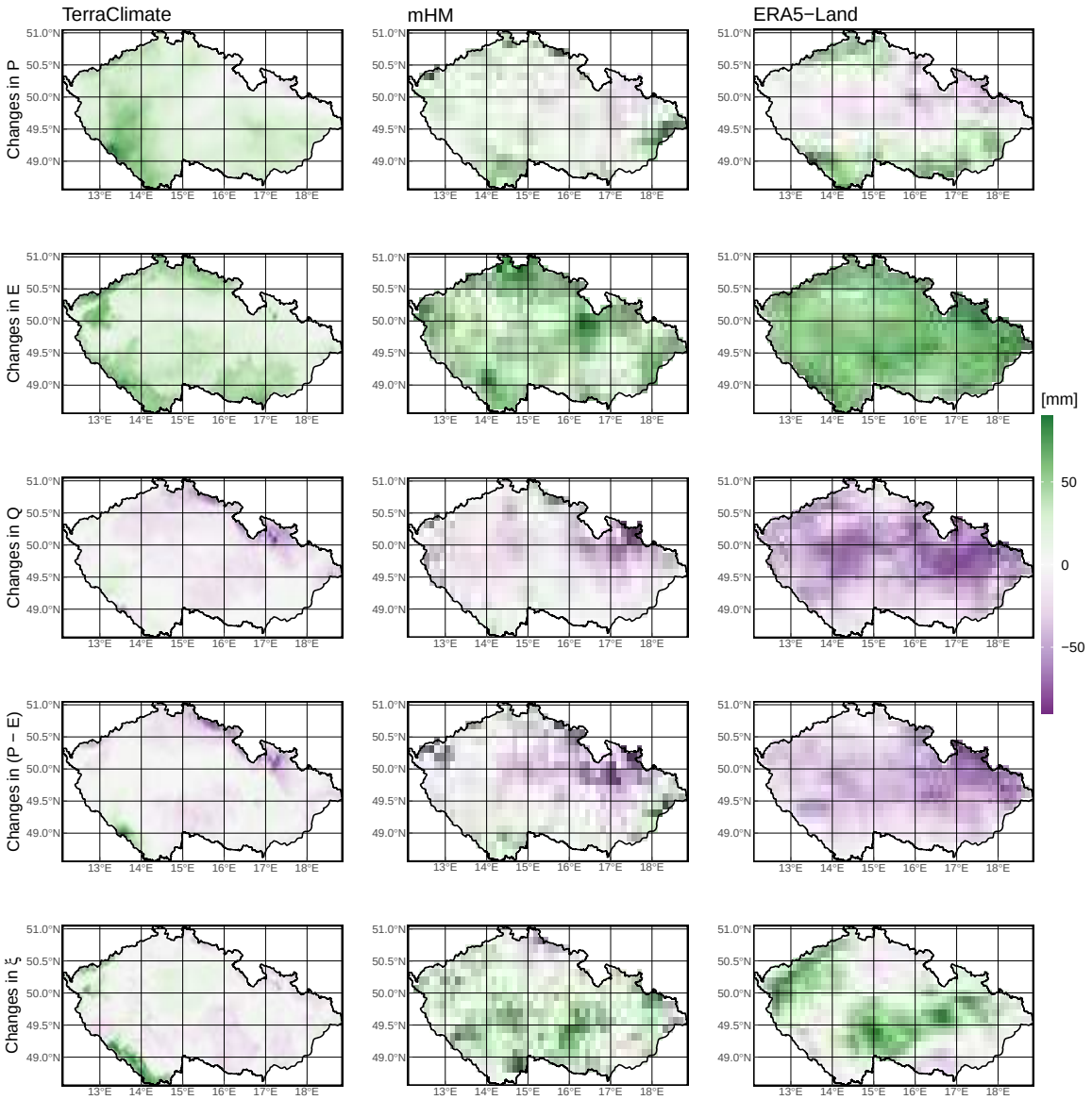


Figure 5.7: Spatial pattern of changes in median water fluxes over Czechia between two 30-year periods: 1961-1990 and 1991-2020. I.e., the value of each grid cell is equal to the median value of 1991-2020 minus the median value of 1961-1990. P is precipitation, E is evapotranspiration, Q is runoff, $P - E$ is precipitation minus evapotranspiration, and ξ is the residual ($P - E - Q$). Left column: TerraClimate (P), TerraClimate (E), and TerraClimate (Q). Middle column: mHM(E-OBS) (P), mHM (E), and mHM (Q). Right column: ERA5-Land (P), ERA5-Land (E), and ERA5-Land (Q).

itation when decreasing. Runoff changes circumscribe winter (increase) and spring (decrease) and are relatively mute during summer and autumn. Regarding water availability, the patterns of $P - E$ reflect those of precipitation. However, the increases in summer and autumn are not as notable. Autumn is a season of spatial homogeneity in TerraClimate because precipitation, evapotranspiration, runoff, and $P - E$ all depict countrywide increases, albeit of smaller magnitude than in other seasons. On the other hand, a distinctive contrast takes place in winter, in which we have a decrease in runoff in spite of an increase in water availability.

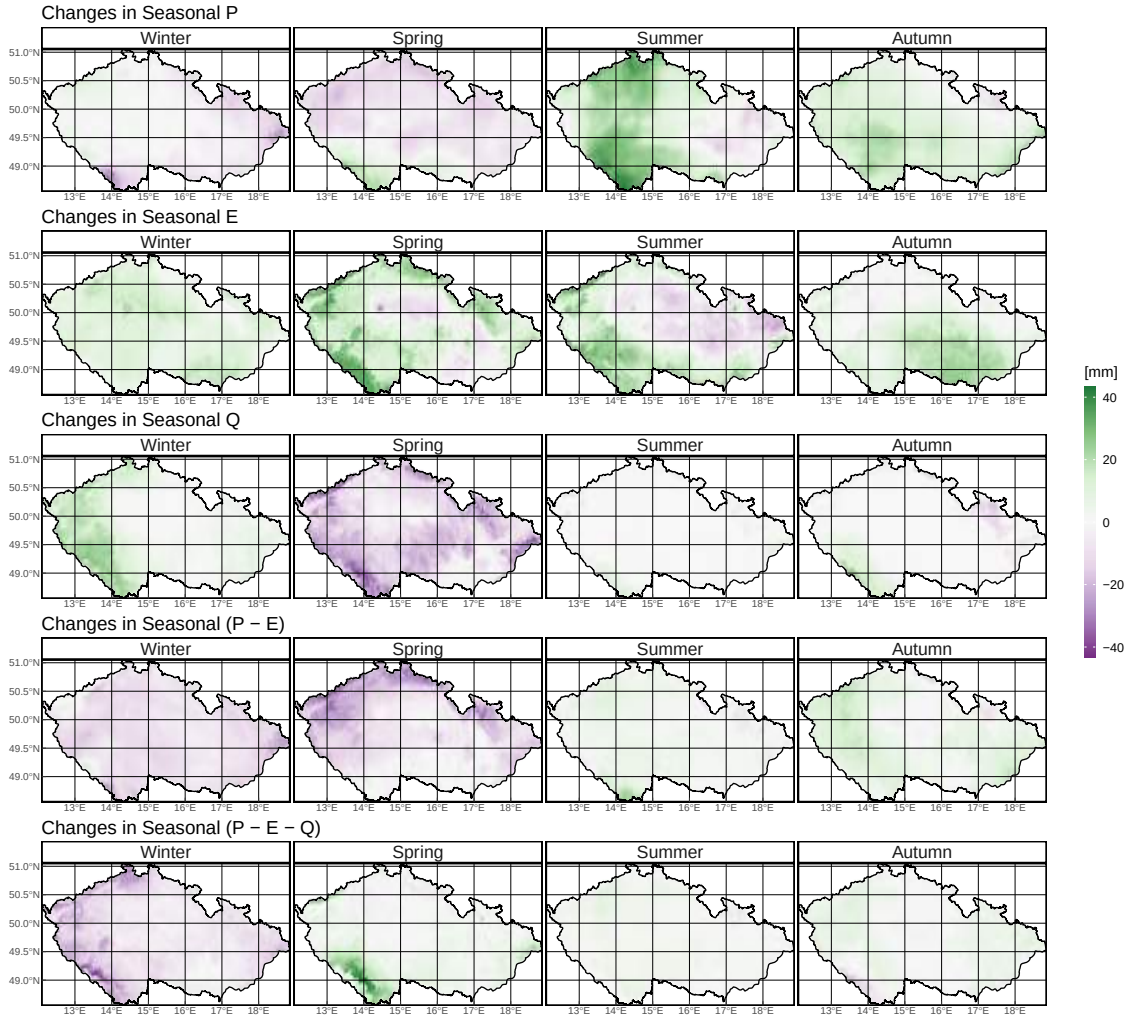


Figure 5.8: TerraClimate spatial pattern of changes in seasonal median water fluxes over Czechia between two 30-year periods: 1961-1990 and 1991-2020. I.e., the value of each grid cell is equal to the seasonal median value of 1991-2020 minus the seasonal median value of 1961-1990. P is precipitation, E is evapotranspiration, and Q is runoff. The seasons are defined as follows: winter as December, January, and February; spring as March, April, and May; summer as June, July, and August; autumn as September, October, and November.

Seasonal spatial patterns of mHM have the least substantial changes, with magnitudes mainly

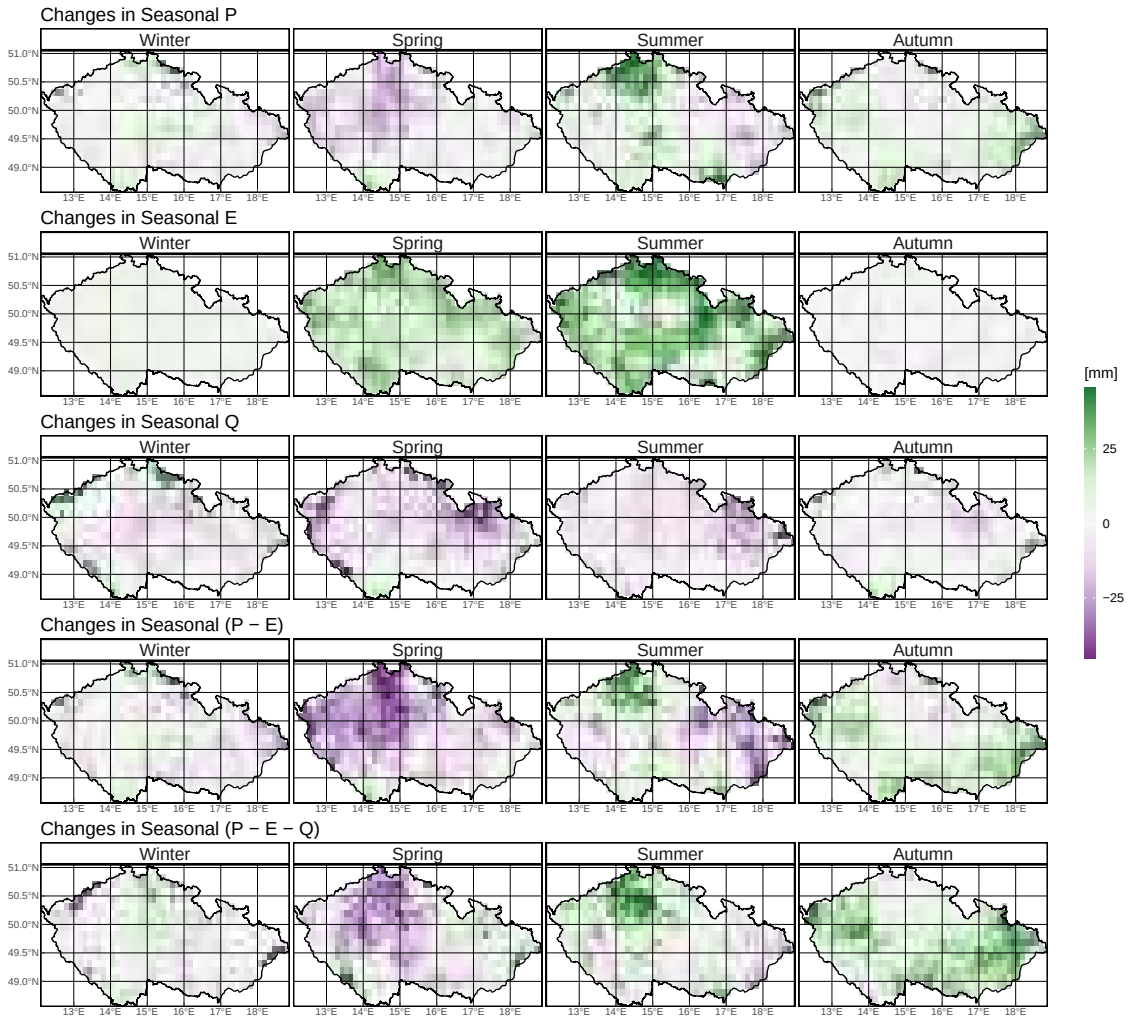


Figure 5.9: mHM spatial pattern of changes in seasonal median water fluxes over Czechia between two 30-year periods: 1961-1990 and 1991-2020. I.e., the value of each grid cell is equal to the seasonal median value of 1991-2020 minus the seasonal median value of 1961-1990. P is precipitation, E is evapotranspiration, and Q is runoff. The seasons are defined as follows: winter as December, January, and February; spring as March, April, and May; summer as June, July, and August; autumn as September, October, and November.

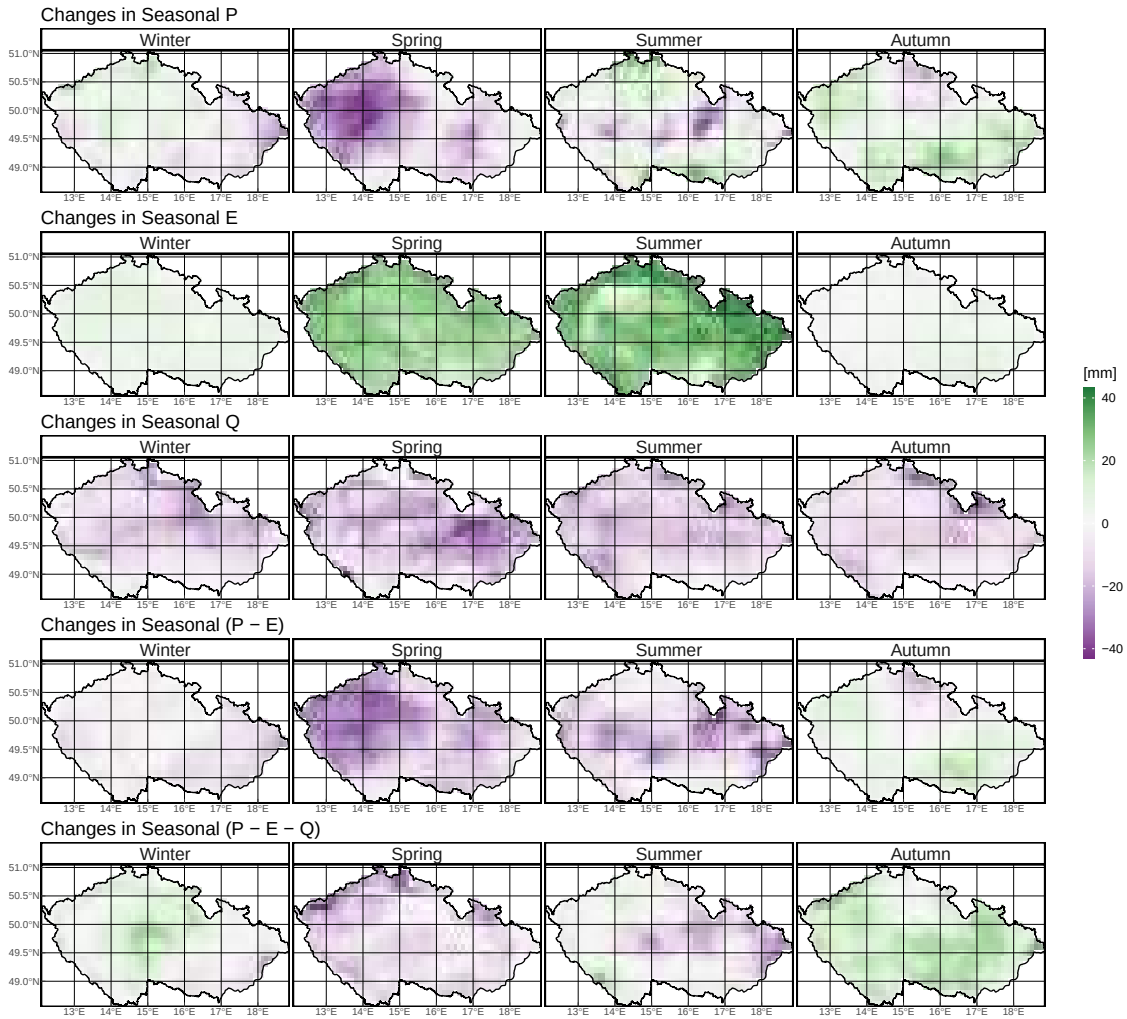


Figure 5.10: ERA5-Land spatial pattern of changes in seasonal median water fluxes over Czechia between two 30-year periods: 1961-1990 and 1991-2020. I.e., the value of each grid cell is equal to the seasonal median value of 1991-2020 minus the seasonal median value of 1961-1990. P is precipitation, E is evapotranspiration, and Q is runoff. The seasons are defined as follows: winter as December, January, and February; spring as March, April, and May; summer as June, July, and August; autumn as September, October, and November.

in the -25 mm to 25 mm range compared to the -40 mm to 40 mm range of TerraClimate and ERA5-Land (Figure 5.9). Precipitation patterns mimic those of TerraClimate except for autumn, where mHM(E-OBS) holds more heterogeneity. Contemporaneously, we observe slightly decreased evapotranspiration. For the rest of the seasons, evapotranspiration presents a widespread pattern of positive changes, with the highest magnitudes in summer. There is a dominant decreasing pattern for runoff across all seasons. In winter, there are pinpoint increases around the Czech borders near the Sudetic, Šumava, and Ore Mountains. $P - E$ has the highest magnitude for decreasing change in spring. There is a mixed pattern of increase and decrease for $P - E$ in winter and summer, yet the extent of decreasing changes is more prominent. Once again, analogous to TerraClimate, we find a season of contrasting runoff (decreasing) and $P - E$ (increasing) changes, but for mHM, it takes place in autumn.

ERA5-Land spatial pattern of changes in seasonal median water fluxes closely resembles those of mHM (Figure 5.10). The previously observed zonal pattern for precipitation change between the two 30-year medians seems to be driven by summer changes. Evapotranspiration changes, unlike TerraClimate or mHM, are increasing across all seasons. With specifically large evapotranspiration increases in summer followed by spring. In opposition, runoff has decreased regardless of the season. The sporadic patches of increased runoff observed in mHM near the Czech borders are nonexistent in ERA5-Land. Similarly, the mixed patterns for $P - E$ for mHM present in winter and summer are missing in ERA5-Land, which only reports decreasing changes. Lastly, we evince contrast in the direction of change between runoff (predominantly decreasing) and $P-E$ (predominantly increasing) in autumn, parallel to that of mHM. While this contrast is present in all data sets, the season differs for mHM and ERA5-Land (autumn) vs. TerraClimate (winter). Moreover, it is also inversed, i.e., TerraClimate has increasing runoff and decreasing $P-E$, but mHM and ERA5-Land have decreasing runoff and increasing $P - E$.

5.4 Discussion

Overall long-term changes in the annual water cycle in Czechia are primarily evident in evapotranspiration. Interestingly, the general agreement among different data sets at low-frequency time scales dissolves as we deepen into seasonal and monthly scales. Higher frequency temporal analysis revealed that while its seasonality modulates changes in precipitation, these changes are overwhelmed by a consistent evapotranspiration increase. This compound behavior results in depleted water availability, as reflected by decreasing runoff and $P - E$. Furthermore, different data combinations estimate

different spatiotemporal patterns of water cycle changes. The observed redistribution of water availability can seriously impact water resources in the region, including the quality and quantity of drinking water, the accessibility of water for irrigation and energy generation, and the health of aquatic ecosystems. Our results herein provide an updated overview of the water cycle in Czechia and map changes in the past 60 years, are essential to assess and ensure the sustainable use and management of water resources in Czechia. Additionally, we have defined and demonstrated the ability of a purely empirical ranking method to benchmark hydroclimatic data fusion and determine the best combination to represent water cycle budget closure that can be applied to any other regional study.

We determined that the best data sets for long-term assessment of water cycle individual components in Czechia based on the selected references are: mHM(E-OBS), ERA5-Land, and TerraClimate for precipitation; ERA5-Land, mHM, and TerraClimate for evapotranspiration; mHM, TerraClimate, and ERA5-Land for runoff. Similar standings for precipitation data were reported by Fallah et al. (2020) and Bandhauer et al. (2022). Fallah et al. (2020) used runoff simulation vs. streamflow observations using different data sets to benchmark precipitation data and found that E-OBS yields a robust agreement, while ERA5, Global Precipitation Climatology Centre (GPCC V.2018; Schneider et al., 2011), and Multi-Source Weighted-Ensemble Precipitation (MSWEP V2; Beck et al., 2019) show good performances. Bandhauer et al. (2022) report that while E-OBS and ERA5 agree qualitatively, ERA5 considerably overestimates mean precipitation over Europe due to too many wet days. These prevalent wet bias in ERA5 has been reported along diverse assessments (e.g., Beštáková et al., 2023; Lavers et al., 2022). NCEP/NCAR R1 had the worst precipitation performance. It was previously reported that, at least regarding extreme precipitation, NCEP/NCAR R1 performed far better than ERA5's predecessors, i.e., ERA40 (Uppala et al., 2005) and ERA-Interim (Dee et al., 2011), (Sun et al., 2018). This disagreement could be attributed to the improvements implemented in ERA5 over its predecessors in model parameterizations, spatial resolution, and input data assimilation. Additionally, the poor performance of NCEP/NCAR R1 might be rooted in its coarse spatial resolution (two grid cells cover Czechia).

Regarding evapotranspiration estimates, ERA5-Land has been reported as an adequate data source to overcome the unavailability of observed agrometeorological data in Europe (Vanella et al., 2022), and its robustness supports its use for drought monitoring (Vicente-Serrano et al., 2023). mHM has undergone extensive evaluation over Europe at multiple spatial scales and has repeatedly shown its ability to capture the observed dynamics of actual evapotranspiration (Hanel et al., 2018; Rakovec et al., 2016a) and its application to determine dominant drought types and their evolution

(Markonis et al., 2021a). While, to our knowledge, there have not been studies focusing on the quality or applications of TerraClimate evapotranspiration to date, it has been calibrated and validated using FLUXNET data (Abatzoglou et al., 2018), a conglomerate of networks gathering and standardizing quality control protocols for station-based evapotranspiration measurements (Pastorello et al., 2020). Most of the abovementioned referenced studies also testify to the quality of runoff data from mHM, TerraClimate, and ERA5-Land because the studies use runoff and streamflow data derived, among other variables, from their evapotranspiration estimates and show that they can capture the streamflow dynamics adequately across a wide range of climate and physiographical characteristics.

Our evaluation of individual water cycle components is cohesive with previous literature. Although the data products assessed herein have been previously analyzed at multiple spatial scales, this is done under a univariate perspective, that does not consider the ability of the data sets to reproduce the water cycle and its changes as a whole in a structurally plausible manner. This is easily denoted by the fact that even though mHM's performance was the best for all water cycle components evaluated using high-quality observational references, the best data set combination ranking is actually TerraClimate exclusive (i.e., all flux estimates from the same data set). Note that the score metric and ranking framework proposed herein serve as a method that can easily and quickly filter out the data set combinations providing implausible results. It should be remarked that this ranking framework acts as an initial assessment to be complemented with additional analyses because the score metric does not account for any biases in the products. Expressly because the aim of our work is not to benchmark the different data sets analyzed herein but to demonstrate how different can become the water cycles depicted by each of them.

It is clear that the story to be told in terms of water cycle changes is not only dependant on the data set of choice but also on the time scale. This kind of differences tend to be overlook when annual averages are being compared but when it comes to annual totals the small discrepancies add up. By further digging into this we unveiled some substantial inconsistency in the ERA5-Land data. It appears that the cumulative sum of the water budget residual in ERA5-Land declines monotonically in time, implying some systematic bias in the water budget closure. Notwithstanding, to our surprise, we found that throughout our analysis, mHM and ERA5-Land (a hydrological model and reanalysis) presented more compatible spatiotemporal patterns than the two hydrological models (mHM and TerraClimate). Regarding hydrological models, their evapotranspiration response is strongly linked to how they represent soil moisture and radiative energy at the surface (Boé and Terray, 2008; Zhao et al., 2013), leading to the visible discrepancies among mHM and TerraClimate. In terms of water

cycle fluxes' magnitude, we report significant ERA5-Land overestimation of precipitation and evapotranspiration, which are in line with previously reported overestimations of summer precipitation over Central Europe (Hassler and Lauer, 2021; Rivoire et al., 2022). These biases in conjunction with the monotonic declining trend we found in the cumulative sum of water budget residual raise further questions about the applicability of ERA5-Land in hydrological studies. Therefore extra caution should be exercised when the widely-used reanalysis data product is employed.

There is agreement among the best-ranked data set combinations that most of the significant changes in Czech water fluxes are localized in spring, particularly in April and May. Notwithstanding, we observe that it is the summer season whose changes determine the spatiotemporal patterns of change between the 1991-2020 and 1961-1990 medians. Declining precipitation and increasing evapotranspiration in spring support reported drying trends over Czechia (Brázdil et al., 2015). In addition to these general patterns, we identified localized increases in winter runoff coupled with decreases and shifts in spring runoff around the Sudetic, Šumava, and Ore Mountains. These changes in mountainous runoff have been previously identified and attributed to decreasing snow cover and earlier snowmelt season (Nedelcev and Jenicek, 2021), which in some Czech catchments also derive in summer low flows (Jenicek and Ledvinka, 2020). Similar seasonal developments of the snow effect on runoff have been reported over multiple mountainous catchments across the world (Berghuijs et al., 2014; Dierauer et al., 2018; Muelchi et al., 2021). Hänsel et al. (2019) remark that seasonal trends are sensitive to shifts in the season definition by one month, which aligns with our monthly analysis because we identified significant changes in months like May and November (peripheral months of spring and autumn as defined herein). Additionally, it could be the reason behind summer, the contiguous season, dominating the long-term precipitation pattern.

The drying regime we report in Czechia, due to the gradual increase in atmospheric evaporative demand over the last 60 years (1961-2020) extends in time and space over central and eastern Europe (Beštáková et al., 2023). Jaagus et al. (2022) reported long-term drying trends for the 1949-2018 period in Slovakia, Hungary, Romania, Moldova, southern Poland, and particularly significant in Czechia. Trnka et al. (2016) described a strong tendency towards increased dryness in most Central Europe. Brázdil et al. (2009) performed one of the longest-record analysis in the region (1881-2006) and exposed an increasing tendency towards more prolonged and more intensive dry episodes. Still, it remains unclear how this long-term shift is linked to the post-2000 seasonal (Potopová et al., 2015), annual (Hanel et al., 2018), and multi-year droughts (Moravec et al., 2021) that have occurred in Central Europe and Czechia in specific. It has been demonstrated, though, that these droughts manifest more as soil moisture deficits than meteorological and hydrological droughts, as they are

related to high evaporative demand during the warm season period (Markonis et al., 2021a). Our results agreement shows that the long-term aridification could be the outcome of the same physical mechanism, i.e., evaporation increase, to the one that dominates the short-term extreme events.

Our study comes with certain limitations that pave the way for future research. A certain limitation is that our analyses do not attribute the observed changes to any potential physical or anthropogenic drivers. It is likely that the evapotranspiration increase is linked to long-term changes in atmospheric circulation patterns that have caused a decline in cloudiness (Lhotka et al., 2020). As it has been shown that global warming is going to disrupt the terrestrial water cycle mainly due to changes in precipitation (Roderick et al., 2014), it is more plausible to attribute the observed intensification to the fluctuations of atmospheric circulation. Yet, this remains to be confirmed by future studies that will determine the factors that contribute most to the hydroclimatic shifts, although drought projections over Czechia (Dubrovsky et al., 2009), and central Europe Hari et al. (2020) indicate an increased drought risk in the future prevalent under different climate change scenarios. Additionally, our work does not investigate the role of water storage (snow and groundwater), as well as land cover or vegetation changes. Lastly, while country-level assessments are essential to improve water resources management and natural hazard policies, the water cycle budget is closed over hydrological units, not administrative boundaries.

5.5 Conclusions

Herein, we have proposed and demonstrated the applicability of a novel benchmarking method based on water cycle budget closure for hydroclimatic data fusion. The method does not enforce closure nor merge multiple data sets into a new one, but instead identifies the best combination of data sets in terms of water cycle budget residual distribution and correlation to referential data. Furthermore, the ranking method presented could easily be applied to any other region and use different referential data sets for evaluation. The ranking method may still be employed using gridded data like GPCC or CRU TS as an evaluation reference in data-scarce areas or when ground-station data is not publicly available. Most importantly, this metric is not constrained by data availability, as any of the variables in the equation evaluation terms can be omitted. This modularity makes it a flexible alternative to traditional approaches.

Using the best water budget data, we demonstrate that Czechia is undergoing water cycle acceleration, evinced by increased atmospheric water demand. Remarkably, the increase in precipitation is not as pronounced as that one in evapotranspiration. While changes in the 30-year median of spa-

tial weight average annual values show a minimum change in water availability, the spatial patterns reveal a prevalent decreasing pattern of runoff across the country. Besides, we identified significant spatial heterogeneity when assessing precipitation at a seasonal scale. Intriguingly, summer patterns are reflected in the spatial difference between the 1991-2020 and the 1961-1990 medians despite most of the significant changes in water cycle components being localized in spring. What is more, the precipitation rain/snow partition effect of less snow and earlier snowmelt around the mountains is reflected in a seasonal shift of runoff (increase in winter and subsequent decrease in spring). This might reflect how sub-seasonal shifts could affect the long-term hydrologic changes.

Based on our results and previous literature, it is safe to state that the depletion of water availability (runoff and $P-E$) over Czechia could prompt a surge in drought frequency. Considering that shifts in evapotranspiration overwhelm those of precipitation, the water cycle in Czechia is mainly driven by changes in energy rather than water availability. Further research is needed to better understand the complex drivers of this drying trend and to develop targeted interventions to address possible factors external to natural variability, like land-use changes and other anthropogenic factors. Although it remains unknown if this drying trend will persist, it should be considered in the planning of effective drought management strategies and water conservation measures to mitigate its adverse impacts for agriculture, energy production, and natural ecosystems in Czechia.

Code Availability

The data compiled herein and the R code for the figures are publicly available at `https://github.com/MiRoVaGo/ugc_cwc`.

Author Contributions

Mijael Rodrigo Vargas Godoy: Conceptualization, Formal analysis, Investigation, Writing - Original Draft. Yannis Markonis: Conceptualization, Supervision, Writing - Review & Editing. Oldrich Rakovec: Investigation, Writing - Review & Editing. Michal Jenicek: Investigation, Writing - Review & Editing. Riya Dutta: Writing - Review & Editing. Rajani Kumar Pradhan: Investigation. Zuzana Bešťáková: Investigation. Jan Kyselý: Writing - Review & Editing. Roman Juras: Writing - Review & Editing. Simon Michael Papalexiou: Writing - Review & Editing. Martin Hanel: Writing - Review & Editing.

Chapter 6

Summary & Conclusions

This dissertation aimed to advance the understanding of multi-source water cycle quantification. The research results improved the understanding and promoted the sensible integration of data from multiple sources. More specifically, this dissertation made contributions in four main aspects, namely:

(1) detailing the chronology of global water cycle quantification, highlighting improvements in consistency over recent years;

(2) enhancing the evaluation of the global water cycle, with a focus on the response to global warming;

(3) developing the *pRecipe* R package;

(4) demonstrating the practical application of a novel data fusion benchmarking to assess regional water cycle changes.

Rise of the data (Chapter 2): The in-depth compilation of the water cycle quantification chronicle highlights key milestones and challenges as we adapted our quantification methods to exploit new technological resources better in each epoch, including the early days of hydrology, model simulations period, satellite era, and age of big data. The current global precipitation and evaporation climatic normal is estimated at circa $500 \times 10^3[km^3/year]$. Overland precipitation is estimated at $110 \times 10^3[km^3/year]$, evaporation at $70 \times 10^3[km^3/year]$, and over the ocean at $390 \times 10^3[km^3/year]$ and $430 \times 10^3[km^3/year]$, respectively. However, despite unprecedented data availability, the quest for robust quantification of the global water cycle continues as uncertainty shrouds the above-reported estimates. Remote sensing satellite products struggle with measurements over

complex topography and cloud cover, climate model simulations are plagued with parametrizations without physical meaning, and observational records are short and heterogeneous. It is clear that no global water cycle data source is without fail, yet in some cases, one data source's strengths cover another's weaknesses. Accordingly, it could be argued that from the early 1900s to the twenty-first century, the paradigm of quantifying the global water cycle has shifted from identifying the best data source per water cycle component to developing the best way to integrate data from various sources. Multi-source quantification acknowledges the inherent unpredictability of the climate system because, analogous to ensemble theory, it formalizes the notion that even under similar local conditions, slight variations of the underlying microphysics may lead to observing a range of different outcomes. Notwithstanding, it is of utmost importance that the research community continuously strives to improve ground station observations, model simulations, reanalyses, and satellite remote sensing measurements individually. More accurate and robust individual data sources will subsequently refine the outcome of multi-source quantification, particularly when leveraging machine learning tools to handle large data sets.

The precipitation-evaporation space (Chapter 3): Climate reanalyses complement observational measurements and offer unprecedented spatiotemporal resolution worldwide. Furthermore, due to the lack of comprehensive observational-based data for crucial water cycle components such as evaporation, reanalysis data is still one of our best tools for researching changes in the global water cycle. Undeniably, these have improved the quantification of the global water cycle. However, their varying performances and uncertainties limit their applicability to the point that diagnostic variables such as precipitation and evaporation should be used with extreme caution. In this context, a framework encompassing precipitation, evaporation, their difference, and their sum was proposed and investigated to further constrain uncertainty by unveiling heretofore overlooked discrepancies in four reanalysis data sets (20CR v3, ERA-20C, ERA5, and NCEP/NCAR R1). It could be argued that introducing a new metric for water cycle acceleration into the current broad spectrum of metrics may lead to inconsistent hydroclimatology analysis. Nevertheless, precipitation plus evaporation is not just an index because it is physically grounded as evinced by good agreement with the range of hydrological sensitivity and, as such, is better suited to describe climate models and reanalyses. For example, a similar tendency of precipitation increases is evident in all reanalyses. In reality, charting the reanalyses in the precipitation–evaporation space, we observe that no reanalysis is alike as they all describe different water cycle dynamics. We displayed how, while precipitation minus evaporation, a critical diagnostic, is not directly observable at the global scale, precipitation plus evaporation is not held back by scale limitations and complements global water cycle research. The

compound behavior of precipitation and evaporation provides a more comprehensive picture of the water balance because it considers both the supply and demand of water or, within the precipitation–evaporation space, both atmospheric water storage (precipitation minus evaporation) and water cycle intensity (precipitation plus evaporation).

A global precipitation climatology toolbox and database (Chapter 4): The *pRecipe* package provides a common starting point for the hydrology scientific community through its homogenized database. By encompassing widely used products from multiple sources and establishing a common ground to start analysis, *pRecipe* guarantees a fully reproducible framework for precipitation research. Currently, the evergrowing *pRecipe* database includes: 20CR v3, CHIRPS v2.0, CMAP, CMORPH, CPC-Global, CRU TS v4.06, EM-Earth, ERA-20C, ERA5, ERA5-Land, FLDAS, GHCN v2, GLDAS CLSM v2.0, GLDAS NOAH v2.0, GLDAS VIC v2.0, GPCC v2020, GPCP v3.2, GPM IMERGM v07, GSMaP v08, JRA-55, MERRA-2, MSWEP v2.8, MSWX-Past, NCEP/DOE R2, NCEP/NCAR R1, PERSIANN-CDR, PREC/L, TerraClimate, TRMM 3B43 v7, and UDel v5.01. In addition to these 30 datasets, which have been homogenized at monthly time steps and a 0.25 spatial resolution, the database also includes several masks such as individual countries, IPCC AR reference regions, and major river basins, with a variety of geographical features such as land cover types, elevation zones, biomes, and Köppen-Geiger climate classifications. Furthermore, *pRecipe* offers a suite of analysis and visualization tools, which include spatial weighted averaging, spatial trend calculation, spatiotemporal subsetting or aggregation, irregular polygon cropping (using shapefiles), time series plots, empirical density histograms, boxplots, maps, and Taylor diagrams. Its versatile methods to work with spatial (raster) and tabular (data.table) data enable the user to export their results at any processing stage in NetCDF files compatible with climate and forecast metadata conventions or CSV files, facilitating a seamless transition for the user into posterior processing. Therefore, the *pRecipe* package constitutes a valuable resource for academics, government agencies, and private-sector professionals because it provides a consistent and transparent approach to precipitation research.

A Multi-Source Water Budget Perspective (Chapter 5): Using *pRecipe* and a novel method for benchmarking hydroclimatic data fusion based on water cycle budget closure, 96 different combinations of multi-source water cycle products (six for precipitation, four for evapotranspiration, and four for runoff) were ranked using local observational data as evaluation references over Czechia. Unfortunately, it is undeniable that spatially comprehensive and sufficiently long observational data records are scarce. In such data-limited areas, the ranking method may still be applied, omitting the corresponding term in the equation (e.g., we used only precipitation and runoff for evaluation). The

empirical metric proposed ranks multi-source data combinations to quickly identify the best combination of data sets to gain insights into regional water cycle changes. Reanalysis and hydrological modeling initially exhibited a general agreement at low-frequency time scales that dissolved as we deepened into seasonal and monthly scales. Unexpectedly, mHM and ERA5-Land (a hydrological model and reanalysis) presented more compatible spatiotemporal patterns than the two hydrological models (mHM and TerraClimate) despite significant ERA5-Land overestimation of precipitation and evapotranspiration. The results report that Czechia is undergoing water cycle changes mainly driven by changes in energy rather than water availability, with notable spatial and temporal variations evinced by increased atmospheric water fluxes. Higher frequency temporal analysis revealed that while its seasonality modulates changes in precipitation, these changes are overwhelmed by a consistent evapotranspiration increase, resulting in depleted water availability.

6.1 Novel contributions

In summary, this dissertation enhanced the understanding of multi-source quantification of the water cycle by proposing new methodologies and tools. The improved understanding and advancements in multi-source data approaches ensure better implementation and more accurate assessment of water cycle changes provided that we understand now results are susceptible to the data sets being used. The novel contributions of this dissertation include:

1. A comprehensive review of the chronological development, data sources, and methods used to quantify the global water cycle.

2. A framework that combines precipitation and evaporation to constrain uncertainties in data sets.

3. A tool that standardizes the download, exploration, processing, and visualization of a comprehensive database of global precipitation data sets.

4. A novel method for benchmarking hydroclimatic data fusion based on water cycle budget closure and observational references.

6.2 Recommendations for future research

Apart from the novel contributions summarized above, several future research lines were identified and recommended based upon the obtained results. These include:

1. The comprehensive review of the chronological development, data sources, and methods used to quantify the global water cycle unveiled the significant heterogeneity of ground-based observations despite a century of continuous technological advancements. This persistent limitation, juxtaposed with the high spatiotemporal variability in global water cycle quantification, calls for a closer examination of methodologies to enhance the temporal and spatial resolution of the available data. The existing limitations in resolution hinder a comprehensive understanding of the intricate dynamics of the water cycle across different regions and timescales. Exploring and implementing advanced technological solutions (e.g., more extensive networks or more sophisticated instruments) and innovative downscaling approaches (e.g., physically informed machine learning) is imperative to overcome this hurdle. By doing so, we can aspire to achieve a more accurate and detailed portrayal of the fluctuations and interactions within the global water cycle, thereby advancing our understanding of this complex system and its nexus implications.

2. Deploying a comprehensive "ground-based" network over oceans is an unrealistic expectation, especially considering the challenges in achieving a similar network over land. However, recognizing that atmospheric water fluxes over the oceans constitute a significant portion of the global water cycle fluxes, the imperative for enhanced precision in oceanic flux estimates becomes apparent. This limitation underscores a critical gap in our current understanding of the global water cycle, emphasizing the necessity to address uncertainties in oceanic precipitation and evaporation estimates. The call to action involves exploring advanced statistical methods and alternative approaches, such as stochastic or machine learning-based methods, to elevate the accuracy of these estimates. By doing so, we aim to overcome the challenges posed by oceanic regions' vast and dynamic nature, ultimately advancing our ability to comprehend and predict the intricate dynamics of the global water cycle, where oceans play a pivotal role.

3. The precipitation-evaporation space emerges as a valuable graphical framework that sheds light on previously unnoticed discrepancies in reanalyses. However, further research must dig into the causes of the unveiled underlying uncertainties. A key aspect to be addressed is the impact of different assimilation schemes, which can introduce variations and potential biases in quantifying water cycle diagnostic fluxes such as precipitation and evaporation. Understanding the intricacies of uncertainty propagation through assimilation schemes is vital for refining data accuracy. Reanalysis data is invaluable due to its high spatiotemporal resolution; thus, enhancing consistency across different reanalyses would propel a more robust understanding of the complexities inherent in the global water cycle.

4. A continuous optimization endeavor for the *pRecipe* package is imperative to ensure its ongoing relevance and efficiency. This optimization process should be dynamic, responsive to emerging technological advancements, and capable of further incorporating additional datasets to enrich its functionality. A logical progression involves the development of a parallel package tailored explicitly for evaporation data. By doing so, both principal water cycle fluxes(precipitation and evaporation) would be encapsulated, providing a comprehensive toolkit for researchers. This dual-package approach facilitates the seamless integration of evaporation data into the precipitation-evaporation space, offering a well-rounded resource for investigating water cycle changes. The concerted efforts toward refining and expanding these packages contribute to the advancement of scientific tools and the accessibility of high-quality data, thereby fostering more robust and fully reproducible research methodologies in the realm of water cycle dynamics.

5. Conducting a comprehensive uncertainty spatial analysis is imperative to pinpoint regions where discrepancies in water cycle flux estimates persist at notable levels. This analysis should delve into the intricacies of each region, considering factors such as local climate dynamics, topographical variations, and the availability of relevant data. By examining these elements, researchers can tailor targeted strategies to alleviate uncertainties in the identified regions. Recognizing that the water cycle operates uniquely in different geographical contexts, addressing uncertainty necessitates a region-specific approach. This approach involves acknowledging the influence of local climate dynamics and complex topography when evaluating the reliability and adequacy of the available data. Through this detailed and context-sensitive approach, researchers can contribute to characterizing the global water cycle more accurately, laying out a more reliable foundation for water resource management and climate change assessments.

6. Expanding the analysis of water cycle changes to encompass socioeconomic factors represents a pivotal step in comprehending the multifaceted dynamics at play. By exploring the intricate interplay between anthropogenic activities, urbanization, and land-use changes, researchers can gain insights into how human interventions influence temporal trends and spatial patterns within the water cycle, contributing to a more holistic perspective on climate change impacts. In tandem, further research should address how water cycle changes in response to global warming drive anthropogenic activities such as urban expansion, population displacement, and land-use change. This research has the potential to illuminate the broader implications for water resource management and food security. As human societies evolve, understanding the intricate connections between societal dynamics and the water cycle becomes crucial for

devising sustainable strategies that effectively address the challenges posed by climate change.

In conclusion, this dissertation has advanced the multi-source water cycle quantification field, making noteworthy contributions that enhance our comprehension and facilitate the integration of data from diverse sources. Exploring the chronology of global water cycle quantification has provided valuable insights into methodologies' historical evolution, emphasizing recent consistency improvements. Evaluating the global water cycle's response to global warming has enriched our understanding of climate change impacts on hydrological processes. The development of the *pRecipe* R package offers a standardized and versatile tool for handling precipitation data. At the same time, applying a novel data fusion benchmarking methodology has demonstrated its practical utility in assessing regional water cycle changes. Collectively, these contributions underscore the importance of a multidimensional approach to water cycle research, combining historical context, methodological innovations, and practical tools. This dissertation serves as a testament to the ongoing pursuit of knowledge in unraveling the complexities of the global water cycle, providing a foundation for informed decision-making in water resource management and climate change mitigation.

Bibliography

Abatzoglou JT, Dobrowski SZ, Parks SA, Hegewisch KC (2018) TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Scientific Data 5(1):170191, DOI 10.1038/sdata.2017.191, number: 1 Publisher: Nature Publishing Group

Abbott BW, Bishop K, Zarnetske JP, Minaudo C, Chapin FS, Krause S, Hannah DM, Conner L, Ellison D, Godsey SE, Plont S, Marçais J, Kolbe T, Huebner A, Frei RJ, Hampton T, Gu S, Buhman M, Sara Sayedi S, Ursache O, Chapin M, Henderson KD, Pinay G (2019) Human domination of the global water cycle absent from depictions and perceptions. Nature Geoscience 12(7):533–540, DOI 10.1038/s41561-019-0374-y, number: 7 Publisher: Nature Publishing Group

Accadia C, Mariani S, Casaioli M, Lavagnini A, Speranza A (2003) Sensitivity of Precipitation Forecast Skill Scores to Bilinear Interpolation and a Simple Nearest-Neighbor Average Method on High-Resolution Verification Grids. Weather and Forecasting 18(5):918–932, DOI 10.1175/1520-0434(2003)018⟨0918:SOPFSS⟩2.0.CO;2, publisher: American Meteorological Society Section: Weather and Forecasting

Adler RF, Huffman GJ, Chang A, Ferraro R, Xie PP, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P, Nelkin E (2003) The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). Journal of Hydrometeorology 4(6):1147–1167, DOI 10.1175/1525-7541(2003)004⟨1147:TVGPCP⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Adler RF, Sapiano MRP, Huffman GJ, Wang JJ, Gu G, Bolvin D, Chiu L, Schneider U, Becker A, Nelkin E, Xie P, Ferraro R, Shin DB (2018) The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation. Atmosphere 9(4):138, DOI 10.3390/atmos9040138, number: 4 Publisher: Multidisciplinary Digital Publishing Institute

Aires F (2014) Combining Datasets of Satellite-Retrieved Products. Part I: Methodology and Water Budget Closure. Journal of Hydrometeorology 15(4):1677–1691, DOI 10.1175/JHM-D-13-0148.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Aires F, Prigent C, Rossow WB (2004) Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 3. Network Jacobians. Journal of Geophysical Research: Atmospheres 109(D10), DOI 10.1029/2003JD004175, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2003JD004175

Albers S, Zipper S, Prosdocimi I (2022) CRAN Task View: Hydrological Data and Modeling. Publisher: Comprehensive R Archive Network (CRAN)

Albrecht F (1960) Jahreskarten des Wärme-und Wasserhaushaltes der Ozeane. Verlag nicht ermittelbar

Alcamo J (2019) Water quality and its interlinkages with the Sustainable Development Goals. Current Opinion in Environmental Sustainability 36:126–140, DOI 10.1016/j.cosust.2018.11.005

ALCAMO J, DÖLL P, HENRICHS T, KASPAR F, LEHNER B, RÖSCH T, SIEBERT S (2003) Development and testing of the WaterGAP 2 global model of water use and availability. Hydrological Sciences Journal 48(3):317–337, DOI 10.1623/hysj.48.3.317.45290, publisher: Taylor & Francis _eprint: https://doi.org/10.1623/hysj.48.3.317.45290

Allan RP (2012) Regime dependent changes in global precipitation. Climate Dynamics 39(3):827–840, DOI 10.1007/s00382-011-1134-x

Allan RP, Barlow M, Byrne MP, Cherchi A, Douville H, Fowler HJ, Gan TY, Pendergrass AG, Rosenfeld D, Swann ALS, Wilcox LJ, Zolina O (2020) Advances in understanding large-scale responses of the water cycle to climate change. Annals of the New York Academy of Sciences 1472(1):49–75, DOI 10.1111/nyas.14337, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/nyas.14337

Allen MR, Ingram WJ (2002) Constraints on future changes in climate and the hydrologic cycle. Nature 419(6903):224–232, DOI 10.1038/nature01092, number: 6903 Publisher: Nature Publishing Group

Arnell NW (1999) A simple water balance model for the simulation of streamflow over a large geographic domain. Journal of Hydrology 217(3):314–335, DOI 10.1016/S0022-1694(99)00023-2

Ashouri H, Hsu KL, Sorooshian S, Braithwaite DK, Knapp KR, Cecil LD, Nelson BR, Prat OP (2015) PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. Bulletin of the American Meteorological Society 96(1):69–83, DOI 10.1175/BAMS-D-13-00068.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Azarderakhsh M, Rossow WB, Papa F, Norouzi H, Khanbilvardi R (2011) Diagnosing water variations within the Amazon basin using satellite data. Journal of Geophysical Research: Atmospheres 116(D24), DOI 10.1029/2011JD015997, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2011JD015997

Backus JW, Stern H, Ziller I, Hughes RA, Nutt R, Beeber RJ, Best S, Goldberg R, Haibt LM, Herrick HL, Nelson RA, Sayre D, Sheridan PB (1957) The FORTRAN automatic coding system. In: Papers presented at the February 26-28, 1957, western joint computer conference: Techniques for reliability on - IRE-AIEE-ACM '57 (Western), ACM Press, Los Angeles, California, pp 188–198, DOI 10.1145/1455567.1455599

Bala G, Caldeira K, Nemani R (2010) Fast versus slow response in climate change: implications for the global hydrological cycle. Climate Dynamics 35(2):423–434, DOI 10.1007/s00382-009-0583-y

Balsamo G, Beljaars A, Scipal K, Viterbo P, Hurk Bvd, Hirschi M, Betts AK (2009) A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System. Journal of Hydrometeorology 10(3):623–643, DOI 10.1175/2008JHM1068.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Bandhauer M, Isotta F, Lakatos M, Lussana C, Båserud L, Izsák B, Szentes O, Tveito OE, Frei C (2022) Evaluation of daily precipitation analyses in E-OBS (v19.0e) and ERA5 by comparison to regional high-resolution datasets in European regions. International Journal of Climatology 42(2):727–747, DOI 10.1002/joc.7269, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.7269

Barnes JC, Bowley CJ (1968) Snow cover distribution as mapped from satellite photography. Water Resources Research 4(2):257–272, DOI 10.1029/WR004i002p00257, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/WR004i002p00257

Baumgartner A, Reichel E (1972) Preliminary results of new investigations of world's water balance. Applied optics 7:1705–1710

Beck HE, Pan M, Roy T, Weedon GP, Pappenberger F, van Dijk AIJM, Huffman GJ, Adler RF, Wood EF (2019) Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. Hydrology and Earth System Sciences 23(1):207–224, DOI 10.5194/hess-23-207-2019, publisher: Copernicus GmbH

Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U, Ziese M (2013) A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. Earth System Science Data 5(1):71–99, DOI 10.5194/essd-5-71-2013, publisher: Copernicus GmbH

Bengtsson L (2010) The global atmospheric water cycle. Environmental Research Letters 5(2):025202, DOI 10.1088/1748-9326/5/2/025202

Berghuijs WR, Woods RA, Hrachowitz M (2014) A precipitation shift from snow towards rain leads to a decrease in streamflow. Nature Climate Change 4(7):583–586, DOI 10.1038/nclimate2246, number: 7 Publisher: Nature Publishing Group

Betts AK, Zhao M, Dirmeyer PA, Beljaars ACM (2006) Comparison of ERA40 and NCEP/DOE near-surface data sets with other ISLSCP-II data sets. Journal of Geophysical Research: Atmospheres 111(D22), DOI 10.1029/2006JD007174, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2006JD007174

Bezner Kerr R, Hasegawa T, Lasco R, Bhatt I, Deryng D, Farrell A, Gurney-Smith H, Ju H, Lluch-Cota S, Meza F, Nelson G, Neufeldt H, Thornton P (2022) Food, Fibre and Other Ecosystem Products. In: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, pp 713–906, doi:10.1017/9781009325844.007

Beštáková Z, Strnad F, Vargas Godoy MR, Singh U, Markonis Y, Hanel M, Máca P, Kyselý J (2023) Changes of the aridity index in Europe from 1950 to 2019. Theoretical and Applied Climatology 151(1):587–601, DOI 10.1007/s00704-022-04266-3

Bhaduri A, Bogardi J, Siddiqi A, Voigt H, Vörösmarty C, Pahl-Wostl C, Bunn SE, Shrivastava P, Lawford R, Foster S, Kremer H, Renaud FG, Bruns A, Osuna VR (2016) Achieving Sustainable Development Goals from a Water Perspective. Frontiers in Environmental Science 4, DOI 10.3389/fenvs.2016.00064

Bishop CM (1996) Neural Networks for Pattern Recognition. Oxford University Press, Oxford, New York

Bonan GB, Oleson KW, Vertenstein M, Levis S, Zeng X, Dai Y, Dickinson RE, Yang ZL (2002) The Land Surface Climatology of the Community Land Model Coupled to the NCAR Community Climate Model. Journal of Climate 15(22):3123–3149, DOI 10.1175/1520-0442(2002)015⟨3123:TLSCOT⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Climate

Bondeau A, Smith PC, Zaehle S, Schaphoff S, Lucht W, Cramer W, Gerten D, Lotze-Campen H, Müller C, Reichstein M, Smith B (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Global Change Biology 13(3):679–706, DOI 10.1111/j.1365-2486.2006.01305.x, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2486.2006.01305.x

Borgomeo E, Hall JW, Salehin M (2018) Avoiding the water-poverty trap: insights from a conceptual human-water dynamical model for coastal Bangladesh. International Journal of Water Resources Development 34(6):900–922, DOI 10.1080/07900627.2017.1331842, publisher: Routledge _eprint: https://doi.org/10.1080/07900627.2017.1331842

Bosilovich MG, Chen J, Robertson FR, Adler RF (2008) Evaluation of Global Precipitation in Reanalyses. Journal of Applied Meteorology and Climatology 47(9):2279–2299, DOI 10.1175/2008JAMC1921.1, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology

Bosilovich MG, Robertson FR, Chen J (2011) Global Energy and Water Budgets in MERRA. Journal of Climate 24(22):5721–5739, DOI 10.1175/2011JCLI4175.1, publisher: American Meteorological Society Section: Journal of Climate

Boé J, Terray L (2008) Uncertainties in summer evapotranspiration changes over Europe and implications for regional climate change. Geophysical Research Letters 35(5), DOI 10.1029/2007GL032417, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2007GL032417

Bralower T, Bice D (2012) Module 4: Introduction to General Circulation Models. In: Earth 103:

Earth in the Future, College of Earth and Mineral Science, The Pennsylvania State University, URL http://creativecommons.org/licenses/by-nc-sa/4.0/

Brocca L, Filippucci P, Hahn S, Ciabatta L, Massari C, Camici S, Schüller L, Bojkov B, Wagner W (2019) SM2RAIN–ASCAT (2007–2018): global daily satellite rainfall data from ASCAT soil moisture observations. Earth System Science Data 11(4):1583–1601, DOI 10.5194/essd-11-1583-2019, publisher: Copernicus GmbH

Bromwich DH, Wilson AB, Bai L, Liu Z, Barlage M, Shih CF, Maldonado S, Hines KM, Wang SH, Woollen J, Kuo B, Lin HC, Wee TK, Serreze MC, Walsh JE (2018) The Arctic System Reanalysis, Version 2. Bulletin of the American Meteorological Society 99(4):805–828, DOI 10.1175/BAMS-D-16-0215.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Brutsaert W (2023) Hydrology: an introduction, second edition edn. Cambridge University Press, New York, NY

Brázdil R, Trnka M, Dobrovolný P, Chromá K, Hlavinka P, Žalud Z (2009) Variability of droughts in the Czech Republic, 1881–2006. Theoretical and Applied Climatology 97(3):297–315, DOI 10.1007/s00704-008-0065-x

Brázdil R, Dobrovolný P, Trnka M, Kotyza O, Řezníčková L, Valášek H, Zahradníček P, Štěpánek P (2013) Droughts in the Czech Lands, 1090–2012 AD. Climate of the Past 9(4):1985–2002, DOI 10.5194/cp-9-1985-2013, publisher: Copernicus GmbH

Brázdil R, Trnka M, Mikšovský J, Řezníčková L, Dobrovolný P (2015) Spring-summer droughts in the Czech Land in 1805–2012 and their forcings. International Journal of Climatology 35(7):1405–1421, DOI 10.1002/joc.4065, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.4065

Brückner E (1905) Die bilanz des kreislaufs des wassers auf der erde. Geographische Zeitschrift 11(8. H):436–445

Budyko MI (1955) Teplowoj Balans Zemnoi Poverkhuorti. Glavnaya geofizicheskaya observatoriya

Budyko MI (1961) The Heat Balance of the Earth's Surface. Soviet Geography 2(4):3–13, DOI 10.1080/00385417.1961.10770761, publisher: Routledge _eprint: https://doi.org/10.1080/00385417.1961.10770761

Budyko MI (1963) Atlas teplovogo balansa zemnogo shara. Glavnaya geofizicheskaya observatoriya

Budyko MI (1970) The water balance of the oceans. In: Symposium on World Water Balance, Gentbrügge, Int. Ass. Scient. Hydrol., vol 1, pp 24–33

Budyko MI (1974) Climate and life. No. 18 in International geophysics series, Academic Press, New York

Burges SJ, Wigmosta MS, Meena JM (1998) Hydrological Effects of Land-Use Change in a Zero-Order Catchment. Journal of Hydrologic Engineering 3(2):86–97, DOI 10.1061/(ASCE)1084-0699(1998)3:2(86), publisher: American Society of Civil Engineers

Byrne MP, O'Gorman PA (2015) The Response of Precipitation Minus Evapotranspiration to Climate Warming: Why the "Wet-Get-Wetter, Dry-Get-Drier" Scaling Does Not Hold over Land. Journal of Climate 28(20):8078–8092, DOI 10.1175/JCLI-D-15-0369.1, publisher: American Meteorological Society Section: Journal of Climate

Bytheway JL, Kummerow CD (2013) Inferring the uncertainty of satellite precipitation estimates in data-sparse regions over land. Journal of Geophysical Research: Atmospheres 118(17):9524–9533, DOI 10.1002/jgrd.50607, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jgrd.50607

Cardak O (2009) Science Students' Misconceptions of the Water Cycle According to their Drawings. Journal of Applied Sciences 9(5):865–873, DOI 10.3923/jas.2009.865.873

Carson D (1982) Current parameterizations of land surface processes in atmospheric general circulation models. Land surface processes in atmospheric general circulation models pp 67–108, publisher: Cambridge University Press New York

Cavalcanti IFA, Carril AF, Penalba OC, Grimm AM, Menéndez CG, Sanchez E, Cherchi A, Sörensson A, Robledo F, Rivera J, Pántano V, Bettolli LM, Zaninelli P, Zamboni L, Tedeschi RG, Dominguez M, Ruscica R, Flach R (2015) Precipitation extremes over La Plata Basin – Review and new results from observations and climate simulations. Journal of Hydrology 523:211–230, DOI 10.1016/j.jhydrol.2015.01.028

Chahine MT (1992a) GEWEX: The Global Energy and Water Cycle Experiment. Eos, Transactions American Geophysical Union 73(2):9–14, DOI 10.1029/91EO00007, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/91EO00007

Chahine MT (1992b) The hydrological cycle and its influence on climate. Nature 359(6394):373–380, DOI 10.1038/359373a0, number: 6394 Publisher: Nature Publishing Group

Chambers DP, Bonin JA (2012) Evaluation of Release-05 GRACE time-variable gravity coefficients over the ocean. Ocean Science 8(5):859–868, DOI 10.5194/os-8-859-2012, publisher: Copernicus GmbH

Chapin FS, Matson PA, Vitousek PM (2011) Principles of Terrestrial Ecosystem Ecology. Springer, New York, NY, DOI 10.1007/978-1-4419-9504-9

Chen F, Li X (2016) Evaluation of IMERG and TRMM 3B43 Monthly Precipitation Products over Mainland China. Remote Sensing 8(6):472, DOI 10.3390/rs8060472, number: 6 Publisher: Multidisciplinary Digital Publishing Institute

Chen M, Xie P, Janowiak JE, Arkin PA (2002) Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations. Journal of Hydrometeorology 3(3):249–266, DOI 10.1175/1525-7541(2002)003⟨0249:GLPAYM⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Cheng L, Trenberth KE, Gruber N, Abraham JP, Fasullo JT, Li G, Mann ME, Zhao X, Zhu J (2020) Improved Estimates of Changes in Upper Ocean Salinity and the Hydrological Cycle. Journal of Climate 33(23):10357–10381, DOI 10.1175/JCLI-D-20-0366.1, publisher: American Meteorological Society Section: Journal of Climate

Cherubim R (1931) Uber Verdunstungsmessung auf See. Ann d Hydrogr u Marit Meteor 59:325

Ciabatta L, Camici S, Massari C, Filippucci P, Hahn S, Wagner W, Brocca L (2020) Soil Moisture and Precipitation: The SM2RAIN Algorithm for Rainfall Retrieval from Satellite Soil Moisture. In: Levizzani V, Kidd C, Kirschbaum DB, Kummerow CD, Nakamura K, Turk FJ (eds) Satellite Precipitation Measurement: Volume 2, Advances in Global Change Research, Springer International Publishing, Cham, pp 1013–1027, DOI 10.1007/978-3-030-35798-6_27

Clapeyron BPE (1834) Memoir sur la puissance motrice de la chaleur. Journal de l'École royale polytechnique pp 153–190

Clark EA, Sheffield J, Vliet MTHv, Nijssen B, Lettenmaier DP (2015) Continental Runoff into the Oceans (1950–2008). Journal of Hydrometeorology 16(4):1502–1520, DOI 10.1175/JHM-D-14-0183.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Clemen RT (1986) Linear constraints and the efficiency of combined forecasts. Journal of Forecasting 5(1):31–38, DOI 10.1002/for.3980050104, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/for.3980050104

Collins M, Knutti R, Arblaster J, Dufresne JL, Fichefet T, Friedlingstein P, Gao X, Gutowski WJ, Johns T, Krinner G (2013) Long-term climate change: projections, commitments and irreversibility. In: Climate Change 2013-The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, pp 1029–1136

Compo G, Slivinski L, Whitaker J, Sardeshmukh P, Allan J, McColl C, Yin X, Vose R, Matsui N, Ashcroft L, Auchmann R, Benoy M, Bessemoulin P, Brandsma T, Brohan P, Brunet M, Comeaux J, Cram T, Crouthamel R, Groisman P, Hersbach H, Jones P, Jonsson T, Jourdain S, Kelly G, Knapp K, Kruger A, Kubota H, Lentini G, Lorrey A, Lott N, Lubker S, Luterbacher J, Marshall G, Maugeri M, Mock C, Mok H, Nordli O, Przybylak R, Rodwell M, Ross T, Schuster D, Srnec L, Valente M, Vizi Z, Wang X, Westcott N, Woollen J, Worley S (2015) The International Surface Pressure Databank version 3. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory DOI 10.5065/D6D50K29

Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose RS, Rutledge G, Bessemoulin P, Brönnimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY, Jones PD, Kruk MC, Kruger AC, Marshall GJ, Maugeri M, Mok HY, Nordli Ross TF, Trigo RM, Wang XL, Woodruff SD, Worley SJ (2011) The Twentieth Century Reanalysis Project. Quarterly Journal of the Royal Meteorological Society 137(654):1–28, DOI 10.1002/qj.776, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.776

Cornes RC, van der Schrier G, van den Besselaar EJM, Jones PD (2018) An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. Journal of Geophysical Research: Atmospheres 123(17):9391–9409, DOI 10.1029/2017JD028200, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2017JD028200

Cox PM, Betts RA, Bunton CB, Essery RLH, Rowntree PR, Smith J (1999) The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. Climate Dynamics 15(3):183–203, DOI 10.1007/s003820050276

Cronshey R (1986) Urban hydrology for small watersheds. Tech. rep., US Dept. of Agriculture, Soil Conservation Service, Engineering Division

Cruz-Alonso V, Pucher C, Ratcliffe S, Ruiz-Benito P, Astigarraga J, Neumann M, Hasenauer H, Rodríguez-Sánchez F (2023) The easyclimate R package: Easy access to high-resolution daily climate data for Europe. Environmental Modelling & Software p 105627, DOI 10.1016/j.envsoft.2023.105627

Dagan G, Stier P (2020) Constraint on precipitation response to climate change by combination of atmospheric energy and water budgets. npj Climate and Atmospheric Science 3(1):34, DOI 10.1038/s41612-020-00137-8

Dai A, Trenberth KE (2002) Estimates of Freshwater Discharge from Continents: Latitudinal and Seasonal Variations. Journal of Hydrometeorology 3(6):660–687, DOI 10.1175/1525-7541(2002)003⟨0660:EOFDFC⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Daley R (1999) Atmospheric data analysis, 1st edn. No. 2 in Cambridge atmospheric and space science series, Cambridge University Press, Cambridge

Dalton J (1799) Experiments and observations to determine whether the quantity of rain and dew is equal to the quantity of water carried off by the rivers and raised by evaporation: With an enquiry into the origin of springs. The Manchester Literary and Philosophical Society

Dastorani MT, Moghadamnia A, Piri J, Rico-Ramirez M (2010) Application of ANN and ANFIS models for reconstructing missing flow data. Environmental Monitoring and Assessment 166(1):421–434, DOI 10.1007/s10661-009-1012-8

DeCicco L, Hirsch R, Lorenz D, Read J, Walker J, Carr L, Watkins D, Blodgett D, Johnson M, Krall A (2022) dataRetrieval: Retrieval Functions for USGS and EPA Hydrology and Water Quality Data

Dee DP (2005) Bias and data assimilation. Quarterly Journal of the Royal Meteorological Society 131(613):3323–3343, DOI 10.1256/qj.05.137, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1256/qj.05.137

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F (2011) The

ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society 137(656):553–597, DOI 10.1002/qj.828, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.828

Devi U, Shekhar MS, Singh GP, Rao NN, Bhatt US (2019) Methodological application of quantile mapping to generate precipitation data over Northwest Himalaya. International Journal of Climatology 39(7):3160–3170, DOI 10.1002/joc.6008, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.6008

Di Baldassarre G, Kooy M, Kemerink JS, Brandimarte L (2013) Towards understanding the dynamic behaviour of floodplains as human-water systems. Hydrology and Earth System Sciences 17(8):3235–3244, DOI 10.5194/hess-17-3235-2013, publisher: Copernicus GmbH

Dickinson RE (1984) Modeling Evapotranspiration for Three-Dimensional Global Climate Models. In: Climate Processes and Climate Sensitivity, American Geophysical Union (AGU), pp 58–72, DOI 10.1029/GM029p0058, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/GM029p0058

Dierauer JR, Whitfield PH, Allen DM (2018) Climate Controls on Runoff and Low Flows in Mountain Catchments of Western North America. Water Resources Research 54(10):7495–7510, DOI 10.1029/2018WR023087, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018WR023087

van Dijk AIJM, Renzullo LJ (2011) Water resource monitoring systems and the role of satellite observations. Hydrology and Earth System Sciences 15(1):39–55, DOI 10.5194/hess-15-39-2011, publisher: Copernicus GmbH

Dirmeyer PA, Gao X, Zhao M, Guo Z, Oki T, Hanasaki N (2006) GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface. Bulletin of the American Meteorological Society 87(10):1381–1398, DOI 10.1175/BAMS-87-10-1381, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Dubach LL, Ng C (1988) Compendium of meteorological space programs, satellites, and experiments

Dubrovsky M, Svoboda MD, Trnka M, Hayes MJ, Wilhite DA, Zalud Z, Hlavinka P (2009) Application of relative drought indices in assessing climate-change impacts on drought conditions in Czechia. Theoretical and Applied Climatology 96(1):155–171, DOI 10.1007/s00704-008-0020-x

Durack P (2015) Ocean Salinity and the Global Water Cycle. Oceanography 28(1):20–31, DOI 10.5670/oceanog.2015.03

Durack PJ, Wijffels SE, Matear RJ (2012) Ocean Salinities Reveal Strong Global Water Cycle Intensification During 1950 to 2000. Science 336(6080):455–458, DOI 10.1126/science.1212222, publisher: American Association for the Advancement of Science

Durand M, Fu LL, Lettenmaier DP, Alsdorf DE, Rodriguez E, Esteban-Fernandez D (2010) The Surface Water and Ocean Topography Mission: Observing Terrestrial Surface Water and Oceanic Submesoscale Eddies. Proceedings of the IEEE 98(5):766–779, DOI 10.1109/JPROC.2010.2043031, conference Name: Proceedings of the IEEE

Döll P, Trautmann T, Gerten D, Schmied HM, Ostberg S, Saaed F, Schleussner CF (2018) Risks for the global freshwater system at 1.5 °C and 2 °C global warming. Environmental Research Letters 13(4):044038, DOI 10.1088/1748-9326/aab792, publisher: IOP Publishing

Eicker A, Jensen L, Wöhnke V, Dobslaw H, Kvas A, Mayer-Gürr T, Dill R (2020) Daily GRACE satellite data evaluate short-term hydro-meteorological fluxes from global atmospheric reanalyses. Scientific Reports 10(1):4504, DOI 10.1038/s41598-020-61166-0, number: 1 Publisher: Nature Publishing Group

Eischeid JK, Baker CB, Karl TR, Diaz HF (1995) The Quality Control of Long-Term Climatological Data Using Objective Data Analysis. Journal of Applied Meteorology and Climatology 34(12):2787–2795, DOI 10.1175/1520-0450(1995)034⟨2787:TQCOLT⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology

Eischeid JK, Pasteris PA, Diaz HF, Plantico MS, Lott NJ (2000) Creating a Serially Complete, National Daily Time Series of Temperature and Precipitation for the Western United States. Journal of Applied Meteorology and Climatology 39(9):1580–1591, DOI 10.1175/1520-0450(2000)039⟨1580:CASCND⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology

Ellis EC, Klein Goldewijk K, Siebert S, Lightman D, Ramankutty N (2010) Anthropogenic transformation of the biomes, 1700 to 2000. Global Ecology and Biogeography 19(5):589–606, DOI 10.1111/j.1466-8238.2010.00540.x, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1466-8238.2010.00540.x

van der Ent RJ, Savenije HHG, Schaefli B, Steele-Dunne SC (2010) Origin and fate of atmospheric moisture over continents. Water Resources Research 46(9), DOI 10.1029/2010WR009127, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2010WR009127

Espinoza V, Waliser DE, Guan B, Lavers DA, Ralph FM (2018) Global Analysis of Climate Change Projection Effects on Atmospheric Rivers. Geophysical Research Letters 45(9):4299–4308, DOI 10.1029/2017GL076968, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2017GL076968

Essery RLH, Best MJ, Betts RA, Cox PM, Taylor CM (2003) Explicit Representation of Subgrid Heterogeneity in a GCM Land Surface Scheme. Journal of Hydrometeorology 4(3):530–543, DOI 10.1175/1525-7541(2003)004⟨0530:EROSHI⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Evans JP, McCabe MF (2010) Regional climate simulation over Australia's Murray-Darling basin: A multitemporal assessment. Journal of Geophysical Research: Atmospheres 115(D14), DOI 10.1029/2010JD013816, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2010JD013816

Falkenmark M, Lindh G (1974) How Can We Cope with the Water Resources Situation by the Year 2015? Ambio 3(3/4):114–122, publisher: [Springer, Royal Swedish Academy of Sciences]

Falkenmark M, Wang-Erlandsson L, Rockström J (2019) Understanding of water resilience in the Anthropocene. Journal of Hydrology X 2:100009, DOI 10.1016/j.hydroa.2018.100009

Fallah A, O S, Orth R (2020) Climate-dependent propagation of precipitation uncertainty into the water cycle. Hydrology and Earth System Sciences 24(7):3725–3735, DOI 10.5194/hess-24-3725-2020, publisher: Copernicus GmbH

Farmer WH, Vogel RM (2016) On the deterministic and stochastic use of hydrologic models. Water Resources Research 52(7):5619–5633, DOI 10.1002/2016WR019129, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2016WR019129

Federer CA, Vörösmarty C, Fekete B (1996) Intercomparison of Methods for Calculating Potential Evaporation in Regional and Global Water Balance Models. Water Resources Research 32(7):2315–2321, DOI 10.1029/96WR00801, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/96WR00801

Fekete BM, Vörösmarty CJ, Grabs W (2002) High-resolution fields of global runoff combining observed river discharge and simulated water balances. Global Biogeochemical Cycles 16(3):15–1–15–10, DOI 10.1029/1999GB001254, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/1999GB001254

Fekete BM, Vörösmarty CJ, Roads JO, Willmott CJ (2004) Uncertainties in Precipitation and Their Impacts on Runoff Estimates. Journal of Climate 17(2):294–304, DOI 10.1175/1520-0442(2004)

017⟨0294:UIPATI⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Climate

Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37(12):4302–4315, DOI 10.1002/joc.5086, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.5086

Flato G, Marotzke J, Abiodun B, Braconnot P, Chou SC, Collins W, Cox P, Driouech F, Emori S, Eyring V, Forest C, Gleckler P, Guilyardi E, Jakob C, Kattsov V, Reason C, Rummukainen M (2013) Evaluation of climate models. In: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, pp 741–866

Flato GM (2011) Earth system models: an overview. WIREs Climate Change 2(6):783–800, DOI 10.1002/wcc.148, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.148

Fläschner D, Mauritsen T, Stevens B (2016) Understanding the Intermodel Spread in Global-Mean Hydrological Sensitivity. Journal of Climate 29(2):801–817, DOI 10.1175/JCLI-D-15-0351.1, publisher: American Meteorological Society Section: Journal of Climate

Friedl MA, McIver DK, Hodges JCF, Zhang XY, Muchoney D, Strahler AH, Woodcock CE, Gopal S, Schneider A, Cooper A, Baccini A, Gao F, Schaaf C (2002) Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment 83(1):287–302, DOI 10.1016/S0034-4257(02)00078-0

Fritzsche R (1906) Niederschlag, Abfluss und Verdunstung auf den Landflächen der Erde. as (Dresden Druck von W. Baensch)

Fuchs T, Rapp J, Rubel F, Rudolf B (2001) Correction of synoptic precipitation observations due to systematic measuring errors with special regard to precipitation phases. Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere 26(9):689–693, DOI 10.1016/S1464-1909(01)00070-3

Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A, Michaelsen J (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific Data 2(1):150066, DOI 10.1038/sdata.2015.66, number: 1 Publisher: Nature Publishing Group

Funk CC, Peterson PJ, Landsfeld MF, Pedreros DH, Verdin JP, Rowland JD, Romero BE, Husak GJ, Michaelsen JC, Verdin AP (2014) A quasi-global precipitation time series for drought monitoring. Tech. Rep. 832, U.S. Geological Survey, DOI 10.3133/ds832, iSSN: 2327-638X Publication Title: Data Series

Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, Randles CA, Darmenov A, Bosilovich MG, Reichle R, Wargan K, Coy L, Cullather R, Draper C, Akella S, Buchard V, Conaty A, Silva AMd, Gu W, Kim GK, Koster R, Lucchesi R, Merkova D, Nielsen JE, Partyka G, Pawson S, Putman W, Rienecker M, Schubert SD, Sienkiewicz M, Zhao B (2017) The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Journal of Climate 30(14):5419–5454, DOI 10.1175/JCLI-D-16-0758.1, publisher: American Meteorological Society Section: Journal of Climate

Gimeno L, Sorí R, Vázquez M, Stojanovic M, Algarra I, Eiras-Barca J, Gimeno-Sotelo L, Nieto R (2022) Extreme precipitation events. WIREs Water 9(6):e1611, DOI 10.1002/wat2.1611, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wat2.1611

Gires A, Tchiguirinskaia I, Schertzer D, Lovejoy S (2013) Development and analysis of a simple model to represent the zero rainfall in a universal multifractal framework. Nonlinear Processes in Geophysics 20(3):343–356, DOI 10.5194/npg-20-343-2013, URL `https://npg.copernicus.org/articles/20/343/2013/`, publisher: Copernicus GmbH

Gleick PH (1993) Water in crisis: a guide to the world's fresh water resources. Oxford Univ. Press, New York, NY

Gokceoglu C, Pourghasemi HR (eds) (2019) Spatial modeling in GIS and R for earth and environmental sciences. Elsevier, Amsterdam, Netherlands Oxford, United Kingdom Cambridge, MA, United States

Gosling SN, Arnell NW (2011) Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis. Hydrological Processes 25(7):1129–1145, DOI 10.1002/hyp.7727, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hyp.7727

Greve P, Orlowsky B, Mueller B, Sheffield J, Reichstein M, Seneviratne SI (2014) Global assessment of trends in wetting and drying over land. Nature Geoscience 7(10):716–721, DOI 10.1038/ngeo2247, number: 10 Publisher: Nature Publishing Group

Gupta VK, Waymire EC (1993) A Statistical Analysis of Mesoscale Rainfall as a Random Cascade. Journal of Applied Meteorology and Climatology 32(2):251–267, DOI 10.1175/1520-0450(1993)032⟨0251:ASAOMR⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology

Gutenstein M, Fennig K, Schröder M, Trent T, Bakan S, Roberts JB, Robertson FR (2021) Intercomparison of freshwater fluxes over ocean and investigations into water budget closure. Hydrology and Earth System Sciences 25(1):121–146, DOI 10.5194/hess-25-121-2021, publisher: Copernicus GmbH

Haddeland I, Clark DB, Franssen W, Ludwig F, Voß F, Arnell NW, Bertrand N, Best M, Folwell S, Gerten D, Gomes S, Gosling SN, Hagemann S, Hanasaki N, Harding R, Heinke J, Kabat P, Koirala S, Oki T, Polcher J, Stacke T, Viterbo P, Weedon GP, Yeh P (2011) Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results. Journal of Hydrometeorology 12(5):869–884, DOI 10.1175/2011JHM1324.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Haddeland I, Heinke J, Biemans H, Eisner S, Flörke M, Hanasaki N, Konzmann M, Ludwig F, Masaki Y, Schewe J, Stacke T, Tessler ZD, Wada Y, Wisser D (2014) Global water resources affected by human interventions and climate change. Proceedings of the National Academy of Sciences 111(9):3251–3256, DOI 10.1073/pnas.1222475110, publisher: Proceedings of the National Academy of Sciences

Hagemann S, Dümenil L (1997) A parametrization of the lateral waterflow for the global scale. Climate Dynamics 14(1):17–31, DOI 10.1007/s003820050205

Hagemann S, Gates LD (2003) Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations. Climate Dynamics 21(3):349–359, DOI 10.1007/s00382-003-0349-x

Halbfaß W (1934) Review of Beitrag zur Methode der kartographischen Darstellungvon Wasserkräften. (Veröff. der Schles. Ges. f. Erdkunde und des Geogr. Instituts der Univ. Breslau, Heft 20.). Geographische Zeitschrift 40(10):391–391, publisher: Franz Steiner Verlag

Hanasaki N, Kanae S, Oki T, Masuda K, Motoya K, Shirakawa N, Shen Y, Tanaka K (2008) An integrated model for the assessment of global water resources – Part 1: Model description and

input meteorological forcing. Hydrology and Earth System Sciences 12(4):1007–1025, DOI 10.5194/hess-12-1007-2008, publisher: Copernicus GmbH

Hanel M, Kožín R, Heřmanovský M, Roub R (2017) An R package for assessment of statistical downscaling methods for hydrological climate change impact studies. Environmental Modelling & Software 95:22–28, DOI 10.1016/j.envsoft.2017.03.036

Hanel M, Rakovec O, Markonis Y, Máca P, Samaniego L, Kyselý J, Kumar R (2018) Revisiting the recent European droughts from a long-term perspective. Scientific Reports 8(1):9499, DOI 10.1038/s41598-018-27464-4, number: 1 Publisher: Nature Publishing Group

Harding R, Best M, Blyth E, Hagemann S, Kabat P, Tallaksen LM, Warnaars T, Wiberg D, Weedon GP, Lanen Hv, Ludwig F, Haddeland I (2011) WATCH: Current Knowledge of the Terrestrial Global Water Cycle. Journal of Hydrometeorology 12(6):1149–1156, DOI 10.1175/JHM-D-11-024.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Hargreaves GH, Samani ZA (1982) Estimating Potential Evapotranspiration. Journal of the Irrigation and Drainage Division 108(3):225–230, DOI 10.1061/JRCEA4.0001390, publisher: American Society of Civil Engineers

Hari V, Rakovec O, Markonis Y, Hanel M, Kumar R (2020) Increased future occurrences of the exceptional 2018–2019 Central European drought under global warming. Scientific Reports 10(1):12207, DOI 10.1038/s41598-020-68872-9, number: 1 Publisher: Nature Publishing Group

Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Scientific data 7(1):1–18, publisher: Nature Publishing Group

Hassler B, Lauer A (2021) Comparison of Reanalysis and Observational Precipitation Datasets Including ERA5 and WFDE5. Atmosphere 12(11):1462, DOI 10.3390/atmos12111462, number: 11 Publisher: Multidisciplinary Digital Publishing Institute

Hawkins E, Smith RS, Gregory JM, Stainforth DA (2016) Irreducible uncertainty in near-term climate projections. Climate Dynamics 46(11):3807–3819, DOI 10.1007/s00382-015-2806-8

Hegerl GC, Black E, Allan RP, Ingram WJ, Polson D, Trenberth KE, Chadwick RS, Arkin PA, Sarojini BB, Becker A, Dai A, Durack PJ, Easterling D, Fowler HJ, Kendon EJ, Huffman GJ, Liu C, Marsh R, New M, Osborn TJ, Skliris N, Stott PA, Vidale PL, Wijffels SE, Wilcox LJ, Willett KM, Zhang X (2015) Challenges in Quantifying Changes in the Global Water Cycle. Bulletin of the

American Meteorological Society 96(7):1097–1115, DOI 10.1175/BAMS-D-13-00212.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Held IM, Soden BJ (2006) Robust Responses of the Hydrological Cycle to Global Warming. Journal of Climate 19(21):5686–5699, DOI 10.1175/JCLI3990.1, publisher: American Meteorological Society Section: Journal of Climate

Hellström C, Chen D, Achberger C, Räisänen J (2001) Comparison of climate change scenarios for Sweden based on statistical and dynamical downscaling of monthly precipitation. Climate Research 19(1):45–55, DOI 10.3354/cr019045

Herrera S, Kotlarski S, Soares PMM, Cardoso RM, Jaczewski A, Gutiérrez JM, Maraun D (2019) Uncertainty in gridded precipitation products: Influence of station density, interpolation method and grid resolution. International Journal of Climatology 39(9):3717–3729, DOI 10.1002/joc.5878, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.5878

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut JN (2020) The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146(730):1999–2049, DOI 10.1002/qj.3803, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25(15):1965–1978, DOI 10.1002/joc.1276, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.1276

Hijmans RJ, Etten Jv, Sumner M, Cheng J, Baston D, Bevan A, Bivand R, Busetto L, Canty M, Fasoli B, Forrest D, Ghosh A, Golicher D, Gray J, Greenberg JA, Hiemstra P, Hingee K, Ilich A, Geosciences IfMA, Karney C, Mattiuzzi M, Mosher S, Naimi B, Nowosad J, Pebesma E, Lamigueiro OP, Racine EB, Rowlingson B, Shortridge A, Venables B, Wueest R (2022) raster: Geographic Data Analysis and Modeling. URL `https://CRAN.R-project.org/package=raster`

Hildebrand PH, Houser P, Schlosser CA (2003) Observing the Global Water Cycle from Space. In: 31st International Conference on Radar Meteorology, Citeseer

Ho L, Alonso A, Eurie Forio MA, Vanclooster M, Goethals PLM (2020a) Water research in support of the Sustainable Development Goal 6: A case study in Belgium. Journal of Cleaner Production 277:124082, DOI 10.1016/j.jclepro.2020.124082

Ho Sp, Anthes RA, Ao CO, Healy S, Horanyi A, Hunt D, Mannucci AJ, Pedatella N, Randel WJ, Simmons A, Steiner A, Xie F, Yue X, Zeng Z (2020b) The COSMIC/FORMOSAT-3 Radio Occultation Mission after 12 Years: Accomplishments, Remaining Challenges, and Potential Impacts of COSMIC-2. Bulletin of the American Meteorological Society 101(7):E1107–E1136, DOI 10.1175/BAMS-D-18-0290.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Hodnebrog Myhre G, Samset BH, Alterskjær K, Andrews T, Boucher O, Faluvegi G, Fläschner D, Forster PM, Kasoar M, Kirkevåg A, Lamarque JF, Olivié D, Richardson TB, Shawki D, Shindell D, Shine KP, Stier P, Takemura T, Voulgarakis A, Watson-Parris D (2019) Water vapour adjustments and responses differ between climate drivers. Atmospheric Chemistry and Physics 19(20):12887–12899, DOI 10.5194/acp-19-12887-2019, publisher: Copernicus GmbH

Hoekstra AY, Mekonnen MM (2012) The water footprint of humanity. Proceedings of the National Academy of Sciences 109(9):3232–3237, DOI 10.1073/pnas.1109936109, publisher: Proceedings of the National Academy of Sciences

Hoeting JA, Madigan D, Raftery AE, Volinsky CT (1999) Bayesian model averaging: a tutorial. Statistical science pp 382–401, publisher: JSTOR

Hollinger J (1991) DMSP special sensor microwave/imager calibration/validation. Tech. rep., NAVAL RESEARCH LAB WASHINGTON DC

Hong Y, Hsu KL, Sorooshian S, Gao X (2004) Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. Journal of Applied Meteorology and Climatology 43(12):1834–1853, DOI 10.1175/JAM2173.1, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology

Hong Y, Adler RF, Hossain F, Curtis S, Huffman GJ (2007) A first approach to global runoff simulation using satellite rainfall estimation. Water Resources Research 43(8), DOI 10.1029/2006WR005739, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2006WR005739

Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden P, Dai X, Maskell K, Johnson CA (2001) Climate Change 2001: The Scientific Basis. Cambridge University Press p 881

Huang J, Yu H, Guan X, Wang G, Guo R (2016) Accelerated dryland expansion under climate change. Nature Climate Change 6(2):166–171, DOI 10.1038/nclimate2837, number: 2 Publisher: Nature Publishing Group

Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB, Adler RF, Gu G, Hong Y, Bowman KP, Stocker EF (2007) The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology 8(1):38–55, DOI 10.1175/JHM560.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Huffman GJ, Adler RF, Bolvin DT, Gu G (2009) Improving the global precipitation record: GPCP Version 2.1. Geophysical Research Letters 36(17), DOI 10.1029/2009GL040000, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2009GL040000

Huffman GJ, Adler RF, Bolvin DT, Nelkin EJ (2010) The TRMM Multi-Satellite Precipitation Analysis (TMPA). In: Gebremichael M, Hossain F (eds) Satellite Rainfall Applications for Surface Hydrology, Springer Netherlands, Dordrecht, pp 3–22, DOI 10.1007/978-90-481-2915-7_1

Huffman GJ, Bolvin DT, Braithwaite D, Hsu K, Joyce R, Xie P, Yoo SH (2015) NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm Theoretical Basis Document (ATBD) Version 4:26

Huffman GJ, Bolvin DT, Braithwaite D, Hsu KL, Joyce RJ, Kidd C, Nelkin EJ, Sorooshian S, Stocker EF, Tan J, Wolff DB, Xie P (2020) Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). In: Levizzani V, Kidd C, Kirschbaum DB, Kummerow CD, Nakamura K, Turk FJ (eds) Satellite Precipitation Measurement: Volume 1, Advances in Global Change Research, Springer International Publishing, Cham, pp 343–353, DOI 10.1007/978-3-030-24568-9_19

Huffman GJ, Adler RF, Behrangi A, Bolvin DT, Nelkin EJ, Gu G, Ehsani MR (2023) The New Version 3.2 Global Precipitation Climatology Project (GPCP) Monthly and Daily Precipitation Products. Journal of Climate 36(21):7635–7655, DOI 10.1175/JCLI-D-23-0123.1, publisher: American Meteorological Society Section: Journal of Climate

Huntington TG, Weiskel PK, Wolock DM, McCabe GJ (2018) A new indicator framework for quantifying the intensity of the terrestrial water cycle. Journal of Hydrology 559:361–372, DOI 10.1016/j.jhydrol.2018.02.048

Hurrell JW, Holland MM, Gent PR, Ghan S, Kay JE, Kushner PJ, Lamarque JF, Large WG, Lawrence D, Lindsay K, Lipscomb WH, Long MC, Mahowald N, Marsh DR, Neale RB, Rasch P, Vavrus S, Vertenstein M, Bader D, Collins WD, Hack JJ, Kiehl J, Marshall S (2013) The Community Earth System Model: A Framework for Collaborative Research. Bulletin of the American Meteorological Society 94(9):1339–1360, DOI 10.1175/BAMS-D-12-00121.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Hänsel S, Ustrnul Z, Łupikasza E, Skalak P (2019) Assessing seasonal drought variations and trends over Central Europe. Advances in Water Resources 127:53–75, DOI 10.1016/j.advwatres.2019.03.005

Ihaka R, Gentleman R (1996) R: A Language for Data Analysis and Graphics. Journal of Computational and Graphical Statistics 5(3):299–314, DOI 10.1080/10618600.1996.10474713, publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/10618600.1996.10474713

IPCC IPoCC (2023) Water Cycle Changes. In: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, pp 1055–1210, DOI 10.1017/9781009157896.010

Jaagus J, Aasa A, Aniskevich S, Boincean B, Bojariu R, Briede A, Danilovich I, Castro FD, Dumitrescu A, Labuda M, Labudová L, Lõhmus K, Melnik V, Mõisja K, Pongracz R, Potopová V, Řezníčková L, Rimkus E, Semenova I, Stonevičius E, Štěpánek P, Trnka M, Vicente-Serrano SM, Wibig J, Zahradníček P (2022) Long-term changes in drought indices in eastern and central Europe. International Journal of Climatology 42(1):225–249, DOI 10.1002/joc.7241, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.7241

Janowiak JE, Xie P (1999) CAMS–OPI: A Global Satellite–Rain Gauge Merged Product for Real-Time Precipitation Monitoring Applications. Journal of Climate 12(11):3335–3342, DOI 10.1175/1520-0442(1999)012⟨3335:COAGSR⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Climate

Jasechko S, Sharp ZD, Gibson JJ, Birks SJ, Yi Y, Fawcett PJ (2013) Terrestrial water fluxes dominated by transpiration. Nature 496(7445):347–350, DOI 10.1038/nature11983, number: 7445 Publisher: Nature Publishing Group

Jenicek M, Ledvinka O (2020) Importance of snowmelt contribution to seasonal runoff and summer low flows in Czechia. Hydrology and Earth System Sciences 24(7):3475–3491, DOI 10.5194/hess-24-3475-2020, publisher: Copernicus GmbH

Jenicek M, Hnilica J, Nedelcev O, Sipek V (2021) Future changes in snowpack will impact seasonal runoff and low flows in Czechia. Journal of Hydrology: Regional Studies 37:100899, DOI 10.1016/j.ejrh.2021.100899

Johnson GC, Chambers DP (2013) Ocean bottom pressure seasonal cycles and decadal trends from GRACE Release-05: Ocean circulation implications. Journal of Geophysical Research: Oceans 118(9):4228–4240, DOI 10.1002/jgrc.20307, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jgrc.20307

Jones PW (1999) First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. Monthly Weather Review 127(9):2204–2210, DOI 10.1175/1520-0493(1999)127⟨2204:FASOCR⟩2.0.CO;2, publisher: American Meteorological Society Section: Monthly Weather Review

Joyce RJ, Janowiak JE, Arkin PA, Xie P (2004) CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. Journal of Hydrometeorology 5(3):487–503, DOI 10.1175/1525-7541(2004)005⟨0487:CAMTPG⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Jung M, Reichstein M, Ciais P, Seneviratne SI, Sheffield J, Goulden ML, Bonan G, Cescatti A, Chen J, de Jeu R, Dolman AJ, Eugster W, Gerten D, Gianelle D, Gobron N, Heinke J, Kimball J, Law BE, Montagnani L, Mu Q, Mueller B, Oleson K, Papale D, Richardson AD, Roupsard O, Running S, Tomelleri E, Viovy N, Weber U, Williams C, Wood E, Zaehle S, Zhang K (2010) Recent decline in the global land evapotranspiration trend due to limited moisture supply. Nature 467(7318):951–954, DOI 10.1038/nature09396, number: 7318 Publisher: Nature Publishing Group

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-Year Reanalysis Project. Bulletin of the American Meteorological Society 77(3):437–472, DOI 10.1175/1520-0477(1996)077⟨0437:TNYRP⟩2.0.CO;2, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Kanamitsu M, Ebisuzaki W, Woollen J, Yang SK, Hnilo JJ, Fiorino M, Potter GL (2002) NCEP–DOE AMIP-II Reanalysis (R-2). Bulletin of the American Meteorological Society 83(11):1631–1644, DOI 10.1175/BAMS-83-11-1631, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Kašpar M, Bližňák V, Hulec F, Müller M (2021) High-resolution spatial analysis of the variability in the subdaily rainfall time structure. Atmospheric Research 248:105202, DOI 10.1016/j.atmosres.2020.105202

Kašpárek L, Kožín R (2022) Changes in precipitation and runoff in river basins in the Czech Republic during the period of intense warming. Vodohospodářské technicko-ekonomické informace 64(2):12–27, DOI 10.46555/VTEI.2022.01.002, publisher: Výzkumný ústav vodohospodářský T. G. Masaryka, veřejná výzkumná instituce

Kemp WP, Burnell DG, Everson DO, Thomson AJ (1983) Estimating Missing Daily Maximum and Minimum Temperatures. Journal of Applied Meteorology and Climatology 22(9):1587–1593, DOI 10.1175/1520-0450(1983)022⟨1587:EMDMAM⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology

Kessler A (1968) Globalbilanzen von Klimaelementen: ein Beitrag zur allgemeinen Klimatologie der Erde. na

Kharin VV, Zwiers FW, Zhang X, Wehner M (2013) Changes in temperature and precipitation extremes in the CMIP5 ensemble. Climatic Change 119(2):345–357, DOI 10.1007/s10584-013-0705-8

Kibler KM, Biswas RK, Juarez Lucas AM (2014) Hydrologic data as a human right? Equitable access to information as a resource for disaster risk reduction in transboundary river basins. Water Policy 16(S2):36–58, DOI 10.2166/wp.2014.307

Kidd C, Huffman G (2011) Global precipitation measurement. Meteorological Applications 18(3):334–353, DOI 10.1002/met.284, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/met.284

Kidd C, Becker A, Huffman GJ, Muller CL, Joe P, Skofronick-Jackson G, Kirschbaum DB (2017) So, How Much of the Earth's Surface Is Covered by Rain Gauges? Bulletin of the American Meteorological Society 98(1):69–78, DOI 10.1175/BAMS-D-14-00283.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Kobayashi S, Ota Y, Harada Y, Ebita A, Moriya M, Onoda H, Onogi K, Kamahori H, Kobayashi C, Endo H, Miyaoka K, Takahashi K (2015) The JRA-55 Reanalysis: General Specifications and Basic Characteristics. Journal of the Meteorological Society of Japan Ser II 93(1):5–48, DOI 10.2151/jmsj.2015-001

Koirala S (2010) Explicit representation of groundwater process in a global-scale land surface model to improve hydrological predictions. PhD Thesis, University of Tokyo

Korzoun VI (1978) World water balance and water resources of the earth. Studies and Reports in Hydrology 25, publisher: UNESCO

Kubota T, Shige S, Hashizume H, Aonashi K, Takahashi N, Seto S, Hirose M, Takayabu YN, Ushio T, Nakagawa K, Iwanami K, Kachi M, Okamoto K (2007) Global Precipitation Map Using Satellite-Borne Microwave Radiometers by the GSMaP Project: Production and Validation. IEEE Transactions on Geoscience and Remote Sensing 45(7):2259–2275, DOI 10.1109/TGRS.2007.895337

Kumar R, Samaniego L, Attinger S (2013) Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. Water Resources Research 49(1):360–379, DOI 10.1029/2012WR012195, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2012WR012195

Kumar S, Allan RP, Zwiers F, Lawrence DM, Dirmeyer PA (2015) Revisiting trends in wetness and dryness in the presence of internal climate variability and water limitations over land. Geophysical Research Letters 42(24):10,867–10,875, DOI 10.1002/2015GL066858, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015GL066858

Kummerow C, Poyner P, Berg W, Thomas-Stahle J (2004) The Effects of Rainfall Inhomogeneity on Climate Variability of Rainfall Estimated from Passive Microwave Sensors. Journal of Atmospheric and Oceanic Technology 21(4):624–638, DOI 10.1175/1520-0426(2004)021⟨0624:TEORIO⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Atmospheric and Oceanic Technology

Kunkee DB, Poe GA, Boucher DJ, Swadley SD, Hong Y, Wessel JE, Uliana EA (2008) Design and Evaluation of the First Special Sensor Microwave Imager/Sounder. IEEE Transactions on Geoscience and Remote Sensing 46(4):863–883, DOI 10.1109/TGRS.2008.917980, conference Name: IEEE Transactions on Geoscience and Remote Sensing

Kyselý J, Beranová R (2009) Climate-change effects on extreme precipitation in central Europe: uncertainties of scenarios based on regional climate models. Theoretical and Applied Climatology 95(3):361–374, DOI 10.1007/s00704-008-0014-8

Kyselý J, Gaál L, Beranová R, Plavcová E (2011) Climate change scenarios of precipitation extremes in Central Europe from ENSEMBLES regional climate models. Theoretical and Applied Climatology 104(3):529–542, DOI 10.1007/s00704-010-0362-z

Landerer FW, Swenson SC (2012) Accuracy of scaled GRACE terrestrial water storage estimates. Water Resources Research 48(4), DOI 10.1029/2011WR011453, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2011WR011453

Lavers DA, Simmons A, Vamborg F, Rodwell MJ (2022) An evaluation of ERA5 precipitation for climate monitoring. Quarterly Journal of the Royal Meteorological Society 148(748):3152–3165, DOI 10.1002/qj.4351, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.4351

Lawford RG (1999) A midterm report on the GEWEX Continental-Scale International Project (GCIP). Journal of Geophysical Research: Atmospheres 104(D16):19279–19292, DOI 10.1029/1999JD900266, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/1999JD900266

L'Ecuyer TS, Stephens GL (2002) An Estimation-Based Precipitation Retrieval Algorithm for Attenuating Radars. Journal of Applied Meteorology and Climatology 41(3):272–285, DOI 10.1175/1520-0450(2002)041⟨0272:AEBPRA⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology

Levizzani V, Cattani E (2019) Satellite Remote Sensing of Precipitation and the Terrestrial Water Cycle in a Changing Climate. Remote Sensing 11(19):2301, DOI 10.3390/rs11192301, number: 19 Publisher: Multidisciplinary Digital Publishing Institute

Lhotka O, Trnka M, Kyselý J, Markonis Y, Balek J, Možný M (2020) Atmospheric Circulation as a Factor Contributing to Increasing Drought Severity in Central Europe. Journal of Geophysical Research: Atmospheres 125(18):e2019JD032269, DOI 10.1029/2019JD032269, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019JD032269

Liang X, Lettenmaier DP, Wood EF, Burges SJ (1994) A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research: Atmospheres 99(D7):14415–14428, DOI 10.1029/94JD00483, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/94JD00483

Lim WH, Roderick ML (2009) An Atlas of the Global Water Cycle: Based on the IPCC AR4 Climate Models, 1st edn. ANU Press, DOI 10.22459/AGWC.07.2009

Loaiciga HA, Valdes JB, Vogel R, Garvey J, Schwarz H (1996) Global warming and the hydrologic cycle. Journal of Hydrology 174(1):83–127, DOI 10.1016/0022-1694(95)02753-X

Lombardo F, Volpi E, Koutsoyiannis D, Serinaldi F (2017) A theoretically consistent stochastic cascade for temporal disaggregation of intermittent rainfall. Water Resources Research 53(6):4586–4605, DOI 10.1002/2017WR020529, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017WR020529

Lorenz C, Kunstmann H (2012) The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis. Journal of Hydrometeorology 13(5):1397–1420, DOI 10.1175/JHM-D-11-088.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Lorenz C, Kunstmann H, Devaraju B, Tourian MJ, Sneeuw N, Riegger J (2014) Large-Scale Runoff from Landmasses: A Global Assessment of the Closure of the Hydrological and Atmospheric Water Balances. Journal of Hydrometeorology 15(6):2111–2139, DOI 10.1175/JHM-D-13-0157.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

L'vovitch M (1945) World Water Regime Elements. Sverdlovsk, Moscow, Russia

L'vovitch M (1970) World water balance (general report). In: World water balance: Proceedings of the Reading Symposium, pp 401–415

Lvovitch MI (1973) The global water balance. Eos, Transactions American Geophysical Union 54(1):28–53, DOI 10.1029/EO054i001p00028, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/EO054i001p00028

Maggioni V, Meyers PC, Robinson MD (2016) A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era. Journal of Hydrometeorology 17(4):1101–1117, DOI 10.1175/JHM-D-15-0190.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Manabe S, Bryan K (1969) Climate Calculations with a Combined Ocean-Atmosphere Model. Journal of the Atmospheric Sciences 26(4):786–789, DOI 10.1175/1520-0469(1969)026⟨0786:CCWACO⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences

Manabe S, Wetherald RT (1975) The Effects of Doubling the CO2 Concentration on the climate of a General Circulation Model. Journal of the Atmospheric Sciences 32(1):3–15, DOI 10.1175/1520-0469(1975)032⟨0003:TEODTC⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences

Marcinek J (1964) Der Abfluß von den Landflächen der Erde und seine Verteilung auf 5° Zonen. Berliner geographische Arbeiten, Verlag für Bauwesen

Marengo JA (2005) Characteristics and spatio-temporal variability of the Amazon River Basin Water Budget. Climate Dynamics 24(1):11–22, DOI 10.1007/s00382-004-0461-6

Markonis Y, Strnad F (2020) Representation of European hydroclimatic patterns with self-organizing maps. The Holocene 30(8):1155–1162, DOI 10.1177/0959683620913924, publisher: SAGE Publications Ltd

Markonis Y, Hanel M, Máca P, Kyselý J, Cook ER (2018) Persistent multi-scale fluctuations shift European hydroclimate to its millennial boundaries. Nature Communications 9(1):1767, DOI 10.1038/s41467-018-04207-7, number: 1 Publisher: Nature Publishing Group

Markonis Y, Papalexiou SM, Martinkova M, Hanel M (2019) Assessment of Water Cycle Intensification Over Land using a Multisource Global Gridded Precipitation DataSet. Journal of Geophysical Research: Atmospheres 124(21):11175–11187, DOI 10.1029/2019JD030855, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019JD030855

Markonis Y, Kumar R, Hanel M, Rakovec O, Máca P, AghaKouchak A (2021a) The rise of compound warm-season droughts in Europe. Science Advances 7(6):eabb9668, DOI 10.1126/sciadv.abb9668, publisher: American Association for the Advancement of Science

Markonis Y, Pappas C, Hanel M, Papalexiou SM (2021b) A cross-scale framework for integrating multi-source data in Earth system sciences. Environmental Modelling & Software 139:104997, DOI 10.1016/j.envsoft.2021.104997

Marques AC, Veras CE, Rodriguez DA (2022) Assessment of water policies contributions for sustainable water resources management under climate change scenarios. Journal of Hydrology 608:127690, DOI 10.1016/j.jhydrol.2022.127690

Mather JR (1962) Average climatic water balance data of the continents: part 1. Africa. Publications in Climatology

Mather JR (1963a) Average climatic water balance data of the continents: part 2. Asia (excluding U.S.S.R.). Publications in Climatology

Mather JR (1963b) Average climatic water balance data of the continents: part 3. U.S.S.R. Publications in Climatology

Mather JR (1963c) Average climatic water balance data of the continents: part 4. Australia, New Zeland, and Oceania. Publications in Climatology

Mather JR (1964a) Average climatic water balance data of the continents: part 5. Europa. Publications in Climatology

Mather JR (1964b) Average climatic water balance data of the continents: part 6. North America (excluding United States). Publications in Climatology

Mather JR (1964c) Average climatic water balance data of the continents: part 7. United States. Publications in Climatology

Mather JR (1965) Average climatic water balance data of the continents: part 8. South America. Publications in Climatology

Mather JR (1969) The average annual water balance of the world. AWRA Symposium

McCabe MF, Ershadi A, Jimenez C, Miralles DG, Michel D, Wood EF (2016) The GEWEX LandFlux project: evaluation of model evaporation using tower-based and globally gridded forcing data. Geoscientific Model Development 9(1):283–305, DOI 10.5194/gmd-9-283-2016, publisher: Copernicus GmbH

McCabe MF, Rodell M, Alsdorf DE, Miralles DG, Uijlenhoet R, Wagner W, Lucieer A, Houborg R, Verhoest NEC, Franz TE, Shi J, Gao H, Wood EF (2017) The future of Earth observation in hydrology. Hydrology and Earth System Sciences 21(7):3879–3914, DOI 10.5194/hess-21-3879-2017, publisher: Copernicus GmbH

McColl KA, Roderick ML, Berg A, Scheff J (2022) The terrestrial water cycle in a warming world. Nature Climate Change DOI 10.1038/s41558-022-01412-7

McGuffie K, Henderson-Sellers A (2001) Forty years of numerical climate modelling. International Journal of Climatology 21(9):1067–1109, DOI 10.1002/joc.632, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.632

McNally A, Arsenault K, Kumar S, Shukla S, Peterson P, Wang S, Funk C, Peters-Lidard CD, Verdin JP (2017) A land data assimilation system for sub-Saharan Africa food and water security applications. Scientific Data 4(1):170012, DOI 10.1038/sdata.2017.12, number: 1 Publisher: Nature Publishing Group

Meigh JR, McKenzie AA, Sene KJ (1999) A Grid-Based Approach to Water Scarcity Estimates for Eastern and Southern Africa. Water Resources Management 13(2):85–115, DOI 10.1023/A:1008025703712

Meinardus W (1934) Eine neue niederschlagskarte der erde. Petermanns Geogr Mitt 80:1–4

Mira A (1964) Physical geographical atlas of the world, Moscow, Russia, 1964. Sovietgeography, Review and Translation 6

Miralles DG, Holmes TRH, De Jeu RaM, Gash JH, Meesters AGCA, Dolman AJ (2011) Global land-surface evaporation estimated from satellite-based observations. Hydrology and Earth System Sciences 15(2):453–469, DOI 10.5194/hess-15-453-2011, publisher: Copernicus GmbH

Mitchell JFB, Wilson CA, Cunnington WM (1987) On CO2 climate sensitivity and model dependence of results. Quarterly Journal of the Royal Meteorological Society 113(475):293–322, DOI 10.1002/qj.49711347517, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49711347517

Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology 25(6):693–712, DOI 10.1002/joc.1181, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.1181

Moazamnia M, Hassanzadeh Y, Nadiri AA, Khatibi R, Sadeghfam S (2019) Formulating a strategy to combine artificial intelligence models using Bayesian model averaging to study a distressed aquifer with sparse data availability. Journal of Hydrology 571:765–781, DOI 10.1016/j.jhydrol.2019.02.011

Moler C, Little J, Bangert S (1982) MATLAB users' guide. University of New Mexico Albuquerque, NM, USA

Monteith J, Unsworth M (2013) Principles of environmental physics: plants, animals, and the atmosphere. Academic Press

Moravec V, Markonis Y, Rakovec O, Svoboda M, Trnka M, Kumar R, Hanel M (2021) Europe under multi-year droughts: how severe was the 2014–2018 drought period? Environmental Research Letters 16(3):034062, DOI 10.1088/1748-9326/abe828, publisher: IOP Publishing

Morice CP, Kennedy JJ, Rayner NA, Winn JP, Hogan E, Killick RE, Dunn RJH, Osborn TJ, Jones PD, Simpson IR (2021) An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set. Journal of Geophysical Research: Atmospheres 126(3):e2019JD032361, DOI 10.1029/2019JD032361

Mortimer A, Ahmed I, Johnson T, Tang L, Alston M (2023) Localizing Sustainable Development Goal 13 on Climate Action to Build Local Resilience to Floods in the Hunter Valley: A Literature Review. Sustainability 15(6):5565, DOI 10.3390/su15065565, number: 6 Publisher: Multidisciplinary Digital Publishing Institute

Mozny M, Trnka M, Vlach V, Vizina A, Potopova V, Zahradnicek P, Stepanek P, Hajkova L, Staponites L, Zalud Z (2020) Past (1971–2018) and future (2021–2100) pan evaporation rates in the Czech Republic. Journal of Hydrology 590:125390, DOI 10.1016/j.jhydrol.2020.125390

Mu Q, Heinsch FA, Zhao M, Running SW (2007) Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sensing of Environment 111(4):519–536, DOI 10.1016/j.rse.2007.04.015

Mu Q, Zhao M, Running SW (2011) Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sensing of Environment 115(8):1781–1800, DOI 10.1016/j.rse.2011.02.019

Muelchi R, Rössler O, Schwanbeck J, Weingartner R, Martius O (2021) River runoff in Switzerland in a changing climate – runoff regime changes and their time of emergence. Hydrology and Earth System Sciences 25(6):3071–3086, DOI 10.5194/hess-25-3071-2021, publisher: Copernicus GmbH

Muller C, Chapman L, Johnston S, Kidd C, Illingworth S, Foody G, Overeem A, Leigh R (2015) Crowdsourcing for climate and atmospheric sciences: current status and future potential. International Journal of Climatology 35(11):3185–3203, DOI 10.1002/joc.4210, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.4210

Munier S, Aires F (2018) A new global method of satellite dataset merging and quality characterization constrained by the terrestrial water budget. Remote Sensing of Environment 205:119–130, DOI 10.1016/j.rse.2017.11.008

Muñoz-Sabater J, Dutra E, Agustí-Panareda A, Albergel C, Arduini G, Balsamo G, Boussetta S, Choulga M, Harrigan S, Hersbach H, Martens B, Miralles DG, Piles M, Rodríguez-Fernández NJ, Zsoter E, Buontempo C, Thépaut JN (2021) ERA5-Land: a state-of-the-art global re-

analysis dataset for land applications. Earth System Science Data 13(9):4349–4383, DOI 10.5194/essd-13-4349-2021, publisher: Copernicus GmbH

Möller F (1951) Quarterly charts of rainfall for the whole earth. Petermanns Geograph Mitt 95:1–7

Nace RL (1968) Water of the World Geological Survey. New Release, backup Publisher: U.S. Dept. of Interior

Nedelcev O, Jenicek M (2021) Trends in seasonal snowpack and their relation to climate variables in mountain catchments in Czechia. Hydrological Sciences Journal 66(16):2340–2356, DOI 10.1080/02626667.2021.1990298, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02626667.2021.1990298

Newman AJ, Clark MP, Craig J, Nijssen B, Wood A, Gutmann E, Mizukami N, Brekke L, Arnold JR (2015) Gridded Ensemble Precipitation and Temperature Estimates for the Contiguous United States. Journal of Hydrometeorology 16(6):2481–2500, DOI 10.1175/JHM-D-15-0026.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Newman AJ, Clark MP, Longman RJ, Gilleland E, Giambelluca TW, Arnold JR (2019) Use of Daily Station Observations to Produce High-Resolution Gridded Probabilistic Precipitation and Temperature Time Series for the Hawaiian Islands. Journal of Hydrometeorology 20(3):509–529, DOI 10.1175/JHM-D-18-0113.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

NOAA US (1987) Space-based remote sensing of the earth: a report to the Congress. NASA

NRC (1986) Global Change in the Geosphere-Biosphere. National Academy Press

Oki T (1999) 1.2 the Global Water Cycle. Global Energy and Water Cycles 134800000:10

Oki T (2006) The Hydrologic Cycles and Global Circulation. In: Encyclopedia of Hydrological Sciences, John Wiley & Sons, Ltd, DOI 10.1002/0470848944.hsa001, section: 2 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/0470848944.hsa001

Oki T, Kanae S (2006) Global Hydrological Cycles and World Water Resources. Science 313(5790):1068–1072, DOI 10.1126/science.1128845, publisher: American Association for the Advancement of Science

Onogi K, Tsutsui J, Koide H, Sakamoto M, Kobayashi S, Hatsushika H, Matsumoto T, Yamazaki N, Kamahori H, Takahashi K, Kadokura S, Wada K, Kato K, Oyama R, Ose T, Mannoji N,

Taira R (2007) The JRA-25 Reanalysis. Journal of the Meteorological Society of Japan Ser II 85(3):369–432, DOI 10.2151/jmsj.85.369

Otto-Bliesner BL, Brady EC, Fasullo J, Jahn A, Landrum L, Stevenson S, Rosenbloom N, Mai A, Strand G (2016) Climate Variability and Change since 850 CE: An Ensemble Approach with the Community Earth System Model. Bulletin of the American Meteorological Society 97(5):735–754, DOI 10.1175/BAMS-D-14-00233.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

O'Gorman PA, Muller CJ (2010) How closely do changes in surface and column water vapor follow Clausius–Clapeyron scaling in climate change simulations? Environmental Research Letters 5(2):025207, DOI 10.1088/1748-9326/5/2/025207

O'Gorman PA, Allan RP, Byrne MP, Previdi M (2012) Energetic Constraints on Precipitation Under Climate Change. Surveys in Geophysics 33(3):585–608, DOI 10.1007/s10712-011-9159-6

Palissy B (1580) Discours admirables. Martin Le Jeune, Paris

Pan M, Wood EF (2006) Data Assimilation for Estimating the Terrestrial Water Budget Using a Constrained Ensemble Kalman Filter. Journal of Hydrometeorology 7(3):534–547, DOI 10.1175/JHM495.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Pan M, Sahoo AK, Troy TJ, Vinukollu RK, Sheffield J, Wood EF (2012) Multisource Estimation of Long-Term Terrestrial Water Budget for Major Global River Basins. Journal of Climate 25(9):3191–3206, DOI 10.1175/JCLI-D-11-00300.1, publisher: American Meteorological Society Section: Journal of Climate

Pan M, Fisher CK, Chaney NW, Zhan W, Crow WT, Aires F, Entekhabi D, Wood EF (2015) Triple collocation: Beyond three estimates and separation of structural/non-structural errors. Remote Sensing of Environment 171:299–310, DOI 10.1016/j.rse.2015.10.028

Papalexiou SM (2018) Unified theory for stochastic modelling of hydroclimatic processes: Preserving marginal distributions, correlation structures, and intermittency. Advances in Water Resources 115:234–252, DOI 10.1016/j.advwatres.2018.02.013

Papalexiou SM, Markonis Y, Lombardo F, AghaKouchak A, Foufoula-Georgiou E (2018) Precise Temporal Disaggregation Preserving Marginals and Correlations (DiPMaC) for Stationary and Nonstationary Processes. Water Resources Research 54(10):7435–7458, DOI 10.1029/2018WR022726, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018WR022726

Papalexiou SM, Serinaldi F, Strnad F, Markonis Y, Shook K (2021) CoSMoS: Complete Stochastic Modelling Solution

Pappas C, Papalexiou SM, Koutsoyiannis D (2014) A quick gap filling of missing hydrometeorological data. Journal of Geophysical Research: Atmospheres 119(15):9290–9300, DOI 10.1002/2014JD021633, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2014JD021633

Parker WS (2016) Reanalyses and Observations: What's the Difference? Bulletin of the American Meteorological Society 97(9):1565–1572, DOI 10.1175/BAMS-D-14-00226.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Pastorello G, Trotta C, Canfora E, Chu H, Christianson D, Cheah YW, Poindexter C, Chen J, Elbashandy A, Humphrey M, Isaac P, Polidori D, Reichstein M, Ribeca A, van Ingen C, Vuichard N, Zhang L, Amiro B, Ammann C, Arain MA, Ardö J, Arkebauer T, Arndt SK, Arriga N, Aubinet M, Aurela M, Baldocchi D, Barr A, Beamesderfer E, Marchesini LB, Bergeron O, Beringer J, Bernhofer C, Berveiller D, Billesbach D, Black TA, Blanken PD, Bohrer G, Boike J, Bolstad PV, Bonal D, Bonnefond JM, Bowling DR, Bracho R, Brodeur J, Brümmer C, Buchmann N, Burban B, Burns SP, Buysse P, Cale P, Cavagna M, Cellier P, Chen S, Chini I, Christensen TR, Cleverly J, Collalti A, Consalvo C, Cook BD, Cook D, Coursolle C, Cremonese E, Curtis PS, D'Andrea E, da Rocha H, Dai X, Davis KJ, Cinti BD, Grandcourt Ad, Ligne AD, De Oliveira RC, Delpierre N, Desai AR, Di Bella CM, Tommasi Pd, Dolman H, Domingo F, Dong G, Dore S, Duce P, Dufrêne E, Dunn A, Dušek J, Eamus D, Eichelmann U, ElKhidir HAM, Eugster W, Ewenz CM, Ewers B, Famulari D, Fares S, Feigenwinter I, Feitz A, Fensholt R, Filippa G, Fischer M, Frank J, Galvagno M, Gharun M, Gianelle D, Gielen B, Gioli B, Gitelson A, Goded I, Goeckede M, Goldstein AH, Gough CM, Goulden ML, Graf A, Griebel A, Gruening C, Grünwald T, Hammerle A, Han S, Han X, Hansen BU, Hanson C, Hatakka J, He Y, Hehn M, Heinesch B, Hinko-Najera N, Hörtnagl L, Hutley L, Ibrom A, Ikawa H, Jackowicz-Korczynski M, Janouš D, Jans W, Jassal R, Jiang S, Kato T, Khomik M, Klatt J, Knohl A, Knox S, Kobayashi H, Koerber G, Kolle O, Kosugi Y, Kotani A, Kowalski A, Kruijt B, Kurbatova J, Kutsch WL, Kwon H, Launiainen S, Laurila T, Law B, Leuning R, Li Y, Liddell M, Limousin JM, Lion M, Liska AJ, Lohila A, López-Ballesteros A, López-Blanco E, Loubet B, Loustau D, Lucas-Moffat A, Lüers J, Ma S, Macfarlane C, Magliulo V, Maier R, Mammarella I, Manca G, Marcolla B, Margolis HA, Marras S, Massman W, Mastepanov M, Matamala R, Matthes JH, Mazzenga F, McCaughey H, McHugh I, McMillan AMS, Merbold L, Meyer W, Meyers T, Miller SD, Minerbi S, Moderow U, Monson RK, Montagnani L, Moore CE, Moors E, Moreaux V, Moureaux C, Munger JW, Nakai

T, Neirynck J, Nesic Z, Nicolini G, Noormets A, Northwood M, Nosetto M, Nouvellon Y, Novick K, Oechel W, Olesen JE, Ourcival JM, Papuga SA, Parmentier FJ, Paul-Limoges E, Pavelka M, Peichl M, Pendall E, Phillips RP, Pilegaard K, Pirk N, Posse G, Powell T, Prasse H, Prober SM, Rambal S, Rannik U, Raz-Yaseef N, Rebmann C, Reed D, Dios VRd, Restrepo-Coupe N, Reverter BR, Roland M, Sabbatini S, Sachs T, Saleska SR, Sánchez-Cañete EP, Sanchez-Mejia ZM, Schmid HP, Schmidt M, Schneider K, Schrader F, Schroder I, Scott RL, Sedlák P, Serrano-Ortíz P, Shao C, Shi P, Shironya I, Siebicke L, Šigut L, Silberstein R, Sirca C, Spano D, Steinbrecher R, Stevens RM, Sturtevant C, Suyker A, Tagesson T, Takanashi S, Tang Y, Tapper N, Thom J, Tomassucci M, Tuovinen JP, Urbanski S, Valentini R, van der Molen M, van Gorsel E, van Huissteden K, Varlagin A, Verfaillie J, Vesala T, Vincke C, Vitale D, Vygodskaya N, Walker JP, Walter-Shea E, Wang H, Weber R, Westermann S, Wille C, Wofsy S, Wohlfahrt G, Wolf S, Woodgate W, Li Y, Zampedri R, Zhang J, Zhou G, Zona D, Agarwal D, Biraud S, Torn M, Papale D (2020) The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Scientific Data 7(1):225, DOI 10.1038/s41597-020-0534-3, number: 1 Publisher: Nature Publishing Group

Pathirana A, Herath S, Yamada T (2003) Estimating rainfall distributions at high temporal resolutions using a multifractal model. Hydrology and Earth System Sciences 7(5):668–679, DOI 10.5194/hess-7-668-2003, publisher: Copernicus GmbH

Pellet V, Aires F, Munier S, Fernández Prieto D, Jordá G, Dorigo WA, Polcher J, Brocca L (2019) Integrating multiple satellite observations into a coherent dataset to monitor the full water cycle – application to the Mediterranean region. Hydrology and Earth System Sciences 23(1):465–491, DOI 10.5194/hess-23-465-2019, publisher: Copernicus GmbH

Pendergrass AG (2018) What precipitation is extreme? Science 360(6393):1072–1073, DOI 10.1126/science.aat1871, publisher: American Association for the Advancement of Science

Pendergrass AG, Hartmann DL (2014) Two Modes of Change of the Distribution of Rain. Journal of Climate 27(22):8357–8371, DOI 10.1175/JCLI-D-14-00182.1, publisher: American Meteorological Society Section: Journal of Climate

Peterson TC, Vose RS (1997) An Overview of the Global Historical Climatology Network Temperature Database. Bulletin of the American Meteorological Society 78(12):2837–2850, DOI 10.1175/1520-0477(1997)078⟨2837:AOOTGH⟩2.0.CO;2, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Petković V, Kummerow CD (2017) Understanding the Sources of Satellite Passive Microwave Rainfall Retrieval Systematic Errors Over Land. Journal of Applied Meteorology and Climatology 56(3):597–614, DOI 10.1175/JAMC-D-16-0174.1, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology

Pfister L, Savenije HHG, Fenicia F (2009) Leonardo Da Vinci's water theory: on the origin and fate of water. No. 9 in IAHS special publication, International Association of Hydrological Sciences, Wallingford

Phillips NA (1956) The general circulation of the atmosphere: A numerical experiment. Quarterly Journal of the Royal Meteorological Society 82(352):123–164, DOI 10.1002/qj.49708235202, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49708235202

Pletzer A, Hayek W (2019) Mimetic Interpolation of Vector Fields on Arakawa C/D Grids. Monthly Weather Review 147(1):3–16, DOI 10.1175/MWR-D-18-0146.1, publisher: American Meteorological Society Section: Monthly Weather Review

Poli P, Hersbach H, Dee DP, Berrisford P, Simmons AJ, Vitart F, Laloyaux P, Tan DGH, Peubey C, Thépaut JN, Trémolet Y, Hólm EV, Bonavita M, Isaksen L, Fisher M (2016) ERA-20C: An Atmospheric Reanalysis of the Twentieth Century. Journal of Climate 29(11):4083–4097, DOI 10.1175/JCLI-D-15-0556.1, publisher: American Meteorological Society Section: Journal of Climate

Pollio MV (1648) De Architectura, Liber Octavus. In: Vitruvius (ed) De Architectura, Elsevier, pp 150 – 172

Postel SL, Daily GC, Ehrlich PR (1996) Human Appropriation of Renewable Fresh Water. Science 271(5250):785–788, DOI 10.1126/science.271.5250.785, publisher: American Association for the Advancement of Science

Potopová V, Štěpánek P, Možný M, Türkott L, Soukup J (2015) Performance of the standardised precipitation evapotranspiration index at various lags for agricultural drought risk assessment in the Czech Republic. Agricultural and Forest Meteorology 202:26–38, DOI 10.1016/j.agrformet.2014.11.022

Povey AC, Grainger RG (2015) Known and unknown unknowns: uncertainty estimation in satellite remote sensing. Atmospheric Measurement Techniques 8(11):4699–4718, DOI 10.5194/amt-8-4699-2015, publisher: Copernicus GmbH

Prein AF, Pendergrass AG (2019) Can We Constrain Uncertainty in Hydrologic Cycle Projections? Geophysical Research Letters 46(7):3911–3916, DOI 10.1029/2018GL081529, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018GL081529

Qian T, Dai A, Trenberth KE, Oleson KW (2006) Simulation of Global Land Surface Conditions from 1948 to 2004. Part I: Forcing Data and Evaluations. Journal of Hydrometeorology 7(5):953–975, DOI 10.1175/JHM540.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL `https://www.R-project.org/`

Rajulapati CR, Papalexiou SM, Clark MP, Pomeroy JW (2021) The Perils of Regridding: Examples Using a Global Precipitation Dataset. Journal of Applied Meteorology and Climatology 60(11):1561–1573, DOI 10.1175/JAMC-D-20-0259.1, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology

Rakovec O, Kumar R, Attinger S, Samaniego L (2016a) Improving the realism of hydrologic model functioning through multivariate parameter estimation. Water Resources Research 52(10):7779–7792, DOI 10.1002/2016WR019430, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2016WR019430

Rakovec O, Kumar R, Mai J, Cuntz M, Thober S, Zink M, Attinger S, Schäfer D, Schrön M, Samaniego L (2016b) Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins. Journal of Hydrometeorology 17(1):287–307, DOI 10.1175/JHM-D-15-0054.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Rakovec O, Samaniego L, Hari V, Markonis Y, Moravec V, Thober S, Hanel M, Kumar R (2022) The 2018–2020 Multi-Year Drought Sets a New Benchmark in Europe. Earth's Future 10(3):e2021EF002394, DOI 10.1029/2021EF002394, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2021EF002394

Raschke E, Karstens U, Nolte-Holube R, Brandt R, Isemer HJ, Lohmann D, Lobmeyr M, Rockel B, Stuhlmann R (1998) The Baltic Sea Experiment BALTEX: A brief overview and some selected results of the authors. Surveys in Geophysics 19(1):1–22, DOI 10.1023/A:1006567924820

Raschke E, Meywerk J, Warrach K, Andrea U, Bergström S, Beyrich F, Bosveld F, Bumke K, Fortelius C, Graham LP, Gryning SE, Halldin S, Hasse L, Heikinheimo M, Isemer HJ, Ja-

cob D, Jauja I, Karlsson KG, Keevallik S, Koistinen J, Lammeren Av, Lass U, Launianen J, Lehmann A, Liljebladh B, Lobmeyr M, Matthäus W, Mengelkamp T, Michelson DB, Napiórkowski J, Omstedt A, Piechura J, Rockel B, Rubel F, Ruprecht E, Smedman AS, Stigebrandt A (2001) The Baltic Sea Experiment (BALTEX): A European Contribution to the Investigation of the Energy and Water Cycle over a Large Drainage Basin. Bulletin of the American Meteorological Society 82(11):2389–2414, DOI 10.1175/1520-0477(2001)082⟨2389:TBSEBA⟩2.3.CO;2, URL `https://journals.ametsoc.org/view/journals/bams/82/11/1520-0477_2001_082_2389_tbseba_2_3_co_2.xml`, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Rasmussen JL (1970) The Atmospheric Water Balance and the Hydrology of Large River Basins1. JAWRA Journal of the American Water Resources Association 6(4):631–639, DOI 10.1111/j.1752-1688.1970.tb00523.x, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1752-1688.1970.tb00523.x

Redelsperger JL, Thorncroft CD, Diedhiou A, Lebel T, Parker DJ, Polcher J (2006) African Monsoon Multidisciplinary Analysis: An International Research Project and Field Campaign. Bulletin of the American Meteorological Society 87(12):1739–1746, DOI 10.1175/BAMS-87-12-1739, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Reichel E (1952) Der Stand des Verdunstungsproblems. Ber Dt Wetterdienst US-Zone 35:155–172

Reichle R (2012) The MERRA-land data product (version 1.2). GMAO Off Note 3

Richardson TB, Forster PM, Andrews T, Boucher O, Faluvegi G, Fläschner D, Hodnebrog Kasoar M, Kirkevåg A, Lamarque JF, Myhre G, Olivié D, Samset BH, Shawki D, Shindell D, Takemura T, Voulgarakis A (2018) Drivers of Precipitation Change: An Energetic Understanding. Journal of Climate 31(23):9641–9657, DOI 10.1175/JCLI-D-17-0240.1, publisher: American Meteorological Society Section: Journal of Climate

Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim GK, Bloom S, Chen J, Collins D, Conaty A, Silva Ad, Gu W, Joiner J, Koster RD, Lucchesi R, Molod A, Owens T, Pawson S, Pegion P, Redder CR, Reichle R, Robertson FR, Ruddick AG, Sienkiewicz M, Woollen J (2011) MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. Journal of Climate 24(14):3624–3648, DOI 10.1175/JCLI-D-11-00015.1, publisher: American Meteorological Society Section: Journal of Climate

Rivoire P, Le Gall P, Favre AC, Naveau P, Martius O (2022) High return level estimates of daily ERA-5 precipitation in Europe estimated using regionalized extreme value distributions. Weather and Climate Extremes 38:100500, DOI 10.1016/j.wace.2022.100500

Robertson FR, Bosilovich MG, Roberts JB, Reichle RH, Adler R, Ricciardulli L, Berg W, Huffman GJ (2014) Consistency of Estimated Global Water Cycle Variations over the Satellite Era. Journal of Climate 27(16):6135–6154, DOI 10.1175/JCLI-D-13-00384.1, publisher: American Meteorological Society Section: Journal of Climate

Rockström J, Falkenmark M, Lannerstad M, Karlberg L (2012) The planetary water drama: Dual task of feeding humanity and curbing climate change. Geophysical Research Letters 39(15), DOI 10.1029/2012GL051688, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2012GL051688

Rodell M, Houser PR, Jambor U, Gottschalck J, Mitchell K, Meng CJ, Arsenault K, Cosgrove B, Radakovich J, Bosilovich M, Entin JK, Walker JP, Lohmann D, Toll D (2004) The Global Land Data Assimilation System. Bulletin of the American Meteorological Society 85(3):381–394, DOI 10.1175/BAMS-85-3-381, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Rodell M, Beaudoing HK, L'Ecuyer TS, Olson WS, Famiglietti JS, Houser PR, Adler R, Bosilovich MG, Clayson CA, Chambers D, Clark E, Fetzer EJ, Gao X, Gu G, Hilburn K, Huffman GJ, Lettenmaier DP, Liu WT, Robertson FR, Schlosser CA, Sheffield J, Wood EF (2015) The Observed State of the Water Cycle in the Early Twenty-First Century. Journal of Climate 28(21):8289–8318, DOI 10.1175/JCLI-D-14-00555.1, publisher: American Meteorological Society Section: Journal of Climate

Roderick ML, Sun F, Lim WH, Farquhar GD (2014) A general framework for understanding the response of the water cycle to global warming over land and ocean. Hydrology and Earth System Sciences 18(5):1575–1589, DOI 10.5194/hess-18-1575-2014, publisher: Copernicus GmbH

Rodgers CD (2000) Inverse Methods for Atmospheric Sounding: Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics, vol 2. WORLD SCIENTIFIC, DOI 10.1142/3171

de Rosnay P, Polcher J (1998) Modelling root water uptake in a complex land surface scheme coupled to a GCM. Hydrology and Earth System Sciences 2(2/3):239–255, DOI 10.5194/hess-2-239-1998, publisher: Copernicus GmbH

Rost S, Gerten D, Bondeau A, Lucht W, Rohwer J, Schaphoff S (2008) Agricultural green and blue water consumption and its influence on the global water system. Water Resources Research 44(9), DOI 10.1029/2007WR006331, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2007WR006331

Roth N, Jaramillo F, Wang-Erlandsson L, Zamora D, Palomino-Ángel S, Cousins SAO (2021) A call for consistency with the terms 'wetter' and 'drier' in climate change studies. Environmental Evidence 10(1):8, DOI 10.1186/s13750-021-00224-0

Rudolf B, Schneider U (2005) Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations. In: Proc. Second Workshop of the Int. Precipitation Working Group, pp 231–247

Rustemeier E, Ziese M, Meyer-Christoffer A, Schneider U, Finger P, Becker A (2019) Uncertainty Assessment of the ERA-20C Reanalysis Based on the Monthly In Situ Precipitation Analysis of the Global Precipitation Climatology Centre. Journal of Hydrometeorology 20(2):231–250, DOI 10.1175/JHM-D-17-0239.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Saha S, Moorthi S, Pan HL, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu H, Stokes D, Grumbine R, Gayno G, Wang J, Hou YT, Chuang Hy, Juang HMH, Sela J, Iredell M, Treadon R, Kleist D, Delst PV, Keyser D, Derber J, Ek M, Meng J, Wei H, Yang R, Lord S, Dool Hvd, Kumar A, Wang W, Long C, Chelliah M, Xue Y, Huang B, Schemm JK, Ebisuzaki W, Lin R, Xie P, Chen M, Zhou S, Higgins W, Zou CZ, Liu Q, Chen Y, Han Y, Cucurull L, Reynolds RW, Rutledge G, Goldberg M (2010) The NCEP Climate Forecast System Reanalysis. Bulletin of the American Meteorological Society 91(8):1015–1058, DOI 10.1175/2010BAMS3001.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Sahoo AK, Pan M, Troy TJ, Vinukollu RK, Sheffield J, Wood EF (2011) Reconciling the global terrestrial water budget using satellite remote sensing. Remote Sensing of Environment 115(8):1850–1865, DOI 10.1016/j.rse.2011.03.009

Salleh A (2016) Climate, Water, and Livelihood Skills: A Post-Development Reading of the SDGs. Globalizations 13(6):952–959, DOI 10.1080/14747731.2016.1173375, publisher: Routledge _eprint: https://doi.org/10.1080/14747731.2016.1173375

Saltikoff E, Kurri M, Leijnse H, Barbosa S, Stiansen K (2017) Maintenance Keeps Radars Running.

Bulletin of the American Meteorological Society 98(9):1833–1840, DOI 10.1175/BAMS-D-16-0095.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Salzmann M (2016) Global warming without global mean precipitation increase? Science Advances 2(6):e1501572, DOI 10.1126/sciadv.1501572, publisher: American Association for the Advancement of Science

Samaniego L, Kumar R, Attinger S (2010) Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. Water Resources Research 46(5), DOI 10.1029/2008WR007327, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2008WR007327

Samaniego L, Thober S, Wanders N, Pan M, Rakovec O, Sheffield J, Wood EF, Prudhomme C, Rees G, Houghton-Carr H, Fry M, Smith K, Watts G, Hisdal H, Estrela T, Buontempo C, Marx A, Kumar R (2019) Hydrological Forecasts and Projections for Improved Decision-Making in the Water Sector in Europe. Bulletin of the American Meteorological Society 100(12):2451–2472, DOI 10.1175/BAMS-D-17-0274.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Samset BH, Myhre G, Forster PM, Hodnebrog Andrews T, Boucher O, Faluvegi G, Fläschner D, Kasoar M, Kharin V, Kirkevåg A, Lamarque JF, Olivié D, Richardson TB, Shindell D, Takemura T, Voulgarakis A (2018) Weak hydrological sensitivity to temperature change over land, independent of climate forcing. npj Climate and Atmospheric Science 1(1):1–8, DOI 10.1038/s41612-017-0005-5, number: 1 Publisher: Nature Publishing Group

Santhi C, Srinivasan R, Arnold JG, Williams JR (2006) A modeling approach to evaluate the impacts of water quality management plans implemented in a watershed in Texas. Environmental Modelling & Software 21(8):1141–1157, DOI 10.1016/j.envsoft.2005.05.013

Scanlon BR, Healy RW, Cook PG (2002) Choosing appropriate techniques for quantifying groundwater recharge. Hydrogeology Journal 10(1):18–39, DOI 10.1007/s10040-001-0176-2

Schlesinger WH (2005) Biogeochemistry, vol 8. Elsevier

Schlosser CA, Houser PR (2007) Assessing a Satellite-Era Perspective of the Global Water Cycle. Journal of Climate 20(7):1316–1338, DOI 10.1175/JCLI4057.1, publisher: American Meteorological Society Section: Journal of Climate

Schmidt W (1915) Strahlung und Verdunstung an freien Wasserflächen; ein Beitrag zum Wärmehaushalt des Weltmeers und zum Wasserhaushalt der Erde. Ann Calender Hydrographie und Maritimen Meteorologie 43:111–124

Schmitt RW (1995) The ocean component of the global water cycle. Reviews of Geophysics 33(S2):1395–1409, DOI 10.1029/95RG00184, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/95RG00184

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese M (2011) GPCC full data reanalysis version 6.0 at 0.5: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. GPCC Data Rep, doi 10

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M, Rudolf B (2014) GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theoretical and Applied Climatology 115(1):15–40, DOI 10.1007/s00704-013-0860-x

Schneider U, Finger P, Meyer-Christoffer A, Rustemeier E, Ziese M, Becker A (2017) Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC). Atmosphere 8(3):52, DOI 10.3390/atmos8030052, number: 3 Publisher: Multidisciplinary Digital Publishing Institute

Schulzweida U (2022) CDO User Guide. Zenodo DOI 10.5281/zenodo.7112925, publisher: Zenodo

Seager R, Naik N, Vecchi GA (2010) Thermodynamic and Dynamic Mechanisms for Large-Scale Changes in the Hydrological Cycle in Response to Global Warming. Journal of Climate 23(17):4651–4668, DOI 10.1175/2010JCLI3655.1, publisher: American Meteorological Society Section: Journal of Climate

Sheffield J, Wood EF (2007) Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. Journal of Geophysical Research: Atmospheres 112(D17), DOI 10.1029/2006JD008288, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2006JD008288

Sheffield J, Goteti G, Wood EF (2006) Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. Journal of Climate 19(13):3088–3111, DOI 10.1175/JCLI3790.1, publisher: American Meteorological Society Section: Journal of Climate

Sheffield J, Ferguson CR, Troy TJ, Wood EF, McCabe MF (2009) Closing the terrestrial water budget from satellite remote sensing. Geophysical Research Letters 36(7), DOI 10.1029/2009GL037338, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2009GL037338

Sheffield J, Wood EF, Munoz-Arriola F (2010) Long-Term Regional Estimates of Evapotranspiration for Mexico Based on Downscaled ISCCP Data. Journal of Hydrometeorology 11(2):253–275, DOI 10.1175/2009JHM1176.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Sheffield J, Wood EF, Pan M, Beck H, Coccia G, Serrat-Capdevila A, Verbist K (2018) Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions. Water Resources Research 54(12):9724–9758, DOI 10.1029/2017WR022437, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2017WR022437

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. In: Proceedings of the 1968 23rd ACM national conference on -, ACM Press, Not Known, pp 517–524, DOI 10.1145/800186.810616

Sherwood S, Fu Q (2014) A Drier Future? Science 343(6172):737–739, DOI 10.1126/science.1247620, publisher: American Association for the Advancement of Science

Shiklomanov IA (1998) World Water Resources: A new appraisal and assessment for the 21st century. UNESCO

Shuttleworth WJ, Wallace JS (1985) Evaporation from sparse crops-an energy combination theory. Quarterly Journal of the Royal Meteorological Society 111(469):839–855, DOI 10.1002/qj.49711146910, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49711146910

Simmons A (2006) ERA-Interim: New ECMWF reanalysis products from 1989 onwards. ECMWF newsletter 110:25–36

Simolo C, Brunetti M, Maugeri M, Nanni T (2010) Improving estimation of missing values in daily precipitation series by a probability density function-preserving approach. International Journal of Climatology 30(10):1564–1576, DOI 10.1002/joc.1992, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.1992

Singh GG, Hilmi N, Bernhardt JR, Cisneros Montemayor AM, Cashion M, Ota Y, Acar S, Brown JM, Cottrell R, Djoundourian S, González-Espinosa PC, Lam V, Marshall N, Neu-

mann B, Pascal N, Reygondeau G, Rocklv J, Safa A, Virto LR, Cheung W (2019) Climate impacts on the ocean are making the Sustainable Development Goals a moving target travelling away from us. People and Nature 1(3):317–330, DOI 10.1002/pan3.26, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pan3.26

Singh P (2020) Linux Development on WSL. In: Singh P (ed) Learn Windows Subsystem for Linux: A Practical Guide for Developers and IT Professionals, Apress, Berkeley, CA, pp 131–168, DOI 10.1007/978-1-4842-6038-8_8

Skliris N, Zika JD, Nurser G, Josey SA, Marsh R (2016) Global water cycle amplifying at less than the Clausius-Clapeyron rate. Scientific Reports 6(1):38752, DOI 10.1038/srep38752, number: 1 Publisher: Nature Publishing Group

Slater LJ, Thirel G, Harrigan S, Delaigue O, Hurley A, Khouakhi A, Prosdocimi I, Vitolo C, Smith K (2019) Using R in hydrology: a review of recent developments and future directions. Hydrology and Earth System Sciences 23(7):2939–2963, DOI 10.5194/hess-23-2939-2019, publisher: Copernicus GmbH

Slivinski LC, Compo GP, Whitaker JS, Sardeshmukh PD, Giese BS, McColl C, Allan R, Yin X, Vose R, Titchner H, Kennedy J, Spencer LJ, Ashcroft L, Brönnimann S, Brunet M, Camuffo D, Cornes R, Cram TA, Crouthamel R, Domínguez-Castro F, Freeman JE, Gergis J, Hawkins E, Jones PD, Jourdain S, Kaplan A, Kubota H, Blancq FL, Lee TC, Lorrey A, Luterbacher J, Maugeri M, Mock CJ, Moore GK, Przybylak R, Pudmenzky C, Reason C, Slonosky VC, Smith CA, Tinz B, Trewin B, Valente MA, Wang XL, Wilkinson C, Wood K, Wyszyński P (2019) Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. Quarterly Journal of the Royal Meteorological Society 145(724):2876–2908, DOI 10.1002/qj.3598, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3598

Slivinski LC, Compo GP, Sardeshmukh PD, Whitaker JS, McColl C, Allan RJ, Brohan P, Yin X, Smith CA, Spencer LJ, Vose RS, Rohrer M, Conroy RP, Schuster DC, Kennedy JJ, Ashcroft L, Brönnimann S, Brunet M, Camuffo D, Cornes R, Cram TA, Domínguez-Castro F, Freeman JE, Gergis J, Hawkins E, Jones PD, Kubota H, Lee TC, Lorrey AM, Luterbacher J, Mock CJ, Przybylak RK, Pudmenzky C, Slonosky VC, Tinz B, Trewin B, Wang XL, Wilkinson C, Wood K, Wyszyński P (2021) An Evaluation of the Performance of the Twentieth Century Reanalysis Version 3. Journal of Climate 34(4):1417–1438, DOI 10.1175/JCLI-D-20-0505.1, publisher: American Meteorological Society Section: Journal of Climate

Sood A, Smakhtin V (2015) Global hydrological models: a review. Hydrological Sciences Journal 60(4):549–565, DOI 10.1080/02626667.2014.950580, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02626667.2014.950580

Speidel D, Agnew A (1982) The natural geochemistry of our environment. Westview Press p 16

Starr VP, Peixoto JP (1958) On the Global Balance of Water Vapor and the Hydrology of Deserts. Tellus 10(2):188–194, DOI 10.3402/tellusa.v10i2.9237, publisher: Taylor & Francis _eprint: https://doi.org/10.3402/tellusa.v10i2.9237

Steen LA (ed) (1990) On the Shoulders of Giants: New Approaches to Numeracy. The National Academies Press, Washington, DC, DOI 10.17226/1532

Stewart RE, Leighton HG, Marsh P, Moore GWK, Ritchie H, Rouse WR, Soulis ED, Strong GS, Crawford RW, Kochtubajda B (1998) The Mackenzie GEWEX Study: The Water and Energy Cycles of a Major North American River Basin. Bulletin of the American Meteorological Society 79(12):2665–2684, DOI 10.1175/1520-0477(1998)079⟨2665:TMGSTW⟩2.0.CO;2, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Stommel H, Stommel E (1979) The year without a summer. Scientific American 240(6):176–187, publisher: JSTOR

Sun Q, Miao C, Duan Q, Ashouri H, Sorooshian S, Hsu KL (2018) A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. Reviews of Geophysics 56(1):79–107, DOI 10.1002/2017RG000574, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017RG000574

Svoboda V, Hanel M, Máca P, Kyselý J (2016) Projected changes of rainfall event characteristics for the Czech Republic. Journal of Hydrology and Hydromechanics 64(4):415–425, DOI 10.1515/johh-2016-0036

Swenson S, Wahr J (2006) Estimating Large-Scale Precipitation Minus Evapotranspiration from GRACE Satellite Gravity Measurements. Journal of Hydrometeorology 7(2):252–270, DOI 10.1175/JHM478.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Syed TH, Famiglietti JS, Chambers DP, Willis JK, Hilburn K (2010) Satellite-based global-ocean mass balance estimates of interannual variability and emerging trends in continental freshwater

discharge. Proceedings of the National Academy of Sciences 107(42):17916–17921, DOI 10.1073/pnas.1003292107, publisher: Proceedings of the National Academy of Sciences

Takata K, Emori S, Watanabe T (2003) Development of the minimal advanced treatments of surface interaction and runoff. Global and Planetary Change 38(1):209–222, DOI 10.1016/S0921-8181(03)00030-4

Tang G, Clark MP, Papalexiou SM (2022) EM-Earth: The Ensemble Meteorological Dataset for Planet Earth. Bulletin of the American Meteorological Society 103(4):E996–E1018, DOI 10.1175/BAMS-D-21-0106.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Tapiador FJ, Navarro A, Moreno R, Sánchez JL, García-Ortega E (2020) Regional climate models: 30 years of dynamical downscaling. Atmospheric Research 235:104785, DOI 10.1016/j.atmosres.2019.104785

Tapley BD, Bettadpur S, Ries JC, Thompson PF, Watkins MM (2004) GRACE Measurements of Mass Variability in the Earth System. Science 305(5683):503–505, DOI 10.1126/science.1099192, publisher: American Association for the Advancement of Science

Tennant W (2004) Considerations when using pre-1979 NCEP/NCAR reanalyses in the southern hemisphere. Geophysical Research Letters 31(11), DOI 10.1029/2004GL019751, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2004GL019751

Ternes T, Joss A, Oehlmann J (2015) Occurrence, fate, removal and assessment of emerging contaminants in water in the water cycle (from wastewater to drinking water). Water Research 72:1–2, DOI 10.1016/j.watres.2015.02.055

Thackeray CW, DeAngelis AM, Hall A, Swain DL, Qu X (2018) On the Connection Between Global Hydrologic Sensitivity and Regional Wet Extremes. Geophysical Research Letters 45(20):11,343–11,351, DOI 10.1029/2018GL079698, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018GL079698

Thornthwaite CW (1948) An Approach toward a Rational Classification of Climate. Geographical Review 38(1):55–94, DOI 10.2307/210739, publisher: [American Geographical Society, Wiley]

Tolasz R, Míková T, Valeriánová A, Voženílek V (2007) Climate atlas of Czechia. Czech Hydrometeorological Institute, Prague 256

Trenberth KE (1998) Atmospheric Moisture Residence Times and Cycling: Implications for Rainfall Rates and Climate Change. Climatic Change 39(4):667–694, DOI 10.1023/A:1005319109110

Trenberth KE, Fasullo JT (2013a) North American water and energy cycles. Geophysical Research Letters 40(2):365–369, DOI 10.1002/grl.50107, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/grl.50107

Trenberth KE, Fasullo JT (2013b) Regional Energy and Water Cycles: Transports from Ocean to Land. Journal of Climate 26(20):7837–7851, DOI 10.1175/JCLI-D-13-00008.1, publisher: American Meteorological Society Section: Journal of Climate

Trenberth KE, Guillemot CJ (1995) Evaluation of the Global Atmospheric Moisture Budget as Seen from Analyses. Journal of Climate 8(9):2255–2272, DOI 10.1175/1520-0442(1995)008⟨2255:EOTGAM⟩2.0.CO;2, publisher: American Meteorological Society Section: Journal of Climate

Trenberth KE, Guillemot CJ (1998) Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses. Climate Dynamics 14(3):213–231, DOI 10.1007/s003820050219

Trenberth KE, Zhang Y (2018) How Often Does It Really Rain? Bulletin of the American Meteorological Society 99(2):289–298, DOI 10.1175/BAMS-D-17-0107.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Trenberth KE, Dai A, Rasmussen RM, Parsons DB (2003) The Changing Character of Precipitation. Bulletin of the American Meteorological Society 84(9):1205–1218, DOI 10.1175/BAMS-84-9-1205, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Trenberth KE, Smith L, Qian T, Dai A, Fasullo J (2007) Estimates of the Global Water Budget and Its Annual Cycle Using Observational and Model Data. Journal of Hydrometeorology 8(4):758–769, DOI 10.1175/JHM600.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Trenberth KE, Fasullo JT, Mackaro J (2011) Atmospheric Moisture Transports from Ocean to Land and Global Energy Flows in Reanalyses. Journal of Climate 24(18):4907–4924, DOI 10.1175/2011JCLI4171.1, publisher: American Meteorological Society Section: Journal of Climate

Trenberth KE, Zhang Y, Gehne M (2017) Intermittency in Precipitation: Duration, Frequency, Intensity, and Amounts Using Hourly Data. Journal of Hydrometeorology 18(5):1393–1412, DOI

10.1175/JHM-D-16-0263.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Trnka M, Balek J, Štěpánek P, Zahradníček P, Možný M, Eitzinger J, Žalud Z, Formayer H, Turňa M, Nejedlík P, Semerádová D, Hlavinka P, Brázdil R (2016) Drought trends over part of Central Europe between 1961 and 2014. Climate Research 70(2-3):143–160, DOI 10.3354/cr01420

Turk JT, Mostovoy GV, Anantharaj V (2010) The NRL-Blend High Resolution Precipitation Product and its Application to Land Surface Hydrology. In: Gebremichael M, Hossain F (eds) Satellite Rainfall Applications for Surface Hydrology, Springer Netherlands, Dordrecht, pp 85–104, DOI 10.1007/978-90-481-2915-7_6

United Nations (2022) World Population Prospects 2022: Summary of Results. Statistical Papers - United Nations (Ser. A), Population and Vital Statistics Report, United Nations, DOI 10.18356/9789210014380

Uppala SM, KÅllberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg LVD, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PaEM, Jenne R, Mcnally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 re-analysis. Quarterly Journal of the Royal Meteorological Society 131(612):2961–3012, DOI 10.1256/qj.04.176, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1256/qj.04.176

Valmassoi A, Keller JD, Kleist DT, English S, Ahrens B, Ďurán IB, Bauernschubert E, Bosilovich MG, Fujiwara M, Hersbach H, Lei L, Löhnert U, Mamnun N, Martin CR, Moore A, Niermann D, Ruiz JJ, Scheck L (2023) Current Challenges and Future Directions in Data Assimilation and Reanalysis. Bulletin of the American Meteorological Society 104(4):E756–E767, DOI 10.1175/BAMS-D-21-0331.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Van Rossum G, Drake Jr FL (1995) Python tutorial, vol 620. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands

VanDerLeeden F, Troise FL, Todd DK (1991) The water encyclopedia, 2nd edn. Geraghty & Miller ground-water series, Lewis, Chelsea, Mich

Vanella D, Longo-Minnolo G, Belfiore OR, Ramírez-Cuesta JM, Pappalardo S, Consoli S, D'Urso G, Chirico GB, Coppola A, Comegna A, Toscano A, Quarta R, Provenzano G, Ippolito M, Castagna A, Gandolfi C (2022) Comparing the use of ERA5 reanalysis dataset and ground-based agrometeorological data under different climates and topography in Italy. Journal of Hydrology: Regional Studies 42:101182, DOI 10.1016/j.ejrh.2022.101182

Vargas Godoy MR, Markonis Y (2023a) pRecipe: A global precipitation climatology toolbox and database. Environmental Modelling & Software 165:105711, DOI 10.1016/j.envsoft.2023.105711

Vargas Godoy MR, Markonis Y (2023b) Water cycle changes in reanalyses: a complementary framework. Scientific Reports 13(1):1–12, DOI 10.1038/s41598-023-31873-5, number: 1 Publisher: Nature Publishing Group

Vargas Godoy MR, Markonis Y, Hanel M, Kyselý J, Papalexiou SM (2021) The Global Water Cycle Budget: A Chronological Review. Surveys in Geophysics 42(5):1075–1107, DOI 10.1007/s10712-021-09652-6

Vargas Godoy MR, Markonis Y, Rakovec O, Jenicek M, Dutta R, Pradhan RK, Bešťáková Z, Kyselý J, Juras R, Papalexiou SM, Hanel M (2024) Water cycle changes in Czechia: a multi-source water budget perspective. Hydrology and Earth System Sciences 28(1):1–19, DOI 10.5194/hess-28-1-2024, publisher: Copernicus GmbH

Vecchi GA, Soden BJ, Wittenberg AT, Held IM, Leetmaa A, Harrison MJ (2006) Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. Nature 441(7089):73–76, DOI 10.1038/nature04744, number: 7089 Publisher: Nature Publishing Group

Vicente-Serrano SM, McVicar TR, Miralles DG, Yang Y, Tomas-Burguera M (2020) Unraveling the influence of atmospheric evaporative demand on drought and its response to climate change. WIREs Climate Change 11(2):e632, DOI 10.1002/wcc.632, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.632

Vicente-Serrano SM, Domínguez-Castro F, Reig F, Tomas-Burguera M, Peña-Angulo D, Latorre B, Beguería S, Rabanaque I, Noguera I, Lorenzo-Lacruz J, El Kenawy A (2023) A global drought monitoring system and dataset based on ERA5 reanalysis: A focus on crop-growing regions. Geoscience Data Journal 10(4):505–518, DOI 10.1002/gdj3.178, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/gdj3.178

Vinukollu RK, Meynadier R, Sheffield J, Wood EF (2011a) Multi-model, multi-sensor estimates of global evapotranspiration: climatology, uncertainties and trends. Hydrological Processes 25(26):3993–4010, DOI 10.1002/hyp.8393, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hyp.8393

Vinukollu RK, Wood EF, Ferguson CR, Fisher JB (2011b) Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches. Remote Sensing of Environment 115(3):801–823, DOI 10.1016/j.rse.2010.11.006

Vörösmarty CJ, Moore III B, Grace AL, Gildea MP, Melillo JM, Peterson BJ, Rastetter EB, Steudler PA (1989) Continental scale models of water balance and fluvial transport: An application to South America. Global Biogeochemical Cycles 3(3):241–265, DOI 10.1029/GB003i003p00241, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/GB003i003p00241

Wahr J, Molenaar M, Bryan F (1998) Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. Journal of Geophysical Research: Solid Earth 103(B12):30205–30229, DOI 10.1029/98JB02844, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/98JB02844

Walker D, Forsythe N, Parkin G, Gowing J (2016) Filling the observational void: Scientific value and quantitative validation of hydrometeorological data from a community-based monitoring programme. Journal of Hydrology 538:713–725, DOI 10.1016/j.jhydrol.2016.04.062

Wambua RM, Mutua BM, Raude JM (2016) Prediction of Missing Hydro-Meteorological Data Series Using Artificial Neural Networks (ANN) for Upper Tana River Basin, Kenya. American Journal of Water Resources 4(2):35–43, DOI 10.12691/ajwr-4-2-2, number: 2 Publisher: Science and Education Publishing

Wang G, Wang D, Trenberth KE, Erfanian A, Yu M, Bosilovich MG, Parr DT (2017) The peak structure and future changes of the relationships between extreme precipitation and temperature. Nature Climate Change 7(4):268–274, DOI 10.1038/nclimate3239, number: 4 Publisher: Nature Publishing Group

Wang K, Dickinson RE (2012) A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. Reviews of Geophysics 50(2), DOI 10.1029/2011RG000373, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2011RG000373

Wang-Erlandsson L, Bastiaanssen WGM, Gao H, Jägermeyr J, Senay GB, van Dijk AIJM, Guerschman JP, Keys PW, Gordon LJ, Savenije HHG (2016) Global root zone storage capacity from satellite-based evaporation. Hydrology and Earth System Sciences 20(4):1459–1481, DOI 10.5194/hess-20-1459-2016, publisher: Copernicus GmbH

Webb RS, Rosenzweig CE, Levine ER (1993) Specifying land surface characteristics in general circulation models: Soil profile data set and derived water-holding capacities. Global Biogeochemical Cycles 7(1):97–108, DOI 10.1029/92GB01822, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/92GB01822

Wehbe Y, Temimi M, Adler RF (2020) Enhancing Precipitation Estimates Through the Fusion of Weather Radar, Satellite Retrievals, and Surface Parameters. Remote Sensing 12(8):1342, DOI 10.3390/rs12081342, number: 8 Publisher: Multidisciplinary Digital Publishing Institute

Weiskel PK, Vogel RM, Steeves PA, Zarriello PJ, DeSimone LA, Iii KGR (2007) Water use regimes: Characterizing direct human interaction with hydrologic systems. Water Resources Research 43(4), DOI 10.1029/2006WR005062

Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, Woo K, Yutani H, Dunnington D, RStudio (2022) ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. URL `https://CRAN.R-project.org/package=ggplot2`

Wild M, Liepert B (2010) The Earth radiation balance as driver of the global hydrological cycle. Environmental Research Letters 5(2):025203, DOI 10.1088/1748-9326/5/2/025203

Williams DN, Doutriaux CM, Drach RS, McCoy RB (2009) The Flexible Climate Data Analysis Tools (CDAT) for Multi-model Climate Simulation Data. In: 2009 IEEE International Conference on Data Mining Workshops, pp 254–261, DOI 10.1109/ICDMW.2009.64, iSSN: 2375-9259

Willmott CJ, Matsuura K (2001) Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 1999). University of Delaware

Willmott CJ, Rowe CM, Mintz Y (1985) Climatology of the terrestrial seasonal water cycle. Journal of Climatology 5(6):589–606, DOI 10.1002/joc.3370050602, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.3370050602

Willmott CJ, Robeson SM, Feddema JJ (1994) Estimating continental and terrestrial precipitation averages from rain-gauge networks. International Journal of Climatology 14(4):403–414, DOI 10.1002/joc.3370140405, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.3370140405

Woodruff SD, Worley SJ, Lubker SJ, Ji Z, Eric Freeman J, Berry DI, Brohan P, Kent EC, Reynolds RW, Smith SR, Wilkinson C (2011) ICOADS Release 2.5: extensions and enhancements to the surface marine meteorological archive. International Journal of Climatology 31(7):951–967, DOI 10.1002/joc.2103, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.2103

Wundt W (1938) Das Bild des Wasserkreislaufs auf Grund früherer und neuer Forschungen. Reichs- und Preuß. Ministerium für Ernährung und Landwirtschaft, Landesanst ...

Wüst G (1922) Verdunstung und Niederschlag auf der Erde. Z Ges f Erdkunde Berlin

Wüst G, Defant A (1936) Schichtung und Zirkulation des atlantischen Ozeans. W. de Gruyter

Wüst G, Brogmus W, Noodt E (1954) Die zonale Verteilung von Salzgehalt, Niederschlag, Verdunstung, Temperatur und Dichte an der Oberfläche der Ozeane. Kieler Meeresforschungen 10(1954):2

Xiao M, Gao M, Vogel RM, Lettenmaier DP (2020) Runoff and Evapotranspiration Elasticities in the Western United States: Are They Consistent With Dooge's Complementary Relationship? Water Resources Research 56(8):e2019WR026719, DOI 10.1029/2019WR026719, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019WR026719

Xie P, Arkin PA (1997) Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs. Bulletin of the American Meteorological Society 78(11):2539–2558, DOI 10.1175/1520-0477(1997)078⟨2539:GPAYMA⟩2.0.CO;2, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Xie P, Chen M, Shi W (2010a) CPC unified gauge-based analysis of global daily precipitation. In: Preprints, 24th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc, vol 2

Xie P, Joyce R, Wu S, Yoo SH, Yarosh Y, Sun F, Lin R (2017) Reprocessed, Bias-Corrected CMORPH Global High-Resolution Precipitation Estimates from 1998. Journal of Hydrometeorology 18(6):1617–1641, DOI 10.1175/JHM-D-16-0168.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Xie SP, Deser C, Vecchi GA, Ma J, Teng H, Wittenberg AT (2010b) Global Warming Pattern Formation: Sea Surface Temperature and Rainfall. Journal of Climate 23(4):966–986, DOI 10.1175/2009JCLI3329.1, publisher: American Meteorological Society Section: Journal of Climate

Xue BL, Wang L, Li X, Yang K, Chen D, Sun L (2013) Evaluation of evapotranspiration estimates for two river basins on the Tibetan Plateau by a water balance method. Journal of Hydrology 492:290–297, DOI 10.1016/j.jhydrol.2013.04.005

Yasunari T (1994) Gewex-related asian monsoon experiment (game). Advances in Space Research 14(1):161–165, DOI 10.1016/0273-1177(94)90365-4

Yin X, Gruber A, Arkin P (2004) Comparison of the GPCP and CMAP Merged Gauge–Satellite Monthly Precipitation Products for the Period 1979–2001. Journal of Hydrometeorology 5(6):1207–1222, DOI 10.1175/JHM-392.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology

Young KC (1992) A Three-Way Model for Interpolating for Monthly Precipitation Values. Monthly Weather Review 120(11):2561–2569, DOI 10.1175/1520-0493(1992)120⟨2561:ATWMFI⟩2.0.CO;2, publisher: American Meteorological Society Section: Monthly Weather Review

Young NE, Anderson RS, Chignell SM, Vorster AG, Lawrence R, Evangelista PH (2017) A survival guide to Landsat preprocessing. Ecology 98(4):920–932, DOI 10.1002/ecy.1730, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ecy.1730

Zaitchik BF, Rodell M, Biasutti M, Seneviratne SI (2023) Wetting and drying trends under climate change. Nature Water 1(6):502–513, DOI 10.1038/s44221-023-00073-w, number: 6 Publisher: Nature Publishing Group

Zhang K, Kimball JS, Nemani RR, Running SW (2010) A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006. Water Resources Research 46(9), DOI 10.1029/2009WR008800, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2009WR008800

Zhang Y, Pan M, Wood EF (2016) On Creating Global Gridded Terrestrial Water Budget Estimates from Satellite Remote Sensing. Surveys in Geophysics 37(2):249–268, DOI 10.1007/s10712-015-9354-y

Zhang Y, Pan M, Sheffield J, Siemann AL, Fisher CK, Liang M, Beck HE, Wanders N, MacCracken RF, Houser PR, Zhou T, Lettenmaier DP, Pinker RT, Bytheway J, Kummerow CD, Wood EF (2018) A Climate Data Record (CDR) for the global terrestrial water budget: 1984–2010. Hydrology and Earth System Sciences 22(1):241–263, DOI 10.5194/hess-22-241-2018, publisher: Copernicus GmbH

Zhao L, Xia J, Xu Cy, Wang Z, Sobkowiak L, Long C (2013) Evapotranspiration estimation methods in hydrological models. Journal of Geographical Sciences 23(2):359–369, DOI 10.1007/s11442-013-1015-9

Zhao M, Golaz JC, Held IM, Ramaswamy V, Lin SJ, Ming Y, Ginoux P, Wyman B, Donner LJ, Paynter D, Guo H (2016) Uncertainty in Model Climate Sensitivity Traced to Representations of Cumulus Precipitation Microphysics. Journal of Climate 29(2):543–560, DOI 10.1175/JCLI-D-15-0191.1, publisher: American Meteorological Society Section: Journal of Climate

Čampulová M, Čampula R, Holešovský J (2022) An R package for identification of outliers in environmental time series data. Environmental Modelling & Software 155:105435, DOI 10.1016/j.envsoft.2022.105435

Řehoř J, Brázdil R, Lhotka O, Trnka M, Balek J, Štěpánek P, Zahradníček P (2021) Precipitation in the Czech Republic in Light of Subjective and Objective Classifications of Circulation Types. Atmosphere 12(11):1536, DOI 10.3390/atmos12111536, number: 11 Publisher: Multidisciplinary Digital Publishing Institute

Appendix A

Other Peer Review Publications During the Ph.D.

A.1 Shayeghi A, Rahmati Ziveh A, Bakhtar A, Teymoori J, Hanel M, **Vargas Godoy MR**, Markonis Y, AghaKouchak A. Assessing drought impacts on groundwater and agriculture in Iran using high-resolution precipitation and evapotranspiration products (2024). Journal of Hydrology. In press.

A.2 Beštáková Z, Strnad F, **Vargas Godoy MR**, Singh U, Markonis Y, Hanel M, Máca P, Kyselý J (2023) Changes of the aridity index in Europe from 1950 to 2019. Theoretical and Applied Climatology 151(1):587–601. https://doi.org/10.1007/s00704-022-04266-3

A.3 Nasreen S, Součková M, **Vargas Godoy MR**, Singh U, Markonis Y, Kumar R, Rakovec O, Hanel M (2022) A 500-year annual runoff reconstruction for 14 selected European catchments. Earth System Science Data 14(9):4035–4056. https://doi.org/10.5194/essd-14-4035-2022

A.4 Pradhan RK, Markonis Y, **Vargas Godoy MR**, Villalba-Pradas A, Andreadis KM, Nikolopoulos EI, Papalexiou SM, Rahim A, Tapiador FJ, Hanel M (2022) Review of GPM IMERG performance: A global perspective. Remote Sensing of Environment 268:112754. https://doi.org/10.1016/j.rse.2021.112754

Theoretical and Applied Climatology (2023) 151:587–601
https://doi.org/10.1007/s00704-022-04266-3

ORIGINAL PAPER

Changes of the aridity index in Europe from 1950 to 2019

Zuzana Bešťáková[1,2] · Filip Strnad[1] · Mijael Rodrigo Vargas Godoy[1] · Ujjwal Singh[1] · Yannis Markonis[1] · Martin Hanel[1] · Petr Máca[1] · Jan Kyselý[1,2]

Received: 19 October 2021 / Accepted: 26 October 2022 / Published online: 29 November 2022

Abstract
The aridity index, also known as the Budyko index, describes spatiotemporal changes in the hydroclimatic system in the long-term perspective. Defined as the ratio between potential evapotranspiration and precipitation, it can be used to determine wet (humid) and dry (arid) regions. In this study, we evaluated the aridity index estimated in different temporal scales, investigated its spatial patterns, and highlighted the long-term changes in Europe using three gridded data sets (CRU, E–OBS, and ERA5). A significant dry region expansion is evident in all data sets since the late 1980s. The extent of the dry regions has increased in Western, Central, and Eastern Europe, especially at low and medium altitudes. The results show the long-term development of the European hydroclimatic system and which areas have changed from wet to dry.

Keywords Aridity index · Budyko curve · Potential evapotranspiration · Precipitation · Drought · Europe

1 Introduction

Drought affects millions of people worldwide each year (Dai 2011). In particular, it has become increasingly widespread in Europe in the recent decades (Markonis et al. 2021; Moravec et al. 2021) and has been causing more and more problems in various socio-economic sectors like agriculture, water resources, and industry (Naumann et al. 2021). Droughts are expected to be more frequent, severe, and prolonged in the future (Ault 2020; Rakovec et al. 2022). Severe droughts significantly impact on the hydroclimatic system (Wilhite 2000; Keyantash and Dracup 2004; Van Loon 2015), so it is crucial to investigate droughts based on the long-term behavior of the hydroclimatic system and seek solutions and mitigation strategies.

The aridity index describes the long-term functioning of the atmosphere, more specifically, the process of receiving and releasing water from the underlying surface hydrological system. We study the potential flow of water to the atmosphere, assuming that we have an unlimited water supply (Budyko 1974; Wang and Alimohammadi 2012; Blöschl et al. 2013; Creed et al. 2014). In an arid/dry environment (d), evapotranspiration prevails over precipitation, and vice versa in a humid/wet environment (w).

The aridity index is one of the primary inputs of the Budyko modeling framework (Budyko 1974; Arora 2002; Gerrits et al. 2009; Blöschl et al. 2013; Zhou et al. 2015; Carmona et al. 2016). Budyko (1974) defined the aridity index as the ratio of potential evapotranspiration (defined in Section 2.1) to precipitation (PET/P).

Many authors applied the aridity index to characterize long-term runoff or actual evapotranspiration (Arora 2002; Zheng et al. 2009; Wang and Alimohammadi 2012; Blöschl et al. 2013; Creed et al. 2014). Among them, Arora (2002) used the aridity index to obtain an analytic formula to estimate the change in runoff under annual changes in precipitation and available energy. Long-term actual evapotranspiration can be calculated based on the aridity index according to the Budyko hypothesis, which describes actual evapotranspiration as a function of the aridity index (Zheng et al. 2009). According to the United Nations Environment Programme (Barrow 1992) and other authors (Lioubimtseva et al. 2005; Diaz-Padilla et al. 2011; Spinoni et al. 2015; Huang et al. 2016; Zhao et al. 2019; Myronidis and Nikolaos 2021), the aridity index can also be defined as the ratio P/PET. In this study, the aridity index is defined as PET/P (Arora 2002; Gerrits et al. 2009; Nyman et al. 2014; Zhou et al. 2015; Carmona et al. 2016; Liu et al. 2019).

✉ Zuzana Bešťáková
bestakova@ufa.cas.cz; bestakova@fzp.czu.cz

[1] Faculty of Environmental Sciences, Czech University of Life Sciences, Prague, Czech Republic

[2] Institute of Atmospheric Physics, Czech Academy of Sciences, Prague, Czech Republic

Studies reporting aridity change on a global scale observed an increase in aridity index (AI) in Europe (Gao and Giorgi 2008; Spinoni et al. 2015; Huang et al. 2016). However, these analyses are performed at a very coarse scale, prohibiting the detection of any regional signal. Regional-scale studies show similar results, but most of them focus on small areas in southern or southeastern Europe (Paltineanu et al. 2007; Gao and Giorgi 2008; Salvati et al. 2013; Pravalie and Bandoc 2015; Spinoni et al. 2015; Huang et al. 2016; Cheval et al. 2017; Myronidis and Nikolaos 2021). There is general agreement on the development of dry areas among authors who study AI-based climate change. However, no study has dealt with the long-term development of AI in Central, Eastern and Western Europe. Other studies focus on different areas or global scales and do not provide regional details. We did not find any study dealing with Central (defined here as 45–55° latitude and 6–24° longitude), Eastern (east of 24° longitude), or Western Europe (west of 6° longitude) specifically.

2 Data and methods

2.1 Potential evapotranspiration

Potential evapotranspiration is defined as the amount of water transpired by a short crop layer of uniform height continuously shading the soil and having sufficient water in the soil (Peng et al. 2017). In contrast, reference evapotranspiration is crop-specific and used at the local scale, especially by agronomists (Allan et al. 1998; Oudin et al. 2005; Seiller and Anctil 2016; Peng et al. 2017; Kohli et al. 2020; Xiang et al. 2020). Potential evapotranspiration can be converted into reference evapotranspiration when multiplying it by the appropriate crop coefficient, which provides an estimate of crop water use. If land use data are unavailable (or not considered), this coefficient cannot be estimated (Seiller and Anctil 2016). Since we do not have land-use data available and focus on spatiotemporal changes in the hydroclimatic system in the long-term perspective, potential evapotranspiration is used to calculate the aridity index. The difference between the potential evapotranspiration and the reference crop evapotranspiration is discussed in detail by (Xiang et al. 2020).

In many studies, potential evapotranspiration and reference evapotranspiration are confused, and in such studies, they are also compared to each other (Winter et al. 1995; Xu and Singh 2000; Tabari 2010; Tegos et al. 2015; Poyen et al. 2016; Seiller and Anctil 2016). The problem was pointed out by Xiang et al. (2020). Xu and Singh (2000) and Xiang et al. (2020) divided the methods for calculating potential and reference evapotranspiration into four groups: mass transfer type, temperature-based type, radiation type, and combination type.

In general, the most often used equations are of the temperature-based type. The most widely used method of this type is the Thornthwaite equation, as it is simple and requires only temperature for calculation (Pereira and Pruitt 2004; Bautista et al. 2009; Chang et al. 2019), but has been found to underestimate reference evapotranspiration (Pereira and Pruitt 2004; Sentelhas et al. 2010; Lakatos et al. 2020; Xiang et al. 2020).

The radiation type equations can be viewed as simplified forms of the Penman equation (Gardelin and Lindstrom 1997; Xiang et al. 2020). Among the widely used radiation-based methods is the Turc equation, which shows good results in different studies, but its calculation requires a large number of variables (Xu et al. 2013; Xiang et al. 2020) that are often unavailable.

The combined type includes energy balance and aerodynamic aspects, with the velocity term and vapor pressure being its two basic terms. Penman and Monteith are the basic equations for the combined type, which were then linked (Penman-Monteith formula) and recommended by FAO as a standardized method (Xiang et al. 2020). However, it leads to better estimates only if the input variables are well measured or estimated (they are sensitive to the imprecision of the data (Weiland et al. 2012; Seiller and Anctil 2016)). Another problem is that the Penman-Monteith formula is an equation for calculating reference crop evapotranspiration and should not be used for potential evapotranspiration (Xiang et al. 2020). Also, Weiland et al. (2012) found better model performance using the simplest formulas versus the combination ones, thereby reducing the number of data-intensive input variables. For these reasons, we decided to use the Oudin formula to calculate potential evapotranspiration, which is defined as:

$$PET = \frac{0.408 R_e (\mathrm{TG} + 5)}{100}, \; if \; TG + 5 > 0$$
$$PET = 0 \; otherwise \tag{1}$$

where PET $[mm \cdot day^{-1}]$ is potential evapotranspiration, R_e $[MJ \cdot m^{-2} \cdot day^{-1}]$ is the solar radiation (this is top-of-atmosphere radiation, calculated only based on the time of year and geographical location—the influence of the atmosphere is not considered), and TG $[°C]$ is mean daily air temperature (Oudin et al. 2005).

We chose the calculation of potential evapotranspiration according to the Oudin formula (which also belongs to the radiation type equations) because it provides the most adequate PET input to rainfall-runoff models. Oudin et al. (2005) tried to identify the most relevant approach to calculating PET and concluded that formulas based only on temperature and radiation provide the best results. Lang

et al. (2017) also pointed to better results of such methods and Tegos et al. (2015) reported that the Oudin formula shows relatively good results in Europe. Another advantage is data simplicity, as the coverage of meteorological stations is not always sufficient, so it is not possible to use formulas with large data requirements. In addition unlike the extended Thornthwaite formula, the Oudin formula is applicable for average daily temperatures already from −5°C (Oudin et al. 2005).

2.2 Aridity index

In line with the original Budyko's concept, we use the following definition of the aridity index:

$$\mathrm{AI}_T = \frac{\mathrm{PET}_T}{P_T}, \tag{2}$$

where $AI_T[-]$ is the dimensionless aridity index, PET_T $[mm]$ is potential evapotranspiration, and P_T $[mm]$ is precipitation aggregated over interval T (Gerrits et al. 2009; Wang and Alimohammadi 2012; Creed et al. 2014; Zhou et al. 2015; Carmona et al. 2016).

2.3 Transitions, states, patterns

We study transitions between wet (w) and dry (d) hydroclimatic systems. Transitions for different aggregations ($T = 1, 10, 20, 30$ years) are defined based on the magnitude of the $AI_T(t-1)$ value in the previous period $t-1$ and the $AI_T(t)$ value in the current period t.

If the value of AI in the previous period is less than one and in the current period greater than one (transition wd), a wet grid cell shifts into the dry state in the successive periods. If the value of AI in the previous period is greater than one and in the current period less than one (transition dw), the dry state changes into the wet state. If the AI value in the previous and current period is less than one (pattern ww), a wet state is attributed to both periods for a given grid cell. If the value of AI in the previous and current periods is greater than one (pattern dd), a dry state is attributed to the given grid cell in both periods. When determining the relative frequencies of transitions, states, and patterns for a larger area, we weighted them by the area of the grid cells at a specific latitude.

2.4 Gridded data sets used for the analyses

This section presents the three gridded data sets applied in this work: Climatic Research Unit (CRU) TS version 4.04, E–OBS version 22.0e, and the European Centre for Medium-range Weather Forecasts (ECMWF) ERA5 reanalysis. The whole study deals with hydrological years starting on November 1 and ending on October 31; this definition is set so that there is no significant year-to-year transfer of precipitation in the snow cover.

E–OBS has a resolution of 0.25° × 0.25° and is available for 1920–2020. The daily precipitation and average daily temperature values come from the E–OBS 22.0e data set released in May 2020 (Cornes et al. 2018). The annual precipitation totals were calculated based on monthly values, and potential evapotranspiration was based on daily temperature values according to the Oudin formula (Oudin et al. 2005; Sarbu and Sebarchievici 2017; Wald 2018). Studies analyzing E–OBS (Hofstra et al. 2009; Herrera et al. 2012; Skok et al. 2016; Navarro et al. 2019) have reservations about the density of the station network or their inhomogeneous coverage. The highest density of stations is in Northern Europe, England, Central Europe, and northern Italy, while the network of stations is sparsest in Southern and Eastern Europe (Cornes et al. 2018; Navarro et al. 2019). The accuracy of E–OBS strongly depends on the density of the station network, and sparsely covered areas have less accurate data. E–OBS was evaluated as better than ERA5 in regions with dense data, while in areas with sparse data both data sets are at the same level (Bandhauer et al. 2022).

CRU is a data set with a resolution of 0.5° × 0.5°, produced at the University of East Anglia and available for 1901–2019. It is a monthly gridded observational data set, spatially interpolated. The interpolation is based on angular-distance weighting. Available PET in this data set is estimated using the Penman-Monteith formula (Harris et al. 2020), which includes the mean, maximum and minimum temperature, vapor pressure, cloud cover, and wind speed (Harris et al. 2014). The Penman-Monteith formula is in fact a reference crop evapotranspiration (mass transfer type; (Xiang et al. 2020)), so it was not used in our study; the Oudin formula was used instead to calculate the potential evapotranspiration.

ERA5 is the fifth generation of ECMWF reanalysis for the global climate and weather. The ERA5 is a grid data set with a resolution of 0.25° × 0.25°, available for 1950–2020. The data is obtained from the Climate Data Store entries (Hersbach et al. 2018; Bell et al. 2020) and applies the Hargreaves-Samani formula for potential evapotranspiration (Hargreaves and Samani 1985). The formula uses the average, maximum, and minimum temperature, evaporation equivalent, and empirical coefficient for the calculation (Rolle et al. 2021). The Hargreaves-Samani formula is also a reference crop evapotranspiration (temperature-based type; Xiang et al. (2020)), and its implementation in ERA5 contains a code error (transpiration is zero in areas without vegetation—coastal and dry regions). The Oudin formula was used in this study also for ERA5.

The primary parameters determined by the different grid models were used to calculate Oudin-type PET and then

compared to the PET calculated by ERA5 and CRU teams (Fig. 1). There is a significant underestimation of available PET in coastal areas in ERA5 and a slight underestimation in the UK and the Po Valley in CRU. On the contrary, the overestimation in ERA5 and CRU is evident mainly in Southern Europe (south of 45° latitude), especially in the Iberian Peninsula.

The spatial resolution of 0.5° × 0.5° of all data sets was used for computation and analysis (the raster package (Hijmans 2021) in R environment was used for aggregation in E–OBS and ERA5). Most of Europe (13° W – 30° E longitude, 30°–70° N latitude) is studied here. Furthermore, the stats package in R environment (Wickham 2016; R Core Team 2021) is used to create empirical quantile-based distribution functions and construct the probability density.

3 Results

3.1 Basic properties

The increase in the mean AI over the whole period is shown in Fig. 2. For two out of the three data sets, the mean value decreased from 1960 to 1969 and considerably increased thereafter. ERA5 increasing trend lags behind the other two starting only at the end of the 1970s.

The mean and standard deviation of AI in Europe for 10-year aggregations for each transition and pattern are shown in Table 1. A decrease in the mean of AI values was found in the *dd* pattern during 1960–1979 and a subsequent sharp increase in the 1980–1989 period in all three data sets (Table 1). A decrease and subsequent gradual increase

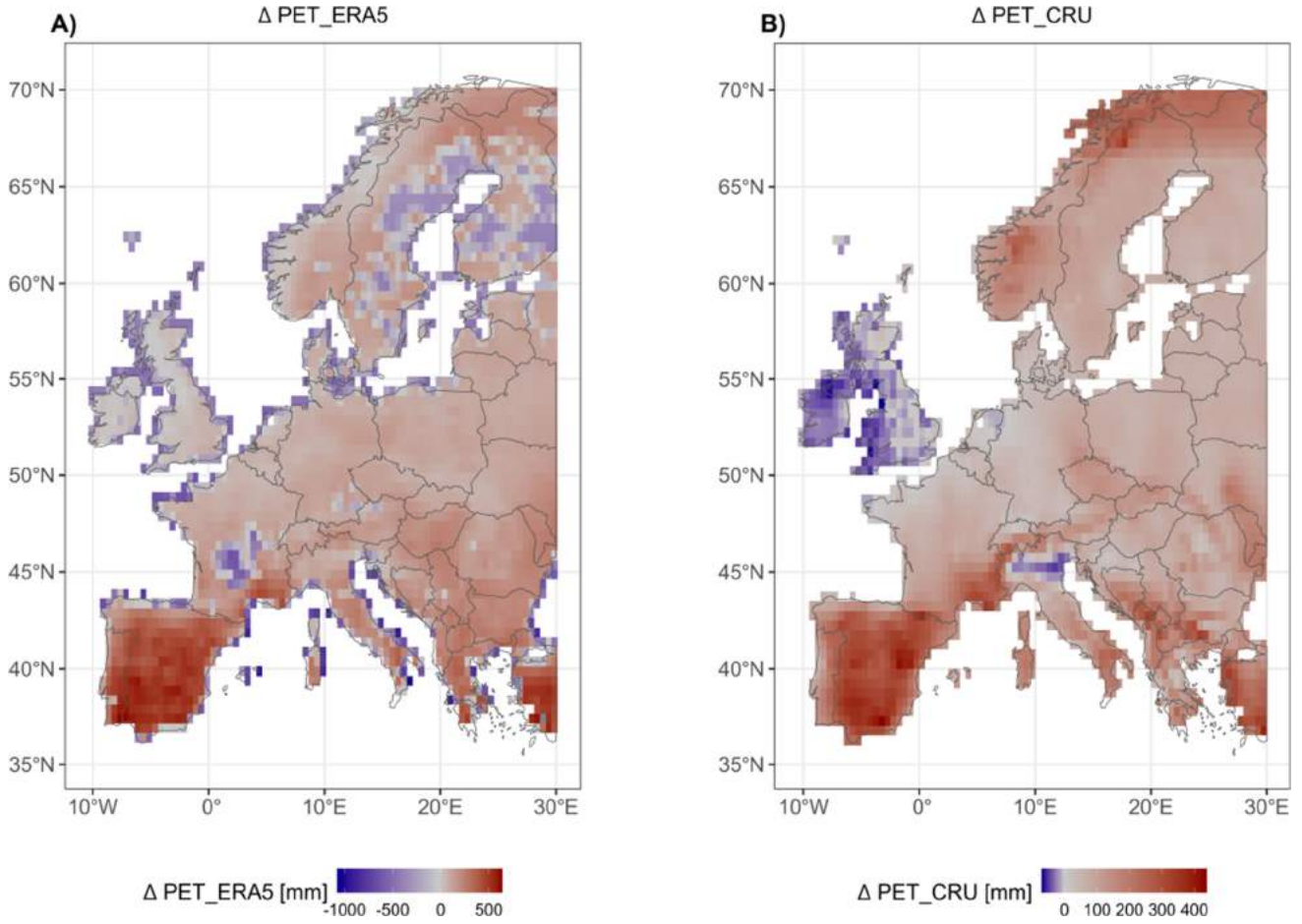


Fig. 1 Differences between mean annual values of A) PET calculated by ERA5 team (Hargreaves-Samani formula) and PET according to the Oudin formula (Δ PET_ERA5, left), B) PET calculated by CRU team (Penman-Monteith formula) and PET according to the Oudin formula (Δ PET_CRU, right), for 1950–2019

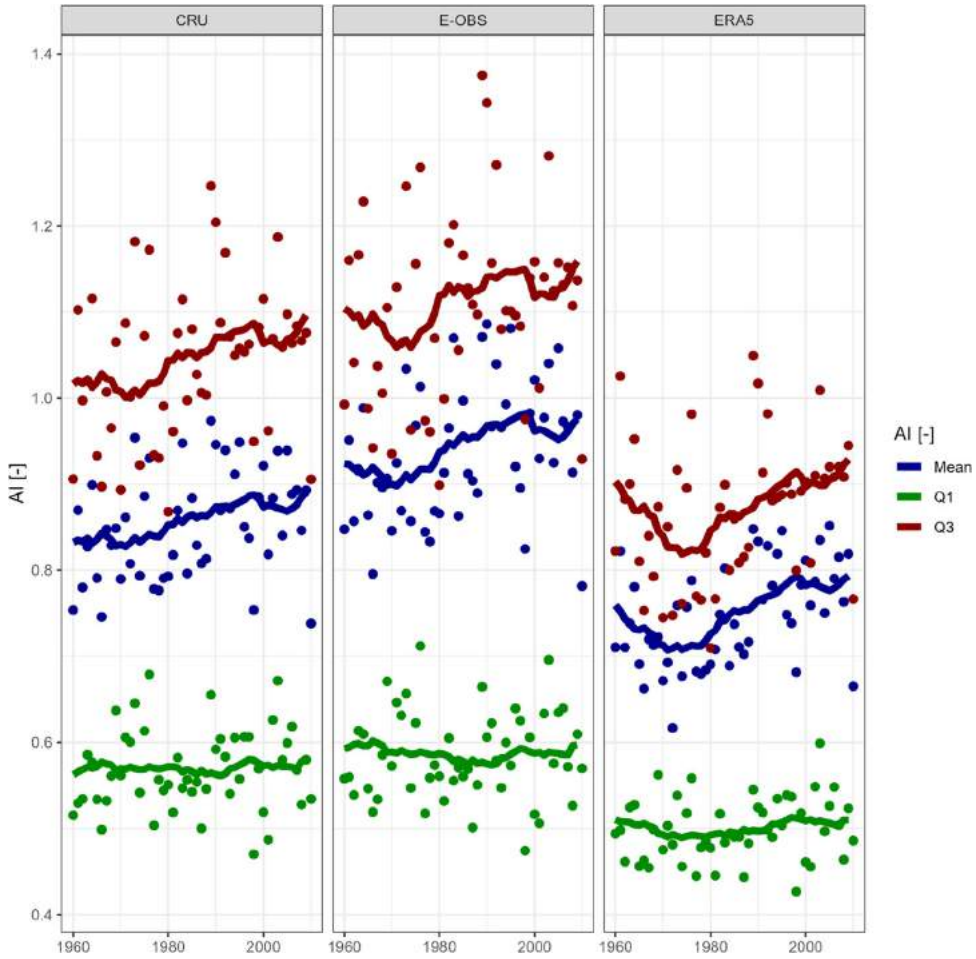


Fig. 2 20-year moving averages of annual mean, the lower quartile (Q1) and upper quartile (Q3) of AI values (solid lines) for all data sets for 1950–2019. Points show values in individual years; the lower and upper quartiles are estimated from the distribution across all grid boxes in the analysed area (Europe)

were also partly found in the dw and wd transitions in all data sets. Standard deviation increased during the 1990–1999 period for the pattern dd in E–OBS and ERA5, and the highest value for CRU was found in the 1980–1989 period. For the ww pattern, the changes in mean values are negligible, and only in ERA5 there is a slight decrease during the 1960–1989 period.

Empirical quantile-based distribution functions for all three data sets are compared in Fig. 3A. There is a considerable underestimation of AI values in ERA5 compared to the other two data sets. Figure 3B shows the different shapes of the probability density functions for each data set. If the AI is around 1, there is a transition from wet to dry and vice versa. The density is plotted on the y-axis.

3.2 Transitions between wet and dry regions

The average AI values (Fig. 4) classified similar proportions of wet and dry regions for the whole period investigated here in CRU and E–OBS, while ERA5 shows wet environments over most of the study area instead. In the CRU and E–OBS data wet regions occur all over Northern Europe (north of 55° latitude) and the UK, while dry regions in Southern Europe, the Pannonian Basin, Moldova, Ukraine, Poland, Czechia, the northeastern part of Germany, and the French Lowlands. In contrast, ERA5 shows wet conditions in the North-German Lowland, the Wielkopolska Lowland, Belarus and northern Ukraine. These additional wet conditions in ERA5 could be caused by the overestimation of precipitation (see Section 4).

Western, Central, and Eastern Europe form a transition strip (white color). The term “transition strip” is defined here as the central latitude strip of Europe (between 45°–55° latitude, and 5°–30° longitude). The most significant changes in the development of dry areas take place within this region.

We identified regions with the largest differences between consecutive mean AI values for 20-year periods (Fig. 5). During 1980–1999, there was a substantial increase in AI in Southern Europe, while in 2000–2019, drying took place mainly in the transition strip. Although the data sets share an overall similar signal for dry region development,

Table 1 Mean and standard deviation of mean values of AI for each transition (dd, ww) and pattern (dw, wd) in individual grid boxes for 10-year periods

10-year	Period	Mean_dd	sd_dd	Mean_dw	sd_dw	Mean_wd	sd_wd	Mean_ww	sd_ww
CRU	1950–1959	1.42	0.36	1.03	0.14	0.98	0.14	0.61	0.22
	1960–1969	1.36	0.33	0.94	0.13	0.95	0.12	0.60	0.23
	1970–1979	1.35	0.32	0.96	0.12	0.97	0.13	0.61	0.23
	1980–1989	1.48	0.43	1.01	0.11	0.98	0.15	0.61	0.23
	1990–1999	1.42	0.41	0.97	0.14	0.98	0.14	0.61	0.24
	2000–2009	1.44	0.40	0.98	0.13	0.97	0.12	0.61	0.23
	2010–2019	1.43	0.40	1.01	0.14	1.01	0.15	0.61	0.24
E-OBS	1950–1959	1.59	0.52	1.00	0.16	1.00	0.15	0.63	0.21
	1960–1969	1.48	0.49	0.95	0.15	0.96	0.13	0.61	0.22
	1970–1979	1.49	0.45	0.95	0.14	0.97	0.15	0.62	0.22
	1980–1989	1.64	0.57	1.01	0.12	0.99	0.16	0.61	0.22
	1990–1999	1.60	0.60	0.98	0.16	0.99	0.16	0.61	0.22
	2000–2009	1.60	0.58	0.98	0.17	0.97	0.16	0.61	0.22
	2010–2019	1.54	0.54	1.01	0.16	1.01	0.16	0.62	0.23
ERA5	1950–1959	1.51	0.46	0.98	0.17	0.98	0.16	0.60	0.22
	1960–1969	1.40	0.44	0.91	0.16	0.94	0.14	0.57	0.20
	1970–1979	1.37	0.33	0.95	0.17	0.97	0.17	0.57	0.20
	1980–1989	1.54	0.44	0.99	0.13	0.95	0.18	0.56	0.20
	1990–1999	1.51	0.47	0.96	0.16	0.98	0.16	0.58	0.21
	2000–2009	1.55	0.44	1.00	0.17	0.99	0.16	0.59	0.21
	2010–2019	1.49	0.47	0.97	0.15	0.97	0.15	0.58	0.21

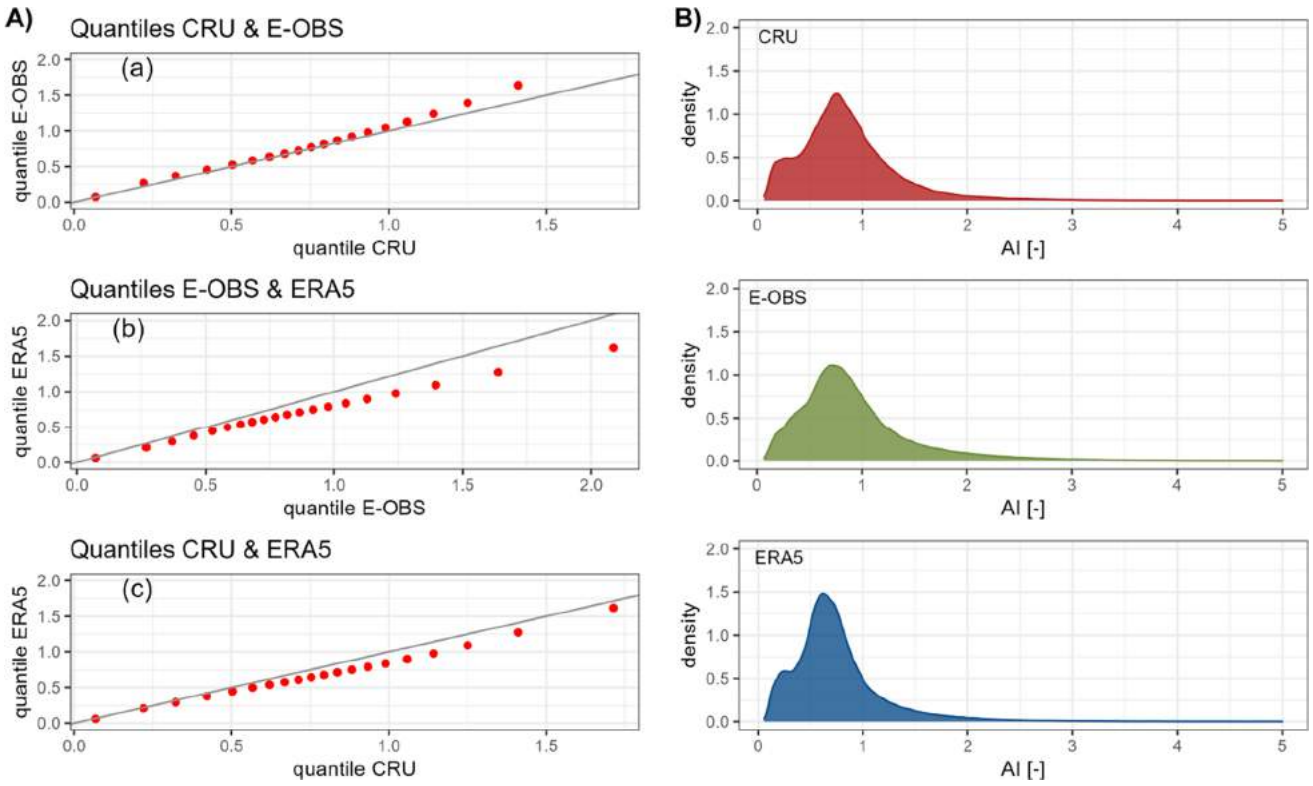


Fig. 3 **A** Quantile-quantile plots of annual AI values for 1950–2019: (a) CRU and E–OBS, (b) E–OBS and ERA5, (c) CRU and ERA5. **B** Probability density of annual AI values for 1950–2019. The x-axis shows the AI values and the y-axis the density

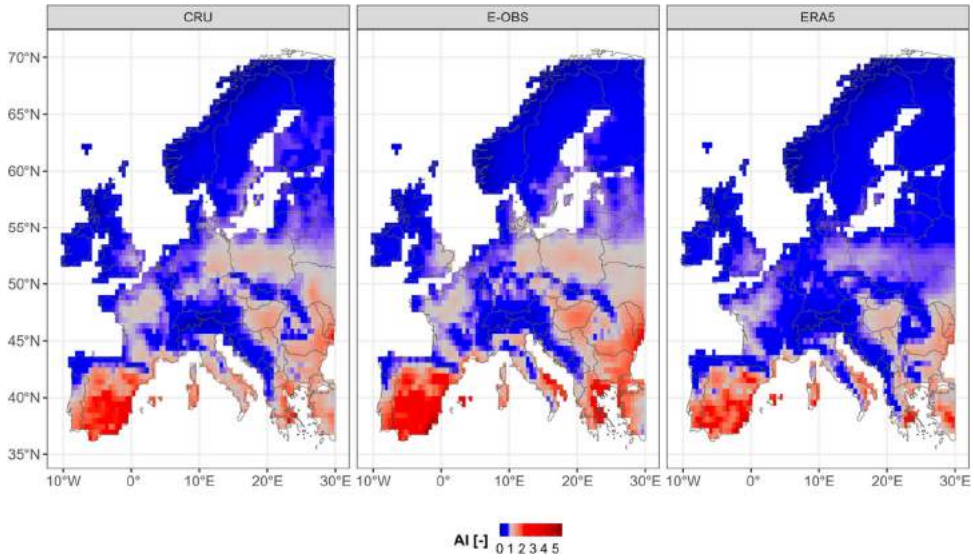


Fig. 4 Average AI values for the whole period (1950–2019) for each grid cell

the location of the hotspots differs; CRU in Western Europe, E–OBS in Central and Western Europe, and ERA5 in Eastern Europe (Figs. 5 and 7).

The 10-year moving average of annual AI values were compared with those of the previous 10-year periods. States (d, w), transitions (wd, dw), and patterns (ww, dd) were assigned to these AI values (Fig. 6). According to the three data sets' median, the probability of occurrence of dry regions has increased by approximately 7% over the last 50 years (Fig. 6). The most intense increase has been observed since the late 1980s (approximately 8%). The above is also reflected in the behavior of the dd and ww patterns. Over the last 50 years, there has been an increase in the dd pattern by approximately 3% and a decrease in ww by approximately 1%. The ww pattern was most frequent in the mid-1980s, followed by a considerable decrease, i.e., an increase in dry regions. The dw transitions decreased from 1970s until mid-1980s, then increased considerably from the mid to late 1980s. The incidence of the wd transition increased considerably from the mid-1980s to the beginning of the twenty-first century and then grew slightly. The dw transition has a declining trend throughout the whole period, and the wd transition has an increasing trend. The three data sets found similar trends in transitions, patterns, and states throughout the period. The median trend (black) almost overlaps with CRU.

After plotting the transitions and patterns (wd, dw, ww, dd) for 20-year periods, we found only negligible changes towards drying in Northern Europe and the southeast UK (Fig. 7).

In Southern Europe, there is a gradual expansion of the dry region. In all data sets, the east coast of the Adriatic Sea is included within a wet region, although it is a dry area in fact due to unique geomorphologic features of the Dinaric Mountains with subterranean rivers and streams. The relatively high precipitation amounts may be related to windward effects on the slopes of the Dinaric Mountains when Mediterranean cyclones influence the region.

A large area of Eastern Europe belongs to the dry area in the CRU and E–OBS, while in ERA5 a significant part of Eastern Europe belongs to the wet area. The spread of drought in Eastern Europe has taken place over the last two decades in CRU and E–OBS, especially in western Ukraine.

Western Europe has been drying up mainly in the French Lowlands; Eastern Europe mainly in northern Ukraine and southern Belarus; and Central Europe from the east, especially in the Pannonian Basin, the Wielkopolska Lowlands, eastern part of the North-German Lowlands, and Czechia (according to CRU and E–OBS).

Southern Europe, the Pannonian Basin, Moldova, Ukraine, Poland, Czechia, the northeastern part of Germany, and the French Lowlands are evaluated as dry areas of Europe according to CRU and E–OBS data sets.

No development of dry regions has taken place at high altitudes. The high mountains remain entirely in the wet region (Figs. 4 and 7), even though drying occurs in these parts of Europe (Fig. 5).

4 Discussion

Our study supports previous results concerning development of aridity in Europe and detects these changes, especially in the central latitude strip of Europe (between 45° and 55° latitude).

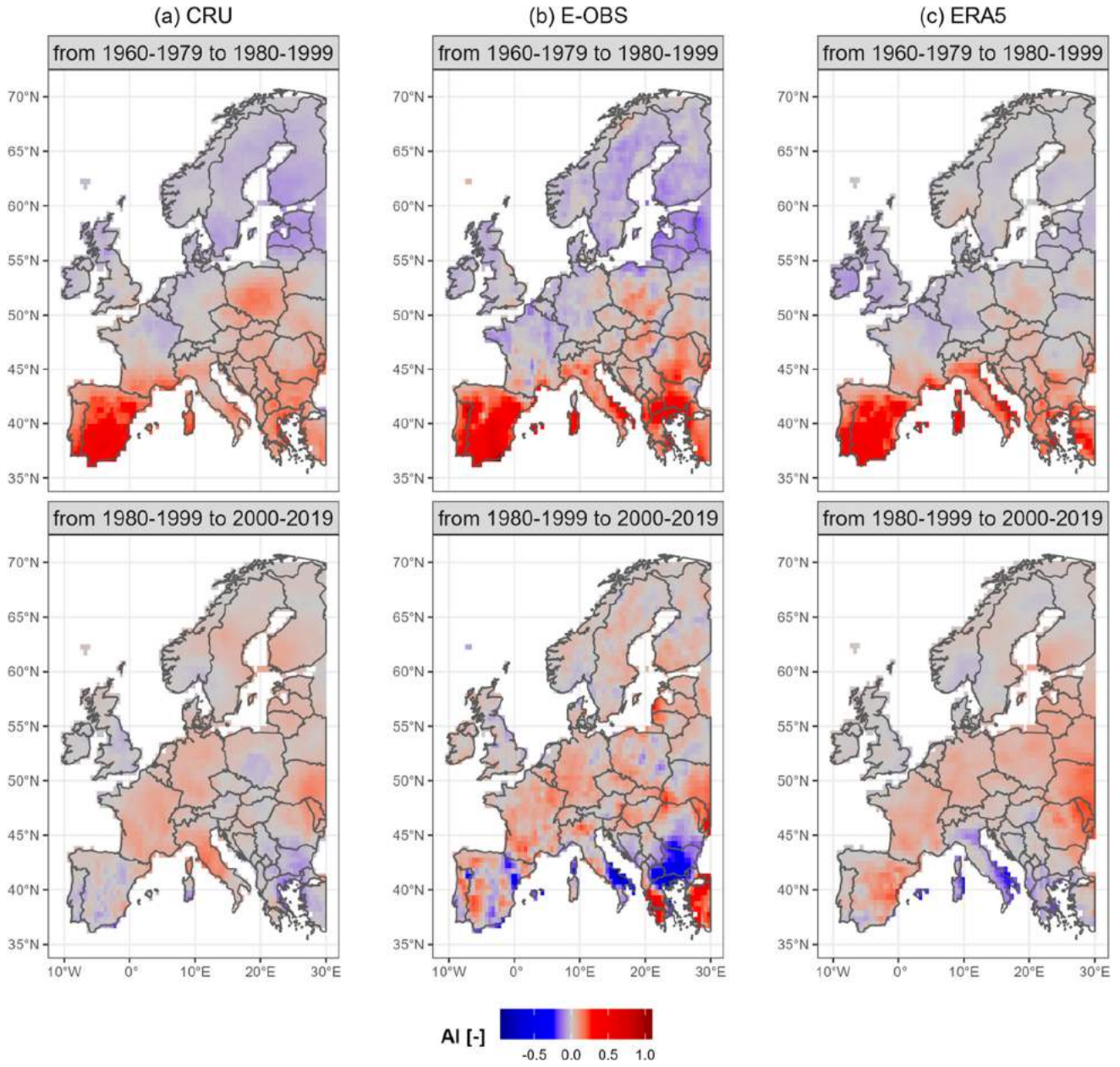


Fig. 5 Differences in mean AI values for the three data sets for consecutive 20-year periods from 1960 to 2019

4.1 Comparison of data sets

The study has identified specific differences between the three data sets related to the different methods for creating them.

Most studies that evaluated precipitation in E–OBS report underestimations in the data (Hofstra et al. 2009; Fibbi et al. 2016; Nechita et al. 2019; Bandhauer et al. 2022). Despite this Nechita et al. (2019) recommend E–OBS for the needs of the current climate data set. According to Laiti et al. (2018), E–OBS shows lower performance in the Alps compared to ADIGE (regional precipitation and temperature data set (Laiti et al. 2018)) and APGD (the Alpine Precipitation Grid Data set (Isotta et al. 2015)), especially in small river catchments.

Our results are consistent with van der Schrier et al. (2013), who compared CRU and E–OBS and concluded that CRU generally has higher temperature values than E–OBS, which is reflected in our study as higher AI values.

Quite few studies compared observational based data set (E–OBS) and reanalysis data set (ERA5) (Bandhauer et al. 2022; Hassler and Lauer 2021; Velikou et al. 2022).

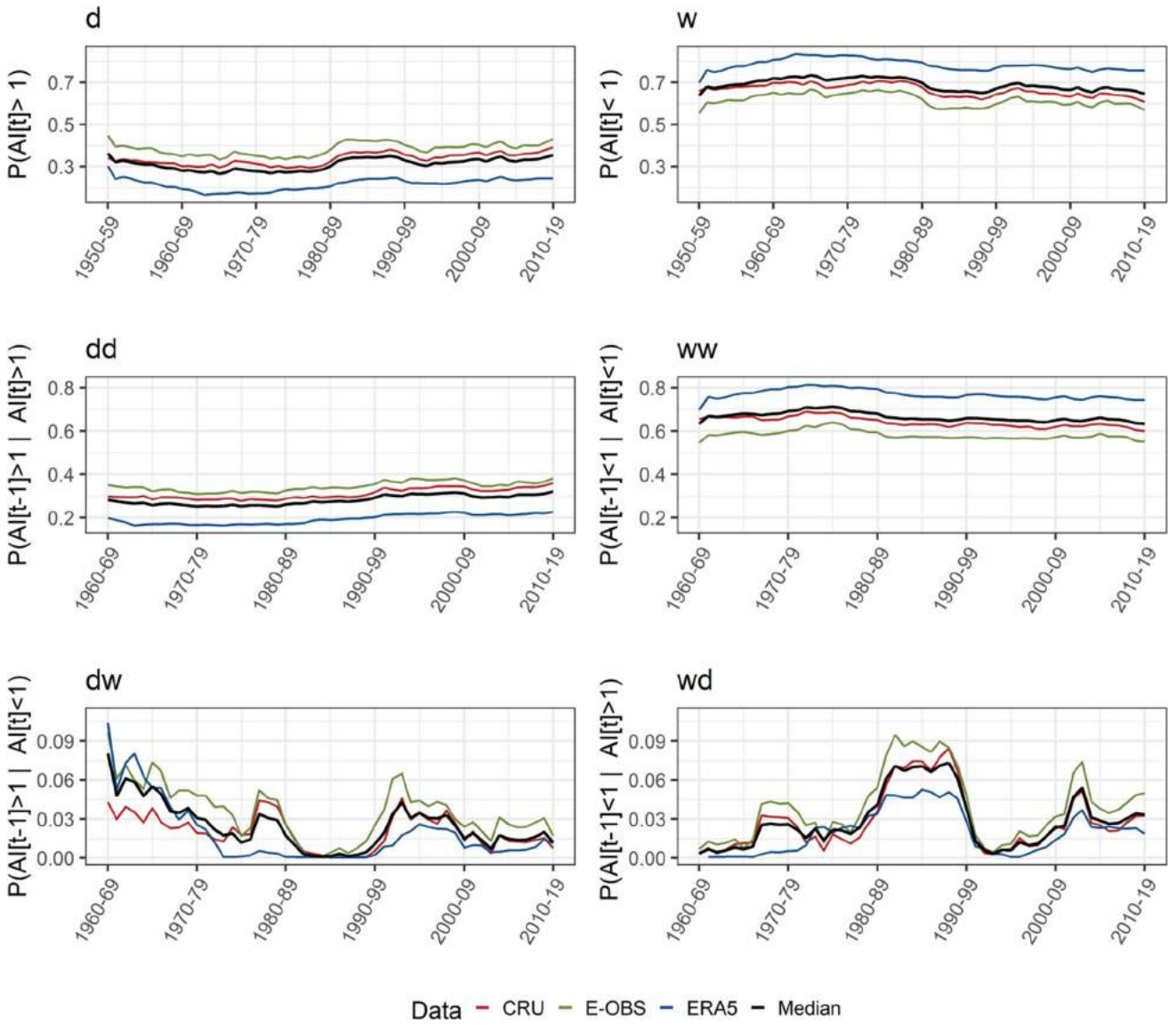


Fig. 6 10-year moving relative frequencies of states (d, w), patterns (dd, ww) and transitions (dw, wd) for 10-year moving periods per unit area for CRU, E–OBS and ERA5. The black line shows the median of all three data sets

Velikou et al. (2022) found underestimated temperatures in ERA5 compared to E–OBS mainly in areas with complex topography, especially in the Alps and the Mediterranean. Bandhauer et al. (2022) and Hassler and Lauer (2021) compared precipitation in E–OBS and ERA5 and concluded that ERA5 overestimated it in Europe (mainly in summer) due to far too many wet days. This is further supported by our results which show considerably wetter conditions in ERA5. We found a substantial difference focused in the area stretching from Central to Eastern Europe.

Bandhauer et al. (2022) emphasized that the underestimation of precipitation in E–OBS is substantially smaller than the overestimation in ERA5. Nevertheless they deemed both E–OBS and ERA5 as useful data sets for Central and Northern Europe.

Some authors use CRU as a reference (Tapiador and Sanchez 2008; Sanchez et al. 2009) and highlight its advantages: higher spatial resolution than other data sets of similar temporal extent, longer temporal coverage than other products of similar spatial resolution, encompassing a more extensive suite of surface climate variables than available elsewhere, and the construction method ensuring that strict temporal fidelity is maintained (Zveryaev 2004). On the contrary, some studies express reservations—especially about the density of the network of stations or their inhomogeneous coverage (Samani 2000; Zveryaev and

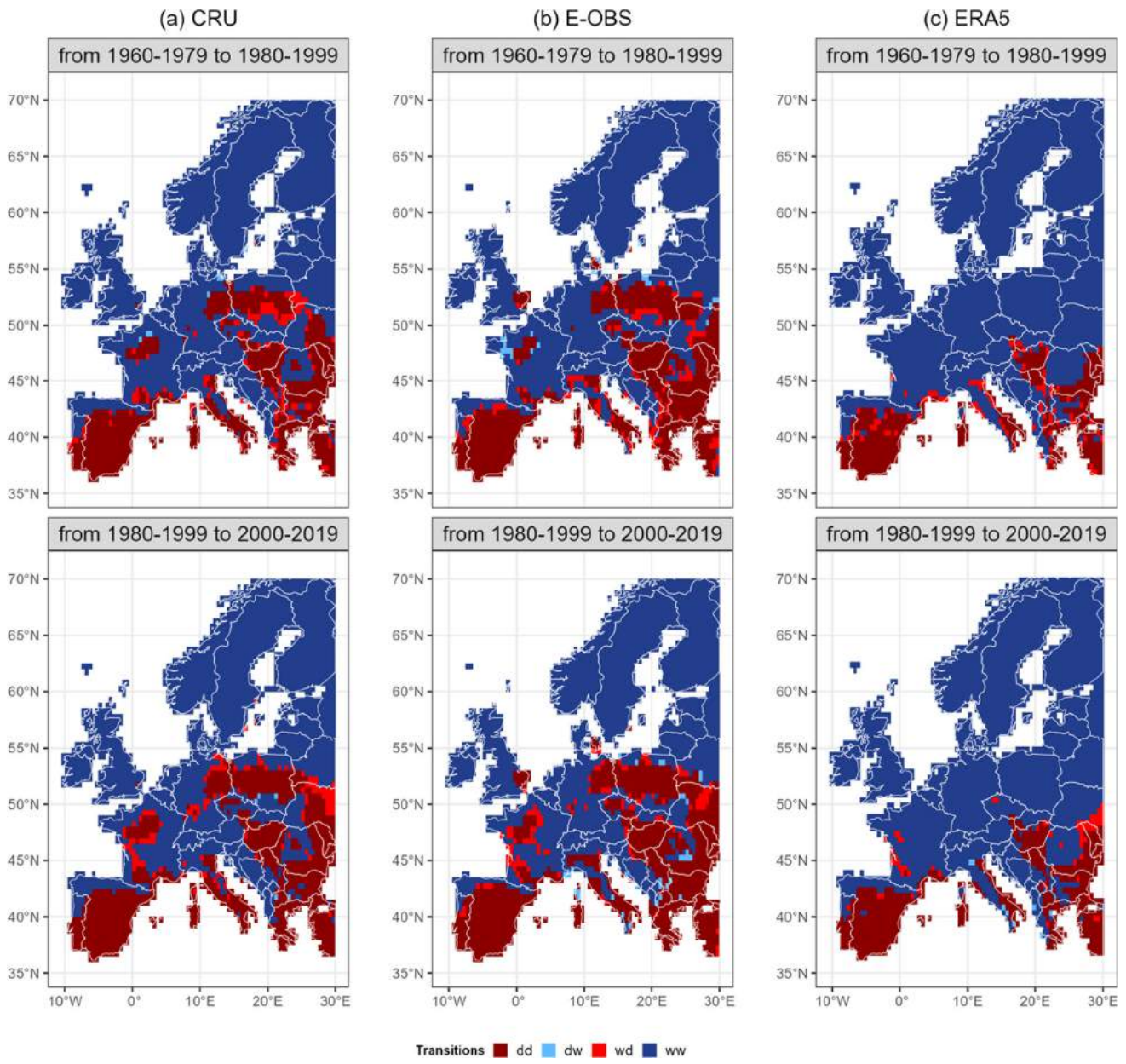


Fig. 7 Transitions (dw, wd), and patterns (dd, ww) for 20-year periods from 1960 to 2019 for the three data sets

Gulev 2009; Tegos et al. 2015; Duan et al. 2019). We do not consider the density of the network of stations to be a problem since the greatest development of drought has taken place in the central latitude strip, and the density of the network is sufficient in this area.

4.2 Temporal development of AI

According to our study, the aridity intensified mainly in the 1950s and from the mid-1980s onward. Unlike existing aridity studies, the results presented herein include the long-term perspective of assessing AI over Europe.

Our results are aligned with those reported by Huang et al. (2016) and Hanel et al. (2018) who identified extreme drought throughout Europe in the early 1950s, the ensuing decline or stagnation, and the renewed considerable increase in dry areas since the 1980s. Nechita et al. (2019) used the CRU, E–OBS, and ROCADA (Birsan and Dumitrescu 2014) data sets to study the Southern Carpathians over 1961–2013. They found wetter climate in the mid-1980s and an increase in extreme temperatures since 1986, manifested as an increase in wd and dd transitions and a decrease in dw and ww transitions in our results. Pan et al. (2021) examined the effect of increasing potential evapotranspiration and

decreasing precipitation on dry regions on the global scale over 1901–2017, and found a global increase in dry areas. Potential evapotranspiration has an increasing trend and weakens the role of precipitation. According to their study, the turning point when potential evapotranspiration exceeded precipitation on a global scale occurred in 1966.

The aridity developed fastest, especially in the period 2000–2019, in the central latitude strip of Europe, namely in northern Ukraine and southern Belarus area, Hungary, Wielkopolska Lowlands, Czechia, North-German Lowlands, and French Lowlands.

Most AI-based drought studies mainly focus on Southern Europe (Gao and Giorgi 2008; Salvati et al. 2013; Spinoni et al. 2015; Huang et al. 2016; Cheval et al. 2017; Myronidis and Nikolaos 2021), in which they report an increase in AI similarly to our study—dry areas are expanding, especially in the Balkan Peninsula and Italy, while drought is deepening in the Iberian Peninsula. Paltineanu et al. (2007) and Pravalie and Bandoc (2015) noticed drought development in Romania and similar results are manifested in our study; most of Romania belongs to the dry area and only the Carpathian region is included in the wet area. Moreover, we observe a more pronounced drought in the Pannonian Basin compared to Romania; a sharp increase in potential evapotranspiration in the Pannonian Basin, especially in its western part at low altitudes, was reported by Lakatos et al. (2020).

Pan et al. (2021) analyzed some areas of Romania as wet catchments, in contrast to our results. This discrepancy may be related to different type of averaging (across catchments vs. grids), a different range of borders between wet and dry catchments in inversely defined AI or a different version of CRU.

AI-based drought studies do not directly address other parts of Europe. We, on the other hand, examined the whole of Europe. No significant changes towards drying were found in Northern Europe. In Southern Europe, there is a gradual variable expansion of the dry region; the largest increase in AI was found during 1980–1999. The most significant changes in the development of dry areas occurred in the transition strip around 50° latitude, and the largest increase in AI was observed there during the period 2000–2019.

Over the past 50 years, the probability of occurrence of dry areas has increased by approximately 7% for the study area. The most intensive increase was observed from the end of the 1980s. The following areas in Europe are classified as dry: Southern Europe, the Pannonian Basin, Moldova, Ukraine, Poland, Czechia, the northeastern part of Germany, and the French Lowlands.

Concerning future scenarios, Gao and Giorgi (2008) estimated further development and an increase in the severity of drought to the north in the dry regions, especially in the central and southern parts of the Iberian, Balkan and Apennine Peninsulas, and on the main islands (Corsica, Sardinia, and Sicily), and then in the Pannonian Basin and Romania. Cheval et al. (2017) reported that the significant relocations to the dry area would occur in southeastern Italy, near the Black Sea, in the eastern part of the Balkan Peninsula, and in the Pannonian Basin. In all of these areas, we find an increase in AI. Furthermore, we observed a significant increase in drought especially in Central Europe, which has not been highlighted as a hot spot in the above-cited studies.

4.3 Drought linked with atmospheric circulation

Physical mechanisms affecting long-term changes in the hydro-climatic system and drought development include atmospheric circulation (e.g. Lhotka et al. (2020)). The dominant mode of climate variability in Europe, especially in the cold half-year, is the North Atlantic Oscillation (NAO). Changes in the intensity of the NAO and the location of its centers of action were reported by many authors (e.g. Kucerova et al. (2017)) and influenced temperature and precipitation trends in Europe over the past decades. They are also related to the shift of storm tracks to the north (Sfica et al. 2021) or northwest (Kucerova et al. 2017) which results in positive atmospheric pressure anomalies over Central Europe (Tomczyk et al. 2019), decrease of cloud cover and development of inland drought (Sfica et al. 2021). Kucerova et al. (2017), Lhotka et al. (2020), and Rehor et al. (2021) reported an increase in anticyclonic circulation types in Central and Eastern Europe, which supports the increase in AI taking place in the central latitude strip. The frequency changes between cyclonic/anticyclonic circulation types as well as between types associated with warm/cold advection and deficit/excess of precipitation are particularly important for droughts and heatwaves in warm half-year, and Lhotka et al. (2020) showed that circulation changes contributed to an increase in drought characteristics over Central Europe since the 1950s. It remains open and subject to follow-up investigation on whether similar changes also played a role in drying trends in other parts of Europe.

Huguenin et al. (2020) linked the record-breaking heatwaves and water shortages in recent years in Central Europe to anthropogenic warming and a weaker jet stream, which allowed the quasi-stationary and high-pressure system to persist for many days. As heatwaves are coupled to soil moisture conditions, these changes (including the

persistence of circulation patterns) will likely influence drought characteristics.

5 Conclusion

Our study examined dry and wet regions in Europe based on the aridity index in space and time over 1950–2019. The main goal was to identify regions where a shift from wet to dry conditions occurred, the period of increased transitions to dry regions, and to compare three widely used continental data sets (CRU, E–OBS, ERA5).

We chose the Oudin formula to calculate potential evapotranspiration for all data sets, because it is the most adequate PET input to rainfall-runoff models. The pronounced change from wet to dry regions since the late 1980s is clearly manifested in all data sets. The main results can be summarized as follows:

1. Significant development of dry regions has been observed in Western, Central, and Eastern Europe since the late 1980s.
2. From the late 1980s to the present, the extent of the dry regions has increased by approximately 7%.
3. There was a slight decrease in the dry-wet transition and a slight increase in the wet-dry transition from 1950 to 2019.
4. Throughout the study period, Northern Europe and the UK were classified as wet regions, while the Iberian Peninsula and the southern tip of the Balkan Peninsula as dry regions.
5. The development of dry areas was mainly found in France, Germany, Poland, Czechia, the Pannonian Basin, and the region between Belarus and Ukraine.

The results demonstrate the long-term development of dry regions since the late 1980s, mainly in Western and Central Europe in all data sets, and in Eastern Europe in CRU and E–OBS. Compared to CRU and E–OBS, ERA5 has low values of aridity index mainly in Eastern Europe, due to the overestimation of precipitation, and should be interpreted with caution.

Acknowledgements We acknowledge E–OBS from the EU-FP6 project UERRA, the Copernicus Climate Change Service and Climate Data Store, and the data providers in the ECA&D project, the Climatic Research Unit, and ECMWF.

Author contribution ZB, PM, JK: conceptualization, review, and editing. ZB: writing—original draft, data processing, visualization. FS: visualization review and editing. MV, US, YM, MH: review and editing.

Funding The study was supported by the Czech Science Foundation, project 20-28560S ("Driving mechanisms of extremes in reanalysis and climate models"). ZB was also supported within student project "Conditional probabilities of transition from arid to humid environment and vice versa in Europe during the period 1766–2015" by the Internal Grant Agency of the Faculty of Environmental Sciences.

Data availability The data sets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Code availability Not applicable.

Declarations

Ethics approval The authors have agreed for authorship, read and approved the manuscript, and given consent for submission and subsequent publication of the manuscript.

Consent to participate All authors consent to participate in the present study.

Consent for publication All authors read the manuscript and agree for the publication.

Conflict of interest The authors declare no competing interests.

References

Allan R, Pereira L, Smith M (1998) Crop evapotranspiration (guidelines for computing crop water requirements), FAO Irrigation and drainage paper 56. Food and Agriculture Organization

Arora VK (2002) The use of the aridity index to assess climate change effect on annual runoff. J Hydrol 265(1):164–177

Ault TR (2020) On the essentials of drought in a changing climate. Science 368(6488, SI):256–260

Bandhauer M, Isotta F, Lakatos M, Lussana C, Baserud L, Izsak B, Szentes O, Tveito OE, Frei C (2022) Evaluation of daily precipitation analyses in E-OBS (v19.0e) and ERA5 by comparison to regional high-resolution datasets in European regions. Int J Climatol 42(2):727–747

Barrow CJ (1992) World Atlas of Desertification (United Nations Environment Programme). Land Degradation & Development 3(4):249

Bautista F, Bautista D, Delgado-Carranza C (2009) Calibration of the equations of Hargreaves and Thornthwaite to estimate the potential evapotranspiration in semi-arid and subhumid tropical climates for regional applications. Atmosfera 22(4):331–348

Bell B, Hersbach H, Berrisford P, Dahlgren P, Horányi A, Muñoz Sabater J, Nicolas J, Radu R, Schepers D, Simmons A, Soci C, Thépaut JN (2020) ERA5 monthly averaged data on single levels from 1950 to 1978 (preliminary version). Copernicus Climate Change Service (C3S) Climate Data Store (CDS)

Birsan MV, Dumitrescu A (2014) ROCADA: Romanian daily gridded climatic dataset (1961-2013) V1.0

Blöschl G, Sivapalan M, Wagener T, Viglione A, Savenije H (2013) Runoff prediction in ungauged basins (synthesis across processes, places and scales). Cambridge University Press, United Kingdom

Budyko MI (1974) Climate and Life. Academic Press, London

Carmona AM, Poveda G, Sivapalan M, Vallejo-Bernal SM, Bustamante E (2016) A scaling approach to Budyko's framework and the complementary relationship of evapotranspiration in humid environments: case study of the Amazon River basin. Hydrol Earth Syst Sci 20(2):589–603

Chang X, Wang S, Gao Z, Luo Y, Chen H (2019) Forecast of daily reference evapotranspiration using a modified daily Thornthwaite Equation and temperature forecasts. Irrig Drain 68(2):297–317

Cheval S, Dumitrescu A, Birsan M-V (2017) Variability of the aridity in the South-Eastern Europe over 1961-2050. Catena 151:74–86

Cornes RC, van der Schrier G, van den Besselaar EJM, Jones PD (2018) An ensemble version of the E-OBS temperature and precipitation data sets. J Gerontol Ser A Biol Med Sci 123(17):9391–9409

Creed IF, Spargo AT, Jones JA, Buttle JM, Adams MB, Beall FD, Booth EG, Campbell JL, Clow D, Elder K, Green MB, Grimm NB, Miniat C, Ramlal P, Saha A, Sebestyen S, Spittlehouse D, Sterling S, Williams MW, Winkler R, Yao H (2014) Changing forest water yields in response to climate warming: results from long-term experimental watershed sites across North America. Global Change Biol 20(10):3191–3208

Dai A (2011) Drought under global warming: a review. WIREs Clim Change 2(1):45–65

Diaz-Padilla G, Sanchez-Cohen I, Guajardo-Panes RA, Del Angel-Perez AL, Ruiz-Corral A, Medina-Garcia G, Ibarra-Castillo D (2011) Mapping of the aridity index and its polulation distribution in Mexico. Revista CHapingo Serire Ciencias Forestales y Del Ambiente 17(SI):267–275

Duan Z, Chen Q, Chen C, Liu J, Gao H, Song X, Wei M (2019) Spatiotemporal analysis of nonlinear trends in precipitation over Germany during 1951-2013 from multiple observation-based gridded products. Int J Climatol 39(4):2120–2135

Fibbi L, Chiesi M, Moriondo M, Bindi M, Chirici G, Papale D, Gozzini B, Maselli F (2016) Correction of a 1 km daily rainfall dataset for modelling forest ecosystem processes in Italy. Meteorol Applications 23:294–303

Gao X, Giorgi F (2008) Increased aridity in the Mediterranean region under greenhouse gas forcing estimated from high resolution simulations with a regional climate model. Global Planet Change 62(3-4):195–209

Gardelin M, Lindstrom G (1997) Priestley-Taylor evapotranspiration in HBV-simulations. Nordic Hydrol 28(4-5):233–246

Gerrits AMJ, Savenije HHG, Veling EJM, Pfister L (2009) Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model. Water Resour Res 45(4):W04403, 1–15

Hanel M, Rakovec O, Markonis Y, Maca P, Samaniego L, Kysely J, Kumar R (2018) Revisiting the recent European droughts from a long-term perspective. Sci Rep 8:9499, 1–11

Hargreaves G, Samani Z (1985) Reference crop evapotranspiration from temperature. Appl Eng Agric 1(2):1–12

Harris I, Jones PD, Osborn TJ, Lister DH (2014) Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. Int J Climatol 34(3):623–642

Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Scient Data 7(1):109, 1–18

Hassler B, Lauer A (2021) Comparison of reanalysis and observational precipitation datasets including ERA5 and WFDE5. Atmosphere 12(11):1462, 1–30

Herrera S, Gutiérrez JM, Ancell R, Pons MR, Frías MD, Fernández J (2012) Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02). Int J Climatol 32(1):74–85

Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A, Muñoz Sabater J, Nicolas J, Peubey C, Radu R, Rozum I, Schepers D, Simmons A, Soci C, Dee D, Thépaut JN (2018) ERA5 monthly averaged data on single levels from 1959 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.f17050d7

Hijmans RJ (2021) raster: Geographic data analysis and modeling. R package version 3.5-2. https://CRAN.R-project.org/package=raster

Hofstra N, Haylock M, New M, Jones PD (2009) Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. J Gerontol Ser A Biol Med Sci 114:D21101, 1–16

Huang J, Ji M, Xie Y, Wang S, He Y, Ran J (2016) Global semi-arid climate change over last 60 years. Clim Dyn 46(3-4):1131–1150

Huguenin MF, Fischer EM, Kotlarski S, Scherrer SC, Schwierz C, Knutti R (2020) Lack of change in the projected frequency and persistence of atmospheric circulation types over Central Europe. Geophys Res Lett 47(9):e2019GL086132, 1–10

Isotta FA, Vogel R, Frei C (2015) Evaluation of European regional reanalyses and downscalings for precipitation in the Alpine region. Meteorologische Zeitschift 24(1):15–37

Keyantash JA, Dracup JA (2004) An aggregate drought index: assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. Water Resour Res 40(9):W09304, 1–13

Kohli G, Lee CM, Fisher JB, Halverson G, Variano E, Jin Y, Carney D, Wilder BA, Kinoshita AM (2020) ECOSTRESS and CIMIS: A comparison of potential and reference evapotranspiration in riverside county, California. Remote Sensing 12(24):4126, 1–12

Kucerova M, Beck C, Philipp A, Huth R (2017) Trends in frequency and persistence of atmospheric circulation types over Europe derived from a multitude of classifications. Int J Climatol 37(5):2502–2521

Laiti L, Mallucci S, Piccolroaz S, Bellin A, Zardi D, Fiori A, Nikulin G, Majone B (2018) Testing the hydrological coherence of high-resolution gridded precipitation and temperature data sets. Water Resour Res 54(3):1999–2016

Lakatos M, Weidinger T, Hoffmann L, Bihari Z, Horvath A (2020) Computation of daily Penman-Monteith reference evapotranspiration in the Carpathian Region and comparison with Thornthwaite estimates. Adv Sci Res 16:251–259

Lang D, Zheng J, Shi J, Liao F, Ma X, Wang W, Chen X, Zhang M (2017) A comparative study of potential evapotranspiration estimation by eight methods with FAO Penman-Monteith method in southwestern China. Water 9(10):734, 1–18

Lhotka O, Trnka M, Kysely J, Markonis Y, Balek J, Mozny M (2020) Atmospheric circulation as a factor contributing to increasing drought severity in Central Europe. J Gerontol Ser A Biol Med Sci 125(18):e2019JD032269, 1–17

Lioubimtseva E, Cole R, Adams JM, Kapustin G (2005) Impacts of climate and land-cover changes in arid lands of central Asia. J Arid Environ 62(2):285–308

Liu J, Xu S, Han X, Chen X, He R (2019) A multi-dimensional hydro-climatic similarity and classification framework based on Budyko theory for continental-scale applications in China. Water 11(2):319, 1–26

Markonis Y, Kumar R, Hanel M, Rakovec O, Maca P, AghaKouchak A (2021) The rise of compound warm-season droughts in Europe. Sci Adv 7(6):eabb9668, 1–7

Moravec V, Markonis Y, Rakovec O, Svoboda M, Trnka M, Kumar R, Hanel M (2021) Europe under multi-year droughts: how severe was the 20142018 drought period? Environ Res Lett 16(3):034062, 1–13

Myronidis D, Nikolaos T (2021) Changes in climatic patterns and tourism and their concomitant effect on drinking water transfers into the region of South Aegean, Greece. Stoch Env Res Risk Assess 35:1725–1739

Naumann G, Cammalleri C, Mentaschi L, Feyen L (2021) Increased economic drought impacts in Europe with anthropogenic warming. Nat Clim Chang 11(6):485–491

Navarro A, Garcia-Ortega E, Merino A, Luis Sanchez J, Kummerow C, Tapiador FJ (2019) Assessment of IMERG precipitation estimates over Europe. Remote Sens 11(21):2470, 1–17

Nechita C, Čufar K, Macovei I, Popa I, Badea ON (2019) Testing three climate datasets for dendroclimatological studies of oaks in the South Carpathians. Sci Total Environ 694:133730, 1–10

Nyman P, Sherwin CB, Langhans C, Lane PNJ, Sheridan GJ (2014) Downscaling regional climate data to calculate the radiative index of dryness in complex terrain. Aust Meteorol Oceanographic J 64(2):109–122

Oudin L, Hervieu F, Michel C, Perrin C, Andreassian V, Anctil F, Loumagne C (2005) Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling. J Hydrol 303(1-4):290–306

Paltineanu C, Mihailescu IF, Seceleanu I, Dragota C, Vasenciuc F (2007) Using aridity indices to describe some climate and soil features in Eastern Europe: a Romanian case study. Theoret Appl Climatol 90(3-4):263–274

Pan N, Wang S, Liu Y, Li Y, Xue F, Wei F, Yu H, Fu B (2021) Rapid increase of potential evapotranspiration weakens the effect of precipitation on aridity in global drylands. J Arid Environ 186:104414, 1–9

Peng L, Li Y, Feng H (2017) The best alternative for estimating reference crop evapotranspiration in different sub-regions of mainland China. Scient Reports 7:5458, 1–19

Pereira AR, Pruitt WO (2004) Adaptation of the Thornthwaite scheme for estimating daily reference evapotranspiration. Agricultural Water Management 66(3):251–257

Poyen FB, Ghosh AK, Kundu P (2016) Review on different evapotranspiration empirical equations. Int J Adv Eng Manag Sci 2:17–24

Pravalie R, Bandoc G (2015) Aridity variability in the last five decades in the Dobrogea Region, Romania. Arid Land Res Manag 29(3):265–287

R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rakovec O, Samaniego L, Hari V, Markonis Y, Moravec V, Thober S, Hanel M, Kumar R (2022) The 2018-2020 multi-year drought sets a new benchmark in Europe. Earths Future 10(3):e2021EF002394, 1–11

Rehor J, Brazdil R, Trnka M, Lhotka O, Balek J, Mozny M, Stepanek P, Zahradnicek P, Mikulova K, Turna M (2021) Soil drought and circulation types in a longitudinal transect over central Europe. Int J Climatol 41(1):E2834–E2850

Rolle M, Tamea S, Claps P (2021) ERA5-based global assessment of irrigation requirement and validation. Plos One 16(4):e0250979, 1–21

Salvati L, Sateriano A, Zitti M (2013) Long-term land cover changes and climate variations—a country-scale approach for a new policy target. Land Use Policy 30:401–407

Samani Z (2000) Estimating solar radiation and evapotranspiration using minimum climatological data. J Irrig Drain Eng 126(4):265–267

Sanchez E, Romera R, Gaertner MA, Gallardo C, Castro M (2009) A weighting proposal for an ensemble of regional climate models over Europe driven by 1961-2000 ERA40 based on monthly precipitation probability density functions. Atmos Sci Lett 10(4):241–248

Sarbu I, Sebarchievici C (2017) Solar heating and cooling systems (fundamentals, experiments and applications). Academic Press, London

Seiller G, Anctil F (2016) How do potential evapotranspiration formulas influence hydrological projections? Hydrol Sci J 61(12):2249–2266

Sentelhas PC, Gillespie TJ, Santos EA (2010) Evaluation of FAO Penman-Monteith and alternative methods for estimating reference evapotranspiration with missing data in Southern Ontario, Canada. Agri Water Manag 97(5):635–644

Sfica L, Beck C, Nita A-I, Voiculescu M, Birsan M-V, Philipp A (2021) Cloud cover changes driven by atmospheric circulation in Europe during the last decades. International Journal of Climatology 41(1):E2211–E2230

Skok G, Zagar N, Honzak L, Zabkar R, Rakovec J, Ceglar A (2016) Precipitation intercomparison of a set of satellite- and raingauge-derived datasets, ERA Interim reanalysis, and a single WRF regional climate simulation over Europe and the North Atlantic. Theoretical and Applied Climatology 123(1-2):217–232

Spinoni J, Vogt J, Naumann G, Carrao H, Barbosa P (2015) Towards identifying areas at climatological risk of desertification using the Koppen-Geiger classification and FAO aridity index. Intern J Climatol 35(9):2210–2222

Tabari H (2010) Evaluation of reference crop evapotranspiration equations in various climates. Water Resour Manage 24(10):2311–2337

Tapiador FJ, Sanchez E (2008) Changes in the European precipitation climatologies as derived by an ensemble of regional models. J Climate 21(11):2540–2557

Tegos A, Malamos N, Koutsoyiannis D (2015) A parsimonious regional parametric evapotranspiration model based on a simplification of the Penman-Monteith formula. J Hydrol 524:708–717

Tomczyk AM, Bednorz E, Polrolniczak M (2019) The occurrence of heat waves in Europe and their circulation conditions. Geografie 124(1):1–17

van der Schrier G, van den Besselaar EJM, Tank AMGK, Verver G (2013) Monitoring European average temperature based on the E-OBS gridded data set. J Gerontol Ser A Biol Med Sci 118(11):5120–5135

Van Loon AF (2015) Hydrological drought explained. Wiley Interdisciplinary Reviews-Water 2(4):359–392

Velikou K, Lazoglou G, Tolika K, Anagnostopoulou C (2022) Reliability of the ERA5 in replicating mean and extreme temperatures across Europe. Water 14(4):543, 1–15

Wald L (2018) Basics in solar radiation at Earth surface. https://www.researchgate.net/publication/322314967

Wang D, Alimohammadi N (2012) Responses of annual runoff, evaporation, and storage change to climate variability at the watershed scale. Water Resour Res 48(5):W05546, 1–16

Weiland FCS, Tisseuil C, Durr HH, Vrac M, van Beek LPH (2012) Selecting the optimal method to calculate daily global reference potential evaporation from CFSR reanalysis data for application in a hydrological model study. Hydrol Earth Syst Sci 16(3):983–1000

Wickham H (2016) ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wilhite DA (2000) Droughts: a global assessment, Chapter 1: Drought as a Natural Hazard: Concepts and Definitions. Routledge, London

Winter TC, Rosenberry DO, Sturrock AM (1995) Evaluation of 11 equations for determining evapotranspiration for a small lake in the North Central United-States. Water Resour Res 31(4):983–993

Xiang K, Li Y, Horton R, Feng H (2020) Similarity and difference of potential evapotranspiration and reference crop evapotranspiration - a review. Agric Water Manag 232:106043, 1–16

Xu CY, Singh VP (2000) Evaluation and generalization of radiation-based methods for calculating evaporation. Hydrol Process 14(2):339–349

Xu J, Peng S, Ding J, Wei Q, Yu Y (2013) Evaluation and calibration of simple methods for daily reference evapotranspiration estimation in humid East China. Archieves of Agronomy and Soil Science 59(6):845–858

Zhao H, Pan X, Wang Z, Jiang S, Liang L, Wang X, Wang X (2019) What were the changing trends of the seasonal and annual aridity indexes in northwestern China during 1961-2015? Atmospheric Res 222:154–162

Zheng H, Zhang L, Zhu R, Liu C, Sato Y, Fukushima Y (2009) Responses of streamflow to climate and land surface change in the headwaters of the Yellow River Basin. Water Res Res 45(7):W00A19, 1–9

Zhou S, Yu B, Huang Y, Wang G (2015) The complementary relationship and generation of the Budyko functions. Geophys Res Lett 42(6):1781–1790

Zveryaev II, Gulev SK (2009) Seasonality in secular changes and interannual variability of European air temperature during the twentieth century. J Geophys Res-Atm 114:D02110, 1–14

Zveryaev II (2004) Seasonality in precipitation variability over Europe. J Gerontol Ser A Biol Med Sci 109(D5):D05103, 1–16

Earth Syst. Sci. Data, 14, 4035–4056, 2022
https://doi.org/10.5194/essd-14-4035-2022

Open Access Earth System Science Data

A 500-year annual runoff reconstruction for 14 selected European catchments

Sadaf Nasreen[1], Markéta Součková[1], Mijael Rodrigo Vargas Godoy[1], Ujjwal Singh[1], Yannis Markonis[1], Rohini Kumar[2], Oldrich Rakovec[1,2], and Martin Hanel[1,3]

[1]Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Suchdol, Prague 16500, Czech Republic
[2]UFZ-Helmholtz Centre for Environmental Research, 04318 Leipzig, Germany
[3]T. G. Masaryk Water Research Institute, p.r.i., Prague 16000, Czech Republic

Correspondence: Martin Hanel (hanel@fzp.czu.cz)

Received: 28 August 2021 – Discussion started: 21 October 2021
Revised: 4 July 2022 – Accepted: 11 July 2022 – Published: 6 September 2022

Abstract. Since the beginning of this century, Europe has been experiencing severe drought events (2003, 2007, 2010, 2018 and 2019) which have had adverse impacts on various sectors, such as agriculture, forestry, water management, health and ecosystems. During the last few decades, projections of the impact of climate change on hydroclimatic extremes have often been used for quantification of changes in the characteristics of these extremes. Recently, the research interest has been extended to include reconstructions of hydroclimatic conditions to provide historical context for present and future extremes. While there are available reconstructions of temperature, precipitation, drought indicators, or the 20th century runoff for Europe, multi-century annual runoff reconstructions are still lacking. In this study, we have used reconstructed precipitation and temperature data, Palmer Drought Severity Index and available observed runoff across 14 European catchments in order to develop annual runoff reconstructions for the period 1500–2000 using two data-driven and one conceptual lumped hydrological model. The comparison to observed runoff data has shown a good match between the reconstructed and observed runoff and their characteristics, particularly deficit volumes. On the other hand, the validation of input precipitation fields revealed an underestimation of the variance across most of Europe, which is propagated into the reconstructed runoff series. The reconstructed runoff is available via Figshare, an open-source scientific data repository, under the DOI https://doi.org/10.6084/m9.figshare.15178107, (Sadaf et al., 2021).

1 Introduction

Global warming has impacted numerous land surface processes (Reinecke et al., 2021) over the last few decades, resulting in more severe droughts, heatwaves, floods and other extreme events. Droughts, in particular, pose a serious threat to Europe's water resources. The flow of many rivers is greatly hampered by prolonged droughts, which restrain the availability of fresh water for agriculture and domestic use. For example, the 2003 drought significantly reduced European river flows by approximately 60 % to 80 % relative to the average (Zappa and Kan, 2007). Likewise, the annual flow levels at several river gauges have decreased by 9 % to 22 % over the last decade (Middelkoop et al., 2001; Krysanova et al., 2008; Uehlinger et al., 2009; Su et al., 2020) due to a lack of rainfall and a warmer climate.

While runoff is a key element related to water security, it is challenging to interpret recent hydroclimate fluctuations (multi-year droughts in particular) considering observed runoff records (Markonis and Koutsoyiannis, 2016; Hanel et al., 2018), which are in general seldom available for years prior to 1900. In this way, the community does not have runoff information on various severe multi-year droughts and pluvial periods, which can be assessed only indirectly using (typically seasonal or annual) reconstruc-

Published by Copernicus Publications.

tions based on various proxy data, such as past tree rings (Nicault et al., 2008; Kress et al., 2010; Cook et al., 2015; Tejedor et al., 2016; Casas-Gómez et al., 2020), speleothem (Vansteenberge et al., 2016), ice cores, sediments (Luoto and Nevalainen, 2017), and documentary and instrumental evidence (Pfister et al., 1999; Brázdil and Dobrovolný, 2009; Dobrovolný et al., 2010; Wetter et al., 2011).

The majority of existing reconstructions focus on temperature (Luterbacher et al., 2004; Xoplaki et al., 2005; Casty et al., 2005; Büntgen et al., 2006; Moberg et al., 2008; Dobrovolný et al., 2010; Trouet et al., 2013; Emile-Geay et al., 2017), precipitation (Wilson et al., 2005; Boch and Spötl, 2011; Wilhelm et al., 2012; Murphy et al., 2018) or droughts (Büntgen et al., 2010; Kress et al., 2014; Cook et al., 2015; Tejedor et al., 2016; Ionita et al., 2017; Brázdil et al., 2018; Hanel et al., 2018) and floods (Wetter et al., 2011; Swierczynski et al., 2012). A few studies have been conducted for the reconstruction of runoff–drought deficit series (Hansson et al., 2011; Kress et al., 2014; Hanel et al., 2018; Moravec et al., 2019; Martínez-Sifuentes et al., 2020). However, these studies are either local or regional or cover a relatively short period. As an example, Hansson et al. (2011) introduced a runoff series for the Baltic Sea from the years 1550 to 1995 using temperature and atmospheric circulation indices. Similarly, Sun et al. (2013) used tree-ring proxies to reconstruct runoff in the Fenhe River basin in China's Shanxi region over the last 211 years. As another example, Caillouet et al. (2017) provide a 140-year dataset of reconstructed streamflow over 662 natural catchments in France since 1871 using the GR6J hydrological model, highlighting several well-known extreme low-flow events. A multi-ensemble modelling approach using GR4J has been applied by Smith et al. (2019) to develop UK-based historical river flows and examine the potential of reconstruction for capturing peak- and low-flow events from 1891 to 2015.

The available reconstructed precipitation and temperature series (or fields) can be used to reconstruct runoff with hydrological (process-based) models (Tshimanga et al., 2011; Armstrong et al., 2020) respecting general physical laws, such as preserving mass balance (e.g. MIKE SHE; Im et al., 2009 or VELMA; Laaha et al., 2017) or data-driven methods which are able to capture complex non-linear relationships (for instance support vector machines, Zuo et al., 2020; Ji et al., 2021; artificial neural networks (ANNs), Senthil Kumar et al., 2005; Hu et al., 2018; Kwak et al., 2020; random forests, Ghiggi et al., 2019; Li et al., 2021; Contreras et al., 2021). While the lack of physical constraints in the data-driven models limits their application under changing boundary conditions (in comparison with those of the model training period), their advantage is that they can often directly use biased reconstructed data as an input series.

The objective of the present study is to provide a multi-century annual runoff reconstruction for 14 European catchments, utilizing the available precipitation (P; Pauling et al., 2006) and temperature (T; Luterbacher et al., 2004) reconstructions and the Old World Drought Atlas self-calibrated Palmer Drought Severity Index (scPDSI) reconstruction (Cook et al., 2015). Specifically, we assessed a conceptual lumped hydrological model (GR1A; Mouelhi et al., 2006) and two data-driven models, long short-term memory neural network (LSTM; Chen et al., 2020) and Bayesian regularized neural network (BRNN; Okut, 2016), for annual runoff simulation over the period 1500–2000.

Section 2 introduces P and T hydroclimatic reconstructions and the scPDSI drought indicator as well as precipitation, temperature and runoff observations. In Sect. 3, we describe the data preprocessing, models, the drought identification methodology and goodness-of-fit assessment. The accuracy of the employed P and T reconstructions, as well as the derived runoff simulations, is evaluated in Sect. 4. Finally, we summarize the advantages and limitations of reconstructed datasets in the Conclusions in Sect. 6.

2 Data

This section presents the data used in this study. To force the models, we investigated the use of precipitation (Pauling et al., 2006) and temperature (Luterbacher et al., 2004) reconstructions for the past half-millennium and scPDSI drought indicator data from the Old World Drought Atlas (Cook et al., 2015). For validating the runoff reconstructions, we used runoff from the Global Runoff Data Center (GRDC; Fekete et al., 1999). The accuracy of atmospheric forcing reconstruction used as model input was assessed using the observational data records of P and T from the Global Historical Climatology Network (GHCN; Menne et al., 2018). The datasets are summarized in Table 1 and are described in more detail below.

2.1 Precipitation

We used reconstructed seasonal precipitation ($0.5° \times 0.5°$) over Europe (30.25–70.75° N, 29.75° W–39.75° E) from 1500 to 2000. Reconstructed precipitation (P) was derived by Pauling et al. (2006) through principal component regression based on documented evidence (i.e. memoirs, annals and newspapers), speleothem proxy records (Proctor et al., 2000) and tree-ring chronologies from the International Tree-Ring Data Bank (ITRDB; Jeong et al., 2021) .

2.2 Temperature

Reconstructed temperature (T) was obtained from Luterbacher et al. (2004), which relies on historical records and seasonal natural proxies (i.e. ice cores from Greenland and tree rings from Scandinavia and Siberia). Reconstructed temperature data are available at the same spatial and temporal resolution as precipitation (see Table 1). We refer to both of these datasets as reconstructed forcings or reconstructed precipitation/temperature fields.

Table 1. Summary of considered datasets.

Reference	Domain	Temporal coverage (CE)*	Spatial resolution	Variables
Pauling et al. (2006)	Europe	1500–2000	0.5° × 0.5°	seasonal precipitation
Luterbacher et al. (2004)	Europe	1500–2000	0.5° × 0.5°	seasonal temperature
Menne et al. (2018)	Global	1760–2010	26 000 point stations	monthly mean temperature
Menne et al. (2018)	Global	1760–2010	20 590 point stations	monthly mean precipitation
Cook et al. (2015)	Europe	0–2012	0.5° × 0.5°	summer Palmer Drought Severity Index

* CE: Common Era.

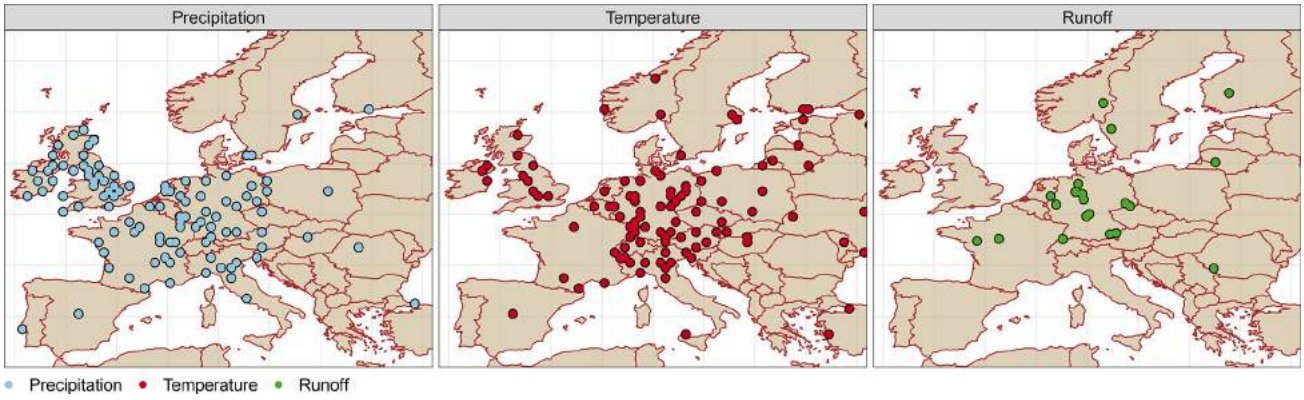


Figure 1. Spatial distribution of the observed GHCN precipitation and temperature stations and GRDC runoff gauges.

2.3 Self-calibrating Palmer Drought Severity Index (scPDSI)

In addition, we used data from the Old World Drought Atlas (OWDA; Cook et al., 2015) which contains information regarding moisture conditions across Europe, specifically the self-calibrated Palmer Drought Severity Index (scPDSI) using summer-related tree-ring proxies over the period 0 to 2012 CE.

2.4 The Global Historical Climatology Network (GHCN)

The GHCN dataset (GHCN; Peterson and Vose, 1997) is one of the largest observational databases, collated by the National Oceanic and Atmospheric Administration (NOAA; Quayle et al., 1999). The GHCN-m dataset contains observed temperature, rainfall and pressure data from 1701 to 2010. Data for the majority of stations are, however, available after 1900. GHCN-m precipitation and temperature from GHCN V2, as well as from the GHCN V4 version (Menne et al., 2012), were used to assess the reconstruction accuracy of the P and T fields as an input into the considered models. We selected 113 precipitation and 144 temperature stations within the European domain (see Fig. 1) with records dating back earlier than 1875. Most stations are geographically concentrated in central Europe, and few stations are located in the eastern and northern areas of Europe (see Table 2). These data, hereafter, are referred to as the GHCN data.

2.5 Observed runoff

The Global Runoff Data Center (GRDC; http://www.bafg.de/GRDC/EN/Home/homepage_node.html, last access: 24 November 2016) provides data for more than 2780 gauging stations in Europe, with the oldest records starting from 1806. Only the GRDC runoff time series with at least 25 years of data prior to 1900 were selected. In total, there were 21 such stations predominantly available in central Europe: 11 in Germany, 2 in France, 2 in Switzerland, 1 in the Czech Republic, 1 in Sweden, 1 in Finland, 1 in Lithuania and 1 in Romania (see Fig. 1). These stations cover 12 European river basins (Rhine, Loire, Elbe, Danube, Wesser, Main, Glama, Slazach, Nemunas, Gota Alv, Inn and Kokemaenjoke), with areas ranging from nearly 6100 km^2 (Kokemaenjoki, Muroleenkoski, Finland) to 576 000 km^2 (Danube, Orsova, Romania). The mean annual discharge (Q_{mean}) varies from 50 to 5 600 m^3 s^{-1} and spans different time periods for each catchment.

The most extensive records were available in Sweden (Vargoens KRV) and Germany (Dresden), containing the longest discharge series of 212 and 208 years, respectively. The gauging station in Köln also provided 195 years of data for the Rhine River. Note that some of the gauging stations are located nearby and therefore have a greater degree of similarity in their runoff time series (e.g. two stations in Basel, Rhine). Detailed information relating to all selected stations is provided in Table 2.

Table 2. Selected study catchments.

Station	River	GRDC no.	Latitude [° N]	Longitude [° E]	Drainage area [km^2]	Mean annual discharge [m^3 s^{-1}]	Start year	Length [year]
Orsova, RO	Danube	6742200	44.7	22.42	576 232	5602	1840	151
Decin, CZ	Elbe	6140400	50.79	14.23	51 123	309	1851	150
Dresden, DE	Elbe	6340120	51.05	13.73	53 096	332	1806	208
Elverum, NO	Gloma	6731401	60.88	11.56	15 426	251	1871	44
Vargoens KRV, SW	Gota Alv	6229500	58.35	12.37	46 885.5	531	1807	212
Wasserburg, DE	Inn	6343100	48.05	12.23	11 983	354	1827	177
Muroleenkoski, FI	Kokemaenjoki	6854104	61.85	23.910	6102	53.1	1863	155
Blois, FR	Loire	6123300	47.58	−0.86	38 240	362	1863	117
Montjean, FR	Loire	6123100	47.58	1.33	110 000	911	1863	117
Schweinfurt Neuer Hafen, DE	Main	6335301	50.03	10.22	12 715	103	1845	156
Würzburg, DE	Main	6335500	49.79	9.92	14 031	108	1824	177
Smalininkai, LT	Nemunas	6574150	55.07	22.57	81 200	531	1812	185
Basel Rheinhalle, CH	Rhine	6935051	47.55	7.61	35 897	1043	1869	140
Basel Schifflaende, CH	Rhine	6935052	47.55	7.58	35 905	1042	1869	127
Köln, DE	Rhine	6335060	50.93	6.96	144 232	2085	1817	195
Rees, DE	Rhine	6335020	51.75	6.39	159 300	2251	1815	183
Burgausen, DE	Salzach	6343500	48.15	12.83	6649	258	1827	174
Hann. Münden, DE	Wesser	6337400	51.42	9.64	12 442	109	1831	182
Bodenwerder, DE	Wesser	6337514	51.97	9.51	15 924	145	1839	175
Vlotho, DE	Wesser	6337100	52.17	8.86	17 618	170	1820	194
Intschede, DE	Wesser	6337200	52.96	9.12	37 720	320	1857	154

2.6 Study area

In the first part of the study, the grid-based reconstruction of precipitation and temperature was verified against the available GHCN data across the European region bounded by (30.25–70.75° N, 29.75° W–39.75° E). The second part focused on 21 specific central European catchments, corresponding to the available long-term GRDC discharge records. The study area and the observational data of the hydroclimatic variables are shown in Fig. 2.

3 Methods

This section is divided into three parts. The first part describes the preprocessing of the reconstructed forcings (i.e. precipitation and temperature) for validation across Europe and the preparation of data for runoff simulation in 21 catchments (Sect. 3.1). The hydrologic and data-driven models used to generate the runoff reconstructions are presented in Sect. 3.2 and 3.3, respectively. Finally, Sect. 3.4 describes the methods for the evaluation of simulated runoff and reconstructed forcings, and Sect. 3.5 presents the methods to identify annual runoff droughts.

3.1 Data preprocessing

Two databases were considered for the analysis and development of the annual runoff reconstruction. The first one was used for evaluating the accuracy of meteorological forcing reconstructions used for hydrological simulations and consists of observed GHCN data for all available European stations with long records (see Sect. 2.4) and values of corresponding grid cells from the reconstructed forcings dataset.

The second database was created as the basis for runoff reconstruction containing the observed runoff data for 21 selected catchments (Table 2) and the corresponding input variables of the models used to generate the multi-century runoff reconstructions. Several input variables were considered for inclusion in models such as reconstructed precipitation and temperature and Old World Drought Atlas scPDSI. The catchment average precipitation, temperature and scPDSI were estimated from the corresponding (gridded) datasets by averaging the relevant grid cells over the catchments. This database was further divided into two parts, calibration (1900–2000) and validation (before 1900), to assess the model's accuracy and to select an appropriate model. The data preprocessing, model selection, and evaluation of the models are depicted in Fig. 2.

3.2 Hydrologic model (GR1A)

We applied the annual timescale hydrologic model, GR1A (Mouelhi et al., 2006), to simulate annual runoff in each catchment. GR1A is a conceptual lumped hydrologic model (Manabe, 1969), considering dynamic storage and antecedent precipitation conditions. The model consists of a simple mathematical equation with a single (optimized) parameter:

$$Q_i = P_i \left\{ 1 - \frac{1}{\left[1 + \left(\frac{0.8P_i + 0.2P_{i-1}}{XE_i} \right)^2 \right]^{0.5}} \right\}, \quad (1)$$

208

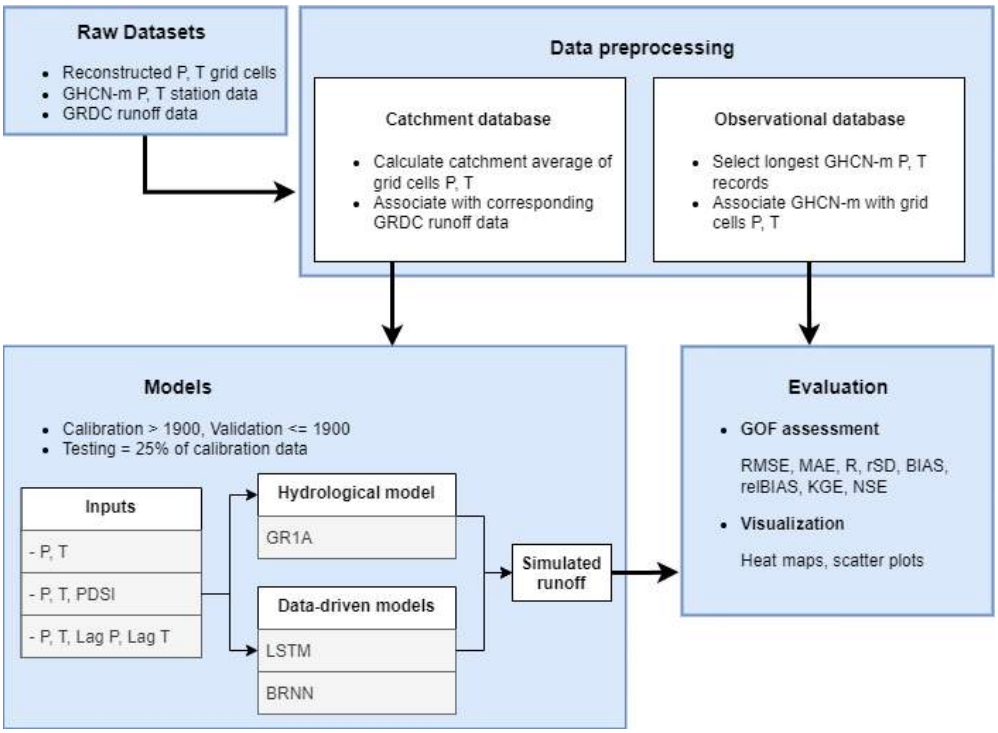


Figure 2. A schematic representation of workflow carried out in the study.

where Q, E and P represent annual runoff, basin average potential evapotranspiration and basin average precipitation, respectively, and i denotes the year. The parameter X is optimized individually for each catchment by maximizing the Nash–Sutcliffe efficiency (NSE) between observed and simulated runoff. Default gradient-based optimization from the R package airGR (Coron et al., 2017) was used. The potential evapotranspiration was calculated using the temperature-based formula (Oudin et al., 2005). Compared to other conceptual models from the GR family (GR4J, GR5J), GR1A is simple to use, and it allows for analysing many variants, particularly defining the best antecedent rainfall, and is potentially useful to predict wet and dry hydrologic conditions (Mouelhi et al., 2006).

3.3 Data-driven models

Artificial neural networks (ANNs; Senthil Kumar et al., 2005; Kwak et al., 2020) can describe non-linear relationships and are widely used for rainfall–runoff prediction. The ANNs consist of artificial neurons organized in layers and connections that route the signal through the network. Each connection has an associated weight that is optimized within the calibration (in the context of ANNs, known as training). There are many types of ANNs which differ in terms of structure and type of connections, as well as direction and functional forms used for neuron activation or training. In the present study, we considered two approaches: long short-term memory (LSTM) neural networks and Bayesian regularized neural networks (BRNNs). These approaches have been commonly used in past rainfall–runoff modelling studies (Hu et al., 2018; Kratzert et al., 2018; Xiang et al., 2020; Ye et al., 2021). We considered combinations of reconstructed forcing, OWDA-based scPDSI and lagged forcing as an input into the network for both model types. Specifically, the network using only reconstructed precipitation and temperature fields is referred to as $[P, T]$, the network with reconstructed forcing and OWDA scPDSI is termed as $[P, T, \text{PDSI}]$; and finally the network which includes 1-year lagged P and T forcing in addition to actual P and T is referred to as $[P, T, \text{lag}]$. We also considered and explored lag times longer than 1 year. However the correlation between precipitation and runoff drops significantly at lag times longer than 1 year and therefore was not included in presented analysis.

Figure A1 shows the architecture of LSTM, which is a modified version of the recurrent neural network (Hochreiter and Schmidhuber, 1997), using backpropagation in time (Werbos, 1990). LSTM is known for efficient simulation of time series with long-term memory (Van Houdt et al., 2020). It generally consists of two unit states (hidden and cell states) and three distinct gates (hidden, input and output). In this process, the cell state saves the long-term memory at the previous unit, while hidden states act as a working memory to process information inside the gates. These gates can determine which information needs to be processed, remembered and transferred in the next state. With LSTM, different ac-

tivation functions, such as hyperbolic tangent and sigmoid, can be used to update unit states. The implementation of the LSTM is carried out by applying the R packages "keras" (Arnold, 2017) and "tensorflow" (Abadi et al., 2016).

The training process of the LSTM is time-consuming due to its inherent complexity. Therefore we also considered the BRNN models that provide fast learning and convergence and were already used to tackle the complex relationship between rainfall and runoff (Ye et al., 2021). BRNNs are based on recurrent neural networks, which are often used to model time-series data (Wang et al., 2007), and the methods are extended with Bayesian regularization (Okut, 2016) to account for uncertainty related to network parameters and input data (Zhang et al., 2011). We trained this model in R using the "brnn" function of the "caret" package (Kuhn, 2015). More details are available in Appendix A3.

To set the optimal hyperparameters of the models (such as the number of neurons and activation functions) and to reduce the likelihood of overfitting during the calibration/training, the model performance was cross-checked considering an independent (or so-called "testing") set. The testing set was for each learning exercise extracted from the calibration data (1900–2000) as a random fraction (25 %). This process of the model development was repeated several times, minimizing the root mean square error (for BRNN) and mean square error (for LSTM) for each catchment individually. The model with the best performance was then chosen for further evaluation.

3.4 Goodness-of-fit assessment

We used a set of seven statistical metrics to assess the performance of simulated runoff, namely Nash–Sutcliffe efficiency (NSE), Pearson correlation (R), standard deviation ratio (rSD), Kling–Gupta efficiency (KGE), root mean square error (RMSE), mean absolute error (MAE), bias (BIAS) and relative bias (relBIAS). The mathematical formulations of these metrics are provided in Appendix A1.

3.5 Runoff drought identification

To check the utility of our reconstruction, we finally explore how well the annual runoff droughts are represented in the simulations. Our study considers annual hydrological droughts, defined as the streamflow/runoff deficit, following the threshold level approach (Yevjevich, 1967; Sung and Chung, 2014; Rivera et al., 2017). This approach is typically used for daily or monthly timescales, considering 0.1 or 0.2 quantile threshold levels. To accommodate the annual scale used here, we defined the start of the drought, when the annual runoff anomaly falls below the 0.33 quantile (regular drought) and the 0.05 quantile (extreme drought). The drought persists until the runoff rises above the threshold again. After that, the difference between runoff and the threshold was determined for each identified drought year, called the runoff deficit. Hydrological drought series can be further assessed to understand the critical aspects of runoff (temporal) dynamics and to classify past droughts in Europe (Wetter and Pfister, 2013; Cook et al., 2015).

4 Results and discussion

In this section, we analyse the 500-year annual reconstruction over space and time across Europe. Firstly, we provide a comparison between the GHCN-observed precipitation and temperature and the corresponding grid cells from Pauling et al. (2006) and Luterbacher et al. (2004) reconstructions. Next, the reconstructed annual runoff series for the selected catchments are evaluated against the corresponding observed GRDC runoff data.

Two distinct model types were investigated, i.e. a process-based conceptual lumped hydrological model (GR1A) and two data-driven models (BRNN and LSTM). While the former takes reconstructed forcing of precipitation and temperature as an input, in the case of the latter, we also considered PDSI and lagged reconstructed precipitation and temperature fields, as shown in Table 3. Statistical metrics, such as NSE, KGE, RMSE, MAE, R, BIAS and relBIAS (Appendix A1), are used to quantify the predictive skills of the models examined.

4.1 Evaluation of reconstructed precipitation and temperature fields

The 500-year annual paleoclimate reconstructions of precipitation (P) and temperature (T) were validated against the GHCN observation data. The map showing the comparison is given in Figs. 3 and 4. The reconstructed data are evaluated against observational P and T across 99 and 94 European sites, respectively. Figure 3 shows that for most of the sites the correlation coefficient (R) of P reconstruction at most of the sites is above 0.5; the relative bias (relBIAS) is between -0.1 and 0.1; KGE and NSE are showing values below 0.5 and 0.6, respectively; the rSD is between 0.7 and 0.9; and RMSE varies between 0 and 150.

The performance of the temperature reconstruction was relatively better, as depicted in Fig. 4. In this case, RMSE between reconstructed and observational T is around 0.2 °C, and rSD fluctuates between 0.95 and 1.05, while R is higher than 0.84, and BIAS is less than 0.5 °C, except for stations located in the Alps. The NSE and KGE values were above 0.5 at the majority of the stations. Low skill observed at some locations can be explained by the unresolved variability of grid-cell average temperature, especially in regions with complex terrain.

It is worth noting that the large spread of goodness-of-fit (GOF) statistics is mainly due to the outlying values at the grid cells located along the boundary of the domain (i.e. the interface between land and sea/ocean) and high elevations (see also Figs. 3 and 4). In general, reconstructed precipita-

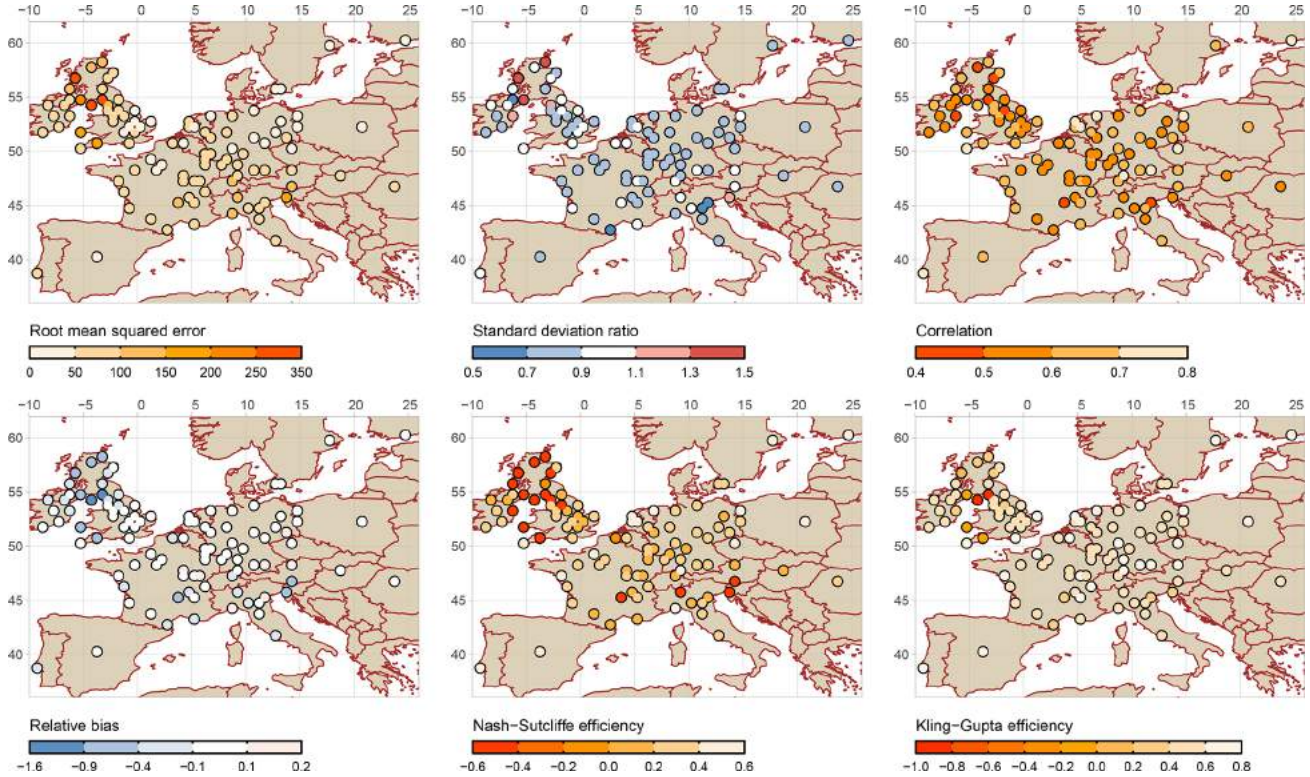


Figure 3. Validation of reconstructed precipitation (Pauling et al., 2006) against GHCN observations.

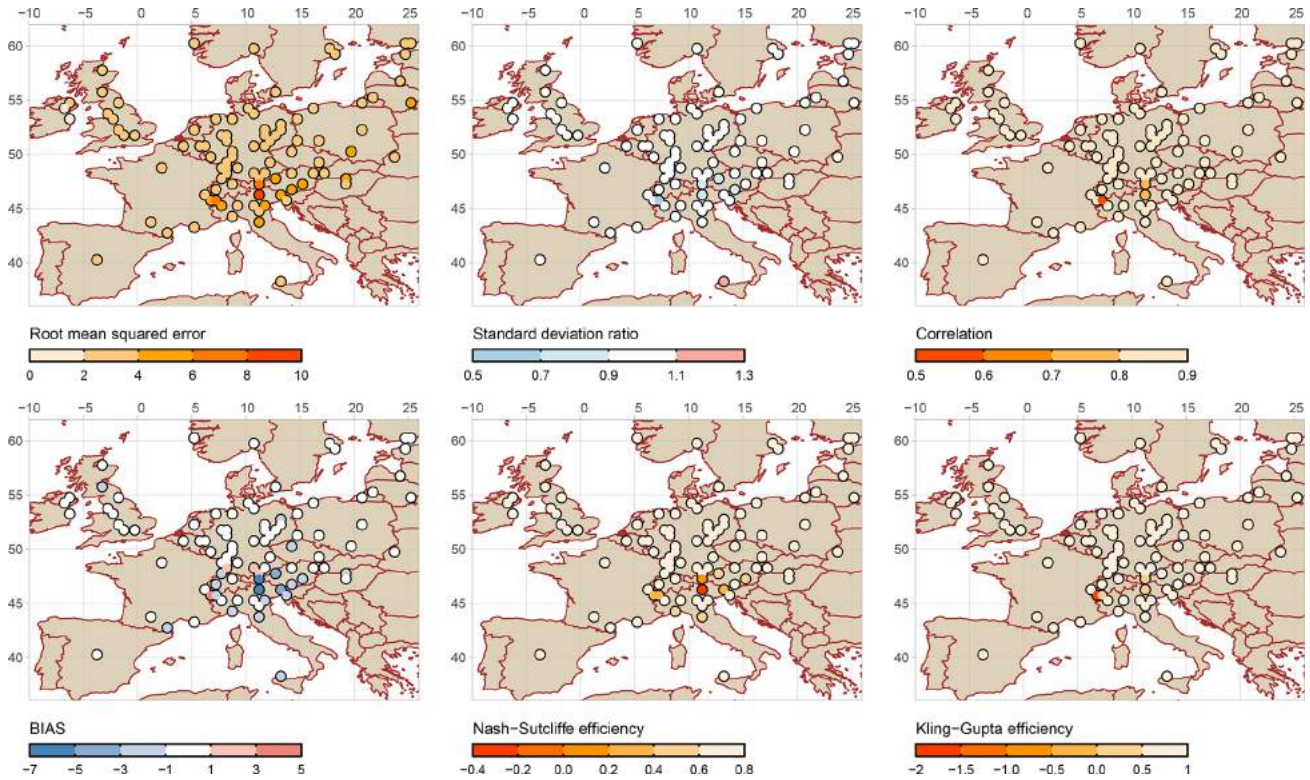


Figure 4. Validation of reconstructed temperature (Luterbacher et al., 2004) against GHCN observations.

tion exhibits greater differences from observations than temperature. This may be because the proxies considered in the reconstruction rely on different seasons and climate conditions. Additionally, the shortest available instrumental data before the 20th century could encounter certain technical errors, such as problems with instrumental tools, station relocation and dating issues (Dobrovolný et al., 2010). Moreover, other studies (e.g. Ljungqvist et al., 2020) stated that the precipitation series employed for the reconstructions were relatively shorter and more erroneous than the temperature series before the 20th century (Pauling et al., 2006; Harris et al., 2014). Finally, the chosen statistical technique (principal component regression) could also possibly contribute to variance inflation with larger timescales (Pauling et al., 2006).

4.2 Assessment of the reconstructed runoff simulations

The GR1A conceptual hydrological model was driven by catchment average P and T and calibrated using observed annual runoff for each catchment separately. The simulated annual runoff series were then compared to the corresponding GRDC observations (for calibration and validation periods), and the results were summarized by means of GOF statistics. As can be seen in Fig. 5, the correlation and NSE statistics for calibration achieve reasonable results at most of the catchments, with a few exceptions (i.e. Kokemenjoki, Goeta, Nemunas and Inn). The catchments with relatively poor skills are located in northern Europe, which is in line with the previous findings by Seiller et al. (2012), who noted that the lumped hydrological models often exhibit larger uncertainties and fail to capture the extreme catchment values (both high and low) in those regions. The low skill for some of the catchments cannot be easily attributed only to bias in reconstructed precipitation and temperature (described in Sect. 4.1) but rather to low station and proxy coverage in some (especially northern) parts of Europe, leading to biased basin-average precipitation and temperature estimate. Another study of Fathi et al. (2019) suggested that the performance of the GR1A model is less efficient than the new Budyko-framework-based SARIMA model in simulating the annual runoff across the Blue Nile and the Danube catchment. This may be due to the simplified nature of the model that does not easily capture the complex relationship between rainfall and runoff variability.

In general, statistical values presented in heat maps (Fig. 5) indicate that the neural network algorithms are more skilled for runoff prediction than the GR1A model. The NSE and R statistics for the BRNN and LSTM models indicate a significant improvement in runoff prediction, as compared to the results obtained through the GR1A model. For instance, for Basel Rheinhalle the NSE increases from 0.27 to 0.73 (BRNN) and 0.75 (LSTM) for calibration and 0.2 to 0.54 (BRNN) and 0.52 (LSTM) for validation. Moreover, including scPDSI from OWDA with reconstructed forcing $[P, T, \text{PDSI}]$ increases the performance slightly more (NSE 0.76 for calibration and 0.57/0.59 for validation, for BRNN/LSTM, respectively), and considering the lagged forcing results in the best performance (NSE 0.75/0.8 for calibration and 0.6/0.54 for validation, for BRNN/LSTM).

Similarly for all sites, the data-driven methods exhibited a strong correlation with the observed runoff, with the GR1A simulations resulting most frequently in lower correlation values. Other metrics (RMSE, MAE, KGE, rSD and relBIAS) are shown in Figs. S1–S5 in the Supplement. Across many study locations, the combination of reconstructed forcings and their 1-year lag performed the best in terms of rapid convergence (the number of iterations needed) and high accuracy from all input combinations for both data-driven models (BRNN, LSTM). For the validation period, the mean NSE (across all catchments) for the GR1A model is 0.16, for the BRNN $[P, T, \text{lag}]$ it is 0.68 and it improves to 0.73 for the LSTM $[P, T, \text{lag}]$. In the case of the mean KGE, GR1A yields 0.62, BRNN $[P, T, \text{lag}]$ is 0.73 and LSTM $[P, T, \text{lag}]$ is 0.78.

To further demonstrate the differences between the individual models, we show the simulated runoff series for all models for those catchments with the highest (Blois–Loire) and lowest (Smalininkai–Nemunas) performance in Fig. 6. The performance of the models is comparable during the calibration period for the Loire River. Clearly, all data-driven models are capable of mimicking the observed runoff, while the GR1A model exhibited certain minor deviations, primarily until 1930. In the validation period, the differences between the models are more visible, in particular, for above-average flows. This can be attributed to different generalization skill of individual models. At the beginning of the validation period (1870–1880), all models failed to simulate the high annual flows.

In the case of Nemunas catchment, the GR1A simulation deviates extremely from the observed data and cannot capture the mean flow level. However, the calibration is poor, even for the data-driven models, and does not simulate the year-to-year variability appropriately. Interestingly, for the validation period, the error in the GR1A model decreases. The performance of the data-driven models is similar in validation and calibration periods. Looking at the GOF statistics, the models considering OWDA-based scPDSI or lagged forcings (e.g. P_{t-1}) perform slightly better in terms of KGE than the other model configurations.

4.3 The annual runoff reconstruction datasets

As a first step, we excluded the catchments that exhibited poor performance in validation (see Fig. 5). As a threshold, we considered validation NSE greater than 0.5 for at least one model, following the approach used by Ayzel et al. (2020). In this step, we excluded 7 catchments (Vlotho–Wesser, Decin–Elbe, Burghausen–Salzach, Smalininkai–Nemunas, Vargoens KRV–Goeta, Elverum–Glama and Muroleekoski–

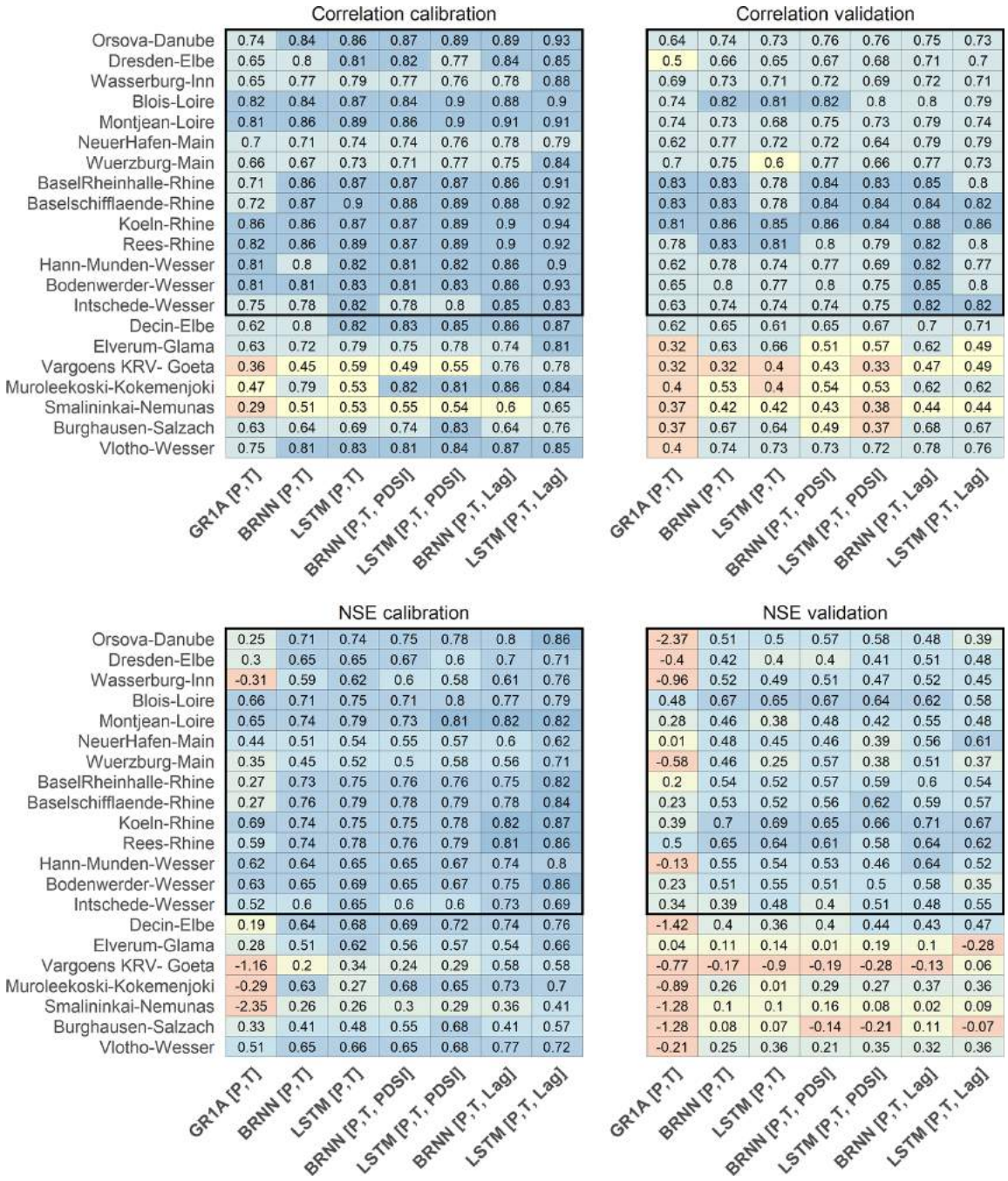


Figure 5. The correlation coefficient (top) and NSE (bottom) for calibration (left) and validation (right) of the considered models for 21 study catchments. The vertical axis represents the catchments (station name and river) and the horizontal axis the considered models. The rectangular black frames represent the catchments with NSE > 0.5 over the validation period.

Kokemenjoki) out of 21, ending up with a set of simulations for 14 catchments (highlighted by the rectangular box in Fig. 5).

Secondly, we identified the best candidate models for each of the 14 selected catchments, considering the GOFs based on the validation NSE and R greater than 0.5 and 0.7, respectively. The best model for each catchment was finally selected from those models considering the remaining validation measures (relBIAS, rSD, KGE, RMSE and MAE) as well. Specifically, we picked the models with consistent good validation measures. This choice is partly subjective, and more formal selection should be explored further. On the other hand, the candidate models all performed comparably in most cases.

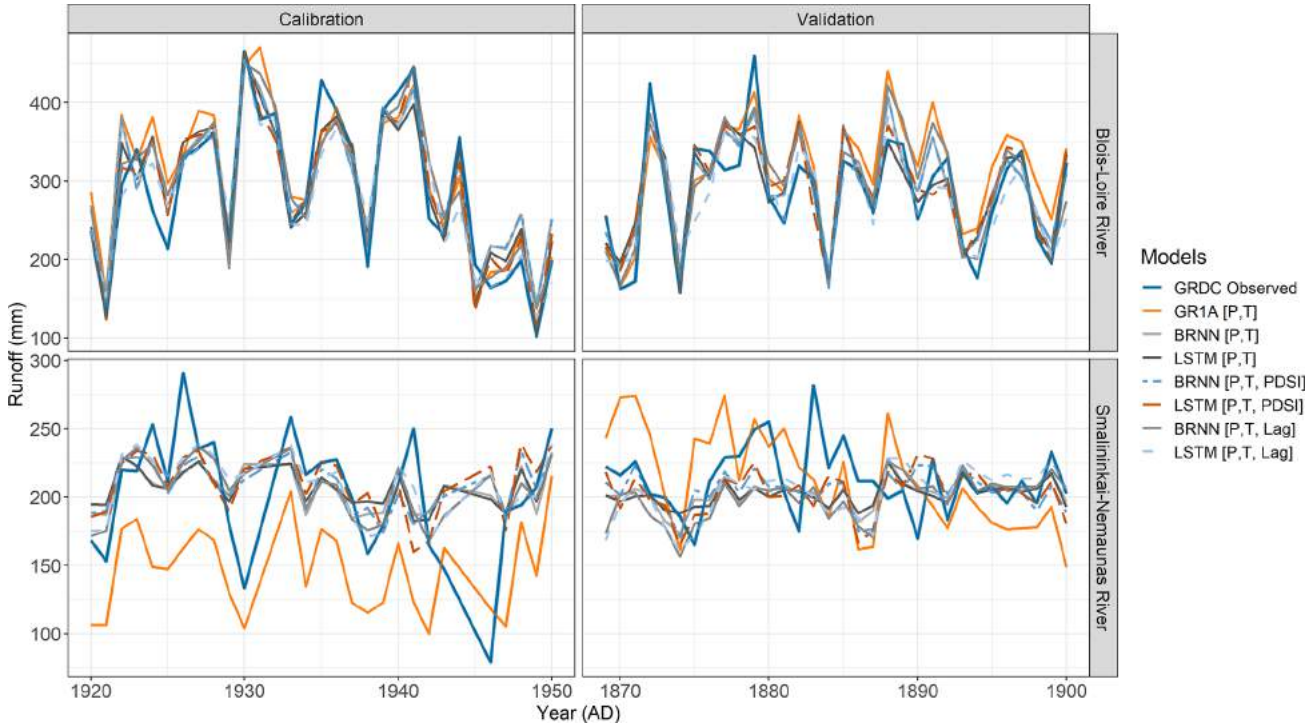


Figure 6. Comparison between the models for the station with the best (Blois–Loire River, top) and the worst (Smalininkai–Nemunas River, bottom) model fit.

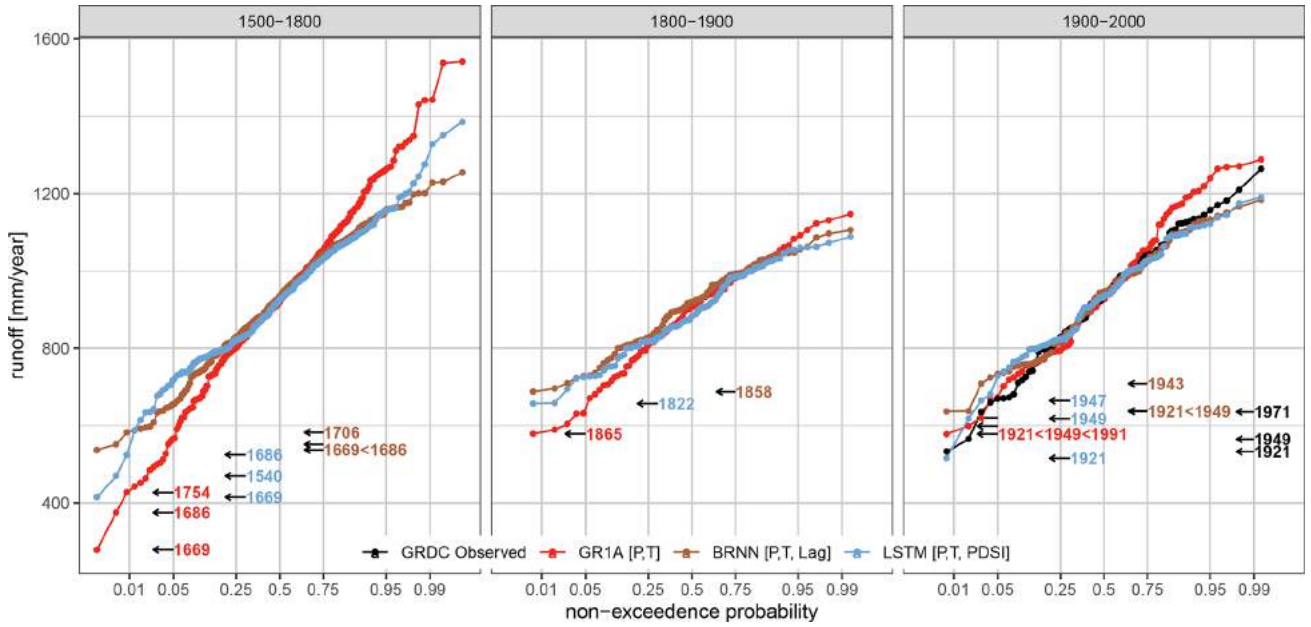


Figure 7. Distribution functions for BRNN [P, T, lag], LSTM [P, T, PDSI], i.e. the best two models, GR1A [P, T]- and GRDC-observed data for the periods 1500–1800, 1800–1900 and 1900–2000 over the Basel Rheinhalle–Rhine catchment. The values on the horizontal axis are transformed using the "probit" function. The coloured labels indicate the most extreme drought years according to each model.

The resulting selected models are shown in Table 3. The combination of reconstructed forcing with 1-year time lags results in the best performance over nine catchments, of which seven employed the BRNN and the remainder the LSTM model. The LSTM with reconstructed forcing and OWDA-scPDSI was the best in just one case, and the remaining time-series reconstructions were most appropriately simulated with the BRNN $[P, T]$ and BRNN $[P, T, \text{PDSI}]$. It should be noted that the differences between the models performing well are small, as noted in Fig. 6 and further demonstrated in Fig. 7. The latter figure compares the cumulative distribution functions of annual runoff for the periods 1500–1800, 1800–1900 and 1900–2000, as simulated by the BRNN $[P, T, \text{lag}]$ and LSTM $[P, T, \text{PDSI}]$ – the two best-performing models – and the GR1A (the most deviating simulation from the best model) with the distribution of the observed annual runoff for the Basel Rheinhalle–Rhine catchment. For the calibration period (1900–2000) in Fig. 7, the models perform well except the GR1A, which generally overestimated the observed maxima. The cumulative distribution of BRNN- and LSTM-simulated runoff values is very similar for the validation period (1800–1900), except for the top and bottom 5 % in 1500–1800. The GR1A simulation showed significant differences for the entire distribution, thus overestimating/underestimating the maxima/minima. Our finding shows that GR1A simulates a Rhine minimum of 279 mm yr^{-1} in Basel, whereas the observed minimum in the past century is greater than 532.6 mm yr^{-1}, inferring that the cumulative distribution function (CDF) has significantly lower/higher runoff values between 1500 and 1800 for BRNN and GR1A, whereas LSTM appears to extrapolate less. The difference from the best model can be expressed in terms of KGE – even here, it was evident that the GR1A model deviated considerably (KGE 0.6–0.7), while the LSTM is very similar to the BRNN (KGE 0.92–0.96). The most severe drought year identified by the models in the period 1500–1800 appears to be 1669 and the year 1921 in the past century (1900–2000) (Fig. 7 left and right panels), while for 1800–1900 the models identified either 1865 (GR1A, LSTM) or 1858 (BRNN). Please note that the 1858 low-water mark is available at Laufenburg Pfister et al. (2006) near Basel and was regarded as one of the worst winter droughts in the last 200 years.

The resulting 14 annual runoff reconstructions are available at https://doi.org/10.6084/m9.figshare.15178107 and are shown in Figs. S6–S8 in the Supplement. As an example, we present only two runoff reconstructions here (Fig. 8). As an additional validation for the reconstructed series, we inspected the scatter plots of the observed and reconstructed runoff (Fig. 9). The simulated series are generally consistent with the observed runoff, especially for the Montjean–Loire, Köln–Rhine and Basel Schifflaende–Rhine catchments, which exhibit the best relationship between the observed and the simulated runoff.

Finally, to check the consistency of our reconstructed dataset, we compared the skill of our simulation with respect to the GRDC runoff observation and the GSWP3-forced GRUN monthly runoff (Ghiggi et al., 2019) datasets. The gridded GRUN datasets were averaged over the respective catchments to enable comparison (Figs. S9 and S10 in the Supplement). Our reconstruction outperforms GRUN data in terms of RMSE, MAE, relBIAS and NSE across the majority of the catchments, while the correlation (reproduction of interannual dynamics) to GRDC runoff is slightly higher for GRUN compared to our reconstruction. The variability, which our data-driven models underestimate (on average by 16.5 %), is overestimated by GRUN (on average by 17.2 %). Since the correlation compensates for the relBIAS, the KGE for our reconstruction and GRUN is comparable. This suggests that GRUN could be used for data-driven model training, provided at least some information on flow characteristics is available in the catchment.

4.4 Identification of low flows, significant hydrological drought events and trends

In the final step of the analysis, we compared the droughts identified in the reconstructions with the GRDC-observed series (Fig. 10). The agreement between the simulated and observed runoff deficit is lower compared to the annual runoff time series. For most of the stations, the simulated deficit is lower than the corresponding observed estimates. This suggests that the reconstructed precipitation and temperature fields do not represent the inter-annual variability correctly. Despite a widespread issue with the representation of inter-annual persistence, Fig. 10 shows that the runoff deficits are simulated reasonably well for the Rees–Rhine and Köln–Rhine catchments.

Furthermore, we contrasted reconstructed drought patterns over the last 500 years with data available from documentary evidence and other sources. In the case of extreme droughts, we considered the $q_{0.05}$ threshold before 2000 CE. Low-flow analysis since 1500 and the large deficit values for catchments (below 5th percentile) are shown in Table 4. In the 16th century, the years 1536, 1540 and 1590 are associated with significant runoff deficits. The event of 1540 has already been reported (Brázdil et al., 2013; Cook et al., 2015; Brázdil et al., 2019) as the worst event of the 16th century and more severe in terms of changing hydrologic conditions. In 1540, almost 90 % of the Rhine and Elbe River catchments (Basel and Cologne) experienced low yearly discharge, which ranked as the greatest low flows in the last 5 centuries (Leggewie and Mauelshagen, 2018). The seasonal precipitation was also deficient and was evident primarily in central Europe and England (Dobrovolný et al., 2010). Wetter and Pfister (2013) stated that the spring and summer of 1540 were likely to have been warmer than the comparable period during the 2003 drought. The simulation shows that the drought during 1540 was evident in most study catch-

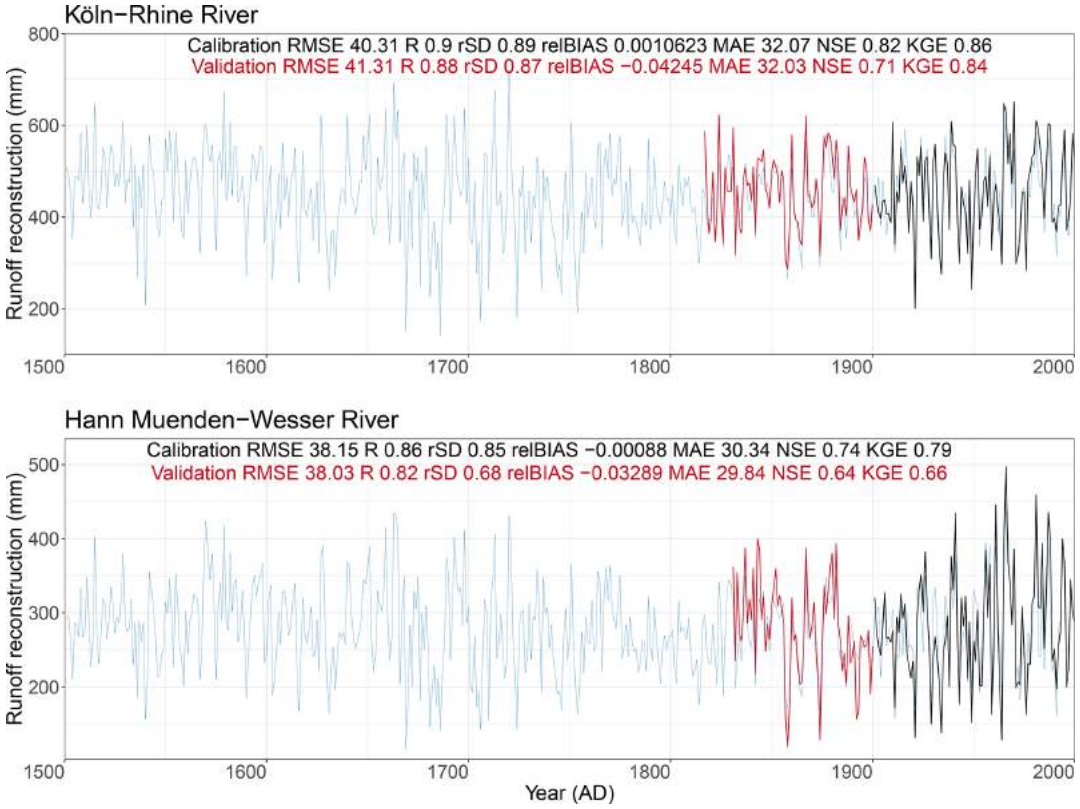


Figure 8. Reconstruction of runoff series for Köln–Main and Hann. Münden Wesser rivers. The blue line corresponds to the reconstructed series, and the black and red lines represent the observed runoff for the calibration and validation period, respectively.

ments, such as the Rhine, Main, Wesser, Loire and Danube, except Wasserburg–Inn.

In the 17th century, the years 1603, 1616, 1631, 1666, 1669, 1676, 1681, 1684 and 1686 were simulated as exceptionally low-flow years. Furthermore, two events (1669 and 1686) were associated with the largest water deficit across several study catchments. Basel Schifflaende–Rhine catchment is a good example of this, which appears to have experienced an extreme runoff deficit during 1669. In the Köln–Rhine catchment, 26 remarkable droughts have been captured over the past 500 years, and the year 1686 reached the largest runoff deficit ($156\,\mathrm{mm\,yr^{-1}}$). The 1616 is considered the driest year of the 17th century, the so-called "drought of the century" (Brázdil et al., 2013), which significantly impacted the major rivers in Europe (e.g. Rhine, Main and Wesser). Brázdil et al. (2018) identified three unusual drought periods (1540, 1616 and 1718–19) over the Czech lands, highlighting the 1616 drought, which caused widespread famine, dried up the Elbe river watershed and altered the climate of neighbouring nations (Switzerland and Germany). The hunger stone of the Elbe River also revealed the exceptionally dry year of 1616 (Brázdil et al., 2013). During the 18th century, a similar level of runoff deficit was simulated in the years 1706 and 1719.

During the 19th century, the years 1863, 1864, 1874, 1893 and 1899 were recognized as drought years in all catchments, while in the 20th century, the driest periods occurred in 1921, 1934, 1949 and 1976. The 1921 drought in the Blois–Loire, Rees–Rhine, Köln–Rhine, Orsova–Danube, Basel Rheinhalle–Rhine and Basel Schifflaende–Rhine catchments was ranked as the most exceptional drought in the 20th century. Three catchments (Basel Rheinhalle–Rhine, Basel Schifflaende–Rhine and Blois–Loire) exhibited a large runoff deficit during the year 1921. A noticeable increase in temperature was experienced across Europe, and certain areas were notably affected by a heatwave in July of that year. The majority of central Europe, southern England and Italy were affected by this drought, where the rainfall was found to have decreased around 50 % to 60 % relative to the average (Bonacina, 1923; Cook et al., 2015). The precipitation totals were recorded as the lowest since 1774, and the year was also ranked top (in terms of deficit rainfall) in the Great Alpine region (Haslinger and Blöschl, 2017), where the rainfall deficit began in winter 1920/21 and lasted until autumn 1921. Also reported in newspapers, the Rhine River (Switzerland), Molesey Weir on the Thames River (United Kingdom) and Loire River (France) all had low river flows in 1921 (Van der Schrier et al., 2021). Monthly runoff anomalies analysed from the GRUN dataset (Ghiggi et al., 2019)

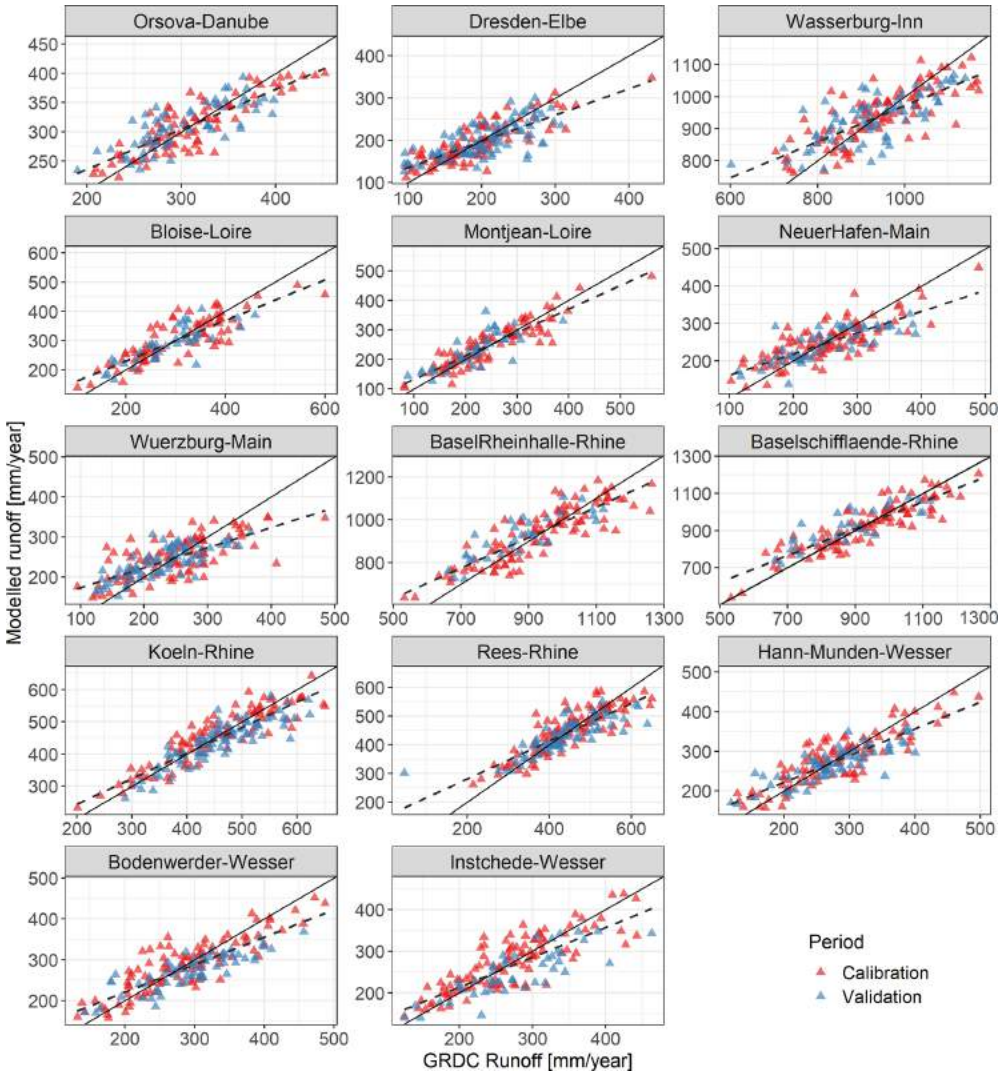


Figure 9. Observed and simulated runoff for 14 selected catchments in the calibration and validation periods. The solid line represents the 1 : 1 relation, and the dashed line corresponds to fitted regression between observed and simulated runoff.

show that August 1976 was the fifth driest month between 1900 and 2014, in agreement with some of our catchment reconstructions signalling the 1976 as a yearly drought in the Köln–Rhine, Hann. Münden–Wesser and Bodenwerder–Wesser.

In summary, the reconstructed annual runoff corresponded well to the majority of extreme drought years (e.g. 1540, 1616, 1669, 1710, 1724 and 1921, as highlighted in Table 4) and previously demonstrated in the OWDA-based PDSI tree-ring reconstructions and previous works (Dobrovolný et al., 2010; Brázdil et al., 2013; Wetter and Pfister, 2013; Cook et al., 2015; Markonis et al., 2018). It is important to note that the presented runoff reconstructions might have missed notably documented dry events, e.g. 1894 (Brodie, 1894), which was associated with unprecedented low levels of rainfall and excessive temperature rises in the south of England, the British Isles and other European regions (Brodie, 1894; Cook et al., 2015; Hanel et al., 2018).

Finally, we assessed the linear trends in the decadal runoff series for several time periods. The reconstructed annual runoff for 1500–2000 for each catchment was first aggregated to 10-year averages and divided by the mean annual runoff. The resulting series are shown in Fig. A2. Although significant negative trends were found for all catchments except for one considering the whole 1500–2000 period, the signal is not clearly linear. Instead, for a number of catchments, there is a period of sustained above-average (Orsova–Danube and Dresden–Elbe) or below-average (Blois and Montjean Loire) annual runoff during approx. 1600–1800, while for the rest, the persistence is weaker, although a low runoff signal is still visible (Basel Rheinhalle, Basel Schifflaende and Köln Rhine). When only the last 50 years is

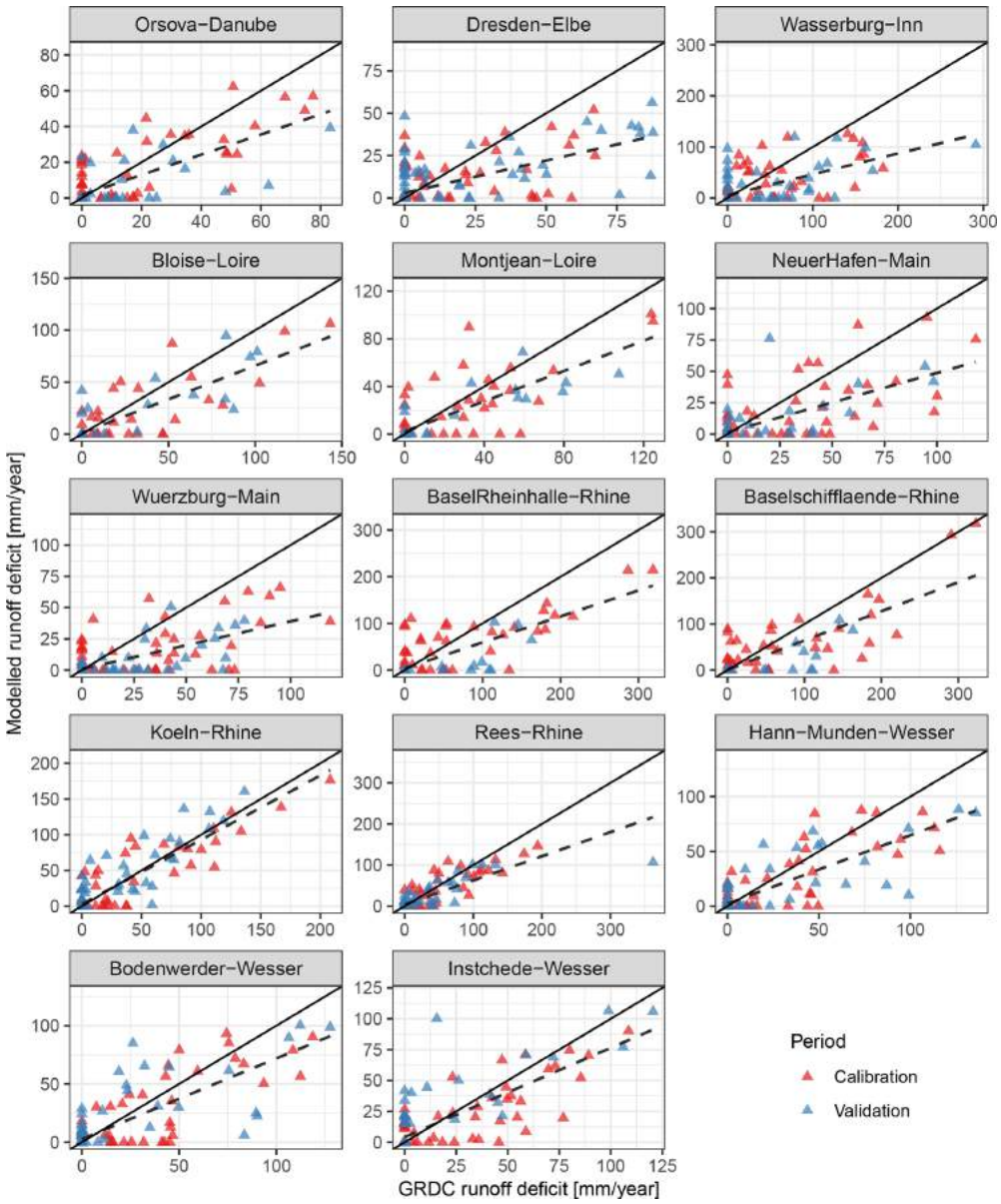


Figure 10. The observed and simulated runoff deficit based on the 33rd percentile threshold for 14 selected catchments during the calibration and validation period. The solid line represents the 1 : 1 relation, and the dashed line corresponds to fitted regression between observed and simulated runoff.

Table 3. Selection of best model for runoff in individual catchments.

Models	Catchments
BRNN [P, T]	Blois–Loire, Rees–Rhine
BRNN [P, T, PDSI]	Würzburg–Main and Orsova–Danube
BRNN [P, T, lag]	Montjean–Loire, Köln–Rhine, Hann. Münden–Wesser, Dresden–Elbe, Basel Rheinhalle–Rhine, Bodenwerder–Wesser, Wasserburg–Inn
LSTM [P, T, lag]	Neuer Hafen–Main, Intschede–Wesser
LSTM [P, T, PDSI]	Basel Schifflaende–Rhine

Table 4. Simulated runoff droughts since 1500. Years in bold indicate extreme droughts below 5 % quantile.

Station name	No of events	Simulated low-flow years	Largest deficit (year)
Orsova–Danube	12	1536, **1540**, **1669**, **1686**, 1704, 1706, **1710**, 1746, 1834, 1943, 1947, 1990	30.33 (1686)
Dresden–Elbe	1	**1669**	2.76 (1669)
Wasserburg–Inn	3	**1669**, **1686**, 1754	27.8 (1669)
Blois–Loire	17	**1540**, 1603, 1631, 1634, **1669**, 1676, **1686**, 1706, **1710**, **1724**,1736, 1754, 1766, 1884, **1921**, 1945, 1949	85.7 (1669)
Montjean–Loire	48	**1540**, 1603, 1607, **1616**, 1630, 1631, 1632, 1633, 1634, 1635, 1661, **1669**, 1670, 1676, 1680, 1681, 1684, 1685, **1686**, 1702, 1704, 1705, 1706, **1710**, 1715, 1717, 1718, 1723, **1724**, 1731, 1736, 1742, 1743, 1744, 1745, 1746, 1753, 1754, 1757, 1785, 1815, 1826, 1834, 1874, 1884, **1921**, 1945, 1949	105.2 (1686)
Neuer Hafen–Main	18	1590, **1616**, **1669**, 1681, 1682, **1686**, 1704, 1706, **1710**, **1724**, 1746, 1754, 1755, 1814, 1865, 1934, 1943, 1964	100.89 (1669)
Würzburg–Main	2	**1540**, **1669**	17.0 (1669)
Basel Rheinhalle–Rhine	21	1536, **1540**, 1590, 1603, **1616**, 1631, 1666, **1669**, 1676, 1681, **1686**, 1704, 1706, **1710**, **1724**, 1736, 1746, 1753, 1754, **1921**, 1949	133.9 (1669)
Basel Schifflaende–Rhine	19	1536, **1540**, 1590, 1603, **1616**, 1666, **1669**, 1676, 1681, 1684, **1686**, 1706, **1710**, **1724**, 1736, 1746, 1754, **1921**, 1949	563 (1669)
Köln–Rhine	28	1536, **1540**, 1590, 1603, **1616**, 1631, 1634, **1669**, 1676, 1681, 1684, **1686**, 1704, 1706, **1710**, **1724**, 1736, 1744, 1745, 1746, 1753, 1754, 1858, 1865, 1874, **1921**, 1949, 1976	157.6 (1686)
Rees–Rhine	18	1536, **1540**, 1603, 1631, 1666, **1669**, 1676, 1681, **1686**, 1704, 1706, **1710**, **1724**, 1736, 1746, 1754, **1921**, 1949	96.0 (1669)
Hann. Münden–Wesser	11	**1540**, **1669**, 1681, **1686**, 1706, **1710**, **1724**, 1911, 1934, 1976, 1991	46.6 (1669)
Bodenwerder–Wesser	15	**1540**, **1616**, 1631, **1669**, 1681, **1686**, 1706, **1710**, **1724**, 1754, 1858, 1874, 1911, 1934, 1976	56.3 (1669)
Intschede–Wesser	18	**1540**, **1616**, 1631, **1669**, 1670, 1676, 1681, 1685, **1686**, 1706, **1710**, 1754, 1814, 1857, 1858, 1865, 1934, 1959	134.4 (1669)

considered, the trends are significantly negative (positive) for seven (two) catchments, with the rest being insignificant.

5 Data availability

The annual runoff reconstructions were prepared using the defined dataset and can be accessed on the public repository Figshare (https://doi.org/10.6084/m9.figshare.15178107, Sadaf et al., 2021). The reconstructed data of precipitation and temperature can be downloaded at https://www.ncdc.noaa.gov/data-access/paleoclimatology-data (last access: 20 Feburary 2020). The monthly global historical climatological network (GHCN) data can be accessed via the link https://www1.ncdc.noaa.gov/pub/data/ghcn/ (last access: 12 May 2019). The data repositories of GRDC runoff are accessible to the public at https://www.bafg.de/GRDC/EN/Home/homepage_node.html (last access: 24 November 2016). All analyses and visualizations were done using R.

6 Conclusions

In this study, hydrological (GR1A) and two data-driven (BRNN and LSTM) models were used to reconstruct the annual runoff during the period 1500–2000, considering various input fields. After comprehensive validation of the simulated series, this work provides annual runoff time series for 14 catchments across Europe. The presented dataset can be used to investigate annual drought duration and severity. The main findings can be summarized as follows:

1. Data-driven methods have proven to be helpful for annual runoff simulations, even when there is high uncertainty in the forcing meteorological data. This contrasts with a conceptual lumped hydrological model, which would require bias correction before hydrological simulation.

2. There is no significant difference between the BRNN- and LSTM-simulated annual runoff, neither in terms of the individual values, nor in relation to the validation metrics.

3. Validation skill metrics suggest that for annual runoff prediction, it is beneficial to consider data-driven models that explicitly account for serial dependence, either through input data (e.g. time-lagged input fields) or directly in the model structure (e.g. LSTM networks).

4. The droughts identified in the reconstructed series correlate well with significant documented events (such as 1540, 1616, 1669, 1710, 1724 and 1921).

The reconstructed annual runoff relies heavily on the consistency of underlying reconstructed precipitation (Pauling et al., 2006) and temperature (Luterbacher et al., 2004) forcing fields. Unfortunately, those cannot be fully verified directly, due to the lack of sufficient long-term observational datasets. With the limited information provided by the GHCN station, we identified several notable deficiencies in the reconstructed forcings, in particular, underestimation of the variance in precipitation reconstruction. Moreover, proxy records that were used for the derivation of precipitation and temperature input fields are spatially heterogeneous, with some regions being better represented than others. This inevitably leads to poor performance over the latter. The skill of precipitation and temperature reconstructions across the selected catchments to derive annual runoff is still fairly good. In addition, the data-driven methods that were used in the paper were capable of removing systematic bias. We cannot be sure, though, that the link between reconstructed forcing and annual runoff is stationary when going back in time. Moreover, when the number of natural proxies included in the derivation of the forcing dataset decreases, the uncertainty increases. The reconstructed data should, therefore, always be considered with caution. Finally, since the runoff reconstruction is annual, dry summers can be compensated for by wet winters masking years with sub-annual dry periods. However, this should be regarded as a resolution- not methodology-related problem. Future research could consider further improvements of the simulations, e.g. by training a meta-model combining the runoff simulations from several fitted models. In addition, since interest is not often focused on the runoff series but on some other indicator (such as PDSI or deficit volume in the case of drought), it is also possible to simulate the drought indices directly, considering either the precipitation and temperature input fields or the simulated runoff. Finally, discrete classifiers (Kolachian and Saghafian, 2021) could also be used to simulate the drought (or water level) classes directly.

Appendix A

A1 Goodness-of-fit assessment

We used several statistical indicators to assess the skill of annual runoff reconstruction. In following definitions, p and o refer to the predicted and observed series, respectively, and i to year.

The standard deviation (SD) ratio (rSD; Ghiggi et al., 2021) is defined as

$$\mathrm{rSD} = \frac{\mathrm{SD_p}}{\mathrm{SD_o}}. \tag{A1}$$

The variability is underestimated when the value is less than 1 and overestimated when the value is greater than 1.

The root mean square error (RMSE; see, for example, Legates and McCabe, 1999)

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(p_i - o_i)^2}{n}} \tag{A2}$$

and mean absolute error (MAE; see e.g. Legates and McCabe, 1999)

$$\mathrm{MAE} = \frac{1}{n} \sum_{i=1}^{n} |(p_i - o_i)| \tag{A3}$$

measure how well predictions fit the observations. MAE and RMSE values can range from zero to infinity, with the former value indicating a perfect fit.

Pearson's correlation coefficient (R) is defined as

$$R = \frac{\sum_{i=1}^{n}(p_i - \overline{p})(o_i - \overline{o})}{\sqrt{\sum_{i=1}^{n}(p_i - \overline{p})^2}\sqrt{\sum_{i=1}^{n}(o_i - \overline{o})^2}}. \tag{A4}$$

The Nash–Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970),

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(p_i - o_i)^2}{\sum_{i=1}^{n}(o_i - \overline{o})^2}, \tag{A5}$$

is alternatively referred to as model efficiency. NSE = 1 corresponds to a perfect match between predicted and observed data, while a value less than 0 indicates that model predictions are on average less accurate than using the long-term mean of the observed time series $\overline{o}$.

Systematic errors can be detected using the absolute bias (BIAS)

$$\mathrm{BIAS} = \overline{p} - \overline{o} \tag{A6}$$

or relative bias (relBIAS)

$$\mathrm{relBIAS} = \frac{\overline{p} - \overline{o}}{\overline{o}}, \tag{A7}$$

which has an ideal value of 0. Positive bias values indicate that the model prediction overestimates observations, whereas negative values indicate underestimated model predictions.

The Kling–Gupta efficiency index (KGE; Gupta et al., 2009)

$$\mathrm{KGE} = 1 - \sqrt{(R-1)^2 + (\mathrm{rSD}-1)^2 + (\mathrm{relBIAS})^2} \tag{A8}$$

is calculated using three primary components, R, rSD and relBIAS, as defined above. relBIAS has a zero ideal value, while rSD and R have an ideal value of 1.

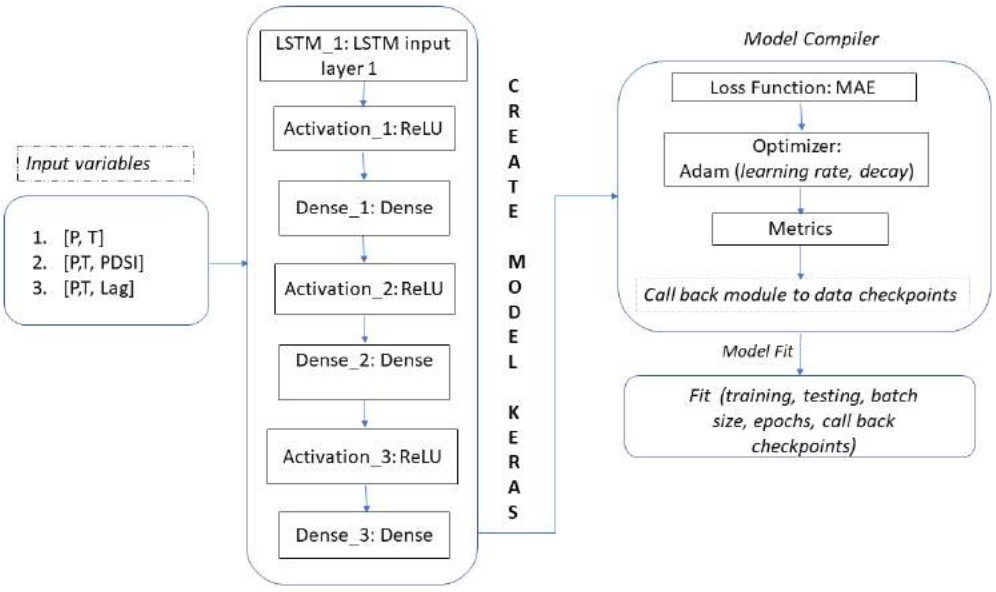


Figure A1. Structure of LSTM neural network model in a Keras environment for runoff predictions.

Table A1. Structure and hyperparameters of two data-driven models (BRNN and LSTM) for runoff predictions.

Training algorithms	Layer types	Activation functions	Hyperparameters
BRNN	input, hidden, output	$g_{k(x)} = \frac{\exp(2x)-1}{\exp(2x)+1}$	Tunelength 20, neurons (1–20)
LSTM	input, hidden, output	rectified linear activation (ReLU) $f(x) = \begin{cases} 0 & \text{when } x < 0 \\ x & \text{when } x \geq 0 \end{cases}$	Learning rate: 0.0001, epochs (30–200), units (5–150), batch input shapes: (1,1,2) for LSTM, (1,1,3) for LSTM [P, T, PDSI], (1,2,2) for LSTM [P, T, lag].

A2 Long short-term memory (LSTM)

To build the LSTM model, we use the Keras environment (Arnold, 2017) with its high-level application programming interface (API) for neural networks and Tensorflow (Abadi et al., 2016). Fig. A1 represents the structure of the LSTM neural model for the rainfall runoff relationship in several catchments. We design our network by stacking one LSTM and two dense layers on top of one other. As shown in Fig. A1, the model configured four distinct input combinations, each of which was normalized to [0, 1] in the training and testing phases. The model parameters choose different batch shapes, units (similar as neurons) and epochs as described in Table A1. The model considers the rectified linear unit (ReLU), using component wise multiplication and defining the dropout parameter as 0.1. According to Kingma and Ba (2014), the optimization algorithm plays a significant role in the algorithm's convergence and optimization. For this reason, Adam's optimizer is considered, as it performs stochastic gradient descent (SGD) more efficiently using the backpropagation algorithm. During compilation, the learning rate is set to 0.001 or 0.002, and the mean square error (MSE) is used to measure model accuracy. In addition, the mean absolute error (MAE) is used as an objective to minimize residues and achieve optimum value. Model checkpoints are used to save the model having minimum loss during the training with minimum loss and better accuracy.

A3 Bayesian regularized neural network (BRNN)

BRNN is a probabilistic technique for handling non-linear problems. Using the caret package, the model "brnn" was designed to work with a two-layer network as described by MacKay (1992) and Foresee and Hagan (1997). BRNN uses the Nguyen and Widrow algorithm to assign initial weights and the Gauss–Newton algorithm to optimize. Model is first trained on the training dataset, and its performance is checked by making a prediction on the testing dataset.

While selecting a model for train control, a simple boot resampling strategy was applied to evaluate performance. We tested the proposed model's predictive ability using a random bootstrap generator, with 75 % of the observations in the training set and 25 % in the testing set. RMSE was utilized as a loss function to compile and verify the model's accuracy. The model was fitted with 20 neurons, one hidden layer and implemented activation function $g_k(x) = \frac{\exp(2x)-1}{\exp(2x)+1}$. Af-

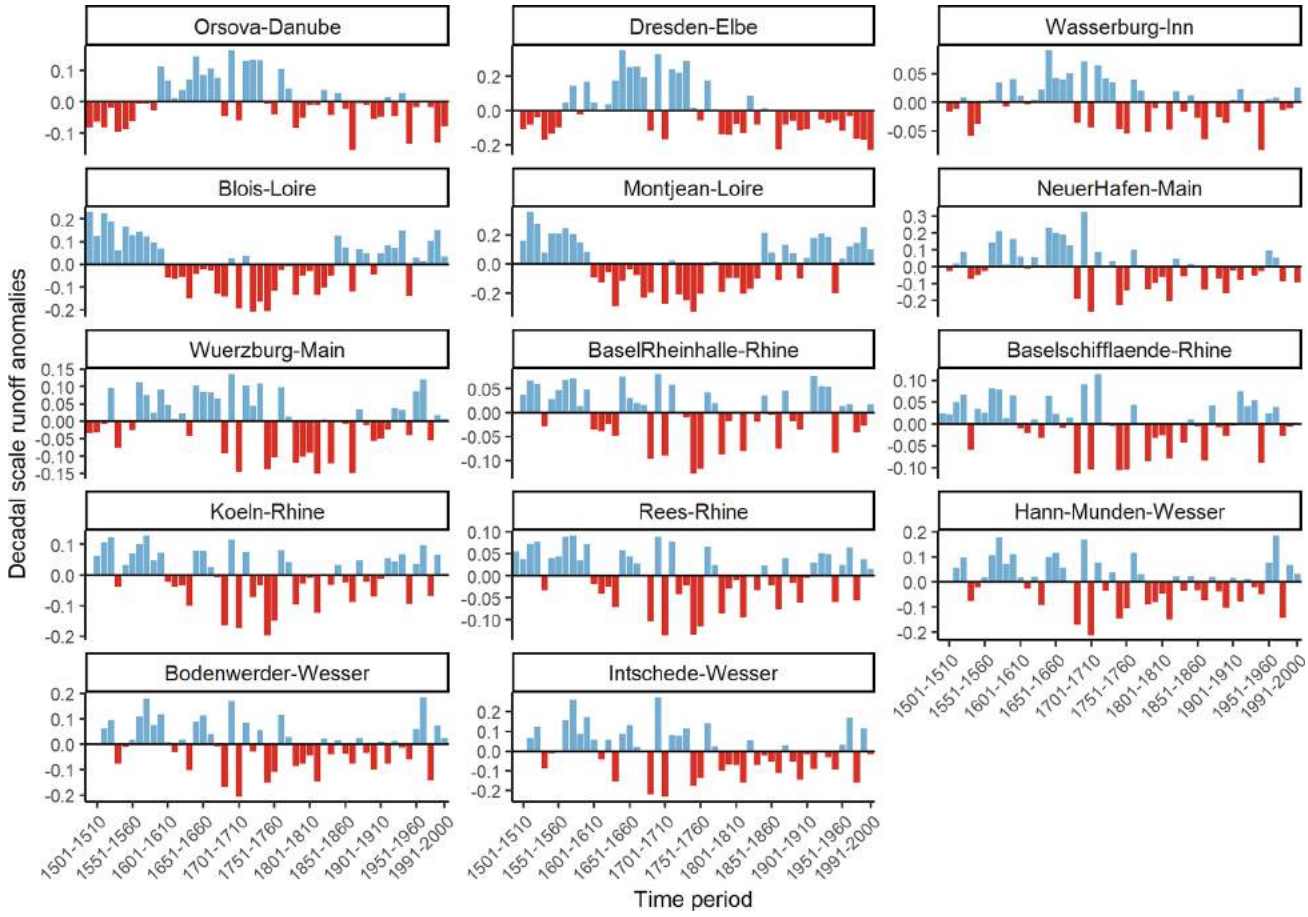


Figure A2. Decadal fluctuation of runoff anomalies in selected catchments over the past 500 years.

ter compilation, the train function automatically selected the best model with the smallest RMSE as the final model.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/essd-14-4035-2022-supplement.

Author contributions. The study was initially designed by RK, MH and YM. Algorithms are coded with the assistance of YM, US and MH. Datasets were collected by MRVG and SN. The research was carried out by SN, MS and MH, who also wrote the paper. OR and RK both helped to revise the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. This work was carried out within the bilateral project XEROS (eXtreme EuROpean drOughtS: multimodel synthesis of past, present and future events). We thank the Global Runoff Data Centre (GRDC) for providing the observed runoff data. We would also like to thank the editor, Christof Lorenz (KIT, Germany), Gionata Ghiggi (EPFL, Switzerland) and the anonymous reviewers for their insightful remarks, which improved the overall quality of the article.

Financial support. This research has been supported by the Grantová Agentura České Republiky (grant no. 1924089J), the Deutsche Forschungsgemeinschaft (grant no. RA 3235/11) and the Fakulta Životního Prostředí, Česká Zemědělská Univerzita v Praze (grant no. 2020B0018).

Review statement. This paper was edited by Christof Lorenz and reviewed by Gionata Ghiggi and two anonymous referees.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moorea, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., and Google brain: Tensorflow: A system for large-scale machine learning, in: 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16, 2–4 November 2016, Savannah, GA, USA, pp. 265–283, 2016.

Armstrong, M. S., Kiem, A. S., and Vance, T. R.: Comparing instrumental, palaeoclimate, and projected rainfall data: Implications for water resources management and hydrological modelling, Journal of Hydrology: Regional Studies, 31, 100728, https://doi.org/10.1016/j.ejrh.2020.100728, 2020.

Arnold, T. B.: kerasR: R interface to the keras deep learning library, Journal of Open Source Software, 2, 296, https://doi.org/10.21105/joss.00296, 2017.

Ayzel, G., Kurochkina, L., and Zhuravlev, S.: The influence of regional hydrometric data incorporation on the accuracy of gridded reconstruction of monthly runoff, Hydrol. Sci. J., 0, https://doi.org/10.1080/02626667.2020.1762886, 1–12, 2020.

Boch, R. and Spötl, C.: Reconstructing palaeoprecipitation from an active cave flowstone, J. Quaternary Sci., 26, 675–687, 2011.

Bonacina, L.: The European drought of 1921, Nature, 112, 488–489, 1923.

Brázdil, R. and Dobrovolný, P.: Historical climate in Central Europe during the last 500 years, The Polish Climate in the European Context: An Historical Overview, Springer, Dordrecht, the Netherlands, p. 41, 2009.

Brázdil, R., Dobrovolný, P., Trnka, M., Kotyza, O., Řezníčková, L., Valášek, H., Zahradníček, P., and Štěpánek, P.: Droughts in the Czech Lands, 1090–2012 AD, Clim. Past, 9, 1985–2002, https://doi.org/10.5194/cp-9-1985-2013, 2013.

Brázdil, R., Kiss, A., Luterbacher, J., Nash, D. J., and Řezníčková, L.: Documentary data and the study of past droughts: a global state of the art, Clim. Past, 14, 1915–1960, https://doi.org/10.5194/cp-14-1915-2018, 2018.

Brázdil, R., Demarée, G. R., Kiss, A., Dobrovolný, P., Chromá, K., Trnka, M., Dolák, L., Řezníčková, L., Zahradníček, P., Limanowka, D., and Jourdain, S.: The extreme drought of 1842 in Europe as described by both documentary data and instrumental measurements, Clim. Past, 15, 1861–1884, https://doi.org/10.5194/cp-15-1861-2019, 2019.

Brodie, F. J.: The great drought of 1893, and its attendant meteorological phenomena, Q. J. Roy. Meteor. Soc., 20, 1–30, 1894.

Büntgen, U., Frank, D. C., Nievergelt, D., and Esper, J.: Summer temperature variations in the European Alps, AD 755–2004, J. Climate, 19, 5606–5623, 2006.

Büntgen, U., Franke, J., Frank, D., Wilson, R., González-Rouco, F., and Esper, J.: Assessing the spatial signature of European climate reconstructions, Clim. Res., 41, 125–130, 2010.

Caillouet, L., Vidal, J.-P., Sauquet, E., Devers, A., and Graff, B.: Ensemble reconstruction of spatio-temporal extreme low-flow events in France since 1871, Hydrol. Earth Syst. Sci., 21, 2923–2951, https://doi.org/10.5194/hess-21-2923-2017, 2017.

Casas-Gómez, P., Sánchez-Salguero, R., Ribera, P., and Linares, J. C.: Contrasting Signals of the Westerly Index and North Atlantic Oscillation over the Drought Sensitivity of Tree-Ring Chronologies from the Mediterranean Basin, Atmosphere, 11, 644, https://doi.org/10.3390/atmos11060644, 2020.

Casty, C., Wanner, H., Luterbacher, J., Esper, J., and Böhm, R.: Temperature and precipitation variability in the European Alps since 1500, Int. J. Climatol., 25, 1855–1880, 2005.

Chen, X., Huang, J., Han, Z., Gao, H., Liu, M., Li, Z., Liu, X., Li, Q., Qi, H., and Huang, Y.: The importance of short lag-time in the runoff forecasting model based on long short-term memory, J. Hydrol., 589, 125359, https://doi.org/10.1016/j.jhydrol.2020.125359, 2020.

Contreras, P., Orellana-Alvear, J., Muñoz, P., Bendix, J., and Célleri, R.: Influence of Random Forest Hyperparameterization on Short-Term Runoff Forecasting in an Andean Mountain Catchment, Atmosphere, 12, 238, https://doi.org/10.3390/atmos12020238, 2021.

Cook, E. R., Seager, R., Kushnir, Y., Briffa, K. R., Büntgen, U., Frank, D., Krusic, P. J., Tegel, W., van der Schrier, G., Andreu-Hayles, L., Baillie, M., Baittinger, C., Bleicher, N., Bonde, N., Brown, D., Carrer, M., Cooper, R., Čufar, K., Dittmar, C., Esper, J., Griggs, C., Gunnarson, B., Günther, B., Gutierrez, E., Haneca, K., Helama, S., Herzig, F., Heussner, K. U., Hofmann, J., Janda, P., Kontic, R., Köse, N., Kyncl, T., Levanič, T., Linderholm, H., Manning, S., Melvin, T. M., Miles, D., Neuwirth, B., Nicolussi, K., Nola, P., Panayotov, M., Popa, I., Rothe, A., Seftigen, K., Seim, A., Svarva, H., Svoboda, M., Thun, T., Timonen, M., Touchan, R., Trotsiuk, V., Trouet, V., Walder, F., Ważny, T., Wilson, R., and Zang, C.: Old World megadroughts and pluvials during the Common Era, Science Advances, 1, e1500561, https://doi.org/10.1126/sciadv.1500561, 2015.

Coron, L., Thirel, G., Delaigue, O., Perrin, C., and Andréassian, V.: The suite of lumped GR hydrological models in an R package, Environ. Modell. Softw., 94, 166–171, 2017.

Dobrovolný, P., Moberg, A., Brázdil, R., Pfister, C., Glaser, R., Wilson, R., van Engelen, A., Limanówka, D., Kiss, A., Halíčková, M., Macková, J., Riemann, D., Luterbacher, J., and Böhm, R.: Monthly, seasonal and annual temperature reconstructions for Central Europe derived from documentary evidence and instrumental records since AD 1500, Climatic Change, 101, 69–107, 2010.

Emile-Geay, J., McKay, N. P., Kaufman, D. S., Von Gunten, L., Wang, J., Anchukaitis, K. J., Abram, N. J., Addison, J. A., Curran, M. A., Evans, M. N., Henley, B. J., Hao, Z., Martrat, B., McGregor, H. V., Neukom, R., Pederson, G. T., Stenni, B., Thirumalai, K., Werner, J. P., Xu, C., Divine, D. V., Dixon, B. C., Gergis, J., Mundo, I. A., Nakatsuka, T., Phipps, S. J., Routson, C. C., Steig, E. J., Tierney, J. E., Tyler, J. J., Allen, K. J., Bertler, N. A. N., Björklund, J., Chase, B. M., Chen, M.-T., Cook, E., de Jong, R., DeLong, K. L., Dixon, D. A., Ekaykin, A. A., Ersek, V., Filipsson, H. L., Francus, P., Freund, M. B., Frezzotti, M., Gaire, N. P., Gajewski, K., Ge, Q., Goosse, H., Gornostaeva, A., Grosjean, M., Horiuchi, K., Hormes, A., Husum, K., Isaksson, E., Kandasamy, S., Kawamura, K., Kilbourne, K. H., Koç, N., Leduc, G., Linderholm, H. W., Lorrey, A. M., Mikhalenko, V., Mortyn, P. G., Motoyama, H., Moy, A. D., Mulvaney, R., Munz, P. M., Nash, D. J., Oerter, H., Opel, T., Orsi, A. J., Ovchinnikov, D. V., Porter, T. J., Roop, H. A., Saenger, C., Sano, M., Sauchyn, D., Saunders, K. M., Seidenkrantz, M.-S., Severi, M., Shao, X., Sicre, M.-A., Sigl, M., Sinclair, K., St. George, S., St. Jacques, J.-M., Thamban, M., Kuwar Thapa, U., Thomas, E.

R., Turney, C., Uemura, R., Viau, A. E., Vladimirova, D. O., Wahl, E. R., White, J. W. C., Yu, Z., Zinke, J., and PAGES2k Consortium: A global multiproxy database for temperature reconstructions of the Common Era, Scientific Data, 4, 170088, https://doi.org/10.1038/sdata.2017.88, 2017.

Fathi, M. M., Awadallah, A. G., Abdelbaki, A. M., and Haggag, M.: A new Budyko framework extension using time series SARIMAX model, J. Hydrol., 570, 827–838, 2019.

Fekete, B. M., Vörösmarty, C. J., and Grabs, W.: Global, composite runoff fields based on observed river discharge and simulated water balances, Tech. Rep. 22, Global Runoff Data Centre, Koblenz, Germany, 1999.

Foresee, F. D. and Hagan, M. T.: Gauss-Newton approximation to Bayesian learning, in: Proceedings of international conference on neural networks (ICNN'97), vol. 3, pp. 1930–1935, IEEE, Houston, TX, USA, 1997.

Ghiggi, G., Humphrey, V., Seneviratne, S. I., and Gudmundsson, L.: GRUN: an observation-based global gridded runoff dataset from 1902 to 2014, Earth Syst. Sci. Data, 11, 1655–1674, https://doi.org/10.5194/essd-11-1655-2019, 2019.

Ghiggi, G., Humphrey, V., Seneviratne, S., and Gudmundsson, L.: G-RUN ENSEMBLE: A Multi-Forcing Observation-Based Global Runoff Reanalysis, Water Resour. Res., 57, e2020WR028787, https://doi.org/10.1029/2020WR028787, 2021.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hanel, M., Rakovec, O., Markonis, Y., Máca, P., Samaniego, L., Kysel, J., and Kumar, R.: Revisiting the recent European droughts from a long-term perspective, Scientific Reports, 8, 1–11, 2018.

Hansson, D., Eriksson, C., Omstedt, A., and Chen, D.: Reconstruction of river runoff to the Baltic Sea, AD 1500–1995, Int. J. Climatol., 31, 696–703, 2011.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations–the CRU TS3.10 Dataset, Int. J. Climatol., 34, 623–642, 2014.

Haslinger, K. and Blöschl, G.: Space-time patterns of meteorological drought events in the European Greater Alpine Region over the past 210 years, Water Resour. Res., 53, 9807–9823, 2017.

Hochreiter, S. and Schmidhuber, J.: Long short-term memory, Neural Comput., 9, 1735–1780, 1997.

Hu, C., Wu, Q., Li, H., Jian, S., Li, N., and Lou, Z.: Deep learning with a long short-term memory networks approach for rainfall-runoff simulation, Water, 10, 1543, https://doi.org/10.3390/w10111543, 2018.

Im, S., Kim, H., Kim, C., and Jang, C.: Assessing the impacts of land use changes on watershed hydrology using MIKE SHE, Environ. Geol., 57, 231, https://doi.org/10.1007/s00254-008-1303-3, 2009.

Ionita, M., Tallaksen, L. M., Kingston, D. G., Stagge, J. H., Laaha, G., Van Lanen, H. A. J., Scholz, P., Chelcea, S. M., and Haslinger, K.: The European 2015 drought from a climatological perspective, Hydrol. Earth Syst. Sci., 21, 1397–1419, https://doi.org/10.5194/hess-21-1397-2017, 2017.

Jeong, J., Barichivich, J., Peylin, P., Haverd, V., McGrath, M. J., Vuichard, N., Evans, M. N., Babst, F., and Luyssaert, S.: Using the International Tree-Ring Data Bank (ITRDB) records as century-long benchmarks for global land-surface models, Geosci. Model Dev., 14, 5891–5913, https://doi.org/10.5194/gmd-14-5891-2021, 2021.

Ji, Y., Dong, H.-T., Xing, Z.-X., Sun, M.-X., Fu, Q., and Liu, D.: Application of the decomposition-prediction-reconstruction framework to medium-and long-term runoff forecasting, Water Supply, 21, 696–709, 2021.

Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, arXiv [preprint], arXiv:1412.6980, 2014.

Kolachian, R. and Saghafian, B.: Hydrological drought class early warning using support vector machines and rough sets, Environ. Earth Sci., 80, 1–15, 2021.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M.: Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks, Hydrol. Earth Syst. Sci., 22, 6005–6022, https://doi.org/10.5194/hess-22-6005-2018, 2018.

Kress, A., Saurer, M., Siegwolf, R. T., Frank, D. C., Esper, J., and Bugmann, H.: A 350 year drought reconstruction from Alpine tree ring stable isotopes, Global Biogeochem. Cy., 24, 1–16, https://doi.org/10.1029/2009GB003613, 2010.

Kress, A., Hangartner, S., Bugmann, H., Büntgen, U., Frank, D. C., Leuenberger, M., Siegwolf, R. T., and Saurer, M.: Swiss tree rings reveal warm and wet summers during medieval times, Geophys. Res. Lett., 41, 1732–1737, 2014.

Krysanova, V., Vetter, T., and Hattermann, F.: Detection of change in drought frequency in the Elbe basin: comparison of three methods, Hydrol. Sci. J., 53, 519–537, 2008.

Kuhn, M.: Caret: classification and regression training, Astrophysics Source Code Library, https://ui.adsabs.harvard.edu/abs/2015ascl.soft05003K (last access: 21 December 2021), pp. ascl–1505, 2015.

Kwak, J., Lee, J., Jung, J., and Kim, H. S.: Case Study: Reconstruction of Runoff Series of Hydrological Stations in the Nakdong River, Korea, Water, 12, 3461, https://doi.org/10.3390/w12123461, 2020.

Laaha, G., Gauster, T., Tallaksen, L. M., Vidal, J.-P., Stahl, K., Prudhomme, C., Heudorfer, B., Vlnas, R., Ionita, M., Van Lanen, H. A. J., Adler, M.-J., Caillouet, L., Delus, C., Fendekova, M., Gailliez, S., Hannaford, J., Kingston, D., Van Loon, A. F., Mediero, L., Osuch, M., Romanowicz, R., Sauquet, E., Stagge, J. H., and Wong, W. K.: The European 2015 drought from a hydrological perspective, Hydrol. Earth Syst. Sci., 21, 3001–3024, https://doi.org/10.5194/hess-21-3001-2017, 2017.

Legates, D. R. and McCabe Jr., G. J.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35, 233–241, 1999.

Leggewie, C. and Mauelshagen, F.: Climate change and cultural transition in Europe, Brill, Leiden, the Netherlands, 2018.

Li, Y., Wei, J., Wang, D., Li, B., Huang, H., Xu, B., and Xu, Y.: A Medium and Long-Term Runoff Forecast Method Based on Massive Meteorological Data and Machine Learning Algorithms, Water, 13, 1308, https://doi.org/10.3390/w13091308, 2021.

Ljungqvist, F. C., Piermattei, A., Seim, A., Krusic, P. J., Büntgen, U., He, M., Kirdyanov, A. V., Luterbacher, J., Schneider, L., Seftigen, K., Stahle, D. W., Villalba, R., Yang, B., and Esper, J.: Ranking of tree-ring based hydroclimate reconstructions of the past millennium, Quaternary Sci. Rev., 230, 106074, https://doi.org/10.1016/j.quascirev.2019.106074, 2020.

Luoto, T. P. and Nevalainen, L.: Quantifying climate changes of the Common Era for Finland, Clim. Dynam., 49, 2557–2567, 2017.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., and Wanner, H.: European seasonal and annual temperature variability, trends, and extremes since 1500, Science, 303, 1499–1503, 2004.

MacKay, D. J.: A practical Bayesian framework for backpropagation networks, Neural Comput., 4, 448–472, 1992.

Manabe, S.: Climate and the ocean circulation: I. The atmospheric circulation and the hydrology of the earth's surface, Mon. Weather Rev., 97, 739–774, 1969.

Markonis, Y. and Koutsoyiannis, D.: Scale-dependence of persistence in precipitation records, Nat. Clim. Change, 6, 399–401, 2016.

Markonis, Y., Hanel, M., Máca, P., Kysel, J., and Cook, E.: Persistent multi-scale fluctuations shift European hydroclimate to its millennial boundaries, Nat. Commun., 9, 1–12, 2018.

Martínez-Sifuentes, A. R., Villanueva-Díaz, J., and Estrada-Ávalos, J.: Runoff reconstruction and climatic influence with tree rings, in the Mayo river basin, Sonora, Mexico, iForest, 13, 98, https://doi.org/10.3832/ifor3190-013, 2020.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G.: An overview of the global historical climatology network-daily database, J. Atmos. Ocean. Tech., 29, 897–910, 2012.

Menne, M. J., Williams, C. N., Gleason, B. E., Rennie, J. J., and Lawrimore, J. H.: The global historical climatology network monthly temperature dataset, version 4, J. Climate, 31, 9835–9854, 2018.

Middelkoop, H., Daamen, K., Gellens, D., Grabs, W., Kwadijk, J. C., Lang, H., Parmet, B. W., Schädler, B., Schulla, J., and Wilke, K.: Impact of climate change on hydrological regimes and water resources management in the Rhine basin, Climatic Change, 49, 105–128, 2001.

Moberg, A., Mohammad, R., and Mauritsen, T.: Analysis of the Moberg et al. (2005) hemispheric temperature reconstruction, Clim. Dynam., 31, 957–971, 2008.

Moravec, V., Markonis, Y., Rakovec, O., Kumar, R., and Hanel, M.: A 250-year European drought inventory derived from ensemble hydrologic modeling, Geophys. Res. Lett., 46, 5909–5917, 2019.

Mouelhi, S., Michel, C., Perrin, C., and Andréassian, V.: Linking stream flow to rainfall at the annual time step: the Manabe bucket model revisited, J. Hydrol., 328, 283–296, 2006.

Murphy, C., Broderick, C., Burt, T. P., Curley, M., Duffy, C., Hall, J., Harrigan, S., Matthews, T. K. R., Macdonald, N., McCarthy, G., McCarthy, M. P., Mullan, D., Noone, S., Osborn, T. J., Ryan, C., Sweeney, J., Thorne, P. W., Walsh, S., and Wilby, R. L.: A 305-year continuous monthly rainfall series for the island of Ireland (1711–2016), Clim. Past, 14, 413–440, https://doi.org/10.5194/cp-14-413-2018, 2018.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I–A discussion of principles, J. Hydrol., 10, 282–290, 1970.

Nicault, A., Alleaume, S., Brewer, S., Carrer, M., Nola, P., and Guiot, J.: Mediterranean drought fluctuation during the last 500 years based on tree-ring data, Clim. Dynam., 31, 227–245, 2008.

Okut, H.: Bayesian regularized neural networks for small n big p data, in: Artificial neural networks-models and applications, in: Artificial Neural Networks, edited by: Rosa, J. L. G., IntechOpen, 21–23, https://doi.org/10.5772/63256, 2016.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2–Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling, J. Hydrol., 303, 290–306, 2005.

Pauling, A., Luterbacher, J., Casty, C., and Wanner, H.: Five hundred years of gridded high-resolution precipitation reconstructions over Europe and the connection to large-scale circulation, Clim. Dynam., 26, 387–405, 2006.

Peterson, T. C. and Vose, R. S.: An overview of the Global Historical Climatology Network temperature database, B. Am. Meteorol. Soc., 78, 2837–2850, 1997.

Pfister, C., Brázdil, R., Glaser, R., Barriendos, M., Camuffo, D., Deutsch, M., Dobrovolný, P., Enzi, S., Guidoboni, E., Kotyza, O., Militzer, S., Rácz, L., and Rodrigo, F. S.: Documentary evidence on climate in sixteenth-century Europe, Climatic Change, 43, 55–110, 1999.

Pfister, C., Weingartner, R., and Luterbacher, J.: Hydrological winter droughts over the last 450 years in the Upper Rhine basin: a methodological approach, Hydrol. Sci. J., 51, 966–985, 2006.

Proctor, C., Baker, A., Barnes, W., and Gilmour, M.: A thousand year speleothem proxy record of North Atlantic climate from Scotland, Clim. Dynam., 16, 815–820, 2000.

Quayle, R. G., Peterson, T. C., Basist, A. N., and Godfrey, C. S.: An operational near-real-time global temperature index, Geophys. Res. Lett., 26, 333–335, 1999.

Reinecke, R., Müller Schmied, H., Trautmann, T., Andersen, L. S., Burek, P., Flörke, M., Gosling, S. N., Grillakis, M., Hanasaki, N., Koutroulis, A., Pokhrel, Y., Thiery, W., Wada, Y., Yusuke, S., and Döll, P.: Uncertainty of simulated groundwater recharge at different global warming levels: a global-scale multi-model ensemble study, Hydrol. Earth Syst. Sci., 25, 787–810, https://doi.org/10.5194/hess-25-787-2021, 2021.

Rivera, J. A., Araneo, D. C., and Penalba, O. C.: Threshold level approach for streamflow drought analysis in the Central Andes of Argentina: a climatological assessment, Hydrol. Sci. J., 62, 1949–1964, 2017.

Sadaf, N., Součková, M., Godoy, M. R. V., Singh, U., Markonis, Y., Kumar, R., Rakovec, O., and Hanel, M.: Supporting data for A 500-year runoff reconstruction for European catchments, figshare [data set], https://doi.org/10.6084/m9.figshare.15178107, 2021.

Seiller, G., Anctil, F., and Perrin, C.: Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions, Hydrol. Earth Syst. Sci., 16, 1171–1189, https://doi.org/10.5194/hess-16-1171-2012, 2012.

Senthil Kumar, A., Sudheer, K., Jain, S., and Agarwal, P.: Rainfall-runoff modelling using artificial neural networks: comparison of network types, Hydrol. Process., 19, 1277–1291, 2005.

Smith, K. A., Barker, L. J., Tanguy, M., Parry, S., Harrigan, S., Legg, T. P., Prudhomme, C., and Hannaford, J.: A multi-objective ensemble approach to hydrological modelling in the UK: an application to historic drought reconstruction, Hydrol. Earth Syst. Sci., 23, 3247–3268, https://doi.org/10.5194/hess-23-3247-2019, 2019.

Su, W., Tao, J., Wang, J., and Ding, C.: Current research status of large river systems: a cross-continental comparison, Environ. Sci. Pollut. R., 27, 39413–39426, 2020.

Sun, J., Liu, Y., Wang, Y., Bao, G., and Sun, B.: Tree-ring based runoff reconstruction of the upper Fenhe River basin, North China, since 1799 AD, Quatern. Int., 283, 117–124, 2013.

Sung, J. H. and Chung, E.-S.: Development of streamflow drought severity–duration–frequency curves using the threshold level method, Hydrol. Earth Syst. Sci., 18, 3341–3351, https://doi.org/10.5194/hess-18-3341-2014, 2014.

Swierczynski, T., Brauer, A., Lauterbach, S., Martín-Puertas, C., Dulski, P., von Grafenstein, U., and Rohr, C.: A 1600 yr seasonally resolved record of decadal-scale flood variability from the Austrian Pre-Alps, Geology, 40, 1047–1050, 2012.

Tejedor, E., de Luis, M., Cuadrat, J. M., Esper, J., and Saz, M. Á.: Tree-ring-based drought reconstruction in the Iberian Range (east of Spain) since 1694, Int. J. Biometeorol., 60, 361–372, 2016.

Trouet, V., Diaz, H., Wahl, E., Viau, A., Graham, R., Graham, N., and Cook, E.: A 1500-year reconstruction of annual mean temperature for temperate North America on decadal-to-multidecadal time scales, Environ. Res. Lett., 8, 024008, 2013.

Tshimanga, R., Hughes, D., and Kapangaziwiri, E.: Initial calibration of a semi-distributed rainfall runoff model for the Congo River basin, Phys. Chem. Earth Pt. A/B/C, 36, 761–774, 2011.

Uehlinger, U. F., Wantzen, K. M., Leuven, R. S., and Arndt, H.: The Rhine river basin, in: Rivers of Europe, edited by: Tockner, K., Academic Press, London, ISBN 978-0-12-369449-2, 2009.

van der Schrier, G., Allan, R. P., Ossó, A., Sousa, P. M., Van de Vyver, H., Van Schaeybroeck, B., Coscarelli, R., Pasqua, A. A., Petrucci, O., Curley, M., Mietus, M., Filipiak, J., Štěpánek, P., Zahradníček, P., Brázdil, R., Řezníčková, L., van den Besselaar, E. J. M., Trigo, R., and Aguilar, E.: The 1921 European drought: impacts, reconstruction and drivers, Clim. Past, 17, 2201–2221, https://doi.org/10.5194/cp-17-2201-2021, 2021.

Van Houdt, G., Mosquera, C., and Nápoles, G.: A review on the long short-term memory model, Artif. Intell. Rev., 53, 5929–5955, 2020.

Vansteenberge, S., Verheyden, S., Cheng, H., Edwards, R. L., Keppens, E., and Claeys, P.: Paleoclimate in continental northwestern Europe during the Eemian and early Weichselian (125–97 ka): insights from a Belgian speleothem, Clim. Past, 12, 1445–1458, https://doi.org/10.5194/cp-12-1445-2016, 2016.

Wang, W., Gelder, P. H. V., and Vrijling, J.: Comparing Bayesian regularization and cross-validated early-stopping for streamflow forecasting with ANN models, IAHS Publications-Series of Proceedings and Reports, 311, 216–221, 2007.

Werbos, P. J.: Backpropagation through time: what it does and how to do it, Proceedings of the IEEE, 78, 1550–1560, 1990.

Wetter, O. and Pfister, C.: An underestimated record breaking event – why summer 1540 was likely warmer than 2003, Clim. Past, 9, 41–56, https://doi.org/10.5194/cp-9-41-2013, 2013.

Wetter, O., Pfister, C., Weingartner, R., Luterbacher, J., Reist, T., and Trösch, J.: The largest floods in the High Rhine basin since 1268 assessed from documentary and instrumental evidence, Hydrol. Sci. J, 56, 733–758, 2011.

Wilhelm, B., Arnaud, F., Sabatier, P., Crouzet, C., Brisset, E., Chaumillon, E., Disnar, J.-R., Guiter, F., Malet, E., Reyss, J.-L., Tachikawa, K., Bard, E., and Delannoy, J.-J.: 1400 years of extreme precipitation patterns over the Mediterranean French Alps and possible forcing mechanisms, Quaternary Res., 78, 1–12, 2012.

Wilson, R. J., Luckman, B. H., and Esper, J.: A 500 year dendroclimatic reconstruction of spring–summer precipitation from the lower Bavarian Forest region, Germany, Int. J. Climatol., 25, 611–630, 2005.

Xiang, Z., Yan, J., and Demir, I.: A rainfall-runoff model with LSTM-based sequence-to-sequence learning, Water Resour. Res., 56, e2019WR025326, https://doi.org/10.1029/2019WR025326, 2020.

Xoplaki, E., Luterbacher, J., Paeth, H., Dietrich, D., Steiner, N., Grosjean, M., and Wanner, H.: European spring and autumn temperature variability and change of extremes over the last half millennium, Geophys. Res. Lett., 32, L15713, https://doi.org/10.1029/2005GL023424, 2005.

Ye, L., Jabbar, S. F., Abdul Zahra, M. M., and Tan, M. L.: Bayesian Regularized Neural Network Model Development for Predicting Daily Rainfall from Sea Level Pressure Data: Investigation on Solving Complex Hydrology Problem, Complexity, 2021, 6631564, https://doi.org/10.1155/2021/6631564, 2021.

Yevjevich, V. M.: Objective approach to definitions and investigations of continental hydrologic droughts, An, Hydrology papers, Colorado State University. Libraries, 23, 1967.

Zappa, M. and Kan, C.: Extreme heat and runoff extremes in the Swiss Alps, Nat. Hazards Earth Syst. Sci., 7, 375–389, https://doi.org/10.5194/nhess-7-375-2007, 2007.

Zhang, X., Liang, F., Yu, B., and Zong, Z.: Explicitly integrating parameter, input, and structure uncertainties into Bayesian Neural Networks for probabilistic hydrologic forecasting, J. Hydrol., 409, 696–709, 2011.

Zuo, G., Luo, J., Wang, N., Lian, Y., and He, X.: Two-stage variational mode decomposition and support vector regression for streamflow forecasting, Hydrol. Earth Syst. Sci., 24, 5491–5518, https://doi.org/10.5194/hess-24-5491-2020, 2020.

Remote Sensing of Environment 268 (2022) 112754

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Review

Review of GPM IMERG performance: A global perspective

Rajani K. Pradhan [a,*], Yannis Markonis [a], Mijael Rodrigo Vargas Godoy [a], Anahí Villalba-Pradas [b], Konstantinos M. Andreadis [c], Efthymios I. Nikolopoulos [d], Simon Michael Papalexiou [a,e,f], Akif Rahim [a], Francisco J. Tapiador [b], Martin Hanel [a]

[a] Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcá 129, Praha-Suchdol 16500, Prague, Czech Republic
[b] Institute of Environmental Sciences, Department of Environmental Sciences, Earth and Space Sciences Group, University of Castilla–La Mancha, Avda. Carlos III s/n, 45071, Toledo, Spain
[c] Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, MA, USA
[d] Mechanical and Civil Engineering Department, Florida Institute of Technology, Melbourne 32901, FL, USA
[e] Civil, Geological and Environmental Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada
[f] Civil Engineering, University of Calgary, Calgary, Canada

ARTICLE INFO

Edited by Dr. Menghua Wang

Keywords:
IMERG
GPM validation
Remote sensing of precipitation
Precipitation extremes
Hydrological modeling

ABSTRACT

Accurate, reliable, and high spatio-temporal resolution precipitation data are vital for many applications, including the study of extreme events, hydrological modeling, water resource management, and hydroclimatic research in general. In this study, we performed a systematic review of the available literature to assess the performance of the Integrated Multi-Satellite Retrievals for GPM (IMERG) products across different geographical locations and climatic conditions around the globe. Asia, and in particular China, are the subject of the largest number of IMERG evaluation studies on the continental and country level. When compared to ground observational records, IMERG is found to vary with seasons, as well as precipitation type, structure, and intensity. It is shown to appropriately estimate and detect regional precipitation patterns, and their spatial mean, while its performance can be improved over mountainous regions characterized by orographic precipitation, complex terrains, and for winter precipitation. Furthermore, despite IMERG's better performance compared to other satellite products in reproducing spatio-temporal patterns and variability of extreme precipitation, some limitations were found regarding the precipitation intensity. At the temporal scales, IMERG performs better at monthly and annual time steps than the daily and sub-daily ones. Finally, in terms of hydrological application, the use of IMERG has resulted in significant discrepancies in streamflow simulation. However, and most importantly, we find that each new version that replaces the previous one, shows substantial improvement in almost every spatiotemporal scale and climatic condition. Thus, despite its limitations, IMERG evolution reveals a promising path for current and future applications.

1. Introduction

During the last three decades, satellite data have become a promising source of precipitation observations at the global scale (Levizzani and Cattani, 2019). They provide continuous measurement of precipitation in both space and time with quasi-global coverage (Derin and Yilmaz, 2014), making them especially important over data-scarce regions (Kidd and Levizzani, 2011). Satellite-based precipitation estimation techniques can be grouped into three main methods; (i) the visible (VIS) and infrared (IR), (ii) the passive microwave (PMW), and (iii) the merged VIS/IR and PMW (Kidd and Levizzani, 2011). As its name implies, the first uses VIS/IR images from geostationary satellites and estimates precipitation based on cloud top temperature. Even though the VIS/IR method provides high spatiotemporal resolution, uncertainties in the indirect relationship between the cloud top temperature and rainfall rate impedes its precipitation estimates (Maggioni et al., 2016; Tapiador et al., 2017). On the other hand, the PMW-based method uses the direct link between microwave scattering and rain/ice particles, and thus provides more direct precipitation estimates (Sun et al., 2018). However, a major drawback is that low orbit PMW satellites present poor sampling in time, which yields significant gaps in the precipitation estimation (Hong et al., 2019). The complementary properties of the

* Corresponding author.
E-mail address: pradhan@fzp.czu.cz (R.K. Pradhan).

https://doi.org/10.1016/j.rse.2021.112754
Received 6 May 2021; Received in revised form 7 October 2021; Accepted 12 October 2021
Available online 1 November 2021

PMW and VIS/IR technologies led to the development of the VIS/IR and PMW merged data products. Even though IR estimates are less skillful, the sparse PMW estimates cannot adequately cover all times, so IR estimates are used to help cover the periods which are not well-represented by the PMW (directly or by morphing). Some of these merged data products include the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network (PERSIANN) (Sorooshian et al., 2000; Hsu et al., 1997), Climate Prediction Center Morphing Technique (CMORPH) (Joyce et al., 2004), Tropical Rainfall Measuring Mission (TRMM) (Huffman et al., 2007), and Global Satellite Mapping of Precipitation (GSMap) (Kubota et al., 2020). Among them, the TRMM mission, which was designed to estimate precipitation over the tropics and subtropics, provided significant information on rainfall and its associated properties (Huffman et al., 2007). The Multi-Satellite Precipitation Analysis (TMPA), the most important data product of TRMM has been one of the widely used precipitation data set for a range of applications (Hamada et al., 2014; Li et al., 2009; Siddique-E-Akbor et al., 2014; Rozante and Cavalcanti, 2008).

Following the success of TRMM, on February 27 2014, NASA (National Aeronautics and Space Administration) and JAXA (Japan Aerospace Exploration Agency) jointly launched the Global Precipitation Measurement (GPM) core observatory satellite (Hou et al., 2013; Liu, 2016). The GPM mission is a constellation of satellites from partner nations and is one of the most accurate and finest spatiotemporal resolution providers of global precipitation measurements (Huffman et al., 2015). GPM not only extended TRMM's coverage from 35° N-S to 65° N-S, but it also carries advanced sensors such as the Dual-Frequency Precipitation Radar (DPR) and GPM Microwave Imager (GMI), which quantify precipitation more accurately and particularly light and solid precipitation (Hou et al., 2014). On the processing side, the Integrated Multi-Satellite Retrievals for GPM (IMERG) algorithm incorporates, merges, and inter-calibrates various IR, microwave (MW), and gauge observations to provide precipitation estimates at relatively high spatial ($0.1°\times0.1°$) and temporal resolution (30 min) (Huffman et al., 2015). Furthermore, IMERG provides three types of data sets: the IMERG-Early run (IMERG-E), IMERG-Late run (IMERG-L), and IMERG-Final run (IMERG-F). IMERG-E and IMERG-L, being near-real-time products, are available with a latency of 4 hours and 14 hours respectively, and can serve as a potential data source for flood forecasting and real-time disaster management (Huffman et al., 2020). IMERG-F is available 3.5 months after observation, and it is mainly aimed for research purposes. Unlike IMERG-E and IMERG-L, IMERG-F incorporates the Global Precipitation Climatology Centre (GPCC) monthly gauge analysis. More detailed information about IMERG precipitation products can be found in Tan et al. (2019). Since IMERG's release in early 2015, a substantial number of studies have used and recommended it for various applications, such as streamflow simulation (Tang et al., 2016b), flood forecasting (Wang et al., 2017b), and analysis of extreme events (Huang et al., 2019). Recently, IMERG version 6 (V06) extended its temporal coverage to the TRMM era and now provides 20-year long data sets from 2000 to present (Huffman, 2020). The high-quality precipitation estimates and long-term coverage of IMERG is expected to provide insights on various hydro-meteorological processes and climatological studies in the future.

A considerable number of studies have evaluated the performance of IMERG precipitation products at various temporal and spatial scales (e.g., Navarro et al., 2019; Shawky et al., 2019; Watters and Battaglia, 2019; Palomino-, ńgel et al., 2019; Prakash et al., 2018b; Tan et al., 2016; Manz et al., 2017). Nonetheless, most of them focus on local/regional domains (e.g., bounded by national boundaries) addressing specific climatic or topographic conditions across the globe. The few studies that investigate the global performance of IMERG present contradictory results (Wang et al., 2018; Liu, 2016; Derin et al., 2019). In this context, the main objective of this study is to review the state-of-the-art of the IMERG precipitation products and summarize the results of the recent efforts to evaluate the IMERG products in a quantitative manner across the globe. We aim to identify the strengths and weaknesses of IMERG products, providing information to the user community and product developers to further improve IMERG algorithms in future versions. The paper is organized as follows: the second section describes the methodology employed herein, briefly discusses data collection, database preparation, and its analysis. The third section presents the results in terms of validation design and geographical distribution of the publications. The fourth section discusses the post-2019 developments, strengths and weaknesses of IMERG products across the globe, identifies limitations, and provides recommendations for future studies. Finally, the last section summarizes the findings, and reports the conclusions.

2. Methodology

To analyze the global performance of GPM IMERG products, we performed an exhaustive literature review using the Google Scholar and Scopus databases. We used the keywords "GPM" and "IMERG", focusing on the period between 2016 and 2019. We limited the used articles to ones that had the evaluation of the performance of IMERG products within their scope. To assure the quality of the scientific articles used in this study, we focused only on articles published in Q1 and Q2 journals according to Scopus (first and second quantile of journals according to their ranking in the hydrology, climate or remote sensing fields). These selection criteria resulted in a 101 articles database with information regarding the performance of IMERG precipitation products across the globe. This information includes a unique identification (id) code, study area, country, continent, surface category, precipitation type, IMERG product, record length, temporal and spatial resolution, validation method, validation data, and statistical metrics (The database is available as a Supplementary File).

The id code was generated as the first three letters of the first author name followed by a two-digit number representing the year of publication. The study area varies widely from small river basins to global-scale analyses; therefore, the country and continent were also reported whenever the study area was not global. The surface category considered two major groups, namely land and ocean. Precipitation type was registered as rainfall or snowfall. IMERG products, as described in the previous section, could be IMERG-E, IMERG-L, or IMERG-F, and their version ranged between IMERG V03 and IMERG V06. The record length or length of the evaluation period was reported in months. The temporal resolution ranged from sub-hourly to yearly, and the spatial resolution was expressed in degrees. Validation methods considered the characteristics of the IMERG product (gridded) and the validation data (gridded or point) being compared, i.e., grid vs. grid or grid vs. point. Validation data reports the source of reference data used (gauge-, radar-, satellite-, or model-based). Statistical metrics were classified into volumetric and categorical indices. The volumetric indices mainly include the correlation coefficient (COR), root mean square error (RMSE), and bias. The key metrics in the categorical indices include the probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI). POD, also known as hit rate, represents the detection capability of the satellite; measures the proportion of the events detected by the satellite to the total number of precipitation events. FAR denotes the fraction of events detected by satellite that is not real or not detected by gauge, while CSI, a function of POD and FAR, represents a balanced score. Furthermore, where available, the database recorded additional metadata like best/worst performances, limitations, year of publication, and the journal of publication.

We followed a two-step evaluation approach: Firstly, the experimental design of the studies was evaluated based on the generated database. Information such as latitude, longitude, and the study area (both country and continents) were used to analyze the geographical distribution of the studies. Then, the publication years and their corresponding counts were employed for the investigation of the chronological evolution. In addition, the information regarding the spatiotemporal resolutions, validation period length, reference data

228

types, statistical metrics, etc., were used for the analysis of additional aspects of the validation design. Secondly, the performance of IMERG was summarised and categorized in terms of the continents, hydrological applications (i.e., streamflow simulation), and extreme events. The statistical metrics such as bias, COR, RMSE, Nash-Sutcliffe coefficient of efficiency (NSE), POD, and FAR were employed to support the results.

Herein we must note that the analysis is not complete. The search criteria were restricted to the "IMERG or GPM" keywords on Google Scholar search engine, and studies that do not have those keywords in their title or abstract were excluded from the database. Only studies whose scope was solely IMERG validation were gathered. In addition, data collection is limited to studies published in Q1 and Q2 journals. There may be more studies in Q3 and Q4 category's journals that were excluded from this study. Another limitation of the study is that we did not consider studies published outside journals, such as conferences, book chapters, and reports. Finally, the study focuses on articles published between 2016 and 2019 (there may be more evaluation studies published in the later 2020). It is worth noting that this study is aimed explicitly at GPM's IMERG products only, so it does not imply to the entire GPM mission (IMERG is one of the products of GPM and GPM provides other data sets as well).

3. Results

3.1. Characteristics of IMERG studies

From the geographical distribution of the 101 studies (Fig. 1) Asia not only has the majority of studies (66) but is also the continent with most studies covering different countries. The Americas hold a total of 21 studies, 12 in North America and 9 in South America, while there were 7 studies in Europe and 2 in Africa. Note that until 2019 no available studies had assessed IMERG performance over Australia. Additionally, 5 studies validated IMERG products at the global level (not shown in Fig. 1). On a per-country basis, IMERG performance was evaluated over a total of 34 different countries. There is an unequal distribution where China, the United States, and India represent around 55% of the studies. Yet, China alone accounts for 40% of the total. Countries like Brazil, Iran, Pakistan, Japan, Myanmar, Malaysia, and the Netherlands have at least two validation studies associated with each, whereas the rest of the countries are featured with a single study. Given the number of studies, IMERG validation is spatially well-distributed over the Asian continent.

The chronology of the studies indicates a growing research interest in the topic (Fig. 2). A total of 14 studies were published in 2016, 21 in 2017, 30 in 2018, and 36 in 2019. Africa has the lowest number of studies, it showed no increment with time, and no studies were published in 2018 and 2019. All other continents show an increase in the number of publications per year. Asia exhibited the highest growth in the number of publications per year. Starting with less than 10 studies in 2016, it has reached 28 in 2019, representing approximately three-quarters of all studies published that year. In Europe, the number of studies remains similar each year, except for 2018, in which the number of studies doubled. North America reported an increasing trend of studies from 2016 to 2018, while a decrease in 2019. Finally, in South America, there is fluctuation in the number of studies with the years without any trend. For the global studies, 2016 and 2018 have the same

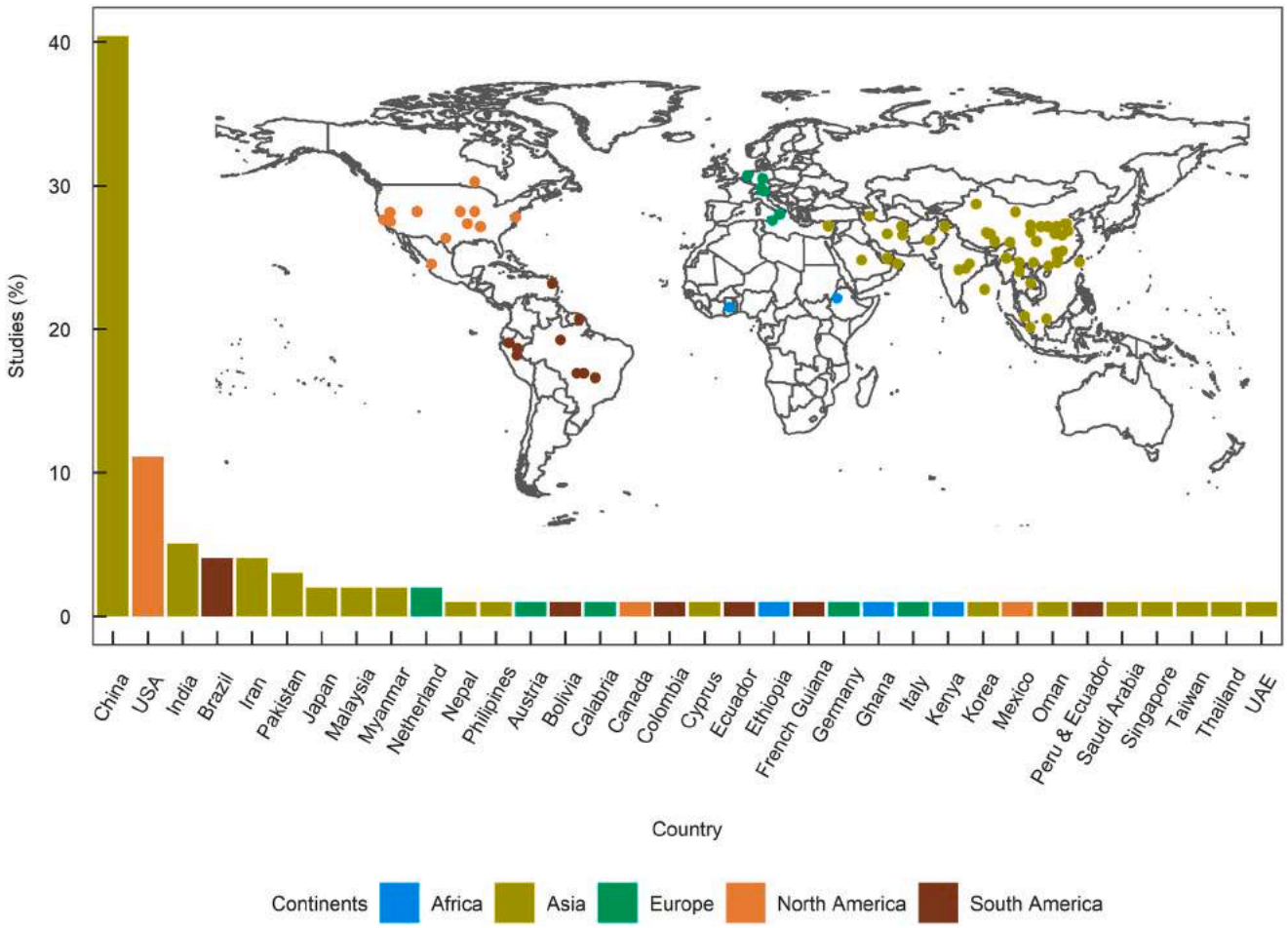


Fig. 1. Geographical distribution of IMERG validation studies across the globe. The points do not represent the extent of the study domain but rather the mean latitude and longitude of the domain.

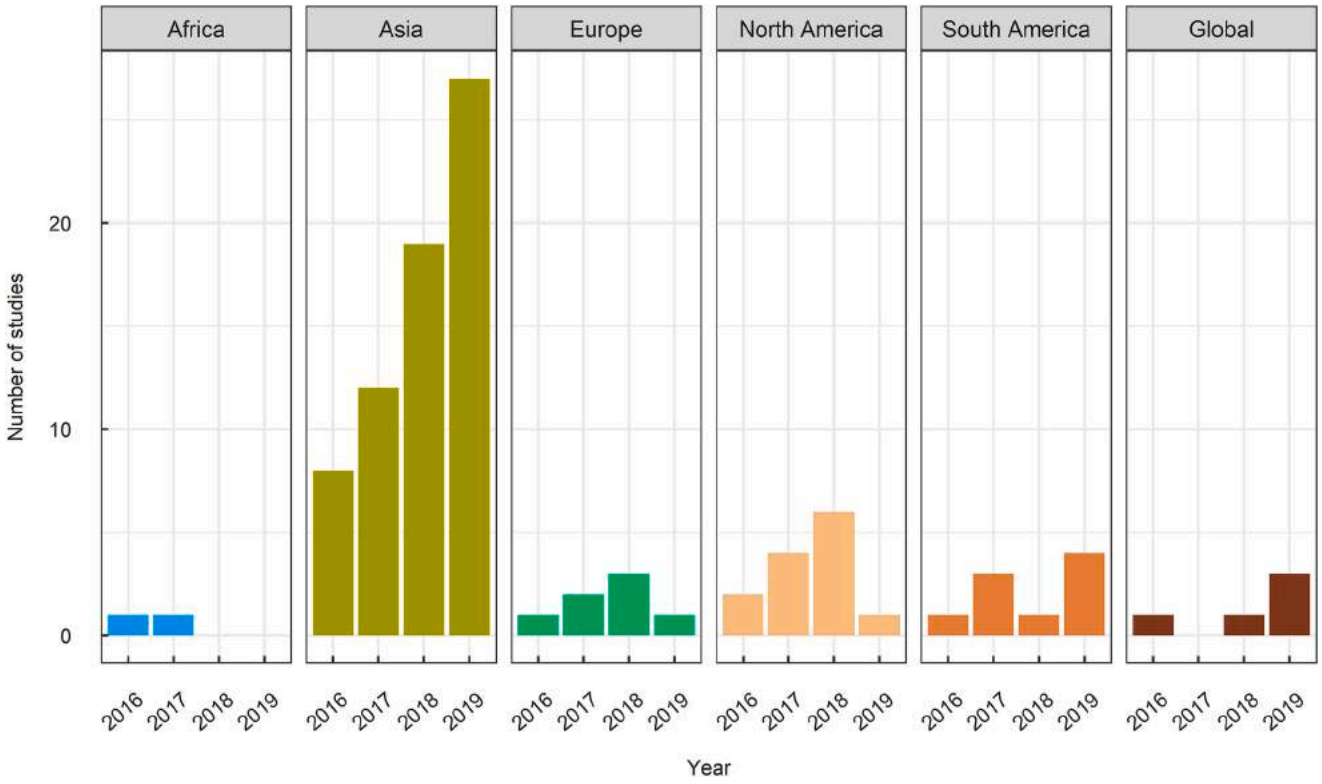


Fig. 2. Number of IMERG validation studies published between 2016 and 2019.

number of publications, but there is a significant increase in 2019.

The temporal resolution at which the validation is performed can significantly affect the results. Therefore, we identified the different temporal resolutions of IMERG products used in the aforementioned studies (Fig. 3). Those using products at daily resolution account for 35%, followed by monthly (22%), sub-daily (21%), annual (13%), and seasonal (9%) resolutions. When combined, the daily, monthly, annual and seasonal scales account for 80% of the total studies, whereas the rest (20%) are sub-daily scales. One possible explanation for this may be the availability of observational data sets. For example, this may be that the availability of observational data with high temporal resolution (sub-daily scale) at a regional scale is scarce. Another possible explanation could be that two precipitation data sets typically have a better agreement when they are upscaled to a coarser resolution in space and time. Therefore, evaluating IMERG at higher temporal scales (i.e., sub-daily) is a bit challenging (Tan et al., 2016). Continent-wise, Asia dominates in most temporal resolutions beyond the sub-daily scale. Studies over Europe are evenly distributed across the different temporal resolutions. In general, the least number of studies were at sub-daily resolutions (attributed to the lack of corresponding reference data), indicating that IMERG's raw resolution (30 min) has still not been adequately evaluated.

In terms of spatial scale, validation of IMERG data has been performed at 0.1°×0.1°, 0.25°×0.25°, 0.5°×0.5°, 1°×1°, 2.5°×2.5°, or 3°×3° resolutions (Fig. 3). The majority of validation studies were at 0.1°×0.1° and 0.25°×0.25° resolution. This could be a consequence of IMERG products' nominal resolution of 0.1°×0.1°, and the fact that most gridded reference data sets used are typically available at 0.25°×0.25° (e.g., TRMM TMPA). We note that studies that evaluated IMERG products using gauges (point vs. pixel-based method) were also considered as evaluated at 0.1°×0.1° resolutions; this also contributes to the higher number of studies on this resolution. Despite Asia being the subject of most of the studies, none of them evaluated IMERG products at resolutions coarser than 0.25°×0.25°; the same holds for Africa, Europe, and South America. In contrast, North America (Tan et al., 2017) and at the global scale (Khan and Maggioni, 2019); despite a smaller number of studies, conducted IMERG evaluation studies at different spatial resolutions. Therefore, more validation studies at multiple spatiotemporal resolutions are needed to better understand and achieve a more in-depth analysis of the IMERG data set properties over different scales.

Generally, the validation period length of the studies is increasing with recent IMERG versions (Fig. 4a). As expected, IMERG V03 has the shortest validation period (median around 12 months), whereas IMERG V05 has the most extended (median around 33 months). In terms of IMERG runs, IMERG-E and IMERG-L have similar validation lengths within the corresponding IMERG versions, unlike the IMERG-F, which has a relatively shorter period length. The main reason for this is that IMERG-F is available at 3.5 months of latency. Although IMERG V06 is available contemporary to the TRMM era (June 2000–present), up to 2019, no study has evaluated this period. Furthermore, the length of the validation period was also reported on a monthly scale and classified into five ranges, each a multiple of 12-month duration (Fig. 4b). Considering the short record of IMERG up through V05 (i.e., available from early 2014 onward), most of the studies' record length falls under the shortest range (0–12 months). Based on the database created, it is evident that the number of studies is inversely related to the validation period. Around 35% of the studies have a validation period length between 0 and 12 months, 31% between 13 and 24 months, 22% between 25 and 36 months, 11% between 37 and 48 months, and 1% between 49 and 60 months. It is interesting that studies using longer validation periods were often associated with coarser temporal resolution (daily and longer), whereas studies with shorter validation periods were associated with sub-daily temporal resolution. In terms of reference data, radar- and model-based data sets are mainly used for short evaluation periods, whereas gauge- and satellite-based data sets are used for long evaluation periods.

The types of reference data sets, and thus their accuracy, play a significant role in the evaluation results. Generally, ground (gauge- and/or radar-based) data sets are preferred as the source of reference data to

230

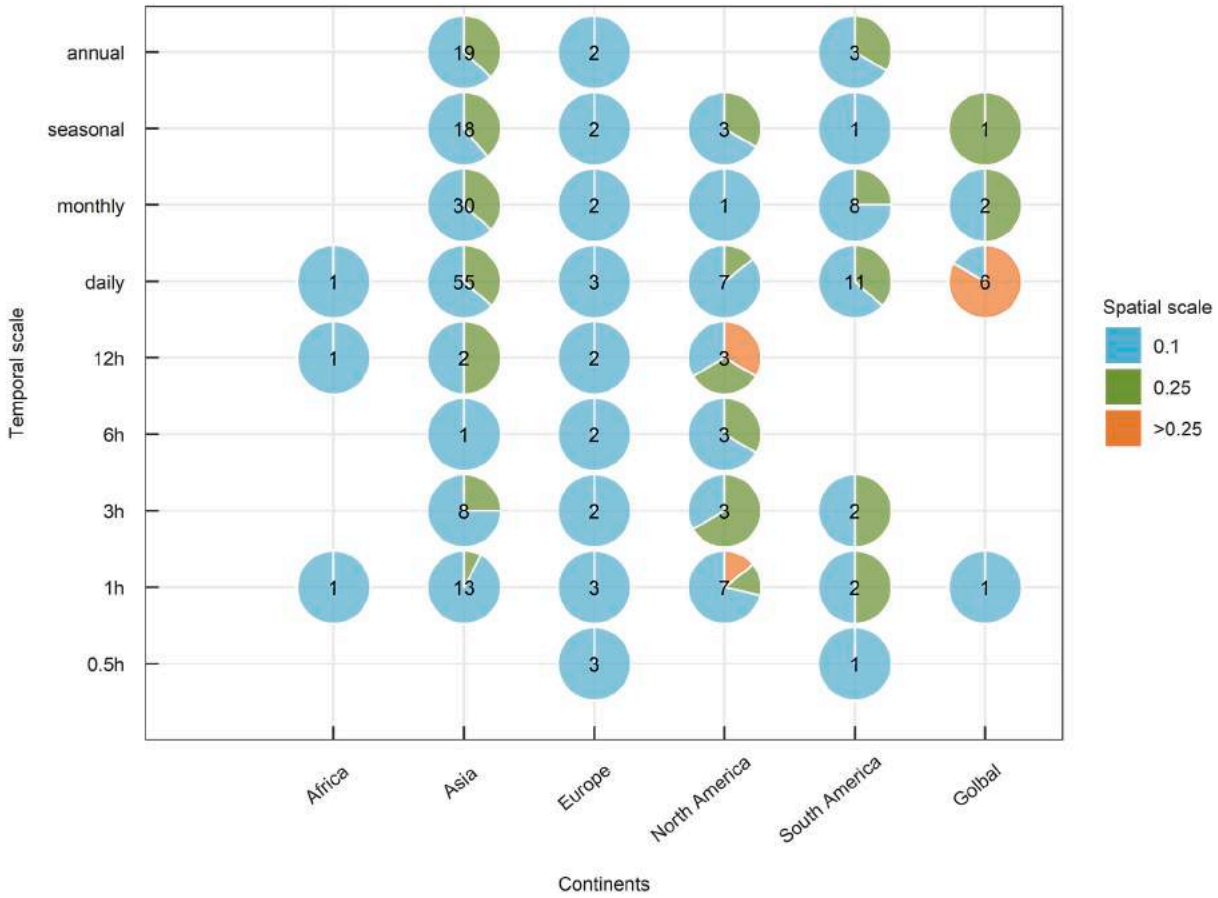


Fig. 3. Spatial and temporal scales of IMERG validation studies per continent. The numbers inside the circles do not represent the actual number of studies, because a single study can have evaluated in multiple temporal scale, and has the possibility of counted more than once.

assess the accuracy of GPM precipitation. However, the low density of ground stations globally forces the scientific community to rely on different sources for evaluation, namely satellite, model, reanalysis, and merged products (Fig. 4b). The most common satellite-based products for IMERG comparison are TMPA, GSMaP, and CMORPH. As GPM is the immediate successor mission of TRMM, most studies compare IMERG using TMPA (i.e., IMERG vs. TMPA) and compare their individual performance versus gauge precipitation (i.e., IMERG vs. TMPA vs. gauge). Only a couple of reference data sets came from different sources other than satellites, namely ERA-Interim (reanalysis) and Weather Research and Forecasting (WRF) (model) when ground observations were unavailable. In addition, radar precipitation data as a reference to evaluate the IMERG is very limited in number as well. The expensive installation and maintenance cost of radars could be the main reason for such fewer studies. Radar evaluation of IMERG is mainly dominated by developed countries/continents like the United States of America (USA) and Europe. The Multi-Radar/Multi-Sensor (MRMS) is the dominant radar product in the USA. Furthermore, all the IMERG evaluation studies against radar assessed the IMERG products on a sub-daily scale.

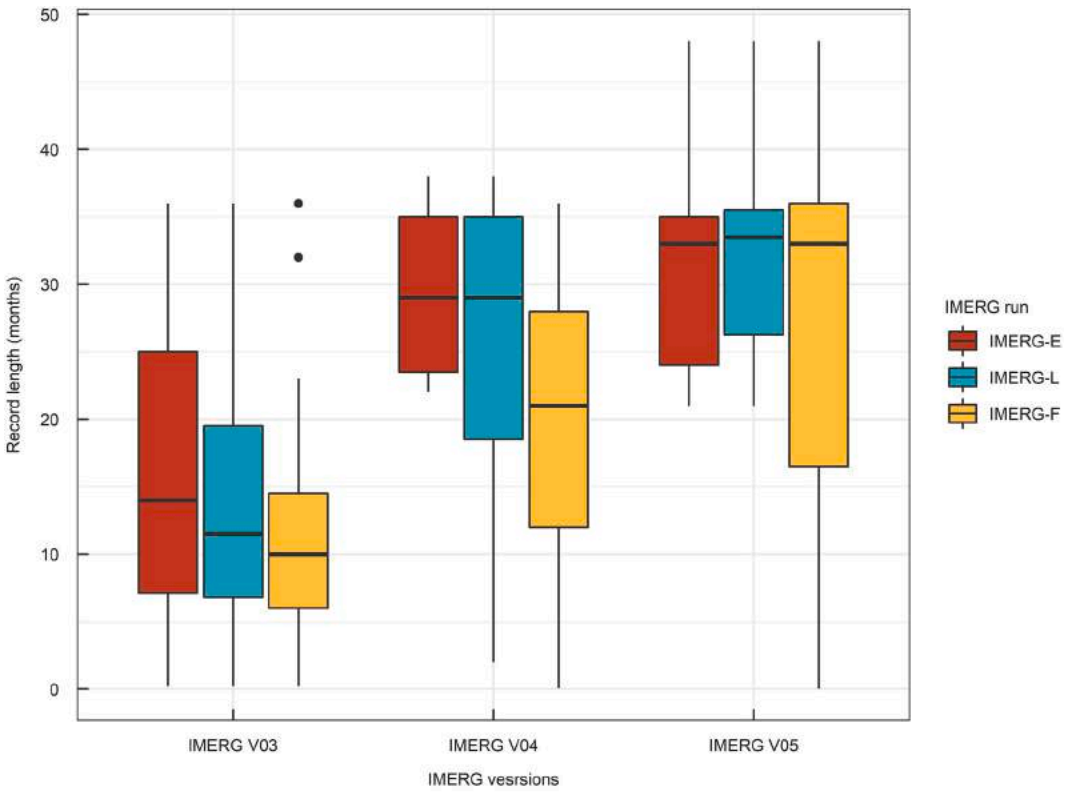
Generally, various statistical metrics were often used to validate satellite precipitation with the reference data sets. The statistical measures used for the IMERG validation can be categorized into two main types; i) Volumetric and ii) Categorical metrics.

The UpSet plot (Fig. 5) shows the number of studies that employed different statistical metrics combinations. The UpSet plot is a standard format to depict the intersection of sets when the number sets are more than three or four, and it was developed by Lex and Gehlenborg (2014). Each bar represents a unique combination of the metrics, and underneath the table shows their combination types. The empty cells (light grey)(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) indicate the particular metric is not part of the intersection, whereas the filled cells (black) indicate it participates in the intersection. From the left to right direction, the number of studies is decreasing. POD, COR, FAR, RMSE, CSI, and RBias are the most frequently used metrics combination (n = 22), followed by POD, COR, FAR, RMSE, CSI (n = 14), and POD, COR, FAR, RMSE (n = 7). Other combinations appear in a very low number of studies, mostly fewer than 3. In addition, the small barplot on the left side represents the unconditional (without combination) metrics used for the studies. From the top, POD is the most reported metrics used in 82 studies, followed by COR (78), FAR (74), and RMSE (72). On the other hand, CSI and RBias fall between 40 and 60.

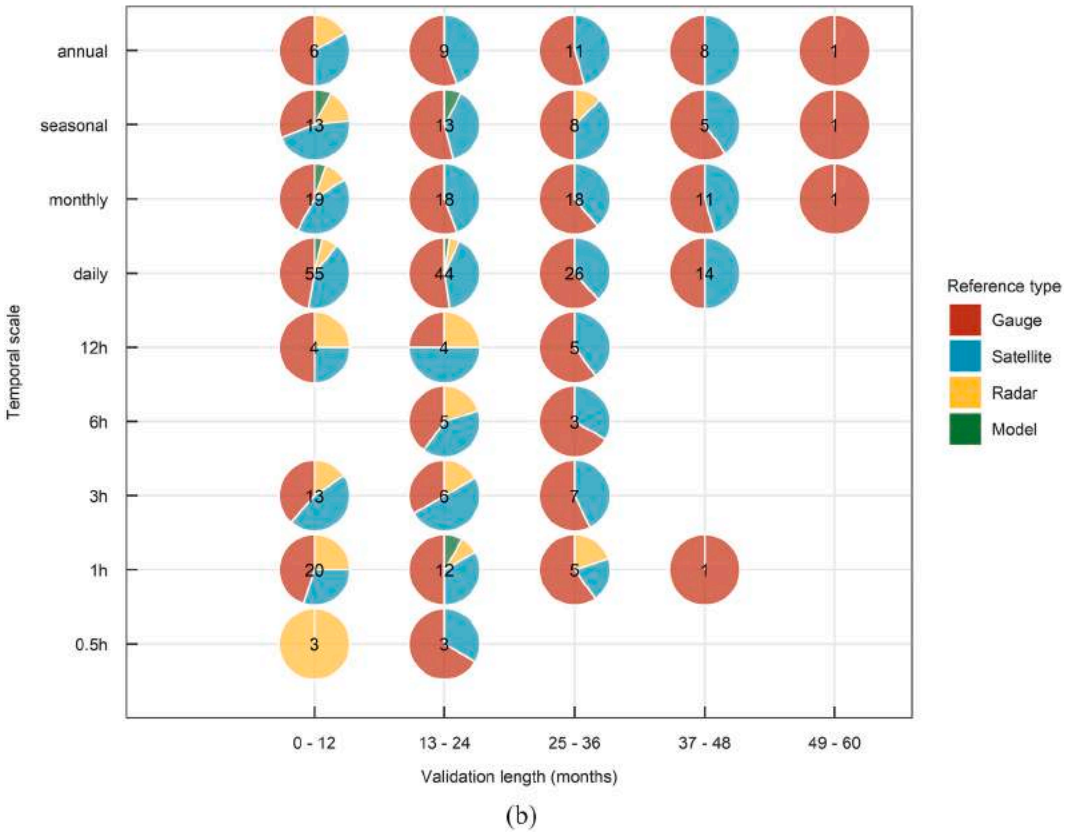
3.2. IMERG's performance by continents

3.2.1. Asia

Asia is characterized by diverse climate patterns and a variety of regional topography. Based on the studies assessing IMERG in China, IMERG captures the overall spatiotemporal behavior of precipitation over the country. However, there are substantial differences in local climatic conditions, which can affect IMERG-F V05 performance (Chen et al., 2018). Geographically, IMERG-F tends to be more accurate in the lower latitudes than mid/high latitudes of China (Chen and Li, 2016). IMERG-E, -L V05 showed more accurate estimates of high-intensity precipitation over wet/humid regions compared to low-intensity precipitation over dry regions (Wu et al., 2018). This is also supported by other studies as well (e.g., Fang et al., 2019; Jiang and Bauer-Gottwein, 2019; Wei et al., 2018; Asong et al., 2017). With regards to the diurnal variation, IMERG-F V06 performs poorly between 06:00 and 10:00 UTC (Xu et al., 2019b). Furthermore, on average, IMERG-F products at daily



(a)



(b)

Fig. 4. (a) Validation length by IMERG versions and runs, (b) Temporal scale versus validation length per reference type of IMERG validation studies. The numbers inside the circles do not represent the actual number of studies, because a single study can have evaluated in multiple references, and has the possibility of counted more than once.

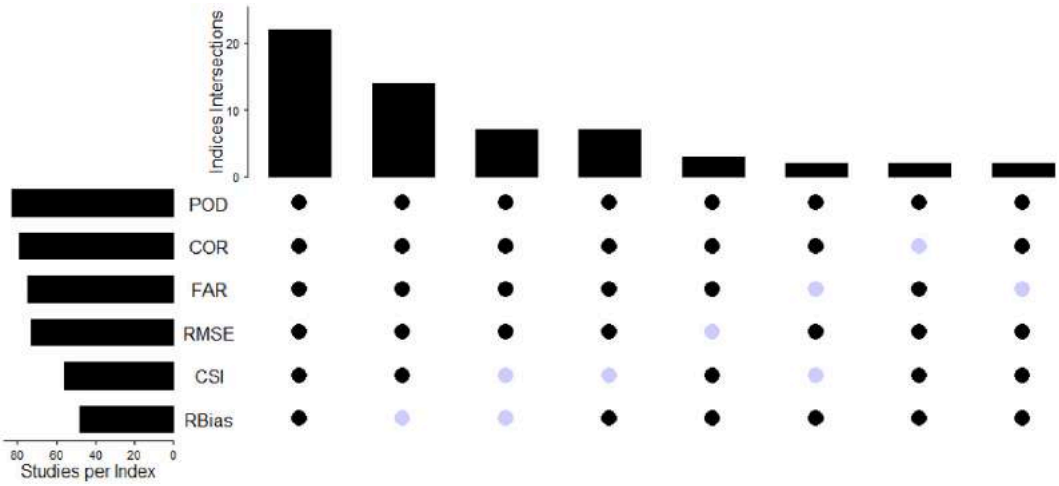


Fig. 5. Statistical metrics used for evaluation of IMERG products (POD = Probability of Detection, COR = Coefficient of Correlation, FAR = False Alarm Ratio, RMSE = Root Mean Square Error, CSI = Critical Success Index, RBias = Relative Bias).

and sub-daily time steps did not perform as well as in the monthly time scale (Xu et al., 2019b,a; Chen et al., 2018; Wang et al., 2019b). When assessing IMERG-F at the seasonal scale, its performance is worse during winter (Chen and Li, 2016). However, it is clear that IMERG-F V03 performance improved compared to TMPA, still leaving substantial room for further improvement over China (Tang et al., 2016a).

3.2.1.1. Eastern China. In eastern China (humid/semi-humid climate), although IMERG-F V04 has good agreement when measuring light precipitation (<8 mm/day), it tends to overestimate high precipitation rates (>64 mm/day), and underestimate precipitation rates between 8 and 64 mm/day. Furthermore, IMERG-F V04 showed an overestimation of up to 17.9% for the 99th percentile of precipitation on wet days (RR99P) and 11.5% relative bias for the R20TOT index (total precipitation sum of daily precipitation that is more than 20 mm) (Ning et al., 2017). The intensity of precipitation at which IMERG over-/-underestimates varies in space according to the topography. For instance, over the Huang-Huai-Hai Plain (eastern coastal region of China), the range of IMERG-F V05 overestimation of precipitation rates lies between 2 and 50 mm/day, and there is an underestimation of heavier precipitation rates (>50 mm/day) (Xu et al., 2019a). On the other hand, over the Huaihe river basin, IMERG-F V05 overestimates precipitation between 0.5 and 25 mm/day and underestimates it above and below that range (Chen et al., 2018).

3.2.1.2. Southern China. In southern China, IMERG-F tends to overestimate the light rainfall, and underestimate the high rainfall. Compared to TMPA, however, the underestimation of light rainfall is lower, and IMERG better captures the probability density function (PDF) and the inter-annual precipitation variability, especially over the lower Mekong river basin (Wang et al., 2017a). In addition, IMERG-F V03, V04 and V05 consistently overestimate precipitation over mountains and underestimates it in coastal regions of Guangdong Province. The positive hit bias and false detection of moderate to heavy precipitation events are attributed to the above-described overestimation (Wang et al., 2019b). IMERG's performance over high elevation (mountains) areas during dry seasons need further improvement. Nonetheless, IMERG estimates denote an improvement over TMPA estimates in terms of light/heavy precipitation detection and hit bias (Wang et al., 2019b).

3.2.1.3. Northwestern China. In the northwest region (arid climate), IMERG-F V05 did not show significant improvement compared to TMPA. IMERG underestimated precipitation at low altitudes and overestimated it at high elevations. On the seasonal scale, IMERG performs better in summer than in winter. Additionally, IMERG seems to suffer from poor detection capability of light rainfall, i.e., 0–2 mm/day. However, it better performs in moderate (>5–10 mm/day) and heavy precipitation events (>25 mm/day) (Wang et al., 2019c). Over the Tianshan mountain, IMERG-F V06 did not show significant improvements compared to the IMERG-F V05 (Anjum et al., 2018). IMERG products perform better in the eastern region compared to the western (e.g., Boertala Valley, Yili Valley, and West Tianshan). Overall, IMERG products are reliable enough to be used in precipitation trend analysis over the Tianshan mountain, but caution should be taken for the western regions (Anjum et al., 2018). Compared to TMPA, IMERG-E, -F V05 products have significant discrepancies over high latitudes and thus can be considered less reliable for the Tianshan mountain (Yang et al., 2019). Overall, IMERG has almost similar performance as TMPA and is significantly affected by the northwestern region's topography and aridity.

3.2.1.4. Tibetan Plateau. IMERG-F V03 and V06 validation studies over the Tibetan Plateau (Ma et al., 2018; Xu et al., 2017; Lu and Yong, 2018) show an overestimation of total precipitation in the southwest regions and an underestimation in the northeast regions. Moreover, a high correlation was observed in the northeast and southeast regions, while a low correlation was reported in the southern regions of the Tibetan Plateau (COR <0.40). Most likely, the complex topography of the Himalayan mountains can be the reason for the low correlation (Ma et al., 2018). Additional findings confirm that the accuracy of IMERG-F V05 decreases as elevation increases, which indicates the direct effect of elevation on IMERG products (Wang et al., 2019b). Furthermore, the detection of light precipitation is particularly affected at elevations above 4500 m (Xu et al., 2017). IMERG-F V05 tends to overestimate trace or light precipitation (0–1 mm/day) and underestimates highly intense precipitation (>50 mm/day) (Wang et al., 2019b). Overestimation could be attributed to the evaporation of light precipitation in the atmosphere before reaching the surface (Wang et al., 2019b). In terms of IMERG versions, IMERG-F V04 did not show significant improvements to its predecessors (i.e., IMERG-F V03). Compared to the IMERG-F V03, IMERG-F V04 showed significant underestimation of daily precipitation's annual average with a relative bias of −60.91% over the Tibetan plateau, which becomes more profound in the winter (−72.33%) (Zhao et al., 2018). Although IMERG-F V03 outperformed IMERG-F V04 over the Tibetan Plateau, yet both products underestimate winter precipitation with relative bias of −6.47% and −70.62% respectively (Wei et al., 2018). Despite the fact that IMERG-F V06 captures the average distribution of total precipitation in space,

233

detection of light rainfall, winter snowfall, and detection of precipitation at high elevations remain major challenges over the Tibetan Plateau (Lu and Yong, 2018).

3.2.1.5. India and Pakistan. Over India, IMERG-F V03 showed a noticeable improvement over GSMaP and TMPA, capturing southwest monsoon mean rainfall and its variability (Prakash et al., 2015). Nevertheless, there was a higher total negative bias and hit bias in IMERG-F V03 and V04 in mountainous regions such as the Himalaya foothills and Western Ghats and underestimation of northeastern orographic precipitation (Prakash et al., 2018b, 2015, 2018a). A large fraction of the FAR and hit bias over the south peninsula can be attributed to the Western Ghats (leeward side or rain shadow region). In contrast, IMERG-E performs better over plains and coastal regions (Singh et al., 2019), significantly improving TMPA systematic error dependency with topography. In addition, there were improvements in rainfall estimates of varying intensities across different topographies over most river basins, except northwest semi-arid basins (Beria et al., 2017). IMERG-F V03 showed reasonable improvement over TMPA capturing heavy precipitation events during the summer monsoon season, especially over Himalaya and northwest India (Prakash et al., 2015). IMERG-F appears to overestimate precipitation in high elevation zones in Pakistan and slightly underestimate it in semi-arid regions. Additionally, it tended to overestimate pre-monsoon and monsoon precipitation but underestimated post-monsoon and winter precipitation (Rahman et al., 2018). Despite overestimation of light precipitation (0–1 mm/day) and underestimation of moderate (1–20 mm/day) to heavy rainfall (>20 mm/day) over the north hill highlands of Pakistan, IMERG-F V04 represents the spatial variation of precipitation better than TMPA (Anjum et al., 2018). Overall, in southeast Asia, the uncertainties about orographic precipitation remain a considerable challenge.

3.2.1.6. Eastern Asia. In Eastern Asian countries, the performance of IMERG varies with space. For instance, in Japan and Korea, IMERG-F's V03 average POD (0.69) for convective rainfall over mountains and coastal regions during pre-monsoon and monsoon season is 8% better than TMPA (POD = 0.61) (Kim et al., 2017). In addition, it outperforms TMPA in both pre- and post-monsoon precipitation, as well as in terms of spatial precipitation patterns. However, contrary to the expectations, TMPA outperformed IMERG-F V04 precipitation estimates at daily and monthly scales both for total and heavy precipitation over Myanmar (Yuan et al., 2017). A poor detection and estimation skill of IMERG-F V04 is found both in light and heavy precipitation with a significant underestimation of total precipitation in the Chindwin river basin of Myanmar (Yuan et al., 2017). Similarly, over the same region, 3B42RT shown the best estimates followed by IMERG-F V05, whereas the near-real-time products (IMERG-E V05 and IMERG-L V05) have the lowest quality (Yuan et al., 2019). Furthermore, despite IMERG-F's V04 better detection of daily precipitation, overall, it did not show significant improvement compared to TMPA over Singapore (Tan and Duan, 2017). In Taiwan, IMERG-F V05 can reproduce different precipitation characteristics like the seasonal variation and temporal bimodal peak of annual precipitation. By validating IMERG with gauge data, it appears again that spatial discrepancies and underestimation bias are higher over mountain regions than plains. Furthermore, in terms of seasonality, IMERG shows poor performance in winter (Huang et al., 2018). Despite underestimating heavy precipitation, IMERG-F V05 agrees with the APHRODITE data product in Japan, the Philippines, and Nepal (Sunilkumar et al., 2019).

3.2.1.7. Western Asia. In western Asia, IMERG-F V03 had acceptable performance compared to TMPA and European Centre for Medium Range Weather Forecasts (ECMWF). Accordingly, it was shown that it could be used as a substitute for ground observations in regions lacking observational precipitation over Iran (Sharifi et al., 2016; Khodadoust Siuki et al., 2017; Maghsood et al., 2020). In the study by Mahmoud et al. (2019), IMERG-F V03 showed good agreement with ground data in the southern, middle, and northern parts of the UAE. However, in the eastern and northeastern parts of the UAE, characterized by mountainous topography and coastal areas, there were errors in detection and estimation. IMERG-F outperformed IMERG-E and IMERG-L in terms of POD, bias, Mean Absolute Error (MAE), and RMSE. However, against expectations, IMERG-E outperformed IMERG-F in terms of correlation. In Oman, it was observed that as precipitation intensity increases, so does IMERG-E, -L, -F V04 underestimation. The Mean Difference (MD) was −3.11 for 2.5–10 mm/day, −12.30 for 10–50 mm/day, and −50.74 for > 50 mm/day intensity classes (Shawky et al., 2019). In Saudi Arabia, IMERG-F V05 also outperformed IMERG-E and IMERG-L, and its uncertainty was higher in the southern and northern parts of the country (Mahmoud et al., 2018).

3.2.2. North America

North America has the second-highest number of IMERG validation studies. It was shown that compared to the Multi-Radar/Multi-Sensor (MRMS) precipitation data set, IMERG-F V03 overestimates drizzle (light rainfall) and underestimates heavy rainfall over CONUS (Contiguous United States) (Tan et al., 2016). Furthermore, compared to TMPA, IMERG-L V03 improved the missed rain bias, and false hits over the same region (Gebregiorgis et al., 2018). Tan et al. (2017) evaluated the IMERG-F V03 product against the MRMS as a function of spatial-temporal scale over the southern United States. They found an enhancement of performance with the increase in the spatial-temporal scale, both capturing the rain occurrence and its estimation. Over the central United States, IMERG-E and -L V05 hourly products show close agreement and higher correlation with NCEP products when the temperature (i.e., hourly) exceeds 280 K (Zhang et al., 2018). IMERG-F V05 performs noticeably well in representing the spatial variability of storms, despite some errors in high-intensity precipitation regions (storm core) (Omranian et al., 2018). In terms of the diurnal and semi-diurnal cycle, IMERG-F V04 agrees with the reference precipitation. However, it overestimates the normalized amplitude over the central US and underestimates it in the western and eastern US mountainous regions (Kirstetter et al., 2018). It also shows substantial differences in the peak of diurnal precipitation for convective and stratiform precipitation of mesoscale convective systems over the Great Plains (Kirstetter et al., 2018). Despite the overestimation of heavy rain events and low performance in mountainous regions, IMERG-F V03 is satisfactory reproducing the spatial distribution and precipitation amount over Canada. In addition, its performance is relatively better in the continental semiarid region than in the humid regions (Asong et al., 2017). In Mexico, IMERG-F V03 underestimates heavy precipitation at daily and hourly scales, but it reduces the error over high-elevation terrains (Mayor et al., 2017). In terms of snowfall, IMERG-F V04 underestimates precipitation compared to SNOTEL with a relative bias between −71% and −82% over the western mountain regions. Furthermore, the discrepancy between IMERG-F V04 and SNOTEL observations increases as daily temperature increases from −14°C and approaches 0°C. Concerning precipitation intensity, the IMERG products have better performance between 0 and 5 mm/day but show significant underestimation at >10 mm/day (Wen et al., 2016). Similarly, Sadeghi et al. (2019) reported the IMERG-F's V04 underestimation of snow accumulation, although it detects the snowfall events comparatively better than the MRMS. It may be concluded that IMERG's current snowfall estimation performance is unreliable for hydrological and climatological applications.

3.2.3. Europe

Europe is the second to last continent with the least published IMERG validation studies (6%). The first study validated IMERG-F V03 over the Netherlands using one year of the data set developed by Gaona et al.

234

(2016). It concluded that IMERG could reproduce the spatiotemporal distribution of precipitation over the nation despite a very small (2%) underestimation across all resolutions (i.e., 0.5h, daily, monthly, and annually). Furthermore, IMERG-F V03 has a small relative bias on the 30 min (−1.51%) and daily (−1.49%) scale, highlighting IMERG's potential in hydrological applications (Gaona et al., 2017). When comparing IMERG-F V05 to radar-based precipitation data, IMERG showed a significant overestimation of precipitation, especially during winter over Germany with low correlation (<0.4), POD (0.38), CSI (0.28), and high FAR (0.48). Furthermore, it showed difficulties in reproducing spatial variability across Germany's diverse topography (Ramsauer et al., 2018). The performance of IMERG-E V04, IMERG-F V05, and IMERG-F V06 was influenced by complex terrain and had problems capturing precipitation over mountainous regions (Chiaravalloti et al., 2018; Ramsauer et al., 2018; Navarro et al., 2019). Additionally, IMERG-F V06 performed better during summer than in the winter (Navarro et al., 2019) and on the monthly scale than in the daily and sub-daily ones (Gaona et al., 2016; Ramsauer et al., 2018).

3.2.4. South America

In South America, IMERG-F V06 effectively represents the spatial pattern of precipitation and shows reasonably better performance than TMPA throughout Brazil (Rozante et al., 2018). In the northeast coast of Brazil, which is characterized by warm rain events, IMERG-F V05 showed significant errors and underestimated daily precipitation. The above could be attributed to the inability of GPM sensors to detect orographically forced warm-rain processes (Gadelha et al., 2019). Large biases appear in the North and Central-west regions, associated with the

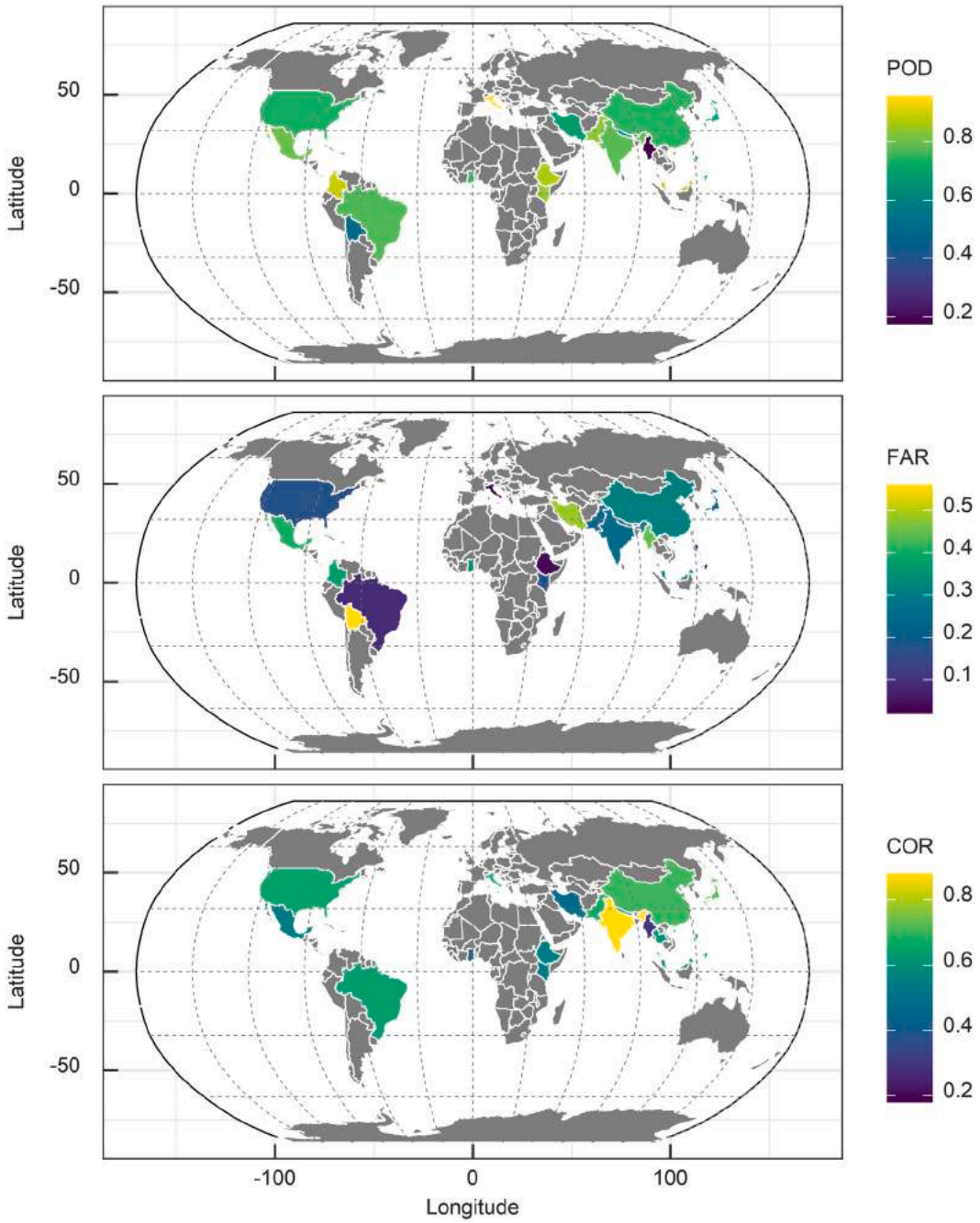


Fig. 6. Spatial distribution of POD, FAR, and correlation values of IMERG evaluation studies across the countries at daily scale. The matrices were calculated based on the daily scale, and median values were considered in case of more than one study available for the same country (e.g., China, Iran, India, USA etc.).

sparse density of gauges. In the Central Plateau of Brazil, IMERG-F V05 reproduced annual and monthly precipitation better than daily precipitation. Furthermore, IMERG-F V05 exhibited strong seasonal variability as numerous errors, and estimation difficulties occur with low and sparse dry season precipitation (Salles et al., 2019). For diurnal precipitation, IMERG-F V03 overestimated the frequency of heavy precipitation over the Negro, Solimões, and Amazon rivers and underestimated dry season precipitation compared to S-band weather radar measurements. The above was attributed to IMERG's difficulties in detecting isolated convective cells and the poor calibration over water surfaces (Oliveira et al., 2016). In the high Andes, IMERG-F V03 efficiently captured rainfall intensity but showed substantial discrepancies with gauge-based observations along the dry Peruvian coastline (Manz et al., 2017).

3.2.5. Africa

Africa has the least IMERG validation studies, with only two works so far (2019). The lack of access to reliable observational data sets could be the reason for such a small number of validation studies. These analyses agree that the performance of IMERG-F V03 and V04 varies with the season, climate, and topography (Dezfuli et al., 2017; Sahlu et al., 2016). For instance, IMERG performs better over Eastern and humid regions of Africa than Southern Sahel, and the discrepancies between IMERG-F V04 and TMPA were higher over mountainous regions (Dezfuli et al., 2017). In addition, despite a slightly superior performance of IMERG-F V03 over the Blue Nile basin, its overall performance was similar to CMORPH (Sahlu et al., 2016). Another finding suggests that IMERG-F V03 detection capability decreases with increasing precipitation, emphasizing the need to improve IMERG's detection capability of heavy precipitation (Sahlu et al., 2016).

3.2.6. Spatial distribution of POD, FAR and COR

The spatial distribution of POD, FAR, and correlation values reported among the reviewed studies are shown in Fig. 6. These values are based on a comparison of IMERG daily precipitation with the corresponding gauge observation on a daily scale. Only the correlation, FAR, and POD values of the most recent version and IMERG final run (IMERG-F) were considered when a single study evaluated successive versions (IMERG V03, IMERG V04, etc.) and various IMERG runs (IMERG-E, -L, and -F) in the same study. More or less, the majority of the studies reported good detection skills (POD $>$ 0.6), fewer false alarms (FAR $<$ 0.5), and reasonable agreement with reference data sets (COR $>$ 0.5). Myanmar is the only country that reported very poor values for both detection (POD = 0.17, FAR = 0.45) and estimation (COR = 0.29) of precipitation. The above is attributed to IMERG's poor performance in n detecting and estimating light and heavy precipitation along the Chindwin river basin of Myanmar (Yuan et al., 2017). However, the study region is characterized by very sparse observation stations, and the evaluation results are based on only four gauge stations.

In terms of continents, IMERG has good detection skills over Africa with POD ranging from 0.73 to 0.84, and FAR up to 0.35, but moderate correlation ranging from 0.42 to 0.54. Similarly, North America (POD = 0.73–0.8, FAR = 0.17) and South America (POD = 0.51–0.89, FAR = 0.12–0.37) revealed a good detection skill, except Bolivia which reports the highest FAR of 0.56. In addition, both the continents reported a reasonable correlation (0.54–0.65) of IMERG with the reference data sets. High spatial variation in both POD and correlation is observed in Asia, varying from the poor results over Myanmar to high values over India. In addition, the East Asian countries such as Thailand, the Philippines, Nepal, and Malaysia have correlation values ranging from 0.5 to 0.69, whereas Japan, India, and China have COR $>$ 0.7. In western Asian countries, Iran and Pakistan present a medium correlation (0.47 and 0.67, respectively). On the other hand, Pakistan and Malaysia report slightly better detection skills ($>$ 0.8) than India (0.78) and China (0.74). Detection skills over Japan, Iran, and Nepal vary from 0.59 to 0.69. In terms of false detection, Iran reported a FAR of 0.51, which could be attributed to the prevailing arid condition, and evaporation of light precipitation before reaching the ground's surface. Moreover, based on the POD, FAR, and COR values among the studies, it can be concluded that IMERG shows its good detection skills throughout various climatic and topographic conditions and has better agreement with the reference observations in overall precipitation estimation.

3.3. IMERG performance by run (IMERG-E, -L, and -F) types

Generally, it is believed that the IMERG-F, which has gauge correction at the monthly scale, has superior performance compared to both -E and -L run products as shown in China (Guo et al., 2016; Tang et al., 2016a), East Asia (Kim et al., 2017), Austria (temperate climate) (Sungmin et al., 2017), Saudi Arabia (Mahmoud et al., 2018), Pakistan (Anjum et al., 2018), and Italy (Chiaravalloti et al., 2018). However, few studies contradicted this and reported either no significant improvement of IMERG-F run or even outperformance by IMERG early run at least in some aspect (i.e., POD, correlation, extreme events). For instance, Maghsood et al. (2020) reported no significant IMERG-F improvement in POD and FAR over Iran at the daily scale, but it does at the monthly scale. In the same study, they report that IMERG-E and -L products are more suitable for extreme precipitation. Moreover, Shawky et al. (2019) found no significant improvements of IMERG-F over IMERG-E in the arid environment of Oman. Mahmoud et al. (2019) reported that IMERG-E outperformed the IMERG-F products in terms of correlation over the UAE. However, in terms of error and bias, IMERG-F outperformed both the early and late products (on average MAE and RMSE decreased by 10% and 11%, and bias from 1.1% to 0.4%). Tan and Santo (2018) observed similar results over Malaysia. This counter-intuitive behavior could be attributed to the sparse gauge availability and consequently the GPCC calibration of IMERG-F over those regions. Moreover, it should be noted that the IMERG-E and IMERG-L runs serve for near-real-time applications (e.g., flood, drought, and crop forecasting), whereas the IMERG-F is best intended for research purposes.

3.4. IMERG performance by versions (IMERG V03, IMERG V04, IMERG V05, and IMERG V06)

The IMERG algorithm and its version are episodically updated, yet only a few studies compared successive IMERG versions (Fig. 7). Xu et al. (2019c) compared the IMERG V04 and IMERG V05 over mainland China and found that IMERG V05 estimates precipitation better except for false precipitation. Wang et al. (2019a) found that IMERG V05 possesses significant enhancements in precipitation estimation compared to the IMERG V03 and V04 over the Guangdong Province, China. Similarly, Satg é et al. (2018) reported IMERG V05's expected improvement compared to its predecessors (i.e., IMERG V03 and V04) throughout Pakistan except for the extreme arid region where the IMERG V04 had the best performance. In addition, Wang et al. (2018) compared the IMERG V03, V04, and V05 at the global level. They found that IMERG V05 significantly improved over the previous IMERG V03 and V04, and the improvements are mainly observed in the estimation of mean oceanic precipitation.

On the other hand, Anjum et al. (2019) revealed no significant improvement of IMERG V06 over IMERG V05 in the Tianshan Mountains, China. Likewise, Derin et al. (2019) reported that when capturing light or heavy precipitation IMERG V06 failed to outperform IMERG V05 over the mountain regions. In addition, IMERG V04 did not shown paramount enhancement compared to its predecessor IMERG V03 (Satg é et al., 2018). The above claim is supported by studies in China (Zhao et al., 2018), and the Tibetan Plateau and Weihe River Basin (Wei et al., 2018). Despite successive IMERG versions aiming towards more accurate estimation of precipitation, its performance varies with the precipitation type and topographical features.

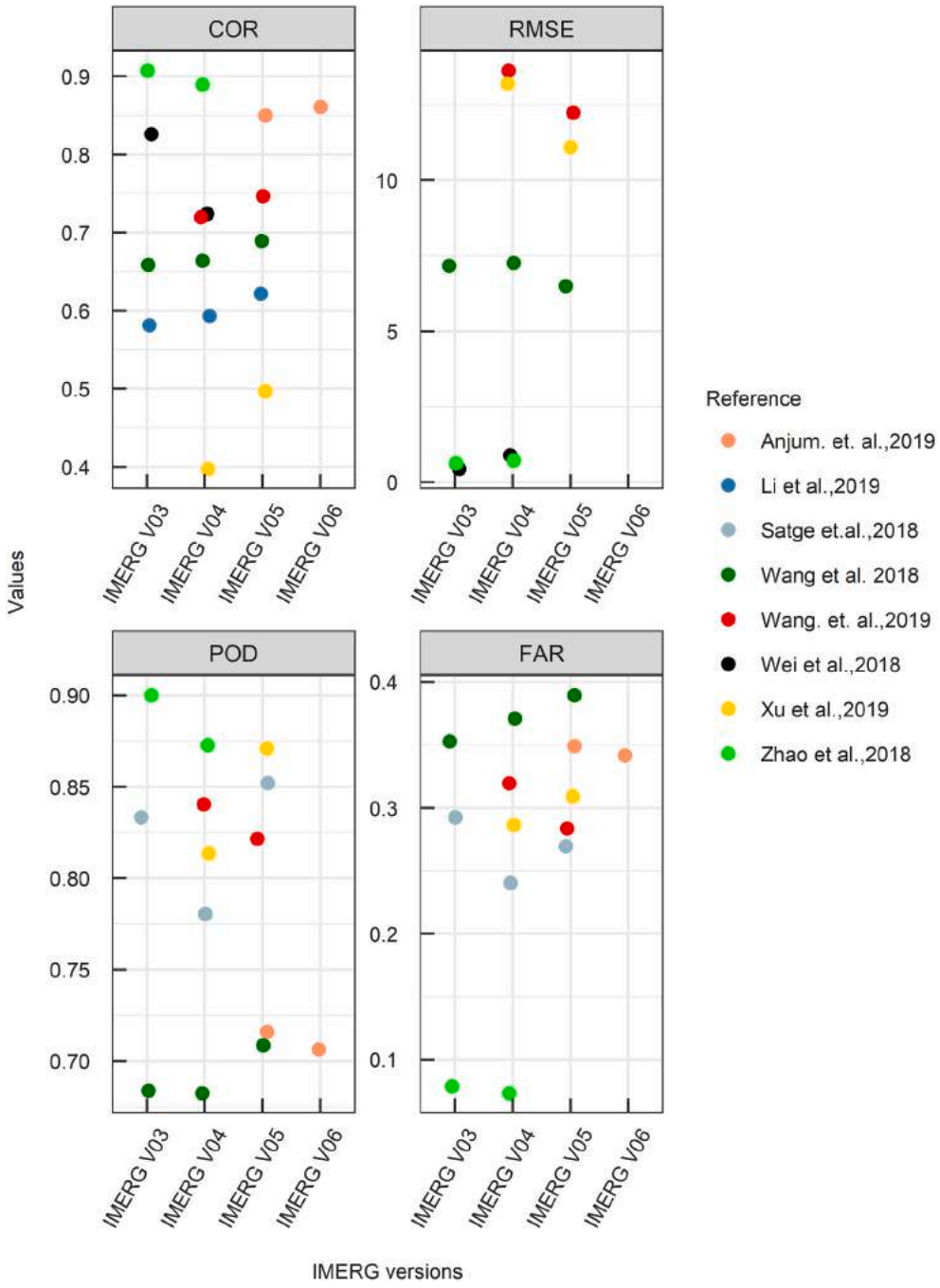


Fig. 7. Performance of IMERG successive versions (IMERG V03, V04, V05 and V06) in term of COR, RMSE, POD, and FAR values across the countries at daily scale.

3.5. IMERG's performance over ocean

Out of the 101 peer review articles in our database, only three studies evaluated the performance of IMERG over the ocean, making it difficult to draw any definitive conclusions. The lack of standard observational data sets could be the possible reason for such a low number of studies over the ocean. Khan and Maggioni (2019) assessed the performance of IMERG-E, -L, -F V05 daily products over the ocean using Ocean-Rain and satellite-based radar products as validation targets. IMERG was found to underestimate Ocean-Rain measurements significantly. However, 80% of the time, IMERG products detect rain. Prakash et al. (2018a) assessed the performance of IMERG-F V04 over the north Indian ocean using hourly moored buoy observations. The authors reported that IMERG performed better over the Arabian sea than over the Bay of Bengal. Despite low errors and good detection capability, IMERG suffered from a high FAR and overestimated rainfall (especially for light and extreme precipitation). Considering IMERG's resolution (0.1° x 0.1°), it is more probable that small-scale showers can occur at any part of the IMERG grid, but not over the exact buoys locations, which are sparser than the gauges over land. Thus, leading to the apparent overestimation of precipitation. Besides, the evaporation of light precipitation in the atmosphere could be another possible reason. Wang et al. (2018) reported that mean oceanic precipitation of IMERG-F V04 and IMERG-F V05 shows significant improvement over IMERG-F V03, and their estimates are close to the ones by Multi-Source Weighted-Ensemble Precipitation (MSWEP) (Beck et al., 2017) and Global Precipitation Climatology Project (GPCP) (Adler et al., 1979-). Even though observational data sets are scarce over the ocean, other global products such as GPCP and

237

MSWEP could be used as a reference for validation, as shown above.

3.6. IMERG's performance in representing extremes

The high spatiotemporal resolution, homogeneous global coverage, and near-real-time availability of satellite-based precipitation data are essential to understand extreme events better. Various recent studies are evaluating IMERG's performance for multiple extreme event applications (Table 1). Omranian et al. (2018) evaluated the capability of IMERG-F V05 to reproduce the precipitation of hurricane Harvey in Texas. They found that IMERG captured the storm with POD >0.82 and FAR <0.2, and precipitation spatial variability with 62% accuracy. Nevertheless, several aspects appear to need improvements, such as underestimation over the coastal region, overestimation in the high-intensity region, and discrepancies between observational data and IMERG precipitation, especially near the storm's center. Huang et al. (2019) analyzed the performance of IMERG-E and -F V05 products to capture six major typhoons over southern China during 2016 and 2017. They found that IMERG captured precipitation spatial variability and areal hourly precipitation. Furl et al. (2018) also revealed a significant underestimation in capturing extreme storms. Interestingly, they observed that the IMERG-E and -L V03 products have lesser underestimation than the IMERG-F V03 products, which can be attributed to the -F run's gauge correction. Wen et al. (2018) evaluated the performance of IMERG-L V04 in terms of Atmospheric Rivers over the western USA. Despite good detection of heavy precipitation events, IMERG significantly underestimates (−40%) the total precipitation volume. Additionally, Fang et al. (2019) found good agreement between IMERG-F V05 and gauge observations regarding extreme precipitation spatial patterns throughout China. IMERG's performance was more consistent over the southeast (humid) than over the northwest (arid) regions. In the same study, an underestimation of extreme precipitation events was reported, implying that although IMERG reproduces the spatial precipitation pattern and volume better than TMPA, limitations in detecting extreme events remain. Zhang et al. (2019) evaluated the performance of IMERG-F V05 products in capturing a 60-year return period extreme precipitation storm over southern China. The authors reported that IMERG products significantly underestimated the event. However, IMERG's performance was reported to vary at different intensities; i.e., IMERG performs poorly when rainfall intensity is above 17 mm/hr and best when the intensity is below 5 mm/hr.

3.7. IMERG's performance in hydrological applications

A large number of studies have evaluated the performance of IMERG products in terms of simulating streamflow (Table 2). However, as the studies used different hydrological models, calibration and validation methods, basin locations, and climatic conditions, the direct comparison of the results cannot be conclusive. Nevertheless, we compiled the Nash-Sutcliffe coefficient of efficiency (NSE) and statistical bias indices from these studies (Fig. 8) to summarise their results. Our main objective here is to highlight the IMERG performance and its different runs (IMERG-E, -L, and -F) in each study separately, rather than inter-comparing different studies results with each other.

Li et al. (2017) found more promising results when the hydrological model parameters were calibrated by gauge and RQPE (Radar-corrected Quantitative Precipitation Estimation) compared to the IMERG-F V04 data set. Furthermore, when the model was calibrated using raw IMERG data, the results were quite problematic over the Ganjiang river basin. In another study, Wang et al. (2017b) evaluated both near real-time and post real-time V03 products using the VIC (Variable Infiltration Capacity) model over the Beijiang River Basin. They found promising results for IMERG-F, while near real-time products showed poor performance (NSE < 0.35). Nonetheless, both products showed reliable flood forecasting results and thus could be considered useful for such applications. The promising performance of IMERG-F V03 products was reported over the Mekong River Basin, suggesting its use for similar mountainous basins (He et al., 2017). Yuan et al. (2017) observed the propagation of IMERG-F V04 error through the Xinanjiang model, which led to significant underestimation of streamflow over the Chindwin river basin in Myanmar; in this context, TMPA showed better results. The evaluation of IMERG-E V05 and IMERG-F V05 products over the Nanliu River Basin in Tropical Humid Southern China showed that IMERG had poor performance estimating streamflow at a daily scale with NSE of <0.4 (Liang et al., 2019). In the Upper Huaihe River Basin, IMERG-F V06 had better performance in flood simulations than IMERG-E V06 and IMERG-L V06 (Su et al., 2019). Nevertheless, due to the significant underestimation of runoff (−16.51%), the authors noted that IMERG-F V05 products should be used cautiously. Yuan et al. (2019) reported significant improvements of IMERG-F V05 performance after model specific input calibration with NSE increasing from 0.66 to 0.84, and the relative bias decreasing from −32.3 to −18.5, enabling the replacement of TMPA in hydrological applications. Tan et al. (2018) found that IMERG-F V05 (NSE = 0.71 and relative bias of −5.3%) outperformed IMERG-E V05 (NSE = 0.70 and relative bias of −27.6%) and IMERG-L (NSE = 0.66 and relative bias of −36.3%) over the Kelantan river basin in Malaysia. Lu and Yong (2018) also reported IMERG-F V06's potential to estimate streamflow on a daily scale over the Yellow river basin of the Tibetan Plateau.

The number of studies validating only IMERG-F products was significantly higher than those including IMERG-E and IMERG-L. In general, IMERG-F outperformed both IMERG-E and IMERG-L products across different basins and hydrological models in terms of NSE, with values ranging between 0.60 and 0.78. This is considered highly acceptable for hydrological simulation (Moriasi et al., 2007). On the other hand, the IMERG-E and IMERG-L products show significant uncertainties between different basins and hydrological models. For example, most studies report NSE values below 0.35 in China, with only a few exceeding 0.65. Relative bias metrics follow a similar behavior to NSE, where IMERG-F outperforms the other products. IMERG-E's streamflow relative bias ranges between −59.4% and 28.4%, while IMERG-L is between −43.9% and 28%, and IMERG-F is between −23.5% and 41.4% (Fig. 8). Succeeding IMERG versions should address these issues since the near-real-time application needs such as flood, landslide, and crop forecasting is only possible for -E and -L runs, whereas the -F runs are mainly for research purposes.

Table 1
List of the studies evaluated the IMERG products in extreme events.

Location	Events	Data	Period	COR	Bias	RMSE	Reference
USA	Hurricane	IMERG-F	Aug-2017	0.61	– –	– –	(Omranian et al., 2018)
China	Typhoon	IMERG-E	2016-2017	0.61	8.38	44.97	(Huang et al., 2019)
		IMERG-F		0.57	13.50	47.50	
China	>90th percentile	IMERG-F	2014-2017	0.63	−22.82	23.52	(Fang et al., 2019)
China	Storm	IMERG-F	May-2017	0.70	−58.77	9.70	(Zhang et al., 2019)
Global	>90th percentile	IMERG-E	2014-2017	– –	−3.18	– –	(Mazzoglio et al., 2019)
		IMERG-L		– –	−3.1	– –	

Table 2
List of the studies evaluated the hydrological applications of IMERG products. (CREST = Coupled Routing and Excess Storage, VIC = Variable Infiltration Capacity, XAJ = Xinanjiang, MGB-IPH = Large Basins Model and Institute of Hydraulic Research, SWAT = Soil and Water Assessment Tool, GXAJ = Grid-based Xinanjiang hydrological model.)

Basin	Location	Model	Data	Period	NSE	Bias	Reference
Ganjiang	China	CREST	IMERG-F	May2014-Sep2014	0.77	−14.09	(Tang et al., 2016b)
Beijiang	China	VIC	IMERG-E	Apr2015-Dec2015	0.34	28.48	(Wang et al., 2017b)
			IMERG-L		0.28	28.5	
			IMERG-F		0.74	10	
Chindwin	Myanmar	XAJ	IMERG-F	Apr2014-Dec2014	0.65 to 0.72	−23.5 to −28.7	(Yuan et al., 2017)
Mahanadi	India	VIC	IMERG-F	Apr2014-Dec2014	0.64	41.4	(Beria et al., 2017)
Amazon	Peru-Ecurdor	MGB-IPH	IMERG-F	Mar2014-June2015	−24.21 to −0.9	– –	(Zubieta et al., 2017)
Mekong	China	XAJ	IMERG-F	May-Oct2015-May-Oct2016	0.53	– –	(He et al., 2017)
Ganjiang	China	CREST	IMERG-F	June2014-Sep2014	0.7	−12.6	(Li et al., 2017)
Yellow	China	VIC	IMERG-F	Jan2015-Dec2015	0.62	−7.2	(Lu and Yong, 2018)
Mishui	China	XAJ	IMERG-E	Apr2014-Dec2015	0.73	−19.52	(Jiang et al., 2018)
			IMERG-L		0.71	−25.23	
			IMERG-F		0.81	−6.53	
Kelantan	Malaysia	SWAT	IMERG-E	Mar2014-Dec2016	0.7	−27.6	(Tan et al., 2018)
			IMERG-L		0.66	−36.3	
			IMERG-F		0.71	−5.3	
Chindwin	Myanmar	GXAJ	IMERG-F	Mar2014-Dec2016	0.84	−18.5	(Yuan et al., 2019)
Huaihe	China	VIC	IMERG-E	Apr2014-Dec2015	0.18	−39.91	(Su et al., 2019)
			IMERG-L		0.16	−43.95	
			IMERG-F		0.64	−16.51	
Nanliu	China	XAJ	IMERG-F	Mar2014-Dec2016	0.28	−7.83	(Liang et al., 2019)
			IMERG-E		0.29	−59.49	
Mun-chi	Thailand	VIC	IMERG-F	Apr2014-Mar2017	−0.98	– –	(Li et al., 2019)

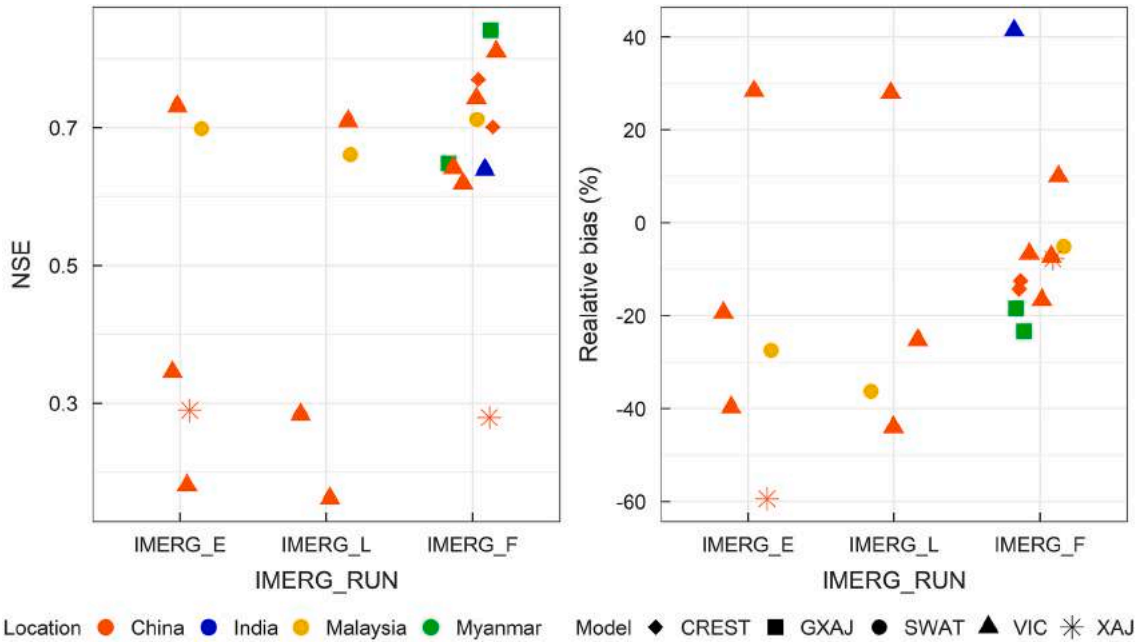


Fig. 8. IMERG performance in hydrological performances (NSE and relative bias). CREST = Coupled Routing and Excess Storage, VIC = Variable Infiltration Capacity, XAJ = Xinanjiang, SWAT = Soil and Water Assessment Tool, GXAJ = Grid-based Xinanjiang hydrological model.

4. Discussion

4.1. Recent developments

Many studies continue to assess the IMERG products and push the boundaries (e.g., spatial coverage and validation length) observed up to 2019. For instance, Islam et al. (2020) assessed the IMERG-F V06 products over Australia (no study had evaluated IMERG over Australia before 2019) on daily, monthly, and annual scales. They found IMERG performed well, despite some discrepancies over regions with high precipitation. It was also revealed that IMERG's best performance takes place during winter in terms of seasonal scale. Similarly, Tang et al. (2020) evaluated the IMERG-F V06 products between 2000 and 2018 (the first study assessing the IMERG for the TRMM period) over China and revealed that IMERG had improved its quality over time, attributed to the increase in the number of passive microwave samples. They also reported that the performance of IMERG deteriorates when it comes to snowfall, and thus further improvement in cold climates is needed.

Most of the studies showed that IMERG robustly represents the spatio-temporal patterns of precipitation (Sharma et al., 2020; Yu et al., 2020, 2020; Zhang et al., 2020b; Peng et al., 2020; Hamza et al., 2020). In addition, studies have given insights into the influence of rainfall type

239

and topography on IMERG's performance (Yang et al., 2020a; Navarro et al., 2020). For instance, Zhang et al. (2020b) revealed that IMERG-F V06 has a high correlation with gauge observations at low elevations and low at high elevations over the Tianshan Mountains, China. Yu et al. (2020) reported a decrease in correlation and increase in RMSE with the elevation over China. In contrast, Zhang et al. (2020a) reported IMERG-E V06 performance decrease from the high-altitude regions to the low-altitude regions in terms of correlation and CSI at the daily scale over the Huang plain. Liu et al. (2020) evaluated the IMERG-F V05 product over the Bali island and concluded that IMERG best detects precipitation events at different altitudes, despite some overestimation in the high altitude. Zhou et al. (2020) evaluated the IMERG-E, -L, and -F V05 products over mainland China and found more consistent performance over southern China than northern China, and more accurate at lower latitude and elevation compared to higher latitudes and elevation. The sparse number of rain gauges and light precipitation observed at the higher latitude could be the possible reason and is in line with Navarro et al. (2020), Retalis et al. (2020), Cui et al. (2020). Besides, the process and mechanism of precipitation formation over the high elevation region could be another possible reason. Moreover, the monthly gauge adjustment to final data sets reduces the elevation and other surface sensitivity uncertainties (Sui et al., 2020).

The difficulties of IMERG-F V06 estimating light rainfall were observed over UAE (Alsumaiti et al., 2020), western Pakistan (Hamza et al., 2020), and Huang plain, China (Peng et al., 2020). In addition, Abebe et al. (2020) observed a significant decline in the detection skill of IMERG-F V06 as the intensity of precipitation increases. Furthermore, Shi et al. (2020) reported that IMERG-E and -L V06 products underestimate light rain events, whereas they overestimate the moderate to heavy precipitation events concerning precipitation detection. Freitas et al. (2020) evaluated the IMERG-F V06 product for sub-daily scale and concluded that IMERG-F has to improve its rainfall intensity and duration estimates over Brazil. Yang et al. (2020b) and Li et al. (2020a) further stress that IMERG-F V06 and V05 have to improve their sub-daily scale performances. Moreover, Afonso et al. (2020) reported IMERG-F V06 better represent the diurnal cycle over the region characterized by deep convective cloud (warm rain process over land) compared to the shallow convection or low-level circulation and recommended further investigation in these perspectives.

In terms of hydrological evaluation, IMERG performance varies with the regions. For instance, Song et al. (2020) evaluated the IMERG-F V06 product over the Quing river basin China and concluded that IMERG has a satisfactory performance simulating the daily streamflow over the humid tropical climate. Similar results were reported for IMERG-F V06 over the Chenab River, Pakistan (Ahmed et al., 2020). Saouabe et al. (2020) evaluated the IMERG-E V06 product in terms of flood modeling in a semi-arid region of Morocco and concluded that IMERG-E has satisfactory performance in the simulation of flood events and can be applied for flood modeling in this climate in the absence of ground observations. In contrast, Mo et al. (2020) reported the unsatisfactory performance of IMERG-E V06 in both estimating precipitation and simulating the runoff over the Xiajia River basin, China, unless the Geographic Difference Analysis (GDA) method is used to correct IMERG data sets. Le et al. (2020) revealed that IMERG-E and -F V06 products outperformed other satellite products, and comparatively, IMERG-F showed better performance than the early product (IMERG-E).

For extreme events, IMERG-F V06 has the potential to capture the storm track and its spatial variation. However, IMERG has difficulties capturing the storm core and underestimates both high precipitation (>90 mm/hr) and accumulated precipitation (Li et al., 2020b). Getirana et al. (2020) evaluated IMERG-E and -F V06 products in terms of monitoring natural disaster/extreme precipitation events over Brazil and revealed the superior performance of IMERG-F compared to IMERG-E. Furthermore, they reported that both products have considerable skills in detecting extreme events (despite a slight underestimation of rain rate) and have the potential application in disaster detection. Li et al. (2020a) and Chen et al. (2020) also recommended the IMERG V05 and V06 near-real-time products for flood forecasting and early warning system.

4.2. Weaknesses and strengths

The latest versions of IMERG products show significant improvement over the TRMM data set. However, some discrepancies remain when compared to ground measurements. Most studies revealed that estimation of light precipitation should be improved (Lu and Yong, 2018; Anjum et al., 2018; Huang et al., 2019). Another common issue is the substantial overestimation or underestimation of precipitation over mountainous regions and its poor performance over complex topographies (Sharifi et al., 2016; Dezfuli et al., 2017; Asong et al., 2017; Kim et al., 2017; Sungmin and Kirstetter, 2018; Huang et al., 2018; Anjum et al., 2018). In addition, IMERG products show substantial bias in dry climates (Tang et al., 2016a; Su et al., 2018; Fang et al., 2019) and over ocean (Liu, 2016; Prakash et al., 2018a). Finally, when it comes to the seasonality, there are discrepancies in winter precipitation (Chen and Li, 2016; Chen et al., 2019; Lee et al., 2019). There is a general agreement that the IMERG algorithm needs further improvements in the aforementioned areas (Tang et al., 2016a; Su et al., 2018; Wang et al., 2018; Prakash et al., 2018a).

On the other hand, IMERG products perform robustly in various cases. For instance, their performance is equally good with the ground observations in the estimation and detection of regional precipitation patterns and their spatial averages (Gaona et al., 2016; Rozante et al., 2018; Palomino-, ńgel et al., 2019). In addition, IMERG has a higher detection capability of snowfall (Sadeghi et al., 2019) and light precipitation (Wang et al., 2017a, 2019a; Yang et al., 2019), compared to other satellite products (e.g., TMPA, CMORPH). Furthermore, IMERG has the potential to detect and estimate hurricane precipitation, indicates its significant applicability for estimation of precipitation during the extremes, and thus can be used for impact modeling studies (Omranian et al., 2018). All these factors reveal promising potential and a wide range of future applications.

4.3. Limitations and future perspectives

It is important to stress that all considered studies suffer from a common limitation: the very short validation period due to the lack of long-term IMERG records when the analyses were performed. Even though the evaluation methods used in most studies are reliable, there are often issues with their underlying assumptions. For instance, there are different approaches to comparing gauge measurements regarding their number, distribution, and density over the validation sites. Additionally, there is significant uncertainty associated with the point-to-area representation for gauge measurements (Dezfuli et al., 2017). The type of interpolation techniques applied to the point measurements might severely affect the evaluation outcome. It is worth mentioning that all the studies reviewed herein have employed a pixel-based approach which may result in problems like "double-penalty error" when comparing two data sets with high spatiotemporal resolution. This is one of the reasons limited studies are evaluating IMERG on sub-daily scales. Therefore, an object-based approach could be a solution to mitigate this effect. In addition, this approach could provide us with more information from the storm's characteristics like size, shape, translation speed, and direction, which cannot be accessed using a pixel-based approach. Some efforts have been made recently like Ayat et al. (2021b), and Cui et al. (2020), but more studies are needed on this topic.

Another limitation associated with the studies examined appears when they use radar observations, reanalysis results, or other satellite products in the absence of rain gauges. In such cases, the reference used may not represent the actual precipitation of the region (assumption of ground truth), and is often associated with significant uncertainties.

240

Even if the studies use gauge-based data as a reference for IMERG validation, gauges also present systematic and random errors. Therefore, we should keep in mind that it is hard to assess the "ground truth" in nature, and thus, we always make some assumptions. In addition, most often, the gauge-based data are not completely independent. For instance, IMERG-F products are adjusted with the GPCC monthly product, which itself uses gauge observations. Thus, the validation results would be biased if there are overlapping stations when compared with IMERG products. Furthermore, most studies seem to be restricted to local/regional climate; thus, the applicability of the results is limited to those particular climate conditions. Finally, most studies did not assess the performance at sub-daily resolutions but instead focused on daily, monthly, and/or annual resolutions.

Considering the identified limitations, gaps, and suggestions in the studies reviewed in this paper, we can provide some recommendations that could potentially improve the performance of IMERG products across the globe:

- Since IMERG recently extended its temporal coverage back to the TRMM era, providing precipitation observations from early 2000–present, it now offers possibilities for longer evaluation of precipitation characteristics across the globe. For instance, the long-term data will help better understand the capacity of IMERG to represent changes in the annual, inter-annual, and seasonal precipitation at the regional scale.
- To fully exploit the advantages of IMERG products' spatiotemporal resolution (30 min and 0.1° x 0.1°) compared to TMPA (3 hourly and 0.25° x 0.25°), future studies should consider sub-daily scales and assess the diurnal and semi-diurnal precipitation in different regions.
- More studies over oceans and mountainous regions could help better understand the effectiveness of IMERG in such conditions. Additionally, as very few studies have evaluated solid precipitation, future studies should focus on the detailed evaluation of snowfall. Since many studies revealed the poor performance of IMERG during the winter, these biases could be related.
- A more detailed evaluation of the hydrological performance of IMERG will be another area of future research. For instance, understanding how the error propagation occurs from precipitation to runoff (Mei et al., 2017; Ehsan Bhuiyan et al., 2019) and their quantitative analysis will help the hydrological community better understand the performance of IMERG products in hydrological applications.
- Evaluating IMERG products at multiple scales simultaneously rather than constraining the analysis to a single spatial and temporal resolution could help us understand how the accuracy and errors vary with spatiotemporal aggregation. Additionally, it will help identify the effective resolution to be used for various hydro-meteorological purposes.
- Even though few studies have evaluated IMERG versions V03, V04, and V05, the results varied with the type of study. Thus, future studies could comprehensively evaluate and answer questions on the effectiveness of the current versions compared to previous ones.
- Along with orographic precipitation and coastal areas evaluation, future studies could also consider evaluating the effect of rainfall intensity and gauge density (Maggioni et al., 2017) on the performance of IMERG products across the climatic regimes and geographical conditions across the globe.
- Even though the gauge-calibrated IMERG final run has advantages over the uncalibrated one, there is still room for improvement. For example, the scarce distribution of gauges over some regions could be solved by exploring the adjustment with other denser observational networks such as E-OBS (Ensembles-OBServation) over Europe in the future (Navarro et al., 2019). It will be beneficial to have more studies over Europe since many dense, well-maintained, observation networks exist.
- As the observational data sets are not free from uncertainty, many authors recommended considering more than one reference type to evaluate IMERG products. For example, the newly available blended data sets such as MSWEP (overland), OceanRain (over Ocean), other radar data sets, and reanalysis model data sets will help for better evaluation.
- Uncertainties from different interpolation methods used when evaluating IMERG (grid) with gauge (point) date are not well represented. Thus, comparing different interpolation techniques and their effect on the IMERG evaluation could provide a more detailed error estimation.
- An additional consideration for future research is the evaluation of IMERG for different types of storms/environmental conditions (e.g., temperature profiles) and microphysical structures (derived from polarimetric radar and NWP) (Bartsotas et al., 2018).
- Another important topic for future research is the effect of different sensors in the final merged products. Although a couple of studies like Tan et al. (2016), and Ayat et al. (2021a) have evaluated this effect in the IMERG final precipitation product, the need for further investigation remains.

5. Conclusions

This study compiled the reported performance of IMERG products across different climatic conditions and geographic locations throughout the globe. Along with IMERG performance representing precipitation, we also investigated the performance of IMERG regarding extreme precipitation events and hydrological application.

In terms of the spatial and temporal distribution of IMERG evaluation, Asia, and China in specific, are dominant in number of studies followed by North and South America, while Africa and Europe recorded the least number. Regarding the spatial and temporal resolutions, $0.1^\circ \times 0.1^\circ$, $0.25^\circ \times 0.25^\circ$ with daily, monthly and annual scales are the most evaluated resolutions of IMERG products. In addition, 12–24 months is the validation period length used in most studies. Studies with longer validation periods were often associated with coarser temporal resolution (daily and longer), whereas studies with shorter validation periods were associated with sub-daily temporal resolution. As expected, the studies at coarser time scales surpass the finer ones, highlighting the need for more research in sub-daily resolutions. Surprisingly, also very few studies exist that investigate how IMERG products perform as they move from daily to monthly scale. Understanding how biases propagate across the time scale-continuum is crucial for the proper validation and application of any data set (Markonis et al., 2021).

IMERG showed better performance compared to the TMPA estimates in the representation of spatio-temporal variability of precipitation across the climatic and geographic conditions. However, IMERG showed significant over/underestimation in different precipitation intensities that varies with region and climatic conditions. When it comes to climate regimes, IMERG tends to more consistent precipitation estimates over humid regions (wet and high intensity precipitation) compared to semi-arid and arid regions (dry and low intensity precipitation), which is especially true for China. Also, IMERG still has difficulties in estimating precipitation over complex terrains and mountainous regions. Orographic precipitation associated with high mountains is the major cause of the poor performance over such conditions. Most of the studies found that the accuracy of IMERG increased significantly with temporal aggregation, i.e., monthly results were reasonably better than daily ones, and annual results were better than monthly ones. IMERG performance is also affected by seasonal variation. Usually, IMERG performs poorly in winter compared to the summer season, which is attributed to the inability to detect light rainfall (more common in winter). Another reason could be that during winter the solid phase of precipitation (snowfall) is more common.

IMERG captures well the spatiotemporal patterns and variability of

extreme precipitation. Nonetheless, IMERG has issues when measuring over the center of the typhoon or hurricanes. In addition, it has some limitations estimating high-intensity precipitation and shows significant underestimation across various geographical locations and climatic conditions. From the perspective of the hydrological application, although most of the studies conclude that there is potential for the application of IMERG products in simulating the streamflow their performances differ depending on the hydrological model used, calibration methods, and basin types. Despite IMERG showing better detection capacity, its performance over the ocean shows substantial over/underestimation of total precipitation. However, the reliability of observational data over the ocean is questionable, and it is too early to draw any conclusions.

Overall, the performance of IMERG varies with climatic conditions, geographical locations, seasons, precipitation types, and intensities. More studies throughout the globe, especially in the regions that are under-represented, are needed for a better evaluation of IMERG's performance. It remains to be seen how the extension of the dataset back to 2000 has influenced the regional and overall performance of IMERG. This will allow also for an investigation of the climatic properties of precipitation, which will be valuable for a better quantification of the global water cycle (Vargas Godoy et al., 2021). What is most promising, though, is that each new version improves the previous one, in almost every validation metric examined in this study. Thus, despite its limitations, IMERG remains one of the most robust alternatives to ground observational records.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

RKP was supported by the Internal Grant Agency (Project no: 2020B0037), Czech University of Life Sciences Prague.

AVP acknowledges funding from project PID2019-108470RB-C21 (Agencia Estatal de Investigacion), and Grant FPI BES-2017-079685 for conducting her PhD.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2021.112754.

References

Abebe, S.A., Qin, T., Yan, D., Gelaw, E.B., Workneh, H.T., Kun, W., Shanshan, L., Biqiong, D., 2020. Spatial and temporal evaluation of the latest high-resolution precipitation products over the upper blue nile river basin, ethiopia. Water 12 (11), 3072.

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., et al., 2003. The version-2 global precipitation climatology project (gpcp) monthly precipitation analysis (1979-present). J. Hydrometeorol. 4 (6), 1147–1167.

Afonso, J.M. d. S., Vila, D.A., Gan, M.A., Quispe, D.P., Barreto, N. d. J. d. C., Huamá n Chinchay, J.H., Palharini, R.S.A., 2020. Precipitation diurnal cycle assessment of satellite-based estimates over brazil. Remote Sens. 12 (14), 2339.

Ahmed, E., Al Janabi, F., Zhang, J., Yang, W., Saddique, N., Krebs, P., 2020. Hydrologic assessment of trmm and gpm-based precipitation products in transboundary river catchment (chenab river, pakistan). Water 12 (7), 1902.

Alsumaiti, T.S., Hussein, K., Ghebreyesus, D.T., Sharif, H.O., 2020. Performance of the cmorph and gpm imerg products over the united arab emirates. Remote Sens. 12 (9), 1426.

Anjum, M.N., Ahmad, I., Ding, Y., Shangguan, D., Zaman, M., Ijaz, M.W., Sarwar, K., Han, H., Yang, M., 2019. Assessment of imerg-v06 precipitation product over different hydro-climatic regimes in the tianshan mountains, north-western china. Remote Sens. 11 (19), 2314.

Anjum, M.N., Ding, Y., Shangguan, D., Ahmad, I., Ijaz, M.W., Farid, H.U., Yagoub, Y.E., Zaman, M., Adnan, M., 2018. Performance evaluation of latest integrated multi-satellite retrievals for global precipitation measurement (imerg) over the northern highlands of pakistan. Atmos. Res. 205, 134–146.

Asong, Z., Razavi, S., Wheater, H., Wong, J., 2017. Evaluation of integrated multisatellite retrievals for gpm (imerg) over southern canada against ground precipitation observations: a preliminary assessment. J. Hydrometeorol. 18 (4), 1033–1050.

Ayat, H., Evans, J.P., Behrangi, A., 2021a. How do different sensors impact imerg precipitation estimates during hurricane days? Remote Sens. Environ. 259, 112417.

Ayat, H., Evans, J.P., Sherwood, S., Behrangi, A., 2021b. Are storm characteristics the same when viewed using merged surface radars or a merged satellite product? J. Hydrometeorol. 22 (1), 43–62.

Bartsotas, N., Anagnostou, E., Nikolopoulos, E., Kallos, G., 2018. Investigating satellite precipitation uncertainty over complex terrain. J. Geophys. Res. Atmos. 123 (10), 5346–5359.

Beck, H.E., Van Dijk, A.I., Levizzani, V., Schellekens, J., Gonzalez Miralles, D., Martens, B., De Roo, A., 2017. Mswep: 3-hourly 0.25 global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. Hydrol. Earth Syst. Sci. 21 (1), 589–615.

Beria, H., Nanda, T., Bisht, D.S., Chatterjee, C., 2017. Does the gpm mission improve the systematic error component in satellite rainfall estimates over trmm?. an evaluation at a pan-india scale. Hydrol. Earth Syst. Sci. 21 (12), 6117–6134.

Chen, C., Chen, Q., Duan, Z., Zhang, J., Mo, K., Li, Z., Tang, G., 2018. Multiscale comparative evaluation of the gpm imerg v5 and trmm 3b42 v7 precipitation products from 2015 to 2017 over a climate transition area of china. Remote Sens. 10 (6), 944.

Chen, F., Li, X., 2016. Evaluation of imerg and trmm 3b43 monthly precipitation products over mainland china. Remote Sens. 8 (6), 472.

Chen, J., Wang, Z., Wu, X., Chen, X., Lai, C., Zeng, Z., Li, J., 2019. Accuracy evaluation of gpm multi-satellite precipitation products in the hydrological application over alpine and gorge regions with sparse rain gauge network. Hydrol. Res. 50 (6), 1710–1729.

Chen, M., Nabih, S., Brauer, N.S., Gao, S., Gourley, J.J., Hong, Z., Kolar, R.L., Hong, Y., 2020. Can remote sensing technologies capture the extreme precipitation event and its cascading hydrological response?. a case study of hurricane harvey using ef5 modeling framework. Remote Sens. 12 (3), 445.

Chiaravalloti, F., Brocca, L., Procopio, A., Massari, C., Gabriele, S., 2018. Assessment of gpm and sm2rain-ascat rainfall products over complex terrain in southern italy. Atmos. Res. 206, 64–74.

Cui, W., Dong, X., Xi, B., Feng, Z., Fan, J., 2020. Can the gpm imerg final product accurately represent mcss' precipitation characteristics over the central and eastern united states? J. Hydrometeorol. 21 (1), 39–57.

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., Chang, C.-H., Chen, H., Delrieu, G., Hsu, Y.C., et al., 2019. Evaluation of gpm-era global satellite precipitation products over multiple complex terrain regions. Remote Sens. 11 (24), 2936.

Derin, Y., Yilmaz, K.K., 2014. Evaluation of multiple satellite-based precipitation products over complex topography. J. Hydrometeorol. 15 (4), 1498–1516.

Dezfuli, A.K., Ichoku, C.M., Huffman, G.J., Mohr, K.I., Selker, J.S., van de Giesen, N., Hochreutener, R., Annor, F.O., 2017. Validation of IMERG precipitation in Africa. J. Hydrometeorol. 18 (10), 2817–2825. https://doi.org/10.1175/JHM-D-17-0139.1.

Ehsan Bhuiyan, M.A., Nikolopoulos, E.I., Anagnostou, E.N., Polcher, J., Albergel, C., Dutra, E., Fink, G., Martí nez-De La Torre, A., Munier, S., 2019. Assessment of precipitation error propagation in multi-model global water resource reanalysis. Hydrol. Earth Syst. Sci. 23 (4), 1973–1994.

Fang, J., Yang, W., Luan, Y., Du, J., Lin, A., Zhao, L., 2019. Evaluation of the TRMM 3B42 and GPM IMERG products for extreme precipitation analysis over China. Atmos. Res. 223, 24–38. http://www.sciencedirect.com/science/article/pii/S0169809518311499.

Freitas, E.d.S., Coelho, V.H.R., Xuan, Y., de Melo, D., Gadelha FA.N., Santos, E.A., Galv a o, C.d.O., Ramos Filho, G.M., Barbosa, L.R., Huffman, G.J., et al., 2020. The performance of the imerg satellite-based product in identifying sub-daily rainfall events and their properties. J. Hydrol. 589, 125128.

Furl, C., Ghebreyesus, D., Sharif, H.O., 2018. Assessment of the performance of satellite-based precipitation products for flood events across diverse spatial scales using gssha modeling system. Geosciences 8 (6), 191.

Gadelha, A.N., Coelho, V.H.R., Xavier, A.C., Barbosa, L.R., Melo, D.C., Xuan, Y., Huffman, G.J., Petersen, W.A., Almeida, C.d.N., 2019. Grid box-level evaluation of imerg over brazil at various space and time scales. Atmos. Res. 218, 231–244.

Gaona, M.F.R., Overeem, A., Brasjen, A., Meirink, J.F., Leijnse, H., Uijlenhoet, R., 2017. Evaluation of rainfall products derived from satellites and microwave links for the netherlands. IEEE Trans. Geosci. Remote Sens. 55 (12), 6849–6859.

Gaona, M.R., Overeem, A., Leijnse, H., Uijlenhoet, R., 2016. First-year evaluation of gpm rainfall over the netherlands: imerg day 1 final run (v03d). J. Hydrometeorol. 17 (11), 2799–2814.

Gebregiorgis, A.S., Kirstetter, P.-E., Hong, Y.E., Gourley, J.J., Huffman, G.J., Petersen, W.A., Xue, X., Schwaller, M.R., 2018. To what extent is the day 1 gpm imerg satellite precipitation estimate improved as compared to trmm tmpa-rt? J. Geophys. Res. Atmos. 123 (3), 1694–1707.

Getirana, A., Kirschbaum, D., Mandarino, F., Ottoni, M., Khan, S., Arsenault, K., 2020. Potential of gpm imerg precipitation estimates to monitor natural disaster triggers in urban areas: The case of rio de janeiro, brazil. Remote Sens. 12 (24), 4095.

Guo, H., Chen, S., Bao, A., Behrangi, A., Hong, Y., Ndayisaba, F., Hu, J., Stepanian, P.M., 2016. Early assessment of integrated multi-satellite retrievals for global precipitation measurement over china. Atmos. Res. 176, 121–133.

Hamada, A., Murayama, Y., Takayabu, Y.N., 2014. Regional characteristics of extreme rainfall extracted from trmm pr measurements. J. Climate 27 (21), 8151–8169.

Hamza, A., Anjum, M.N., Masud Cheema, M.J., Chen, X., Afzal, A., Azam, M., Kamran Shafi, M., Gulakhmadov, A., 2020. Assessment of imerg-v06, trmm-3b42v7,

242

sm2rain-ascat, and persiann-cdr precipitation products over the hindu kush mountains of pakistan, south asia. Remote Sens. 12 (23), 3871.
He, Z., Yang, L., Tian, F., Ni, G., Hou, A., Lu, H., 2017. Intercomparisons of rainfall estimates from trmm and gpm multisatellite products over the upper mekong river basin. J. Hydrometeorol. 18 (2), 413–430.
Hong, Y., Tang, G., Ma, Y., Huang, Q., Han, Z., Zeng, Z., Yang, Y., Wang, C., Guo, X., 2019. Remote sensing precipitation: sensors, retrievals, validations, and applications. Observ. Meas. 1–23. Li, X., Vereecken, H., Eds.
Hou, A., Kakar, R., Neeck, S., Azarbarzin, AA., Kummerow, CD., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2013. The global precipitation measurement (GPM) mission. Bull. Am. Meteorol. Soc 95, 701–722.
Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The global precipitation measurement mission. Bull. Am. Meteorol. Soc 95 (5), 701–722.
Hsu, K.-l., Gao, X., Sorooshian, S., Gupta, H.V., 1997. Precipitation estimation from remotely sensed information using artificial neural networks. J. Clim. Appl. Meteorol. 36 (9), 1176–1190.
Huang, C., Hu, J., Chen, S., Zhang, A., Liang, Z., Tong, X., Xiao, L., Min, C., Zhang, Z., 2019. How well can imerg products capture typhoon extreme precipitation events over southern china? Remote Sens. 11 (1), 70.
Huang, W.-R., Chang, Y.-H., Liu, P.-Y., 2018. Assessment of imerg precipitation over taiwan at multiple timescales. Atmos. Res. 214, 239–249.
Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., Yoo, S.-H., 2015. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm Theoretical Basis Document (ATBD) Version 4, 26.
Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K.-L., Joyce, R.J., Kidd, C., Nelkin, E.J., Sorooshian, S., Stocker, E.F., Tan, J., et al., 2020. Integrated multi-satellite retrievals for the global precipitation measurement (gpm) mission (imerg). In: Satellite Precipitation Measurement. Springer, pp. 343–353.
Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P., Stocker, E.F., 2007. The trmm multisatellite precipitation analysis (tmpa): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8 (1), 38–55.
Huffman, G.J.B., Jan.2020. Reaching for 20 Years with the IMERG Multi -Satellite Products Library Catalog: NASA NTRS. https://ntrs.nasa.gov/search.jsp?R=20200002265.
Islam, M.A., Yu, B., Cartwright, N., 2020. Assessment and comparison of five satellite precipitation products in australia. J. Hydrol. 590, 125474.
Jiang, L., Bauer-Gottwein, P., 2019. How do gpm imerg precipitation estimates perform as hydrological model forcing?. evaluation for 300 catchments across mainland china. J. Hydrol. 572, 486–500.
Jiang, S., Ren, L., Xu, C.-Y., Yong, B., Yuan, F., Liu, Y., Yang, X., Zeng, X., 2018. Statistical and hydrological evaluation of the latest integrated multi-satellite retrievals for gpm (imerg) over a midlatitude humid basin in south china. Atmos. Res. 214, 418–429.
Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. Cmorph: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol. 5 (3), 487–503.
Khan, S., Maggioni, V., 2019. Assessment of level-3 gridded global precipitation mission (gpm) products over oceans. Remote Sens. 11 (3), 255.
Khodadoust Siuki, S., Saghafian, B., Moazami, S., 2017. Comprehensive evaluation of 3-hourly trmm and half-hourly gpm-imerg satellite precipitation products. Int. J. Remote Sens. 38 (2), 558–571.
Kidd, C., Levizzani, V., 2011. Status of satellite precipitation retrievals. Hydrol. Earth Syst. Sci. 15 (4).
Kim, K., Park, J., Baik, J., Choi, M., 2017. Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far -East Asia. Atmos. Res. 187, 95–105.
Kirstetter, P.-E., et al., 2018. Evaluation of diurnal variation of gpm imerg-derived summer precipitation over the contiguous us using mrms data. Q. J. R. Meteorol. Soc.
Kubota, T., Aonashi, K., Ushio, T., Shige, S., Takayabu, Y.N., Kachi, M., Arai, Y., Tashima, T., Masaki, T., Kawamoto, N., et al., 2020. Global satellite mapping of precipitation (gsmap) products in the gpm era. Satellite Precipitation Measurement 1, 355–373.
Le, M.-H., Lakshmi, V., Bolten, J., Du Bui, D., 2020. Adequacy of satellite-derived precipitation estimate for hydrological modeling in vietnam basins. J. Hydrol. 586, 124820.
Lee, J., Lee, E.-H., Seol, K.-H., 2019. Validation of integrated multisatellite retrievals for gpm (imerg) by using gauge-based analysis products of daily precipitation over east asia. Theor. Appl. Climatol. 137 (3), 2497–2512.
Levizzani, V., Cattani, E., 2019. Satellite remote sensing of precipitation and the terrestrial water cycle in a changing climate. Remote Sens. 11 (19), 2301.
Lex, A., Gehlenborg, N., 2014. Points of view: sets and intersections. Nat. Methods 11 (8), 779.
Li, L., Hong, Y., Wang, J., Adler, R.F., Policelli, F.S., Habib, S., Irwn, D., Korme, T., Okello, L., 2009. Evaluation of the real-time trmm-based multi-satellite precipitation analysis for an operational flood prediction system in nzoia basin, lake victoria, africa. Nat. Hazards 50 (1), 109–123.
Li, N., Tang, G., Zhao, P., Hong, Y., Gou, Y., Yang, K., 2017. Statistical assessment and hydrological utility of the latest multi-satellite precipitation analysis imerg in ganjiang river basin. Atmos. Res. 183, 212–223.
Li, R., Shi, J., Ji, D., Zhao, T., Plermkamon, V., Moukomla, S., Kuntiyawichai, K., Kruasilp, J., 2019. Evaluation and hydrological application of trmm and gpm precipitation products in a tropical monsoon basin of thailand. Water 11 (4), 818.
Li, X., Chen, Y., Wang, H., Zhang, Y., 2020a. Assessment of gpm imerg and radar quantitative precipitation estimation (qpe) products using dense rain gauge observations in the guangdong-hong kong-macao greater bay area, china. Atmos. Res. 236, 104834.
Li, Z., Chen, M., Gao, S., Hong, Z., Tang, G., Wen, Y., Gourley, J.J., Hong, Y., 2020b. Cross-examination of similarity, difference and deficiency of gauge, radar and satellite precipitation measuring uncertainties for extreme events using conventional metrics and multiplicative triple collocation. Remote Sens. 12 (8), 1258.
Liang, Z., Chen, S., Hu, J., Huang, C., Zhang, A., Xiao, L., Zhang, Z., Tong, X., 2019. Hydrologic evaluation of integrated multi-satellite retrievals for gpm over nanliu river basin in tropical humid southern china. Water 11 (5), 932.
Liu, C.-Y., Aryastana, P., Liu, G.-R., Huang, W.-R., 2020. Assessment of satellite precipitation product estimates over bali island. Atmos. Res. 244, 105032.
Liu, Z., 2016. Comparison of integrated multisatellite retrievals for GPM (IMERG) and TRMM multisatellite precipitation analysis (TMPA) monthly precipitation products: initial results. J. Hydrometeorol. 17 (3), 777–790.
Lu, D., Yong, B., 2018. Evaluation and hydrological utility of the latest gpm imerg v5 and gsmap v7 precipitation products over the tibetan plateau. Remote Sens. 10 (12), 2022.
Ma, Y., Yang, Y., Han, Z., Tang, G., Maguire, L., Chu, Z., Hong, Y., 2018. Comprehensive evaluation of ensemble multi-satellite precipitation dataset using the dynamic bayesian model averaging scheme over the tibetan plateau. J. Hydrol. 556, 634–644.
Maggioni, V., Meyers, P.C., Robinson, M.D., 2016. A review of merged high-resolution satellite precipitation product accuracy during the tropical rainfall measuring mission (trmm) era. J. Hydrometeorol. 17 (4), 1101–1117.
Maggioni, V., Nikolopoulos, E.I., Anagnostou, E.N., Borga, M., 2017. Modeling satellite precipitation errors over mountainous terrain: the influence of gauge density, seasonality, and temporal resolution. IEEE Trans. Geosci. Remote Sens. 55 (7), 4130–4140.
Maghsood, F.F., Hashemi, H., Hosseini, S.H., Berndtsson, R., 2020. Ground validation of gpm imerg precipitation products over iran. Remote Sens. 12 (1), 48.
Mahmoud, M.T., Al-Zahrani, M.A., Sharif, H.O., 2018. Assessment of global precipitation measurement satellite products over saudi arabia. J. Hydrol. 559, 1–12.
Mahmoud, M.T., Hamouda, M.A., Mohamed, M.M., 2019. Spatiotemporal evaluation of the gpm satellite precipitation products over the united arab emirates. Atmos. Res. 219, 200–212.
Manz, B., Pá ez-Bimos, S., Horna, N., Buytaert, W., Ochoa-Tocachi, B., Lavado-Casimiro, W., Willems, B., 2017. Comparative ground validation of imerg and tmpa at variable spatiotemporal scales in the tropical andes. J. Hydrometeorol. 18 (9), 2469–2489.
Markonis, Y., Pappas, C., Hanel, M., Papalexiou, S.M., 2021. A cross-scale framework for integrating multi-source data in earth system sciences. Environ. Modell. Software 104997.
Mayor, Y.G., Tereshchenko, I., Fonseca-Herná ndez, M., Pantoja, D.A., Montes, J.M., 2017. Evaluation of error in imerg precipitation estimates under different topographic conditions and temporal scales over mexico. Remote Sens. 9 (5), 503.
Mazzoglio, P., Laio, F., Balbo, S., Boccardo, P., Disabato, F., 2019. Improving an extreme rainfall detection system with gpm imerg data. Remote Sens. 11 (6), 677.
Mei, Y., Anagnostou, E.N., Shen, X., Nikolopoulos, E.I., 2017. Decomposing the satellite precipitation error propagation through the rainfall-runoff processes. Adv. Water Res. 109, 253–266.
Mo, C., Zhang, M., Ruan, Y., Qin, J., Wang, Y., Sun, G., Xing, Z., 2020. Accuracy analysis of imerg satellite rainfall data and its application in long-term runoff simulation. Water 12 (8), 2177.
Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.
Navarro, A., García-Ortega, E., Merino, A., Sá nchez, J.L., Kummerow, C., Tapiador, F.J., 2019. Assessment of imerg precipitation estimates over europe. Remote Sens. 11 (21), 2470.
Navarro, A., García-Ortega, E., Merino, A., Sá nchez, J.L., Tapiador, F.J., 2020. Orographic biases in imerg precipitation estimates in the ebro river basin (spain): the effects of rain gauge density and altitude. Atmos. Res. 244, 105068.
Ning, S., Song, F., Udmale, P., Jin, J., Thapa, B.R., Ishidaira, H., 2017. Error analysis and evaluation of the latest gsmap and imerg precipitation products over eastern china. Adv. Meteorol. 2017.
Oliveira, R., Maggioni, V., Vila, D., Morales, C., 2016. Characteristics and diurnal cycle of gpm rainfall estimates over the central amazon region. Remote Sens. 8 (7), 544.
Omranian, E., Sharif, H.O., Tavakoly, A.A., 2018. How well can global precipitation measurement (gpm) capture hurricanes?. case study: hurricane harvey. Remote Sens. 10 (7), 1150.
Palomino-, Á., ngel, S., Anaya-Acevedo, J.A., Botero, B.A., 2019. Evaluation of 3b42v7 and imerg daily-precipitation products for a very high-precipitation region in northwestern south america. Atmos. Res. 217, 37–48.
Peng, F., Zhao, S., Chen, C., Cong, D., Wang, Y., Ouyang, H., 2020. Evaluation and comparison of the precipitation detection ability of multiple satellite products in a typical agriculture area of china. Atmos. Res. 236, 104814.
Prakash, S., Kumar, M.R., Mathew, S., Venkatesan, R., 2018a. How accurate are satellite estimates of precipitation over the north indian ocean? Theor. Appl. Climatol. 134 (1-2), 467–475.
Prakash, S., Mitra, A.K., AghaKouchak, A., Liu, Z., Norouzi, H., Pai, D.S., 2018b. A preliminary assessment of GPM -based multi-satellite precipitation estimates over a monsoon dominated region. J. Hydrol. 556, 865–876. http://www.sciencedirect.com/science/article/pii/S0022169416000470.
Prakash, S., Mitra, A.K., Momin, I.M., Pai, D.S., Rajagopal, E.N., Basu, S., 2015. Comparison of TMPA -3B42 versions 6 and 7 precipitation products with gauge-

243

based data over India for the southwest monsoon period. J. Hydrometeorol. 16 (1), 346–362.
Rahman, K.U., Shang, S., Shahid, M., Li, J., 2018. Developing an ensemble precipitation algorithm from satellite products and its topographical and seasonal evaluations over pakistan. Remote Sens. 10 (11), 1835.
Ramsauer, T., Weiß, T., Marzahn, P., 2018. Comparison of the gpm imerg final precipitation product to radolan weather radar data over the topographically and climatically diverse germany. Remote Sens. 10 (12), 2029.
Retalis, A., Katsanos, D., Tymvios, F., Michaelides, S., 2020. Comparison of gpm imerg and trmm 3b43 products over cyprus. Remote Sens. 12 (19), 3212.
Rozante, J.R., Cavalcanti, I.F.A., 2008. Regional eta model experiments: salljex and mcs development. J. Geophys. Res. Atmos. 113 (D17).
Rozante, J.R., Vila, D.A., Barboza Chiquetto, J., Fernandes, A.d.A., Souza Alvim, D., 2018. Evaluation of trmm/gpm blended daily products over brazil. Remote Sens. 10 (6), 882.
Sadeghi, L., Saghafian, B., Moazami, S., 2019. Evaluation of imerg and mrms remotely sensed snowfall products. Int. J. Remote Sens. 40 (11), 4175–4192.
Sahlu, D., Nikolopoulos, E.I., Moges, S.A., Anagnostou, E.N., Hailu, D., 2016. First evaluation of the Day -1 IMERG over the upper Blue Nile basin. J. Hydrometeorol. 17 (11), 2875–2882.
Salles, L., Satg é, F., Roig, H., Almeida, T., Olivetti, D., Ferreira, W., 2019. Seasonal effect on spatial and temporal consistency of the new gpm-based imerg-v5 and gsmap-v7 satellite precipitation estimates in brazil's central plateau region. Water 11 (4), 668.
Saouabe, T., El Khalki, E.M., Saidi, M.E.M., Najmi, A., Hadri, A., Rachidi, S., Jadoud, M., Tramblay, Y., et al., 2020. Evaluation of the gpm-imerg precipitation product for flood modeling in a semi-arid mountainous basin in morocco. Water 12 (9), 2516.
Satg é, F., Hussain, Y., Bonnet, M.-P., Hussain, B.M., Martinez-Carvajal, H., Akhter, G., Uagoda, R., 2018. Benefits of the successive gpm based satellite precipitation estimates imerg-v03,-v04,-v05 and gsmap-v06,-v07 over diverse geomorphic and meteorological regions of pakistan. Remote Sens. 10 (9), 1373.
Sharifi, E., Steinacker, R., Saghafian, B., 2016. Assessment of gpm-imerg and other precipitation products against gauge data under different topographic and climatic conditions in iran: preliminary results. Remote Sens. 8 (2), 135.
Sharma, S., Chen, Y., Zhou, X., Yang, K., Li, X., Niu, X., Hu, X., Khadka, N., 2020. Evaluation of gpm-era satellite precipitation products on the southern slopes of the central himalayas against rain gauge data. Remote Sens. 12 (11), 1836.
Shawky, M., Moussa, A., Hassan, Q.K., El-Sheimy, N., 2019. Performance assessment of sub-daily and daily precipitation estimates derived from gpm and gsmap products over an arid environment. Remote Sens. 11 (23), 2840.
Shi, J., Yuan, F., Shi, C., Zhao, C., Zhang, L., Ren, L., Zhu, Y., Jiang, S., Liu, Y., 2020. Statistical evaluation of the latest gpm-era imerg and gsmap satellite precipitation products in the yellow river source region. Water 12 (4), 1006.
Siddique-E-Akbor, A., Hossain, F., Sikder, S., Shum, C., Tseng, S., Yi, Y., Turk, F., Limaye, A., 2014. Satellite precipitation data-driven hydrological modeling for water resources management in the ganges, brahmaputra, and meghna basins. Earth Interact. 18 (17), 1–25.
Singh, A.K., Tripathi, J., Singh, K., Singh, V., Sateesh, M., 2019. Comparison of different satellite-derived rainfall products with imd gridded data over indian meteorological subdivisions during indian summer monsoon (ism) 2016 at weekly temporal resolution. J. Hydrol. 575, 1371–1379.
Song, Y., Zhang, J., Meng, X., Zhou, Y., Lai, Y., Cao, Y., 2020. Comparison study of multiple precipitation forcing data on hydrological modeling and projection in the qujiang river basin. Water 12 (9), 2626.
Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H.V., Imam, B., Braithwaite, D., 2000. Evaluation of persiann system satellite-based estimates of tropical rainfall. Bull. Am. Meteorol. Soc 81 (9), 2035–2046.
Su, J., Lü, H., Zhu, Y., Cui, Y., Wang, X., 2019. Evaluating the hydrological utility of latest imerg products over the upper huaihe river basin, china. Atmos. Res. 225, 17–29.
Su, J., Lü, H., Zhu, Y., Wang, X., Wei, G., 2018. Component analysis of errors in four gpm-based precipitation estimations over mainland china. Remote Sens. 10 (9), 1420.
Sui, X., Li, Z., Ma, Z., Xu, J., Zhu, S., Liu, H., 2020. Ground validation and error sources identification for gpm imerg product over the southeast coastal regions of china. Remote Sens. 12 (24), 4154.
Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.-L., 2018. A review of global precipitation data sets: data sources, estimation, and intercomparisons. Rev. Geophys. 56 (1), 79–107.
Sungmin, O., Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J., Petersen, W.A., 2017. Evaluation of gpm imerg early, late, and final rainfall estimates using wegenernet gauge data in southeastern austria. Hydrol. Earth Syst. Sci. 21 (12).
Sungmin, O., Kirstetter, P.-E., 2018. Evaluation of diurnal variation of gpm imerg-derived summer precipitation over the contiguous us using mrms data. Q. J. R. Meteorol. Soc. 144, 270–281.
Sunilkumar, K., Yatagai, A., Masuda, M., 2019. Preliminary evaluation of gpm-imerg rainfall estimates over three distinct climate zones with aphrodite. Earth Space Sci. 6 (8), 1321–1335.
Tan, J., Huffman, G.J., Bolvin, D.T., Nelkin, E.J., 2019. Imerg v06: changes to the morphing algorithm. J. Atmos. Ocean. Technol. 36 (12), 2471–2482.
Tan, J., Petersen, W.A., Kirstetter, P.-E., Tian, Y., 2017. Performance of imerg as a function of spatiotemporal scale. J. Hydrometeorol. 18 (2), 307–319.
Tan, J., Petersen, W.A., Tokay, A., 2016. A novel approach to identify sources of errors in imerg for gpm ground validation. J. Hydrometeorol. 17 (9), 2477–2491.
Tan, M.L., Duan, Z., 2017. Assessment of GPM and TRMM precipitation products over Singapore. Remote Sens. 9 (7), 720.
Tan, M.L., Samat, N., Chan, N.W., Roy, R., 2018. Hydro-meteorological assessment of three gpm satellite precipitation products in the kelantan river basin, malaysia. Remote Sens. 10 (7), 1011.
Tan, M.L., Santo, H., 2018. Comparison of gpm imerg, tmpa 3b42 and persiann-cdr satellite precipitation products over malaysia. Atmos. Res. 202, 63–76.
Tang, G., Clark, M.P., Papalexiou, S.M., Ma, Z., Hong, Y., 2020. Have satellite precipitation products improved over last two decades? a comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. Remote Sens. Environ. 240, 111697 http://www.sciencedirect.com/science/article/pii/S0034425720300663.
Tang, G., Ma, Y., Long, D., Zhong, L., Hong, Y., 2016a. Evaluation of GPM Day -1 IMERG and TMPA version -7 legacy products over Mainland China at multiple spatiotemporal scales. J. Hydrol. 533, 152–167.
Tang, G., Zeng, Z., Long, D., Guo, X., Yong, B., Zhang, W., Hong, Y., 2016b. Statistical and hydrological comparisons between trmm and gpm level-3 products over a midlatitude basin: is day-1 imerg a good successor for tmpa 3b42v7? J. Hydrometeorol. 17 (1), 121–137.
Tapiador, F., Navarro, A., Levizzani, V., Garcí a-Ortega, E., Huffman, G., Kidd, C., Kucera, P., Kummerow, C., Masunaga, H., Petersen, W., et al., 2017. Global precipitation measurements for validating climate models. Atmos. Res. 197, 1–20.
Vargas Godoy, M.R., Markonis, Y., Hanel, M., Kyselý, J., Papalexiou, S.M., 2021. The global water cycle budget: a chronological review. Surv. Geophys. 1–33.
Wang, C., Tang, G., Han, Z., Guo, X., Hong, Y., 2018. Global intercomparison and regional evaluation of GPM IMERG Version -03, Version -04 and its latest version -05 precipitation products: similarity, diffrence and improvements. J. Hydrol. 564, 342–356.
Wang, D., Wang, X., Liu, L., Wang, D., Huang, H., Pan, C., 2019a. Evaluation of tmpa 3b42v7, gpm imerg and cmpa precipitation estimates in guangdong province, china. Int. J. Climatol. 39 (2), 738–755.
Wang, S., Liu, J., Wang, J., Qiao, X., Zhang, J., 2019b. Evaluation of gpm imerg v05b and trmm 3b42v7 precipitation products over high mountainous tributaries in lhasa with dense rain gauges. Remote Sens. 11 (18), 2080.
Wang, W., Lu, H., Zhao, T., Jiang, L., Shi, J., 2017a. Evaluation and comparison of daily rainfall from latest gpm and trmm products over the mekong river basin. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. 10 (6), 2540–2549.
Wang, X., Ding, Y., Zhao, C., Wang, J., 2019c. Similarities and improvements of gpm imerg upon trmm 3b42 precipitation product under complex topographic and climatic conditions over hexi region, northeastern tibetan plateau. Atmos. Res. 218, 347–363.
Wang, Z., Zhong, R., Lai, C., Chen, J., 2017b. Evaluation of the gpm imerg satellite-based precipitation products and the hydrological utility. Atmos. Res. 196, 151–163.
Watters, D., Battaglia, A., 2019. The summertime diurnal cycle of precipitation derived from imerg. Remote Sens. 11 (15), 1781.
Wei, G., Lü, H., Crow, T.W., Zhu, Y., Wang, J., Su, J., 2018. Evaluation of satellite-based precipitation products from imerg v04a and v03d, cmorph and tmpa with gauged rainfall in three climatologic zones in china. Remote Sens. 10 (1), 30.
Wen, Y., Behrangi, A., Chen, H., Lambrigtsen, B., 2018. How well were the early 2017 california atmospheric river precipitation events captured by satellite products and ground-based radars? Q. J. R. Meteorol. Soc. 144, 344–359.
Wen, Y., Behrangi, A., Lambrigtsen, B., Kirstetter, P.-E., 2016. Evaluation and uncertainty estimation of the latest radar and satellite snowfall products using snotel measurements over mountainous regions in western united states. Remote Sens. 8 (11), 904.
Wu, L., Xu, Y., Wang, S., 2018. Comparison of tmpa-3b42rt legacy product and the equivalent imerg products over mainland china. Remote Sens. 10 (11), 1778.
Xu, F., Guo, B., Ye, B., Ye, Q., Chen, H., Ju, X., Guo, J., Wang, Z., 2019a. Systematical evaluation of gpm imerg and trmm 3b42v7 precipitation products in the Huang-Huai-Hai plain, China. Remote Sens. 11 (6), 697.
Xu, J., Ma, Z., Tang, G., Ji, Q., Min, X., Wan, W., Shi, Z., 2019b. Quantitative evaluations and error source analysis of fengyun-2-based and gpm-based precipitation products over mainland china in summer, 2018. Remote Sens. 11 (24), 2992.
Xu, R., Tian, F., Yang, L., Hu, H., Lu, H., Hou, A., 2017. Ground validation of gpm imerg and trmm 3b42v7 rainfall products over southern tibetan plateau based on a high-density rain gauge network. J. Geophys. Res. Atmos. 122 (2), 910–924.
Xu, S., Shen, Y., Niu, Z., 2019c. Evaluation of the imerg version 05b precipitation product and comparison with imerg version 04a over mainland china at hourly and daily scales. Adv. Space Res. 63 (8), 2387–2398.
Yang, M., Li, Z., Anjum, M.N., Gao, Y., 2019. Performance evaluation of version 5 (v05) of integrated multi-satellite retrievals for global precipitation measurement (imerg) over the tianshan mountains of china. Water 11 (6), 1139.
Yang, M., Liu, G., Chen, T., Chen, Y., Xia, C., 2020a. Evaluation of gpm imerg precipitation products with the point rain gauge records over sichuan, china. Atmos. Res. 246, 105101.
Yang, X., Lu, Y., Tan, M.L., Li, X., Wang, G., He, R., 2020b. Nine-year systematic evaluation of the gpm and trmm precipitation products in the shuaishui river basin in east-central china. Remote Sens. 12 (6), 1042.
Yu, C., Hu, D., Liu, M., Wang, S., Di, Y., 2020. Spatio-temporal accuracy evaluation of three high-resolution satellite precipitation products in china area. Atmos. Res. 241, 104952.
Yuan, F., Zhang, L., Soe, K.M.W., Ren, L., Zhao, C., Zhu, Y., Jiang, S., Liu, Y., 2019. Applications of trmm-and gpm-era multiple-satellite precipitation products for flood simulations at sub-daily scales in a sparsely gauged watershed in myanmar. Remote Sens. 11 (2), 140.
Yuan, F., Zhang, L., Win, K.W.W., Ren, L., Zhao, C., Zhu, Y., Jiang, S., Liu, Y., 2017. Assessment of gpm and trmm multi-satellite precipitation products in streamflow

simulations in a data-sparse mountainous watershed in myanmar. Remote Sens. 9 (3), 302.

Zhang, A., Xiao, L., Min, C., Chen, S., Kulie, M., Huang, C., Liang, Z., 2019. Evaluation of latest gpm-era high-resolution satellite precipitation products during the may 2017 guangdong extreme rainfall event. Atmos. Res. 216, 76–85.

Zhang, J., Lin, L.-F., Bras, R.L., 2018. Evaluation of the quality of precipitation products: a case study using wrf and imerg data over the central united states. J. Hydrometeorol. 19 (12), 2007–2020.

Zhang, L., Li, X., Cao, Y., Nan, Z., Wang, W., Ge, Y., Wang, P., Yu, W., 2020a. Evaluation and integration of the top-down and bottom-up satellite precipitation products over mainland china. J. Hydrol. 581, 124456.

Zhang, Y., Hanati, G., Danierhan, S., Liu, Q., Xu, Z., 2020b. Evaluation and comparison of daily gpm/trmm precipitation products over the tianshan mountains in china. Water 12 (11), 3088.

Zhao, H., Yang, S., You, S., Huang, Y., Wang, Q., Zhou, Q., 2018. Comprehensive evaluation of two successive v3 and v4 imerg final run precipitation products over mainland china. Remote Sens. 10 (1), 34.

Zhou, Z., Guo, B., Xing, W., Zhou, J., Xu, F., Xu, Y., 2020. Comprehensive evaluation of latest gpm era imerg and gsmap precipitation products over mainland china. Atmos. Res. 246, 105132.

Zubieta, R., Getirana, A., Espinoza, J.C., Lavado-Casimiro, W., Aragon, L., 2017. Hydrological modeling of the peruvian-ecuadorian amazon basin using gpm-imerg satellite-based precipitation dataset. Hydrol. Earth Syst. Sci. 21 (7), 3543–3555.