Czech University of Life Sciences Prague
Department of Water Resources and Environmental Modeling
Environmental Modelling

The spatial interpolation of point temperature measurements in Amalia pilot site

Master thesis of:
Rodrigo Manzanero Villoria

Supervisor:
doc. Ing. Petr Maca, Ph.D.

31.03.2026

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Environmental Sciences

DIPLOMA THESIS ASSIGNMENT

Bachelor of Science Rodrigo Manzanero Villoria

Environmental Modelling

Thesis title

The spatial interpolation of point temperature measurements in Amalia pilot site

---

Objectives of thesis

The thesis aims to create a spatial field of temperature datasets for the Amalia pilot site. A suitable spatial interpolation model will be selected for this purpose. The leave-one-out procedure will be applied to calibrate and validate the spatial interpolation model.

Methodology

The methodology consists of the following steps:

1. Data preparation: available time series of available measurements from the Amalia site will be prepared
2. The selected spatial interpolation models suitable for temperature data will be implemented
3. The leave-one-out analyses will be performed to calibrate and validate the tested spatial interpolation models
4. The final fields of temperature data will be simulated using the best spatial interpolation model

The proposed extent of the thesis

standard

Keywords

interception, semi-distributed hydrological model, Rutter model

---

Recommended information sources

BEVEN, K. J. *Rainfall-runoff modelling : the primer.* Chichester: Wiley-Blackwell, 2012. ISBN 978-0-470-71459-1.

Goovaerts, P. (1997). Geostatistics for natural resources evaluation. New York (N.Y.): Oxford university press.

Web Science, Scopus and others

Expected date of thesis defence

2025/26 LS – FZP

Thesis supervisor

doc. Ing. Petr Máca, Ph.D.

Supervising department

Department of Water Resources and Environmental Modeling

Electronic approval: 18. 12. 2024

prof. Ing. Martin Hanel, Ph.D.

Head of department

Electronic approval: 18. 12. 2024

prof. RNDr. Michael Komárek, Ph.D.

Dean

Prague on 02. 12. 2025

I hereby declare that I have independently elaborated the diploma/final thesis with the topic of: "The spatial interpolation of point temperature measurements in Amalia pilot site" and that I have cited all the information sources that I used in the thesis and that are also listed at the end of the thesis in the list of used information sources. I am aware that my diploma thesis is subject to Act No. 121/2000 Coll., on copyright, on rights related to copyright and on amendment of some acts, as amended by later regulations, particularly the provisions of Section 35(3) of the act on the use of the thesis. I am aware that by submitting the diploma/final thesis I agree with its publication under Act No. 111/1998 Coll., on universities and on the change and amendments of some acts, as amended, regardless of the result of its defence. With my own signature, I also declare that the electronic version is identical to the printed version and the data stated in the thesis has been processed in relation to the GDPR.

Prague, 31.03.2026.

Rodrigo Manzanero Villoria

Author's Acknowledgement

I would like to express my sincere appreciation to doc. Ing. Petr Máca, Ph.D., for his professional advice, dedication, and support during the preparation of this thesis. He strongly encouraged me to pursue ambitious and technically challenging work.

I would like to thank Ing. Marta Kuželková for her assistance in providing the updated dataset, answering questions regarding its structure, and taking me on a field trip to the Amálie pilot farm.

I would also like to express my appreciation to the Faculty of Environmental Sciences, particularly the International Relations Office, for providing valuable opportunities during my study period, which greatly enriched my academic experience.

Computational resources were supplied by the project *e-Infrastruktura CZ* (e-INFRA CZ LM2018140), provided within the programme Projects of Large Research, Development and Innovations Infrastructures.

Dedicated to my parents and my brother.

Abstract

Temperature fields exhibit strong spatial variability driven by environmental and topographic factors. Accurate spatial interpolation of high-frequency temperature observations is therefore essential for understanding local thermal dynamics. This study evaluates geostatistical approaches for interpolating temperature measured at 15-minute intervals from a sensor network deployed at the Amálie pilot farm in the Czech Republic. Two empirical variogram estimators were compared and four theoretical variogram models were fitted to characterize the spatial covariance structure. The spatial domain was divided into four landscape units to account for differences in monitoring network density and environmental conditions. In some landscapes, seasonal and diurnal stratification was applied to define regime-specific subsets, while others were treated as stationary. Two geostatistical interpolation methods—Ordinary kriging (OK) and kriging with external drift (KED)—were implemented to reconstruct spatial temperature fields. Elevation and other terrain attributes derived from a Digital Elevation Model (DEM) were tested as auxiliary predictors. Results indicate that incorporating topographic information improves interpolation accuracy in some landscapes, while OK performs better in others. The optimal neighborhood size was selected using leave-one-out cross-validation (LOOCV), and model performance was evaluated using the mean absolute error (MAE) and root mean square error (RMSE). The optimal variogram models and interpolation strategies were identified for each landscape unit, enabling the reconstruction of high-resolution temperature fields across the study area at 15-minute temporal resolution.

Keywords: spatial interpolation, variogram modeling, ordinary kriging, kriging with external drift, high-frequency temperature observations

Abstract

Teplotní pole vykazují výraznou prostorovou variabilitu řízenou environmentálními a topografickými faktory. Přesná prostorová interpolace vysokofrekvenčních teplotních pozorování je proto nezbytná pro pochopení lokální tepelné dynamiky. Tato studie hodnotí geostatistické přístupy k interpolaci teplot měřených v 15minutových intervalech ze sítě senzorů rozmístěné na pilotní farmě Amálie v České republice. Byly porovnány dva empirické odhady variogramu a čtyři teoretické variogramové modely byly proloženy za účelem charakterizace prostorové kovarianční struktury. Prostorová doména byla rozdělena do čtyř krajinných jednotek, aby se zohlednily rozdíly v hustotě monitorovací sítě a environmentálních podmínkách. V některých krajinných jednotkách byla použita sezónní a denní stratifikace k definování podmnožin specifických pro daný režim, zatímco ostatní byly považovány za stacionární. K rekonstrukci prostorových teplotních polí byly použity dvě geostatistické interpolační metody – obyčejný kriging (OK) a kriging s externím driftem (KED). Jako pomocné prediktory byly testovány nadmořská výška a další terénní atributy odvozené z digitálního modelu terénu (DEM). Výsledky ukazují, že zahrnutí topografických informací zlepšuje přesnost interpolace v některých krajinných jednotkách, zatímco v jiných dosahuje lepších výsledků metoda OK. Optimální velikost sousedství byla zvolena pomocí křížové validace typu "leave-one-out" (LOOCV) a výkonnost modelu byla hodnocena pomocí průměrné absolutní chyby (MAE) a odmocniny střední kvadratické chyby (RMSE). Pro každou krajinnou jednotku byly identifikovány optimální variogramové modely a interpolační strategie, což umožnilo rekonstrukci teplotních polí s vysokým prostorovým rozlišením v celé studované oblasti při časovém rozlišení 15 minut.

Klíčová slova: prostorová interpolace, modelování variogramu, obyčejný kriging, kriging s externím driftem, vysokofrekvenční teplotní pozorování

Contents

List of Figures **vii**

List of Tables **x**

1 Introduction **1**

2 Objectives **4**

3 Study area **5**

3.1 Smart Landscape on the Amálie Pilot Farm 5

3.2 Physical Characteristics 5

3.3 Digital Elevation Model 6

4 Data **7**

4.1 Original Dataset Description 7

4.2 Data Preparation and Structural Assessment 7

4.3 Co-located Sensors Aggregation 9

4.4 Data Quality Assessment and Exploratory Analysis 9

4.5 Final Dataset for Variogram Diagnosis and Modeling 12

5 Methodology **16**

5.1 Theoretical Background 16

5.2 Spatial Regime Diagnosis and Modeling Framework 26

5.3 Computational Implementation 36

6 Results **37**

6.1 Thermal Characteristics of the Aggregated Dataset 37

6.2 Diagnostic Spatial Structure 38

6.3 Variogram Estimation Stability 40

6.4 Temporal Behavior of Variogram Parameters 41

6.5 Regime-Specific Data Availability 46

6.6 Bootstrap Assessment of Variogram Parameter Stability 47

6.7 Final Variogram Models 51

6.8 Kriging Cross-Validation Performance 56

7 Discussion **67**

8 Conclusions **70**

Bibliography **72**

Appendix A: Bootstrap Stability Figures **1**

Appendix B: Additional Variogram Models **3**

Appendix C: Additional Spatial Interpolation Maps **7**

List of Figures

4.1 Boxplots of aggregated temperature observations by landscape and measurement level (T1–T3) 11
4.2 Skewness of aggregated temperature observations by landscape and measurement level (T1–T3) 11
4.3 Digital Elevation Model (DEM) and spatial distribution of temperature sensors across the Amálie pilot farm 13
4.4 Spatial distribution of temperature sensors across the Amálie pilot farm 14

6.1 Representative empirical variograms of spatial structure in the monitoring networks 39
6.2 Temporal behaviour of nugget and sill parameters in the Agriculture landscape 42
6.3 Temporal behaviour of nugget and sill parameters in the Drainage landscape 43
6.4 Temporal behaviour of nugget and sill parameters in the BP landscape 44
6.5 Temporal behaviour of nugget and sill parameters in the KL landscape 45
6.6 Bootstrap stability diagnostics of the partial sill parameter for the Agriculture landscape 48
6.7 Bootstrap stability diagnostics of the partial sill parameter for the Drainage landscape 49
6.8 Bootstrap stability diagnostics of the partial sill parameter for the KL landscape 49
6.9 Bootstrap stability diagnostics of the partial sill parameter for the BP landscape 50
6.10 Representative variogram model fits for the Agriculture landscape during the warm-midday regime 53
6.11 Representative variogram model fits for the Agriculture landscape during the cold-night regime 54
6.12 Representative variogram model fits for the Drainage landscape during the warm-midday regime 55

6.13 Representative variogram model fits for the BP landscape under stationary conditions . . . 55
6.14 MAE comparison across kriging models, landscapes, and regimes . . . 58
6.15 RMSE comparison across kriging models, landscapes, and regimes . . 59
6.16 Spatial prediction of surface temperature (T3) during the warm season in Agriculture . . . 60
6.17 Spatial prediction of surface temperature (T3) during the cold season in Agriculture . . . 61
6.18 Spatial prediction of surface temperature (T3) during the warm season in Drainage . . . 62
6.19 Spatial prediction of surface temperature (T3) during the cold season in Drainage . . . 63
6.20 Spatial prediction of temperature at measurement levels (T1–T3) under stationary conditions in KL . . . 64
6.21 Spatial prediction at measurement levels (T1–T3) under stationary conditions in BP . . . 65

A.1 Bootstrap stability diagnostics of the nugget parameter for the Zemedelska landscape . . . 1
A.2 Bootstrap stability diagnostics of the nugget parameter for the Drainage landscape . . . 1
A.3 Bootstrap stability diagnostics of the nugget parameter for the KL landscape . . . 2
A.4 Bootstrap stability diagnostics of the nugget parameter for the BP landscape . . . 2
B.1 Representative variogram model fits for the Agriculture landscape during the warm-transition regime . . . 3
B.2 Representative variogram model fits for the Agriculture landscape during the warm-night regime . . . 3
B.3 Representative variogram model fits for the Agriculture landscape during the cold-midday regime . . . 4
B.4 Representative variogram model fits for the Agriculture landscape during the cold-transition regime . . . 4
B.5 Representative variogram model fits for the Drainage landscape during the warm-transition regime . . . 4
B.6 Representative variogram model fits for the Drainage landscape during the warm-night regime . . . 5
B.7 Representative variogram model fits for the Drainage landscape during the cold-midday regime . . . 5

B.8 Representative variogram model fits for the Drainage landscape during the cold-transition regime 5
B.9 Representative variogram model fits for the Drainage landscape during the cold-night regime 6
B.10 Representative variogram model fits for the KL landscape under stationary conditions 6
C.1 Spatial prediction of subsurface (soil) temperature (T2) during the warm season in Agriculture 8
C.2 Spatial prediction of deeper soil temperature (T1) during the warm season in Agriculture 9
C.3 Spatial prediction of subsurface (soil) temperature (T2) during the cold season in Agriculture 10
C.4 Spatial prediction of deeper soil temperature (T1) during the cold season in Agriculture 11
C.5 Spatial prediction of subsurface (soil) temperature (T2) during the warm season in Drainage 12
C.6 Spatial prediction of subsurface (soil) temperature (T2) during the cold season in Drainage 13
C.7 Spatial prediction of deeper soil temperature (T1) during the warm season in Drainage 14
C.8 Spatial prediction of deeper soil temperature (T1) during the cold season in Drainage 15

List of Tables

4.1 Depth ranges of temperature sensors (T1, T2, T3) across different landscapes 8
4.2 Occurrence of co-located sensor installations across landscapes. 9
4.3 Data completeness across landscapes and measurement levels 10
4.4 Spatial characteristics of the monitoring networks across landscapes . 12
4.5 Temporal coverage and completeness of temperature observations across landscapes (T1–T3) 15

5.1 Variogram fitting configuration thresholds determined during the pilot variogram analysis. 28

6.1 Descriptive statistics of temperature observations across landscapes (T1–T3) 37
6.2 Variogram fitting diagnostics from parameter stability analysis 40
6.3 Summary of regime stratification decisions across landscapes 46
6.4 Number of timestamps available for each landscape–regime combination 46
6.5 Selected timestamp pool sizes from bootstrap stability diagnostics for each landscape 51
6.6 Final variogram models selected for each landscape, regime, and temperature measurement 52
6.7 Best-performing kriging models based on cross-validation 57

CHAPTER 1

Introduction

Spatial interpolation methods comprise a group of techniques that allow the prediction of a variable at unsampled spatial locations based on observations from nearby known locations (Cressie, 1990). In recent literature, they are often referred to as spatial predictive methods (Li, 2019). In spatial statistics, kriging has become synonymous with optimal spatial prediction based on spatial autocorrelation among observations (Lam, 1983). Originally developed by Danie G. Krige (1951) for ore reserve estimation in mining applications, kriging has since evolved into a fundamental tool in geostatistical analysis.

Since its development, spatial interpolation has advanced considerably, with major progress in both theoretical and computational approaches. In recent years, machine learning techniques such as Random Forests and neural networks have been increasingly applied to spatial prediction problems, in some cases achieving predictive performance comparable to traditional geostatistical approaches (Hengl et al., 2018; Zhang et al., 2022). Other modeling frameworks, including spatial auto–regressive models (SAR) and Markov Random Fields (MRF), further extend the range of available predictive tools (Xie et al., 2017). Noteworthy, a comprehensive review by Li and Heap (2014) identified and classified 25 commonly used spatial interpolation methods applied in environmental sciences, illustrating both the diversity of available approaches and the continuing methodological development of the field.

Despite these advances, spatial interpolation remains an active area of research, with ongoing challenges related to capturing complex relationships between geographical location and the target variable, as well as minimizing information loss during spatial representation (Zhang et al., 2022).

A key difficulty in spatial interpolation lies in selecting an appropriate method for a given dataset, as model performance depends strongly on data characteristics, sampling design, and environmental context (Huang and Chen, 2007; Li and Heap,

2011). In this context, the aim of this thesis is not to identify a universally optimal interpolation method, but rather to develop a consistent and reproducible workflow for systematically evaluating the stages of the interpolation process. Such frameworks are essential for ensuring transparency and comparability in spatial prediction studies (Meyer et al., 2019), while enabling the methodology to be replicated and extended to other study areas.

One of the challenges addressed in this thesis is the spatial interpolation of high spatio-temporal resolution temperature observations. Such data are critical for understanding microclimatic processes, which are influenced by environmental factors such as topography, soil properties, and vegetation (Bramer et al., 2018), and which in turn shape the ecology of plants and animals and broader ecosystem dynamics (Kemppinen et al., 2024). For example, Kuželková et al. (2024) documented significant fluctuations in soil moisture regimes within forest landscapes during a multi-year monitoring period, highlighting the strong influence of microclimatic variability. Accurate spatial reconstruction of temperature fields therefore provides valuable insights for environmental monitoring, agricultural management, and agroforestry systems. In particular, agroforestry practices such as those implemented in the Amálie pilot farm contribute to improving soil quality, reducing erosion, and enhancing ecosystem resilience (Chalupecká et al., 2026).

A fundamental concept in geostatistical modeling is the variogram, which describes the spatial dependence structure of a variable. In practice, two forms are commonly distinguished: the experimental variogram, computed from observed data, and the theoretical variogram model, which provides a parametric representation of spatial dependence (Goovaerts, 1997; Webster and Oliver, 2007). These components form the basis of kriging and related geostatistical interpolation methods.

In this thesis, two empirical variogram estimators are considered: the classical estimator proposed by Matheron and the robust estimator introduced by Cressie and Hawkins. Additional robust alternatives have been proposed in the literature to address datasets affected by outliers (Genton, 1998); however, these are beyond the scope of this study. Four theoretical variogram models are evaluated: spherical, Gaussian, exponential, and Matérn. In addition, two geostatistical interpolation approaches from the kriging family are implemented and compared: Ordinary kriging (OK) and kriging with external drift (KED). In the latter, environmental variables are incorporated as auxiliary predictors to improve interpolation accuracy (Li, 2019), with digital elevation models commonly used in temperature interpolation. KED is therefore widely recommended for incorporating such auxiliary information in environmental applications (Li and Heap, 2011; Appelhans et al., 2015).

Following the definition of the research objectives, the thesis is organized into several chapters. First, the study area in which the application is developed is described, located within the Smart Landscape I pilot project (Czech University of Life Sciences Prague, 2024). Second, the dataset used in the analysis is explored and assessed in order to produce a final dataset suitable for spatial modeling. Exploratory data analysis techniques, including graphical methods such as box plots, are used to identify potential outliers and data inconsistencies (Goovaerts, 1997; Webster and Oliver, 2007). More recently, computational approaches for outlier detection in geospatial data have been proposed, enabling more robust data preprocessing in complex datasets (Xie et al., 2017).

The methodology chapter is divided into three main sections. The first section presents the theoretical background, including the formulation of the experimental variogram, the theoretical variogram models, and the spatial interpolation methods used in this study. The second section focuses on spatial regime diagnosis and modeling, describing the process from the initial stages of variogram analysis to the estimation of the final regime-specific variogram prior to the implementation of kriging. The third section describes the computational implementation, outlining how the analyses were executed using high-performance computing resources, which facilitate the processing of large spatial datasets (Li, 2019).

The subsequent chapter presents the results, including the selection of variogram models, the implementation of kriging cross-validation, and the reconstruction of temperature fields using the best-performing spatial interpolation model. Recent advances include spatial cross-validation strategies that account for spatial autocorrelation, such as block cross-validation and related approaches (Roberts et al., 2017; Ploton et al., 2020). Furthermore, machine learning frameworks such as Random Forest have been used in combination with spatial and non-spatial cross-validation strategies to improve model generalization and predictive performance (Meyer et al., 2019).

Finally, the thesis concludes with a discussion of the results, their implications, and the main conclusions of the study.

CHAPTER 2

Objectives

The main objective of this thesis is to reconstruct spatial temperature fields for the Amálie pilot farm by identifying the most suitable geostatistical interpolation model for each landscape unit.

The specific objectives are:

- to compare two empirical variogram estimators and four theoretical variogram models in order to characterize the spatial covariance structure;
- to determine the optimal neighborhood size using leave-one-out cross-validation (LOOCV);
- to evaluate the performance of the interpolation models using the mean absolute error (MAE) and root mean square error (RMSE).

CHAPTER 3

Study area

3.1 Smart Landscape on the Amálie Pilot Farm

There are currently three Smart Landscape pilot projects in the Czech Republic focused on drought and flood mitigation (Czech University of Life Sciences Prague, 2024). This diploma thesis focuses on Smart Landscape I, which comprises a combination of agricultural and forested land in the Rakovník region, where all land plots are owned by Česká zemědělská univerzita v Praze (ČZU).

The pilot project is in close cooperation with the Ministry of the Environment, the Ministry of Agriculture, and the Ministry of Regional Development.

The basic principles of the Smart Landscape concept include the optimization of water management, the minimization of the impacts of hydrological extremes such as droughts and floods, the pursuit of sustainable intensification of agriculture and forestry, the reduction of soil erosion, the enhancement of biodiversity, and the increase in landscape value.

In general terms, it represents a system that integrates water management and soil protection based on the principles of sustainable land use.

The study area has an average slope of 10% and an average elevation of 420 m.a.s.l., and is divided into two main watersheds: Brejlský Potok (BP) and Karlův Luh (KL).

3.2 Physical Characteristics

Climate

According to meteorological data from 1993 to 2023, the average annual precipitation in Central Bohemia at the Nové Strašecí station is 595 mm. At the Kněževes station, the average annual temperature for the period 1996–2010 is 8.4 °C (Czech

Hydrometeorological Institute, 2023). Both stations are located in the Rakovník District.

3.3 Digital Elevation Model

The Digital Elevation Model (DEM) used in this study was derived from the Digital Relief Model of the Czech Republic – 5th Generation (DMR 5G), obtained from the Geoportal of the Czech Office for Surveying, Mapping and Cadastre (ČÚZK) (ČÚZK, 2023). DMR 5G represents a bare-earth Digital Terrain Model (DTM) generated from nationwide airborne laser scanning (LiDAR) data, with vegetation and buildings removed during processing.

The dataset provides a reported vertical accuracy of approximately ±0.18 m in open terrain and ±0.30 m in forested areas. The raster DEM used in this study has a spatial resolution of 2 m × 2 m and is referenced to the S–JTSK / Krovak East North coordinate system with elevations in the Baltic Vertical Datum (Bpv).

CHAPTER 4

Data

4.1 Original Dataset Description

The original Amalia dataset that contains the records of the Temperature-Moisture-Sensors (TMS) were provided by the Department of Water Resources and Environmental Modeling (KVHEM) at the Česká zemědělská univerzita v Praze (ČZU).

Air, surface, and soil temperatures are measured simultaneously by three independent sensors (Wild et al., 2019). The TMS sensors are distributed across the pilot farm, covering areas with varying landscape characteristics and sensor densities, including the northern forested area, the southern forested area, and the central agricultural land.

The original dataset consisted of approximately 20.35 million timestamped observations, 289 unique sensors, 329 metadata entries, 15-minute temporal resolution and three temperature measurements (air, surface, and soil). The spatial coordinates were stored in the S-JSTK / Krovak East North projection (epsg:5514).

4.2 Data Preparation and Structural Assessment

The preliminary checks revealed that there were exact duplicate rows, repeated (ID, DTM) timestamp pairs, conflicting values at identical timestamps and co-located sensors (multiple loggers installed at identical coordinates). To deal with these issues a systematic preprocessing prior to spatial modeling was required.

In fact, two types of duplicates were identified; exact duplicates that were removed using row-level uniqueness filtering and conflicting duplicates that were resolved using median aggregation. By using the median, we preserved temporal

continuity, avoided arbitrary selection and maintained one unique record per sensor and timestamp. No remaining duplicates (ID, DTM) pairs and 289 unique sensors were preserved afterwards.

Furthermore, seven groups were defined in the metadata. However, a spatial bounding-box analysis revealed that the monitoring networks *BP*, *podrobny*, and *podrobny_hluboke* occupy the same spatial domain and were therefore merged into a unified landscape, hereafter referred to as BP. The final landscape structure consisted of five landscapes: BP, KL, Drainage, Agriculture, and Agroforestry. This restructuring was necessary because variogram modeling assumes spatial homogeneity within a domain. In addition, the landscapes differ in topography and monitoring installation design; therefore, combining physically distinct areas would distort the spatial structure.

Besides, this decision was supported by the sensor installation depths, which revealed that the temperature measurements did not correspond to identical physical depths across landscapes. The depth ranges associated with each temperature measurement (T1, T2, and T3) vary depending on the landscape, as summarized in Table 4.1. T1 showed substantial variability between landscapes, ranging from shallow depths (-8 cm) to deeper soil layers (-118 cm). T2 ranged from subsurface depths to near-surface conditions depending on the landscape. T3 represents surface or near-surface temperature measurements, typically between 5 and 15 cm above or at the surface. Negative values represent depths below the soil surface, while positive values represent sensor height above the surface.

Table 4.1: Depth ranges of temperature sensors (T1, T2, T3) across different landscapes.

	T1 Depth (cm)		T2 Depth (cm)		T3 Height (cm)	
Landscape	Min	Max	Min	Max	Min	Max
BP	-118	-8	-108	2	5	15
KL	-23	-8	-13	2	15	15
Drainage	-60	-20	-60	-20	5	5
Agriculture	-8	-8	2	2	15	15
Agroforestry	-50	-8	-50	-2	5	15

Although, temperature measurements (T1, T2, and T3) are internally consistent within landscapes but not physically equivalent across landscapes. Therefore, variogram models must be landscape-specific, cross-landscape comparison should be interpreted it carefully and depth heterogeneity may partially explain differences in spatial structure and nugget variance.

4.3 Co-located Sensors Aggregation

The landscapes exhibited replicated sensor installation at identical spatial coordinates. Table 4.2 summarizes the occurrence of co-located sensors across landscapes. BP contained 21 co-located stations out of 83 total locations, whereas KL exhibited replicated sensors at all 10 spatial locations. In contrast, co-location was less frequent in Agroforestry and Drainage (6 out of 42 and 1 out of 43 locations, respectively), and no replicated sensors were observed in Agriculture. These replicated installations represent vertical sensor stacks and were aggregated prior to variogram estimation to prevent artificial inflation of the nugget effect due to coincident spatial observations.

Table 4.2: Occurrence of co-located sensor installations across landscapes.

Landscape	Total Locations	Co-located Locations	Max Replicates	Mean Replicates
BP	83	21	7	3.48
KL	10	10	3	2.10
Agroforestry	42	6	3	3.67
Drainage	43	1	3	3.00
Agriculture	35	0	1	–

Co-located sensor measurements were aggregated using the median, resulting in a dataset containing a single observation per spatial coordinate and timestamp. This step eliminated zero-distance pairs between replicated sensors and ensured that the dataset was suitable for variogram estimation and kriging while preserving the full temporal resolution of the observations. The median was preferred over the mean because of its robustness to outliers and sensor noise, providing a more stable representation of local thermal conditions. In geostatistical analysis, classical statistics such as the mean have robust counterparts, including the median, which are less sensitive to extreme values and therefore more suitable when data may contain outliers (Goovaerts, 1997).

4.4 Data Quality Assessment and Exploratory Analysis

Data completeness was high across all landscapes and temperature levels, with valid observations exceeding 91% in all cases (Table 4.3). Given the large number of observations per sensor, the small proportion of missing values is not expected to affect variogram estimation or kriging performance.

Table 4.3: Data completeness across landscapes and temperature measurement levels.

Temperature Level	Landscape	Total Observations	Valid Observations	Valid (%)
T1	BP	6116639	5754741	94.08
	KL	1188848	1169848	98.40
	Agroforestry	1706215	1706215	100.00
	Drainage	1428186	1345668	94.22
	Agriculture	2474397	2473630	99.97
T2	BP	6116639	5605846	91.65
	KL	1188848	1169848	98.40
	Agroforestry	1706215	1706215	100.00
	Drainage	1428186	1345668	94.22
	Agriculture	2474397	2474206	99.99
T3	BP	6116639	5754741	94.08
	KL	1188848	1169848	98.40
	Agroforestry	1706215	1706215	100.00
	Drainage	1428186	1343296	94.06
	Agriculture	2474397	2474206	99.99

An exploratory assessment of the temperature distribution was performed using boxplots and skewness coefficients in order to identify potential outliers and evaluate whether data transformation was required prior to semivariance estimation. Examination of skewness and box-and-whisker plots is recommended in geostatistical data analysis to detect asymmetry and extreme values that may influence variogram estimation (Kerry and Oliver, 2007; Oliver and Webster, 2014). The results are presented in Figures 4.1 and 4.2.

The boxplots indicate moderate variability across landscapes and temperature measurements. Although several extreme values were observed, these were considered consistent with expected environmental conditions (e.g., seasonal maxima and winter minima) rather than indicative of sensors malfunction.

The skewness coefficient describes the asymmetry of a distribution relative to its mean. Negative values indicate left-skewed distributions, whereas positive values indicate right-skewed distributions. Across landscapes and temperature measurements, skewness values remained close to zero, ranging from -0.73 to 0.51 (Figure 4.2). This indicates approximately symmetric to only moderately skewed distributions, suggesting that transformation of the temperature data was not necessary prior to variogram estimation. Data transformations can sometimes be applied to reduce skewness, but they may complicate the interpretation and back-transformation of kriging predictions. In addition, the dataset was analysed separately across different landscapes, which represent distinct physical environ-

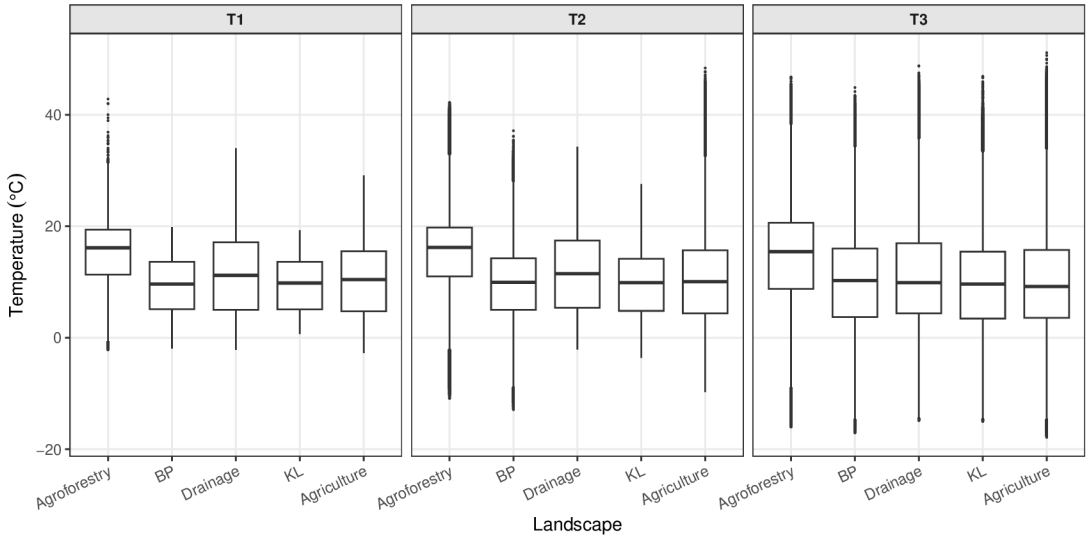

FIGURE 4.1: Boxplots of aggregated temperature observations by landscape and measurement level (T1–T3).

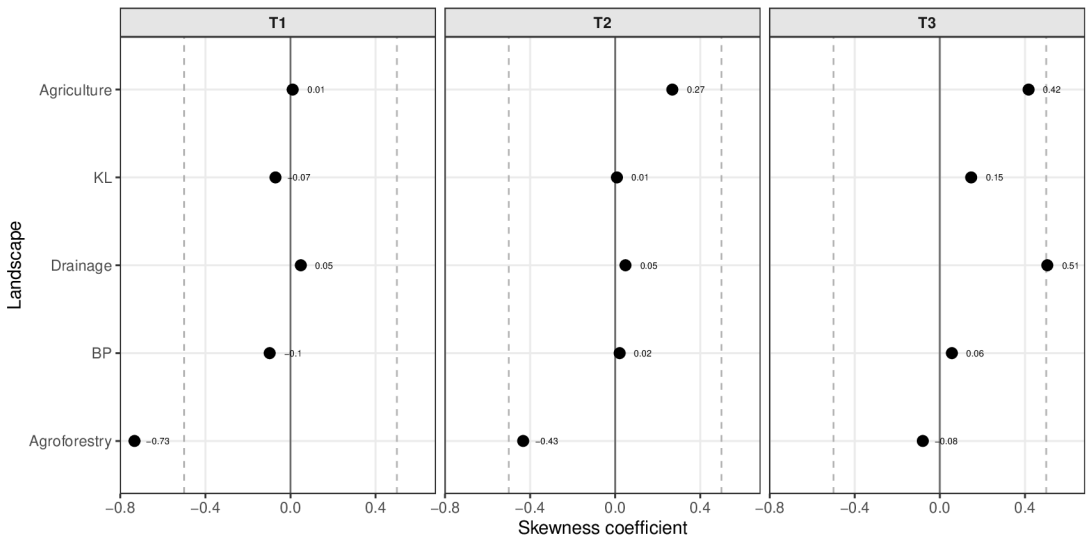

FIGURE 4.2: Skewness coefficients of aggregated temperature observations by landscape and measurement level (T1–T3). Dashed lines indicate reference thresholds for moderate asymmetry.

ments and monitoring conditions, thereby increasing the homogeneity of the data subsets used for variogram modeling (Goovaerts, 1997).

4.5 Final Dataset for Variogram Diagnosis and Modeling

The dataset prepared for spatial modeling is summarized in Tables 4.4 and 4.5. These tables describe the spatial configuration of the monitoring networks and the temporal coverage of the observations after exploratory data analysis and aggregation of co-located sensors.

Table 4.4: Spatial characteristics of the monitoring networks across landscapes.

Landscape	Unique Locations	Elevation Min (m)	Elevation Max (m)	Elevation Range (m)	Max Interstation Distance (km)
BP	83	372.86	445.50	72.64	1.63
KL	10	394.00	438.00	44.00	1.62
Agroforestry	42	422.82	430.96	8.14	0.13
Drainage	43	436.44	465.66	29.22	1.91
Agriculture	35	343.97	476.46	132.49	3.96

Table 4.4 summarizes the spatial configuration of the characteristics of the monitoring networks across landscapes. Noticeable differences in spatial extent were observed among the sites. The Agriculture landscape covers the largest domain (3.96 km), whereas KL, BP and Drainage represent relatively compact landscape units ($\approx$ 1.6–1.9 km).

The KL landscape contains only 10 effective spatial locations due to the presence of co-located sensors, resulting in a comparatively sparse spatial configuration. In contrast, the Agroforestry cluster contains 42 locations within a maximum inter-station distance of only 0.13 km, representing a micro-scale experimental field rather than a landscape-scale network. For this reason, Agroforestry was excluded from the kilometre-scale spatial interpolation analysis, although its sensors remain part of the overall monitoring system deployed within the study area.

Figures 4.3 and 4.4 illustrate the spatial distribution of temperature sensors across the Amálie pilot farm, overlaid on the DEM and an orthophoto, respectively, highlighting the spatial extent of each monitoring network.

Since geostatistical modeling operates on spatial support points, the number of unique spatial coordinates represents the effective spatial sample size for variogram estimation and kriging.

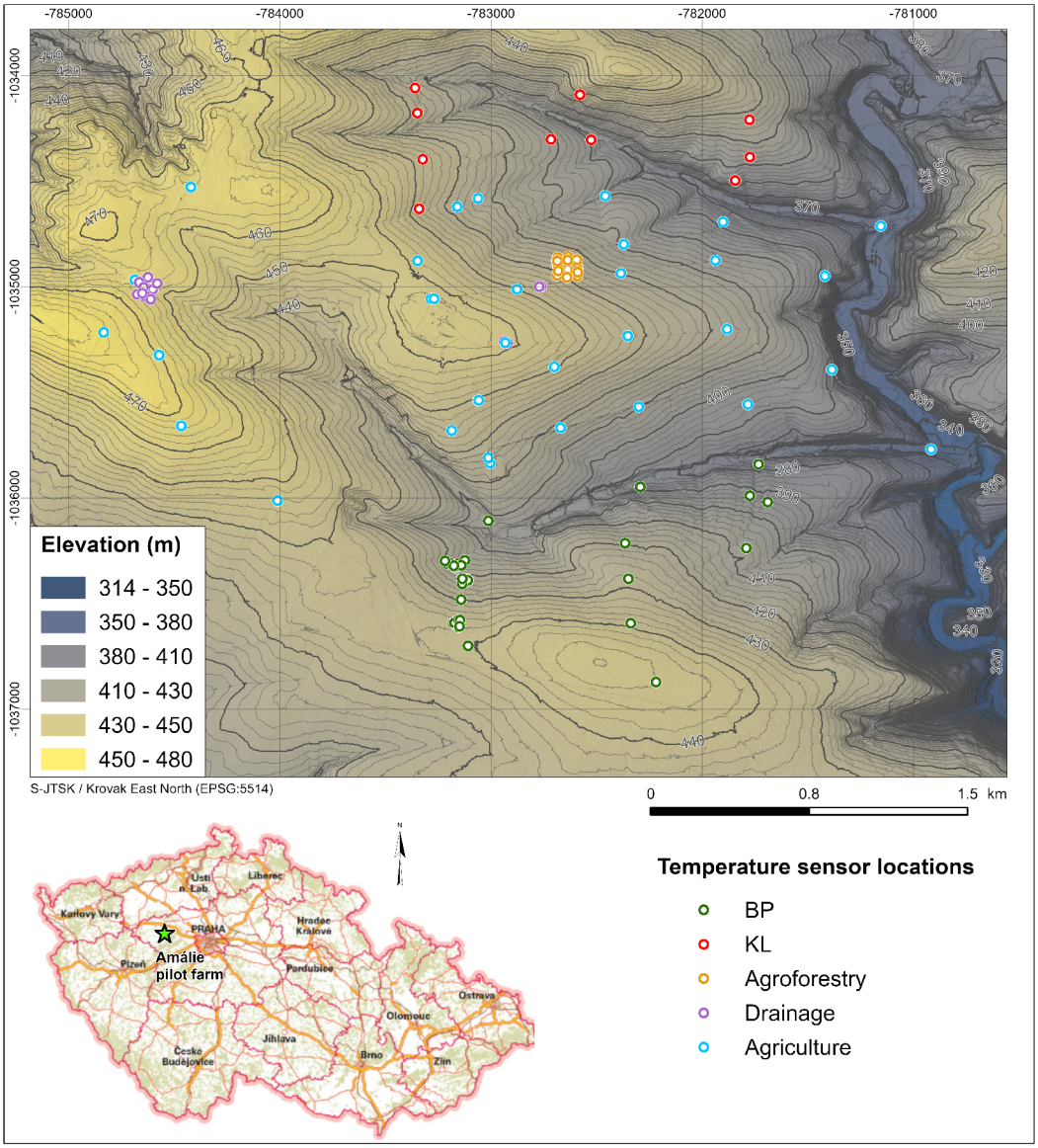

FIGURE 4.3: Digital Elevation Model (DEM) of the Amálie pilot farm illustrating topographic variability with contour lines and the distribution of temperature sensor locations. Sensors are grouped into five landscape units used in the spatial interpolation analysis. The inset map shows the location of the study area within the Czech Republic. Coordinates are displayed in the S-JTSK / Krovak East North projection (EPSG:5514).

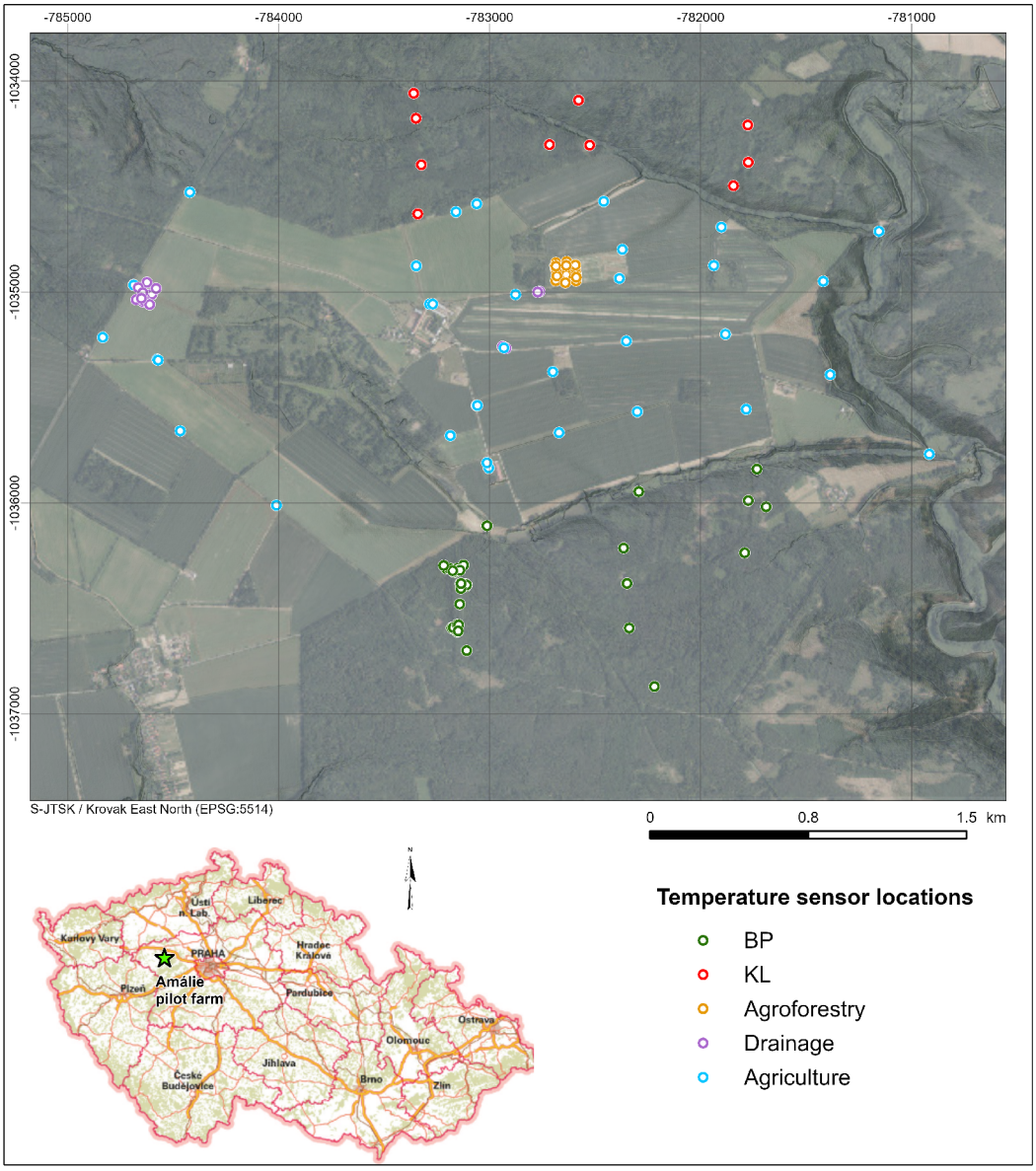

FIGURE 4.4: Spatial distribution of temperature sensors across the Amálie pilot farm shown over an orthophoto. Sensor grouping into landscape units is consistent with Fig. 4.3.

Table 4.5: Temporal coverage and completeness of temperature observations across landscapes and measurement levels (T1–T3).

Landscape	Variable	Locations	15-min Timestamps	Missing (%)
BP	T1	83	134213	5.92
BP	T2	83	134213	8.35
BP	T3	83	134213	5.92
KL	T1	10	119548	1.60
KL	T2	10	119548	1.60
KL	T3	10	119548	1.60
Agroforestry	T1	42	50026	0.00
Agroforestry	T2	42	50026	0.00
Agroforestry	T3	42	50026	0.00
Drainage	T1	43	42060	5.78
Drainage	T2	43	42060	5.78
Drainage	T3	43	42060	5.94
Agriculture	T1	35	80202	0.03
Agriculture	T2	35	80202	0.01
Agriculture	T3	35	80202	0.001

Table 4.5 summarizes the temporal coverage of the temperature observations for each landscape and temperature measurement (T1–T3). The monitoring periods vary slightly among landscapes, reflecting differences in installation dates and operational timelines of the sensor networks.

Most landscapes exhibit continuous observations with high temporal completeness, with missing data percentages generally below 6%. The longest time series are observed in BP, spanning from December 2020 to September 2024 with more than 134,000 timestamps at 15-minute resolution. In contrast, the other landscapes were installed later and therefore contain shorter time series.

The number of observations per spatial location remains relatively consistent within each landscape, indicating a balanced temporal sampling across monitoring sites.

CHAPTER 5

Methodology

5.1 Theoretical Background

Theoretical Framework of the Variogram

Spatially distributed variables such as temperature are strongly influenced by geographic and topographic factors, including elevation, latitude, longitude, and proximity to large water bodies (Hudson and Wackernagel, 1994; Holdaway, 1996).

Because these interactions and variations are complex, spatial variability is described using a stochastic process. A realization of this process, which in this case is temperature, can be represented as $z(x)$, where $Z(x)$ denotes the random variable describing the temperature at location x, as described by Oliver and Webster (2014).

We assume second-order stationarity, meaning that the statistical properties of the process do not depend on absolute location. Under this assumption, the process can be written as

$$Z(x) = \mu + \varepsilon(x),$$

This corresponds to the special case of the general decomposition $Z(x) = m(x) + Y(x)$ introduced later in the kriging framework, where the mean is constant ($m(x) = \mu$), and the residual process is $Y(x) \equiv \varepsilon(x)$. The residual function $\varepsilon(x)$ is assumed to have zero mean and covariance

$$C(h) = \mathbb{E}\big[\varepsilon(x)\,\varepsilon(x+h)\big],$$

which is equivalent to

$$C(h) = \mathbb{E}\big[(Z(x)-\mu)(Z(x+h)-\mu)\big] = \mathbb{E}\big[Z(x)Z(x+h)\big] - \mu^2,$$

where h is the separation vector between locations.

Following Matheron (1962), spatial dependence can be expressed using the semivariance:

$$\gamma(h) = \frac{1}{2}\,\mathrm{Var}[Z(x) - Z(x+h)] = \frac{1}{2}\,\mathbb{E}\big[(Z(x) - Z(x+h))^2\big]\,. \tag{5.1}$$

This expression defines the *theoretical semivariogram* (commonly referred to as the variogram).

The variogram characterizes how spatial variability changes with distance by quantifying the variance of differences between observations as a function of their separation. It therefore provides a concise representation of the spatial structure of the variable under study.

To link the variogram with the covariance function, we start from the definition given in Equation 5.1.

Expanding the square and taking expectations yields

$$\gamma(h) = \frac{1}{2}\left[\mathbb{E}\big[Z(x)^2\big] + \mathbb{E}\big[Z(x+h)^2\big] - 2\,\mathbb{E}\big[Z(x)Z(x+h)\big]\right].$$

Under second-order stationarity, the expected value and variance are constant,

$$\mathbb{E}[Z(x)] = \mu, \qquad \mathrm{Var}(Z(x)) = \sigma^2,$$

and the covariance between two locations depends only on their separation vector h,

$$C(h) = \mathrm{Cov}\big(Z(x), Z(x+h)\big).$$

It follows that

$$\mathbb{E}[Z(x)^2] = \sigma^2 + \mu^2.$$

Substituting these gives

$$\gamma(h) = \frac{1}{2}\left[2(\sigma^2 + \mu^2) - 2\big(C(h) + \mu^2\big)\right] = \sigma^2 - C(h).$$

Since $C(0) = \sigma^2$,

$$\gamma(h) = C(0) - C(h).$$

To account for microscale variability or measurement error, a nugget variance c_0 is included:

$$\gamma(h) = c_0 + \big[C(0) - C(h)\big].$$

Expressing the covariance function in terms of the correlation function $\rho(h)$,

$$C(h) = c\,\rho(h),$$

where c_0 represents the nugget variance and $c = C(0)$ is the partial sill. Substituting this expression into the previous equation yields

$$\gamma(h) = c_0 + c\big[1 - \rho(h)\big].$$

Letting $f(h) = 1 - \rho(h)$, the general form of a theoretical variogram model becomes

$$\gamma(h) = c_0 + c\,f(h). \tag{5.2}$$

In this study, temperature observations are available at repeated timestamps.

Let $Z_t(x)$ denote the spatial random field observed at time t.

While the temperature realization varies over time, we assume the spatial dependence structure remains approximately constant.

This assumption implies that the spatial covariance structure is invariant over time and can be written as

$$C_t(h) = C_s(h). \tag{5.3}$$

where $C_t(h)$ is the spatial covariance at time t and $C_s(h)$ denotes the common spatial covariance function.

Under this assumption, spatial variograms estimated at different timestamps can be combined in order to obtain a stable estimate of the spatial covariance structure used in the kriging framework described later.

Experimental Variogram

Classical Estimator (Matheron)

The expression in the Equation 5.1 defines the *theoretical semivariogram.* In practice, however, the variogram must be estimated from observations $z(x_i)$. The *experimental variogram* (or sample variogram) is computed using Matheron's method of moments:

$$\hat{\gamma}_{\mathrm{M}}(h) = \frac{1}{2\,m(h)} \sum_{j=1}^{m(h)} \left[z(x_j) - z(x_j + h)\right]^2, \tag{5.4}$$

where $m(h)$ is the number of data pairs separated by lag h. The experimental variogram provides semivariance estimates at discrete lag distances.

Robust Estimator (Cressie–Hawkins)

In addition to the classical method-of-moments estimator proposed by Matheron (1962), a robust alternative was introduced by Cressie and Hawkins (1980) to reduce the influence of outliers from the underlying processes (Webster and Oliver, 2007). The Cressie–Hawkins estimator of the semivariance is defined as

$$\hat{\gamma}_{\mathrm{CH}}(h) = \frac{\left(\frac{1}{m(h)} \sum_{j=1}^{m(h)} |z(x_j) - z(x_j + h)|^{1/2}\right)^4}{0.457 + \dfrac{0.494}{m(h)} + \dfrac{0.045}{m(h)^2}}, \tag{5.5}$$

where $m(h)$ is the number of data pairs separated by lag h.

The square root transformation reduces the influence of extreme values, while the denominator provides a bias correction for small sample sizes.

Although formulated differently, the Cressie–Hawkins estimator aims to approximate the same theoretical semivariogram defined in Equation 5.1, but with increased robustness to outliers and heavy-tailed distributions.

Theoretical Variogram Models

Based on the semivariogram definition in Equation 5.1 and the general variogram model given in Equation 5.2, four commonly used theoretical variogram models were considered in this thesis: spherical, exponential, Gaussian, and Matérn. In all cases, c_0 represents the nugget variance, c the partial sill, and a the range parameter.

Spherical Model

The spherical variogram model is defined as

$$\gamma(h) = \begin{cases} c_0 + c\left[\dfrac{3h}{2a} - \dfrac{1}{2}\left(\dfrac{h}{a}\right)^3\right], & 0 < h \leq a, \\ c_0 + c, & h > a. \end{cases} \tag{5.6}$$

This model reaches the sill exactly at the finite distance a, and is therefore considered a bounded model (Goovaerts, 1997).

Exponential Model

The exponential variogram model is given by

$$\gamma(h) = c_0 + c\left[1 - \exp\left(-3\frac{h}{a}\right)\right]. \tag{5.7}$$

The sill is approached asymptotically as h increases, meaning the model has an infinite theoretical range but a practical range approximately equal to a (Webster and Oliver, 2007).

Gaussian Model

The Gaussian variogram model is expressed as

$$\gamma(h) = c_0 + c\left[1 - \exp\left(-3\left(\frac{h}{a}\right)^2\right)\right]. \tag{5.8}$$

This model represents a very smooth spatial process, with correlation decreasing rapidly near the origin and approaching the sill asymptotically (Webster and Oliver, 2007).

Matérn Model

The Matérn variogram model provides additional flexibility through a smoothness parameter ν, which controls the smoothness of the spatial field (Webster and Oliver, 2007).

$$\gamma(h) = c_0 + c\left[1 - \rho(h)\right], \tag{5.9}$$

where the Matérn correlation function is

$$\rho(h) = \frac{1}{2^{\nu-1}\Gamma(\nu)}\left(\frac{2\sqrt{\nu}h}{a}\right)^{\nu} K_\nu\left(\frac{2\sqrt{\nu}h}{a}\right), \tag{5.10}$$

$\Gamma(\nu)$ is the gamma function, and $K_\nu(\cdot)$ is the modified Bessel function of the second kind. Smaller values of ν correspond to rougher spatial fields, while larger values produce smoother fields. The exponential and Gaussian models can be interpreted as special cases of the Matérn family (Webster and Oliver, 2007).

Spatial Interpolation Methods

An isometric map is often referred to as a map of interpolated temperature, where, in our case, the value of such a variable is unknown in some locations and can be predicted through a function that allows the representation of a continuous surface (Lam, 1983).

Moreover, spatial interpolation methods can be broadly classified into exact and approximate approaches. In this thesis, only exact methods were evaluated. A classic overview of this distinction is provided by Lam (1983) in “Spatial interpolation methods: A review”, where interpolation techniques are grouped accordingly.

A more comprehensive comparison of interpolation methods, distinguishing between non-geostatistical, geostatistical, and hybrid approaches, as well as between univariate and multivariate, and deterministic and stochastic methods, was later presented by Li and Heap (2014).

In the present study, one of these approaches was tested, belonging to the geostatistical family and specifically to the class of least-squares methods of spatial interpolation, commonly known as kriging.

The choice of an interpolation model depends on the characteristics of the data, the desired prediction accuracy, and the computational effort involved, as the modeler must formulate a hypothesis about the underlying spatial surface (Goovaerts, 1997; Webster and Oliver, 2007).

To predict the value of a spatial random variable $Z(x)$ at an unsampled location x_0, observations $z(x_1), z(x_2), \ldots, z(x_N)$ at sampled locations $x_1, x_2, \ldots, x_N$ are used (Webster and Oliver, 2007). Many spatial interpolation methods can be expressed within a common linear estimation framework:

$$\hat{Z}(x_0) = \sum_{i=1}^{N} \lambda_i(x_0)\, z(x_i), \tag{5.11}$$

where $\hat{Z}(x_0)$ is the predicted value at location x_0, $z(x_i)$ are observed values at neighboring sampled locations, $\lambda_i(x_0)$ are weights assigned to each observation, and N is the number of neighboring points used in the estimation.

The difference between interpolation methods lies primarily in how the weights $\lambda_i(x_0)$ are determined.

Kriging

Kriging is well known by geostaticians for a family of generalized least-squares interpolation algorithms (Goovaerts, 1997). The method provides the best linear unbiased prediction (BLUP) of a spatial random variable by accounting for the spatial dependence structure described by the variogram.

The value of the spatial random variable $Z(x)$ at an unsampled location x_0 is estimated as a linear combination of neighboring observations (Goovaerts, 1997; Webster and Oliver, 2007). In its most general form, the kriging estimator is written as

$$\hat{Z}(x_0) - m(x_0) = \sum_{i=1}^{n} \lambda_i(x_0) \left[Z(x_i) - m(x_i) \right], \tag{5.12}$$

where $\lambda_i(x_0)$ are weights assigned to the observed values $z(x_i)$, which are realizations of the random variables $Z(x_i)$, and $m(x_0)$ and $m(x_i)$ are the ex-

pected values (mean function) of $Z(x)$ at the prediction and observation locations, respectively.

The estimation error at location x_0 is defined as

$$\hat{Z}(x_0) - Z(x_0),$$

and kriging aims to minimize the variance of this error,

$$\sigma_E^2(x_0) = \mathrm{Var}\big[\hat{Z}(x_0) - Z(x_0)\big], \tag{5.13}$$

subject to the constraint of unbiasedness,

$$\mathbb{E}\Big[\hat{Z}(x_0) - Z(x_0)\Big] = 0.$$

The form of the kriging estimator depends on the model adopted for the random function. In general, the random field can be decomposed as

$$Z(x) = m(x) + Y(x), \tag{5.14}$$

where $m(x)$ represents the deterministic trend component (mean function) and $Y(x)$ is the residual random function.

The residual component $Y(x)$ is assumed to be second-order stationary with zero mean,

$$\mathbb{E}[Y(x)] = 0,$$

and covariance function

$$\mathrm{Cov}\big(Y(x), Y(x+h)\big) = C(h),$$

where $C(h)$ depends only on the separation vector h.

In the special case where the mean is constant, $m(x) = \mu$, the decomposition reduces to

$$Z(x) = \mu + Y(x),$$

which corresponds to the second-order stationary model introduced in Section 5.1.

Under this formulation, the expected value of the random variable equals the trend component,

$$\mathbb{E}[Z(x)] = m(x).$$

The way in which the trend component $m(x)$ is modeled determines the type of kriging algorithm applied.

Simple kriging (SK) was not considered appropriate because it assumes that the global mean of the process is known. Temperature fields exhibit strong temporal variability, meaning that the mean temperature changes across timestamps. Using the sample mean at each timestamp would therefore violate the assumption of a known mean required by SK.

Ordinary Kriging (OK) OK assumes that the mean of the process is constant but unknown within a local neighborhood $N(x_0)$ around the prediction location x_0 (Goovaerts, 1997; Webster and Oliver, 2007),

$$m(x) = \mu, \qquad \forall\, x \in N(x_0),$$

where μ is an unknown constant specific to the local neighborhood. This approach relaxes the assumption of a globally known mean, as in SK, while still requiring local stationarity of the residual process.

The OK estimator at an unsampled location x_0 is expressed as a linear combination of neighboring observations,

$$\hat{Z}(x_0) = \sum_{i=1}^{n} \lambda_i(x_0)\, z(x_i), \tag{5.15}$$

where $\lambda_i(x_0)$ are weights assigned to the observed values $z(x_i)$.

Because the mean is unknown, the estimator must satisfy the unbiasedness condition,

$$\mathbb{E}\Big[\hat{Z}(x_0) - Z(x_0)\Big] = 0,$$

which, under the assumption of a constant local mean, leads to the constraint

$$\sum_{i=1}^{n} \lambda_i(x_0) = 1. \tag{5.16}$$

The kriging weights are obtained by minimizing the estimation variance subject to this constraint. Using the variogram model described in Section 5.1 to represent spatial dependence, the OK system in scalar form becomes

$$\sum_{j=1}^{n} \lambda_j(x_0)\, \gamma(x_i - x_j) + \psi = \gamma(x_i - x_0), \qquad i = 1, \ldots, n, \tag{5.17}$$

together with the constraint given in Equation 5.16, where ψ is a Lagrange multiplier introduced to enforce the unbiasedness condition.

In matrix form, the OK system can be written as

$$\begin{bmatrix} \boldsymbol{\Gamma} & \mathbf{1} \\ \mathbf{1}^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}(x_0) \\ \psi \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma}(x_0) \\ 1 \end{bmatrix}, \tag{5.18}$$

where

$$\boldsymbol{\Gamma} = \begin{bmatrix} \gamma(x_1 - x_1) & \cdots & \gamma(x_1 - x_n) \\ \vdots & \ddots & \vdots \\ \gamma(x_n - x_1) & \cdots & \gamma(x_n - x_n) \end{bmatrix}, \quad \boldsymbol{\lambda}(x_0) = \begin{bmatrix} \lambda_1(x_0) \\ \vdots \\ \lambda_n(x_0) \end{bmatrix},$$

$$\boldsymbol{\gamma}(x_0) = \begin{bmatrix} \gamma(x_1 - x_0) \\ \vdots \\ \gamma(x_n - x_0) \end{bmatrix}, \quad \mathbf{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Solving this system yields the kriging weights $\lambda_i(x_0)$ and the Lagrange multiplier ψ, which are then used in Equation 5.15 to compute the prediction.

Kriging with External Drift (KED) KED extends OK by allowing the mean of the process to vary in space as a function of auxiliary variables (Goovaerts, 1997; Webster and Oliver, 2007). The random field is decomposed as

$$Z(x) = m(x) + Y(x),$$

where $m(x)$ is a deterministic trend (drift) function and $Y(x)$ is a zero-mean, second-order stationary residual process with covariance function $C(h)$.

The drift is modeled as a linear combination of p auxiliary variables,

$$m(x) = \beta_0 + \sum_{k=1}^{p} \beta_k \, q_k(x), \tag{5.19}$$

where $q_k(x)$ are known auxiliary variables at location x (e.g., spatial coordinates, elevation, slope, or aspect), and β_k are unknown regression coefficients.

The KED estimator retains the linear form

$$\hat{Z}(x_0) = \sum_{i=1}^{n} \lambda_i(x_0) \, z(x_i), \tag{5.20}$$

but the weights $\lambda_i(x_0)$ are computed so that the estimator is unbiased with respect to the spatially varying mean in Equation 5.19.

Unbiasedness requires that the estimator reproduces the drift, leading to the constraints

$$\sum_{i=1}^{n} \lambda_i(x_0) \, q_k(x_i) = q_k(x_0), \qquad k = 0, 1, \ldots, p, \tag{5.21}$$

where $q_0(x) = 1$ ensures reproduction of the intercept term.

Minimizing the estimation variance subject to the constraints in Equation 5.21 leads to the KED system

$$\sum_{j=1}^{n} \lambda_j(x_0)\,\gamma(x_i - x_j) + \sum_{k=0}^{p} \psi_k\, q_k(x_i) = \gamma(x_i - x_0), \quad i = 1, \ldots, n, \tag{5.22}$$

together with the $p + 1$ constraint equations

$$\sum_{i=1}^{n} \lambda_i(x_0)\, q_k(x_i) = q_k(x_0), \quad k = 0, \ldots, p,$$

where ψ_k are Lagrange multipliers.

In matrix form, the KED system becomes

$$\begin{bmatrix} \boldsymbol{\Gamma} & \mathbf{Q} \\ \mathbf{Q}^\top & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}(x_0) \\ \boldsymbol{\psi} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma}(x_0) \\ \mathbf{q}(x_0) \end{bmatrix}, \tag{5.23}$$

where $\mathbf{Q}$ is the matrix of auxiliary variables at observation locations,

$$\mathbf{Q} = \begin{bmatrix} 1 & q_1(x_1) & \ldots & q_p(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & q_1(x_n) & \ldots & q_p(x_n) \end{bmatrix},$$

$\mathbf{q}(x_0)$ is the vector of auxiliary variables at the prediction location, and $\boldsymbol{\psi}$ is the vector of Lagrange multipliers.

When only the intercept term is included ($p = 0$), KED reduces to OK.

The incorporation of auxiliary environmental predictors in kriging has been shown to improve prediction accuracy in many environmental applications (Hengl et al., 2004). For example, Hudson and Wackernagel (1994) demonstrated that including elevation as an auxiliary variable significantly improves temperature interpolation. Similarly, Willmott and Matsuura (1995) exploited the strong relationship between air temperature and elevation to improve interpolation of annually averaged air temperature across the United States. These studies highlight the importance of topographic predictors in temperature modeling, which motivates the use of elevation and other DEM-derived variables as auxiliary predictors in the present analysis.

5.2 Spatial Regime Diagnosis and Modeling Framework

Variogram Diagnosis Phase

The variogram diagnosis phase aimed to identify the spatial covariance structure for each landscape unit prior to constructing regime variograms and kriging models. It was assumed that the functional form of the spatial covariance structure remains invariant over time, while its parameters (nugget, sill, and range) may vary across timestamps and seasons.

Under this assumption, the spatial covariance structure at a given time step can be expressed as

$$C(h, 0) = C_s(h), \tag{5.24}$$

where h denotes the spatial lag distance, 0 represents zero temporal lag, and $C_s(h)$ is the spatial covariance function describing the dependence between observations separated by distance h.

During the diagnosis phase, empirical variograms were computed using the classical method-of-moments estimator proposed by Matheron (Equation 5.4).

For variogram modeling, the exponential model defined in Equation 5.7 was selected to represent the spatial dependence structure.

The primary objective of the diagnosis phase was to enable consistent comparison of variogram parameters across timestamps. Introducing multiple estimators or several variogram model families would likely have increased parameter variability and made the resulting parameter time series more difficult to interpret.

Prior to the main diagnosis phase, a pilot analysis was conducted to evaluate suitable configuration settings and to refine the methodological strategy applied during the final variogram diagnosis stage.

This strategy ensured that temporal changes in spatial dependence could be attributed to variations in environmental conditions rather than inconsistencies in variogram estimation.

Pilot Variogram Analysis

A pilot variogram analysis was conducted to diagnose the spatial covariance structure of each landscape and to determine appropriate lag configurations, binning strategies, and fitting thresholds for empirical variogram estimation. This preliminary step was necessary to ensure stable variogram estimation and to enable consistent comparison of variogram parameters across timestamps, while also supporting the development of a scalable pipeline for regime modeling.

The geometric design of lag bins required careful consideration because each landscape exhibits different spatial extents and maximum interstation distances, which directly influence the spatial covariance structure. Choosing appropriate lag classes is important because excessively narrow bins may produce noisy experimental variograms, while overly broad bins can smooth the variogram and obscure important spatial structure (Lark, 2000).

Configuration parameters including lag width, maximum lag distance, minimum number of pairs per bin, and the minimum number of bins required for variogram fitting were therefore tuned during this pilot phase to obtain reliable empirical variograms across landscapes.

For the Agriculture and BP landscapes, stable variogram estimates were obtained using manually tuned fixed-width lag bins. In contrast, the Drainage and KL landscapes presented greater challenges due to their spatial configuration, requiring the use of geometry-aware lag bins constructed from pairwise distance quantiles. In this approach, lag edges were fixed across timestamps based on the empirical distribution of interstation distances. Unlike equal-width bins, this quantile-based produces lag classes containing a similar number of observation pairs, ensuring sufficient statistical support within each bin and resulting in more stable empirical variograms. Because the lag boundaries remain fixed for each landscape, variograms estimated at different timestamps remain directly comparable. Two strategies were therefore explored for constructing empirical variograms.

Variogram model parameters were estimated using weighted least squares (WLS), implemented through nonlinear optimization because theoretical variogram models are nonlinear functions of their parameters (Cressie, 1985; Jian et al., 1996; Oliver and Webster, 2014). The weighting scheme assigns larger influence to lag bins supported by a greater number of observation pairs:

$$w(h) = \sqrt{n_{\text{pairs}}(h)} \tag{5.25}$$

where $n_{\text{pairs}}(h)$ denotes the number of observation pairs within lag h. Weighting by the number of observation pairs gives greater influence to more reliable lag estimates while downweighting bins supported by fewer pairs, which helps stabilize variogram fitting.

Landscape-specific fitting thresholds determined during the pilot analysis are summarized in Table 5.1. These thresholds were selected empirically in order to maximize model convergence rates during the automated variogram fitting procedure.

Table 5.1: Variogram fitting configuration thresholds determined during the pilot variogram analysis.

Landscape	Binning Strategy	Minimum Bins	Minimum Pairs per Bin
Agriculture	fixed-width	6	15
BP	fixed-width	5	20
KL	fixed-edges	4	5
Drainage	fixed-edges	4	20

Main Diagnosis Phase

The diagnosis phase aimed to analyze the four landscapes by sampling 30 stratified days across seasons at 2-hour temporal resolution (12 timestamps per day). For each timestamp, the empirical semivariogram was estimated, the exponential model was fitted, and the corresponding model parameters were stored when fitting procedure successfully converged.

To ensure that variogram estimation was performed only on timestamps with sufficient spatial support, a diagnostic manifest table was constructed. This step prevented the inclusion of timestamps with insufficient spatial coverage, which could otherwise lead to unstable or meaningless variogram estimates.

By creating a reproducible subset of timestamps for the diagnosis phase, we ensured that the sampling was balanced across seasons, that each selected day contained the complete 2-hour temporal grid, and that each selected timestamp exhibited adequate spatial coverage. The diagnostic manifest was constructed according to the following procedure:

1. A coverage table was built for the 2-hour temporal grid for each temperature measurement (T1–T3). For each landscape and timestamp, observations with finite values were retained and the number of unique spatial locations (N_s) was computed.

2. Minimum spatial coverage thresholds were defined by calculating the 25th percentile of N_s across all 2-hour timestamps. Timestamps below this threshold were removed, thereby excluding the weakest 25% of spatial coverage cases.

3. Eligible timestamps were identified by intersecting the valid timestamps across the three temperature measurements (T1–T3).

4. Eligible days were then determined by retaining only those days that contained the complete set of 12 timestamps within the 2-hour grid.

5. Seasonal stratification was performed by randomly sampling 30 eligible days using a fixed random seed to ensure reproducibility.

6. Finally, the selected days were expanded back to their corresponding timestamps to construct the final diagnostic manifest.

Following this procedure, the diagnostic manifest remains scalable and can easily be extended to include a larger number of sampled days without requiring modifications to the analysis script.

Parameter Stability Analysis

Parameter stability diagnostics were conducted to evaluate the robustness of variogram estimation during the diagnosis phase. For each timestamp and landscape, empirical semivariograms were computed and fitted using the selected model family. Diagnostic metrics were recorded, including the proportion of failed model fits, the number of lag bins used in the empirical variogram, and the spatial support (number of observations N).

These diagnostics provide an indication of whether the spatial sampling geometry is sufficient for reliable variogram estimation and help identify landscapes where spatial structure may be weak or poorly constrained.

Regime Definition and Modeling

The variogram parameters obtained during the diagnosis phase provide the basis for identifying potential temporal regimes in the spatial covariance structure. Regime definition relies on examining how these parameters evolve across timestamps and determining whether systematic temporal patterns exist.

The definition of temporal regimes requires both statistical and physical interpretation of variogram parameters. In particular, regime stratification must address the following questions:

- Does the spatial structure vary systematically with hour of the day?
- Do different months exhibit similar diurnal patterns?
- Are observed parameter changes large relative to background variability?
- Is there a consistent separation between nighttime, midday, and transition periods?
- Is this separation seasonally dependent?

To address these questions, a systematic diagnostic plotting scheme was developed. These plots provide a structured basis for interpreting parameter stability and for determining whether regime stratification is justified for each landscape and temperature measurement (T1–T3). Regime stratification is adopted only when patterns are systematic, physically interpretable, and sufficiently stable across timestamps. Otherwise, stationary variogram models are retained in order to preserve parameter robustness.

The nugget represents micro-scale variability and measurement noise. Higher nugget values indicate increased small-scale heterogeneity, turbulence effects, or variability associated with sensor-scale processes, whereas a low nugget values suggest a smoother spatial field at short distances.

The partial sill represents the structured component of spatial variance, while the total sill (nugget plus partial sill) corresponds to the total variance of the spatial field. Higher sill values indicate stronger spatial contrasts and greater variability, whereas lower sill values imply a more spatially homogeneous temperature field.

The range parameter represents the distance at which spatial correlation becomes negligible. Beyond this distance, observations can be considered effectively uncorrelated. Identifying such regimes allows variogram parameters to be estimated within more homogeneous temporal conditions, thereby improving the stability of kriging predictions.

Bootstrap Determination of Regime Timestamp Pool Size

After defining regime-specific subsets, variogram parameters were estimated using a bootstrap resampling strategy. For each landscape–regime combination, timestamps belonging to the same regime were pooled to increase the number of spatial observation pairs available for variogram estimation. Reliable variogram estimation requires a sufficient number of spatial observation pairs at each lag distance (Goovaerts, 1997).

Unlike the diagnosis phase, where semivariograms were fitted independently for each timestamp, regime-specific estimation relies on pooling information across multiple timestamps belonging to the same regime. This pooling increases the effective spatial sample size and improves the stability of variogram fitting. Pooling observations to increase the number of available measurements has been shown to improve the reliability of statistical estimates when individual samples are limited (Bilonick and Nichols, 1983).

Bootstrap resampling was therefore used to determine the minimum number of timestamps required to obtain stable variogram parameter estimates. Candidate pool sizes were evaluated by repeatedly sampling timestamps within each regime

subset and estimating variogram parameters across bootstrap realizations.

Because the target application is kriging of 15-minute temperature fields, fixed lag bins (with identical lag width and origin) were defined across all timestamps within each regime. This ensures that spatial support remains consistent across bootstrap samples and allows variogram parameters estimated from different resamples to remain directly comparable.

Bootstrap Stability Procedure

Bootstrap resampling was used to evaluate the stability of variogram parameters under different timestamp pool sizes. Reliable estimation of variograms requires a sufficiently large number of observations and spatial pairs to obtain stable estimates (Lark, 2000). The bootstrap procedure therefore provides a practical way to evaluate how the stability of fitted variogram parameters changes as the number of timestamps included in the regime pool increases.

To determine the minimum number of timestamps required for stable regime-level variogram estimation, a bootstrap resampling strategy was implemented. For each landscape-regime combination, a sequence of candidate sample sizes

$$n = \{50, 100, 200, 400, 800\}$$

timestamps was evaluated. For each candidate sample size n, the following procedure was applied:

1. n timestamps were randomly sampled from the corresponding regime.

2. For each timestamp t, an empirical semivariogram $\hat{\gamma}_t(h_k)$ was computed using the predefined fixed lag bins.

3. Empirical semivariograms were pooled across timestamps using a pair-count weighted average:

$$\bar{\gamma}(h_k) = \frac{\sum_t m_t(h_k)\,\hat{\gamma}_t(h_k)}{\sum_t m_t(h_k)} \tag{5.26}$$

   where $m_t(h_k)$ denotes the number of observation pairs contributing to lag h_k at timestamp t.

   This weighting ensures that timestamps with greater spatial support contribute proportionally to the pooled empirical semivariogram.

   Importantly, empirical semivariograms were averaged prior to model fitting rather than averaging fitted parameters. This approach avoids bias introduced by nonlinear model estimation and preserves the empirical covariance structure before parametric modeling.

4. The pooled empirical semivariogram was then fitted with an exponential model using weighted nonlinear least squares in order to evaluate parameter stability under controlled sampling conditions.

5. Steps 1–4 were repeated $R = 20$ times for each candidate sample size n.

Two semivariogram estimators were considered for empirical computation:

- the classical Matheron estimator
- the robust Cressie–Hawkins estimator

This resampling framework allows the relationship between sample size and variogram parameter stability to be evaluated explicitly.

Stability Criterion

For each variogram parameter (nugget, partial sill, and range), the mean and standard deviation across bootstrap replicates were computed. Parameter stability was quantified using the relative standard deviation (RSD):

$$RSD_\theta = \frac{SD(\hat{\theta}_n)}{Mean(\hat{\theta}_n)} \tag{5.27}$$

where $\hat{\theta}_n$ denotes the variogram parameter estimate obtained using n sampled timestamps, and $\theta \in \text{nugget}, \text{psill}, \text{range}$.

The partial sill was adopted as the primary stability criterion because it represents the spatial component of spatial variance and directly governs kriging uncertainty. A stabilization threshold of

$$RSD_{psill} \leq 0.20$$

was adopted to indicate adequate parameter stability. This criterion ensures that the relative variability of the partial sill does not exceed 20% of its mean value across bootstrap replicates.

For each landscape, the smallest sample size n satisfying this stability criterion, together with acceptable convergence rates across bootstrap replicates, was selected as the final timestamp pool size for regime-level variogram estimation.

Both the classical Matheron estimator and the robust Cressie–Hawkins estimator were evaluated during the bootstrap stability analysis. In practice, the Matheron estimator produced more consistent convergence across landscapes and sample sizes, whereas the Cressie–Hawkins estimator occasionally exhibited unstable behavior under limited spatial support.

Regime Variogram Estimation

Following the bootstrap-based determination of regime timestamp pool sizes, regime-specific variograms were estimated separately for each landscape–regime combination.

For each regime subset, empirical semivariograms were computed by pooling the selected timestamps using the fixed lag structure defined in the previous sections. Both the classical Matheron estimator and the robust Cressie–Hawkins estimator were evaluated in order to assess the robustness of empirical variogram estimation.

The pooled empirical semivariograms were subsequently fitted using four theoretical variogram model families:

- Exponential
- Gaussian
- Spherical
- Matérn

These model families represent different types of spatial covariance structures characterized by varying smoothness and range behavior, thereby providing flexibility in capturing regime-specific spatial dependence. Their mathematical formulations were presented previously in section 5.1.

Candidate variogram models were compared by minimizing the root mean square error (RMSE) between the theoretical variogram model and the pooled empirical semivariogram:

$$RMSE = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(\hat{\gamma}(h_k) - \gamma(h_k;\theta)\right)^2} \tag{5.28}$$

where $\hat{\gamma}(h_k)$ denotes the empirical semivariogram at lag h_k, $\gamma(h_k;\theta)$ the fitted theoretical model, and K the number of lag bins.

As a secondary diagnostic, the Akaike Information Criterion (AIC) was used to compare fitted variogram models. The AIC is defined as

$$AIC = 2p - 2\ln(L) \tag{5.29}$$

where p denotes the number of estimated model parameters and L is the likelihood of the fitted model (Akaike, 1974). The first term ($2p$) penalizes model complexity, whereas the second term ($-2\ln(L)$) reflects the goodness of fit of the model. Lower AIC values therefore indicate models that achieve a better balance

between goodness of fit and model parsimony. The AIC has also been used for variogram model comparison in geostatistical studies (Jian et al., 1996).

In this study, AIC was computed using the built-in `AIC()` function in `R`.

The resulting candidate variogram models were subsequently evaluated through kriging cross-validation in order to identify the most suitable model configuration for each landscape–regime subset.

Kriging Implementation

Temperature measurements are available at repeated timestamps t, forming a sequence of spatial fields $Z_t(x)$. Although the temperature realization varies through time due to diurnal and seasonal processes, the spatial covariance structure is assumed to remain approximately invariant.

Formally, under the assumption stated in Equation 5.3, empirical semivariograms computed at individual timestamps can be pooled to estimate a stable spatial variogram. Pooling across timestamps increases the number of observation pairs contributing to each lag class, thereby improving the stability of the variogram estimation.

Kriging Models

Two kriging variants were evaluated for spatial interpolation of the 15-minute temperature fields: OK and KED.

OK assumes an unknown but locally constant mean, which is more appropriate for spatial interpolation of temperature fields at individual timestamps (Goovaerts, 1997; Webster and Oliver, 2007).

In addition, KED was evaluated to account for potential deterministic spatial trends driven by environmental covariates such as topographic variables derived from the Digital Elevation Model (DEM). KED extends OK by allowing the mean of the process to vary as a function of auxiliary predictors. This formulation is mathematically equivalent to Regression kriging (RK) and closely related to Universal kriging (UK) approaches (Hengl et al., 2007).

Drift Variables

Several drift configurations were evaluated to represent possible deterministic spatial trends in temperature fields.

Four drift configurations were tested in this study:

(i) spatial coordinates (x, y),

(ii) elevation derived from the DEM,

(iii) spatial coordinates combined with elevation (x, y, DEM), and

(iv) spatial coordinates together with multiple topographic predictors derived from the DEM, including elevation, slope, and aspect.

Aspect was represented through its sine and cosine components to preserve its circular nature.

These drift configurations were used to evaluate how different representations of topographic and spatial trends influence the performance of KED. The specific implementation of these drift models is described in the following section.

Neighborhood Definition

Local kriging neighbourhoods were defined to control the number of observations used in each interpolation. The maximum number of neighbouring observations was set to 20.

For models without drift variables such as OK, a minimum of one neighbour was required. For kriging models including drift variables, the minimum of neighbours was constrained by the number of drift parameters to ensure solvability of the kriging system.

Specifically, the minimum neighbourhood size was defined as $k \geq p + 1$, where p denotes the number of drift parameters.

Cross-Validation Strategy

The optimal neighbourhood size was determined using leave-one-out cross-validation (LOOCV). In this approach, observations are sequentially removed from the dataset and predicted using the remaining observations within the neighbourhood, allowing prediction errors to be evaluated for all locations (Willmott and Matsuura, 1995). The procedure was repeated for several candidate values of k, representing the number of neighbouring observations used for prediction.

This procedure was repeated for each timestamp, landscape-regime combination, temperature variable, and drift configuration. For each configuration, cross-validation was performed on a representative sample of 100 timestamps to obtain stable performance estimates.

Performance Metrics

Prediction accuracy was evaluated using several statistical metrics derived from cross-validation:

- Mean Error (ME), measuring prediction bias
- Mean Absolute Error (MAE)

- Root Mean Square Error (RMSE)
- Willmott's Index of Agreement (WIL)
- Variance ratio between predicted and observed values

For each timestamp, these metrics were computed from the leave-one-out prediction. The resulting values were then summarized across the 100 sampled timestamps to obtain a representative statistics for each landscape-regime-variable-model configuration.

Among these metrics, the MAE was used as the primary criterion for selecting the optimal neighbourhood size.

Using multiple timestamps ensures that model performance reflects a range of spatial and thermal conditions rather than a single temporal realization.

5.3 Computational Implementation

The main scripts used in this study were developed and initially tested in RStudio (version 2026.01.1) on a local workstation using the R statistical computing environment (R Core Team, 2024). However, the evaluation of multiple variogram estimators, theoretical variogram models, bootstrap resampling procedures, and interpolation configurations across numerous timestamps required substantial computational resources.

To efficiently handle these analyses, computations were executed on the MetaCentrum CESNET high-performance computing (HPC) infrastructure (CESNET, 2024). The HPC environment enabled parallel execution of model evaluation and resampling procedures, significantly reducing computation time and allowing systematic comparison of interpolation configurations.

CHAPTER 6

Results

6.1 Thermal Characteristics of the Aggregated Dataset

As shown previously in Figure 4.4, the monitored landscapes differed substantially in spatial extent and station density, with Agriculture representing the largest domain and KL the sparsest network. Temperature conditions also varied among landscapes.

Table 6.1: Descriptive statistics of temperature observations across landscapes and measurement levels (T1–T3).

Landscape	Variable	Mean	Median	SD	IQR	Min	Max
BP	T1	9.33	9.62	4.62	8.50	-1.94	19.84
BP	T2	9.63	9.94	5.66	9.25	-12.91	37.12
BP	T3	10.07	10.25	7.97	12.28	-17.06	44.88
KL	T1	9.38	9.81	4.58	8.53	0.66	19.28
KL	T2	9.57	9.88	5.29	9.34	-3.62	27.56
KL	T3	9.67	9.62	7.85	12.00	-15.03	46.88
Agroforestry	T1	14.65	16.12	5.94	8.06	-2.12	42.81
Agroforestry	T2	14.90	16.19	6.50	8.75	-10.88	42.19
Agroforestry	T3	14.96	15.44	8.69	11.88	-16.00	46.75
Drainage	T1	11.25	11.19	6.60	12.12	-2.19	34.00
Drainage	T2	11.55	11.50	6.62	12.06	-2.12	34.25
Drainage	T3	11.13	9.88	9.25	12.56	-14.88	48.75
Agriculture	T1	10.18	10.44	6.16	10.75	-2.75	29.12
Agriculture	T2	10.24	10.06	7.45	11.31	-9.75	48.38
Agriculture	T3	10.03	9.19	8.92	12.19	-17.88	51.12

Table 6.1 summarizes the thermal characteristics of the aggregated dataset across landscapes and measurement levels (T1–T3). Mean temperatures ranged

from approximately 9 to 15 °C, indicating clear differences among landscape units. The highest mean temperature was observed in Agroforestry, whereas BP and KL landscapes showed lower overall thermal conditions.

Across most landscapes, variability increased from T1 to T3, with surface-level measurements (T3) exhibiting the largest standard deviations and widest temperature ranges. This pattern reflects the stronger influence of atmospheric forcing and surface energy exchanges, whereas deeper measurements (T1) showed lower variability due to the thermal buffering effect of the soil.

The observed temperature ranges are consistent with expected environmental conditions in the study region, including winter minima and summer maxima. Median values were generally close to the means, which is consistent with the approximately symmetric distributions identified during the exploratory analysis as discussed in Section 4.4.

Although Agroforestry is included here to represent the full monitored temperature range of the study area, it was excluded from the subsequent kilometre-scale interpolation analysis because its dense micro-scale configuration differs substantially from the other landscape networks. The remaining landscapes exhibited clear differences in thermal variability and spatial extent, suggesting that their spatial covariance structures should be evaluated separately in the following sections.

6.2 Diagnostic Spatial Structure

Based on the diagnostic variogram analysis described in Section 5.2, Figure 6.1 presents representative empirical semivariograms illustrating the range of spatial behaviours observed across the different landscapes. The examples were selected from the stratified sampling of timestamps used in the diagnosis phase and demonstrate differences in spatial dependence, ranging from well-defined spatial structure to weak or unstable variograms.

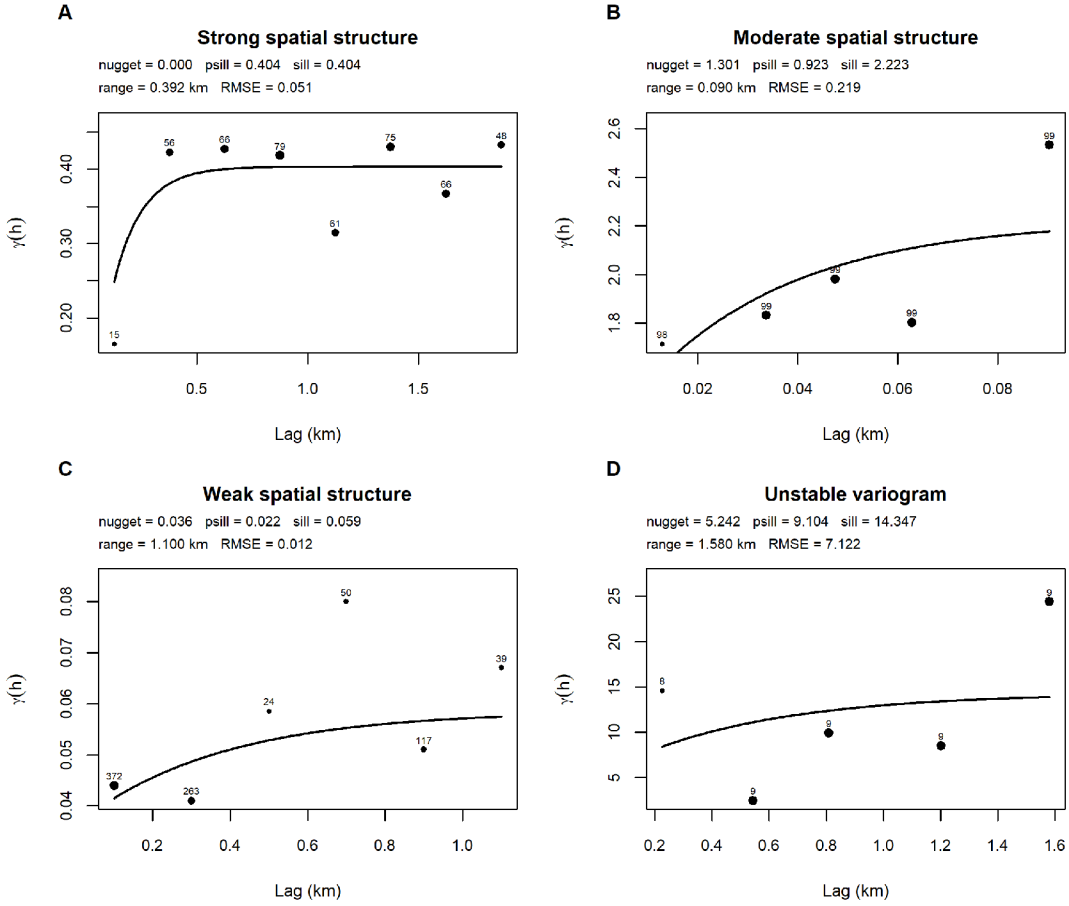

FIGURE 6.1: Representative empirical variograms illustrating different types of spatial structure observed in the temperature monitoring networks. Empirical semivariance was estimated using the Matheron estimator, and a simple exponential model was fitted for diagnostic purposes. Numbers above points indicate the number of point pairs within each lag class. Panels A–D correspond to representative examples selected from the Agriculture, Drainage, BP, and KL landscapes, each representing a distinct monitoring network geometry.

Panel A shows a case of strong spatial structure, where semivariance increases rapidly with lag distance and approaches a clear sill. The empirical points follow a smooth increasing pattern, and the fitted exponential model closely approximates the observed structure. This behaviour indicates a stable spatial covariance structure with a well-defined correlation range of approximately 0.4 km.

Panel B illustrates moderate spatial dependence. Semivariance increases with distance, but the spatial structure is less pronounced and the empirical semivariance values exhibit greater variability. Although a spatial trend is still evident, the relationship between lag distance and semivariance is weaker than in Panel A.

Panel C represents weak spatial structure. The empirical semivariogram

shows only a small increase in semivariance with lag distance, and the sill is reached rapidly. This pattern indicates limited spatial variability and suggests that temperature differences between locations are relatively small.

Panel D demonstrates an unstable variogram, where empirical semivariance values vary irregularly across lag classes and no clear spatial structure is evident. Such behaviour can occur when spatial gradients are weak, when spatial variability is low relative to measurement noise, or when the monitoring network provides limited spatial support.

Together, these examples confirm that the temperature observations exhibit varying degrees of spatial dependence, supporting the use of geostatistical interpolation while highlighting the importance of evaluating empirical variogram stability prior to covariance model fitting.

6.3 Variogram Estimation Stability

Diagnostic outputs from the variogram diagnosis phase were examined to assess the stability of variogram parameter estimation and to identify failures in the model fitting procedure. Table 6.2 summarizes the fitting diagnostics for each landscape and temperature measurement (T1–T3), including the proportion of failed model fits, the average number of lag bins used in the empirical variogram, and the mean number of spatial observations (N).

Table 6.2: Variogram fitting diagnostics from the parameter stability analysis.

Landscape	Variable	Fail Rate	Mean Lag Bins	Mean Spatial Support (N)
BP	T1	0.083	5.57	45.03
BP	T2	0.319	5.57	43.87
BP	T3	0.008	5.57	45.03
Drainage	T1	0.192	5.00	41.40
Drainage	T2	0.175	5.00	41.40
Drainage	T3	0.044	5.00	41.40
KL	T1	0.356	5.00	10.00
KL	T2	0.527	5.00	10.00
KL	T3	0.475	5.00	10.00
Agriculture	T1	0.067	7.67	32.97
Agriculture	T2	0.061	7.67	32.97
Agriculture	T3	0.058	7.67	32.97

Agriculture. The Agriculture landscape exhibits highly stable variogram estimation, with failure rates of approximately 6% across all temperature measurements. The average number of lag bins is approximately 7.7, indicating

strong spatial support and well-defined empirical variogram structures. Among all landscapes, this network provides the most reliable spatial information for variogram modeling.

Drainage. The Drainage landscape shows intermediate behaviour, with failure rates between 17% and 19% for T1 and T2. These values indicate moderate spatial instability, potentially related to local heterogeneity in spatial temperature structure or limitations in the sampling geometry.

BP. The BP landscape shows a variable-specific instability for the T2 measurement. Despite adequate spatial support ($N \approx 44$), the failure rate reaches approximately 32%. This behaviour may indicate frequent pure nugget structures, extremely short spatial ranges relative to the sensor spacing, or nearly flat empirical variograms. These diagnostics suggest that the T2 temperature measurement exhibits weak spatial structure at many timestamps.

KL. The KL landscape exhibits a mean spatial support of only $N = 10$ locations and failure rates between 35% and 53%. This indicates insufficient spatial support for reliable variogram estimation. With such a limited number of spatial locations, the variogram cloud becomes sparse, resulting in a reduced number of lag classes and unstable model fitting. Under these conditions the exponential model becomes poorly constrained, leading to an unstable nugget–range trade-off. Consequently, variogram diagnosis for the KL landscape is strongly limited by the sampling geometry and regime-level decisions for this network must be interpreted with caution.

These diagnostics provide important guidance for the subsequent regime modeling stage by identifying landscapes where spatial structure can be reliably estimated and those where kriging predictions may be constrained by limited spatial support.

6.4 Temporal Behavior of Variogram Parameters

To evaluate potential temporal variability in the spatial covariance structure, diagnostic variogram parameters (nugget and sill) were examined as functions of hour-of-day and month for each landscape and temperature level. These diagnostics provide insight into whether spatial temperature variability exhibits systematic diurnal or seasonal patterns that justify the adoption of regime-based variogram modeling.

Agriculture

Diagnostic plots of the nugget and sill parameters for the Agriculture landscape (Figure 6.2) reveal a clear and consistent diurnal modulation of the spatial covariance structure across all temperature measurements (T1–T3). Both parameters exhibit low values during nighttime hours, followed by a gradual increase during the morning transition period and culminating in a pronounced peak between approximately 09:00 and 15:00.

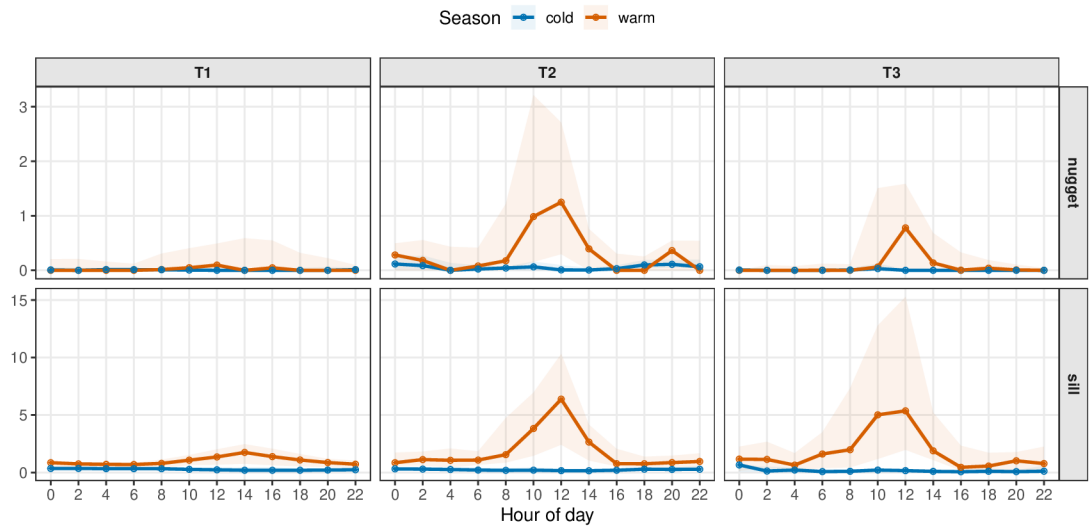

FIGURE 6.2: Temporal behaviour of nugget and sill parameters in the Agriculture landscape. Lines represent median values across seasons and ribbons indicate the interquartile range.

This midday amplification reflects periods of enhanced surface heating and stronger spatial temperature contrasts across the landscape. Rather than indicating a simple binary day–night separation, the observed pattern suggests a smooth cyclic structure composed of three distinct phases. Based on these diagnostics, a three-phase diurnal regime was defined:

- Night: 00:00–05:00 and 20:00–22:00
- Midday peak: 09:00–15:00
- Transition periods: 06:00–08:00 and 16:00–19:00

The timing of these phases corresponds closely to the observed minima and maxima in the variogram parameters and reflects shifts in the surface energy balance and atmospheric mixing conditions.

In addition to the diurnal modulation, clear seasonal differences in the magnitude of the variogram parameter are observed. Variogram parameters are

substantially larger during the warm season, indicating stronger spatial temperature contrasts, whereas the cold season exhibits relatively low and stable parameter values. Importantly, the timing of the diurnal peaks remains consistent across seasons, suggesting that the temporal structure is stable while its magnitude varies seasonally.

Based on these diagnostics, a binary seasonal stratification was defined separating a warm period (March–October) from a cold period (November–February). Combining the diurnal and seasonal components results in six regime-specific variogram models (three diurnal phases × two seasonal groups) for the Agriculture landscape.

Drainage

The Drainage landscape exhibits a similarly coherent temporal structure (Figure 6.3). Both nugget and sill parameters display clear diurnal patterns for the T1 and T2 temperature measurements, characterized by low nighttime values and pronounced midday peaks occurring between approximately 09:00 and 15:00.

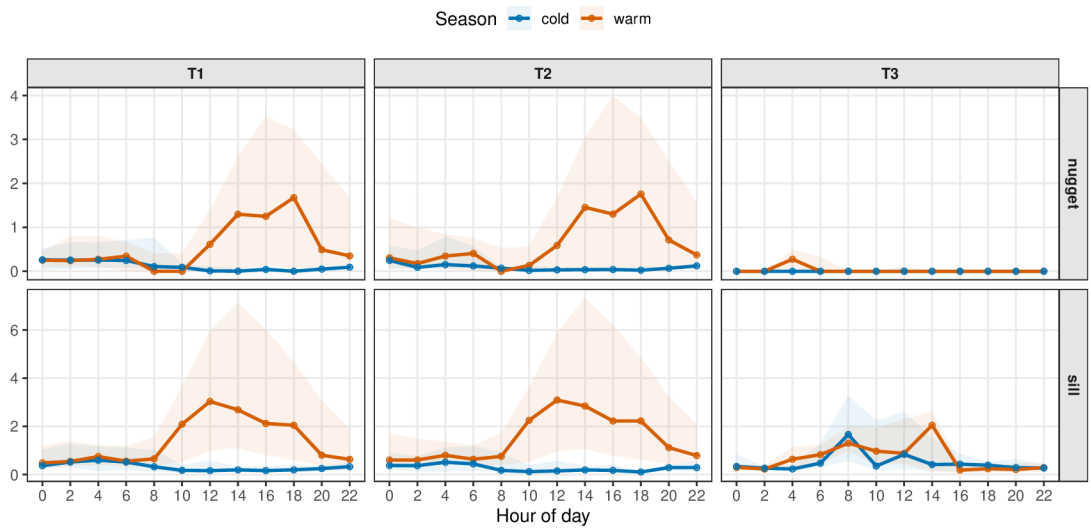

FIGURE 6.3: Temporal behaviour of nugget and sill parameters in the Drainage landscape. Graphical conventions follow those defined in Figure 6.2.

These peaks are particularly evident during the warm season, when increased surface heating produces stronger spatial temperature contrasts across the landscape. During the cold season, variogram parameters remain substantially lower and exhibit weaker diurnal modulation.

The T3 measurement shows greater variability compared to T1 and T2, which likely reflects the stronger sensitivity of near-surface temperatures to local micro-

climatic conditions. Nevertheless, the overall timing of the midday amplification remains consistent with the patterns observed in the other temperature levels.

Based on these diagnostics, the same three-phase diurnal regime defined for the Agriculture landscape was adopted:

- Night: 00:00–05:00 and 20:00–22:00
- Midday peak: 09:00–15:00
- Transition periods: 06:00–08:00 and 16:00–19:00

Seasonal differences in parameter magnitude further justify the introduction of a seasonal stratification separating a warm period (May–September) from a cold period (October–April). The combination of these diurnal and seasonal components results in six regime-specific variogram models for the Drainage landscape.

BP

In contrast to the previous landscapes, diagnostic plots for the BP landscape reveals substantially weaker and less coherent temporal patterns (Figure 6.4). Although occasional increases in nugget or sill values are observed, these peaks occur irregularly and are not consistently reproduced across temperature measurements or seasons.

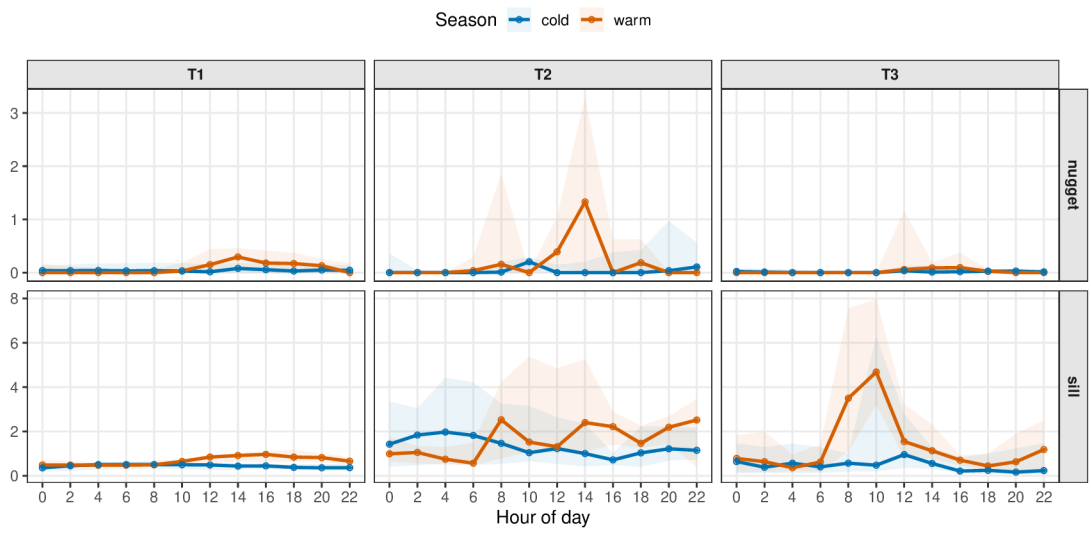

FIGURE 6.4: Temporal behaviour of nugget and sill parameters in the BP landscape. Graphical conventions follow those defined in Figure 6.2.

The diurnal structure is therefore less pronounced than in the Agriculture and Drainage landscapes. Some midday amplification can be observed sporadically, particularly in the warm season, but these patterns vary considerably between variables and do not form a stable or interpretable temporal structure.

Because neither diurnal nor seasonal patterns are consistently reproduced across the temperature measurements, regime-based stratification was not adopted for the BP landscape. Instead, stationary variogram models were retained to ensure parameter stability and to avoid introducing unnecessary model complexity.

KL

The KL landscape exhibits the weakest and most irregular temporal patterns among the analyzed landscapes (Figure 6.5). While occasional midday increases in variogram parameters can be observed, these patterns are inconsistent across temperature measurements and display substantial variability between hours and seasons.

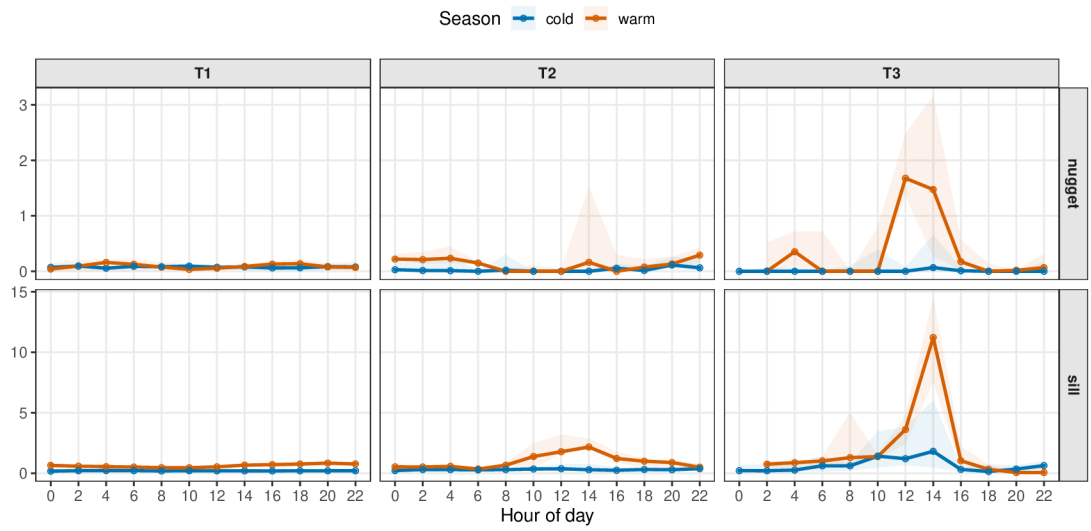

FIGURE 6.5: Temporal behaviour of nugget and sill parameters in the KL landscape. Graphical conventions follow those defined in Figure 6.2.

This instability is largely attributable to the limited spatial support of the KL network, which contains only ten spatial locations. With such a sparse sampling configuration, empirical variograms are based on a small number of spatial pairs, resulting in unstable parameter estimates and relatively high variogram fitting failure rates.

Because of this limitation, introducing temporal regimes would likely increase model complexity without improving parameter reliability. Consequently, both

diurnal and seasonal regime stratifications were rejected for the KL landscape, and a stationary variogram model was retained.

Table 6.3: Summary of regime stratification decisions across landscapes.

Landscape	Regime Decision
Agriculture	3 diurnal × 2 seasonal
Drainage	3 diurnal × 2 seasonal
BP	Stationary
KL	Stationary

This regime structure balances model flexibility and robustness by introducing temporal stratification only in landscapes where stable and physically interpretable patterns are observed.

6.5 Regime-Specific Data Availability

After defining the regime stratification, the number of available timestamps for each landscape–regime combination was quantified in order to determine the feasible sampling space for bootstrap resampling. Table 6.4 summarizes the total number of timestamps available within each regime-specific subset.

Table 6.4: Number of timestamps available for each landscape–regime combination prior to bootstrap-based assessment of variogram parameter stability.

Landscape	Regime	Number of Timestamps
BP	stationary	117354
KL	stationary	111216
Drainage	cold_midday	5343
Drainage	cold_night	7594
Drainage	cold_transition	5346
Drainage	warm_midday	4256
Drainage	warm_night	6115
Drainage	warm_transition	4263
Agriculture	cold_midday	6748
Agriculture	cold_night	9640
Agriculture	cold_transition	6748
Agriculture	warm_midday	14753
Agriculture	warm_night	21056
Agriculture	warm_transition	14744

The stationary landscapes (BP and KL) contain the full temporal record, resulting in more than 100,000 timestamps for each.

In contrast, landscapes exhibiting regime-specific spatial behavior (Drainage and Agriculture) are subdivided into six temporal regimes (three diurnal phases ×

two seasonal groups). As a result, the number of timestamps per subset is smaller, ranging from approximately 4,000 to 21,000 depending on the regime.

Despite this subdivision, each regime retains a sufficiently large temporal sample to support bootstrap-based variogram estimation.

6.6 Bootstrap Assessment of Variogram Parameter Stability

Bootstrap resampling was used to evaluate the stability of variogram parameter estimation across candidate timestamp pool sizes. The stabilization behaviour of the partial sill parameter was used as the primary diagnostic criterion, following the methodology described in section 5.2.

For the Agriculture landscape, stabilization of the partial sill parameter was generally achieved at pooling levels of approximately $n \approx 200$ across regimes and temperature variables. In contrast, nugget variability continued to decrease with increasing pooling and only approached stable values around $n \approx 400$.

This behaviour is consistent with the nugget parameter's sensitivity to microscale variability and fluctuations among short-distance observation pairs, which typically require larger effective sample sizes for stable estimation. To ensure robust estimation of both large-scale spatial variance and the nugget component, $n = 400$ was therefore retained as the operational pooling criterion for the Agriculture landscape.

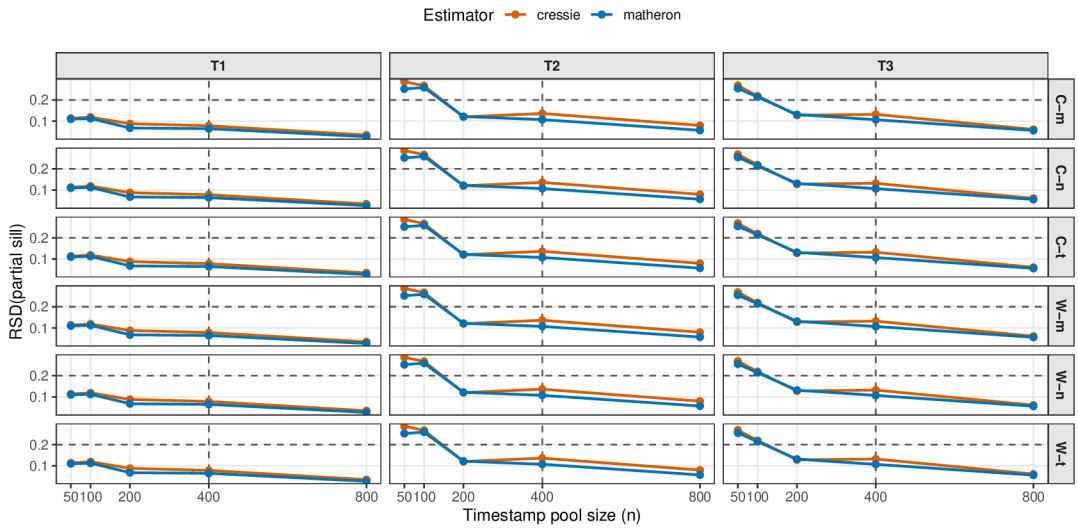

FIGURE 6.6: Bootstrap stability diagnostics of the partial sill parameter for the Agriculture landscape. Curves show the relative standard deviation of the partial sill across candidate timestamp pool sizes for the Matheron and Cressie–Hawkins estimators. The horizontal dashed line marks the stability threshold ($RSD = 0.20$), and the vertical dashed line indicates the selected pool size. Regime labels are abbreviated as follows: C-m = cold-midday, C-n = cold-night, C-t = cold-transition, W-m = warm-midday, W-n = warm-night, and W-t = warm-transition.

For the Drainage landscape, stabilization of the partial sill parameter was achieved at approximately $n \approx 200$ across all regimes and temperature measurements. Nugget variability exhibited slightly different behaviour among temperature levels. The T1 and T2 measurements showed stable nugget estimates even at relatively small pooling levels ($n = 50$–100), whereas the T3 measurement required slightly larger pooling ($n \approx 100$) to achieve comparable stability.

Beyond these pooling levels, additional increases in n did not substantially reduce parameter variability. These results indicate that, for regime-stratified landscapes such as Drainage, moderate timestamp pooling is sufficient to ensure stable estimation of both the large-scale spatial variance (partial sill) and the microscale variance component (nugget).

The KL landscape exhibited the strongest instability among all study areas. While the T1 measurement achieved partial sill stabilization at moderate pooling levels, the T2 measurement showed persistent identifiability issues under both estimators, characterized by low convergence rates and highly variable parameter estimates. Increasing the pooling level did not consistently improve stability for this variable, suggesting heterogeneous spatial covariance patterns that are difficult to represent using a strictly stationary variogram model.

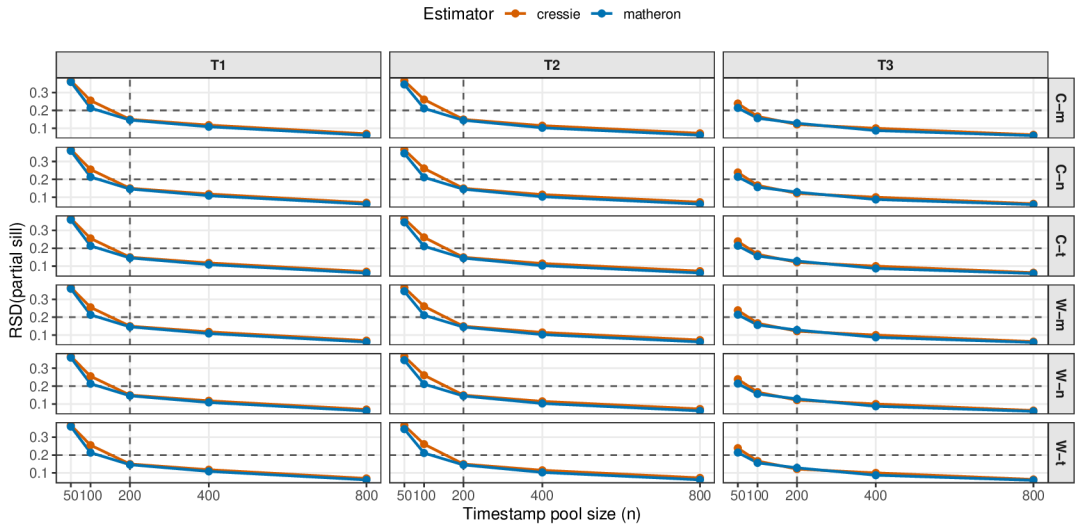

FIGURE 6.7: Bootstrap stability diagnostics of the partial sill parameter for the Drainage landscape. Graphical conventions follow those defined in Figure 6.6.

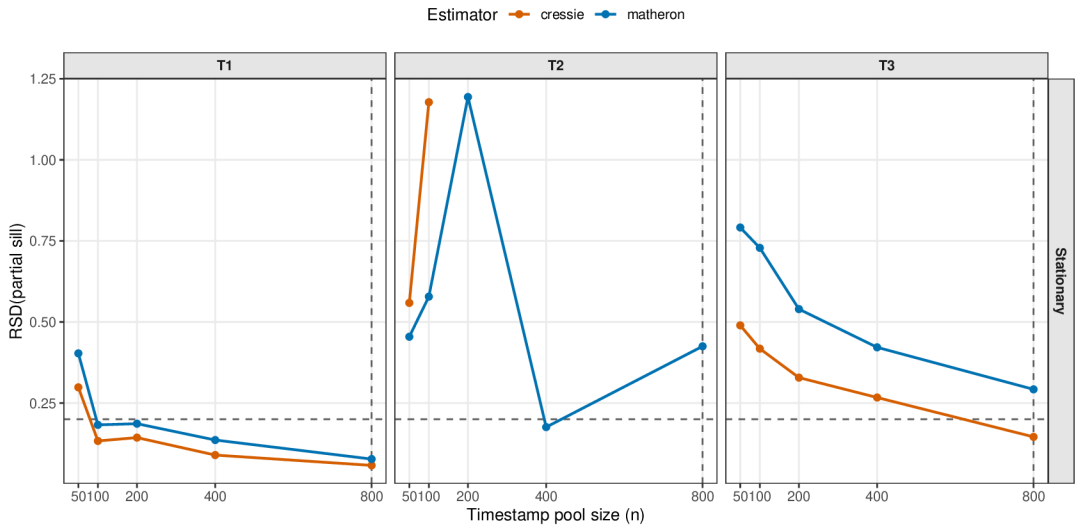

FIGURE 6.8: Bootstrap stability diagnostics of the partial sill parameter for the KL landscape. Graphical conventions follow those defined in Figure 6.6.

For the T3 measurement, the Cressie–Hawkins estimator showed progressive stabilization with increasing pooling and achieved acceptable variability at $n = 800$, whereas the Matheron estimator remained unstable across pooling levels. Consequently, a pooling level of $n = 800$ using the Cressie estimator was adopted for the KL landscape in order to obtain robust partial sill estimates across temperature variables, while acknowledging residual instability for the T2 measurement.

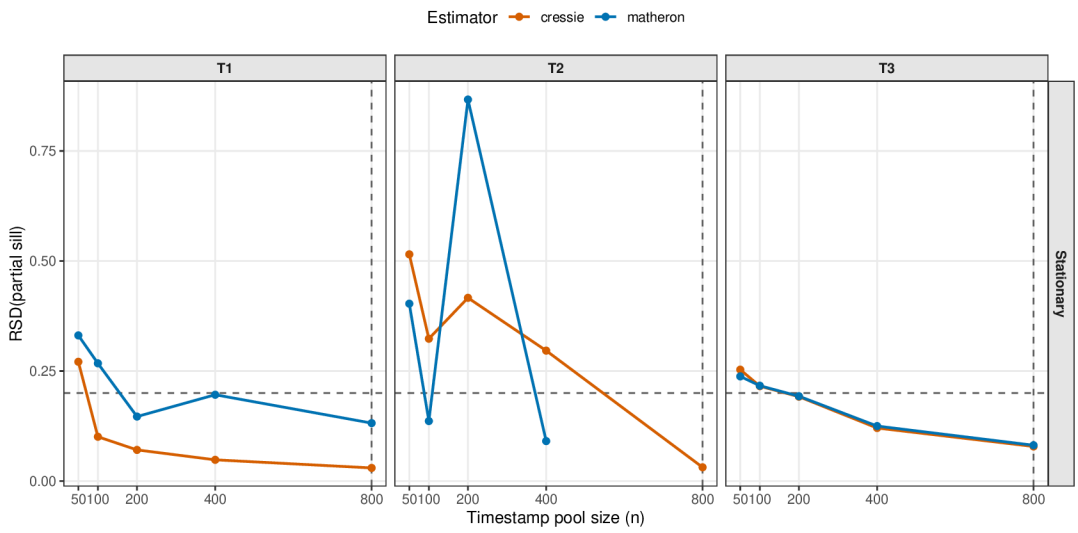

FIGURE 6.9: Bootstrap stability diagnostics of the partial sill parameter for the BP landscape. Graphical conventions follow those defined in Figure 6.6.

For the BP landscape, stabilization behaviour differed among temperature variables. The T1 and T3 measurements exhibited stable partial sill estimates at intermediate pooling levels ($n \approx 400$). However, the T2 measurement demonstrated persistent instability, with acceptable partial sill variability and satisfactory convergence achieved only at the largest pooling level ($n = 800$) when using the robust Cressie–Hawkins estimator.

In contrast, the classical Matheron estimator exhibited consistently lower convergence rates and unstable parameter behaviour for this subset. To ensure stable variogram parameter estimation across all temperature measurements, a pooling level of $n = 800$ was therefore selected for the BP landscape.

The bootstrap diagnostics revealed different stabilization thresholds among landscapes. Table 6.5 summarizes the smallest timestamp pool sizes satisfying the stability criterion for each landscape.

The Drainage landscape required the smallest timestamp pool size to achieve stable variogram parameter estimates, whereas the KL and BP landscapes required substantially larger pools. This behaviour is consistent with the spatial sampling

Table 6.5: Selected timestamp pool sizes determined from bootstrap stability diagnostics for each landscape.

Landscape	Selected Pool Size (n)
Drainage	200
Agriculture	400
KL	800
BP	800

characteristics identified during the variogram diagnosis phase, where limited spatial support and weaker spatial structure were observed for these landscapes.

These pool sizes were subsequently used in the final regime-specific variogram estimation stage. A complete set of bootstrap diagnostics, including nugget parameter variability across pooling levels, is provided in Appendix A.

6.7 Final Variogram Models

Following the bootstrap stability analysis, final variogram models were estimated for each landscape–regime–temperature combination using the selected timestamp pool sizes. For each subset, several theoretical variogram families (exponential, spherical, Gaussian, and Matérn) were fitted to the pooled empirical variograms. Model selection was based primarily on root mean square error (RMSE), while Akaike Information Criterion (AIC) was used as a secondary metric to support discrimination among competing models.

A comparison between the classical Matheron semivariance estimator and the robust Cressie–Hawkins estimator indicated that the Matheron estimator consistently provided lower RMSE values and more stable parameter estimates across all landscapes and regimes. Consequently, the final variogram models were derived using the Matheron estimator.

During this stage, the upper bound imposed on the range parameter during the bootstrap stability analysis was relaxed. This adjustment was necessary because the previously imposed cap occasionally constrained the range estimates and prevented the models from adequately representing the spatial correlation structure. Allowing a larger range during the final fitting stage enabled the variogram models to better capture the scale of spatial dependence present in the data.

Table 6.6 summarizes the final variogram parameters selected for each landscape, regime, and temperature measurements. Regime abbreviations are used to reduce table width and are explained in the caption.

Table 6.6: Final variogram models selected for each landscape, regime, and temperature measurement using the Matheron semivariance estimator. Regime codes are: S = Stationary, C-m = Cold-midday, C-n = Cold-night, C-t = Cold-transition, W-m = Warm-midday, W-n = Warm-night, and W-t = Warm-transition. Model selection was based primarily on RMSE, with AIC used as a secondary criterion.

Regime	Variable	Model	Nugget	Psill	Range	ν	RMSE	AIC
BP								
S	T1	Spherical	0.433	0.277	0.503	–	0.291	66.599
S	T2	Gaussian	2.983	0.053	0.001	–	0.454	79.720
S	T3	Gaussian	0.387	0.868	0.616	–	0.650	77.996
KL								
S	T1	Gaussian	0.293	0.129	1.314	–	0.091	42.554
S	T2	Spherical	0.636	0.070	0.886	–	0.124	45.751
S	T3	Spherical	1.409	0.286	0.953	–	0.196	50.293
Drainage								
C-m	T1	Gaussian	0.452	0.280	0.181	–	0.034	36.907
C-m	T2	Gaussian	0.546	0.166	0.181	–	0.038	38.110
C-m	T3	Spherical	0.336	0.835	0.066	–	0.178	53.221
C-n	T1	Gaussian	0.411	0.151	0.181	–	0.021	32.546
C-n	T2	Gaussian	0.521	0.155	0.181	–	0.030	35.982
C-n	T3	Matérn	0.050	0.630	0.021	2.000	0.097	47.104
C-t	T1	Gaussian	0.505	0.111	0.181	–	0.026	34.412
C-t	T2	Gaussian	0.653	0.077	0.181	–	0.038	38.233
C-t	T3	Spherical	0.000	0.858	0.033	–	0.131	50.146
W-m	T1	Matérn	1.764	7.675	0.181	2.000	0.287	59.109
W-m	T2	Matérn	1.888	7.156	0.181	2.000	0.294	59.347
W-m	T3	Spherical	0.528	1.726	0.058	–	0.206	55.752
W-n	T1	Gaussian	0.516	0.559	0.181	–	0.061	43.608
W-n	T2	Gaussian	0.589	0.813	0.181	–	0.061	43.652
W-n	T3	Spherical	0.230	0.378	0.070	–	0.094	47.904
W-t	T1	Gaussian	1.316	1.612	0.181	–	0.101	48.657
W-t	T2	Gaussian	1.627	1.795	0.181	–	0.106	49.121
W-t	T3	Spherical	0.421	0.401	0.065	–	0.138	51.758
Agriculture								
C-m	T1	Matérn	0.158	0.757	3.750	2.000	0.045	58.632
C-m	T2	Spherical	0.423	0.210	1.308	–	0.041	57.832
C-m	T3	Spherical	0.617	0.420	1.261	–	0.069	65.197
C-n	T1	Matérn	0.171	0.733	3.750	2.000	0.053	61.548
C-n	T2	Matérn	0.489	0.786	3.750	2.000	0.055	62.366
C-n	T3	Matérn	0.367	1.578	3.750	2.000	0.046	59.460
C-t	T1	Matérn	0.178	0.735	3.750	2.000	0.053	61.610
C-t	T2	Matérn	0.454	0.557	3.750	2.000	0.058	63.632
C-t	T3	Spherical	0.365	0.468	3.750	–	0.038	56.580
W-m	T1	Matérn	0.894	1.892	3.750	2.000	0.183	80.632
W-m	T2	Matérn	3.110	5.715	3.750	2.000	0.321	92.265
W-m	T3	Matérn	2.276	7.281	3.750	0.300	0.485	97.622
W-n	T1	Matérn	0.261	0.807	3.750	0.300	0.084	68.443
W-n	T2	Gaussian	0.659	0.885	3.750	–	0.085	70.738
W-n	T3	Matérn	0.509	3.979	3.750	2.000	0.162	79.746
W-t	T1	Matérn	0.689	1.659	3.750	2.000	0.170	76.684
W-t	T2	Matérn	1.223	1.933	3.750	2.000	0.118	75.823
W-t	T3	Matérn	1.837	3.540	3.750	2.000	0.348	86.088

The variogram model selection shows clear differences among landscapes. For the stationary landscapes (BP and KL), the best-performing models were primarily Gaussian and spherical. These models provided the lowest RMSE values while maintaining stable parameter estimates, indicating relatively simple spatial covariance structures without the need for more flexible model forms.

In the Drainage landscape, model performance was more heterogeneous across regimes and temperature variables. Gaussian models were frequently selected for T1 and T2, whereas spherical and Matérn models provided better fits for several regimes of the T3 measurement. This variability reflects the stronger regime-dependent changes in spatial temperature structure identified during the temporal diagnostics.

For the Agriculture landscape, the Matérn model was selected in the majority of regimes, particularly during the warm season. The flexibility of the Matérn covariance function allows it to capture smoother spatial transitions and longer correlation ranges that arise during periods of strong daytime heating. During the cold regimes, spherical models occasionally provided comparable or slightly better fits, indicating weaker spatial gradients and more abrupt spatial decorrelation under colder conditions.

Overall, these results highlight that landscapes exhibiting regime-dependent spatial variability tend to favour more flexible covariance models, whereas stationary landscapes are adequately described by simpler Gaussian or spherical structures.

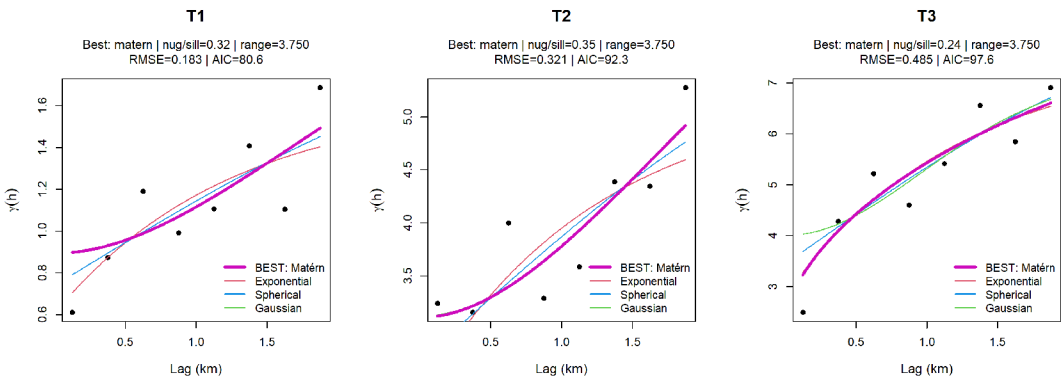

FIGURE 6.10: Representative variogram model fits for the Agriculture landscape during the warm-midday regime. Points represent the pooled empirical semivariogram for the three temperature measurements (T1–T3). Thin curves correspond to candidate theoretical models (exponential, spherical, Gaussian, and Matérn), while the thick curve indicates the selected best-fitting model based on RMSE.

During the warm-midday regime, the empirical variograms exhibit large sill

values and a clear increase with lag distance across all temperature measurements. The Matérn model provides the best fit for T1–T3, indicating a smooth spatial covariance structure and relatively long spatial correlation ranges. This behaviour is consistent with enhanced surface heating during midday hours, which produces stronger spatial temperature gradients across the landscape.

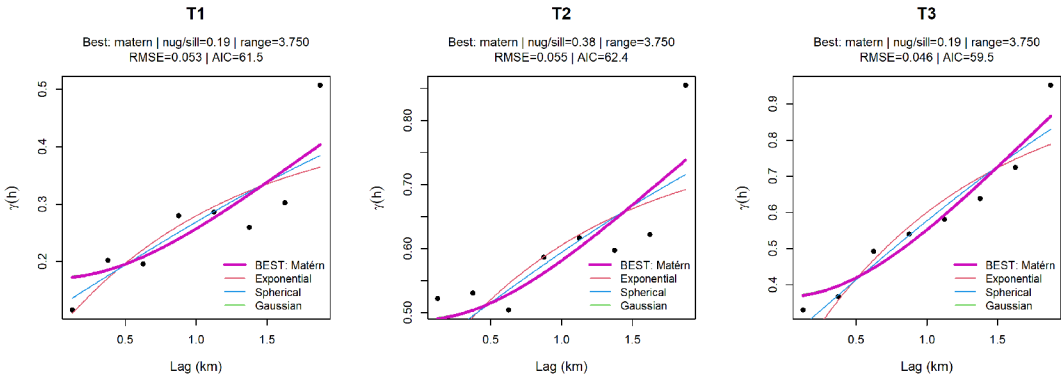

FIGURE 6.11: Representative variogram model fits for the Agriculture landscape during the cold-night regime. Graphical conventions follow those described in Figure 6.10.

In contrast to the warm regime, the cold-night variograms display substantially lower sill values, reflecting weaker spatial temperature contrasts. Although the Matérn model remains the best-performing covariance structure for all temperature levels, the overall magnitude of spatial variability is reduced. This pattern is consistent with more homogeneous nighttime atmospheric conditions and weaker surface-driven temperature gradients.

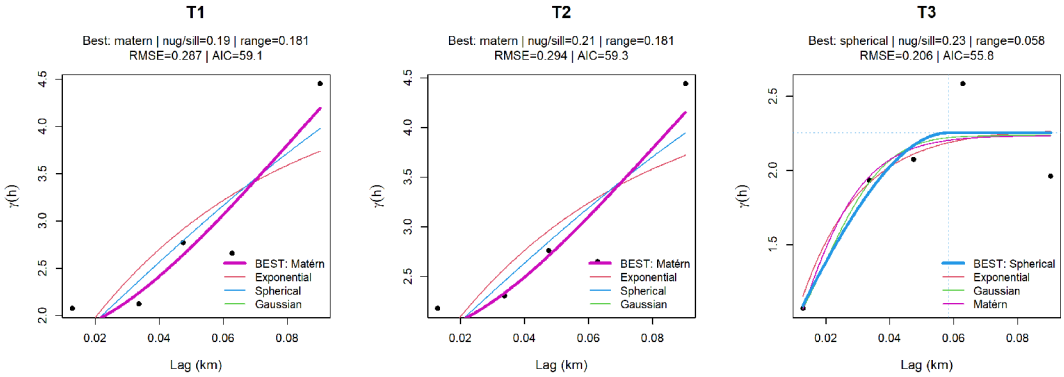

FIGURE 6.12: Representative variogram model fits for the Drainage landscape during the warm-midday regime. Graphical conventions follow those described in Figure 6.10. The vertical dashed line marks the estimated correlation range and the horizontal dashed line indicates the sill.

For the Drainage landscape, the warm-midday regime also exhibits pronounced spatial variability. The Matérn model provides the best fit for the T1 and T2 temperature measurements, whereas the spherical model yields the lowest RMSE for T3. Compared with the Agriculture landscape, the estimated spatial ranges are shorter, indicating a more rapid decay of spatial temperature correlation within the Drainage landscape.

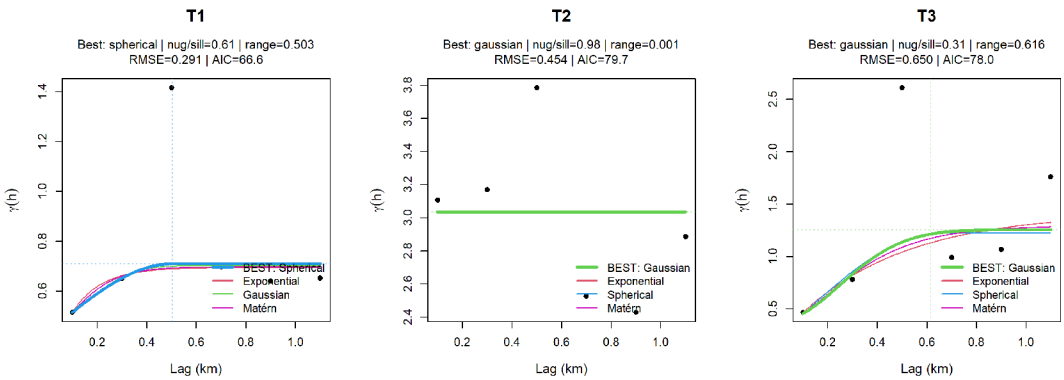

FIGURE 6.13: Representative variogram model fits for the BP landscape under stationary conditions. Graphical conventions follow those described in Figure 6.10. The vertical dashed line marks the estimated correlation range and the horizontal dashed line indicates the sill.

The BP landscape exhibits a stationary spatial structure without clear regime-dependent behaviour. The best-performing covariance models differ among temperature measurements: the spherical model provides the best fit for T1, whereas Gaussian models are selected for T2 and T3. The relatively flat variogram for

T2, characterized by a dominant nugget component and a very short estimated range, indicates weak spatial dependence. These results support the decision to retain a stationary variogram model for this landscape rather than introducing regime-specific stratification.

These fitted models constitute the final covariance structures used in the subsequent kriging interpolation and validation analysis. The complete set of final variogram models for all landscape–regime–temperature combinations is provided in Appendix B.

6.8 Kriging Cross-Validation Performance

The predictive performance of the kriging models was evaluated using leave-one-out cross-validation for each landscape, regime, and temperature measurement. Model comparison was based primarily on the mean absolute error (MAE), with the RMSE used as a secondary criterion. In addition to Ordinary kriging (OK), several kriging with external drift (KED) variants were evaluated using terrain-based predictors. The results were summarized across 100 timestamps per subset in order to identify the best-performing interpolation model for each landscape–regime–temperature combination.

Table 6.7 summarizes the best-performing kriging models based on cross-validation. The results reveal clear differences in kriging model performance among landscapes.

For the stationary landscapes (BP and KL), OK consistently provided the lowest prediction errors across all temperature variables. This result indicates that the spatial temperature patterns in these landscapes can be adequately captured using the variogram-based covariance structure alone, without the need for additional drift predictors.

In the Drainage landscape, OK also produced the best predictive performance across all regimes and temperature measurements. Although several KED variants were evaluated, none consistently improved the cross-validation errors relative to OK. This suggests that the spatial variability of temperature in this landscape is primarily driven by the spatial correlation structure captured by the variogram model rather than by the auxiliary predictors considered in this study.

In contrast, the Agriculture landscape shows a different behaviour. While OK performed well for the colder regimes and subsurface and deeper soil temperature measurements (T1 and T2), the KED model using elevation as predictor (KED_dem) consistently achieved lower errors for the T3 measurement and for most warm regimes. This result indicates that elevation plays an important role

Table 6.7: Best-performing kriging models based on cross-validation. Model selection was primarily based MAE, with RMSE used as a secondary criterion. Regime codes: S = stationary; C-m = cold-midday; C-n = cold-night; C-t = cold-transition; W-m = warm-midday; W-n = warm-night; W-t = warm-transition.

Landscape	Regime	Var	Model	Drift	MAE	RMSE
Stationary landscapes						
BP	S	T1	OK	–	0.508	0.660
BP	S	T2	OK	–	1.051	1.293
BP	S	T3	OK	–	0.436	0.737
KL	S	T1	OK	–	0.469	0.585
KL	S	T2	OK	–	0.586	0.743
KL	S	T3	OK	–	0.688	0.918
Drainage landscape						
Cold regimes						
Drainage	C-m	T1	OK	–	0.567	0.674
Drainage	C-m	T2	OK	–	0.601	0.724
Drainage	C-m	T3	OK	–	0.555	0.738
Drainage	C-n	T1	OK	–	0.547	0.645
Drainage	C-n	T2	OK	–	0.592	0.710
Drainage	C-n	T3	OK	–	0.337	0.446
Drainage	C-t	T1	OK	–	0.550	0.646
Drainage	C-t	T2	OK	–	0.593	0.706
Drainage	C-t	T3	OK	–	0.435	0.582
Warm regimes						
Drainage	W-m	T1	OK	–	1.076	1.323
Drainage	W-m	T2	OK	–	1.074	1.328
Drainage	W-m	T3	OK	–	0.574	0.838
Drainage	W-n	T1	OK	–	0.670	0.797
Drainage	W-n	T2	OK	–	0.698	0.833
Drainage	W-n	T3	OK	–	0.315	0.427
Drainage	W-t	T1	OK	–	0.860	1.025
Drainage	W-t	T2	OK	–	0.885	1.050
Drainage	W-t	T3	OK	–	0.420	0.586
Agriculture landscape						
Cold regimes						
Agriculture	C-m	T1	OK	–	0.343	0.455
Agriculture	C-m	T2	OK	–	0.542	0.683
Agriculture	C-m	T3	KED	elev_dem	0.512	0.674
Agriculture	C-n	T1	OK	–	0.355	0.471
Agriculture	C-n	T2	OK	–	0.554	0.685
Agriculture	C-n	T3	KED	elev_dem	0.374	0.482
Agriculture	C-t	T1	OK	–	0.346	0.461
Agriculture	C-t	T2	OK	–	0.513	0.639
Agriculture	C-t	T3	KED	elev_dem	0.389	0.505
Warm regimes						
Agriculture	W-m	T1	KED	elev_dem	0.742	0.937
Agriculture	W-m	T2	KED	elev_dem	1.379	1.740
Agriculture	W-m	T3	KED	elev_dem	1.399	1.806
Agriculture	W-n	T1	KED	elev_dem	0.482	0.628
Agriculture	W-n	T2	KED	elev_dem	0.613	0.796
Agriculture	W-n	T3	KED	elev_dem	0.566	0.724
Agriculture	W-t	T1	KED	elev_dem	0.613	0.787
Agriculture	W-t	T2	KED	elev_dem	0.782	0.995
Agriculture	W-t	T3	KED	elev_dem	0.938	1.202

in explaining spatial temperature variability in this landscape, particularly during periods of stronger thermal gradients.

Overall, the results demonstrate that the benefit of external drift predictors depends strongly on the landscape characteristics. In landscapes with relatively homogeneous spatial structure, OK is sufficient, whereas landscapes with stronger topographic control on temperature variability can benefit from the inclusion of terrain-based predictors.

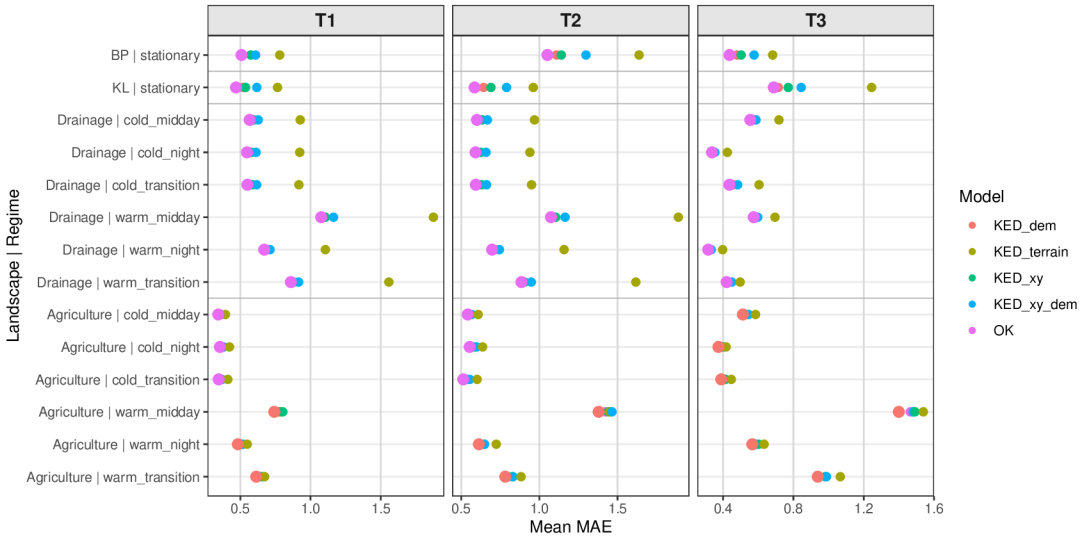

FIGURE 6.14: Comparison of MAE across kriging models for all landscapes, regimes, and temperature measurements. Points represent the average cross-validation error over 100 timestamps for each subset. OK provides the lowest errors for the stationary landscapes and the Drainage landscape, whereas KED using elevation (*KED_dem*) improves predictive performance for several regimes of the Agriculture landscape, particularly for the T3 temperature measurement.

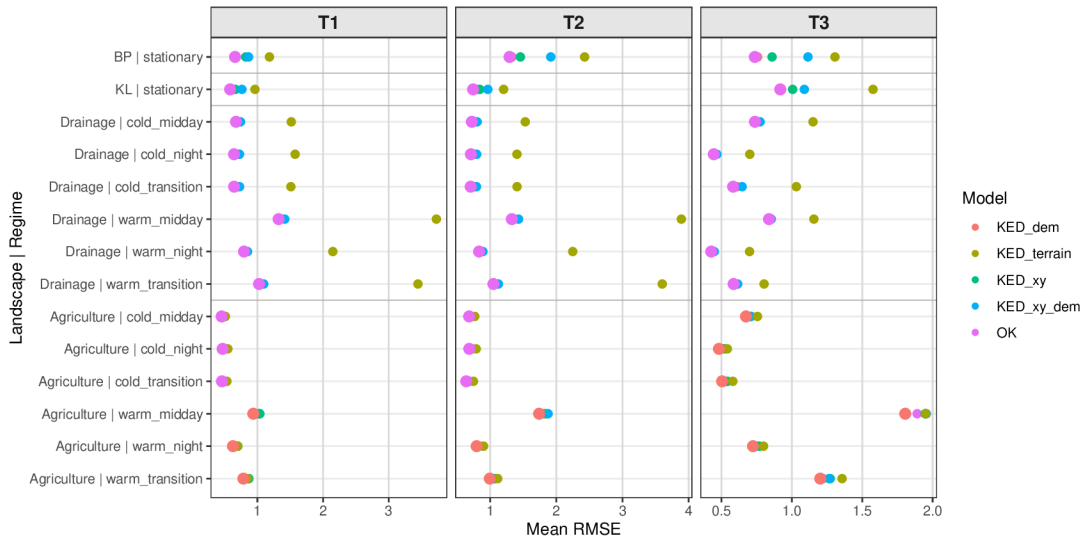

FIGURE 6.15: Comparison of RMSE across kriging models for all landscapes, regimes, and temperature measurements. The overall patterns are consistent with the MAE results: OK generally performs best for the stationary landscapes and the Drainage landscape, while KED using elevation provides improvements in the Agriculture landscape. Larger prediction errors are observed for the T3 measurement during warm regimes in Agriculture, reflecting stronger spatial variability during these conditions.

The MAE and RMSE comparisons reveal consistent patterns across landscapes and regimes. Prediction errors are generally lowest for the T3 temperature measurement during night regimes in the Drainage landscape, indicating relatively stable spatial temperature patterns under these conditions. In contrast, the largest errors occur for the T3 measurement in the Agriculture landscape, particularly during warm regimes, reflecting stronger spatial gradients and greater variability associated with daytime heating. These patterns are consistent with the regime-dependent spatial structures previously identified in the variogram analysis.

To illustrate the spatial patterns captured by the interpolation models, six multipanel representative temperature maps are presented for selected landscapes and thermal regimes. These examples include both KED and OK predictions, together with their corresponding prediction standard deviation maps. The figures highlight how spatial temperature structures vary across landscapes and regimes, while also demonstrating how prediction uncertainty increases with distance from the sensor network and in areas with limited observational support.

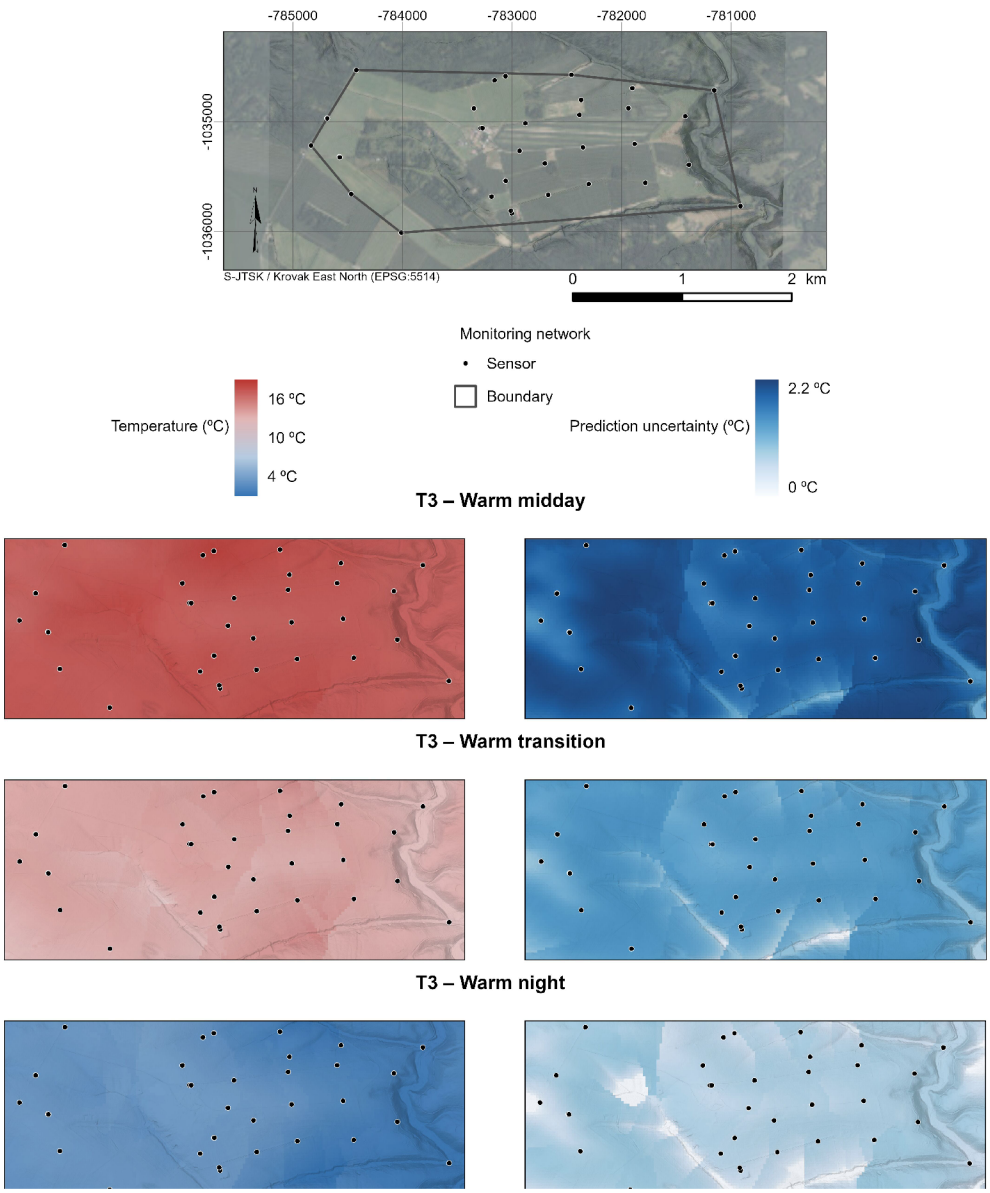

FIGURE 6.16: Spatial prediction of surface temperature (T3) for three representative regimes (warm-midday, warm-transition, and warm-night) in the Agriculture landscape using KED, with DEM-derived elevation as the drift variable. Left panels show predicted temperature, and right panels show the associated prediction standard deviation. A common color scale is used across all regimes for both temperature and uncertainty to allow direct comparison.

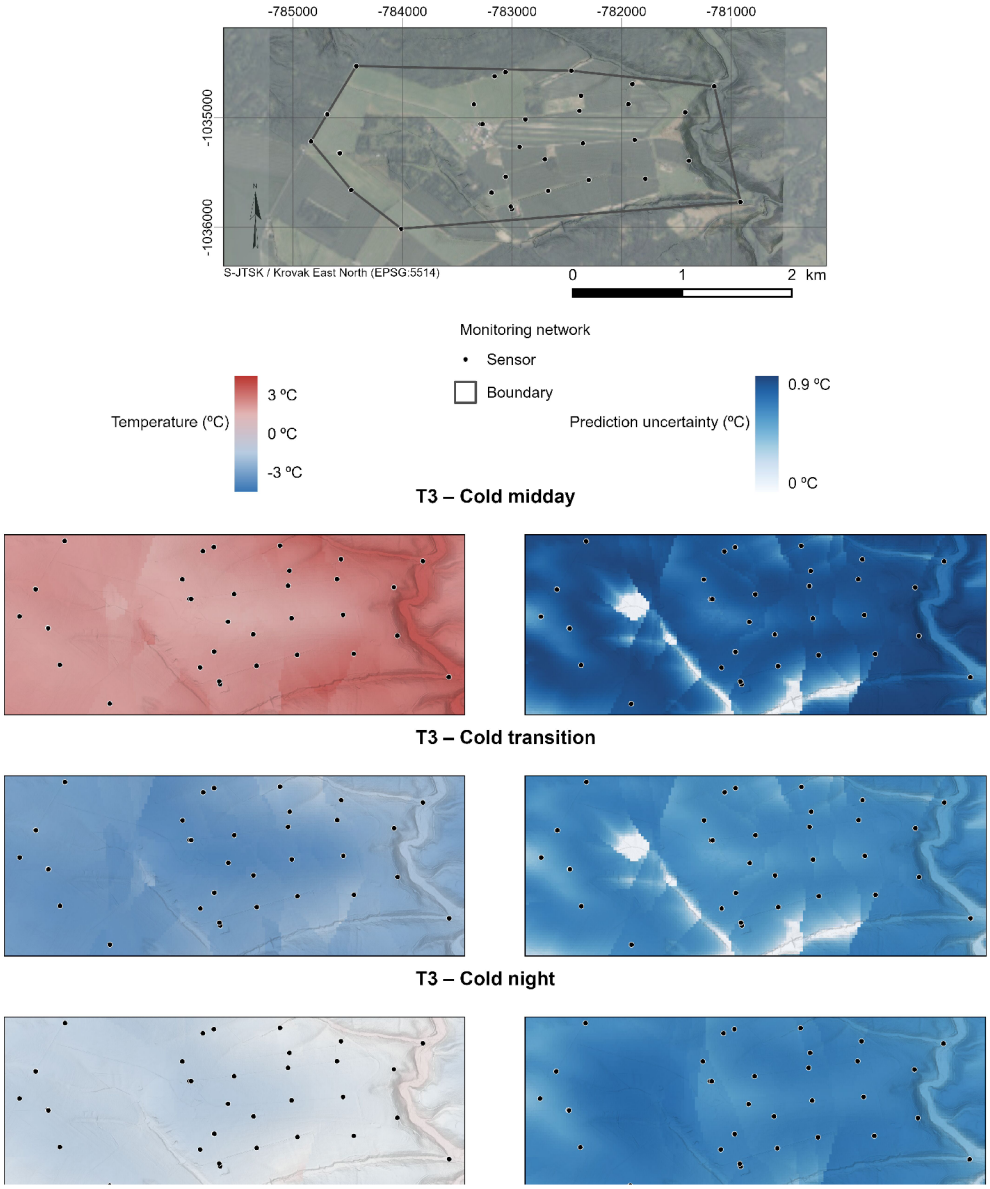

FIGURE 6.17: Spatial prediction of surface temperature (T3) for three representative regimes (cold-midday, cold-transition, and cold-night) in the Agriculture landscape using KED, with DEM-derived elevation as the drift variable. Graphical conventions follow those described in Figure 6.16.

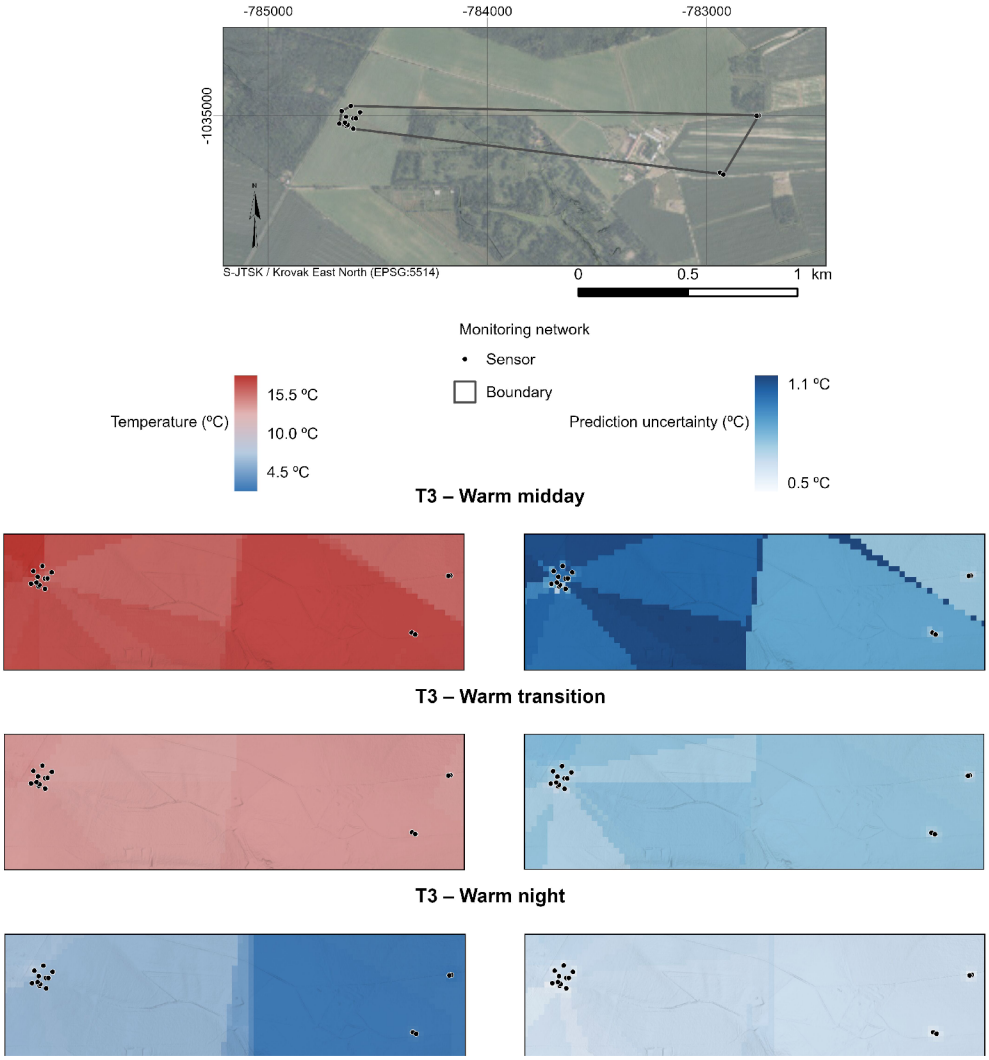

FIGURE 6.18: Spatial prediction of surface temperature (T3) for three representative regimes (warm-midday, warm-transition, and warm-night) in the Drainage landscape using OK. Graphical conventions follow those described in Figure 6.16.

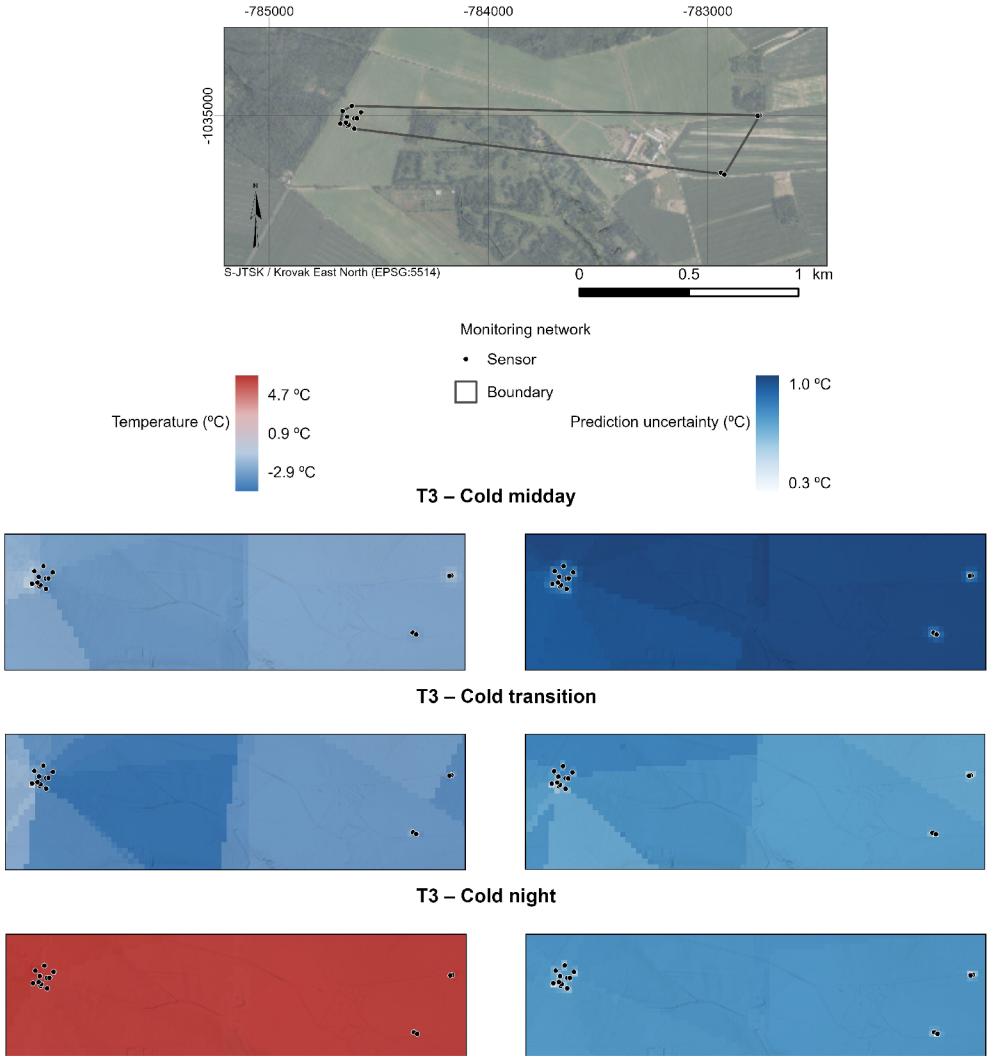

FIGURE 6.19: Spatial prediction of surface temperature (T3) for three representative regimes (cold-midday, cold-transition, and cold-night) in the Drainage landscape using OK. Graphical conventions follow those described in Figure 6.16.

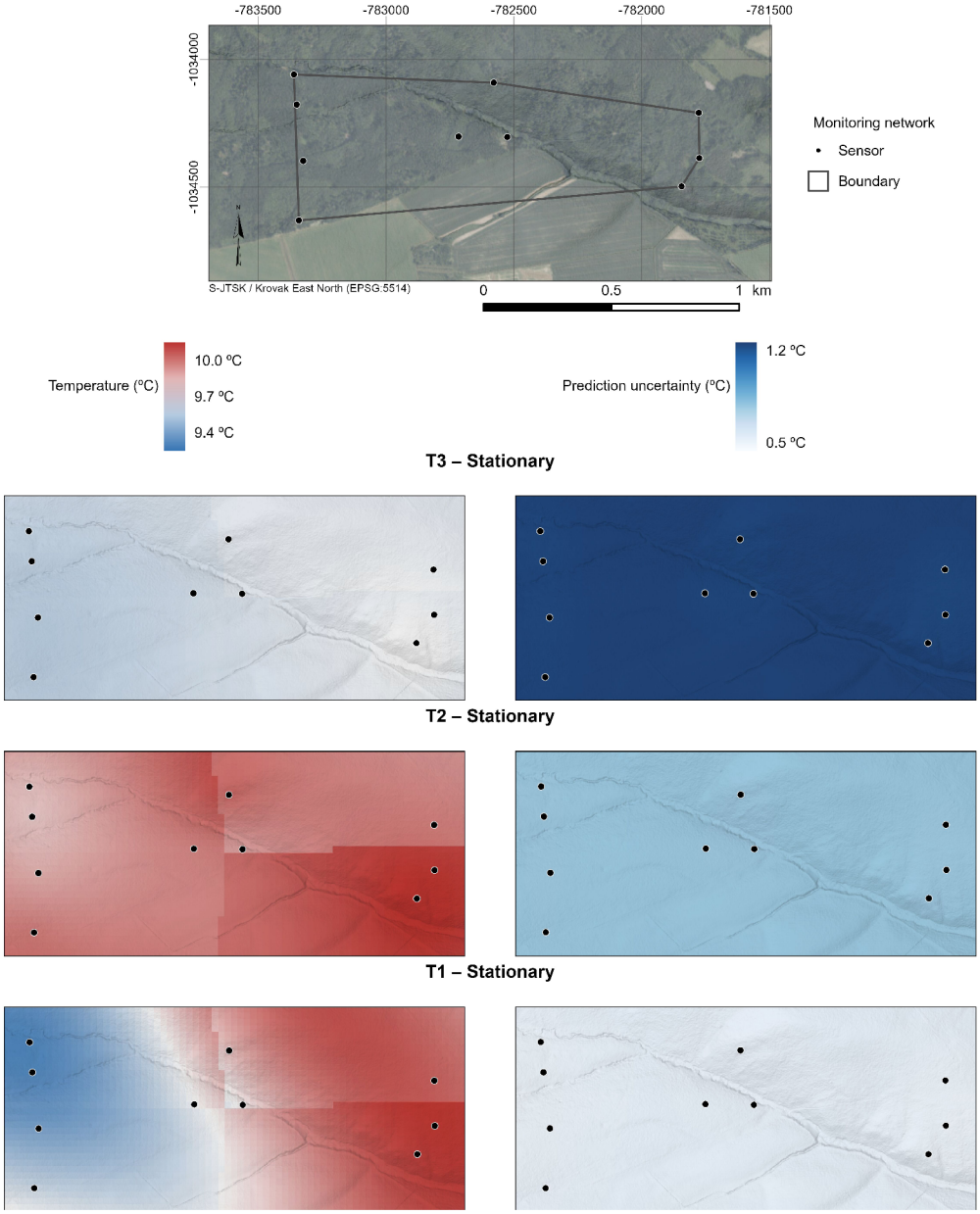

FIGURE 6.20: Spatial prediction of temperature at measurement levels (T1–T3) under stationary conditions in the KL landscape using OK. Graphical conventions follow those described in Figure 6.16.

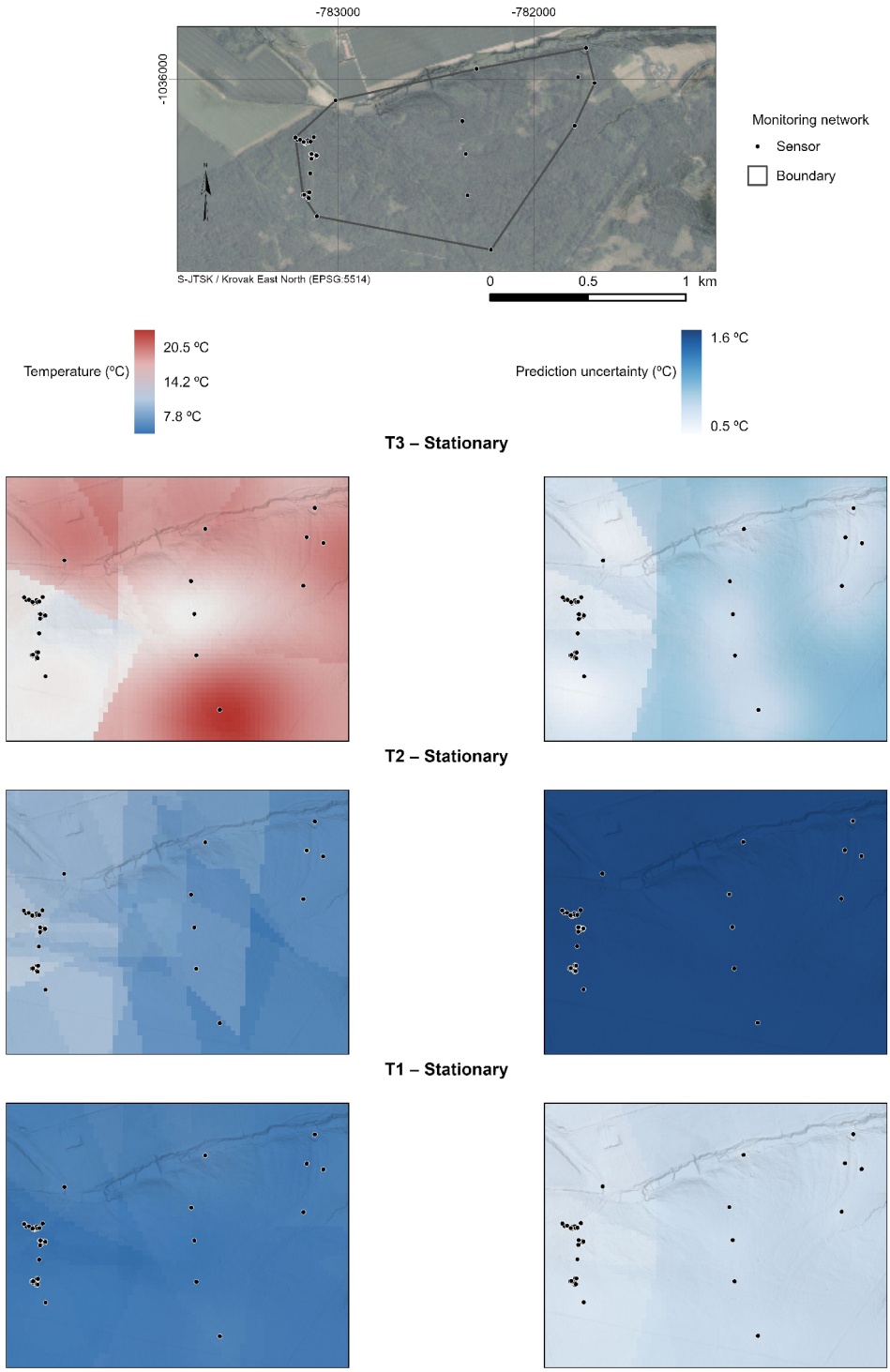

FIGURE 6.21: Spatial prediction of temperature at measurement levels (T1–T3) under stationary conditions in the BP landscape using OK. Graphical conventions follow those described in Figure 6.16.

The Drainage landscape exhibited a relatively short variogram range ($\approx$ 0.18 km), indicating that spatial correlation decays rapidly with distance. As a consequence, interpolation becomes weakly constrained outside the main sensor cluster, which is reflected in the increase in prediction standard deviation away from the observation network.

Overall, these representative examples illustrate how the interpolation models capture the dominant spatial temperature patterns while highlighting the influence of sensor distribution and spatial correlation on prediction uncertainty across landscapes and thermal regimes. Additional spatial interpolation maps for all landscape–regime–temperature combinations are provided in Appendix C.

CHAPTER 7

Discussion

The regime-specific approach adopted in this thesis aims to partially address the spatio-temporal variability of temperature fields. Similar challenges have been discussed in spatio-temporal geostatistical studies such as Rouhani and Myers (1990), who investigated temporal variability in spatial covariance structures. In the present study, temporal variability was addressed through regime-based variogram modeling, where timestamps belonging to similar seasonal and diurnal conditions were grouped together. After pooling the number of timestamps suggested by the bootstrap stability analysis (see Figures 6.6 and 6.7), the resulting regime variogram models for the Drainage and Agriculture landscapes exhibited relatively stable parameter estimates, particularly for the surface temperature variable (T3) (see Figures 6.10, 6.11, and 6.12). This behaviour is consistent with the interpretation proposed by Rouhani and Myers (1990), where temporal observations can be treated as multiple realizations of the same spatial process within a given temporal regime. Similar seasonal differences in spatial variability have also been reported by Holdaway (1996), who observed that semivariance values varied between warm and cold seasons in temperature interpolation studies.

Following the estimation of regime variograms, the selection of appropriate interpolation methods becomes an important methodological consideration. In this thesis, kriging with external drift (KED) was preferred over Regression kriging (RK) because it provides a more straightforward implementation, although both approaches are mathematically equivalent. Hengl et al. (2007) demonstrated this equivalence and provided a formal proof showing that RK and KED produce identical predictions under the same assumptions. However, they also recommend using at least ten observations per predictor to avoid overfitting. The monitoring networks across the landscapes of the Amálie pilot farm contain between 10 and 83 spatial locations (see Table 4.2). The methodological approach adopted in this study justified splitting the data into landscape units and pooling timestamps

based on regime stability in order to obtain reliable regime variograms, even in cases where the minimum number of spatial observations was relatively small.

Both Ordinary kriging (OK) and KED were applied for spatial interpolation. KED, incorporating elevation as an auxiliary variable, consistently outperformed OK in the Agriculture landscape. In particular, all warm-season regimes exhibited lower prediction errors when elevation was included in the model, while during the cold regime, the surface temperature variable (T3) also showed improved predictive performance (see Table 6.7). These findings are consistent with Hudson and Wackernagel (1994), who demonstrated that incorporating elevation as an external drift variable improves the interpolation of mean January temperatures in Scotland compared with OK.

One possible explanation is that the Agriculture landscape covers a relatively larger spatial domain, where elevation gradients exert a stronger influence on temperature variability. In contrast, in other landscapes where elevation variability is limited, OK provided more stable predictions. Seasonal differences may also contribute to the observed patterns. During warm regimes, deeper soil temperature (T1) and subsurface (soil) temperature (T2) may be influenced by elevation-related factors such as solar radiation exposure and terrain-driven microclimatic variability. During colder regimes, however, these effects are likely buffered by reduced solar radiation and more homogeneous atmospheric conditions, which may explain why T1 and T2 were better predicted using OK.

In landscapes where KED provided the best predictive performance, the average MAE varied substantially between regimes and temperature measurements. For warm regimes, MAE values ranged approximately from 0.48 to 1.39, whereas for cold regimes they ranged from 0.37 to 0.51. In particular, the lowest prediction error was observed for the Agriculture cold-night regime, which produced an RMSE of 0.50 (see Table 6.7). These results are consistent with findings reported by Willmott and Matsuura (1995), who showed that incorporating elevation information derived from a Digital Elevation Model (DEM) into temperature interpolation substantially reduces prediction errors compared with traditional interpolation approaches. The improvement in prediction accuracy observed in the present study, therefore, supports the idea that topographic variables can play an important role in explaining spatial variability in temperature fields.

The visualization of spatial prediction results in this thesis was designed to display both predicted temperature fields and their associated uncertainty through the kriging standard deviation (see Figures 6.16–6.21), as well as those provided in Appendix C. Using the standard deviation allows the uncertainty maps to be expressed in the same temperature units as the predictions, which facilitates interpretation and direct comparison between interpolation methods. A

similar visualization strategy was adopted by Hengl et al. (2004), who combined predictions and uncertainty when mapping soil variables using RK. Their study also demonstrated that incorporating auxiliary variables such as elevation, slope, curvature, compound topographic index, stream power index, and viewshed can substantially improve prediction accuracy. In the present work, only terrain variables derived from a DEM were used as auxiliary predictors, as described in Section 5.2; however, many other environmental covariates could potentially be incorporated in future studies to further improve spatial prediction performance.

CHAPTER 8

Conclusions

This thesis developed and evaluated a geostatistical workflow for reconstructing high-resolution spatial temperature fields from a 15-minute monitoring network at the Amálie pilot farm in the Czech Republic. Three temperature measurements representing different microclimatic layers—surface, subsurface, and deeper soil—were interpolated independently in order to characterize the spatial structure of the local microclimate.

A central methodological component of this study was the identification of an appropriate variogram modeling strategy. The study area exhibits considerable intra-variability due to the coexistence of different landscapes and research activities within the farm. Addressing this heterogeneity required the definition of landscape-specific spatial domains and, in some cases, the application of regime-based stratification to account for seasonal and diurnal variability in the temperature field.

The analysis compared two empirical variogram estimators and four theoretical variogram models in order to characterize the spatial covariance structure of the temperature observations. Two geostatistical interpolation methods—Ordinary kriging (OK) and kriging with external drift (KED)—were evaluated for reconstructing spatial temperature fields. Elevation and other terrain attributes derived from a Digital Elevation Model (DEM) were incorporated as auxiliary predictors to assess their contribution to interpolation accuracy.

The results demonstrate that the optimal interpolation approach varies among landscape units. In some areas, the incorporation of topographic predictors improved prediction accuracy, while in others OK provided more stable results. These findings highlight the importance of adapting variogram modeling and interpolation strategies to the spatial characteristics of each landscape.

The relatively small spatial extent of the study area and the uneven distribution of monitoring stations represent additional challenges for variogram estimation and

spatial prediction. Nevertheless, the proposed methodological framework provides a systematic approach for evaluating variogram models, determining the minimum number of timestamps required to obtain stable variogram parameter estimates through bootstrap resampling, selecting optimal neighborhood configurations using leave-one-out cross-validation (LOOCV), and reconstructing high-frequency temperature fields.

Future work could explore alternative interpolation approaches, including non-geostatistical, hybrid, or machine-learning-based methods, as well as the integration of additional environmental covariates to further improve prediction performance.

Bibliography

Akaike, H., (1974): A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716–723. DOI: https://doi.org/10.1109/TAC.1974.1100705.

Appelhans, T., Mwangomo, E., Hardy, D.R., Hemp, A., Nauss, T., (2015): Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. Spatial Statistics, 14, 91–113. DOI: https://doi.org/10.1016/j.spasta.2015.05.008.

Bilonick, R.A., Nichols, D.G., (1983): Temporal variations in acid precipitation over New York State—What the 1965–1979 USGS data reveal. Atmospheric Environment, 17(6), 1063–1072. DOI: https://doi.org/10.1016/0004-6981(83)90329-3.

Bramer, I., Anderson, B.J., Bennie, J., Bladon, A.J., De Frenne, P., Hemming, D., Hill, R.A., Kearney, M.R., Körner, C., Korstjens, A.H., Lenoir, J., Maclean, I.M., Marsh, C.D., Morecroft, M.D., Ohlemüller, R., Slater, H.D., Suggitt, A.J., Zellweger, F., Gillingham, P.K. (2018): Advances in monitoring and modelling climate at ecologically relevant scales. *Advances in Ecological Research*. Vol. 58. Academic Press, 101–161. DOI: https://doi.org/10.1016/bs.aecr.2017.12.005.

CESNET, (2024): MetaCentrum: large infrastructure for e-science in the Czech Republic (on-line), available at <https://www.metacentrum.cz>.

Chalupecká, H., Jačka, L., Kuželková, M., Kovář, M., Rous, J., Abbasizadeh, H., Komárek, J., Preininger, D., Lojka, B., (2026): The early-stage effect of alley cropping system on the soil moisture and temperature of drought-prone silt-loam soil in European temperate transitional climate. Agroforestry Systems, 100(3), 78. DOI: https://doi.org/10.1007/s10457-026-01436-4.

Cressie, N., (1985): Fitting variogram models by weighted least squares. Mathematical Geology, 17(5), 563–586. DOI: https://doi.org/10.1007/BF01032109.

Cressie, N., (1990): The origins of kriging. Mathematical Geology, 22(3), 239–252. DOI: https://doi.org/10.1007/BF00889887.

Cressie, N., Hawkins, D.M., (1980): Robust estimation of the variogram. Mathematical Geology, 12(2), 115–125. DOI: https://doi.org/10.1007/BF01035243.

ČÚZK, (2023): Digital relief model of the Czech Republic – 5th generation (DMR 5G) (on-line), Accessed: 13 March 2026, available at <https://geoportal.cuzk.cz>.

Czech Hydrometeorological Institute, (2023): Historical meteorological data – monthly observation summaries (on-line), Accessed: 13 March 2026, available at <https://intranet.chmi.cz/historicka-data/pocasi/mesicni-data/mesicni-prehledy-pozorovani>.

Czech University of Life Sciences Prague, (2024): Smart Landscape (on-line), Accessed: 13 March 2026, available at <https://cvpk.czu.cz/en/r-17791-smart-landscape>.

Genton, M.G., (1998): Highly robust variogram estimation. Mathematical Geology, 30(2), 213–221. DOI: https://doi.org/10.1023/A:1021728614555.

Goovaerts, P., (1997): Geostatistics for natural resources evaluation. Oxford University Press, New York. ISBN: ISBN: 9780195115383.

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., (2007): About regression-kriging: From equations to case studies. Computers & Geosciences, 33, 1301–1315. DOI: https://doi.org/10.1016/j.cageo.2007.05.001.

Hengl, T., Heuvelink, G.B.M., Stein, A., (2004): A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma, 120, 75–93. DOI: https://doi.org/10.1016/j.geoderma.2003.08.018.

Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G.B., Gräler, B., (2018): Random Forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ, 6, e5518. DOI: https://doi.org/10.7717/peerj.5518.

Holdaway, M.R., (1996): Spatial modeling and interpolation of monthly temperature using kriging. Climate Research, 6(3), 215–225. DOI: https://doi.org/10.3354/cr006215.

Huang, H.-C., Chen, C.-S., (2007): Optimal geostatistical model selection. Journal of the American Statistical Association, 102(479), 1009–1024. DOI: https://doi.org/10.1198/016214507000000491.

Hudson, G., Wackernagel, H., (1994): Mapping temperature using kriging with external drift: Theory and an example from Scotland. International Journal of Climatology, 14(1), 77–91. DOI: https://doi.org/10.1002/joc.3370140107.

Jian, X., Olea, R.A., Yu, Y., (1996): Semivariogram modeling by weighted least squares. Computers & Geosciences, 22(4), 387–397. DOI: https://doi.org/10.1016/0098-3004(95)00095-X.

Kemppinen, J., Lembrechts, J.J., Van Meerbeek, K., Carnicer, J., Chardon, N.I., Kardol, P., Lenoir, J., Liu, D., Maclean, I.M., et al., (2024): Microclimate, an important part of ecology and biogeography. Global Ecology and Biogeography. DOI: https://doi.org/10.1111/geb.13834.

Kerry, R., Oliver, M.A., (2007): Determining the effect of asymmetric data on the variogram. II. Outliers. Computers & Geosciences, 33(10), 1233–1240. DOI: https://doi.org/10.1016/j.cageo.2006.11.009.

Kuželková, M., Jačka, L., Kovář, M., Hradílek, V., Máca, P., (2024): Tree trait-mediated differences in soil moisture regimes: a comparative study of beech, spruce, and larch in a drought-prone area of Central Europe. European Journal of Forest Research, 143, 319–332. DOI: https://doi.org/10.1007/s10342-023-01628-y.

Lam, N.S.-N., (1983): Spatial interpolation methods: A review. The American Cartographer, 10(2), 129–150. DOI: https://doi.org/10.1559/152304083783914958.

Lark, R.M., (2000): Estimating variograms of soil properties by the method-of-moments and maximum likelihood. European Journal of Soil Science, 51(4), 717–728. DOI: https://doi.org/10.1046/j.1365-2389.2000.00345.x.

Li, J., (2019): A critical review of spatial predictive modeling process in environmental sciences with reproducible examples in R. Applied Sciences, 9(10), 2048. DOI: https://doi.org/10.3390/app9102048.

Li, J., Heap, A.D., (2011): A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. Ecological Informatics, 6(3–4), 228–241. DOI: https://doi.org/10.1016/j.ecoinf.2010.12.003.

Li, J., Heap, A.D., (2014): Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling & Software, 53, 173–189. DOI: https://doi.org/10.1016/j.envsoft.2013.12.008.

Matheron, G., (1962): Traité de géostatistique appliquée. Éditions Technip, Paris.

Meyer, H., Reudenbach, C., Wöllauer, S., Nauss, T., (2019): Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction. Ecological Modelling, 411, 108815. DOI: https://doi.org/10.1016/j.ecolmodel.2019.108815.

Oliver, M.A., Webster, R., (2014): A tutorial guide to geostatistics: Computing and modelling variograms and kriging. Catena, 113, 56–69. DOI: https://doi.org/10.1016/j.catena.2013.09.006.

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., et al., (2020): Spatial validation reveals poor predictive performance of large-scale ecological mapping models. Nature Communications, 11(1), 4540. DOI: https://doi.org/10.1038/s41467-020-18321-y.

R Core Team (2024): *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., et al., (2017): Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913–929. DOI: https://doi.org/10.1111/ecog.02881.

Rouhani, S., Myers, D.E., (1990): Problems in space-time kriging of geohydrological data. Mathematical Geology, 22(5), 611–623. DOI: https://doi.org/10.1007/BF00890600.

Webster, R., Oliver, M.A., (2007): Geostatistics for environmental scientists. John Wiley & Sons, Chichester.

Wild, J., Kopecký, M., Macek, M., Šanda, M., Jankovec, J., Haase, T., (2019): Climate at ecologically relevant scales: A new temperature and soil moisture logger for long-term microclimate measurement. Agricultural and Forest Meteorology, 268, 40–47. DOI: https://doi.org/10.1016/j.agrformet.2019.01.005.

Willmott, C.J., Matsuura, K., (1995): Smart interpolation of annually averaged air temperature in the United States. Journal of Applied Meteorology, 34(12), 2577–2586. DOI: https://doi.org/10.1175/1520-0450(1995)034<2577:SIOAAA>2.0.CO;2.

Xie, Y., Eftelioglu, E., Ali, R., Tang, X., Li, Y., Doshi, R., Shekhar, S., (2017): Transdisciplinary foundations of geospatial data science. ISPRS International Journal of Geo-Information, 6(12), 395. DOI: https://doi.org/10.3390/ijgi6120395.

Zhang, M., Yu, D., Li, Y., Zhao, L. (2022): Deep geometric neural network for spatial interpolation. *Proceedings of the 30th International Conference on Advances in Geographic Information Systems.* Seattle, Washington: ACM, 1–4. DOI: https://doi.org/10.1145/3557915.3561008.