



Česká zemědělská univerzita v Praze

Fakulta životního
prostředí

Statistická korekce denních srážkových úhrnů z klimatických modelů

Disertační práce

2016

Autor práce:
Jan Hnilica

Školitel:
doc. Ing. Zdeněk Chára, CSc.

Prohlášení

Prohlašuji, že jsem tuto disertační práci vypracoval samostatně pod odborným vedením školitele a uvedl jsem veškeré použité literární prameny. Tištěná a elektronická verze práce se doslovně shodují.

.....
Jan Hnilica

Poděkování

Rád bych na tomto místě poděkoval vedoucímu práce doc. Ing. Zdeňku Chárovi, CSc. za cenné připomínky a trpělivý přístup. Dále bych chtěl poděkovat Soně, Vojtovi a Zuzance.

Abstrakt

Odhad budoucích klimatických změn a zhodnocení jejich potenciálního dopadu představují jeden z klíčových úkolů současné vědy. Nezastupitelnou úlohu při predikci vývoje klimatu sehrávají regionální klimatické modely, které doplňují velkoprostorové globální simulace a umožňují vykreslit meteorologické jevy středních měřítek. Jejich prostorové rozlišení je ale příliš hrubé pro kvalitní predikci meteorologických veličin v lokálním měřítku, která je při posuzování dopadu klimatické změny nezbytná. Výstupy z regionálních modelů, zejména teploty a srážkové úhrny, jsou ve svých statistických ukazatelích odchyleny od měřených hodnot a před použitím v dopadových studiích jsou obvykle podrobovány statistické korekci chyb.

Korekce chyb je v oboru klimatologie aktuálně diskutovaným tématem a korekční metody prodělaly značný rozvoj od jednoduché aditivní metody až po sofistikovanější přístupy, jakým je např. kvantilové mapování. Bez ohledu na tento vývoj jsou však aktuálně užívané metody spojeny s řadou problémů, které vnáší nejistoty do navazujících studií.

Tato práce se soustředí na dva problematické aspekty spjaté s korekcemi denních srážkových úhrnů. Prvním problémem spočívá v tom, že korekce jsou aplikovány na datové soubory s odlišnými statistickými vlastnostmi, než mají data, na kterých byly kalibrovány. V rámci práce jsou odvozeny a validovány metody vykazující vůči nestacionárním podmínkám zvýšenou odolnost. Druhým řešeným problémem je korekce prostorových vztahů mezi srážkovými záznamy z různých lokalit. Práce obsahuje odvození a validaci procedury, korigující kompletní závislostní (tj. korelační a kovarianční) vazby v mnohorozměrných srážkových datech.

Klíčová slova: regionální klimatický model, korekce chyb, nestacionarita, korelace, kovariance

Abstract

Climate change prediction and evaluation of its impact currently represent one of the key challenges for the science community. Regional climate models (RCM) have been recently established as a main source of the data for climate change assessment studies. Nevertheless, RCM outputs suffer from systematic errors caused primarily by their low spatial resolution and cannot be used directly without any form of bias correction.

The bias correction is an actual topic in climatology and several correction methods were developed, ranging from the simple additive method to more advanced approaches (e.g. quantile mapping). However, despite this progress, the bias correction methods suffer from several difficulties, which bring another source of uncertainty into the climate change impact assessment studies.

This thesis is focused on two problematic points connected with the bias correction of daily precipitation data. The first one is a non-stationarity between calibration and application periods. New correction methods are developed, showing an increased resistance to non-stationary conditions. The second problem is related to the correction of a dependence (i.e. correlation and covariance) structure of multivariate precipitation data. A new procedure is proposed, correcting the complete dependence structure of the model data. All newly introduced methods are validated using measured and RCM-simulated data; the validation demonstrates their suitable applicability.

Keywords: regional climate model, bias correction, non-stationarity, correlation, covariance

Obsah

Úvod.....	8
LITERÁRNÍ REŠERŠE	10
1 Globální klimatické modely	10
1.1 Základní komponenty globálních modelů	10
1.2 Validace globálních modelů	12
1.3 Typy experimentů	13
1.4 Nejistoty klimatických studií	14
1.5 Vývoj klimatických modelů	16
1.6 Systematické chyby globálních modelů a metody regionalizace	17
2 Regionální klimatické modely.....	19
2.1 Faktory určující kvalitu regionálních modelů	19
2.2 Modifikace základního konceptu vnoření	21
2.3 Problematické aspekty konceptu vnoření.....	23
3 Korekce chyb	25
3.1 Úvod	25
3.2 Přehled nejčastěji používaných metod	29
3.3 Ostatní (nepřímé) korekční metody.....	36
3.4 Empirické porovnání metod	37
3.5 Vyjádření pravděpodobnostního rozdělení srážek	38
3.6 Závěrečné poznámky.....	41
PRAKTICKÁ ČÁST.....	43
4 Použitá data.....	43
5 Lineární povaha korekce a podobnost metod	46
6 Nové lineární metody.....	53
6.1 Úvod	53
6.2 Odvození korekčních metod.....	54
6.3 Validace a porovnání s ostatními metodami	56
6.4 Výsledky a diskuze.....	57
6.5 Dodatek - QM pod vlivem nestacionarity	65
7 Korekce závislostní struktury	68
7.1 Úvod	68
7.2 Odvození metody	68
7.3 Standardní způsob korekce.....	73
7.4 Porovnání metod PCC a QM.....	74
7.5 Kombinace metod PCC a QM.....	77

8 Souhrn a závěry	84
Seznam literatury	86
Seznam obrázků.....	95
Seznam tabulek	96
Seznam publikací.....	97

Úvod

Regionální klimatické modely se v posledním desetiletí staly jedním z hlavních nástrojů pro výzkum klimatického systému. Podrobnější prostorové a časové rozlišení těchto modelů ve srovnání s modely globálními přináší dodatečné informace ke globálním klimatickým projekcím. Hlavní efekt regionálních modelů spočívá v lepším vystižení meteorologických jevů středních měřítek, které formují podobu regionálního klimatu.

Jedním z hlavních výzkumných směrů v současné klimatologii je predikce potenciální klimatické změny zapříčiněné zejména antropogenní činností. Projekce regionálních klimatických modelů jsou hlavním zdrojem dat pro tento typ studií. Nicméně ani podrobnější prostorové rozlišení regionálních modelů nevede k realistické simulaci meteorologických veličin v lokálním měřítku, což se týká zejména teplot a srážkových úhrnů. Tyto veličiny simulované regionálním modelem vykazují systematické odchylky od pozorovaných hodnot a nemohou být přímo využity jako vstup do navazujících hydrologických, ekologických či zemědělských studií.

Z těchto důvodů bylo vyvinuto množství metod určených ke korekci modelových výstupů. Tyto korekční metody jsou hojně využívány v klimatologických studiích, zároveň jsou však předmětem značné kritiky ze strany části odborné veřejnosti. Primárním důvodem je jejich statistická povaha. Hlavní devizou regionálních modelů při získávání podrobnějších informací o klimatickém systému je především to, že modely pracují v souladu s fyzikálními zákony determinujícími vývoj atmosférického a oceánského proudění. Aplikací statistických metod na konečné výstupy modelu je do procesu vnesena nehomogenita s těmito zákony a konečné výsledky mohou být nekonzistentní s reálně dosažitelnými daty, přestože některé statistické parametry korigovaných souborů vykazují s realitou lepší shodu než neupravené modelové výstupy.

Statistická povaha korekcí je zásadním problémem a je také obtížně odstranitelná. Nicméně i v rámci statistických metod existuje řada dílčích problémů, které je možné řešit, umožnit tak dosažení realističtějších výsledků korekce a tím také zvýšit hodnověrnost dopadových studií. Disertační práce se soustředí na dva z nich. Prvním problémem spočívá v tom, že korekce jsou aplikovány na datové soubory s odlišnými statistickými vlastnostmi, než mají data, na kterých byly kalibrovány. Tato nestacionarita vnáší nejistotu do výsledků navazujících studií. Druhým řešeným problémem je fakt, že korekce jsou obvykle kalibrovány a aplikovány odděleně pro jednotlivé prostorové body (meteorologické stanice, modelové grid-boxy). Takový postup však nijak neopravuje vzájemné korelační a kovarianční vazby mezi jednotlivými datovými soubory, což může vést k chybám při použití korigovaných dat např. v hydrologických studiích.

Cílem práce je odvodit a validovat nové metody korekce umožňující řešit dva výše popsané problémy, nebo alespoň dosahovat realističtějších výsledků v porovnání s běžně používanými metodami.

V první části práce je provedena stručná literární rešerše problematiky globálního a regionálního klimatického modelování a dále podrobnější rešerše metod používaných při korekci modelových srážkových úhrnů.

Praktická část začíná představením datových podkladů, používaných pro validaci metod, a dále je rozdělena do tří víceméně nezávislých studií. První z nich navazuje na literární rešerši a poukazuje na lineární povahu a vzájemnou podobnost jednotlivých korekčních metod. Lineární povaha je uvedena teoretickým rozбором a demonstrována na měřených a

modelových datech. Další dvě studie pak těchto poznatků využívají. V první z nich jsou odvozeny nové metody korekce, vykazující zvýšenou odolnost proti nestacionárním podmínkám. Druhá se pak soustředí na prostorové vazby mezi datovými soubory a obsahuje odvození procedury, která koriguje závislostní (tj. korelační a kovarianční) struktury modelových dat. Závěrečná kapitola obsahuje stručný souhrn disertační práce a několik poznámek k nově odvozeným metodám.

LITERÁRNÍ REŠERŠE

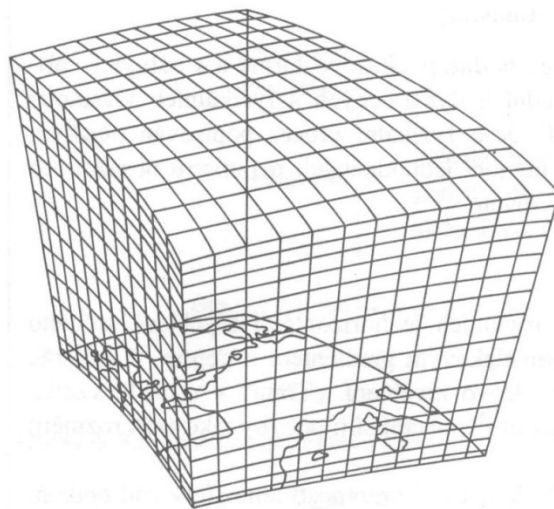
1 Globální klimatické modely

Klimatické modely jsou základním nástrojem a zdrojem dat pro výzkum klimatu. Globální klimatický model je počítačový program simulující vývoj klimatického systému Země ze zadaného počátečního stavu. Jde o systém interagujících komponent, z nichž každá simuluje část klimatického systému, přičemž struktura modelu zhruba odpovídá struktuře systému samotného. Čtyřmi základními komponentami jsou model cirkulace atmosféry, model cirkulace oceánu, model mořského ledu a model zemského povrchu. Od devadesátých let minulého století jsou klimatické modely doplňovány o modely chemické, zachycující vliv chemických sloučenin, které mají efekt na energetickou bilanci klimatického systému, především oxidu uhličitého a aerosolů.

V klimatologických textech bývá frekventovaně používána zkratka GCM, která může v závislosti na kontextu označovat dvě odlišné věci. Zprvce to může být „global climate model“, tedy kompletní globální klimatický model jako souhrn všech jeho komponent. Zadruhé se může jednat o „general circulation model“. Tímto termínem se označují první dvě výše uvedené komponenty klimatických modelů, tedy model cirkulace atmosféry a model cirkulace oceánu. V tomto textu bude zkratka GCM užívána k označení kompletního globálního klimatického modelu. Jednotlivé komponenty globálního modelu jsou stručně představeny v následujících odstavcích.

1.1 Základní komponenty globálních modelů

Model cirkulace atmosféry je založen na nelineární soustavě diferenciálních rovnic, která popisuje globální atmosférické proudění. Tyto rovnice bývají označovány jako „primitive equations“ a jejich výsledná soustava nemá známé analytické řešení. Z tohoto důvodu je základní složkou modelu atmosféry tzv. dynamické jádro, které zajišťuje numerickou integraci časového vývoje soustavy. Numerické výpočty jsou realizovány buď metodou konečných diferencí, nebo metodami spektrálními.



Obr. 1.1 Schématická ilustrace výpočetní sítě modelu cirkulace atmosféry (převzato z Houghton, 1998).

Na dynamické jádro je napojeno množství dodatečných výpočtů, jejichž cílem je zachytit výsledný efekt fyzikálních dějů, které buď nejsou obsaženy v rovnicích dynamického jádra, nebo nastávají v menším prostorovém či časovém měřítku, než je výpočetní krok modelu. Tyto dodatečné výpočty bývají označovány jako tzv. parametrizace. Vstupními údaji pro parametrizace jsou hodnoty veličin počítané dynamickým jádrem, typicky jde o teplotu, vlhkost vzduchu a rychlost proudění. Mezi hlavní parametrizované procesy patří vznik srážek, formace a rozpouštění oblaků typu cirrus a stratus a konvektivní proudy formující oblaka typu cumulus. Dále jsou parametrizace aplikovány na absorpci a emisi elektromagnetického záření molekulami vzduchu a aerosoly a také na konverze mezi jednotlivými typy vodních částic (ledové krystaly, sníh, voda). Pro zhodnocení kvality používaných parametrizací bylo uskutečněno několik projektů, jejich výčet spolu s odkazy na příslušnou literaturu uvádí Bader et al. (2008).

Výpočetní oblast modelů zahrnuje atmosféru kolem celé zeměkoule, ve vertikálním směru se v závislosti na konkrétním modelu jedná o 20 až 30 km vrstvu. Předmětem simulací je tedy troposféra a spodní část stratosféry, ve kterých je obsaženo více než 95 % atmosférické hmoty a prakticky všechna atmosférická voda. Rozlišení modelů je limitováno dostupným výpočetním výkonem. Horizontální rozlišení se pohybuje v jednotkách stupňů zeměpisné šířky, což odpovídá vzdálenosti uzlů výpočetní sítě 100 a více km. Obvyklé vertikální rozlišení je zhruba 1 km. Časový krok modelu je z důvodu zachování numerické stability výpočtu omezen jeho horizontálním rozlišením a pohybuje se v desítkách minut.

Struktura *modelu cirkulace oceánu* je podobná struktuře modelu atmosféry. Vývoj oceánského proudění je stejně jako v případě atmosféry popsán sadou diferenciálních rovnic popisujících dynamiku kontinua, proto je základní částí oceánského modelu opět dynamické jádro, na které je napojena soustava parametrizací. Oceán má členitější hranice, než je tomu u atmosféry, proto je výpočetní síť oceánských modelů jemnější se vzdáleností uzlových bodů menší než 100 km. V oblastech okolo rovníku, do cca 5° zeměpisné šířky, je výpočetní síť dále zjemněna. Stejně jako u modelu atmosféry nalezneme i v oceánském modelu disproporci mezi horizontálním a vertikálním rozlišením.

Nejdůležitějšími parametrizacemi v případě oceánského modelu jsou vliv turbulentního míchání vody v malých měřítcích (m) a vliv vírů středních měřítek (km). Oba tyto jevy přispívají ke vstřebávání tepla a následné stratifikaci oceánské vody, která má vliv na cirkulaci oceánu v dlouhodobém horizontu desítek let. Tyto parametrizace výrazně přispívají k celkovým nejistotám spojeným s klimatickým modelováním a také výrazně odlišují globální modely mezi sebou (Schopf et al., 2002). Další parametrizace se týkají vlivu turbulence generované kontaktem oceánské vody se dnem a míchání vody vlivem přílivu (Lee et al., 2006). Spojením atmosférického a oceánského modelu vzniká tzv. „coupled atmosphere-ocean general circulation model“ (AOGCM).

Děje probíhající na *zemském povrchu*, obzvláště výměna tepla a vodní páry s atmosférou, zásadně ovlivňují energetickou bilanci klimatického systému. Vliv těchto dějů, a s nimi souvisejících zpětných vazeb, na variabilitu teplot v Evropě dokumentují např. Seneviratne et al. (2006). Pevný povrch je navíc velmi heterogenní a modelování těchto dějů je tudíž značně náročné. Model pevného povrchu zahrnuje (parametrizuje) všechny hlavní procesy kontrolující výměnu vody a tepla mezi zemským povrchem a atmosférou. Mezi tyto procesy patří evaporace z vody a z půdy, transpirace rostlin, intercepce, sublimace, tání sněhu,

povrchový odtok, vsakování vody a emise dlouhovlnného záření. Obecně platí, že validace modelů povrchu je problematická, zejména kvůli nedostatku pozemních měření evapotranspirace a toků tepla. Probíhá tedy obvykle s využitím odhadů těchto jevů získaných pomocí měření jiných veličin (Bader et al., 2008). Rostoucí roli v procesech kalibrace a validace modelů sehrávají satelitní měření, která umožňují získat prostorově distribuované odhady tepelných toků na velkých plochách. Při modelování energetické bilance systému sehrává podstatnou roli vliv sněhu a zemského ledu kvůli vysoké míře odrazivosti dopadajícího elektromagnetického záření. Modely reprezentují sníh buď jako homogenní vrstvu s prostorově uniformním nárůstem a ubýváním tloušťky, nebo berou v úvahu distribuci hloubky sněhu v subgridovém měřítku (Liston, 2004). Pevninské ledovce jsou obvykle reprezentovány jako prostorově stacionární, nicméně náhled na tuto problematiku se mění především vlivem satelitních pozorování, která dokládají jejich zjevnou dynamiku (Rignot a Kanagaratnam, 2006). Modelování dynamiky ledovců je však podmíněno efektivnější parametrizací povrchových a subglaciálních hydrologických procesů. Z pohledu klasické hydrologie obsahují modely zemského povrchu jednoduché velkoplošné srážko-odtokové modely, realizované jako soustavy lineárních nádrží charakterizovaných různými dobami zdržení vody, tzv. runoff routing models (Arora a Boer, 1999, Ducharne et al., 2003). Modely zemského povrchu zatím obvykle neobsahují popis proudění podzemní vody.

Model mořského ledu simuluje pohyb ledových vrstev, aktuální rozsah a tloušťku ledu a transport tepla a soli v ledové hmotě. Led samotný je v modelu reprezentován jako kontinuum se specifickými reologickými vlastnostmi. Výčet používaných reologických modelů podává Bader et al. (2008). Z termodynamického hlediska je led zjednodušeně schematizován jako několik ledových vrstev se specifickými hodnotami tepelné vodivosti. Obsah soli je parametrizován. Hodnoty albeda jsou odhadovány semiempirickým vztahem jako funkce tloušťky ledu a sněhu, teploty povrchové vrstvy a dalších faktorů (Curry et al. 1995).

1.2 Validace globálních modelů

Cílem vývoje globálních klimatických modelů je postupné snižování empiricismu obsaženém v jejich struktuře a jeho nahrazení fyzikálně realistickým popisem modelovaných procesů. Z tohoto důvodu je validace modelů obvykle zaměřena na kontrolu těchto dílčích procesů spíše než na hodnocení celkových energetických bilancí.

Validace bývá prováděna několika způsoby. Jedním z nich je kontrola schopnosti modelu spontánně generovat vznik a frekvenci výskytu některých klimatických jevů, například ENSO (El Niño Southern Oscillation), nebo ověření schopnosti modelu reprodukovat významné oceánské proudy, jako je např. Atlantická termohalinní cirkulace. Vzhledem k tomu, že GCM jsou konstruovány pro obecné použití v klimatologii bez primárního zaměření na určitou oblast výzkumu, nejsou důsledkem těchto testů optimalizace modelu pro predikci daného jevu, neboť tyto ad hoc zásahy obvykle vedou ke zhoršení schopnosti předikovat jevy jiné (Bader et al., 2008).

K validaci slouží také simulace, při kterých je model využit k reprodukci klimatu 20. století. Modelové výstupy jsou pak přímo porovnávány s měřenými záznamy, popř. s upravenými a prostorově interpolovanými měřenými daty, tzv. re-analýzami, viz např. Uppala et al. (2005). Tato srovnání jsou samozřejmě prováděna s ohledem na to, že cílem modelování klimatu není predikce přesného průběhu počasí, ale vystihnoutí jeho dlouhodobých charakteristik a

trendů. Objektem porovnání jsou tedy statistické ukazatele datových souborů, obvykle obecné a centrální momenty nebo celkové pravděpodobnostní rozdělení, popřípadě vývoj těchto charakteristik v čase.

Dalším nástrojem vhodným pro posuzování kvality klimatických modelů jsou satelity. Příkladem může být projekt ISCCP – International Satellite Cloud Climatology Project (Rossow a Schiffer, 1991), v rámci kterého jsou od roku 1983 zpracovávána satelitní měření poskytující přehled o distribuci a kvalitě oblačnosti. Tato data jsou následně využívána pro validaci atmosférické části modelů.

1.3 Typy experimentů

Jedním z možných využití klimatických modelů je výzkum vlivu potenciálně se zvyšující atmosférické koncentrace skleníkových plynů na klimatický systém. Do kategorie skleníkových plynů („greenhouse gases“, GHG) spadá množství chemických sloučenin, mezi něž patří především oxid uhličitý, metan, troposférický ozon, oxid dusný, fluorované uhlovodíky, halony a další látky. V rámci klimatického modelování bývá jejich efekt často zastupován tzv. ekvivalentní koncentrací CO₂, která je definována jako taková koncentrace CO₂, jejíž vliv na radiční procesy v atmosféře je roven součtu vlivů všech skleníkových plynů (Kalvová et al., 2002). Z hlediska koncentrace skleníkových plynů a aerosolů, uvažované dotace klimatického systému globálním zářením a také z hlediska aplikované úrovně těchto vlivů v čase lze simulace klimatických modelů rozdělit do několika kategorií.

Tzv. *kontrolní běhy* modelů se vyznačují konstantní úrovní těchto působení v čase a slouží především k analýze vnitřní variability klimatického systému. Po inicializaci z počátečních a okrajových podmínek je model integrován do dosažení ustáleného stavu. Pokud předepsané koncentrace CO₂ odpovídají soudobým měřeným koncentracím, je také vyžadována dobrá shoda charakteristik ustáleného stavu s aktuálním klimatem.

Druhou skupinou jsou tzv. „*idealized experiments*“, které spočívají v navyšování koncentrací CO₂ podle předem definovaného postupu. Patří sem tzv. 2×CO₂ experimenty, při kterých dojde k jednorázovému zdvojnásobení koncentrace CO₂, po kterém se model opět integruje do dosažení rovnovážného stavu. Tyto simulace bývají prováděny obvykle na jednodušších typech klimatických modelů, tzv. EMIC – Earth System Model of Intermediate Complexity (např. Eby et al. 2013) a slouží k ověření citlivosti modelů na uměle vyvolané vzruchy. Obvyklou mírou této citlivosti je tzv. „*equilibrium climate sensitivity*“ definovaná jako změna průměrné globální teploty vzduchu při zemském povrchu zjištěná po dosažení rovnovážného stavu u 2×CO₂ experimentů. Ta se u většiny EMIC modelů pohybuje v rozmezí 1.5 až 4 stupně. V případě komplexních globálních klimatických modelů jsou k ověření citlivosti prováděny experimenty s 1% ročním navyšováním ekvivalentní koncentrace CO₂. Mírou citlivosti je pak změna globálního průměru teploty při dosažení dvojnásobku původní koncentrace CO₂. Toto citlivostní kritérium bývá označováno jako „*transient climate response*“ a jeho použití je motivováno příliš dlouhou dobou, která je v případě globálních klimatických modelů potřeba k dosažení ustáleného stavu při 2×CO₂ experimentech (Kalvová et al., 2002).

Třetí kategorií jsou *experimenty s realistickým zvyšováním koncentrace GHG* a časově proměnlivými hodnotami dalších vstupů, jako jsou množství dopadajícího globálního záření a množství antropogenních a jiných (vulkanických) aerosolů, odpovídající pozorovaným hodnotám v průběhu 20. století. Koncentrace GHG pro odhad vývoje budoucího klimatu jsou následně aplikovány podle zvolených emisních scénářů, kterými mohou být scénáře SRES

(Special Report on Emission Scenarios, viz Nakicenovic a Stewart (2000)) nebo nověji scénáře koncentrací RCP (Representative Concentration Pathways, viz van Vuuren et al. (2011)).

S klimatickým modelováním souvisí termín „klimatický scénář“, který obecně představuje nějakou reprezentaci budoucího klimatu bez specifikace způsobu, který byl při tvorbě tohoto odhadu použit. Obvyklým podkladem pro tvorbu klimatických scénářů jsou projekce klimatických modelů podle zvolených emisních scénářů. Přehled dalších metod tvorby klimatických scénářů nabízí např. Kalvová et al. (2002).

1.4 Nejistoty klimatických studií

S klimatickými studii je spjata řada nejistot. Foley (2010) rozlišuje nejistoty spjaté se samotným klimatickým systémem a dále nejistoty dané omezenou schopností tento systém modelovat.

Do první kategorie patří především nejistota ohledně dalšího vývoje lidské společnosti, který je prakticky nepředpověditelný a jeho možné alternativy jsou simulovány pomocí výše zmíněných variantních scénářů (SRES, RCP). Nejistoty jsou dále posilovány množstvím pozitivních a negativních zpětných vazeb v klimatickém systému. Například potenciální oteplení vlivem vzrůstajícího množství skleníkových plynů v atmosféře může mít za následek tání permafrostu, v jehož důsledku dojde k uvolňování metanu, který bude následně akcelarovat další oteplování (Anisimov, 2007). Naopak vzrůstající množství antropogenních aerosolů může mít za následek částečné ochlazování atmosféry vlivem jejich schopnosti odrážet dopadající krátkovlnnou radiaci. Dalším faktorem posilujícím nejistoty je vulkanická činnost, která může mít značný (a náhlý) vliv na globální klima.

Do druhé kategorie nejistot patří především počáteční a okrajové podmínky modelového běhu a použité parametrizace. Klimatický systém vykazuje chaotické chování, které je zapříčiněno nelinearitou řídicích rovnic dynamického jádra. Vývoj systému je sice možné opakovaně přesně reprodukovat z téhož výchozího stavu, na druhou stranu i nepatrné odchylky ve výchozím stavu vedou k velkým rozdílům ve stavu výsledném. Z těchto faktů vyplývá závislost výsledků modelového běhu na počátečních měřených hodnotách, které nejsou k dispozici v dostatečné kvalitě a prostorovém rozlišení. Okrajové podmínky jsou dány například dotací systému extraterestrickou radiací nebo předepsanými typy vegetačního krytu. Zásadním faktorem podílejícím se značnou měrou na celkové nejistotě jsou pak parametrizace jevů malého měřítko, neboť používané vztahy jsou obvykle obtížně validovatelné a navíc nelze automaticky předpokládat, že budou platit i při změně klimatických podmínek.

Uvedené faktory mají za následek rozdíly ve výsledcích jednotlivých modelů a také rozdíly mezi běhy jednoho modelu za pozměněných podmínek. K diferenciaci mezi modely přispívají i zvolené numerické metody a různé prostorové a časové rozlišení. Z těchto důvodů je nutné interpretovat výsledky jednotlivých běhů modelu s určitou opatrností. V klimatologických studiích jsou proto individuální simulace nahrazovány tzv. ensembly, které sestávají z většího množství modelových výsledků, ze kterých je odvozována střední hodnota a možné rozpětí předpokládaných změn klimatu.

Ensemble může být tvořen buď výsledky více modelů, nebo výsledky jednoho modelu dosaženými za pozměněných podmínek. Mitchell a Hulme (1999) upozorňují, že sestavení ensamble z několika modelů má smysl pouze v případě, že modely jsou kvalitní (prokazují

dobrou schopnost simulovat aktuální klima) a vnitřně nezávislé (používají rozdílná numerická a parametrizační schémata). Autoři dále poukazují na to, že v praxi jsou modely vybírány spíše podle jejich dostupnosti než na základě výše uvedených kritérií.

Další otázkou ohledně používání ensemblů je, zda výsledky jednotlivých modelů vážit na základě nějakého kvalitativního kritéria. Populární metodou vyhodnocování ensemblů, založenou na vážení výsledků jednotlivých modelů je tzv. REA (Reliability Ensemble Averaging), kterou publikovali Giorgi a Mearns (2002). Vstupními údaji do výpočtu jsou projekce jednotlivých klimatických modelů, ze kterých jsou odečteny změny posuzované meteorologické veličiny ΔX_i vzhledem k referenčnímu (dnešnímu) klimatu. Celková změna je podle REA dána váženým průměrem jednotlivých změn

$$\overline{\Delta X} = \frac{\sum_i R_i \Delta X_i}{\sum_i R_i} \quad (1.1)$$

Váhy R_i přisuzované jednotlivým modelům jsou dány součinem dvou kritérií. Prvním kritériem je chyba modelu při reprodukci klimatu 20. století (B_i), druhým je odlehlost projekce modelu od průměru celého ensmbu (D_i). Obecný předpis pro výpočet vah je dán vztahem

$$R_i = \left(\left(\frac{e}{|B_i|} \right)^m \left(\frac{e}{|D_i|} \right)^n \right)^{1/(mn)} \quad (1.2)$$

kde e je míra přirozené variability posuzované veličiny, vypočítaná z pozorovaných časových řad 20. století po odečtení lineárního trendu. Parametry m a n umožňují nastavit vzájemnou váhu oběma posuzovacím kritériím, v původní práci Giorgi a Mearns (2002) jsou oba rovny jedné. Kritérium D_i je dáno vzdáleností projekce modelu od ensmblového průměru (1.1), proto musí být jeho výpočet prováděn iterativně, konvergence však nastává po několika iteracích.

K vyčíslení míry rozptýlení vstupního souboru ΔX_i používají autoři obdobu směrodatné odchylky, kde aritmetický průměr je nahrazen váženým průměrem (1.1). Z míry rozptýlení je teoreticky možné počítat i klasický odhad intervalu spolehlivosti, k tomu je však třeba znát typ rozdělení náhodné veličiny ΔX . Pro určení typu rozdělení je však obvykle k dispozici málo modelů ensmbu.

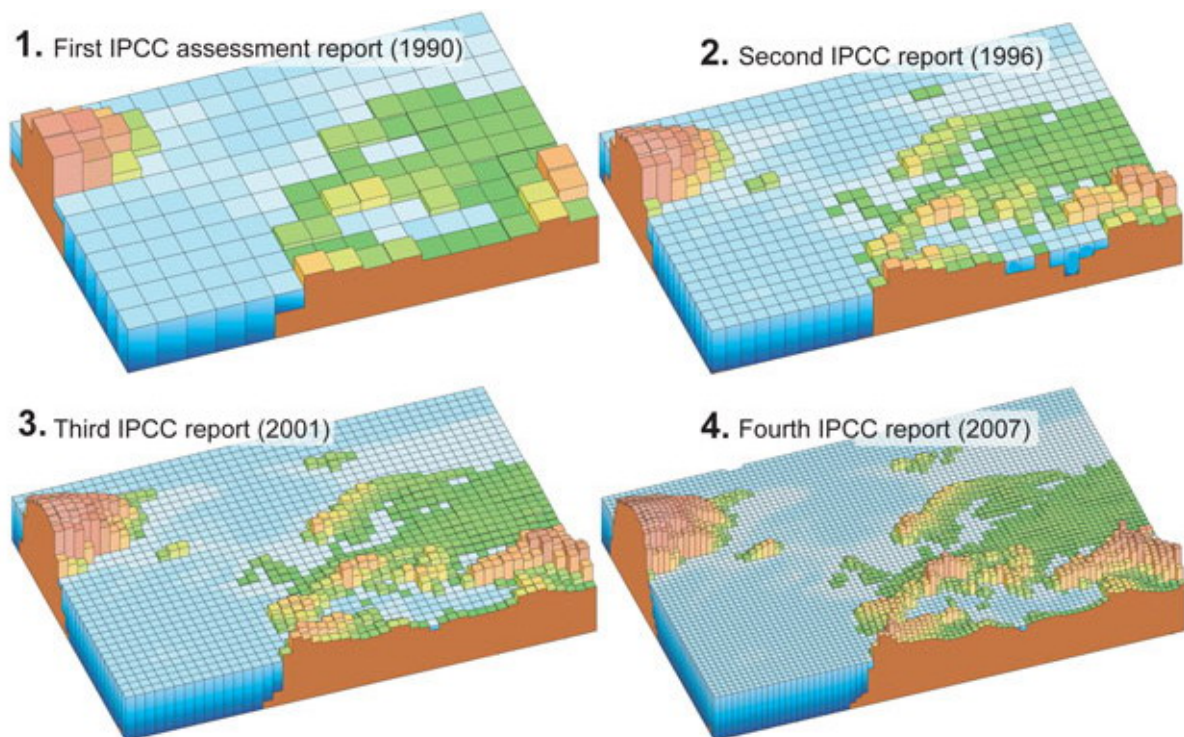
Užitečnost pravděpodobnostního odhadu klimatické změny pro následné rozhodovací procesy zmiňují např. Raisanen a Palmer (2001), kteří zároveň nabízejí jednoduchý postup pro odhad pravděpodobnosti překročení určité hodnoty ΔX jako podílu modelů ensmbu, které danou hodnotu překročí. Giorgi a Mearns (2003) tento postup modifikovali v souladu s metodou váženého průměrování REA. Definují spolehlivost jednotlivého modelu P_i jako podíl jeho váhy podle (1.2) na sumě vah celého ensmbu

$$P_i = \frac{R_i}{\sum_j R_j} \quad (1.3)$$

a pravděpodobnost překročení hodnoty ΔX definují jako součet spolehlivostí modelů, které překročení této hodnoty připouští.

1.5 Vývoj klimatických modelů

První pokusy o predikci vývoje klimatu založené na výpočtech cirkulace atmosféry, byly provedeny v 50. letech, nedlouho po etablování numerických modelů pro předpověď počasí. Zjednodušená reprezentace oceánu v raných klimatických modelech však způsobovala neadekvátně rychlé reakce modelovaného systému na vnesené impulsy. Propojení modelu atmosféry s trojrozměrným modelem cirkulace oceánu, který bere v úvahu oceánskou cirkulaci a tepelnou kapacitu mořských hlubin, bylo provedeno v 70. letech v Geophysical Fluid Dynamics Laboratory na Princetonské univerzitě. Plné využití spojených atmosféricko-oceánských modelů nastalo vzhledem k dostupnému výpočetnímu výkonu až v 90. letech, vývoj rozlišení modelů ilustruje obrázek 1.2. Využití klimatických modelů k predikci důsledků možného antropogenního vlivu na klimatický systém bylo navrženo již v roce 1971 v publikaci „Inadvertent Climate Modification“, vydané v Massachusetts Institute of Technology (SMIC, 1971). Zvyšující se úsilí v této oblasti v 80. letech spolu s prudkým rozvojem modelů vyústilo v ustanovení Mezivládního panelu pro změny klimatu (IPCC) v roce 1988. Detailní přehled vývoje výzkumu klimatu poskytuje Solomon et al. (2007), kapitola 1.



Obr. 1.2 Vývoj rozlišení globálních klimatických modelů používaných při tvorbě hodnotících zpráv IPCC (zdroj: www.ipcc.ch).

Tabulka 1.1 představuje ilustrativní výběr z celkem 23 globálních modelů použitých při tvorbě čtvrté hodnotící zprávy IPCC z roku 2007. Pátá hodnotící zpráva IPCC z roku 2013 byla sestavena na základě výsledků 54 modelů, jejichž přehled nabízí Stocker et al. (2013) a které představují seznam aktuálně používaných globálních modelů.

Tab. 1.1 Přehled vybraných globálních klimatických modelů, použitých pro vytvoření čtvrté hodnotící zprávy IPCC (Solomon et al. 2007).

Model	Instituce, stát	Rozlišení atmosférické / oceánské části [° z. délka × ° z. šířka]	Reference
CCSM3	National Center for Atmospheric Research, USA	1.4° × 1.4° / 0.3° - 1° × 1°	Collins et al. (2006)
CNRM-CM3	Météo France, Centre National de Recherches Météorologiques, France	1.9° × 1.9° / 0.5° - 2° × 2°	Déqué et al. (1994)
ECHAM5/MPI-OM	Max Planck Institute for Meteorology, Germany	1.9° × 1.9° / 1.5° × 1.5°	Roeckner et al. (2003)
GFDL-CM2.1	Geophysical Fluid Dynamics Laboratory, Princeton, USA	2.0° × 2.5° / 0.3° - 1° × 1°	Delworth et al. (2006)
UKMO-HadCM3	Hadely Centre for Climate Prediction and Research/Met Office, UK	1.3° × 1.9° / 0.3° - 1° × 1°	Johns et al. (2006)

1.6 Systematické chyby globálních modelů a metody regionalizace

Simulace globálních klimatických modelů jsou vysoce náročné na výpočetní výkon, což je dáno numerickým řešením řídicích rovnic modelů cirkulace atmosféry a oceánu, a také délkou prováděných integrací. V důsledku této výpočetní náročnosti mají modely hrubé prostorové rozlišení, které se odráží ve značně zjednodušené reprezentaci zemského povrchu. V případě atmosférické části nelze ani při použití nejvýkonnějších počítačů dosáhnout vyššího rozlišení než v řádu stovek km, což nedostačuje ke kvalitnímu zachycení meteorologických jevů, které určují regionální či lokální klima. Procesy nastávající v menším prostorovém měřítku (vznik konvektivní oblačnosti), nebo procesy ovlivněné členitostí terénu (orografické srážky) jsou modelem vystihnuty nedostatečně. Tento stav způsobuje disproporci mezi potřebami vědecké komunity zkoumající dopady potenciální změny klimatu a mezi informací, kterou je globální model schopen v dostatečné kvalitě poskytovat. Modelové výstupy veličin vykazujících lokální proměnlivost (srážky, teploty, vlhkosti vzduchu) jsou ve svých statistických ukazatelích systematicky odchýleny od pozorovaných hodnot a nemohou být v následných studiích použity bez nějaké formy korekce. Metody přenosu velkorozměrové informace globálního modelu do podrobnějších měřítek bývají označovány jako regionalizace, popř. anglickým termínem „downscaling“.

V zásadě se nabízí dva odlišné přístupy. Oba jsou založeny na tom, že globální model dobře vystihuje velkoprostorové charakteristiky stavu atmosféry (např. geopotenciální výšky) a že mezi těmito charakteristikami a hodnotami lokálních veličin lze nalézt jisté vazby (Leung et al., 2003).

Statistický downscaling spočívá v nalezení funkčního vztahu mezi velkoprostorovými charakteristikami (prediktory) a cílovými lokálními proměnnými (prediktandy). K odvození a validaci funkce je nutná analýza historických měřených dat a kontrolních běhů modelu.

Lokální proměnné pro budoucí klima, obvykle teploty a srážkové úhrny, jsou pak získány jako funkční hodnoty prediktorů poskytnutých globálním modelem.

Dynamický downscaling spočívá v numerické simulaci prováděné modelem založeným na stejných principech jako je model globální, avšak s jemnějším prostorovým a časovým rozlišením, při kterém je možné vystihnout meteorologické jevy utvářející charakter regionálního či lokálního klimatu. V důsledku výše zmiňované výpočetní náročnosti mohou být tyto podrobnější simulace prováděny buď na omezené oblasti, či na omezeném časovém úseku. Výstupy globálních modelů jsou aplikovány jako počáteční a okrajové podmínky na hranicích integrační oblasti.

Oba popsané principy mají specifické výhody a omezení a nelze určit, který z nich poskytuje kvalitnější výsledky. Statistický downscaling je nenáročný na výpočetní čas, po nalezení funkčního vztahu je získání hodnot prediktandů snadné a rychlé. Na druhé straně není zcela samozřejmé, že je nalezený vztah při jeho aplikaci platný. Funkce je odvozena z datových podkladů reprezentujících současné nebo minulé klimatické podmínky, a posléze aplikována na prediktandy reprezentující podmínky odlišné. Tato nestacionarita mezi kalibrací a aplikací vnáší další nejistotu do navazujících studií.

Dynamický downscaling tímto problémem netrpí, neboť simulace podrobnějších modelů kontinuálně zachycuje změny klimatu reprezentované globálním modelem. Nevýhodou dynamického přístupu je především výpočetní náročnost, která se pohybuje na úrovni modelu globálního, a dále pak skutečnost, že výstupy podrobnějších modelů částečně zachovávají systematické odchylky od měřených hodnot. Na rozdíl od statistického přístupu dovolují podrobnější modely analyzovat chování klimatického systému v regionálním měřítku. Prostorové rozlišení, které se u těchto modelů pohybuje v desítkách km, však není dostatečné pro vystihnout lokální variability meteorologických veličin, což platí především pro srážkové úhrny. Proto bývají výstupy regionálních modelů obvykle podrobeny *korekci chyb* („biass correction“). Ta je v jistém smyslu analogií statistického downscalingu, ovšem jako prediktory zde většinou vystupují samotné lokální veličiny, které jsou předmětem korekce. Nejde tedy o přenos velkorozměrové informace do regionálního měřítka, ale o systematickou opravu modelových hodnot tak, aby jejich statistické charakteristiky odpovídaly měřeným hodnotám. Obdobná korekce modelových hodnot probíhá i při standardní předpovědi počasí (Wilks, 2011).

2 Regionální klimatické modely

Jak již bylo uvedeno výše, termínem dynamický downscaling se označuje simulace vývoje klimatu prováděná modely s vyšším prostorovým a časovým rozlišením, které jsou založeny na stejných principech jako globální klimatické modely (numerické řešení rovnic dynamického jádra). Data globálních modelů přitom slouží jako počáteční a okrajové podmínky integrace. Cílem simulací je vystihnout meteorologické procesy, jejichž rozměry jsou pod rozlišovací schopností globálních modelů a které určují klima v regionálním měřítku. Vyšší prostorové rozlišení modelu přispívá k lepšímu vystižení heterogenity zemského povrchu, která je jednou z hlavních příčin tvorby jevů regionálního a lokálního měřítku. Spolu s rozlišením také obecně roste kvalita numerické simulace. Vyšší časové rozlišení také způsobuje realističtější variabilitu simulovaných jevů ve výsledných časových řadách.

Dynamický downscaling může být realizován několika způsoby. Podle rozsahu výpočetní oblasti a způsobu zahušťování výpočetní sítě lze rozlišit tři základní kategorie modelů:

- *regionální klimatické modely* (RCM, též „limited area models“ (LAM) nebo „nested models“) mají výpočetní doménu omezenou na zájmovou oblast
- tzv. *stretched-grid* modely jsou globální modely, které mají v zájmové oblasti zahuštěnou výpočetní síť
- tzv. *time-slice* modely jsou rovnoměrně zahuštěné globální modely, simulující omezený časový úsek běhu globálního modelu.

Všechny tři uvedené kategorie představují modely atmosférické cirkulace doplněné o model zemského povrchu. Globální model poskytuje informace o stavu atmosféry na okrajích integrační oblasti a o teplotě na hladině oceánu. Okrajové podmínky však mohou být dány i pozorovanými hodnotami či re-analýzami, jednak za účelem studia samotného klimatického systému, popřípadě pro zhodnocení kvality downscalingu. Experimenty s několika sety okrajových podmínek pro posouzení jejich vlivu na nejistoty regionálních klimatických predikcí prováděli např. Pan et al. (2001).

Z výše uvedených kategorií modelů jsou nejvyužívanější modely regionální, o kterých bude pojednávat další text. Za fundamentální práce na poli regionálního klimatického modelování jsou považovány publikace Dickinson et al. (1989) a Giorgi (1990) na území USA, v Evropě např. Jones et al. (1995). Koncept vnořeného regionálního modelu řízeného modelem globálním však vznikl v souvislosti s předpovídáním počasí již před nástupem klimatického modelování a regionální klimatické modely jsou s modely pro předpověď počasí strukturálně shodné. Principiální rozdíl mezi uvedenými úlohami spočívá v tom, že simulace počasí na několik dní dopředu je problémem počátečních podmínek, kdežto největší nejistotu do delších simulací klimatických modelů vnášejí podmínky okrajové (Rummukainen, 2010).

2.1 Faktory určující kvalitu regionálních modelů

Horizontální rozlišení dnešních regionálních modelů se pohybuje v desítkách km, vertikální rozlišení se oproti globálním modelům příliš nemění a zůstává cca 1 km. Časové rozlišení se zvyšuje spolu s horizontálním, což je nutná podmínka numerické stability dynamického jádra, a v současnosti se pohybuje v jednotkách minut. Dalším faktorem, který se vztahuje k rozlišení modelů a ovlivňuje výsledky downscalingu, je poměr rozlišení řídicího globálního a vnořeného regionálního modelu. Antic et al. (2004) na základě experimentálních výsledků uvádí, že tento poměr by neměl přesáhnout číslo 12, Rummukainen (2010) udává jako

optimální poměr 6 – 8. Podobné experimenty byly prováděny i se stretched-grid modely, kdy byl sledován maximální možný poměr mezi nehrubší a nejjemnější částí sítě, který by podle Bader et al. (2008) neměl přesáhnout číslo 3.

Přes podrobnější prostorové rozlišení regionálních modelů je stále nutné parametrizovat řadu meteorologických jevů. Patří mezi ně především tvorba a mikrofyzika oblačnosti, transfer radiace a konvektivní procesy malých měřítek. Podrobnější modelování přináší nutnost změny v parametrizačních schématech oproti globálnímu modelu, protože cirkulace atmosféry v regionálním měřítku má s lokálními proměnnými jiné statistické vztahy, což platí obzvláště pro parametrizaci konvektivních procesů (Wang a Kotamarthi, 1014). V současnosti je pro každý subgridový proces k dispozici několik parametrizačních schémat, což přispívá k diferenciaci mezi jednotlivými regionálními modely. V souvislosti se zvyšováním rozlišení také vyvstává otázka, zda je objem jednoho elementu výpočetní sítě dostatečně velký k tomu, aby ospravedlnil stochastické modelování, kterým parametrizace jsou, neboť zachycují celkový efekt heterogenního děje uvnitř tohoto elementu (Bader et al., 2008).

Počáteční a okrajové podmínky, které poskytuje globální model, jsou tři složky vektoru rychlosti proudění vzduchu, teplota, tlak a obsah vodních par, popř. teplota hladiny oceánu. Tyto hodnoty jsou k dispozici pro každý výpočetní krok globálního modelu v každém uzlovém bodě jeho výpočetní sítě. První sada slouží jako počáteční podmínky, další jako podmínky okrajové. Na jemnější síť vnořeného modelu musí být interpolovány, v horizontálním směru bývá prováděna bilineární interpolace, ve vertikálním směru pak interpolace lineární. Vzhledem ke kratšímu výpočetnímu kroku vnořeného modelu je nutné provádět lineární interpolaci také po časové ose mezi dvěma poskytnutými okrajové podmínky řídicím modelem (Giorgi, 1990).

Frekvence poskytování okrajových podmínek je faktorem, který ovlivňuje kvalitu výsledků simulace. Antic et al. (2004) experimentálně zjistili, že časová perioda obnovení okrajové podmínky vyšší než dvanáct hodin silně negativně ovlivňuje kvalitu simulovaných srážkových úhrnů. Giorgi (1990) konstatuje, že dostatečná frekvence je individuální a závisí na výskytu rychlých disturbancí na okrajích integrační oblasti, přičemž podává doporučení umisťovat okraje modelové domény do oblastí mimo výskyt těchto jevů. Denis et al. (2002) doporučují použít pro přibližný výpočet potřebného času obnovení poměr rozměru (km) a rychlosti (km/h) jevů, které by měl regionální model zachytit. Čas obnovení by pak měl být menší než čtvrtina tohoto poměru. Například tropický synoptický systém o průměru 1000 km a rychlosti 50 km/h vyžaduje čas obnovení maximálně 5 hodin, aby byl modelem dobře zachycen.

Okrajové podmínky jsou aplikovány ve směšovací vrstvě (v literatuře „relaxation zone“, „buffer zone“ apod.), která zasahuje několik uzlových bodů podél okraje výpočetní domény směrem dovnitř regionálního modelu. V této vrstvě dochází ke smísení hodnot okrajové podmínky ψ^* a hodnot vnitřního řešení regionálního modelu ψ z předchozího kroku, hodnoty směšovací vrstvy ψ_{sv} jsou počítány podle vztahu

$$\psi_{sv} = \psi + \gamma(\psi^* - \psi) \quad (2.1)$$

kde parametr γ nabývá na okraji výpočetní domény hodnoty 1, směrem dovnitř směšovací vrstvy jeho hodnota klesá k nule. Ve směšovací vrstvě navíc často probíhá dodatečné vyhlazování hodnot. Šířka směšovací vrstvy bývá obvykle 4 – 10 uzlových bodů. Staniforth (1997) doporučuje realizovat vrstvu co možná nejširší, protože se tím minimalizuje riziko porušení dynamické rovnováhy pomocí uměle dodaného gradientu, což může vyústit v dodatečnou falešnou akceleraci jednotlivých složek proudění.

Velikost výpočetní oblasti je dalším faktorem, který má vliv na výsledek simulace. Pokud je výpočetní oblast příliš malá, systém je pod neúměrně velkou kontrolou okrajových podmínek a nedojde k rozvinutí cílových jevů středního a malého měřítka. Příliš velká výpočetní oblast má naopak za následek nežádoucí odklon velkoprostorových charakteristik stavu atmosféry vygenerovaných regionálním modelem od řídicích dat globálního modelu (Pan et al., 2001, Leung et al., 2003). Velikostí výpočetní oblasti jsou nejvíce ovlivněny simulované srážkové úhrny. Jacob a Podzun (1997) při experimentech s regionálním klimatickým modelem REMO zjistili, že modely s různou velikostí výpočetní domény produkují odlišné sumy a pravděpodobnostní spektra srážkových úhrnů.

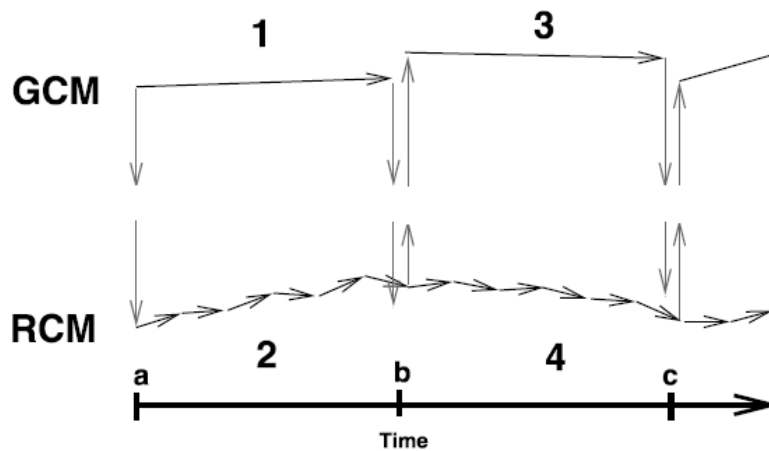
Problém přílišné odchylky řešení regionálního modelu od okrajové podmínky lze potlačit pomocí techniky označované jako „spectral nudging“, která spočívá v částečné aplikaci okrajové podmínky uvnitř výpočetní domény vnořeného modelu. Spectral nudging je analogií mísení okrajové podmínky a vnitřního řešení ve směšovací vrstvě modelu podle vztahu (2.1) s tím rozdílem, že parametr γ zde není aplikován na hodnoty řešení v okrajové oblasti, ale na koeficienty jednotlivých frekvencí diskrétní Fourierovy transformace řešení v celém rozsahu výpočetní oblasti. Tento postup publikovali von Storch et al. (2000), v jejich pojetí je parametr γ závislý na výšce a jeho hodnota klesá směrem k zemskému povrchu. Přízemní atmosféra v regionálním modelu tak není ovlivňována okrajovou podmínkou a mohou se rozvinout cirkulační charakteristiky dané lokální orografií, využitím území apod.

Spectral nudging bývá označován za poněkud kontroverzní techniku, protože sice na jedné straně dobře zabraňuje přílišným odchýlkám výsledků regionálního modelu od řídicích dat, na druhé straně ale způsobuje zhoršení v simulaci srážkových extrémů a obecně všech jevů vysokého rozlišení (Rummukainen, 2010). Aplikace se doporučuje na modely s velkou výpočetní oblastí, menší modely nemají tendenci tolik divergovat od výsledků globálního modelu, neboť malá výpočetní oblast je sama o sobě pod dostatečnou kontrolou okrajových podmínek. Von Storch et al. (2000) obecně nedoporučují používat spectral nudging na modely, jejichž smyslem je modelovat dynamiku lokálních jevů spíše než dlouhodobé klima a dále na modely s obousměrným doplňováním, tzv. „two-way nesting“ (viz dále).

2.2 Modifikace základního konceptu vnoření

Spojený systém globální – regionální model bývá obvykle realizován jako jednosměrný, tzn. výsledné řešení regionálního modelu není poskytováno zpět globálnímu modelu. Regionální procesy v exponovaných oblastech však mohou mít na globální proudění značný vliv. Detailní řešení regionálního modelu pro takovou oblast proto může sloužit globálnímu modelu jako okrajová podmínka do dalších iterací. Tento přístup bývá označován jako two-way nesting, což bude v tomto textu překládáno jako obousměrné doplňování. Je zřejmé, že užitečnost tohoto přístupu silně závisí na vhodné volbě modelové oblasti. Základní princip spočívá v tom, že globální model poskytne okrajové podmínky modelu regionálnímu, který provede takový počet iterací, aby obsáhl jeden krok globálního modelu. Výsledný stav regionálního

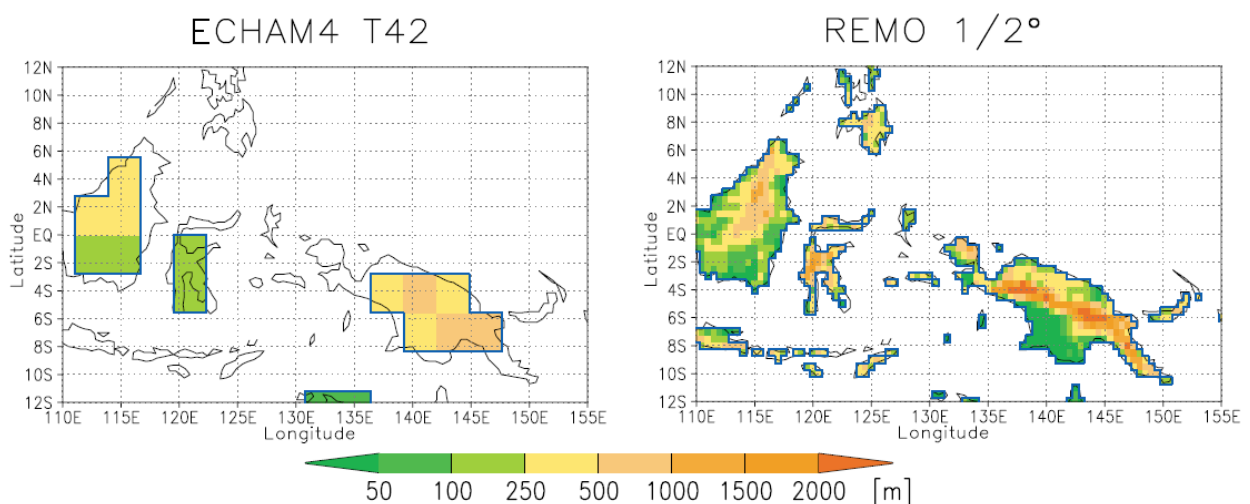
modelu je průměrován a interpolován na uzly výpočetní sítě řídicího modelu, kde nahradí jeho aktuální stav. Ten je naopak převzat regionálním modelem a celý mechanismus se opakuje. Celý proces je schematicky znázorněn na obr. 2.1.



Obr. 2.1 Mechanismus obousměrného doplňování (převzato z Lorenz a Jacob, 2005).

Díky interpolaci a průměrování hodnot vnořeného modelu je podstatná část jeho informace ztracena, nicméně přesto dochází k mírné modifikaci stavu globálního modelu. Systém je výpočetně značně náročný, neboť vyžaduje simultánní chod obou modelů.

Mechanismus obousměrného doplňování použili Lorenz a Jacob (2005) při spojení globálního modelu ECHAM4 a regionálního modelu REMO, vyvinutých v Max Planck Institute for Meteorology v Německu. Regionální model byl umístěn do oblasti jihovýchodní Asie na rozhraní Indického a Tichého oceánu, pokrývající Filipíny a Indonésii. Tato oblast, označovaná jako tzv. „maritime continent“, je jednou z klíčových pro rozvoj globálního proudění atmosféry a vyznačuje se vysokými teplotami moře, intenzivními tepelnými toky, častým výskytem silných konvekci a vysokými srážkovými úhrny. Obrázek 2.2 ilustruje rozdíl v prostorovém rozlišení sledované oblasti dosaženém oběma modely. Výsledkem studie bylo jednoznačné zjištění, že systém s obousměrným doplňováním vede ke kvalitnější simulaci teplotních a srážkových polí daleko za hranicemi oblasti regionálního modelu.



Obr. 2.2 Prostorové rozlišení oblasti „maritime continent“ dosažené globálním a regionálním modelem (převzato z Lorenz a Jacob, 2005).

Jinou modifikací klasického systému globální – regionální model je systém vícenásobného vnoření. Feser et al. (2011) uvádí, že numerické modely obecně řeší meteorologické procesy v rozměru zhruba od čtyřnásobku vzdálenosti uzlů výpočetní sítě. Při snaze o zachycení detailních jevů tedy může nastat situace, kdy dojde k příliš vysokému poměru mezi rozlišením globálního a regionálního modelu. Řešením je systém vícenásobného vnoření, kdy globální model řídí model středního měřítka, který poskytuje okrajové podmínky cílovému regionálnímu modelu. Obě vnoření jsou principiálně shodná, mateřský model poskytuje vždy informaci o rychlosti a směru proudění, teplotě, obsahu vodních par a tlaku. Výhodou je, že okrajové podmínky mohou být regionálnímu modelu poskytovány v kratších intervalech. Tento postup použili Im et al. (2006) pro kontrolní simulaci aktuálního klimatu na korejském poloostrově, kdy střední model o rozlišení 60 km řídil regionální model o rozlišení 20 km. Výsledky studie ukázaly, že průměrné hodnoty generované oběma vnořenými modely jsou shodné, ale vnitřní model vykazuje zřetelně lepší prostorové rozložení simulovaných veličin. Vedlejším výsledkem bylo zjištění, že pro letní srážky na korejském poloostrově toto zlepšení zaregistrováno nebylo, neboť primárním mechanismem jejich vzniku jsou monzunové konvekce a tropické bouře. Tyto jevy vykazují náhodný charakter a prakticky je nelze předpovídat, proto regionální model v tomto období pro srážky nepřinášel žádné zlepšení. Toto zjištění podporuje obecnou teorii, že vnořený model přináší efekt pouze v lokalitách a situacích, kdy jsou sledované veličiny výrazně ovlivněny lokální orografií terénu.

2.3 Problematické aspekty konceptu vnoření

Na kvalitu systému globální – regionální model má vliv řada faktorů. Mezi ty základní patří faktory již popsané v předchozích odstavcích – poměr rozlišení globálního a regionálního modelu, frekvence poskytování okrajových podmínek a velikost domény vnořeného modelu. Kromě těchto faktorů vyjmenovávají Denis et al. (2002) řadu dalších, které také výrazně ovlivňují kvalitu výsledného systému. Jedná se zejména o

1. *matematickou formulaci vnoření* – diskrétní aproximace nelineární soustavy parciálních diferenciálních rovnic je sama o sobě problematická a vnořený model matematickou náročnost dále umocňuje, komplexně se tímto bodem zabývá Staniforth (1997), který rozebírá otázky strategií aplikace okrajových podmínek a jejich potřebného počtu,
2. *spin-up time* – čas, který model potřebuje po inicializaci k ustavení své vlastní dynamické rovnováhy,
3. *konzistenci fyzikálních parametrizací* – rozdílná parametrizační schémata produkují mírně odlišné výsledky, tudíž použití různých schémat na sousedících uzlových bodech systému může vést k nefyzikálnímu gradientu odhadované veličiny, tento problém je důležitý z toho důvodu, že rozdílná měřítka modelů obvykle použití různých schémat vyžadují,
4. *interpolační chyby* – interpolace okrajových podmínek jsou nutné kvůli napojení modelů, v důsledku příliš zjednodušené interpretace povrchu v globálním modelu často dochází k extrapolacím pod jeho vertikální rozsah,
5. *kvalitu řídicích dat* – ta je zřejmě nejdůležitějším faktorem determinujícím kvalitu výsledné simulace, globální model je hlavním zdrojem predikce, regionální model (v případě jednosměrného systému) pouze vykresluje jevy menších měřítek, platí pravidlo „garbage in → garbage out“ (Foley, 2010),
6. *systematické chyby vnořeného modelu*.

Uvedený výčet potenciálních problémů přináší jistou skepsi ohledně možného využití regionálních modelů. Denis et al. (2002) proto přišli s experimentem testujícím samotnou podstatu konceptu regionálního modelování. Takzvaný „Big brother experiment“ hodnotí schopnost regionálního modelu produkovat jevy středních měřítek, které nejsou obsaženy v informaci okrajové podmínky, přitom je však designován tak, aby jeho výsledky nebyly ovlivňovány chybami řídicího globálního modelu nebo chybami měřených hodnot. První fází experimentu je nasimulování referenčního klimatu pomocí globálního modelu s vysokým rozlišením, nebo pomocí regionálního modelu s dostatečně velkou výpočetní oblastí (big brother model – BB). Z výsledků této simulace jsou potom odfiltrovány jevy středního měřítka až po úroveň rozlišení danou běžným globálním modelem. Filtrované výsledky potom slouží jako okrajové podmínky pro regionální model (little brother model – LB), který je ve všech ohledech identický jako BB model (prostorové rozlišení, parametrizační schémata, použité numerické metody), ale jeho výpočetní doména je menší a je umístěna uvnitř domény BB modelu. Nakonec dojde k porovnání klimatu nasimulovaného LB modelem s referenčním klimatem vytvořeným BB modelem. Protože výstupy LB modelu nejsou porovnávány s měřenými daty, jsou z hodnocení eliminovány chyby řídicího modelu a chyby měření. Případné rozdíly mezi výslednými klimatickými projekcemi jsou tak způsobeny pouze vnitřními problémy systému regionální – globální model. Filtrování referenčního klimatu probíhá rozkladem výsledků BB modelu pomocí diskrétní kosinové transformace, kdy všechny jevy s vlnovými délkami menšími než daná prahová hodnota (v případě studie Denis et al. (2002) se jednalo o 500 km) jsou ze signálu odstraněny. Podrobný popis použitého filtru podává Denis et al. (2001). Výsledky big brother experimentu provedeného na území USA s modely o prostorovém rozlišení 45 km ukazují, že uvnitř domény vnořeného modelu dochází ke kvalitní regeneraci jevů malých a středních měřítek a potvrzují potenciál regionálních modelů přinášet hodnověrné dodatečné informace k výsledkům globálních experimentů.

Obecně lze konstatovat, že vyšší rozlišení přináší lepší výsledky (shodu s pozorováním) u veličin vykazujících vysokou prostorovou variabilitu, jako jsou srážky nebo teploty, nemá však výraznější vliv na simulaci proudění ve volné atmosféře. Tento fakt byl několikrát potvrzen v experimentálních studiích. Např. Mo et al. (2005) při experimentech se dvěma modely s různým rozlišením potvrdili, že vyšší rozlišení má zřetelný vliv na kvalitu simulace monsunových dešťů a celkově přináší realističtější časovou a prostorovou variabilitu srážek, nezpůsobuje však žádné zlepšení v případě rozložení tlakových polí. K podobným závěrům došli Hay et al. (2006), Feser et al. (2011) a další autoři.

Regionální klimatické modely mají široké spektrum využití, Rummukainen (2010) uvádí např.

- *studium klimatických procesů* (model simuluje i veličiny, které se neměří),
- *predikce velikosti potenciální klimatické změny,*
- *zdroj dat pro podrobné dopadové studie,*
- *měsíční a sezónní předpovědi počasí,*
- *testovací nástroj pro parametrizace globálních modelů.*

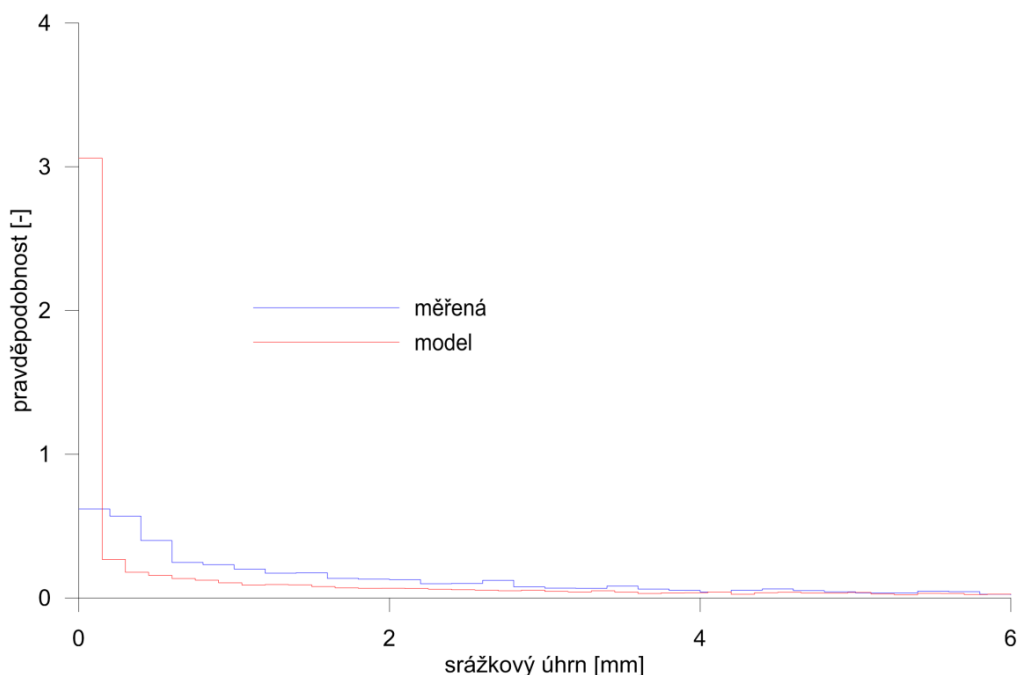
3 Korekce chyb

Regionální klimatické modelování přináší efekt v podobě zlepšené simulace meteorologických jevů středních měřítek, nicméně výpočetní síť regionálních modelů stále není dostatečně podrobná k zachycení lokálních charakteristik meteorologických veličin v blízkosti zemského povrchu. Rozlišení v řádu desítek kilometrů neumožňuje detailní vystihnout orografie terénu, která určuje proudění přízemních vrstev atmosféry a formuje tak lokální rozložení teplot a srážkových úhrnů. Kromě hrubého rozlišení přispívá k nepřesnostem i nedokonalá vnitřní struktura modelu, který je (přes svou značnou složitost a sofistikovanost) stále velmi zjednodušenou aproximací klimatického systému a vyžaduje množství parametrizací. Dále je nutno brát v úvahu, že data generovaná modelem představují průměrné hodnoty pro oblast příslušnou celému elementu výpočetní sítě o rozloze v řádu stovek km². Porovnání výstupů regionálních modelů s bodovými měřenými hodnotami z meteorologických stanic tudíž ukazuje, že modelová data jsou zatížena systematickou chybou, která komplikuje jejich využití v hydrologických, ekologických či zemědělských studiích. Přímé použití syrových modelových dat v navazujících studiích je sice možné (viz např. Arnell et al., 2003), ale naprostá většina autorů přistupuje k dodatečné úpravě dat, tzv. korekci chyb („bias correction“).

3.1 Úvod

Účelem korekce chyb je transformace souboru modelových hodnot tak, aby statistické vlastnosti korigovaného souboru lépe odpovídaly pozorováním v daném místě. Kritéria používaná pro zhodnocení podobnosti korigovaných a měřených souborů zahrnují obvyklé míry polohy a rozptýlení (střední hodnoty, směrodatné odchylky), popř. shodu celkového rozložení pravděpodobnosti. Dále bývá hodnocena schopnost modelu reprodukovat extrémní hodnoty. V případě srážkových úhrnů bývá navíc posuzována frekvence srážkových dní a rozložení období sucha.

Předmětem této studie je korekce denních srážkových úhrnů, které jsou obecně považovány za veličinu nejvíce postiženou systematickou chybou. Rozborem literatury lze dospět k závěru, že odchylky modelových výstupů od měřených hodnot obvykle vykazují jisté společné rysy bez ohledu na analyzovaný regionální model. Typickým znakem modelových dat je velmi malý počet suchých dní (nulových hodnot), který bývá kompenzován příliš vysokým počtem malých srážkových úhrnů. Tento tzv. „drizzling effect“ (mrholení) je ilustrován na obr. 3.1, který obsahuje porovnání histogramů měřených srážkových úhrnů z meteorologické stanice v Českých Budějovicích a prostorově a časově odpovídajícího výstupu z regionálního klimatického modelu REMO, vyvinutého v Max Planck Institute for Meteorology, viz Jacob (2001). Efekt mrholení nelze automaticky považovat za chybu modelu, protože jak již bylo uvedeno, modelová hodnota může být pokládána za průměr pro oblast reprezentovanou jedním elementem výpočetní sítě. Je zřejmé, že pokud se pokusíme aproximovat modelové hodnoty tím, že spočítáme denní průměry z dostupných meteorologických stanic na příslušné oblasti, dojde ve výsledné časové řadě také ke ztrátě nulových hodnot a vzniku množství malých úhrnů, což ilustrovali např. Hay et al. (2002).



Obr. 3.1 Histogramy měřených denních srážkových úhrnů ze stanice České Budějovice a výstupů modelu REMO za období 1961 – 1997. Pro lepší zřetelnost rozdílů v oblasti počátku jsou histogramy vyneseny pouze do hodnoty 6 mm, zobrazená část pokrývá zhruba 90 % hustoty pravděpodobnosti obou veličin.

Průměrováním lze tedy vysvětlit vznik mrholení, nepřítomnost nulových hodnot a navíc i nízkou četnost extrémů v modelových datech. Nicméně dlouhodobé průměry měřených a modelových hodnot se obvykle neshodují, což již vypovídá o chybě modelu. Podrobnější rozbor časových řad také ukazuje, že modelová data vykazují příliš hladký roční chod bez obvyklé sezónní či měsíční variability. Celkově lze tedy shrnout, že systematická odchylka modelových dat od měřených hodnot, ať už vzniklá průměrováním či vnitřními chybami modelu, se projevuje:

- absencí suchých dní
- příliš velkým počtem malých úhrnů
- malým počtem extrémních hodnot
- nad/pod hodnocením dlouhodobých průměrů
- nízkou sezónní variabilitou.

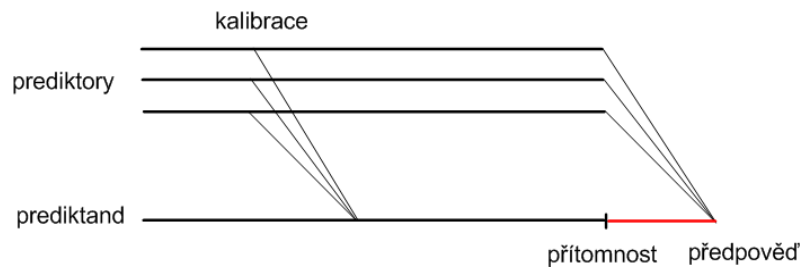
Korekce je v praxi realizována prostřednictvím tzv. transferní funkce f , která vyjadřuje vztah mezi jedním nebo více prediktory (x) a cílovou veličinou, tzv. prediktandem (y).

$$(x_1, x_2, \dots, x_n) \xrightarrow{f} y$$

Prediktory jsou v tomto případě projekce modelu pro zájmové budoucí období, obvykle jde o běžné meteorologické veličiny, může však jít také o jiné modelové výstupy, reprezentující stav atmosféry (např. geopotenciální výšky). Prediktandem jsou pak srážkové úhrny. Modelové srážkové úhrny (postížené systematickou chybou) buď mohou být z kalkulací zcela vyřazeny, nebo mohou také patřit mezi prediktory. Druhý případ je obvyklejší, modelové srážky dokonce často bývají jediným prediktorem.

Transferní funkce používané pro korekci modelových dat historicky vycházejí ze statistické předpovědi počasí, která byla založena na regresním vztahu mezi skupinou prediktorů a

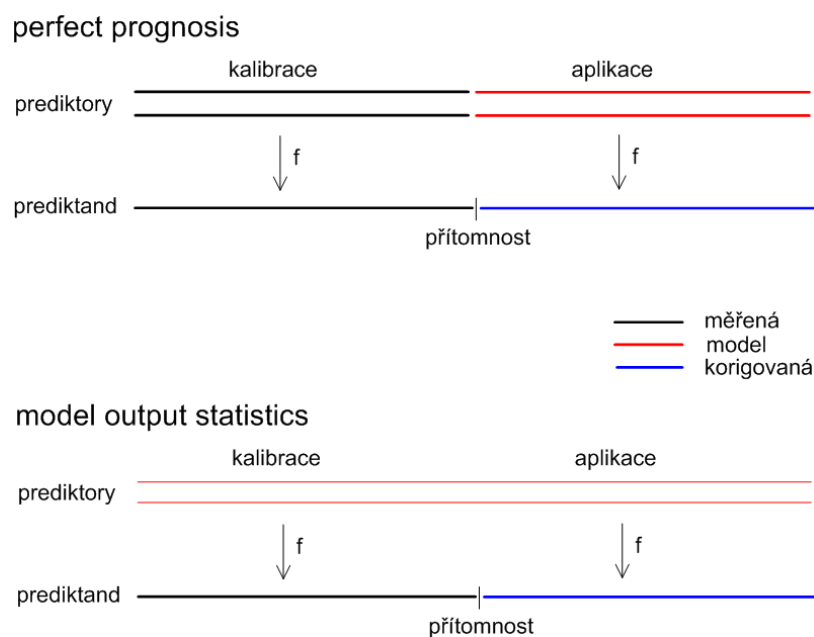
prediktandem, posunutým o určitý časový úsek dopředu, jak je schematicky znázorněno na obr. 3.2.



Obr. 3.2 Schéma statistické předpovědi počasí.

Statistická předpověď se dnes používá pouze pro velmi krátkodobé předpovědi v řádu několika hodin, tzv. nowcasting, kde by numerická předpověď byla příliš pomalá nebo neproveditelná vzhledem k charakteru úlohy. Příkladem může být predikce pohybu hurikánu pomocí tlakových výšek v jeho okolí, podrobnosti uvádí Wilks (2011).

Transfenní funkce pro statistickou předpověď je kalibrována s pomocí historických záznamů časových řad prediktorů a prediktandu a stejný postup je samozřejmě aplikován i při korekcích modelových chyb. Zde se ale nabízejí dva možné přístupy ke kalibraci, jejichž schematické znázornění je na obr. 3.3. Prvním z nich spočívá v odvození transfenní funkce z historických měřených záznamů, tento přístup bývá v literatuře označován jako „perfect prognosis“. Výhodou tohoto přístupu je, že odvozená funkce věrně odráží reálný vztah mezi prediktory a prediktandem. Může být tedy aplikována na výstupy libovolného modelu a je platná pro jakoukoliv délku předpovědi. Navíc je zřejmé, že zkvalitnění modelu (snížení systematické chyby prediktoru) vede automaticky ke zlepšení korekce, neboť prediktor se posune do stavu podobnějšímu reálným podmínkám, ze kterých byla korekční funkce odvozena. Nevýhodou tohoto přístupu je jednak to, že jako prediktory lze používat pouze běžně měřené veličiny, hlavním nedostatkem perfect prognosis je ale skutečnost, že korekční funkce v sobě nemá zakomponovanu informaci o systematické chybě konkrétního modelu.



Obr. 3.3 Rozdíly mezi perfect prognosis a model output statistics.

Druhým možným přístupem je kalibrace korekční funkce na základě archivu měřených hodnot prediktandu a modelových hodnot prediktorů. Tento přístup bývá označován jako „model output statistics“ (MOS) a jeho výhody a nevýhody jsou zřejmě právě opačné v porovnání s perfect prognosis. Zásadní předností je skutečnost, že korekční funkce zde zahrnuje systematickou chybu modelu. Ta je specifická pro každý model a je tedy nutné kalibrovat transferní funkci pro každý model zvlášť. Jako prediktoru lze použít jakoukoliv veličinu simulovanou modelem, kromě standardních veličin tedy i např. vertikální rychlosti proudění, geopotenciální výšky apod. Jako prediktoru je možné využít i tu veličinu, která je předmětem korekce, čehož se v případě srážek často využívá, jak již bylo uvedeno.

V oblasti korekcí chyb numerických modelů proudění atmosféry jednoznačně dominuje MOS přístup, což se týká nejen modelů klimatických ale i standardních modelů pro předpověď počasí, kde jsou korekce nutné v principu ze stejných důvodů jako v klimatologii (hrubé rozlišení, nedokonalé parametrizace). Předpověď počasí spočívá v krátkodobé predikci přesného stavu atmosféry a je silně závislá na počátečních podmínkách. Vzhledem k tomuto charakteru úlohy je nutné kalibrovat transferní funkci zvlášť pro různé délky předpovědi, protože systematická chyba modelu, obsažená v rovnici korekce, se s délkou předpovědi zvyšuje. Při změně (upgradu) modelu dojde ke změně systematické chyby a je tedy nutná nová kalibrace, která vyžaduje re-forecasting historických časových řad. Kvůli frekvenci upgradování modelů představuje tato skutečnost problém, nicméně v praxi je možné provádět re-forecasting pouze bodově s frekvencí jednou za několik dní, vzhledem k silné autokorelaci v časových řadách meteorologických prvků. Transferní funkce samotná je u předpovědí počasí obvykle založena na vícenásobné lineární regresii (Wilks, 2011).

Při korekcích výstupů z regionálních klimatických modelů se ke kalibraci využívá kontrolních běhů modelu, které zahrnují simulace obvykle pro druhou polovinu 20. století. Modelování klimatu je úlohou okrajových podmínek a jeho cílem je vystihnout dlouhodobé charakteristiky počasí, nikoliv jeho konkrétní vývoj v dané lokalitě. Systematická odchylka modelu bývá z tohoto pohledu považována za stacionární (např. Chen et al., 2011) a kalibrace pro různé délky předpovědí se zde neprovádí. Problémem při korekcích výstupů klimatických modelů je spíše nestacionarita způsobená změnou samotného klimatu. Primární motivací k numerickému modelování klimatického systému je odhad jeho změny způsobené zejména antropogenními vlivy. Korekce jsou tedy většinou prováděny na datových souborech, jejichž parametry se od kalibračních dat odlišují. Tato situace přispívá k nejistotám ve vyhodnocování dopadových studií. Analogický problém je spojen se statistickým downscalingem výstupů z globálních klimatických modelů.

3.2 Přehled nejčastěji používaných metod

Předmětem korekcí jsou ve většině případů denní teploty a denní srážkové úhrny, neboť tyto veličiny představují obvyklý vstup do hydrologických či zemědělských modelů. Tato kapitola se věnuje přehledu nejpoužívanějších metod (transfernihých funkcí) korekce denních srážkových sum. V případě, že je metoda používána i pro korekci modelových teplot, jsou zmíněny případné rozdíly mezi implementacemi pro obě veličiny. Před výčtem metod je uveden seznam používaných symbolů a dolních indexů.

Symboly

X	datový soubor
x	konkrétní hodnota z datového souboru
$\mu(\cdot)$	střední hodnota
$\sigma^2(\cdot)$	rozptyl
$cv(\cdot)$	variační koeficient

Dolní indexy

O	měřená data (observed)
M	modelová data
T	korigovaná (transformovaná) data
C	kalibrační (kontrolní) období
P	projekce modelu, budoucí období

Přírůstková metoda

V případě přírůstkové metody se v pravém smyslu nejedná o korekci modelové chyby, ale spíše o odvození klimatického scénáře s využitím modelových dat. Tato jednoduchá metoda bývá frekventovaně zmiňována a hodnocena spolu s klasickými korekcemi (např. v Themessl et al., 2012), proto byla zařazena i do tohoto přehledu. Její podstatou je modifikace souboru měřených dat (reprezentujících současné klima) o faktor, který představuje potenciální změnu sledované veličiny v budoucnosti. Tento modifikační faktor je odečten z modelové projekce jako diference mezi současným a budoucím obdobím. Projekce modelu pro budoucí období tedy nejsou korigovány pro následné využití v dopadových studiích, ale slouží pouze pro odvození změny sledované veličiny, tzv. klimatického signálu.

Modifikační faktor δ je pro teploty a srážky odvozen odlišným způsobem. V případě teplot je počítán jako rozdíl středních hodnot projekčního a kontrolního období modelu a měřené teploty jsou následně modifikovány aditivně:

$$\begin{aligned}\delta &= \mu(X_{MP}) - \mu(X_{MC}) \\ x_T &= x_O + \delta\end{aligned}\tag{3.1}$$

Modifikovaný soubor teplot je tedy oproti výchozímu měřenému souboru posunut o hodnotu představující hypotetický nárůst či úbytek teplotního průměru v projekčním období. Tento postup by v případě srážek vedl ke zcela nereálnému scénáři kvůli ztrátě nulových hodnot, proto je srážkový faktor počítán jako podíl středních hodnot modelových dat a měřené srážky jsou modifikovány multiplikativně:

$$\begin{aligned}\delta &= \mu(X_{MP}) / \mu(X_{MC}) \\ x_T &= \delta x_O\end{aligned}\tag{3.2}$$

Modifikace může být provedena za celé období najednou, nebo je možné spočítat modifikační faktory zvlášť pro jednotlivé fáze roku. Druhý případ je častější, obvyklé je dělení ročního cyklu na sezóny nebo měsíce, méně často i na jednotlivé dny. To platí i pro korekční metody popsané v dalších kapitolách, kdy je v rámci daného členění roku pro každou roční fázi kalibrována individuální transferní funkce.

Ze vztahů (3.1) a (3.2) je možné odvodit následující vlastnosti výsledných modifikovaných souborů. Pro teploty platí

$$\begin{aligned}\mu(X_T) &= \mu(X_O) + \delta \\ \sigma^2(X_T) &= \sigma^2(X_O)\end{aligned}$$

Pro srážky pak

$$\begin{aligned}\mu(X_T) &= \delta\mu(X_O) \\ \sigma^2(X_T) &= \delta^2\sigma^2(X_O)\end{aligned}$$

Předpokládejme, že kontrolní běh modelu je kvalitní aproximací současného klimatu v tom smyslu, že $\mu(X_{MC}) = \mu(X_O)$. Potom pro střední hodnoty modifikovaných souborů srážek i teplot platí, že $\mu(X_T) = \mu(X_{MP})$. Pokud tedy zvolený model dobře vystihuje střední hodnotu sledované veličiny při simulaci současného stavu, zajišťuje přírůstková metoda dobrou shodu středních hodnot výsledného modifikovaného souboru a projekce modelu pro budoucí období.

Rozptyl modifikovaného souboru teplot se oproti výchozím měřeným datům nezmění. U srážek ke změně dojde, avšak relativní míra rozptýlení zůstává i v případě srážek stejná – pro variační koeficient platí $cv(X_T) = cv(X_O)$. Vzhledem k invarianci rozptylu k posunutí zmiňuje Kalvová et al. (2002) alternativní způsob aditivní modifikace teplot. Ten spočívá v separaci deterministické složky časové řady pozorovaných dat a následné korekci podle vztahu $x_T = \bar{x} + \delta(x_O - \bar{x})$, kde \bar{x} je průměrná hodnota deterministické složky pro daný kalendářní den.

Přírůstková metoda je hojně využívaným nástrojem k odvození klimatických scénářů, její použití nalezneme např. v Déqué (2007), Graham et al. (2007), Moore et al. (2008) a v mnoha dalších studiích. Je ale intuitivně zřejmé, že případná změna v průměrných hodnotách sledované veličiny by se v reálném klimatickém systému projevila složitěji a komplexněji než pouhým posunem datového souboru. Navíc modifikovaná srážková data přímo přebírají některé důležité vlastnosti výchozího měřeného souboru, např. frekvenci a distribuci srážkových dní.

Metody, popsané v dalších kapitolách již nespočívají v modifikaci pozorovaných hodnot o klimatický signál, ale ve skutečné korekci modelových hodnot z projekčního období. Společným rysem nejpoužívanějších metod je MOS přístup ke kalibraci – korekční funkce je kalibrována s využitím kontrolního běhu modelu, simulujícího zpravidla několik dekád druhé poloviny 20. století, a měřených dat z tohoto období. Dalším společným prvkem je využití modelových srážek jako jediného prediktoru, metody s touto charakteristikou bývají označovány jako přímé. Použití srážek jako jediného prediktoru vychází z předpokladu, že modelové srážky v jistém smyslu sdružují všechny ostatní relevantní prediktory. Tuto teorii podporuje také skutečnost, že při objektivním hledání vhodných modelových prediktorů v rámci nepřímých metod jsou často ve výsledku právě srážky prediktorem s nejvyšší vahou a významem (Themessl et al., 2011).

Lineární korekce

Stejně jako u přírůstkové metody se i zde jedná o lineární transformaci s pomocí korekčních faktorů, odvozených ze středních hodnot datových souborů. Korekční faktory jsou ale v tomto případě odvozeny z rozdílu měřených a modelových dat v kontrolním období a posléze aplikovány na modelová data z projekčního období. Neindikují tedy změnu modelované veličiny, ale již jednoduchým způsobem vyjadřují chybu modelu při simulaci současného klimatu. Teploty a srážky jsou i zde korigovány rozdílným způsobem, ze stejných důvodů jako v případě přírůstkové metody. Aditivní korekce teplot tedy probíhá podle vztahů

$$\begin{aligned}\delta &= \mu(X_O) - \mu(X_{MC}) \\ x_T &= x_{MP} + \delta\end{aligned}\tag{3.3}$$

srážky jsou korigovány multiplikativně podle

$$\begin{aligned}\delta &= \mu(X_O) / \mu(X_{MC}) \\ x_T &= \delta x_{MP}\end{aligned}\tag{3.4}$$

Korekční faktory bývají odvozovány zpravidla pro každý měsíc. Z funkcí (3.3) a (3.4) vyplývají tyto vlastnosti korigovaných souborů. Pro teploty

$$\begin{aligned}\mu(X_T) &= \mu(X_{MP}) + \delta \\ \sigma^2(X_T) &= \sigma^2(X_{MP})\end{aligned}$$

pro srážky pak

$$\begin{aligned}\mu(X_T) &= \delta \mu(X_{MP}) \\ \sigma^2(X_T) &= \delta^2 \sigma^2(X_{MP})\end{aligned}$$

Obecným cílem korekce je potlačení systematických odchylek modelových dat a dosažení lepší shody modelového výstupu s pozorovanými hodnotami. Účinnost korekce však samozřejmě nelze hodnotit pomocí korigovaného souboru z projekčního období, neboť zde nejsou k dispozici měřené hodnoty. Ty nemohou být pro srovnávací účely ani odhadovány, protože není jasné, zda model dobře vystihl vývoj klimatického systému a situaci dále komplikuje zatemňující vliv nestacionarity na korekční funkci. Proto při hodnocení korekčních metod bývají často zkoumány vlastnosti souboru, který vznikne aplikací korekční funkce na modelová data z kalibračního období, tj. na data ze kterých byla funkce odvozena. Tento test umožňuje zhodnotit míru shody mezi korigovanými a měřenými daty bez rušivého vlivu nestacionarity, tj. jedná se o užitečné kritérium při vzájemném porovnávání korekčních metod, i když samozřejmě nepodává žádnou informaci o výsledku korekce v projekčním období. Také je možné nahlížet na něj jako na zhodnocení korekce při hypotetické situaci nulové nestacionarity v obou časových řadách. V dalším textu budou vlastnosti korekčních metod posuzovány tímto způsobem.

Pokud tedy použijeme lineární korekční funkce (3.3) a (3.4) na kalibrační modelová data, zjistíme, že pro multiplikativní i aditivní variantu platí $\mu(X_T) = \mu(X_O)$. Lineární transformace tedy koriguje chybu střední hodnoty. Pro centrální momenty už ale rovnost nenastává, tzn. $\sigma^2(X_T) \neq \sigma^2(X_O)$. Aditivní korekce dokonce rozptyl nemění vůbec, pro teploty tedy zůstává $\sigma^2(X_T) = \sigma^2(X_{MC})$.

Lineární metodu použili např. Lenderink et al. (2007) při hydrologických projekcích klimatické změny na řece Rýnu.

Lineární korekce s úpravou frekvence srážkových dní

Tuto metodu publikovali Schmidli et al. (2006) pod názvem „local intensity scaling“ (LOCI). V původní práci byla metoda koncipována jako statistický downscaling srážkových dat z globálních klimatických modelů, je však frekventovaně využívána také pro korekci chyby regionálních modelů. Stejně jako u předchozí metody se jedná o lineární transformaci korigující chybu střední hodnoty, která je v tomto případě navíc adaptována na problém mrhnutí, tzn. upravuje počet suchých dní a odstraňuje přebytek drobných srážkových úhrnů.

Jednoduchá lineární transformace podle (3.4) zajišťuje rovnost středních hodnot $\mu_T = \mu_O = \mu$. Označíme počet srážkových dní v souboru jako n . Efekt mrhnutí regionálních modelů obvykle způsobí, že $n_M \gg n_O$. Střední hodnota obecně zastupuje jednotlivá čísla při výpočtu jejich součtu, z čehož pro celkové množství srážkové vody plyne, že $\mu n_M \gg \mu n_O$. Tento nepoměr může být příčinou znehodnocení výsledku dopadových studií, kdy jsou korigovaná data používána např. pro výpočet hydrologických bilancí apod.

Úprava počtu srážkových dní probíhá prostřednictvím tzv. prahových hodnot. Prahovou hodnotu měřených dat označíme ξ_0 a každá hodnota $x_0 > \xi_0$ představuje srážkový den v souboru X_0 . Pokud jsou používána měřená data z jednotlivých meteorologických stanic, obvykle bývá $\xi_0 = 0$. Při použití měřených dat interpolovaných na pravidelný grid (nebo re-analýz) je však možné a účelné nastavit $\xi_0 > 0$, protože taková data vykazují obdobný efekt mrhnutí jako data modelová. Prahová hodnota modelových dat je hodnota ξ_M taková, že četnost jejího překročení v souboru X_{MC} je stejná jako četnost srážkových dní v souboru X_0 .

Korekční faktor je definován jako

$$\delta = \frac{\mu(X_0; x_0 > \xi_0) - \xi_0}{\mu(X_{MC}; x_{MC} > \xi_M) - \xi_M} \quad (3.5)$$

Vlastní korekce pak probíhá podle

$$x_T = \max(0; \xi_0 + \delta(x_{MP} - \xi_M)) \quad (3.6)$$

Analýza souboru vzniklého korekcí kalibračních modelových hodnot je snadná pro případ, kdy $\xi_0 = 0$. Ze způsobu odvození korekčního faktoru (3.5) je zřejmé, že δ je kladné reálné číslo. Pokud je tedy modelová hodnota $x_{MC} < \xi_M$, potom je výraz $\delta(x_{MC} - \xi_M)$ záporný a podle korekčního vzorce je $x_T = 0$. Soubory X_T a X_0 mají tedy stejný počet srážkových dní. Rovnost $\mu(X_T) = \mu(X_0)$ pak již přímo plyne ze vztahu (3.6) a vlastností střední hodnoty. Pro $\xi_0 = 0$ tedy metoda koriguje chybu střední hodnoty a počtu srážkových dní.

Pro $\xi_0 > 0$ je situace komplikovanější. Pro každé $x_{MC} < \xi_M$ je výraz $\delta(x_{MC} - \xi_M)$ samozřejmě stále záporné číslo, ale jde o číslo libovolně malé ve své absolutní hodnotě. Není tedy zaručeno, že výraz $\xi_0 + \delta(x_{MC} - \xi_M)$ bude v tomto případě také záporný. Počet srážkových dní v souborech X_T a X_0 se tedy obecně nerovná a nerovnají se ani střední hodnoty těchto souborů. V každém případě však dochází k výraznému zlepšení těchto ukazatelů v porovnání se syrovými modelovými daty X_{MC} , jak bylo autorem tohoto textu ověřeno na několika pokusných souborech. Uvedená analýza vychází z korekční rovnice ve tvaru (3.6), publikované v originální práci Schmidli et al. (2006). Rovnosti středních hodnot a frekvence srážkových dní pro $\xi_0 > 0$ lze snadno dosáhnout modifikací postupu ve tvaru

$$\delta = \frac{\mu(X_O; x_O > \xi_O)}{\mu(X_{MC}; x_{MC} > \xi_M)} \quad \text{a posléze } x_T = \begin{cases} \delta x_{MP} & x_{MP} > \xi_M \\ 0 & x_{MP} \leq \xi_M \end{cases}$$

Mocninná korekce

Lineární transformace obecně nekorigují chybu v rozptýlení modelových hodnot. Při aditivní korekci ke změně rozptylu modelových hodnot nedojde vůbec. Multiplikační korekce, používaná v případě srážek, sice rozptyl korigovaného souboru mění, ale protože je korekční faktor δ definován pomocí středních hodnot (viz definice 3.4), nedojde v případě rozptylu ke shodě mezi korigovanými a měřenými soubory. Lze snadno odvodit, že pokud bychom v případě multiplikační korekce v definici korekčního faktoru nahradili střední hodnoty směrodatnými odchylkami, došlo by naopak k přesné shodě rozptylů měřeného a korigovaného souboru. Tato záměna by ale jistě vedla k destrukci shody středních hodnot, proto tato úvaha v praxi nemá smysl. Střední hodnota je považována za základní ukazatel, neboť se z něj dá usuzovat na celkovou změnu zkoumané veličiny. Rozptýlení dat ale hraje důležitou roli při podrobnějším rozboru potenciální změny klimatu, proto vyvstává potřeba korekční funkce se schopností simultánní úpravy střední hodnoty i rozptylu. Z toho důvodu se zavádí mocninná korekce ve tvaru

$$x_T = ax_{MP}^b \quad (3.7)$$

Odvození parametrů a, b není již tak přímočaré jako v případě lineární korekce. Označíme $x^* = x_{MC}^b$ a nalezneme kořen rovnice

$$g(b) = \frac{\sigma(X^*)}{\mu(X^*)} - \frac{\sigma(X_O)}{\mu(X_O)} \quad (3.8)$$

Funkce $g(b)$ je spojitá, pro hledání jejího kořenu je k dispozici mnoho numerických metod, vystačíme zde např. s půlením intervalu. Výsledná hodnota b tedy zajišťuje rovnost variačních koeficientů $cv(X^*) = cv(X_O)$. Parametr a potom nalezneme tak, aby korigoval střední hodnotu souboru X^* . Jeho výpočet je tudíž totožný jako v případě multiplikačního faktoru podle (3.4), tzn. $a = \mu(X_O)/\mu(X^*)$.

Pro variační koeficient platí (na rozdíl od rozptylu), že $cv(aX) = cv(X)$ pro jakékoliv $a > 0$. To znamená, že $cv(X_T) = cv(X^*) = cv(X_O)$. Vzhledem ke způsobu odvození a také platí, že $\mu(X_T) = \mu(X_O)$. Ze současné rovnosti variačních koeficientů a středních hodnot plyne, že $\sigma^2(X_T) = \sigma^2(X_O)$, mocninná transformace tedy simultánně koriguje střední hodnotu i rozptyl výsledného souboru.

Parametry a, b pro mocninnou korekci lze odvodit i jiným způsobem. Protože mocninná funkce ve tvaru $y = ax^b$ je pro kladná x monotónní, znamená to, že $F_X(x) = F_Y(y)$, tzn. hodnoty x a y představují stejný kvantil náhodných veličin X a Y . Protože cílem korekce je podobnost korigovaných a měřených dat, je přirozené požadovat shodu nějakého vybraného korigovaného a měřeného kvantilu. Pokud tedy ze souborů X_{MC} a X_O vybereme odpovídající si kvantily q_M a q_O , získáme tím argument a funkční hodnotu transformační funkce. Protože hledáme dva parametry, provedeme tento výběr dvakrát a dostaneme soustavu rovnic ve tvaru

$$\begin{aligned} aq_{M1}^b &= q_{O1} \\ aq_{M2}^b &= q_{O2} \end{aligned}$$

Z této soustavy už lehce dopočítáme, že

$$b = \frac{\ln\left(\frac{q_{O2}}{q_{O1}}\right)}{\ln\left(\frac{q_{M2}}{q_{M1}}\right)}$$

Odpovídající kvantily lze vybrat prostým setříděním obou souborů a následným výběrem prvků na stejné relativní pozici. Kvalita výsledných parametrů je tudíž závislá na tom, jak dobře výběry reprezentují skutečné skryté rozdělení pravděpodobnosti. Obě uvedená odvození použili Leander a Buishand (2007) s konstatováním, že získané parametry korekční funkce se u těchto postupů příliš neliší.

Shabalova et al. (2003) použili při korekci srážkových sum zajímavé odvození parametrů, které rovněž vede na mocninný tvar korekční funkce. Jejich postup bude nyní stručně popsán, protože má úzký vztah ke korekční metodě kvantilového mapování uvedené v další části tohoto textu. Myšlenka je založena na předpokladu, že transformované veličiny je možno dobře popsat Weibullovým rozdělením pravděpodobnosti s hustotou definovanou předpisem

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^\beta} \quad (3.9)$$

pro $x \geq 0$. Parametry tohoto rozdělení α, β jsou kladná reálná čísla a mají význam měřítka a tvaru. Distribuční funkce rozdělení má tvar

$$F(x) = 1 - e^{-(x/\alpha)^\beta} \quad (3.10)$$

kvantilová funkce potom

$$F^{-1}(p) = \alpha(-\ln(1-p))^{1/\beta} \quad (3.11)$$

Nejdříve odhadneme parametry rozdělení souboru měřených dat (α_O, β_O) a modelových dat z kontrolního období (α_M, β_M) . Opět budeme vyžadovat monotónní korekci, tzn. argument a výsledek transformace představují stejné kvantily v rámci svých rozdělení. Zároveň je vyžadována podobnost souborů X_T a X_O . Tyto požadavky je možné zapsat jako

$$F_O(x_O) \approx F_T(x_T) = p = F_M(x_M)$$

Měřenou hodnotu můžeme podle (3.10) a (3.11) vyjádřit pomocí modelové hodnoty jako

$$x_O = \alpha_O(-\ln(1-p))^{1/\beta_O} \approx \alpha_O \left(-\ln\left(e^{(-x_M/\alpha_M)^{\beta_M}}\right)\right)^{1/\beta_O} = \alpha_O \left(\frac{x_M}{\alpha_M}\right)^{\beta_M/\beta_O} = x_T$$

Korekce pak probíhá podle mocninné funkce (3.7) s parametry

$$\begin{aligned} a &= \alpha_O \alpha_M^{-(\beta_M/\beta_O)} \\ b &= \beta_M/\beta_O \end{aligned} \quad (3.12)$$

Odvození parametrů založené na Weibullově rozdělení použili Shabalova et al. (2003) pro korekci desetidenních srážkových sum při hydrologickém modelování průtoků na řece Rýn. Vícedenní srážkové sumy mohou být předmětem korekcí právě při projekcích povodní a extrémních průtoků na větších tocích, protože tyto jevy bývají spjaty spíše s déletrvajícimi dešti než s okamžitými událostmi. Předpoklad Weibullova rozdělení může být pro vícedenní sumy považován za příliš svazující, proto je možné při hledání parametru b upravit rovnici (3.8) tak, aby $cv(\sum X^*) = cv(\sum X_O)$, stejným způsobem je potom upraveno hledání

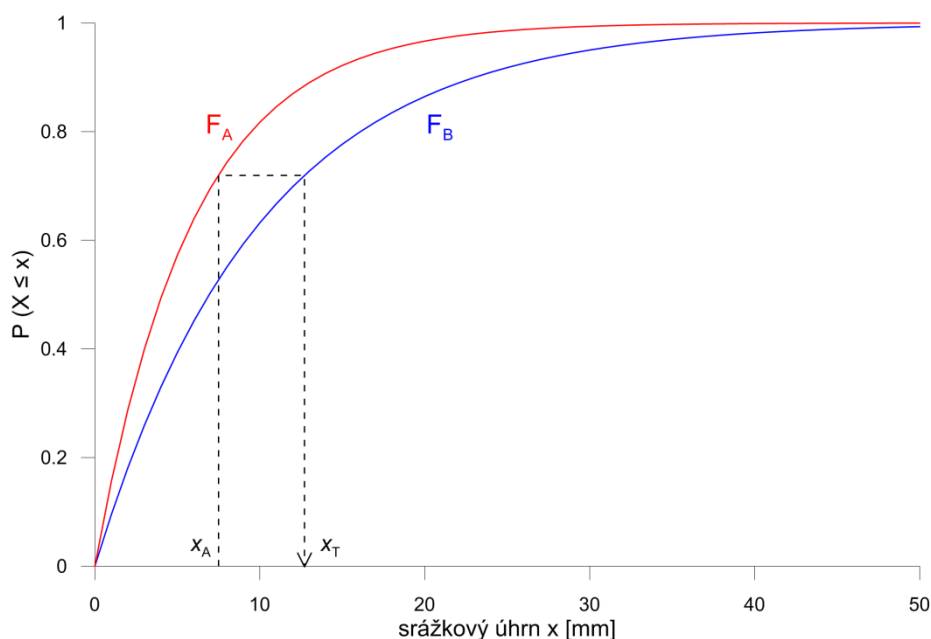
parametru a . Výsledná korekce potom zachovává střední hodnotu a rozptyl pro vícedenní sumy.

Kvantilové mapování

V oblasti korekce chyb klimatických modelů je kvantilové mapování zřejmě nejpoužívanější metodou. Jeho princip je velmi jednoduchý. Mějme dvě náhodné veličiny X_A a X_B , jejichž pravděpodobnostní rozdělení jsou popsána distribučními funkcemi F_A a F_B . Formule mapování kvantilů z rozdělení A na rozdělení B má tvar

$$x_T = F_B^{-1}(F_A(x_A)) \quad (3.13)$$

Celý proces je pro větší názornost schematicky znázorněn na obr. 3.4.



Obr. 3.4 Princip kvantilového mapování.

Z definice (3.13) je patrné, že pro jakékoliv η platí

$$P(X_T \leq \eta) = F_A(F_A^{-1}(F_B(\eta))) = F_B(\eta)$$

To znamená, že formule kvantilového mapování způsobí identitu distribučních funkcí veličin X_T a X_B . Z toho také vyplývá identita celkových rozdělení pravděpodobnosti, která jsou distribučními funkcemi jednoznačně určena. Navíc platí, že korekce podle (3.13) je vždy jednoznačná, protože ať jsou rozdělení veličin X_A a X_B jakákoliv, jejich distribuční funkce jsou vždy neklesající s oborem hodnot $H = \langle 0, 1 \rangle$.

Pro potřebu korekce modelových dat má formule kvantilového mapování tvar

$$x_T = F_O^{-1}(F_{MC}(x_{MP})) \quad (3.14)$$

Z výše uvedeného rozboru vyplývá, že ve výsledku dojde k přesné shodě rozdělení korigovaných a pozorovaných dat. Rovnají se hustoty rozdělení, distribuční funkce a s tím i veškeré obecné i centrální momenty, které rozdělení pozorovaných hodnot má. Tato celková shoda je také příčinou výsadního postavení kvantilového mapování mezi všemi ostatními

metodami korekce. Kvantilové mapování je v literatuře označováno mnoha způsoby, např. „probability mapping“ (Ines a Hansen, 2006), „quantile-quantile plot“ (Déqué, 2007), „quantile mapping“ (Piani et al., 2010a, Wilcke et al., 2013), nebo „histogram equalization“ (Rojas et al., 2011).

3.3 Ostatní (nepřímé) korekční metody

Výše uvedené metody (lineární korekce, mocninná korekce a kvantilové mapování) patří v oblasti opravy srážkových dat mezi nejpoužívanější. Vedle těchto metod, jejichž společným znakem je použití modelových srážek jako jediného prediktoru, existují tzv. metody nepřímé, které jako prediktorů využívají i dalších veličin. Tento přístup je obvyklý v oblasti statistického downscalingu, který se ale soustřeďuje spíše na nalezení vztahů mezi velkoprostorovým a lokálním klimatem a transferní funkce obvykle odvozuje perfect-prognosis postupem, viz např. Wilby a Wigley (1997), Schmidli et al. (2007). Kalibrace vztahů pro korekce regionálních modelů probíhá obvykle MOS přístupem. V následujícím textu jsou pro ilustraci příklad zmíněny dvě nepřímé metody.

Mnohonásobná lineární regrese ve tvaru $x_T = \alpha + \sum_i \beta_i x_i$ je přístupem obvyklým také při postprocessingu předpovědi počasí. Celý proces spočívá ve výběru vhodných prediktorů a v následném odhadu koeficientů lineární kombinace. Koeficienty jsou obvykle hledány metodou nejmenších čtverců. Výběr prediktorů představuje závažnější problém. Prediktorem mohou být jednak veličiny přímo simulované modelem, výběrem vhodných srážkových prediktorů pro statistický downscaling se zabývají např. Wilby a Wigley (2000). Další možností je jako prediktorů využít tzv. hlavních komponent. Analýza hlavních komponent je součástí použité metodiky v kapitole 7, kde je také uveden její stručný rozbor.

Obecně lze dodat, že korekce založená na vícenásobné lineární regresi redukuje rozptyl časové řady v důsledku nižší variability prediktorů. Z toho důvodu je k výsledku někdy přidáván umělý šum, který odpovídá rozptýlení původních dat nevysvětlenému regresním modelem. Šum bývá počítán jako reziduum po odečtení regresní funkce od původních dat.

Metoda analogů je další možnou nepřímou metodou. Její princip spočívá v tom, že cirkulační charakteristiky vygenerované modelem pro určitý den jsou porovnávány s historickými záznamy počasí. V rámci tohoto porovnání je nalezen nejbližší stav atmosféry (analog) a modelové lokální proměnné jsou nahrazeny měřenými záznamy z nalezeného analogu. Existují různé modifikace tohoto postupu, kdy je např. analog vybírán náhodně mezi několika nejpodobnějšími stavy, což může zvýšit variabilitu korigovaných srážek. Metodu analogů je možné vnímat jako speciální případ techniky resamplingu, který je (vzhledem k podmíněnosti atmosférickými stavy) vhodný pro downscaling či korekci chyb. Slabinou této metody je nutnost disponovat rozsáhlými archivy počasí, což je obzvláště limitující v případě extrémních hodnot. Stěžejním bodem celého postupu je pak hodnocení podobnosti stavů atmosféry. Kromě cirkulačních proměnných jsou zde využívány i hodnoty hlavních komponent, viz např. Zorita a von Storch (1999).

3.4 Empirické porovnání metod

V literatuře je možné nalézt několik prací, zabývajících se porovnáním nejpoužívanějších korekčních metod pro denní srážková data. Studie se liší výběrem metod a způsobem jejich testování, nicméně společným cílem těchto prací bylo empiricky zhodnotit a porovnat míru shody mezi korigovanými a pozorovanými daty, dosaženou jednotlivými metodami.

Themessl et al. (2011) porovnávali vícenásobnou lineární regresi, metodu analogů, lineární korekci doplněnou o úpravu frekvence srážek (LOCI) a kvantilové mapování. Testování metod probíhalo systémem cross-validace s pomocí poměrně krátkých (jedenáctiletých) časových řad. Kritérii pro zhodnocení podobnosti pozorovaných a korigovaných dat byly obvyklé míry polohy (průměr, medián), rozptyl, průměrná intenzita srážkových dní a frekvence (četnost) srážek. Dále byla posuzována schopnost korigovaných řad reprodukovat vysoké srážkové úhrny, což bylo realizováno porovnáním kvantilů q_{95} . Výsledkem studie bylo konstatování, že přímé metody (LOCI a kvantilové mapování) v podstatě ve všech ohledech převyšují metody nepřímé, přičemž nejlepšími výsledky bylo dosaženo s pomocí kvantilového mapování. Z porovnání nepřímých metod vyšla lépe metoda analogů, nejméně kvalitní výsledky přinesla v této studii vícenásobná lineární regrese.

Teutschbein a Seibert (2012) posuzovali pouze metody přímé, konkrétně jednoduchou lineární transformaci, LOCI, mocninnou transformaci, kvantilové mapování a pro úplnost i přírůstkovou metodu. Použili výstupy z jedenácti regionálních modelů a měřená data za období 1961 – 1990. Testování probíhalo bez rozdělení na kalibrační a validační období, korekce byly aplikovány přímo na kalibrační data. Autoři použili značný počet kritérií pro posuzování shody korigovaných a měřených dat, např. průměrné hodnoty, kvantily q_{10} a q_{90} , rozptyl, variační koeficient, intenzitu a frekvence srážkových dní a nakonec i celkovou shodu distribučních funkcí. Výsledkem studie bylo zjištění, že ačkoliv všechny metody dobře korigují střední hodnoty, v případě ostatních kritérií je jejich úspěšnost proměnlivá. Při zohlednění všech kritérií bylo nakonec nejlepší metodou uznáno kvantilové mapování.

Berg et al. (2012) poukázali na to, že pokud s pomocí histogramů sestrojíme korekční funkci kvantilového mapování, je její výsledný průběh obvykle dosti podobný přímce. Korekční funkci kvantilového mapování proto regresně aproximovali lineární funkcí a tu porovnávali s klasickou multiplikativní lineární korekcí podle (3.4). Výsledkem studie bylo sice zjištění, že funkce odvozená z kvantilového mapování dosahuje v průměru menší chyby, přesto si autoři všímají podobnosti výsledků dosažených oběma metodami.

Teutschbein a Seibert (2013) empiricky hodnotili vliv nestacionarity na účinnost korekčních metod. Jak autoři zmiňují, rozdělit datové řady na kalibrační a validační období není při hodnocení korekčních metod příliš běžnou praxí, na rozdíl např. od hydrologie. Důvodem může být potřebná délka datových souborů pro podchycení přirozené variability srážek. V tomto případě autoři disponovali 40-letými časovými řadami měřených a modelových dat, se kterými provedli tzv. Differential Split Sample Test (DSST). Ten spočíval v tom, že data byla po jednotlivých letech rozdělena podle celkové sumy měřených srážek na sušší a vlhčí polovinu, tyto poloviční datové soubory pak představovaly kalibrační a validační období. Korekčními metodami, zahrnutými do testování byly lineární korekce, LOCI, mocninná korekce, kvantilové mapování a přírůstková metoda, testovací kritéria obsahovala obvyklé popisné statistiky. Nejlepší metodou bylo shledáno kvantilové mapování, dále pak mocninná transformace a lineární metody. Přírůstková metoda byla vyhodnocena pro nestacionární podmínky jako nejméně vhodná.

3.5 Vyjádření pravděpodobnostního rozdělení srážek

Kvantilové mapování bylo ve všech srovnávacích studiích vyhodnoceno jako nejpřesnější a nejúčinnější korekční metoda, i když rozdíly mezi jednotlivými metodami byly v některých případech nevelké. Z dostupné literatury lze také soudit, že je metodou nejpoužívanější. Na rozdíl od ostatních metod ale kvantilové mapování nepředstavuje celistvý návod, jak při korekci postupovat. Je to proto, že korekční formule (3.14) je velmi obecná a neříká nic o tom, jakým způsobem odhadovat hustotu rozdělení. Vyžaduje pouze schopnost alespoň numerického výpočtu distribuční a kvantilové funkce.

V případě teplot je situace jednoduchá, všeobecně akceptovaným modelem pravděpodobnostního rozložení je zde normální rozdělení. Srážky jsou z tohoto pohledu mnohem komplikovanější náhodnou veličinou. Ze své podstaty jsou srážkové úhrny veličinou spojitou. Je známo, že základní rozdíl mezi diskrétními a spojitými veličinami spočívá v tom, že v případě spojitých veličin je pravděpodobnost jakékoliv konkrétní hodnoty nulová. Smysl má pouze sledovat, s jakou pravděpodobností se spojitá veličina pohybuje v určitém intervalu. Srážkové úhrny jsou z tohoto pohledu veličinou atypickou. Nula (suchý den) je ve srážkových řadách výjimečnou hodnotou, vyskytuje se opakovaně a počet suchých dní a jejich rozmístění v rámci časové řady je významnou součástí popisu srážek. Z tohoto důvodu jsou v některých případech nuly při korekcích ošetřovány zvlášť, např. aplikací prahových hodnot, a korekce jsou posléze prováděny pouze s nenulovými prvky.

Jak již bylo uvedeno v předchozích kapitolách, vyznačují se denní srážkové úhrny hustotou pravděpodobnosti s charakteristickým exponenciálním poklesem (viz obr. 3.1), v této obecné charakteristice se shodují měřené i modelové srážky. Měřená data obvykle obsahují zhruba polovinu či více nulových hodnot, ovšem i po odstranění nul ze souboru vykazuje výsledná křivka zbylých nenulových dat exponenciální pokles.

Obecně jsou pro vyjádření hustoty pravděpodobnosti k dispozici modely parametrické a neparametrické. V případě parametrického vyjádření je logické, že vybíráme z modelů, které umožňují zmíněný exponenciální pokles vystihnout. V následujících bodech je stručně uveden přehled používaných parametrických vyjádření.

1. *Exponenciální rozdělení* je základním modelem s hustotou definovanou předpisem

$$f(x) = \lambda e^{-\lambda x} \quad (3.15)$$

pro $x \geq 0$, parametr λ je kladné reálné číslo. Pro modelování srážek toto rozdělení použili např. Shoji a Kitaura (2006), Monjo et al. (2012) a další autoři.

2. *Smíšené exponenciální rozdělení* je vážený součet více exponenciálních rozdělení s hustotou ve tvaru

$$f(x) = \sum_i \theta_i f_i(x) \quad (3.16)$$

kde pro váhy θ_i platí $\sum_i \theta_i = 1$ a $f_i(x)$ jsou exponenciální hustoty ve tvaru (3.15) s parametry λ_i . Pro modelování srážek je obvyklý součet dvou exponenciálních rozdělení, který použil např. Schoof (2008). Podle některých autorů (např. Wilks, 2011) se jedná o nejlepší aproximaci denních srážkových úhrnů, Woolisher a Roldan (1982) poukazují zejména na schopnost tohoto rozdělení vystihnout vysoké srážkové úhrny.

3. *Logaritmicko-logistické rozdělení* je podle svého názvu používáno pro ty veličiny, jejichž logaritmus má logistické rozdělení. Hustota má tvar

$$f(x) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{(1 + (x/\alpha)^\beta)^2} \quad (3.17)$$

pro $x \geq 0$. Parametry α, β jsou kladná reálná čísla a mají význam měřítka a tvaru. Rozdělení se používá v hydrologických studiích pro modelování průtoků a srážek, dále také při analýzách životnosti populací a organismů v ekologii. Pro popis srážek ho použili např. Monjo et al. (2014).

4. *Gama rozdělení* je jedním ze syntetických Pearsonových rozdělení. Bývá používáno jako univerzální rozdělení nezáporných náhodných veličin. V případě denních srážkových úhrnů se jedná zřejmě o nejpoužívanější, všeobecně akceptovaný model. Hustota tohoto rozdělení je definována předpisem

$$f(x) = \frac{(x/\beta)^{\alpha-1} \exp(-x/\beta)}{\beta \Gamma(\alpha)} \quad (3.18)$$

pro $x \geq 0$. Parametry α, β jsou opět kladná reálná čísla a definují tvar a měřítko křivky hustoty. Při $\alpha = 1$ přechází gama rozdělení na rozdělení exponenciální. Rozborem literatury lze nalézt množství aplikací gama rozdělení na srážkové úhrny přímo ve spojitosti s kvantilovým mapováním, příkladem mohou být práce Baigorria et al. (2007), Block et al. (2009), Teutschbein a Seibert (2012) a mnoho dalších. Je třeba ale poznamenat, že proti automatickému používání gama rozdělení na denní srážky se staví Vlček a Huth (2009), kteří na testech provedených s pomocí dat z 90 meteorologických stanic upozornili na odklony části dat od gama rozdělení.

5. *Weibullovo rozdělení* pro modelování srážek použili např. Shabalova et al. (2003), Castellvi et al. (2004) a další autoři. Hustota je definována vztahem (3.9). Při hodnotě parametru $\beta = 1$ přechází v exponenciální rozdělení.

Uvedený přehled samozřejmě nepokrývá kompletní výčet používaných parametrických modelů, poukazuje ale na to, že všechny používané křivky umožňují vystihnout typický tvar hustoty s exponenciálním poklesem.

Kromě teoretické (parametrické) křivky je možné rozdělení popsat i způsoby neparametrickými. Nejjednodušší možností představuje histogram. Jeho nevýhodou je ale subjektivní volba počátku a šířky intervalu. Z výpočetního hlediska je problémem také jeho nespojitost a obecně nemožnost získat derivace hustoty. Sofistikovanější přístup k neparametrickému popisu rozdělení pravděpodobnosti nabízí tzv. jádrové odhady hustoty („kernel density estimators“), které odstraňují některé nevýhody histogramu.

K jádrovým odhadům se lze od histogramů jednoduše dobrat. Vzorec pro výpočet hustoty při použití histogramu je $f(x) = n_x / (nh)$, kde h je šířka intervalů, n je celkový počet hodnot a n_x je počet hodnot ve stejném intervalu jako x . Histogram je možné interpretovat tak, že body na ose posuneme do středů příslušných intervalů s_i a na každý bod umístíme obdélník o rozměrech $h \times (1/nh)$. Histogram je pak součtem těchto dílčích obdélníků. Jednotlivé obdélníky se středem s_i lze zapsat jako funkci

$$g(x) = \frac{1}{nh} w\left(\frac{x - s_i}{h}\right)$$

kde w je váhová funkce

$$w(x) = \begin{cases} 1 & |x| \leq 0.5 \\ 0 & \text{jinak} \end{cases}$$

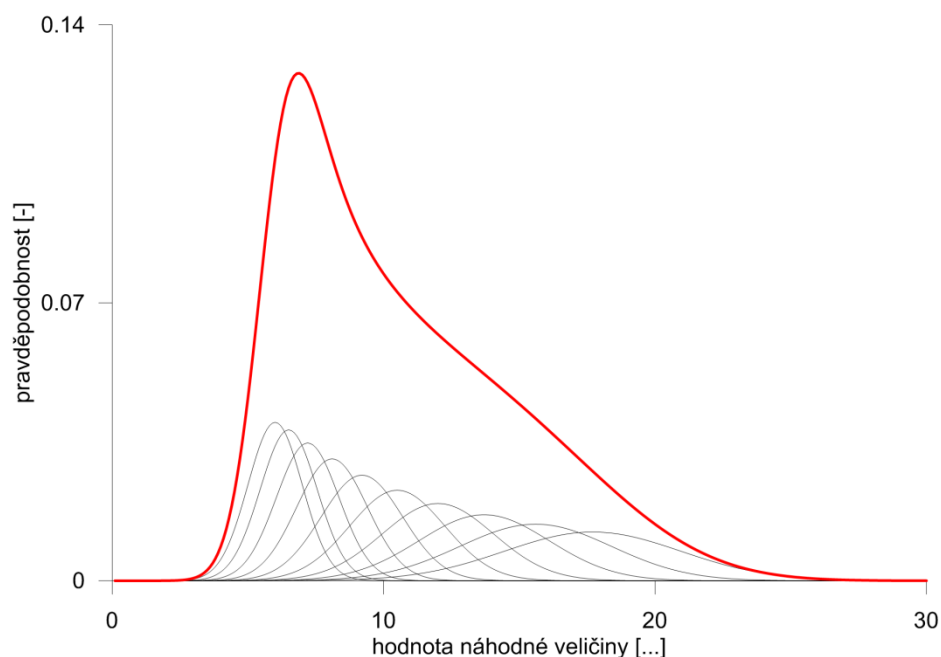
Pokud nebudeme body x_i posouvat do středů intervalů, ale položíme obdélníky do místa jejich původní polohy, dostaneme odhad hustoty ve tvaru

$$f(x) = \frac{1}{nh} \sum_{i=1}^n w\left(\frac{x - x_i}{h}\right)$$

Tato funkce je označována jako tzv. naivní odhad („naive estimator“), viz Silverman (1986). Nahradíme-li váhovou funkci $w(x)$ obecnou funkcí $K(x)$, dostaneme jádrový odhad hustoty ve tvaru

$$f(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (3.19)$$

Funkce $K(x)$ je tzv. jádro (kernel), obvykle se jedná o symetrickou hustotu, vyhovující podmínce $\int_{-\infty}^{\infty} K(x) dx = 1$. Parametr h bývá označován jako šířka okna (smoothing parameter, bandwidth). Můžeme si všimnout, že váhová funkce $w(x)$ použitá v definici naivního odhadu je hustotou spojitého rovnoměrného rozdělení na intervalu $\langle x - 0.5, x + 0.5 \rangle$, naivní odhad je tedy i z tohoto pohledu speciálním příkladem jádrového odhadu.



Obr. 3.5 Ilustrace jádrového odhadu (červená křivka) vytvořeného součtem dílčích Gaussovských kernelů s proměnlivou šířkou okna.

Z rovnice jádrového odhadu plyne, že pokud je jádrová funkce $K(x)$ hustotou, bude výsledný odhad také hustotou a navíc zdědí vlastnosti spojitosti a diferencovatelnosti po jádrové funkci. Nicméně ani při jádrovém odhadu se nevyhneme subjektivním volbám. Výběr konkrétní jádrové funkce má jen malý vliv na výsledný odhad, při výběru některého z běžně používaných hladkých a symetrických kernelů (Gaussian, Epanechnikov) budou výsledné odhady podobné. Volba šířky okna h je výrazně obtížnější problém, jehož popis ale přesahuje možnosti tohoto textu, podrobně se touto problematikou zabývá Silverman (1986). Je také

nutné zmínit, že při aplikaci jádrových odhadů na přirozeně ne-negativní náhodné veličiny, jakými jsou i srážkové úhrny, může symetrický kernel způsobit nenulové odhady hustoty i v záporných hodnotách (tzv. „boundary problem“). Rozborem a řešením všech problémů, které přináší aplikace jádrových odhadů na srážková data, se podrobně zabývají Rajagopalan et al. (1997).

Vyvstává tedy otázka, zda při kvantilovém mapování použít parametrické či neparametrické odhady hustoty. Obecně lze konstatovat, že aplikace jakékoliv parametrické křivky vždy alespoň částečně potlačí individuální charakteristiky daného výběru. Neparametrické odhady dovolují datům projevit se „samy za sebe“. To je vhodné obzvláště pro vícemodální rozdělení, kdy při parametrickém popisu by bylo nutné komplikovaně volit a odhadovat mix několika rozdělení. Nicméně pro účely korekce modelových dat bývá preferován parametrický přístup, přičemž argumenty podporující tuto volbu jsou v zásadě dvojího typu. První z nich spočívá v tom, že neparametrický odhad může snadno být přeurčený („overfitted“). Při příliš podrobném neparametrickém odhadu vykazuje křivka hustoty kolísání a nepravidelnosti, které vyjadřují spíše přirozenou variabilitu kolem skutečné hustoty, poplatnou aktuálnímu náhodnému výběru. Pro potlačení této přirozené variability by bylo potřeba extrémně dlouhých časových řad. Druhým argumentem často uváděným proti používání neparametrických odhadů je tzv. problematika „new extremes“. Pokud jsou pro odhad hustoty použity histogramy, je zřejmé, že $F_0^{-1}(1) \approx \max(X_0)$. To znamená, že při kvantilovém mapování podle (3.14) nelze obdržet korigovanou hodnotu vyšší než je maximální hodnota měřených dat, bez ohledu na modelové hodnoty z projekčního období.

Na obě tyto námitky však lze nalézt protiargumenty. Neparametrické vyjádření podle (3.19) dovoluje kontrolovat míru shlazení výsledné křivky volbou šířky okna h . Není tedy nutné obdržet přeurčenou hustotu. Parametrické metody samozřejmě přinášejí hladkou křivku samy o sobě, ale míru shlazení nelze na rozdíl od neparametrických odhadů nijak kontrolovat. Problém nových extrémů je spjat s použitím histogramu. Při aplikaci kernelové funkce s neomezeným definičním oborem (např. Gaussovský kernel, viz Rajagopalan et al., 1997) obdržíme

$$\lim_{p \rightarrow 1} F_0^{-1}(p) = \infty,$$

což znamená, že nové extrémy zde vznikat mohou. Jiným řešením tohoto problému je extrapolace korekční funkce, viz např. Boe et al. (2007).

3.6 Závěrečné poznámky

Jak již bylo zmíněno, kalibrace korekčních funkcí bývá prováděna obvykle odděleně pro jednotlivé sezóny či měsíce, kvůli kompenzaci chyby RCM při simulování ročního chodu srážek. Setkat se lze i s podrobnějším dělením roku např. na dny, kdy je kalibrační interval ohraničen klouzavým oknem kolem daného dne. Je však třeba vzít v úvahu, že rozdělením ročního chodu se snižuje počet hodnot ke kalibraci korekčních funkcí příslušejících jednotlivým intervalům, např. separace roku na měsíce sníží velikost datového souboru dvanáctkrát. Při použití kvantilového mapování spočívá kalibrace v popisu celého rozdělení pravděpodobnosti, který se při snižování velikosti kalibračního souboru může stát problematickým, celý proces se také stane náchylnějším na ovlivnění extrémními hodnotami. Empirickou analýzu nutné velikosti datového souboru pro robustní odhad parametrů rozdělení provedli Berg et al. (2012).

Použití prahových hodnot pro úpravu frekvence srážkových dní je sice primárně spjata s korekční metodou LOCI (kap. 3.2), ale v principu může být tento krok vložen před kteroukoliv korekční metodu. Při použití mocninné korekce způsobuje efekt mrhnutí identický problém jako v případě lineární korekce a aplikace prahových hodnot by tedy zřejmě měla být součástí postupu. Kvantilové mapování sice teoreticky řeší problém mrhnutí bez nutnosti zvláštního ošetření frekvence srážkových dní, ovšem v praxi je tento krok vhodný minimálně při parametrickém vyjádření hustoty rozdělení. Některé teoretické modely (např. gama rozdělení) vůbec nedovolují odhadovat parametry hustoty pro datové výběry obsahující nulové hodnoty. Jiné modely to sice umožňují, ovšem výsledné parametry (popisující velmi strmou křivku) dosahují často extrémních hodnot a celý proces kvantilového mapování je pak velmi citlivý na jakékoliv numerické chyby. Otázkou je rovněž, jakým způsobem prahové hodnoty odhadovat, tzn. jakou informaci z kalibračního do projekčního období přenášet. Prahová hodnota může být buď představována přímo hodnotou srážky, nebo může jít o hodnotu distribuční funkce, ze které je práh pro projekční období posléze dopočítán.

Závěrem je nutné poznamenat, že rozbořením literatury lze nalézt práce, stavící se ke korekcím chyb kriticky a vyslovující závažné námitky proti jejich používání při dopadových studiích. Pokud vezmeme za příklad hydrologickou studii, sled použitých procedur bude vypadat takto:

GCM → RCM → korekce chyb → hydrologický model

Je zřejmé, že jediným nefyzikálně založeným členem celého řetězce je právě korekce chyb. Přitom právě fyzikální podstata je hlavní devizou regionálních klimatických modelů v porovnání se statistickým downscalingem. Korigovaná data bezesporu vykazují dobrou shodu celkových statistických ukazatelů s pozorovanými hodnotami z jednotlivých meteorologických stanic. Shody je však dosaženo proto, že korekce jsou kalibrovány a prováděny individuálně pro jednotlivé prostorové body (modelové grid-boxy, meteorologické stanice) a jednotlivé meteorologické veličiny. Vliv takovéto procedury na korelační a kovarianční struktury modelových dat je nejasný, což může mít značný dopad na výsledky např. již zmiňovaných hydrologických studií. Dalším problematickým aspektem je omezená délka kalibračních intervalů, při které nedojde k podchycení celého spektra možných hodnot, a nakonec je zde otázka zatemňujícího vlivu nestacionarity časových řad mezi kalibrací a aplikací korekční funkce.

Pokud vezmeme v úvahu tyto problematické aspekty a pokud dále uvážíme výrazný vliv, jaký mají korekce na výslednou podobu klimatických scénářů a na změnu celého klimatického signálu, je jistá opatrnost zcela na místě. Ehret et al. (2012) za jediné zcela korektní řešení problému systematických chyb považuje zvýšení prostorového rozlišení modelů až na úroveň, zachycující konvektivní procesy v malém měřítku (rozlišení < 4 km) a dále sofistikovanější práce s modelovými ensembly na všech stupních modelového řetězce. V praxi je však takové rozlišení v současné době nedosažitelné a není ani zřejmé, zda by modely s takovým rozlišením produkovaly data použitelná přímo do dalších studií bez nutnosti post-processingu. Rovněž by zřejmě nebyla zcela eliminována nutnost používat v klimatických modelech různé typy parametrizací, které jsou samy o sobě statistické povahy a nutno podotknout, že statistické povahy je i vyhodnocování ensemblových simulací. Z těchto důvodů jsou korekce chyb pro naprostou většinu autorů jedinou možností pro vytvoření realistického klimatického scénáře a tudíž i nedílnou součástí klimatologických studií.

PRAKTICKÁ ČÁST

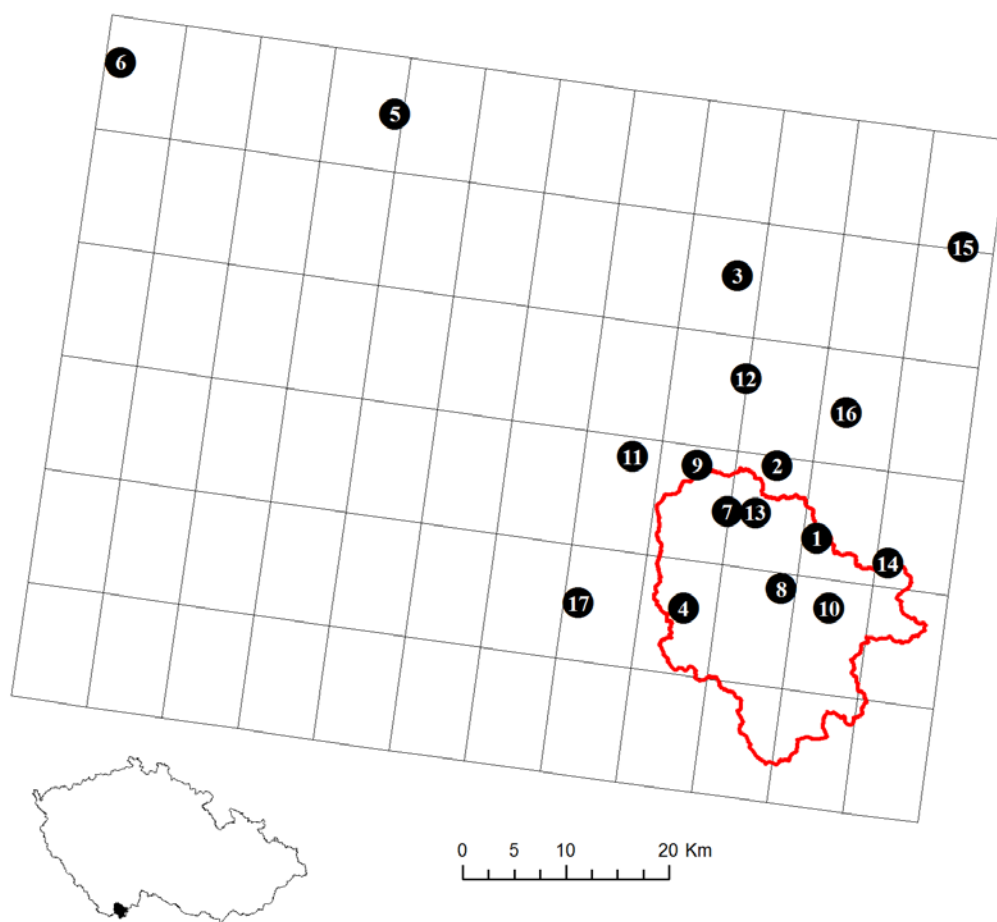
4 Použitá data

V disertační práci byly použity datové podklady dvojího typu. Prvním typem jsou měřená data ze sedmnácti meteorologických stanic v povodí řeky Malše a v jejím blízkém okolí, doplněná simulacemi regionálního klimatického modelu REMO (např. Jacob, 2001) s prostorovým rozlišením cca 0.1° zeměpisné šířky / délky. Měřená data tedy v tomto případě reprezentují úhrny zachycené jednotlivými srážkoměry, modelová data pak vyjadřují simulace srážkových úhrnů pro jednotlivé modelové grid-boxy. Měřená a simulovaná data byla pro analýzy k dispozici za časové období 1961–1997. Přehled stanic je uveden v tabulce 4.1.

Tab. 4.1 Přehled meteorologických stanic z povodí Malše.

	stanice	zeměpisná šířka	zeměpisná délka	nadm. výška (m n. m.)
1	Benešov nad Černou	48° 43′	14° 37′	681
2	Besednice	48° 47′	14° 33′	592
3	České Budějovice	48° 57′	14° 28′	388
4	Dolní Dvořiště	48° 39′	14° 27′	621
5	Husinec	49° 02′	13° 59′	536
6	Churáňov	49° 04′	13° 36′	1118
7	Kaplice	48° 44′	14° 29′	586
8	Malonty	48° 41′	14° 34′	688
9	Netřebice	48° 47′	14° 27′	616
10	Pohorská Ves	48° 40′	14° 38′	750
11	Přídolí	48° 46′	14° 21′	659
12	Římov	48° 50′	14° 29′	474
13	Soběnov	48° 44′	14° 32′	526
14	Staré Hutě	48° 43′	14° 42′	792
15	Třeboň	49° 00′	14° 46′	429
16	Trhové Sviny	48° 50′	14° 38′	468
17	Vyšší Brod	48° 38′	14° 18′	559

Mapa zájmové oblasti je na obrázku 4.1.



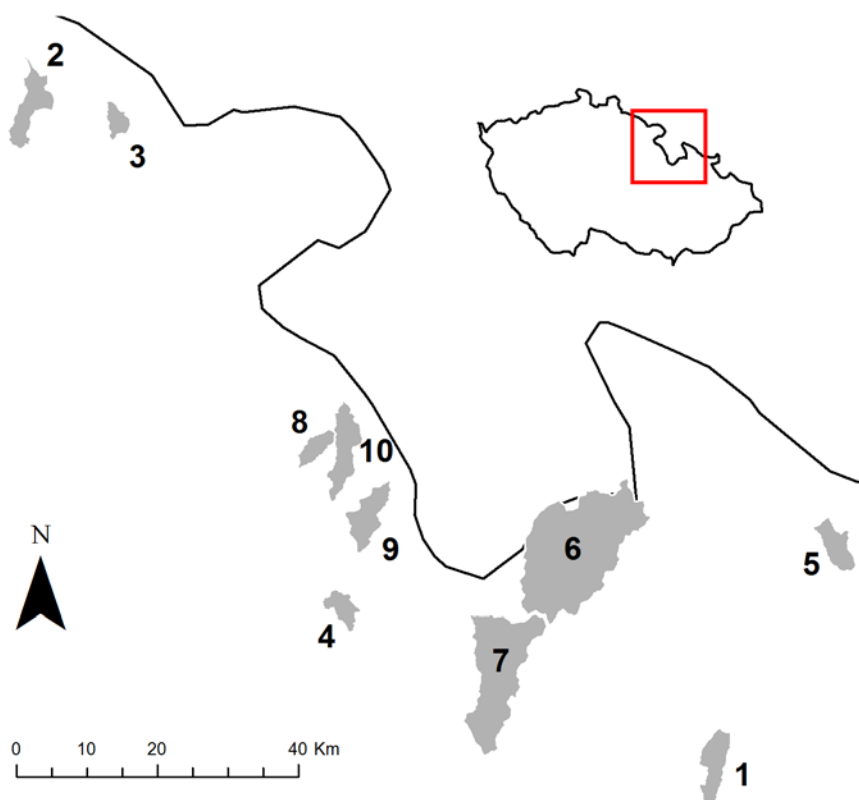
Obr. 4.1 Povodí Malše (červený obrys), použité meteorologické stanice (černé body, čísla odpovídají číslům stanic v tabulce 4.1), REMO grid-boxy a poloha povodí Malše v České Republice.

Druhý typ dat reprezentuje průměrné hodnoty srážkových úhrnů, připadající na plochy deseti vybraných povodí v severovýchodní části České republiky. Měřené průměry pochází z datového souboru poskytnutého Českým hydrometeorologickým ústavem, který byl získán interpolací z více než 500 meteorologických stanic na pravidelnou síť o prostorovém rozlišení 25×25 kilometrů. Originální staniční data byla nejprve podrobena kontrole, byly doplněny chybějící hodnoty a interpolace byla prováděna metodou IDW („inverse distance weighting“). Srážkové úhrny byly dále korigovány s ohledem na průměrnou nadmořskou výšku celého grid-boxu. Detailnější informace o odvození těchto dat jsou k dispozici v publikaci Štěpánek et al. (2011). K výpočtu průměrných úhrnů na plochu vybraných povodí byl použit vážený průměr grid-boxů zasahujících na plochu povodí s váhami odpovídajícími podílu plochy jednotlivých grid-boxů na ploše příslušného povodí. Simulace klimatického modelu pochází z projektu EURO-CORDEX (Giorgi et al., 2009), použita byla data z regionálního modelu ALADIN53 s rozlišením 0.11 stupně, který byl řízen globálním modelem CNRM-CM5. Také modelové srážkové úhrny reprezentují průměrné hodnoty na plochu jednotlivých povodí a k jejich odvození z originálních modelových dat bylo použito stejného váženého průměrování jako v případě měřených dat. Měřená i modelová data byla k dispozici pro období 1970 – 2005. Přehled povodí je uveden v tabulce 4.2.

Tab. 4.2 Přehled vybraných povodí.

ID	vodní tok	plocha povodí [km ²]
1	Huntava	27.0
2	Čistá	30.2
3	Babí	9.7
4	Potočnice	13.9
5	Opavice	20.7
6	Morava	217.2
7	Březná	126.5
8	Kněžská	16.0
9	Rokytenka	27.6
10	Zdobnice	72.2

Mapa povodí je na obrázku 4.2.



Obr. 4.2 Mapa vybraných povodí (šedé plochy) a poloha oblasti v České republice. Čísla u jednotlivých povodí odpovídají jejich identifikátorům (sloupec ID) v tabulce 4.2.

5 Lineární povaha korekce a podobnost metod

Tato kapitola navazuje na literární rešerši korekčních metod. Obsahuje úvahy autora práce doplněné praktickou analýzou metod s využitím měřených a modelových dat. Tyto úvahy vedou k poznatku o lineární povaze korekčních metod, objasňují její příčiny a poukazují na podobnost mezi jednotlivými korekčními metodami.

Analýza je omezena pouze na metody přímé, které jsou z uvedeného spektra nejpoužívanější a také se obvykle ve srovnávacích studiích jeví jako úspěšnější (viz kapitola 3.4). Při vyloučení přírůstkové metody zbývají tři základní postupy: lineární korekce, mocninná korekce a kvantilové mapování. Úprava frekvence srážkových dní spojená s metodou LOCI nebude v rámci této práce považována za samostatný postup, protože jak již bylo uvedeno, tento krok je možné (a účelné) vsunout před kteroukoliv z korekčních metod. Uvedené pořadí metod – lineární, mocninná a kvantilové mapování – odráží i úroveň sofistikovanosti od triviální lineární korekce dosahující pouze rovnosti středních hodnot, přes mocninnou metodu korigující také rozptyl modelových dat, až po kvantilové mapování, které přináší shodu celého rozdělení pravděpodobnosti.

Výchozím bodem pro následující úvahy jsou parametrická rozdělení používaná k popisu denních srážkových úhrnů, jejichž stručný přehled je uveden v kapitole 3.5. Jak již bylo uvedeno, pravděpodobnostní rozložení denních srážkových úhrnů vykazuje charakteristický exponenciální pokles a k jeho popisu jsou užívány teoretické modely, které dovolují tento pokles vystihnout. Nejjednodušším takovým modelem je pochopitelně exponenciální rozdělení. Pro přehlednost zopakujme, že toto rozdělení s parametrem λ je definováno hustotou ve tvaru

$$f(x) = \lambda e^{-\lambda x}$$

Z definice hustoty je možné vyjádřit distribuční funkci

$$F(x) = 1 - e^{-\lambda x} \quad (5.1)$$

a nakonec kvantilovou funkci ve tvaru

$$F^{-1}(p) = -\frac{1}{\lambda} \ln(1 - p) \quad (5.2)$$

Nyní je možné analyzovat vlastnosti kvantilového mapování s exponenciálně rozdělenými daty. Připomeňme, že obecná definice kvantilového mapování je dána formulí

$$x_T = F_O^{-1}(F_M(x_M)) \quad (5.3)$$

Pro exponenciálně rozdělené náhodné veličiny tedy podle (5.1), (5.2) a (5.3) dostáváme

$$x_T = -\frac{1}{\lambda_O} \ln(e^{-\lambda_M x_M}) = \frac{\lambda_M}{\lambda_O} x_M \quad (5.4)$$

Vztah (5.4) ukazuje, že kvantilové mapování s exponenciálně rozdělenými daty vede na lineární tvar korekční funkce. Z definice hustoty exponenciálního rozdělení je možné vyjádřit střední hodnotu tohoto rozdělení jako

$$\mu(X) = \int_0^{\infty} x \lambda e^{-\lambda x} dx = \frac{1}{\lambda} \quad (5.5)$$

Spojením (5.4) a (5.5) je obdržen konečný vztah

$$x_T = \frac{\mu(X_O)}{\mu(X_M)} x_M, \quad (5.6)$$

což je předpis pro lineární korekční metodu, v kapitole 3.2 uvedený jako vztah (3.4). Z této analýzy tedy plyne, že multiplikativní lineární korekce a kvantilové mapování jsou pro exponenciálně rozdělené náhodné veličiny zcela totožné korekční metody.

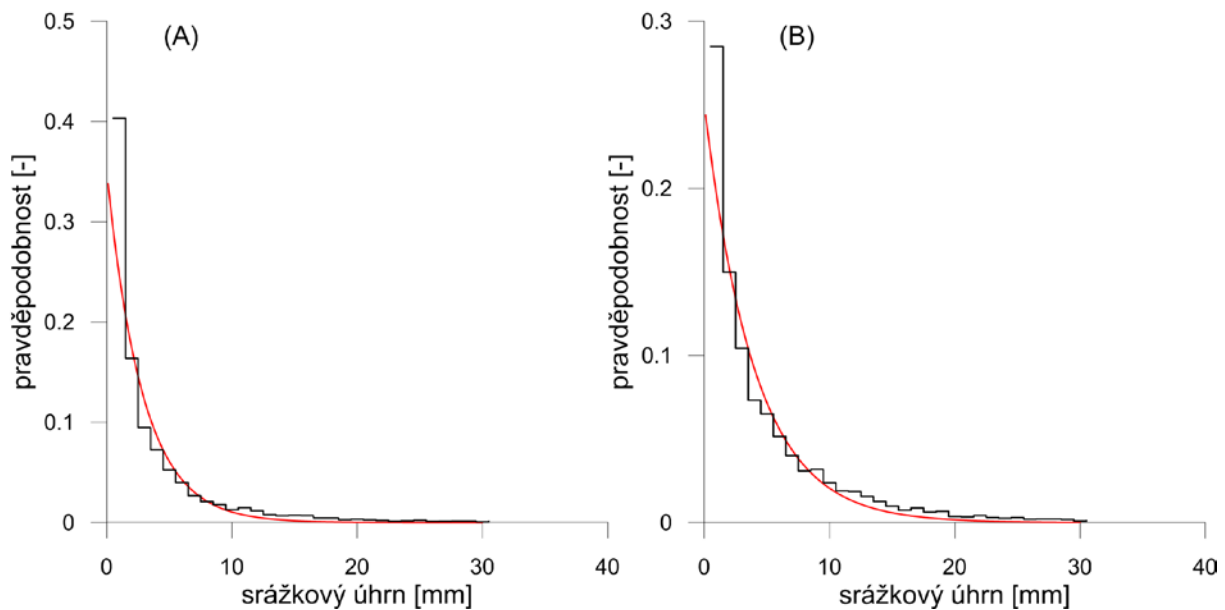
Z rozboru vlastností multiplikativní lineární korekce provedeného v kapitole 3.2 vyplývá, že tato metoda sice přináší shodu středních hodnot, v případě rozptylů však shoda nenastává. Pokud ale má být lineární korekční metoda zároveň speciálním případem kvantilového mapování, jak bylo právě ukázáno, musí být jejím výsledkem shoda celých rozdělení a rovnost rozptylů nutně musí nastat také. Tento rozpor lze objasnit výpočtem rozptylu exponenciálního rozdělení, při kterém je obdržen vztah

$$\sigma^2(X) = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \quad (5.7)$$

Pro rozptyl náhodné veličiny při multiplikativní korekci platí vztah $\sigma^2(aX) = a^2 \sigma^2(X)$. Při aplikaci tohoto vztahu na korekční formuli (5.6) a při současném uvážení (5.5) a (5.7) ihned vychází, že rovnosti rozptylů je dosaženo rovněž. Nakonec je možné doplnit další zajímavost exponenciálního rozdělení, kterou je hodnota variačního koeficientu. Ze vzorců pro střední hodnotu a rozptyl tohoto rozdělení plyne, že $cv = 1$ pro jakékoliv hodnoty parametrů, relativní míra rozptýlení se tedy korekcí nijak nemění.

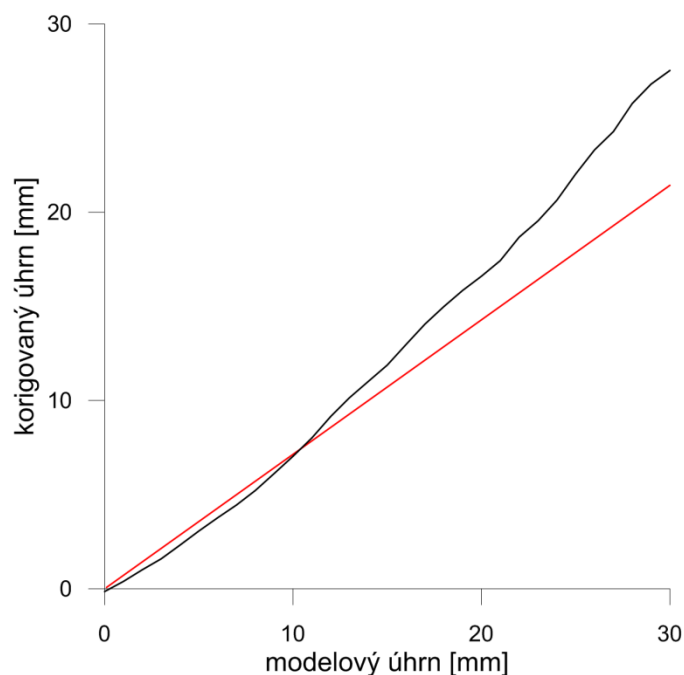
Z výše uvedeného tedy plyne, že pokud se při popisu srážkových úhrnů nedopustíme výrazné chyby použitím exponenciálního rozdělení, nedopustíme se výrazné chyby ani použitím lineární korekce a dosáhneme velmi přesné shody celých rozdělení pravděpodobnosti. Tento závěr lze inverzně interpretovat v tom smyslu, že chyba lineární korekční metody je úměrná nepřesnosti exponenciálního rozdělení použitého při popisu denních srážkových úhrnů.

Pro měřená srážková data z meteorologické stanice v Českých Budějovicích byl proveden odhad parametru exponenciálního rozdělení. Stejný odhad byl proveden také pro časově a prostorově odpovídající výstup z regionálního klimatického modelu REMO. Dále byly pro měřená i modelová data spočítány histogramy. Je nutné doplnit, že v případě měřených dat byly odhady hustoty prováděny pouze pro nenulové hodnoty. V případě modelových dat byla nejprve aplikována prahová hodnota k odstranění přebytku malých úhrnů a odhady byly provedeny pro zbylá nenulová data. Křivky rozdělení tedy popisují soubory, které by byly použity pro případnou kalibraci korekční funkce. Obrázek 5.1 ukazuje exponenciální hustoty spolu s histogramy pro měřená (A) a modelová (B) data.



Obr. 5.1 Histogramy (černé linky) a odhady exponenciálního rozdělení (červené linky) pro (A) měřená data ze stanice České Budějovice a (B) prostorově a časově odpovídající výstup modelu REMO.

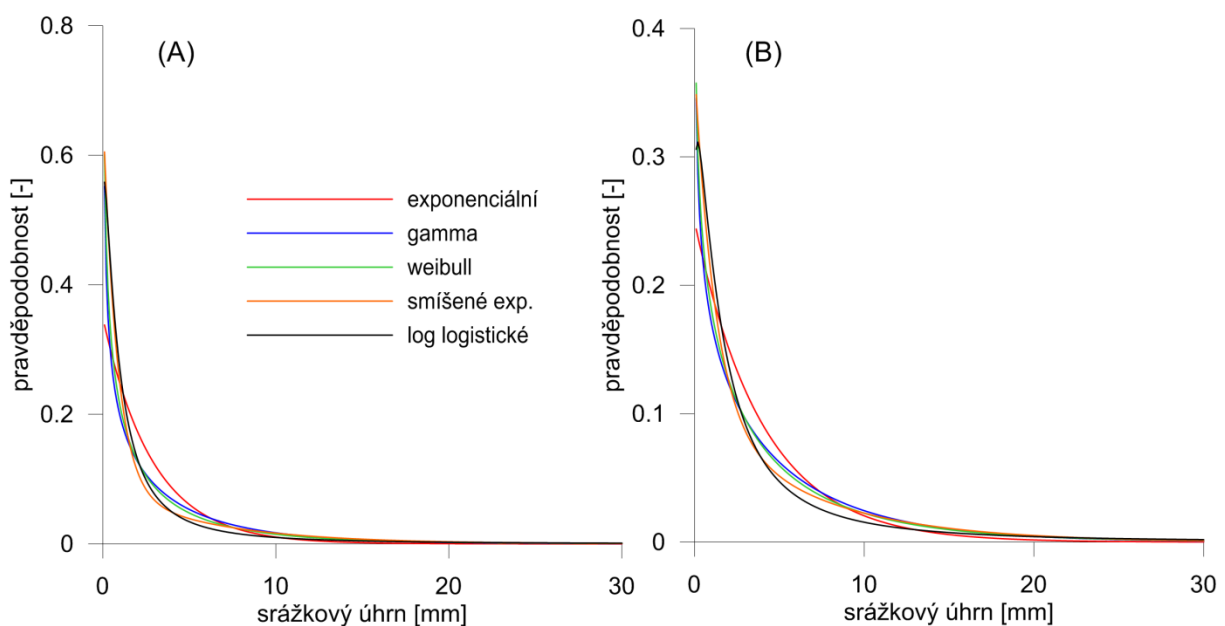
Tvar rozdělení obou souborů je jistě alespoň zhruba vystihnout histogramem, který je (i přes problémy se subjektivitou a nespojitostí) užitečným nástrojem analýzy dat. Pro obrázek 5.1 byly použity histogramy o šířce intervalu 1.0 mm, neboť nejlépe vystihují plynulý průběh hustoty bez rušivého kolísání, které nastává při použití užších intervalů. Z obrázku je patrné, že průběh histogramu je odhadem exponenciální hustoty v obou případech dobře vystihnout. Histogramové i exponenciální odhady hustot z obrázku 5.1 byly použity k odvození funkcí kvantilového mapování. Tyto funkce jsou zobrazeny na obrázku 5.2. Červeně vyobrazená funkce je odvozena z exponenciálních hustot (a je tedy funkcí lineární korekce, jak bylo ukázáno výše), černě vyobrazená funkce je numericky odvozena z histogramů.



Obr. 5.2 Funkce kvantilového mapování odvozené z hustot vyobrazených na obr. 5.1. Červeně vyobrazená funkce je odvozena z exponenciálních hustot, černě vyobrazená je odvozena z histogramů.

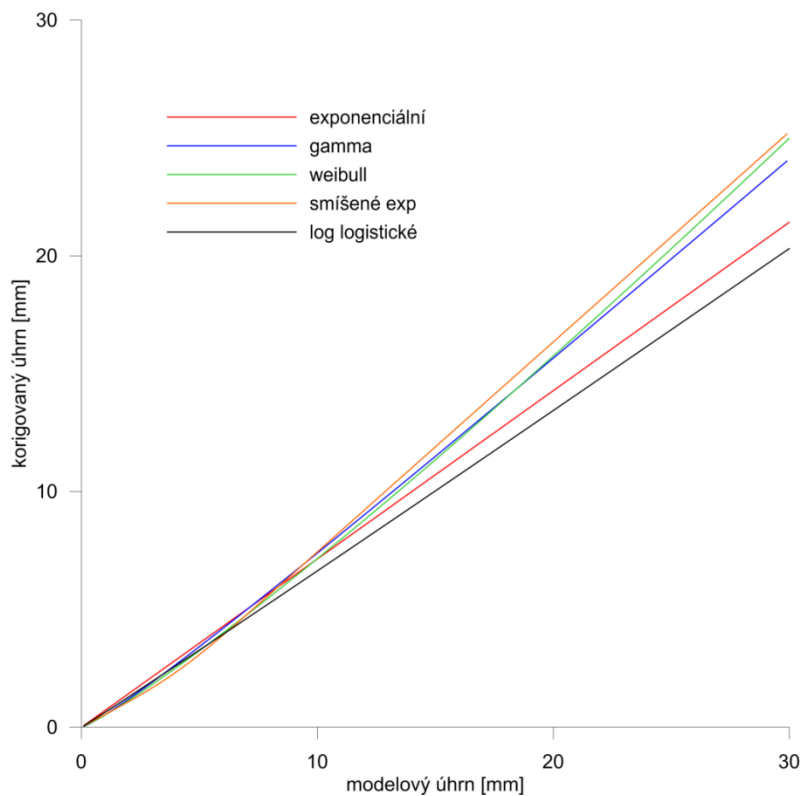
Z obrázku 5.2 je patrné, že funkce odvozená z histogramů má průběh dosti podobný lineární funkci odvozené z exponenciálních hustot. Je nutné podotknout, že neparametrická histogramová funkce vede k přesné shodě rozdělení korigovaných a měřených dat. Její lineární průběh a podobnost s lineární korekční funkcí jsou dány exponenciálním tvarem hustoty rozdělení srážkových dat, který je obecně platný pro denní srážkové úhrny jakožto náhodnou veličinu. Stejně porovnání histogramových a lineárních korekčních funkcí bylo provedeno pro data z většiny stanic z tabulky 4.1, přičemž výsledky byly velmi podobné jako pro data z Českých Budějovic a podávaly stejnou informaci o lineární povaze kvantilového mapování.

Exponenciální rozdělení je nicméně příliš jednoduchým jednoparametrickým modelem a k popisu srážkových úhrnů se obvykle nevyužívá. Pro praktické využití tedy přicházejí v úvahu spíše plastičtější modely s více parametry, jejichž stručný přehled byl podán v kapitole 3.5, jedná se o smíšené exponenciální rozdělení, logaritmicko-logistické rozdělení, gama rozdělení a weibullovo rozdělení. Parametry uvedených rozdělení byly odhadnuty opět pro měřená data z meteorologické stanice v Českých Budějovicích a pro odpovídající výstupy z regionálního klimatického modelu REMO. Všechny odhady byly prováděny metodou maximální věrohodnosti. Obr. 5.3 obsahuje výsledné pravděpodobnostní křivky pro (A) měřená a (B) modelová data.



Obr. 5.3 Hustoty rozdělení srážkových úhrnů z Českých Budějovic podle různých parametrických modelů pro (A) měřená a (B) modelová data.

Z obrázku je patrné, že křivky jednotlivých typů rozdělení jsou si dosti podobné. Tato podobnost je logická, neboť se jedná o různé teoretické modely popisující stejný datový soubor. Mírný odklon nalezneme v případě měřených i modelových dat u exponenciálního rozdělení, což je dáno relativní jednoduchostí (a tedy i menší plastičností) tohoto rozdělení v porovnání s ostatními modely. Lze tedy očekávat, že funkce kvantilového mapování, odvozené z uvedených teoretických modelů, budou mít navzájem podobný průběh a zároveň se nebudou příliš lišit od funkce odvozené z exponenciálního rozdělení, tedy od funkce lineární korekce. Tato skutečnost je demonstrována na obrázku 5.4, který uvedené funkce znázorňuje.



Obr. 5.4 Korekční funkce kvantilového mapování odvozené z jednotlivých rozdělení na obrázku 5.3.

Z obrázku je patrné, že všechny korekční funkce mají silně lineární charakter. Odlišnost mezi jednotlivými funkcemi spočívá v mírných rozdílech ve sklonu jednotlivých přímek. Je ovšem nutné podotknout, že tyto odlišné sklony mohou způsobit rozdíly v korekci extrémních hodnot. Tento fakt způsobuje problém při posuzování extrémů v korigovaných datech, protože sklony korekčních přímek jsou zároveň citlivé na hodnoty odhadnutých parametrů rozdělení. Mírné odchylky v hodnotách parametrů způsobí sice pouze nepatrnou odchylku korekční funkce v oblastech kolem počátku, pro extrémní hodnoty jsou však změny větší. Proto je při posuzování srážkových extrémů v korigovaných datech na místě jistá opatrnost. Naproti tomu v okolí počátku jsou si korekční funkce velmi podobné. Pokud uvážíme, že na intervalu $\langle 0 - 10 \rangle$ leží zhruba 97 % celkové masy všech (nenulových!) hodnot, je zřejmé, že naprostá většina hodnot je korigována téměř totožně bez ohledu na použitý pravděpodobnostní model.

Funkce odvozená z logaritmicko-logistického rozdělení na obrázku 5.4 je lineární korekci nejpodobnější. Pro získání detailního teoretického popisu této funkce je možné provést stejnou analýzu kvantilového mapování jako v případě exponenciálního rozdělení. Hustota logaritmicko-logistického rozdělení s parametry α, β je dána definicí (3.17), distribuční a kvantilová funkce jsou pak ve tvaru

$$F(x) = \frac{1}{1 + (x/\alpha)^{-\beta}}$$

$$F^{-1}(p) = \alpha \left(\frac{p}{1-p} \right)^{1/\beta}$$

Dosažením těchto vztahů do formule kvantilového mapování (5.3) obdržíme korekční funkci ve tvaru

$$x_T = px_M^q, \quad (5.8)$$

kde $p = \alpha_0/\alpha_M^q$ a $q = \beta_M/\beta_0$

Tento rozbor tedy ukazuje, že kvantilové mapování ve spojitosti s log-logisticky rozdělenými daty vede na mocninný tvar korekční funkce. Nyní místo logaritmicko-logistického rozdělení použijeme Weibullovo rozdělení s parametry označenými opět jako α, β . Definice hustoty, distribuční a kvantilové funkce jsou dány vztahy (3.9), (3.10) a (3.11). Lze snadno ověřit, že kvantilové mapování s Weibullovým rozdělením nejen že vede opět na mocninnou korekci, ale dokonce ke zcela identickému vztahu (5.8) jako v případě logaritmicko-logistického rozdělení.

Dále je možné si povšimnout, že korekční formule (5.8) je také shodná se vztahem, ke kterému dospěli Shabalova et al. (2003), viz kapitola 3.2, vztah (3.12). Myšlenkový postup, který byl při jeho odvození použit, je založen na tom, že mocninná korekční funkce je pro kladná x_T monotónní a tedy hodnoty modelových a korigovaných dat představují stejné kvantily v rámci svých rozdělení, tzn.

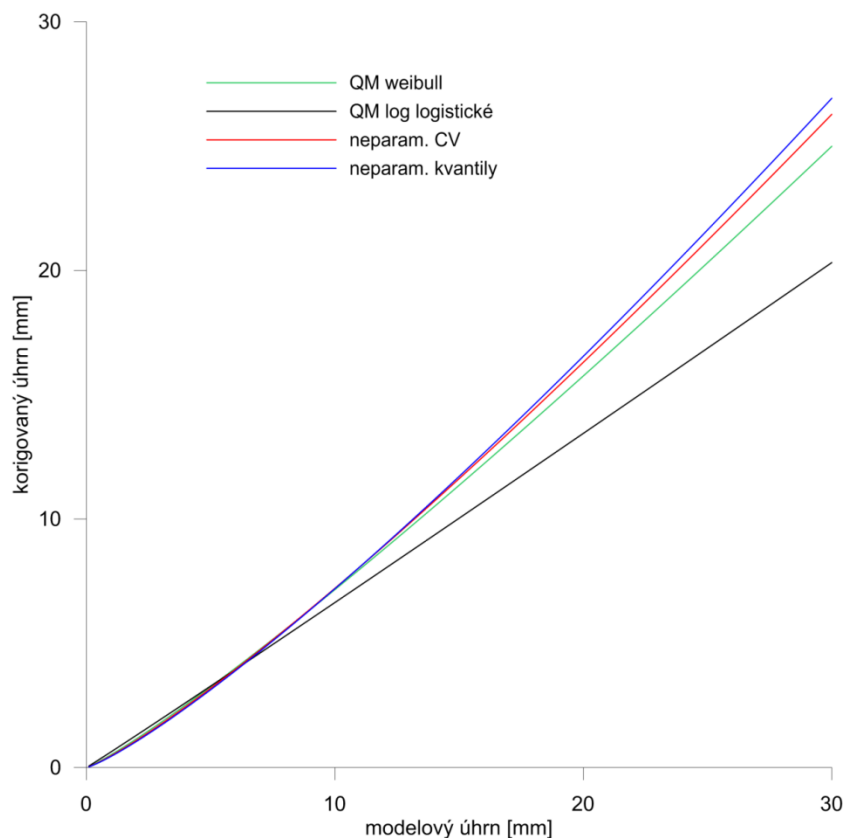
$$F_T(x_T) = p = F_M(x_M)$$

Pokud vyžadujeme přesnou shodu distribučních funkcí korigovaných a měřených dat a funkci F_T v rovnici nahradíme funkcí F_0 , obdržíme přímo formuli kvantilového mapování (5.3).

Poslední poznámka k mocninné transformaci se týká gama rozdělení. Leander a Buishand (2007) kalibrovali parametry mocninné transformace bez vazby na konkrétní parametrický model rozdělení, jejich způsob kalibrace ale zaručuje současnou rovnost středních hodnot a rozptylů korigovaných a měřených dat, viz kapitola 3.2. Výběrová střední hodnota a rozptyl jsou však dostatečné údaje pro odhad parametrů gama rozdělení metodou momentů. Je tedy zřejmé, že tímto způsobem odhadu obdržíme pro korigovaná a měřená data zcela shodné parametry gama rozdělení. Tuto „rovnost“ je třeba brát poněkud s rezervou, nicméně zcela jistě dojde k velmi dobré shodě.

Z výše uvedeného rozboru je patrné, že mezi korekčními metodami existuje mnoho vazeb a že za určitých podmínek tyto metody přecházejí jedna v druhou. Tyto podmínky vzájemného splývání metod jsou dány uvažovaným typem rozdělení pravděpodobnosti při kvantilovém mapování. Protože jsou si křivky rozdělení pravděpodobnosti pro jednotlivé teoretické modely velmi podobné, jak ukazuje obrázek 5.3, jsou si podobné i korekční funkce všech metod, což je ilustrováno na obrázku 5.4. Protože zároveň jedna z nich je čistá přímka, je výsledný charakter korekce lineární bez ohledu na použitou metodu.

Závěrem této části práce bylo prošetřeno, zda se výrazněji odlišují parametry mocninné funkce odhadnuté z teoretických pravděpodobnostních modelů od parametrů odhadnutých některým ze dvou postupů prezentovaných v kapitole 3.2, které žádné předpoklady o pravděpodobnostním rozdělení nevyžadují. Připomeňme, že první z postupů je založený na rovnosti variačních koeficientů, druhý pak na shodě vybraných kvantilů, oba tyto postupy použili Leander a Buishand (2007). Obrázek 5.5 nabízí porovnání mocninné funkce získané z pravděpodobnostních modelů s funkcemi získanými uvedenými ne-parametrickými postupy (opět pro data ze stanice České Budějovice).



Obr. 5.5 Porovnání mocninných korekčních funkcí.

Jak lze očekávat, mocninné funkce jsou si podobné, pouze funkce založená na logaritmicko-logistickém rozdělení se odlišuje ve vyšších částech pravděpodobnostního spektra. K tomu lze poznamenat, že proces odvození funkce je značně citlivý na odhad parametrů pravděpodobnostního modelu. Při jiné metodě odhadu (např. použitím metody momentů) jsou výsledky mírně odlišné, jak bylo autorem práce ověřeno, nicméně lineární povaha a podobnost jednotlivých metod zůstávají zachovány.

Lze ještě dodat, že kvantilové mapování je běžně používanou korekční metodou i při korekci teplotních dat. V tomto případě vede rovněž k lineárnímu tvaru korekční funkce. To je dáno tím, že průměrné denní teploty jsou náhodnou veličinou dobře popsatelnou normálním rozdělením, přičemž konverze mezi dvěma normálními rozděleními je beze zbytku realizovatelná lineární funkcí, podrobněji viz např. Anděl (2007).

6 Nové lineární metody

6.1 Úvod

Jak již bylo v rámci literární rešerše několikrát zmíněno, jedním z problematických faktorů při aplikacích korekcí na modelová data je potenciální nestacionarita v časových řadách srážek. Při praktické aplikaci v rámci studií zkoumajících dopady klimatické změny jsou korekční funkce kalibrovány s použitím kontrolních běhů modelu, zahrnujících zpravidla několik dekád druhé poloviny dvacátého století, a posléze aplikovány na modelová data, simulující vývoj klimatického systému v budoucnosti. Korekční funkce jsou tedy aplikovány na datové soubory s odlišnými charakteristikami, než byly soubory použité k jejich kalibraci, což je dále umocněno tím, že modelové simulace jsou pro budoucí období zpravidla zatíženy vybraným scénářem vývoje koncentrace skleníkových plynů. Není tedy zcela zřejmé, že funkce v pozměněných podmínkách budou stále účinně korigovat chyby modelu. Navíc zůstává otázkou, zda sama chyba modelu je v čase stacionární, nebo se v odlišných podmínkách projevuje jinak.

Rozborem literatury lze nalézt několik studií, zabývajících se problematikou nestacionarity a jejího vlivu na účinnost korekčních metod. Výskyt nestacionarity ve srážkových řadách potvrdili např. Wilby (1997), Trenberth et al. (2003), aktuálně např. Chen et al. (2015). Maraun (2012) analyzoval efekt nestacionarity na korekce pomocí takzvané pseudo-reality, kdy měřená data pro kalibraci i následnou validaci korekcí nahradil klimatickou simulací, přičemž studie přinesla zjištění, že korekce mají v obecnosti pozitivní efekt na výsledky i při nestacionárních podmínkách. Teutschbein a Seibert (2013) porovnávali účinnost běžně používaných korekčních metod pod vlivem nestacionarity. Jejich studie již byla zmíněna v kapitole 3.4, na tomto místě lze stručně připomenout, že potvrdila kvantilové mapování jako obecně nejefektivnější korekční metodu.

Tato část disertační práce vychází ze závěrů předchozí kapitoly. Lineární povaha korekce ukazuje, že lineární funkcí lze dosáhnout velmi dobré shody mezi pravděpodobnostními rozděleními měřených a korigovaných dat. Použití lineární funkce je zároveň motivováno přesvědčením, že jednodušší funkce s menším počtem parametrů bude obecně méně náchylná k chybám způsobeným nestacionaritou než přesná funkce kvantilového mapování. Tento předpoklad motivoval několik autorů k náhradě korekční funkce její regresní přímkou, což lze nalézt např. v pracích Piani et al. (2010, b) nebo Berg et al. (2012). Další možností je použít ke korekci lineární korekční metodu, detailní rozbor v minulé kapitole ovšem ukázal, že lineární metoda dává přesné výsledky pouze pro exponenciálně rozdělená data.

Gama rozdělení je pro popis srážek nepochybně lepším modelem, což plyne například z faktu, že exponenciální rozdělení je jeho speciálním případem. Cílem této části disertační práce je proto odvodit lineární korekční funkci založenou na konverzi mezi dvěma gama rozdělenými datovými soubory. Je již předem zřejmé, že taková funkce nemůže být dokonalou korekcí v tom smyslu, že její aplikace na kalibrační modelová data nezpůsobí přesnou shodu s kalibračními měřenými daty. To je dáno tím, že přesná konverzní funkce (daná kvantilovým mapováním) se od lineárního průběhu odchyluje, jak je vidět např. na obrázku 5.4. Je ovšem otázkou, zda je schopnost naprosto přesné korekce kalibračních dat výhodná pro nestacionární podmínky. V následujících částech této kapitoly je popsáno odvození korekčních funkcí, následované jejich validací a porovnáním s jinými metodami korekce.

6.2 Odvození korekčních metod

Prahové hodnoty

Jak bylo uvedeno výše, metodika vychází z předpokladu, že pravděpodobnostní rozdělení srážkových úhrnů je dobře popsateľné pomocí gama rozdělení. Zopakujme, že hustota gama rozdělení s parametry α a β je dána formulí

$$f(x) = \frac{(x/\beta)^{\alpha-1} \exp(-x/\beta)}{\beta \Gamma(\alpha)}$$

pro $x \geq 0$, parametry $\alpha > 0$ a $\beta > 0$ mají význam tvaru a měřítka, symbol $\Gamma(\cdot)$ označuje gama funkci. Skutečnost, že datový soubor X má gama rozdělení s parametry α a β bude v dalším textu označována zkráceným zápisem ve tvaru $X \sim \Gamma(\alpha, \beta)$. Cílem odvození je nalézt takovou lineární korekční funkci, při které dojde k co nejlepší shodě hustoty korigovaných a měřených dat.

Prvním krokem při odvozování parametrů je adaptace na tzv. efekt mrholení v modelových datech, který byl popsán v kapitole 3.1 (vysoký počet malých srážkových úhrnů doprovázený malým počtem suchých dní). Tento problém je řešen stejně jako v metodě „linear intensity scaling“ (LOCI, kapitola 3.2) prostřednictvím tzv. prahových hodnot. Modelový soubor je pomocí prahové hodnoty nejprve transformován na soubor s obdobným počtem suchých dní jaký mají měřená data. Následné korekce jsou pak prováděny pouze na zbylých nenulových hodnotách.

Odlišné nakládání s nulovými a nenulovými hodnotami má kromě odstranění efektu mrholení ještě další praktickou motivaci. Pro odhad parametrů gama rozdělení je v této práci využívána metoda maximální věrohodnosti. Tato metoda dává obecně lepší odhady než metoda momentů, obzvláště pro rozsáhlejší soubory, nicméně její použití je možné pouze pro datové soubory neobsahující nuly. Jakákoliv nulová hodnota v datovém souboru totiž vede k nulové hodnotě věrohodnostní funkce a znemožňuje další výpočty. Proto je nutné omezit odhady na nenulové hodnoty. Kromě tohoto praktického důvodu je tu ještě další (teoretický) pohled na věc spočívající v nulové pravděpodobnosti diskretních hodnot při použití spojitého rozdělení (viz druhý odstavec, kapitola 3.5).

Protože je předpokládáno gama rozdělení, jsou nejprve odhadnuty parametry rozdělení nenulových hodnot modelových dat a prahová hodnota x^0 je odhadována ze vztahu

$$x^0 = F_M^{-1} \left(\frac{\Delta_O - \Delta_M}{1 - \Delta_M} \right) \quad (6.1)$$

kde F_M^{-1} značí kvantilovou funkci nenulových modelových dat, Δ_O a Δ_M značí podíly nul v měřených (soubor X_O) a modelových (soubor X_M) datech z kalibračního období. Následně jsou všechny modelové hodnoty nižší než prahová hodnota transformovány na nulu. Výsledkem této konverze je tedy soubor, v dalším textu označovaný jako X_M^0 , který má podíl nulových hodnot srovnatelný s měřenými daty.

Tento způsob odvození samozřejmě předpokládá, že modelová data efekt mrholení vykazují, tzn. že podíl nul Δ_O je vyšší než Δ_M . Tento předpoklad však nemusí být vždy splněn, obzvláště při použití hodnot interpolovaných na pravidelnou prostorovou mřížku může být podíl nul v měřených datech výrazně snížen. V takovém případě je nutné úvodní krok buď vynechat, nebo stanovit pro modelová i měřená data nějakou alternativní hodnotu, obvykle např. 0.1

mm, což je hodnota udávaná jako minimální úhrn, který jsou schopné meteorologické stanice zaznamenat.

Prahovou hodnotu je možné odvodit i jinými způsoby, např. setříděním datových souborů a vybráním modelové hodnoty na příslušné relativní pozici. Takováto modifikace však nevede k výraznějším rozdílům ve výsledných korigovaných datech, jak bylo autorem ověřeno. Protože je zde prezentovaná metoda založena na gama rozdělení, je k získání prahových hodnot využíván vztah (6.1).

Konverze z X_M na X_M^0 je užitečnou korekcí sama o sobě, v některých případech je možné již tímto krokem dosáhnout dobré shody mezi korigovanými a měřenými daty.

Korekce modifikací měřítka

K odvození lineární korekční funkce bylo využito faktu, že pokud platí $X \sim \Gamma(\alpha, \beta)$, potom $AX \sim \Gamma(\alpha, A\beta)$, kde A je reálné číslo. Multiplikační lineární transformace gama rozdělené náhodné veličiny mění měřítka jejího rozdělení. Je tedy možné nalézt takovou hodnotu parametru A , při které dojde k minimalizaci rozdílu mezi hustotami měřených a korigovaných dat. Velikost tohoto rozdílu je funkcí parametru A , v dalším textu bude označována jako účelová funkce (UF), v souladu s terminologií běžně užívanou v numerické matematice.

Pro vyjádření účelové funkce se nabízí několik způsobů, v této práci byl použit prostý integrál absolutního rozdílu měřené a modelové hustoty ve tvaru

$$UF(A) = \int_0^q \left| f_0(x) - \frac{(x/A\beta_M)^{\alpha_M-1} \exp(-x/A\beta_M)}{A\beta_M\Gamma(\alpha_M)} \right| dx \quad (6.2)$$

kde $f_0(\cdot)$ je hustota měřených dat, α_M a β_M jsou parametry rozdělení odhadnuté ze souboru X_M^0 a A je hledaný korekční parametr. Horní mez integrace, ve vzorci (6.2) označená jako q , byla stanovena jako vyšší z 99% kvantilů měřených a modelových dat. Cílem je tedy nalézt hodnotu A splňující

$$\frac{dUF}{dA} = 0$$

Pro minimalizaci výrazu (6.2) byla použita gradientní metoda, výsledkem procesu je optimální hodnota A , která je využita k transformaci modelových dat ve tvaru

$$x_T = \begin{cases} Ax_M & x_M > x^0 \\ 0 & x_M \leq x^0 \end{cases} \quad (6.3)$$

Korekční metoda podle vzorce (6.3) bude označena jako LM1.

Korekce modifikací měřítka a tvaru

V rámci korekce samozřejmě není třeba omezovat se na čistě multiplikační funkci, při zachování požadavku na lineární korekci je možné aplikovat obecnější funkci ve tvaru $x_T = Ax_M + B$. Pokud je zachováno vyjádření účelové funkce jako integrálu rozdílu hustot, její tvar nyní dostane podobu

$$UF(A) = \int_0^q \left| f_0(x) - \frac{(x/\beta_M^*)^{\alpha_M^*-1} \exp(-x/\beta_M^*)}{\beta_M^*\Gamma(\alpha_M^*)} \right| dx \quad (6.4)$$

kde parametry modelového (korigovaného) rozdělení α_M^* a β_M^* závisí na aktuálních hodnotách A a B . Hodnoty parametrů A a B byly opět hledány pomocí gradientní metody tak, aby bylo současně splněno

$$\frac{\partial UF}{\partial A} = 0, \quad \frac{\partial UF}{\partial B} = 0$$

Bylo třeba nahradit analytické derivace numerickými, což bylo provedeno jako náhrada tečny sečnou. Během minimalizačního procesu byl modelový datový soubor X_M^0 přepočítáván a parametry α_M^* a β_M^* byly odhadovány po každém kroku v kartézském systému $A \times B$. Minimalizace (6.4) vede k optimálním hodnotám A a B . Korekce modelových hodnot z projekčního období posléze probíhá ve tvaru

$$x_T = \begin{cases} Ax_M + B & x_M > x^0 \\ 0 & x_M \leq x^0 \end{cases} \quad (6.5)$$

Procedura (6.5) mění měřítko i tvar hustoty modelového souboru, v dalším textu bude označována jako LM2.

Závěrem metodologické části je nutné připojit technickou poznámku týkající se odhadu parametrů gama rozdělení metodou maximální věrohodnosti. Po vyjádření derivací věrohodnostní funkce a jejich položení rovno nule není možné nalézt analytické řešení výsledné soustavy a k hodnotám parametrů je třeba se dobrat numericky. V odborné literatuře je možné dohledat dva numerické způsoby (tzv. „Newton-Rapson method“ a „scoring method“), které vedou k přibližně stejným výsledkům. Celý proces je do detailu popsán v publikaci Choi a Wette (1969).

6.3 Validace a porovnání s ostatními metodami

Účinnost nově odvozených metod byla porovnána s běžně používanými metodami. Pro porovnání byla vybrána metoda kvantilového mapování, nadále označovaná jako QM, vzorec (3.14), a lineární korekce, v dalším textu označovaná jako LS – „linear scaling“, vzorec (3.4). Při praktické realizaci QM a LS byly použity stejné prahové hodnoty jako v případě nově odvozených metod. Ve spojení s aplikací prahových hodnot mají uvedené metody tvar

$$\text{QM: } x_T = \begin{cases} F_O^{-1}(F_M(x_M)) & x_M > x^0 \\ 0 & x_M \leq x^0 \end{cases}$$

$$\text{LS: } x_T = \begin{cases} x_M \frac{\mu(X_O)}{\mu(X_M)} & x_M > x^0 \\ 0 & x_M \leq x^0 \end{cases}$$

Lineární korekce je tedy v tomto provedení vlastně jakousi obdobou metody LOCI, popsané v kapitole 3.2, nicméně v dalším textu bude označována jako LS.

Pro testování metod byla použita srážková staniční data z povodí Malše spolu s výstupy modelu REMO, přehled stanic je uveden v tabulce 4.1. Pro kalibraci i validaci byla zvolena velmi krátká šestiletá období. Metody byly kalibrovány s použitím dat z let 1961–1966 a následně aplikovány na data z let 1992–1997. Kalibrace s takto časově omezenými daty by v reálné klimatologické studii byla nevhodná, protože šestileté řady zcela jistě nedokáží

zachytit kompletní přirozenou variabilitu na dané lokalitě. Pro demonstraci metod je ale právě toto účelem, neboť dojde k zesílení efektu nestacionarity mezi kalibračním a validačním obdobím, čemuž napomáhá i dlouhý časový interval mezi obdobími. Pro každou meteorologickou stanici a odpovídající výstup klimatického modelu byly s využitím kalibračních dat nalezeny parametry všech čtyř korekčních metod LM1, LM2, QM a LS. Metody byly posléze aplikovány na modelová data z validačního období a výsledné korigované soubory byly porovnány s měřenými validačními daty.

Jako testovací kritéria pro posouzení shody měřených a korigovaných dat byly určeny: střední hodnota (μ), rozptyl (σ^2), 90% kvantil (q^{90}) a celková shoda rozdělení (r). Posouzení celkové shody rozdělení bylo prováděno pomocí histogramů, aby výsledky nebyly ovlivněny předpoklady o typu rozdělení. Pro měřená i korigovaná data z validačního období byly spočteny histogramy o stejném počátku (0) a stejné šířce intervalu (1 mm). Chyba korigovaného rozdělení pak byla počítána podle vztahu

$$er = s \frac{1}{2} \sum_i |f_{0i} - f_{Ti}| \quad (6.6)$$

kde er je chyba rozdělení korigovaných dat, s je šířka histogramových intervalů, f_{0i} a f_{Ti} jsou hodnoty i -tého intervalu histogramů měřených a korigovaných dat. Výraz (6.6) může být snadno interpretován jako celkové množství odlišně umístěné hustoty při porovnání dvou histogramů. Hodnoty er se pohybují v rozmezí $\langle 0,1 \rangle$, kdy 0 znamená přesnou shodu histogramů, 1 naopak zcela nepřekrývající se histogramy. Přehled všech kritérií spolu s výpočtem chyby pro jednotlivá kritéria je uveden v tabulce 6.1.

Tab. 6.1 Přehled kritérií použitých pro zhodnocení chyby korekčních metod ve validačním období.

Kritérium	Výpočet chyby
μ	$e\mu = \mu_T - \mu_0$
σ^2	$e\sigma = \sigma_T^2 - \sigma_0^2$
q^{90}	$eq = q_T^{90} - q_0^{90}$
r	er , definice 6.6

6.4 Výsledky a diskuze

Tabulka 6.2 obsahuje kompletní výsledky pro všechny meteorologické stanice. Pro každou stanici a každé kritérium je uvedena chyba dosažená jednotlivými korekčními metodami. Zvýrazněné hodnoty označují nejmenší chybu, která byla v daném řádku (tedy pro danou stanici a dané kritérium) dosažena napříč spektrem korekčních metod. Chyby jsou uváděny i se znaménkem, záporné číslo znamená podhodnocení daného ukazatele, kladné číslo nadhodnocení (viz výpočet chyb, tabulka 6.1). Při porovnání chyb pro danou metodu je ale samozřejmě uvažována jejich absolutní hodnota.

Tab. 6.2 Chyby korekčních metod ve validačním období při korekcích výstupů modelu REMO na jednotlivé stanice na povodí Malše.

stanice		LS	QM	LM1	LM2	stanice		LS	QM	LM1	LM2
Benešov	eμ	-0.039	-0.033	-0.061	-0.029	Pohorská Ves	eμ	-0.175	-0.148	-0.185	-0.178
	eσ	-9.803	-3.437	-9.767	-7.515		eσ	-6.798	1.377	-6.611	-6.270
	eη	0.243	-0.602	0.236	0.459		eη	-0.928	-1.802	-0.909	-0.875
	er	0.255	0.231	0.212	0.207		er	0.212	0.229	0.176	0.171
Besednice	eμ	0.168	0.194	0.141	0.164	Přídolí	eμ	-0.025	0.040	-0.040	-0.024
	eσ	-6.145	5.361	-6.220	-3.860		eσ	-9.152	2.525	-9.112	-3.657
	eη	0.876	0.798	0.847	1.103		eη	0.721	0.197	0.720	0.883
	er	0.276	0.239	0.231	0.218		er	0.214	0.177	0.169	0.171
České Budějovice	eμ	0.005	0.031	-0.003	-0.044	Římov	eμ	-0.091	-0.072	-0.110	-0.036
	eσ	-7.244	3.601	-6.732	-4.975		eσ	-9.233	-3.749	-9.288	-8.140
	eη	0.415	0.198	0.510	0.605		eη	0.194	-0.400	0.170	0.341
	er	0.172	0.181	0.191	0.166		er	0.224	0.247	0.198	0.181
Dolní Dvořiště	eμ	-0.178	-0.150	-0.180	-0.138	Soběnov	eμ	0.095	0.111	0.088	0.076
	eσ	-9.264	-3.774	-9.123	-8.837		eσ	-8.105	5.527	-7.882	-5.772
	eη	-0.340	-1.400	-0.312	-0.259		eη	0.636	0.498	0.664	0.868
	er	0.195	0.248	0.171	0.169		er	0.173	0.156	0.171	0.144
Husinec	eμ	-0.263	-0.256	-0.287	-0.268	Staré Hutě	eμ	-0.081	-0.082	-0.097	-0.089
	eσ	-10.330	-6.228	-10.610	-5.462		eσ	-5.801	-0.401	-5.474	-4.572
	eη	-0.551	-0.800	-0.606	-0.484		eη	-0.295	-0.801	-0.265	-0.201
	er	0.181	0.181	0.172	0.154		er	0.196	0.203	0.135	0.140
Churáňov	eμ	0.083	0.170	0.106	0.077	Třeboň	eμ	0.216	0.299	0.192	0.219
	eσ	16.282	-0.376	17.303	12.196		eσ	1.624	18.770	1.256	2.784
	eη	-0.467	0.800	-0.383	-0.625		eη	0.344	-0.002	0.283	0.439
	er	0.174	0.214	0.194	0.237		er	0.206	0.212	0.199	0.173
Kaplice	eμ	0.040	0.079	0.021	0.037	Trhové Sviny	eμ	-0.228	-0.220	-0.244	-0.232
	eσ	-7.766	3.865	-7.888	-6.191		eσ	-6.820	-2.776	-6.734	-6.051
	eη	0.425	0.097	0.394	0.582		eη	-0.858	-1.201	-0.853	-0.780
	er	0.219	0.226	0.184	0.177		er	0.217	0.209	0.173	0.164
Malonty	eμ	0.062	0.099	0.046	0.065	Vyšší Brod	eμ	0.157	0.575	0.130	0.125
	eσ	-8.370	5.982	-7.632	-3.778		eσ	-0.358	4.374	-0.461	2.008
	eη	0.332	-0.201	0.440	0.738		eη	0.667	-0.102	0.636	0.856
	er	0.204	0.146	0.158	0.131		er	0.207	0.158	0.181	0.139
Netřebice	eμ	0.033	0.092	0.021	0.030						
	eσ	-6.935	8.620	-6.829	-4.871						
	eη	0.584	0.197	0.594	0.783						
	er	0.220	0.176	0.179	0.159						

Úspěšnost korekční metody může být vyjádřena počtem všech případů, kdy metoda byla vyhodnocena jako nejlepší pro dané kritérium v dané stanici. Tabulka tedy podává následující celkové výsledky: lineární scaling (LS) byl vyhodnocen jako nejlepší metoda ve třech případech, kvantilové mapování (QM) ve 23 případech, první lineární metoda (LM1) ve 12 případech a druhá lineární metoda (LM2) ve 30 případech.

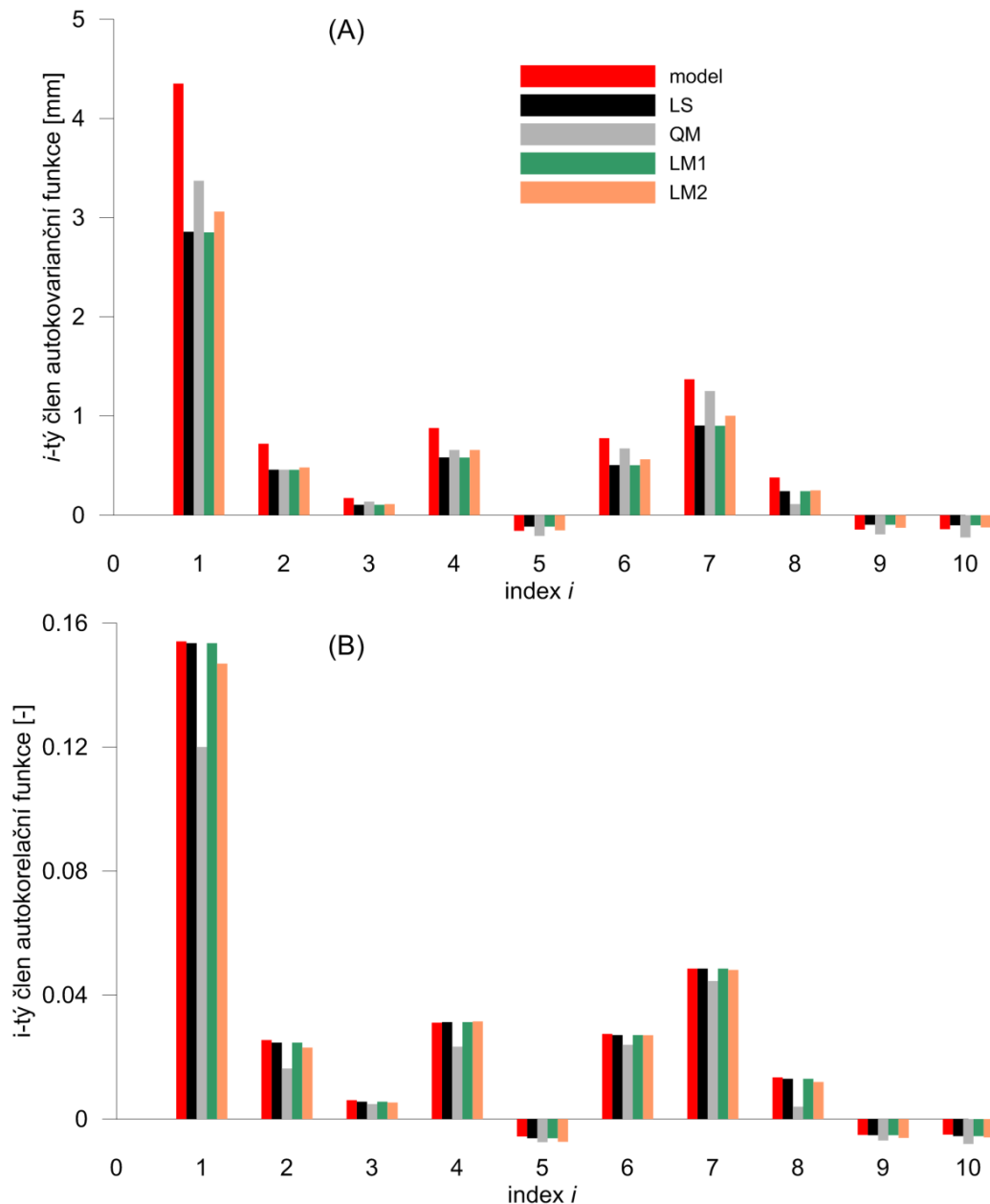
Je nutno poznamenat, že kvantilové mapování přináší obecně nejlepší výsledky při korekci rozptýlení modelových dat, neboť dosahuje nejmenší chyby $e\sigma$ v 11 případech. V případě dvou stanic (Churáňov a Staré Hutě) dosahuje QM dokonce řádově menší chyby rozptylu než ostatní metody. Nicméně v případě ostatních stanic není rozdíl v chybě rozptýlení tak výrazný a v několika případech QM při korekci rozptylu selhává. Metoda LM2 je naopak dominantní v případě korekce celkové chyby rozdělení er , kdy dosahuje nejmenší chyby ve 14 případech. Jako nejméně vhodná do nestacionárních podmínek se jeví metoda klasické lineární korekce (LS). Celkové výsledky tedy naznačují, že kvantilové mapování nemusí být nutně nejlepší volbou, pokud mezi kalibračním a validačním obdobím nastanou změny v charakteru datových souborů. V takovém případě je užitečnější zachytit lineární trend korekce, než vyjádřit precizní korekční vztah platný v kalibračním období. Z porovnání lineárních korekcí je dále zřejmé, že korekce odvozená z předpokladu gama rozdělení (obzvláště LM2 metoda) je výrazně úspěšnější než metoda LS předpokládající exponenciálně rozdělená data.

Dále byl zkoumán efekt korekcí na data z pohledu časové řady. Bylo prošetřeno, jak korekce ovlivňují rozložení suchých period ve výsledných řadách. Jediný krok v rámci celého korekčního procesu, který má na tento ukazatel vliv, je aplikace prahové hodnoty. Byly porovnány délky suchých období v měřených a korigovaných datech z validačního období. Délku suchého období můžeme považovat za dobu čekání na déšť. Ze statistické teorie je známo, že délka čekání na nějakou událost mívá exponenciální rozdělení. K ověření tohoto faktu byly spočteny délky suchých období, vyneseny jejich histogramy a nakonec bylo ověřeno, že průběhy histogramů jsou velmi dobře vystihnutele odhadem exponenciální hustoty. Porovnání exponenciálních hustot ukázalo, že rozložení sucha v korigovaných datech je podobné jako v datech měřených, u všech stanic došlo aplikací prahových hodnot k podstatnému zlepšení tohoto parametru ve srovnání s nekorigovanými daty. Například ve stanici Benešov nad Černou mají délky sucha v měřených datech parametr exponenciálního rozdělení roven 0.29, syrová modelová data 0.76 a data po korekci (tedy po aplikaci prahové hodnoty) 0.38. Velmi podobné výsledky byly zjištěny u všech stanic a nejsou zde uvedeny. Obecně lze dodat, že data z klimatického modelu REMO na všech stanicích vykazují klasický efekt mrhnutí v kalibračním i validačním období, proto aplikace prahových hodnot výrazně přibližuje celkový podíl suchých dní i jejich rozložení v čase reálným hodnotám. Je však nutné poznamenat, že rozložení sucha v korigovaných datech závisí z větší části na kvalitě samotného klimatického modelu, spíše než na určení prahové hodnoty. Bylo ověřeno, že různé prahové hodnoty mají na podíl suchých dní a jejich rozložení v čase podobný efekt.

Dále bylo zkoumáno, jak korekce ovlivňují hodnoty autokovarianční a autokorelační funkce časových řad. Z teoretického hlediska je možné jednoduchým dosazením zjistit, že čistě lineární korekce ve tvaru $y = Ax + B$ vede k

$$c_y^k = A^2 c_x^k \quad (6.7)$$

kde c_x^k a c_y^k jsou k -té členy autokovariančních funkcí časových řad x a y . Lineární korekce tedy mění měřítko funkce, nikoliv však poměr mezi jednotlivými funkčními členy. Stejným způsobem lze odvodit, že autokorelační funkce není lineární korekcí ovlivněna vůbec. Toto chování je ilustrováno na obrázku 6.1, na kterém je vyobrazeno prvních 10 členů autokovarianční a autokorelační funkce modelových a korigovaných dat z validačního období pro stanici Benešov nad Černou.



Obr. 6.1 Členy autokovarianční (A) a autokorelační (B) funkce modelových a korigovaných dat z validačního období ze stanice Benešov nad Černou.

Z obrázku je patrné, že vliv korekcí zhruba odpovídá teoretickým předpokladům. Odchytky jsou patrné u metody QM, kde se dají vysvětlit odklonem funkce od čistě lineárního průběhu. Malé ovlivnění autokorelační funkce v případech lineárních metod lze vysvětlit jako vliv působení prahových hodnot. V průběhu jejich aplikace je určitá část modelových dat (často i několik desítek procent) převedena na nulu. Efekt této transformace na autokovarianci a autokorelaci v časových řadách je ale celkem marginální. Je to proto, že aplikace prahových hodnot je selektivní proces, při kterém jsou původně *podobné* hodnoty (menší než prahová hodnota) transformovány na *identickou* hodnotu (nulu), což nemůže výrazně ovlivnit průběh autokovariance ani autokorelace.

Dalším bodem analýzy, který se týká jen nově odvozených metod, bylo prošetření, zda je proces hledání korekčních parametrů A a B citlivý na různé formulace účelové funkce. Při popisu odvození metod bylo zvoleno vyjádření účelové funkce jako integrálu z rozdílu hustot,

viz vztahy (6.2) a (6.4). Odchylka rozdělení mezi modelovými a měřenými daty však může být vyjádřena mnoha způsoby. Pro porovnání s používaným vyjádřením bylo vybráno několik odlišných formulací, jejichž přehled je uveden v tabulce 6.3.

Tab. 6.3 Různé formulace účelové funkce (rozdílu hustoty měřených dat f_O a hustoty modelových dat f_M) použité pro testování stability procesu hledání parametrů.

Formulace účelové funkce

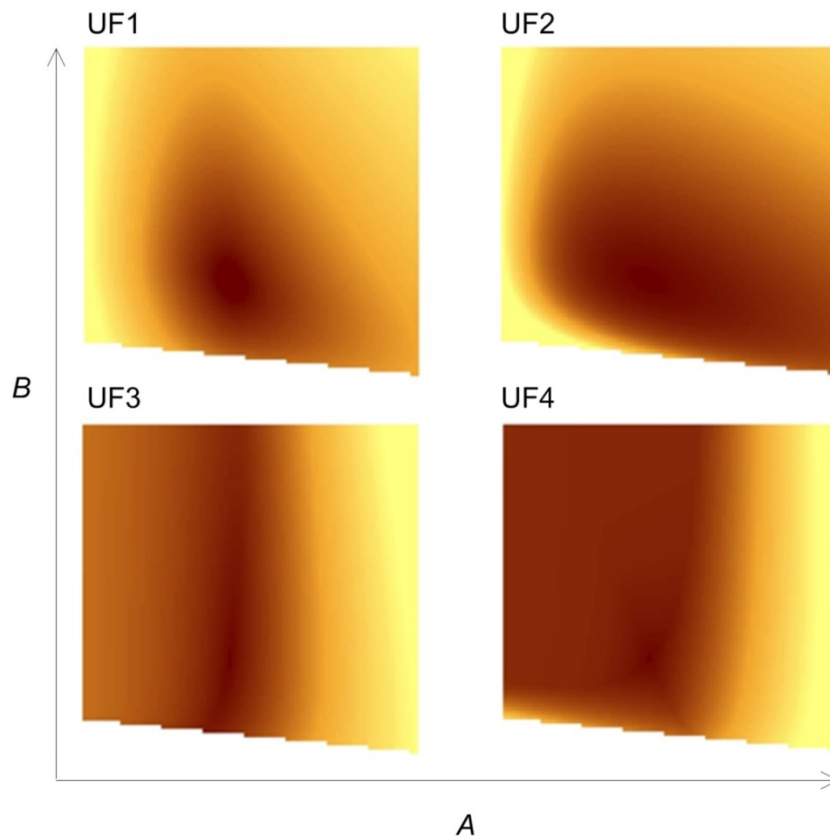
$$UF1 = \int |f_O - f_M|$$

$$UF2 = \sum_i (f_{Oi} - f_{Mi})^2$$

$$UF3 = \sum_i \frac{|f_{Oi} - f_{Mi}|}{f_{Oi}}$$

$$UF4 = \max \frac{|f_{Oi} - f_{Mi}|}{f_{Oi}}$$

Formulace UF1 je základní účelová funkce, popsaná v metodické části v definicích (6.2) a (6.4), zápis v tabulce 6.3 je zkrácen, a použita při výpočtu výše prezentovaných výsledků. Vyjádření UF2, UF3 a UF4 jsou diskrétními formulacemi počítanými přes stejný interval jako UF1 s krokem 0.01 mm. Na datech z několika stanic bylo ověřeno, že formulace účelové funkce nemá podstatný vliv na hodnoty výsledných parametrů. Tato skutečnost je demonstrována na obrázku 6.2, který pro kalibrační data ze stanice České Budějovice ukazuje kolísání všech čtyř účelových funkcí z tabulky 6.3 na stejném výřezu parametrického prostoru $A \times B$. K obrázku není připojena legenda, protože hodnoty účelových funkcí jsou vyjádřeny v odlišných jednotkách a pohybují se v odlišných řádech, nicméně pro všechny platí, že tmavší barva znamená nižší hodnotu účelové funkce. Podstatnou informací tedy je, že minima funkcí (tedy nejtmaší místa) leží pro všechna vyjádření přibližně na stejné kombinaci parametrů A a B .



Obr. 6.2 Průběh různých vyjádření účelových funkcí pro kalibrační data ze stanice České Budějovice. Pro podrobnosti viz text.

Závěrem byla testována účinnost nových metod při korekci simulací regionálního modelu ALADIN53 na měřená data z vybraných povodí, uvedených v tabulce 4.2. Měřená data i modelové simulace v tomto případě představují průměrné hodnoty na plochu uvedených povodí, v případě měřených dat jsou navíc ještě odvozena ze staničních dat interpolovaných na pravidelnou prostorovou mřížku, viz popis dat v kapitole 4. Pro kalibraci korekcí byla použita data z let 1970 – 1987, validace probíhala na datech z let 1988 – 2005. Oproti předchozímu testování korekcí se staničními měřenými daty a výstupy modelu REMO tedy došlo k několika změnám: byla použita prostorově interpolovaná data, kalibrační a validační období jsou delší a následují ihned po sobě. Výběr korekčních metod a použitá hodnotící kritéria se oproti předchozímu testu nijak nezměnily. Kompletní výsledky jsou uvedeny v tabulce 6.4 a ukazují, že oproti prvnímu testu došlo ke změně v hodnocení metod. Nejlepší metodou je v tomto případě kvantilové mapování, které dosáhlo nejmenší chyby v 18 případech, kdežto metoda LM2 dosáhla nejmenší chyby pouze ve 14 případech

Tab. 6. 4 Chyby korekčních metod ve validačním období při korekcích výstupů modelu ALADIN53 na jednotlivá povodí z tabulky 4.2.

povodí		LS	QM	LM1	LM2	povodí		LS	QM	LM1	LM2
Huntava	$\epsilon\mu$	-0.196	-0.211	-0.210	-0.156	Morava	$\epsilon\mu$	-0.196	-0.215	-0.223	-0.228
	$\epsilon\sigma$	-5.652	-2.531	-5.689	-3.780		$\epsilon\sigma$	-10.366	-8.743	-10.728	-10.379
	$\epsilon\eta$	-0.858	-0.901	-0.874	-0.593		$\epsilon\eta$	-0.903	-0.701	-0.973	-0.942
	$\epsilon\tau$	0.121	0.124	0.127	0.120		$\epsilon\tau$	0.155	0.126	0.152	0.144
Čistá	$\epsilon\mu$	-0.196	-0.213	-0.225	-0.400	Břežná	$\epsilon\mu$	-0.190	-0.207	-0.176	-0.176
	$\epsilon\sigma$	-6.198	-0.530	-6.545	-8.326		$\epsilon\sigma$	-6.614	-4.717	-6.435	-6.435
	$\epsilon\eta$	-0.972	-0.851	-1.054	-1.523		$\epsilon\eta$	-0.399	-0.348	-0.356	-0.356
	$\epsilon\tau$	0.135	0.098	0.137	0.116		$\epsilon\tau$	0.102	0.095	0.103	0.103
Babí potok	$\epsilon\mu$	-0.209	-0.216	-0.235	-0.190	Kněžská	$\epsilon\mu$	-0.106	-0.118	-0.111	-0.111
	$\epsilon\sigma$	-5.400	0.093	-5.686	-4.067		$\epsilon\sigma$	-5.218	-1.958	-5.283	-5.283
	$\epsilon\eta$	-0.956	-1.202	-1.019	-0.782		$\epsilon\eta$	-0.061	-0.220	-0.076	-0.076
	$\epsilon\tau$	0.128	0.122	0.127	0.127		$\epsilon\tau$	0.089	0.104	0.091	0.091
Potočnice	$\epsilon\mu$	-0.133	-0.130	-0.152	-0.129	Rokytenka	$\epsilon\mu$	-0.146	-0.144	-0.174	-0.150
	$\epsilon\sigma$	-5.593	-3.189	-5.840	-5.203		$\epsilon\sigma$	-7.082	-3.457	-7.400	-6.059
	$\epsilon\eta$	-0.152	-0.202	-0.206	-0.096		$\epsilon\eta$	-0.410	-0.402	-0.480	-0.292
	$\epsilon\tau$	0.087	0.091	0.088	0.085		$\epsilon\tau$	0.130	0.119	0.134	0.109
Opavice	$\epsilon\mu$	0.005	-0.003	-0.001	-0.007	Zdobnice	$\epsilon\mu$	-0.143	-0.151	-0.127	-0.123
	$\epsilon\sigma$	-2.530	-3.757	-2.623	-2.637		$\epsilon\sigma$	-5.613	-4.060	-5.377	-5.377
	$\epsilon\eta$	-0.367	-0.351	-0.381	-0.391		$\epsilon\eta$	-0.445	-0.446	-0.398	-0.314
	$\epsilon\tau$	0.087	0.095	0.085	0.087		$\epsilon\tau$	0.112	0.119	0.111	0.102

Posun v úspěšnosti metod je dán dvěma faktory. Zaprvé, při interpolaci dat na prostorovou mřížku a následném průměrování na plochu povodí dochází k odklonu rozdělení interpolovaných datových souborů od původních rozdělení staničních dat. Centrální limitní věta ukazuje, že náhodná veličina daná součtem nezávislých, stejně rozdělených náhodných veličin s rostoucím počtem sčítanců konverguje k normálnímu rozdělení. To se v tomto případě neděje, neboť jednotlivé dílčí veličiny (data z originálních meteorologických stanic) mají rozdělení mírně odlišná, je porušena podmínka nezávislosti a sčítanců není dostatečný počet. Nicméně zcela jistě dochází alespoň k mírnému odklonu od původních rozdělení. Nově odvozené lineární metody jsou přitom založeny na předpokladu gama rozdělení, který obvykle dobře vyhovuje originálním staničním datům.

Druhým faktorem negativně ovlivňujícím účinnost lineárních korekcí je snížená míra nestacionarity. Pro ověření tohoto faktu byla nestacionarita kvantifikována s využitím histogramů z kalibračního a validačního období. Přitom bylo bráno v úvahu, že nestacionarita, ať již ve formě trendu, či pouhé přirozené variability nastává v časových řadách měřených i modelových dat. Pro vyjádření nestacionarity *nezachycené* modelem byly spočteny histogramy měřených a modelových dat z kalibračního i validačního období, přitom je podstatné, že všechny histogramy mají stejný počátek a šířku intervalu. *Nezachycená* nestacionarita (*nn*) pak byla vyjádřena jako

$$nn = s \frac{1}{4} \sum_i |(f_{OVi} - f_{OCi}) - (f_{MVi} - f_{MCi})| \quad (6.8)$$

kde s je šířka histogramových intervalů, f jsou hodnoty histogramů, indexy C a V označují kalibrační a validační období. Minimální hodnota kritéria (6.8) je 0, která značí buď absenci nestacionarity, nebo její precizní detekci modelem, maximální možná hodnota je 1. Hodnoty nn pro jednotlivé časové řady z prvního i druhého testu jsou uvedeny v tabulce 6.5.

Tab. 6.5 Hodnoty nezachycené nestacionarity pro časové řady použité v obou testech.

hodnoty nn pro data z prvního testu		hodnoty nn pro data z druhého testu	
stanice	nn	povodí	nn
Benešov nad Černou	0.117	Huntava	0.056
Besednice	0.135	Čistá	0.057
České Budějovice	0.099	Babí potok	0.063
Dolní Dvořiště	0.136	Potočnice	0.062
Husinec	0.096	Opavice	0.052
Churáňov	0.097	Morava	0.090
Kaplice	0.123	Březná	0.058
Malonty	0.099	Kněžná	0.055
Netřebice	0.114	Rokytenka	0.063
Pohorská Ves	0.130	Zdobnice	0.069
Přídolí	0.107		
Římov	0.143		
Soběnov	0.089		
Staré Hutě	0.120		
Třeboň	0.133		
Trhové Sviny	0.122		
Vyšší Brod	0.098		

Je třeba poznamenat, že kritérium (6.8) je citlivé na zvolenou šířku intervalu použitých histogramů, proto je nutné výsledky pro konkrétní časové řady interpretovat s rezervou. Nicméně protože při hodnocení nestacionarity všech časových řad z obou testů byla použita stejná šířka intervalu 0.1 mm, dovoluje tabulka 6.5 porovnat míru nezachycené nestacionarity v datech z obou testů. Toto porovnání jasně vypovídá o tom, že hodnoty nestacionarity ve druhém testu byly v průměru zhruba poloviční než v případě prvního testu. Tyto výsledky potvrzují teorii, že odvozené lineární metody jsou úspěšnější až při určité míře nestacionarity, při nepříliš odlišných podmínkách zůstává efektivnější metodou kvantilové mapování.

6.5 Dodatek - QM pod vlivem nestacionarity

Text tohoto dodatku k šesté kapitole přímo nesouvisí s nově odvozenými metodami. Jeho obsahem je velmi stručný rozbor vlastností kvantilového mapování pod vlivem nestacionárních podmínek, který (podle mínění autora této práce) přispívá k argumentům podporujícím náhradu QM nějakou jednodušší (lineární) korekční funkcí.

Výchozím bodem úvahy je skutečnost, že výsadní postavení QM mezi korekčními metodami je dáno schopností této metody korigovat celé rozdělení pravděpodobnosti, na rozdíl od jednodušších metod opravujících pouze některé vybrané momenty. Z toho důvodu by mělo být prošetřeno, jak je korekce celého rozdělení nestacionárními podmínkami ovlivnitelná.

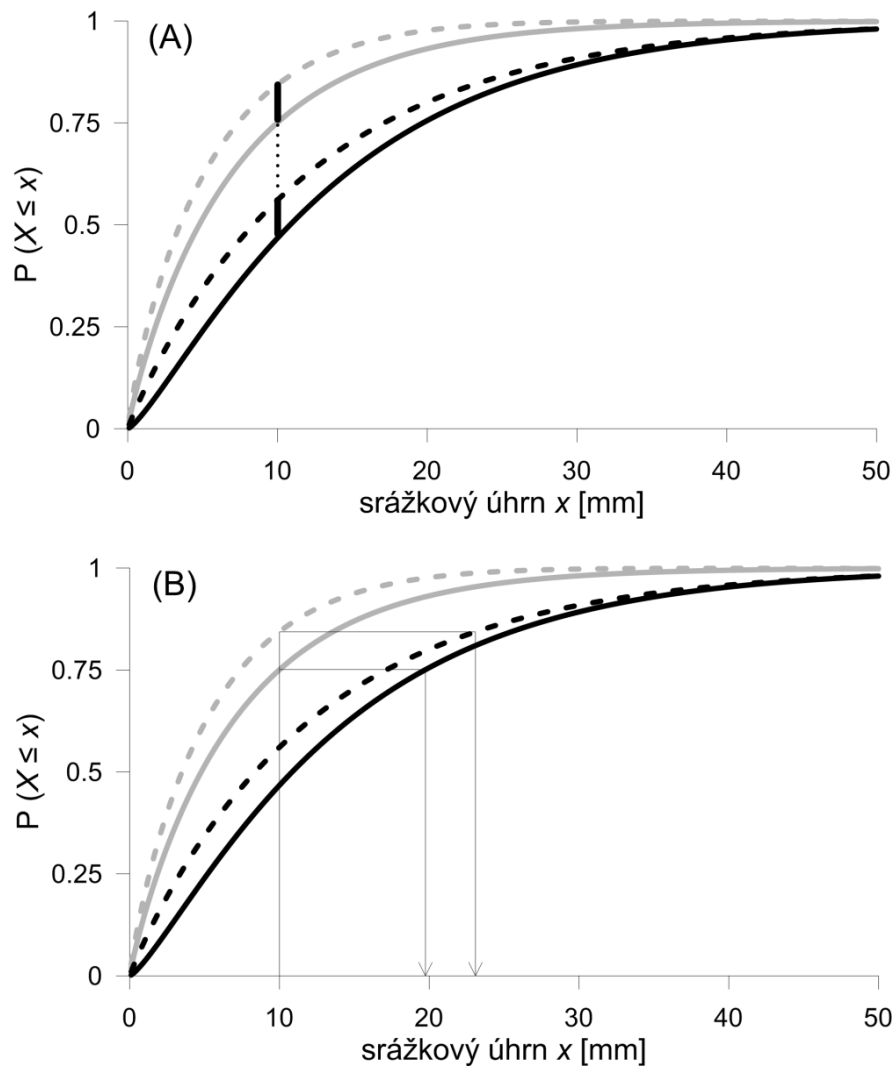
Nestacionarita bude chápána jako jakákoliv odchylka mezi distribučními funkcemi, na kterých je korekční funkce kalibrována, a distribučními funkcemi z validačního (potažmo budoucího zájmového) období, na kterých je korekce aplikována. Nestacionarita je tedy funkcí dvou změn: 1) změny distribuční funkce měřených dat a 2) změny distribuční funkce modelových dat. Do úvah o nestacionaritě tedy vstupuje další proměnná – otázka, jak kvalitně se klimatickému modelu podařilo zachytit změny probíhající v reálných datech.

Nějaké změny lze jistě předpokládat, neboť modely jsou primárně konstruovány k tomu, aby simulovaly měnící se klima. Také lze konstatovat, že pokud by model zcela selhal a změna modelové distribuční funkce by naprosto neodpovídala realitě, korekce chyby by v takové situaci příliš nepomohla. Smysl má tedy pouze zabývat se případem, kdy klimatický model probíhající změny alespoň přibližně zachytí.

Nejprve je analyzován hypotetický případ, kdy změny měřené distribuční funkce jsou modelem zachyceny naprosto precizně. To znamená, že difference mezi kalibrační a validační distribuční funkcí jsou stejné pro měřená i modelová data *pro každou hodnotu srážkového úhrnu*, tzn.

$$F_{OV}(x) - F_{OC}(x) = F_{MV}(x) - F_{MC}(x); \forall x \quad (6.9)$$

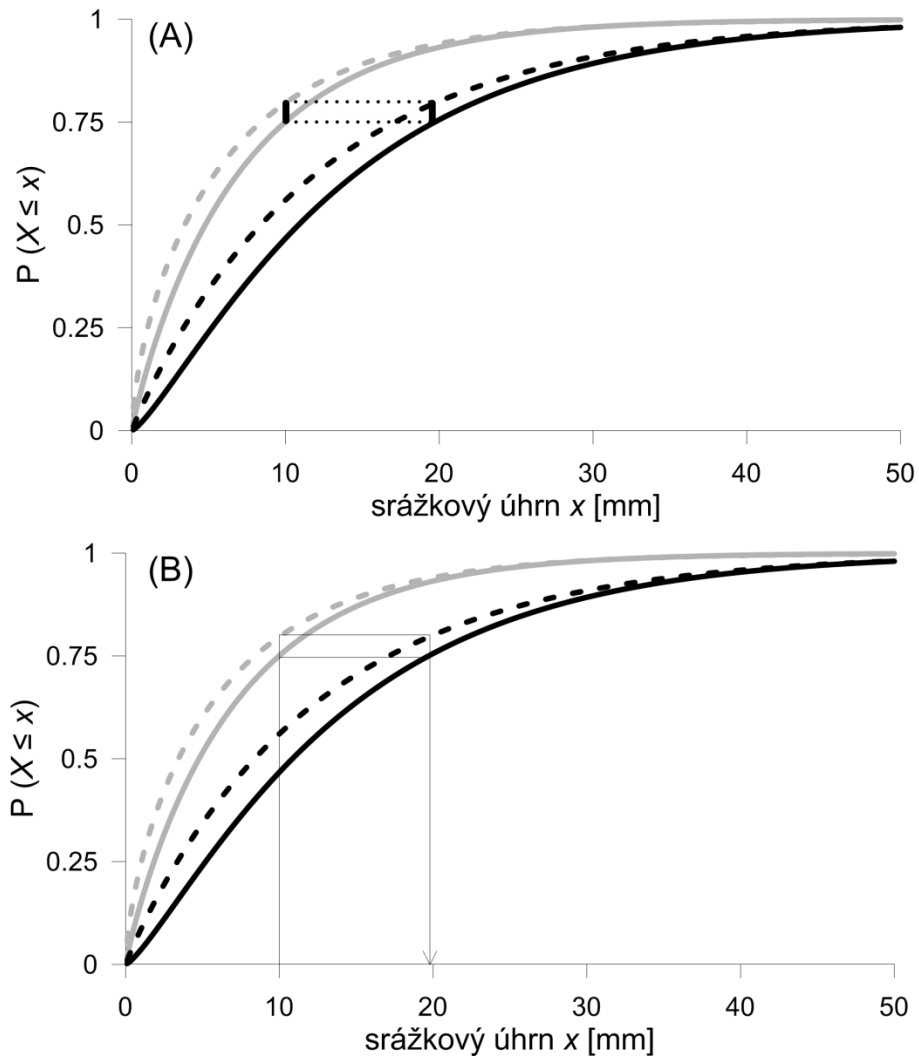
kde symboly O a M znamenají měřená a modelová data, symboly C a V znamenají kalibrační a validační období. Situace odpovídající rovnici (6.9) je znázorněna na obrázku 6.3(A) a účinek tohoto uspořádání na QM je znázorněn na obrázku 6.3(B).



Obr. 6.3 Precizní detekce nestacionarity měřených dat modelem. (A) Diference mezi kalibračními (plné čáry) a validačními (čárkované čáry) distribučními funkcemi měřených dat (černá barva) a modelových dat (šedá barva). (B) Vliv na kvantilové mapování: korekce modelové hodnoty $x = 10$ v kalibračním a validačním období.

Z obrázku 6.3(B) plyne, že funkce kvantilového mapování v tomto případě selhává. Modelová hodnota $x = 10$ mm, která by měla být ve validačním období korigována na hodnotu cca 23 mm, je korigována podle kalibrační funkce na hodnotu menší než 20 mm. Totéž platí i pro všechny ostatní hodnoty, výsledkem bude zcela zkreslená distribuční funkce korigovaných dat ve validačním období, která neodpovídá měřeným datům.

Otázkou tedy je, jakým způsobem by musel model reagovat na změnu měřených dat, aby korigovaná distribuční funkce byla ve validačním období platná. Odpověď přináší obrázek 6.4, který demonstruje případ, kdy diference mezi kalibrační a validační distribuční funkcí jsou stejné pro měřená i modelová data *pro každou hodnotu pravděpodobnosti*:



Obr. 6.4 Detekce nestacionarity měřených dat vyhovující kvantilovému mapování. (A) Diference mezi kalibračními (plné čáry) a validačními (čárkované čáry) distribučními funkcemi měřených dat (černá barva) a modelových dat (šedá barva). (B) Korekce modelové hodnoty $x = 10$ v kalibračním a validačním období.

Pro uspořádání na obrázku 6.4 tedy platí, že

$$F_{OV}(x_1) - F_{OC}(x_1) = F_{MV}(x_2) - F_{MC}(x_2), \text{ kde } x_2 = F_{MC}^{-1}(F_{OC}(x_1)) \quad (6.10)$$

Z výše uvedeného tedy plyne, že k tomu, aby kvantilové mapování ve validačním či projekčním období skutečně korigovalo celé rozdělení pravděpodobnosti je zapotřebí, aby model detekoval vývoj reálné nestacionarity a navíc tato detekce musí být v souladu s rovnicí (6.10). Tato podmínka je však zcela uměle vytvořená ad hoc pro účely kvantilového mapování. Detekce nestacionarity splňující rovnici (6.9), která by odpovídala intuitivní představě o korektním a bezchybném výkonu modelu, vede při praktické aplikaci k deformaci výsledné distribuční funkce. Efektivitu kvantilového mapování je tedy nutno brát v nestacionárních podmínkách s rezervou, o čemž kromě tohoto rozboru svědčí i výsledky kapitoly 6.4.

7 Korekce závislostní struktury

7.1 Úvod

Jedním z problematických faktorů při korekci chyb klimatických modelů je skutečnost, že korekce jsou zpravidla kalibrovány a aplikovány odděleně pro jednotlivé prostorové body (meteorologické stanice, modelové grid-boxy) a jednotlivé meteorologické veličiny (srážky, teploty). Výsledkem tohoto procesu je, že jednotlivé korigované soubory sice vykazují dobrou shodu statistických ukazatelů s měřenými daty, ale závislostní (tj. korelační a kovarianční) struktury korigovaných a měřených dat se navzájem liší (Ehret et al., 2012). Tato skutečnost může mít značný vliv na následné hydrologické, zemědělské či ekologické studie, hodnotící dopad klimatické změny.

Chyba ve vztazích mezi meteorologickými veličinami může ovlivnit modelování hydrologických procesů (Liu et al., 2014, Teng et al., 2015), realistický vztah mezi teplotami a srážkami je vyžadován také zemědělskými modely pro simulaci růstu plodin (Hoffmann a Rath, 2012).

Lze nalézt několik studií zabývajících se vlivem korekce na vztah mezi srážkami a teplotami. Wilcke et al.(2013) ukázali, že kvantilové mapování nemá zřetelný efekt na korelace mezi veličinami. Hoffmann a Rath (2012) odvodili proceduru, upravující vztah srážek a teplot po provedení klasické korekce. Piani a Haerter (2012) rozšířili kvantilové mapování pro použití na dvourozměrné datové soubory. V jejich pojetí byla korekce provedena nejprve klasicky pro jednotlivé teplotní soubory, srážková data pak byla korigována po částech spadajících do individuálních teplotních kvantilů.

Uvedené studie se zabývaly úpravou vztahů modelových srážek a teplot v rámci jedné meteorologické stanice. Chybně vystižené *prostorová* závislostní struktura mezi srážkovými daty na daném povodí může negativně ovlivnit např. modelování povrchového odtoku v hydrologických studiích (Liu et al., 2014). Cílem této části disertační práce je proto odvodit a validovat proceduru, umožňující korekci prostorových korelací a kovariancí u mnoho-rozměrných srážkových souborů.

7.2 Odvození metody

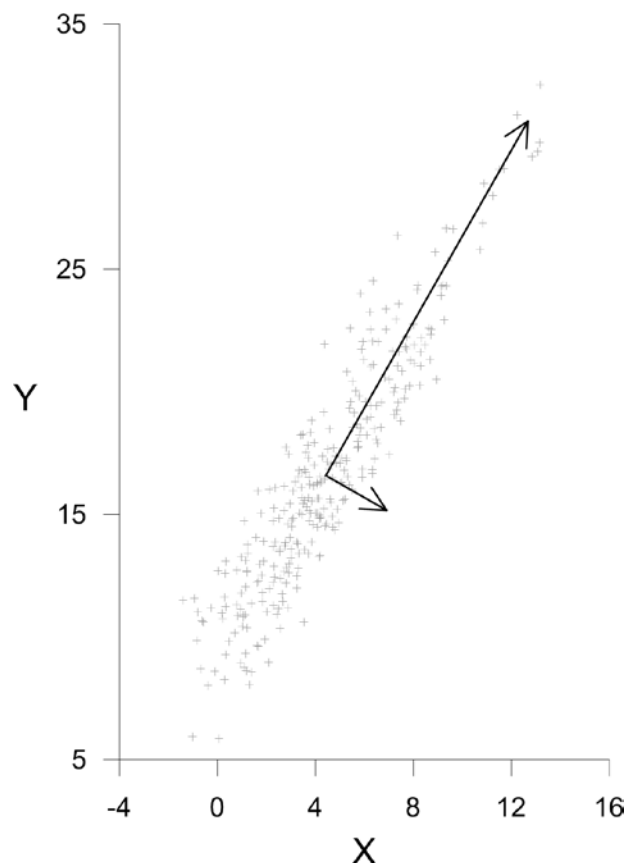
Metoda pro korekci závislostní struktury je založena na tzv. hlavních komponentách, proto úvodem této kapitoly budou stručně shrnuty základní teoretické poznatky.

Uvažujme n -dimenzionální datový soubor, kterým mohou být např. časové řady srážek z n meteorologických stanic za určité období. Hlavní komponenty jsou lineární kombinace jednotlivých originálních veličin, analýzou n -rozměrného souboru obdržíme n hlavních komponent. Koeficienty hlavních komponent jsou dány vlastními vektory kovarianční matice originálního datového souboru. Pokud jsou vlastní vektory seřazeny v sestupném pořadí podle příslušných vlastních čísel, potom první hlavní komponenta určuje směr, ve kterém datový soubor vykazuje maximální rozptyl. Další hlavní komponenty pak vždy definují směr, ve kterém data vykazují maximální možnou část zbytkového rozptylu nevysvětleného předchozími hlavními komponentami.

Kovarianční matice (někdy bývá označována jako varianční, jde o matici složenou z kovariancí jednotlivých složek náhodného vektoru) je symetrická, takže vlastní vektory definují ortogonální bázi n -rozměrného vektorového prostoru. Z toho plyne, že jednotlivé hlavní komponenty jsou na sebe kolmé, a tedy vzájemně nekorelované. Kovarianční matice je

pozitivně-semidefinitní (důkaz viz např. Anděl, 2007, kapitola 3.2), což znamená, že její vlastní čísla jsou nezáporná a lze ukázat, že jsou rovna rozptylům příslušných hlavních komponent. Kompletní teorie analýzy hlavních komponent je vyčerpávajícím způsobem podána např. v Joliffe (2002).

Obvyklou motivací pro použití analýzy hlavních komponent v praktických úlohách je snaha o snížení dimenze datového souboru; původní proměnné jsou nahrazeny několika nejvyššími hlavními komponentami tak, aby byla vysvětlena většina rozptylu původních dat. Hlavní komponenty ale mohou být využity jako alternativní souřadná soustava, která reflektuje skutečnou polohu a tvar datového souboru. Tuto souřadnou soustavu získáme tak, že vlastní vektory umístíme do bodu středních hodnot datového souboru a upravíme jejich délku tak, aby byla rovna odmocnině z příslušného vlastního čísla, tzn. směrodatné odchylce ve směru daném vlastním vektorem. Definujme tedy souřadný systém hlavních komponent jako posloupnost vlastních vektorů kovarianční matice, seřazených sestupně podle velikosti příslušného vlastního čísla, posunutých do bodu střední hodnoty a s délkou upravenou výše popsaným způsobem. Takový systém pro syntetický dvourozměrný datový soubor je znázorněn na obrázku 7.1.

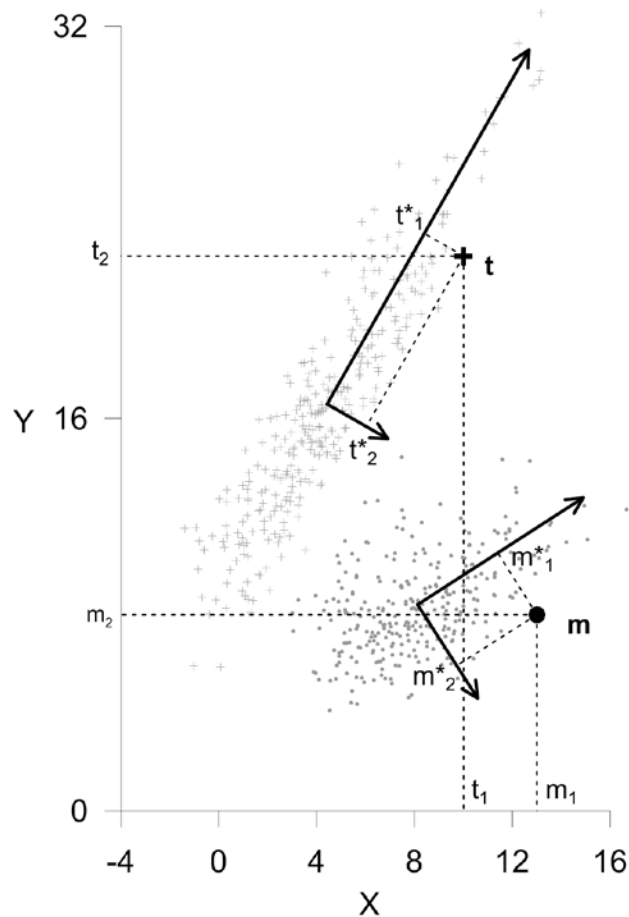


Obr. 7.1 Demonstrace alternativní souřadné soustavy, založené na hlavních komponentách, na příkladu dvourozměrného syntetického datového souboru.

Korekce závislostní struktury je založená na těchto alternativních souřadných soustavách. Nechť \mathcal{O} a \mathcal{M} reprezentují alternativní souřadné soustavy měřených a modelových dat. Původní souřadná soustava, daná kanonickou bází, bude označena jako \mathcal{C} . Korekce spočívá v konverzi modelových dat mezi souřadnými systémy \mathcal{M} a \mathcal{O} . Celý proces bude v textu

popsán pro obecnou n -rozměrnou úlohu, zatímco obrázek 7.2 demonstruje jednoduchý dvourozměrný příklad.

Nechť $\mathbf{m} = (m_1, \dots, m_n)$ je vektor modelových dat, obsahující srážkové úhrny z n meteorologických stanic v jednom dni. Vektor \mathbf{m} reprezentuje bod v souřadném systému \mathcal{C} (na obrázku 7.2 tučně zvýrazněný). Korekce spočívá v nalezení souřadnic bodu \mathbf{m} v souřadném systému \mathcal{M} a v jeho následném nahrazení bodem \mathbf{t} (tučně zvýrazněný křížek na obrázku 7.2), který má identické souřadnice v systému \mathcal{O} .



Obr. 7.2 Demonstrace korekce závislostní struktury na příkladu syntetických datových souborů, modelová data jsou označena tečkami, měřená data křížky, alternativní souřadné systémy úsečkami se šipkami. Celý proces je detailně popsán v textu.

Korekční procedura může být schematicky rozdělena do tří kroků:

1. Přejít ze systému \mathcal{C} do systému \mathcal{M} , jehož výsledkem je vektor (m_1^*, \dots, m_n^*) obsahující souřadnice bodu \mathbf{m} v systému \mathcal{M}
2. Nahrazení bodu \mathbf{m} bodem $\mathbf{t} = (t_1^*, \dots, t_n^*)$ tak, že $t_i^* = m_i^*$, nicméně bod \mathbf{t} je hledán v \mathcal{O}
3. Přejít z \mathcal{O} zpět do \mathcal{C} , výsledkem je vektor (t_1, \dots, t_n) obsahující souřadnice \mathbf{t} v \mathcal{C} , což jsou výsledné korigované hodnoty.

Celý proces je možno stručně zapsat jako

$$\mathbf{t} = \mathbf{e}_0 + \mathbf{O}\mathbf{M}^{-1}(\mathbf{m} - \mathbf{e}_M) \quad (7.1)$$

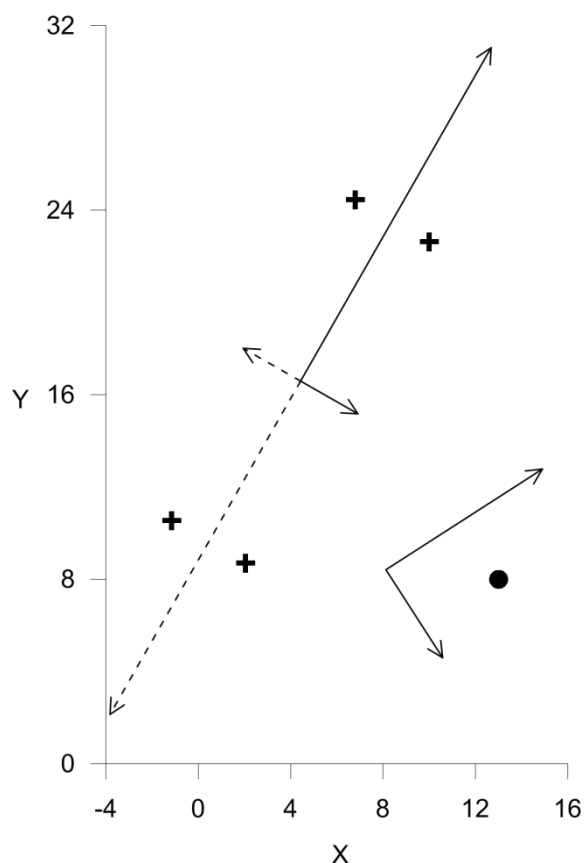
kde

- \mathbf{m} je vstupní sloupcový vektor modelových dat

- \mathbf{e}_O a \mathbf{e}_M jsou sloupcové vektory středních hodnot jednotlivých měřených a modelových proměnných
- \mathbf{M} je matice, jejíž jednotlivé sloupce jsou tvořeny vektory báze \mathcal{M} , tudíž \mathbf{M}^{-1} (inverzní matice k \mathbf{M}) je maticí přechodu z báze \mathcal{C} do báze \mathcal{M}
- \mathbf{O} je matice, jejíž sloupce jsou tvořeny vektory báze \mathcal{O} , tudíž \mathbf{O} je maticí přechodu z báze \mathcal{O} do báze \mathcal{C}
- \mathbf{t} je výsledný sloupcový vektor korigovaných dat.

Popsaná metoda bude označována jako PCC („principal components correction“). Je třeba dodat, že hlavní komponenty jsou zde využity v čistě algebraickém smyslu jako definice alternativní souřadné soustavy, statistická analýza hlavních komponent zde není prováděna. Jelikož z odhadnutých hlavních komponent nejsou vyvozovány žádné závěry statistické povahy, není vyžadována např. normalita rozdělení jednotlivých datových souborů.

Popsaná procedura je spojena s netriviálním praktickým problémem. Výsledkem standardních numerických algoritmů pro odhad vlastních čísel a vlastních vektorů jsou směry hlavních vektorů, což platí i pro algoritmus Jacobiho diagonalizace, použitý v této studii. *Orientace* vypočtených vlastních vektorů však zůstávají zcela nahodilé, což de facto znemožňuje použít originální výsledky numerické procedury jako definici souřadného systému. Obrázek 7.3 využívá stejných souřadných systémů jako obrázek 7.2 pro demonstraci uvedeného problému ve dvou-dimenzionálním prostoru.



Obr. 7.3 Demonstrace problému s orientacemi vlastních vektorů.

Z obrázku je patrné, že pro danou orientaci modelových vlastních vektorů je možné záměnami orientací pozorovaných vlastních vektorů dosáhnout čtyř odlišných výsledků korekce modelového bodu, pro celý soubor by pak došlo k zrcadlovému převrácení podle

některých os souřadného systému měřených dat. Pro obecnou n -dimenzionální úlohu je takovýchto výsledků 2^n a není tedy samozřejmě možné provádět vizuální posouzení jako v případě dvoudimenzionální úlohy. Před vlastní definicí souřadných systémů je tedy nutné synchronizovat orientace jejich dílčích os tak, aby dané úloze vyhovovaly.

Tento problém byl řešen následující úvahou. Lze předpokládat, že modelová a měřená data jsou si podobná v tom smyslu, že korekce závislostní struktury modelových dat sestává z pouze mírné rotace souřadného systému doplněné o změnu měřítka jednotlivých os (individuálně ke každé ose). Nyní tento předpoklad aplikujme na korekci nějaké syntetické extrémní hodnoty, např. bodu o souřadnicích $\mathbf{m} = (1000, 0, \dots, 0)$. Pokud je tento bod korigován podle vztahu (7.1) a pokud platí výše uvedené předpoklady, leží výsledná hodnota \mathbf{t} relativně blízko původního modelového bodu \mathbf{m} . Naopak nesprávné výsledky, dosažené při souřadném systému s jakkoliv nevhodně zvolenými orientacemi souřadných os budou ležet od původního bodu \mathbf{m} značně daleko. Aby bylo dosaženo bezpečné synchronizace všech os, byly syntetické extrémní seskupeny do testovací matice ve tvaru

$$\mathbf{P} = \begin{pmatrix} 1000 & 0 & \dots & 0 \\ 0 & 1000 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1000 \end{pmatrix}$$

Jak bylo uvedeno, v n -dimenzionálním vektorovém prostoru existuje 2^n různých variant orientací os souřadného systému. Tyto varianty však nemusí být testovány všechny, ale orientace každé osy souřadného systému měřených dat může být s použitím uvedené testovací matice posuzována individuálně. Nechť $\mathbf{O}_{i,j}$ je matice, jejíž sloupce jsou dány vektory báze \mathcal{O} s danou j -tou orientací i -té osy, kde $i \in (1, \dots, n)$ a $j \in (1, 2)$. Potom korekce testovací matice \mathbf{P} probíhá v souladu se (7.1) a s využitím matice $\mathbf{O}_{i,j}$ podle

$$\mathbf{P}_{i,j} = \mathbf{O}_{i,j} \mathbf{M}^{-1} \mathbf{P} \quad (7.2)$$

Vektory středních hodnot \mathbf{e}_O a \mathbf{e}_M ze (7.1) jsou při použití extrémních hodnot zanedbatelné, proto nebyly uvažovány. Chyba posuzované j -té orientace i -té osy je potom počítána podle

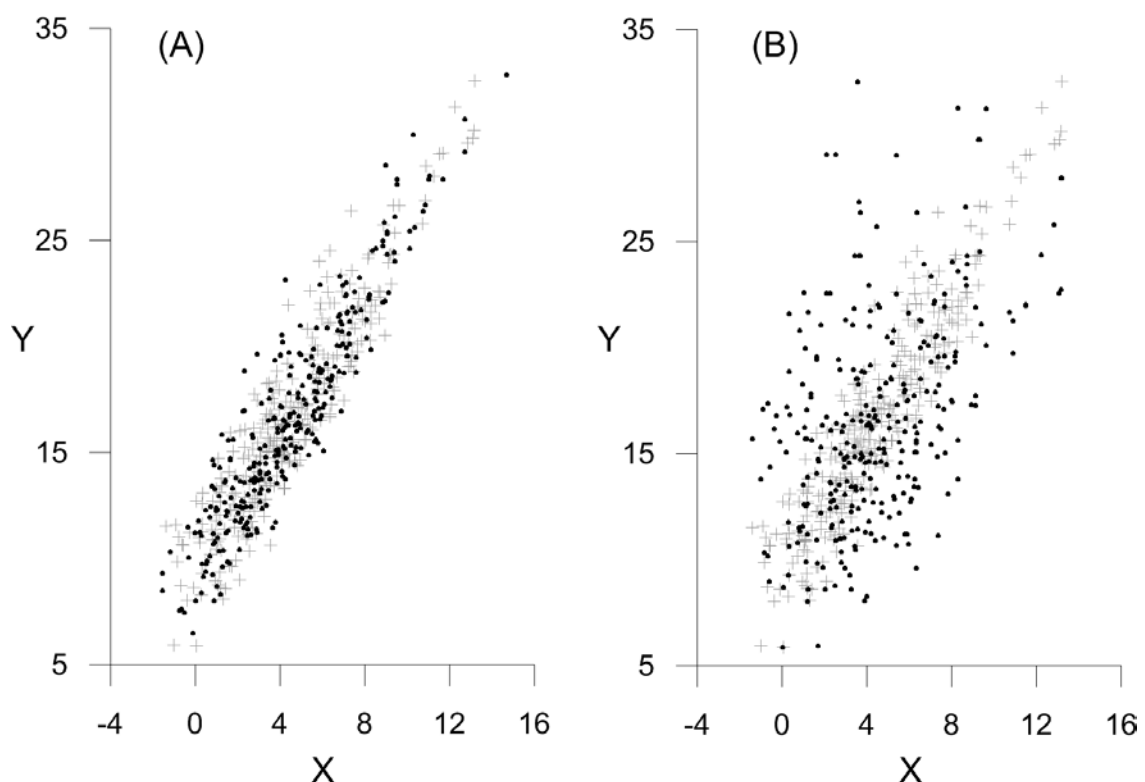
$$\text{err}_{i,j} = \|\mathbf{P} - \mathbf{P}_{i,j}\| \quad (7.3)$$

kde $\|\cdot\|$ označuje maticovou normu, v této studii byla použita Frobeniova norma. Vztah (7.3) tedy hodnotí „vzdálenost“ korigované matice $\mathbf{P}_{i,j}$ od původní matice \mathbf{P} v závislosti na posuzované orientaci i -té osy. V souladu s předpokladem o podobnosti modelových a měřených dat je pro každou osu báze \mathcal{O} vybrána orientace s nižší chybou. Test je prováděn pro každou z n os, kompletní synchronizace souřadných systémů tedy vyžaduje $2n$ výpočtů vzorců (7.2) a (7.3).

7.3 Standardní způsob korekce

Kromě validace nově odvozené metody bylo cílem práce také její porovnání se některou ze standardně využívaných metod. Pro toto porovnání byla vybrána metoda kvantilového mapování (QM) jako nejpůvodnější zástupce běžně používaných metod. Korekční formule QM je dána vzorcem (3.13), stručný rozbor metody je proveden v kapitole 3.2, na lineární povahu kvantilového mapování a na jeho podobnost s ostatními korekčními metodami je poukázáno v kapitole 5.

Pro ilustrační účely byl syntetický modelový datový soubor, použitý na obrázku 7.2 pro demonstraci nově odvozené metody, korigován oběma korekčními metodami PCC a QM. Obrázek 7.4 ukazuje výsledky korekce pro (A) PCC a (B) QM. Korigované hodnoty jsou označeny tečkami, měřené hodnoty křížky.



Obr. 7.4 Výsledky korekce syntetického datového souboru z obrázku 7.2 dosažené (A) metodou PCC a (B) metodou QM.

Výsledky ukazují, že v případě PCC má korigovaný soubor stejný tvar a leží ve stejné oblasti jako hodnoty měřené. Korelační koeficient (0.92) a kovariance (12.86) korigovaných a měřených hodnot jsou zcela identické. Naproti tomu data korigovaná metodou QM zhruba zachovávají korelaci modelových dat z obrázku 7.2, která byla rovna 0.48 (srovnej tvar modelového datového souboru z obrázku 7.2 a souboru korigovaného metodou QM z obrázku 7.4). Kovariance dat korigovaných metodou QM je rovna 6.8, což znamená oproti původním modelovým datům posun směrem k měřeným hodnotám, nicméně oproti požadované hodnotě 12.86 je zhruba poloviční.

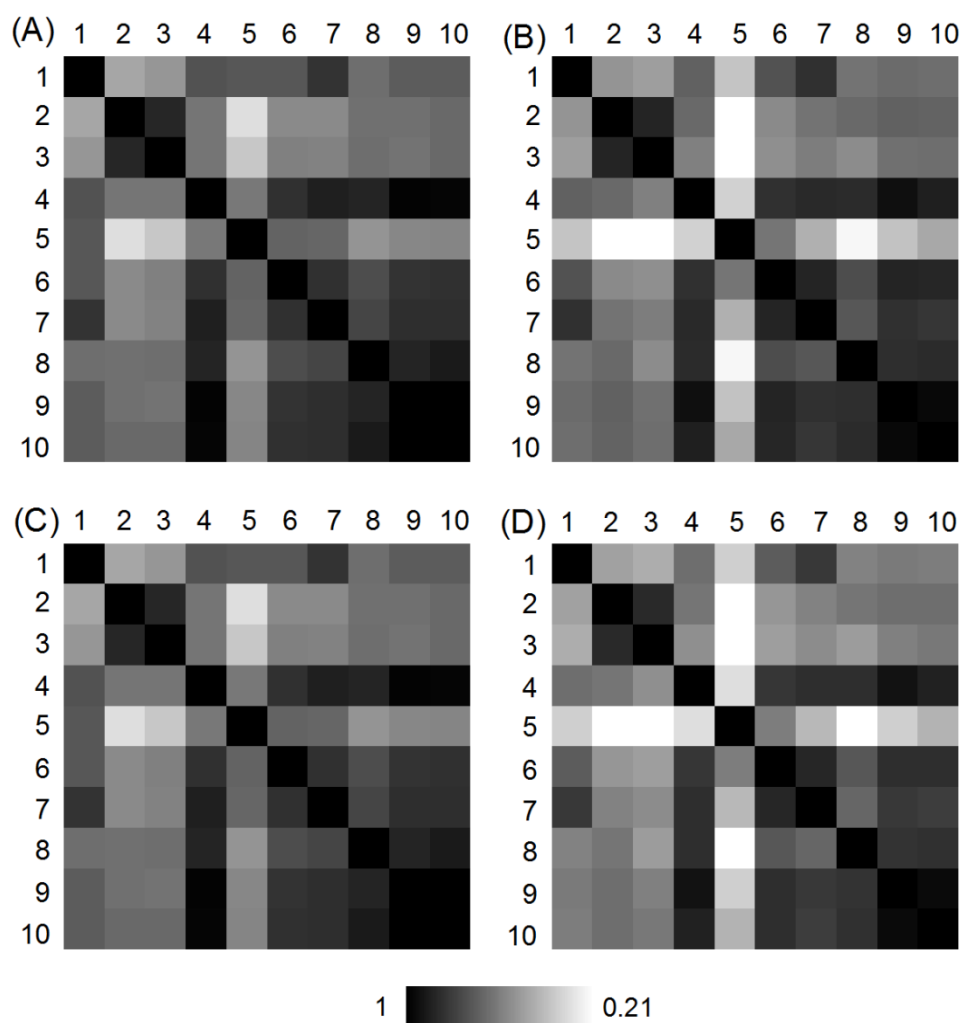
Korekce syntetických dat, prezentovaná na obrázku 7.4, zřetelně demonstruje rozdílný efekt korekčních metod PCC a QM na závislostní strukturu korigovaných dat. V dalším textu jsou prezentovány a diskutovány výsledky dosažené aplikací těchto metod na reálná měřená a

modelová data. Tyto výsledky ukazují, že PCC sice beze zbytku koriguje závislostní struktury, ale vykazuje nízkou účinnost při korekci individuálních rozdílů. Podrobný rozbor a návrh řešení tohoto problému je obsahem následujících kapitol.

7.4 Porovnání metod PCC a QM

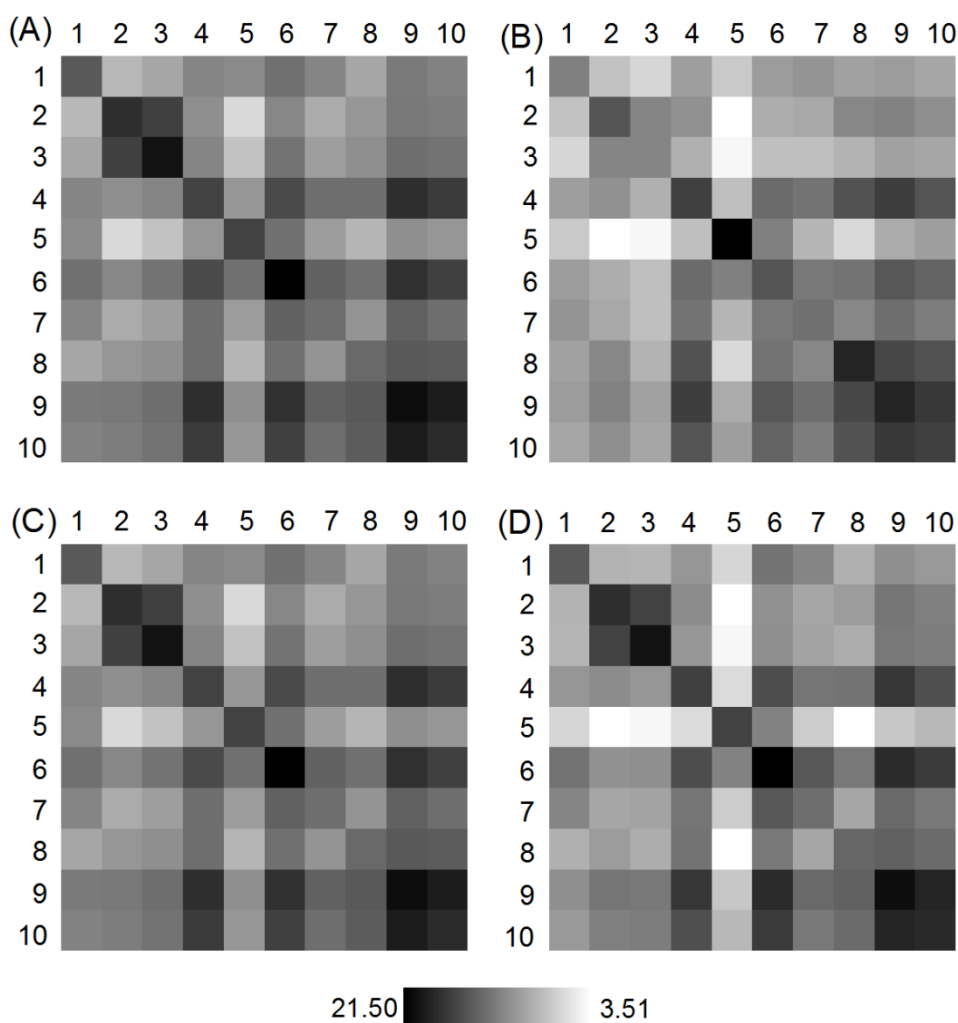
K porovnání metod byly použity simulace regionálního modelu ALADIN53, které byly korigovány na měřená data z vybraných povodí uvedených v tabulce 4.2. Účelem tohoto testu bylo detailně porovnat účinnost obou metod na korekci závislostní struktury a na korekci rozdílů jednotlivých datových souborů. Z tohoto důvodu nebyla provedena separace datových sad na kalibrační a validační období, aby porovnání nebylo ovlivněno zatemňujícím účinkem nestacionarity. Kalibrační – validační test byl proveden posléze a jeho výsledky jsou prezentovány v kapitole 7.5.

Deseti-dimenzionální modelový soubor reprezentující průměrné srážky na jednotlivá povodí byl korigován pomocí obou metod: PCC a QM. Odhad hustot rozdělení pro kvantilové mapování byl realizován s využitím histogramů o šířce intervalu 0.01 mm. Nejprve byly analyzovány závislostní struktury. Obrázek 7.5 ukazuje korelační matice pro (A) měřená data, (B) modelová data, (C) data korigovaná metodou PCC a (D) data korigovaná metodou QM.



Obr. 7.5 Korelační matice pro (A) měřená data, (B) modelová data, (C) data korigovaná metodou PCC a (D) data korigovaná metodou QM. Řádky i sloupce matic jsou označeny identifikátorem příslušného povodí, viz tabulka 4.2.

Kovarianční matice jsou stejným způsobem zobrazeny na obrázku 7.6.



Obr. 7.6 Kovarianční matice pro (A) měřená data, (B) modelová data, (C) data korigovaná metodou PCC a (D) data korigovaná metodou QM. Řádky i sloupce matic jsou označeny identifikátorem příslušného povodí, viz tabulka 4.2.

Z obrázků 7.5 a 7.6 je možno vyčíst, že v případě korelací i kovariancí jsou matice měřených dat na pozici (A) identické s maticemi dat korigovaných metodou PCC na pozici (C). Detailní analýza pak potvrdila, že nelze nalézt detekovatelný rozdíl mezi korelacemi či kovariancemi měřených a PCC-korigovaných dat. Lze tedy konstatovat, že metoda PCC plně koriguje závislostní strukturu modelových dat. Výsledky dosažené metodou QM jsou zcela odlišné. Obrázek 7.5 ukazuje, že QM zhruba zachovává korelační strukturu modelových dat, což lze vidět porovnáním matic na pozicích (B) a (D). Modelované kovariance jsou metodou QM změněny, jak lze vidět porovnáním matic (B) a (D) na obrázku 7.6. Obecně lze říci, že kovariance jsou mírně posunuty směrem k pozorovaným hodnotám, nicméně v kovarianční struktuře po korekci metodou QM přetrvávají značné nepřesnosti, jak lze zjistit porovnáním matic na pozicích (A) a (D).

Výsledky metody PCC při korekci korelační a kovarianční struktury vyplývají přímo z definice této metody. PCC posouvá celý mnohorozměrný soubor měřených dat takovým způsobem, že jeho výsledná poloha je zcela v souladu se souřadným systémem daným hlavními komponentami měřených dat. Tento souřadný systém je však dán vlastními vektory a

vlastními čísly kovarianční matice, což znamená, že je přímo odvozen z kovarianční struktury měřených dat. Tato přímá souvislost je tedy příčinou identických kovariančních struktur měřených a korigovaných dat. Rovnost korelačních matic pak plyne z precizní korekce kovarianční matice. Pro kovarianci platí

$$\text{cov}(X, X) = \sigma^2(X) \quad (7.4)$$

kde $\text{cov}(\cdot)$ značí kovarianci. Ze vztahu (7.4) plyne, že diagonální prvky kovarianční matice jsou rovny rozptylům jednotlivých proměnných, což také znamená, že metoda PCC přirozeně koriguje rozptyl jednotlivých datových souborů. Ze současné rovnosti rozptylů a kovariancí již přímo plyne rovnost korelací, neboť korelační koeficient mezi náhodnými veličinami X a Y je dán vztahem

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\sigma^2(X)\sigma^2(Y)}} \quad (7.5)$$

kde $\rho(\cdot)$ značí korelační koeficient.

Výsledky metody QM jsou dány lineární povahou této metody, která byla prezentována a zdůvodněna v kapitole 5. Z výsledků kapitoly 5 plyne, že korekční funkce QM může být zhruba aproximována lineární funkcí. Jak je známo, pro náhodné veličiny X a Y platí

$$\rho(aX + b, cY + d) = \rho(X, Y) \quad (7.6)$$

kde a, b, c, d jsou reálná čísla. Lineární transformace tedy nemění korelaci mezi proměnnými, což je zcela v souladu s výsledky prezentovanými na obrázku 7.5, ze kterých je patrné, že QM modelové korelace nijak podstatně nezměnilo. Kovariance lineárně transformovaných veličin X a Y je pak dána vztahem

$$\text{cov}(aX + b, cY + d) = ac \text{cov}(X, Y) \quad (7.7)$$

Lineární korekce tedy kovariance mění (jak je možné pozorovat na obrázku 7.6) a její efekt závisí na sklonech korekčních funkcí. Obecně lze vypožorovat, že lineární korekce částečně opravuje modelové kovariance a posunuje je k měřeným hodnotám. To lze vysvětlit tím, že kovariance je (na rozdíl od korelace) veličina závislá na měřítku, a lineární korekce upravují měřítko („velikost“) datových souborů tím, že korigují jejich individuální rozdělení. Nekorigují ale intenzitu vztahů mezi proměnnými, z čehož plyne, že modelové kovariance jsou metodou QM korigovány pouze částečně a korelace nejsou korekcí nijak významně dotčeny. Mírné ovlivnění korelací je dáno odchylkami QM korekčních funkcí od přesného lineárního průběhu.

Dalším bodem analýzy bylo prošetření statistických ukazatelů individuálních datových souborů, reprezentujících korigované srážky na jednotlivých povodích, resp. jejich shody s ukazateli souborů měřených dat. Data korigovaná metodou PCC mají identické střední hodnoty a rozptyly jako data měřená. Obě tyto identity plynou z definice metody PCC. Rovnost středních hodnot vyplývá z toho, že počátky alternativních souřadných systémů modelových i měřených dat jsou položeny do bodů středních hodnot, viz definice metody ve vztahu 7.1. Rovnost rozptylů je pak dána precizní korekcí modelových kovariancí, jak bylo vysvětleno výše. Dále bylo ověřeno, že data korigovaná metodou QM těchto rovností

dosahují samozřejmě také. Dalším krokem bylo zhodnocení celkové shody rozdělení mezi korigovanými a měřenými soubory. K vyjádření a kvantifikaci této shody byl použit dvou-výběrový Kolmogorovův-Smirnovův test. Tento test je založen na statistice ve tvaru

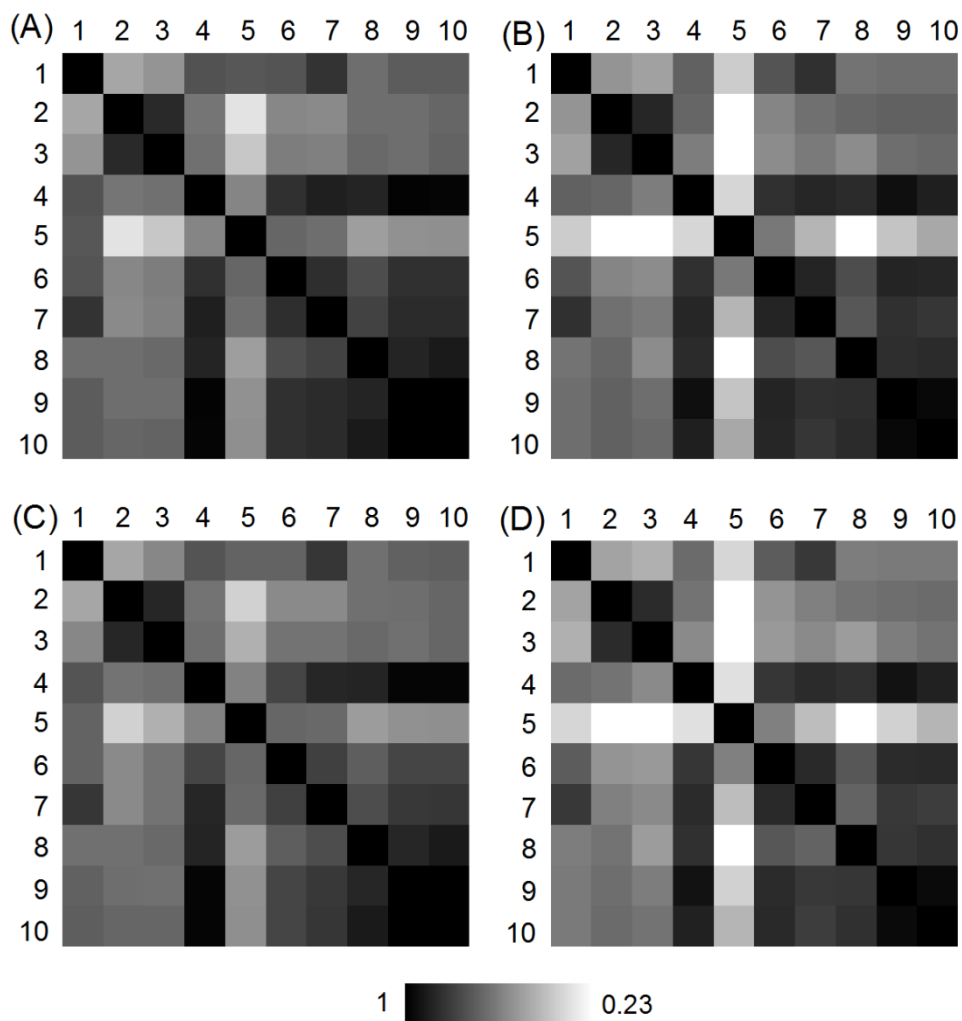
$$D = \sup_x |F_O(x) - F_T(x)| \quad (7.8)$$

kde $F_O(x)$ a $F_T(x)$ jsou distribuční funkce měřených a korigovaných dat, $\sup(\cdot)$ pak značí supremum. Testovací statistika tedy vyjadřuje maximální rozdíl distribučních funkcí obou porovnávaných souborů. V případě metody QM test pro všechna povodí prokázal, že korigovaná a měřená data mohou být považována za náhodné výběry z identických rozdělení. V případě metody PCC byla ve všech deseti případech hypotéza o shodě rozdělení zamítnuta. Tato skutečnost není nikterak překvapivá, protože metoda PCC je navržena pro korekci závislostní struktury, zatímco metoda QM pro korekci jednotlivých rozdělení. Podrobnější analýzou bylo zjištěno, že nekorigované modelové datové soubory mají průměrnou hodnotu D rovnou 0.22, kdežto soubory korigované metodou PCC mají průměrné D rovno 0.079. To značí, že PCC přináší určitý posun a částečně zlepšuje individuální rozdělení, nicméně i přes tuto skutečnost zůstávají ve výsledných souborech značné chyby v kvantilech. Dále je nutné si v souvislosti s rozdělením jednotlivých souborů uvědomit, že PCC spočívá v mírné rotaci modelových dat (doplněné o změny měřítek os). Srážkové úhrny jsou ale veličinou ohraničenou zdola nulovou hodnotou, takže rotovaný soubor zcela jistě vykazuje některé záporné hodnoty. Je tedy možné konstatovat, že ačkoliv data korigovaná metodou PCC vykazují shodné závislostní struktury, střední hodnoty a rozptyly jako data měřená, kvůli chybám v kvantilech a přítomnosti záporných hodnot nejsou přímo využitelná pro dopadové studie. Řešení tohoto problému je popsáno v následující kapitole.

7.5 Kombinace metod PCC a QM

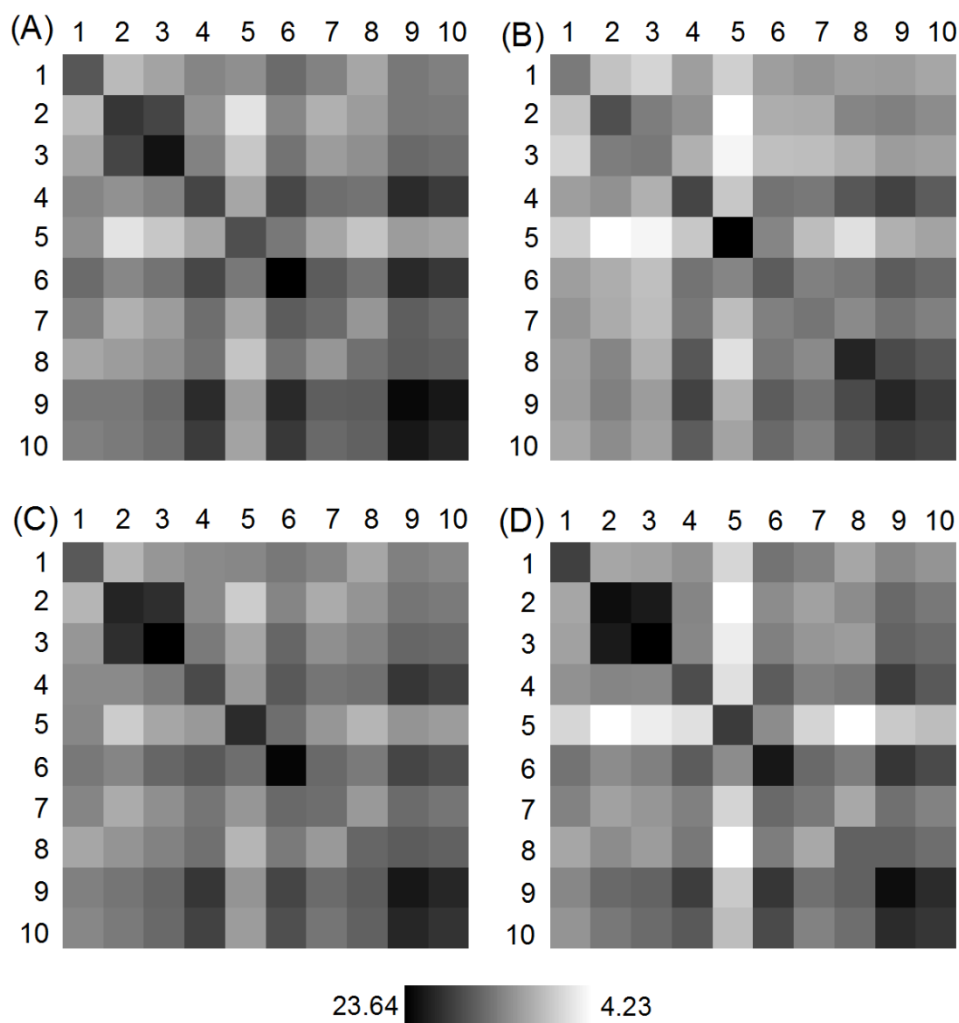
Při řešení problému s rozdělení dat korigovaných metodou PCC lze využít toho, že metody PCC a QM mají doplňující se vlastnosti. Zatímco PCC koriguje závislostní strukturu modelových dat a zanechává v jednotlivých souborech chyby v kvantilech, efekt metody QM je zcela opačný. Výhody obou metod mohou být spojeny následujícím postupem: vícerozměrná data jsou nejprve korigována metodou PCC a následně jsou jednotlivé datové soubory korigovány metodou QM. Podstata této kombinované procedury spočívá v tom, že QM opraví jednotlivá rozdělení a při tom výrazně neporuší korelační a kovarianční struktury, které byly korigovány v prvním kroku metodou PCC. To je dáno lineární povahou QM, která byla diskutována výše. Kombinovaná procedura tedy koriguje jak závislostní struktury, tak jednotlivá rozdělení. Aplikací QM je také vyřešen problém se zápornými hodnotami, které byly do dat vneseny rotací v rámci procedury PCC. V dalším textu bude kombinovaná metoda označována jako PCC+QM.

Ověření účinnosti kombinované procedury bylo nejprve provedeno stejně jako v případě výše prezentované analýzy jednotlivých metod, tedy na kompletních datových souborech bez rozdělení na kalibraci a validaci. Tímto testem byly potvrzeny teoretické závěry a bylo ověřeno, že kombinovaná procedura koriguje závislostní struktury i jednotlivá rozdělení. Dalším krokem bylo provedení kalibračního / validačního testu. Kombinovaná metoda PCC+QM a dále pro porovnání i samotné QM byly kalibrovány s použitím dat z let 1970 – 1987 a aplikovány na data z let 1988 – 2005. Obrázek 7.7 ukazuje korelační matice z validačního období pro (A) měřená data, (B) modelová data, (C) data korigovaná kombinovanou procedurou PCC+QM a (D) data korigovaná standardním QM.



Obr. 7.7 Korelační matice dat z validačního období pro (A) měřená data, (B) modelová data, (C) data korigovaná kombinovanou metodou PCC+QM a (D) data korigovaná metodou QM. Řádky i sloupce matic jsou označeny identifikátorem příslušného povodí, viz tabulka 4.2.

Z obrázku je patrné, že zatímco kombinovaná procedura PCC+QM modelové korelace z velké části opravuje, metoda QM chyby modelu zachovává. Stejný efekt je viditelný při pohledu na kovarianční matice z validačního období, které jsou stejným způsobem znázorněny na obrázku 7.8.

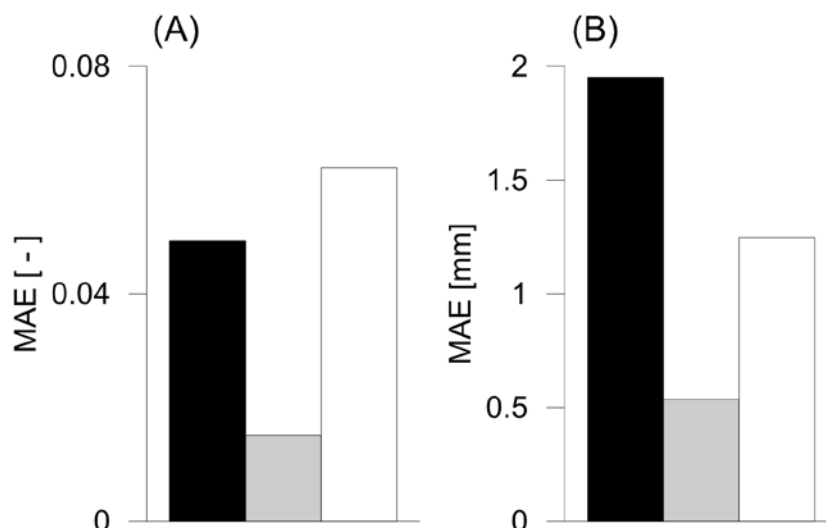


Obr. 7.8 Kovarianční matice dat z validačního období pro (A) měřená data, (B) modelová data, (C) data korigovaná kombinovanou metodou PCC+QM a (D) data korigovaná metodou QM. Řádky i sloupce matic jsou označeny identifikátorem příslušného povodí, viz tabulka 4.2.

Výsledky prezentované na obrázcích 7.7 a 7.8 jsou sumarizovány na obrázku 7.9, který ukazuje hodnoty střední absolutní chyby (MAE, „mean absolute error“), která byla pro hodnoty korelačních koeficientů a hodnoty kovariancí spočítána pro výsledky obou korekčních metod podle vzorce

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |c_O - c_T| \quad (7.9)$$

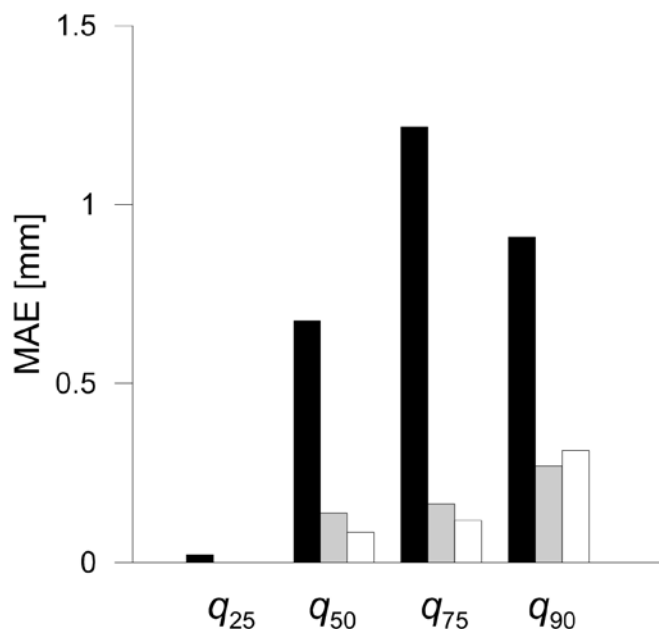
kde c_O a c_T jsou jednotlivé hodnoty korelačních koeficientů, resp. kovariancí, měřených a korigovaných dat.



Obr. 7.9 Střední absolutní chyby (MAE) korelačních koeficientů (A) a kovariancí (B) pro modelová data (černé sloupce), data korigovaná metodou PCC+QM (šedé sloupce) a data korigovaná metodou QM (bílé sloupce).

Výsledky potvrzují teoretické předpoklady. Zatímco metoda QM zhruba zachovává modelové korelace a mírně koriguje modelové kovariance, kombinovaná metoda PCC+QM přináší podstatnou korekci modelové chyby závislostních struktur. Zbývající chyby ve výsledcích PCC+QM jsou zapříčiněny nestacionaritou časových řad mezi kalibrační a validační fází a odklonem korekčních funkcí QM od přesného lineárního průběhu.

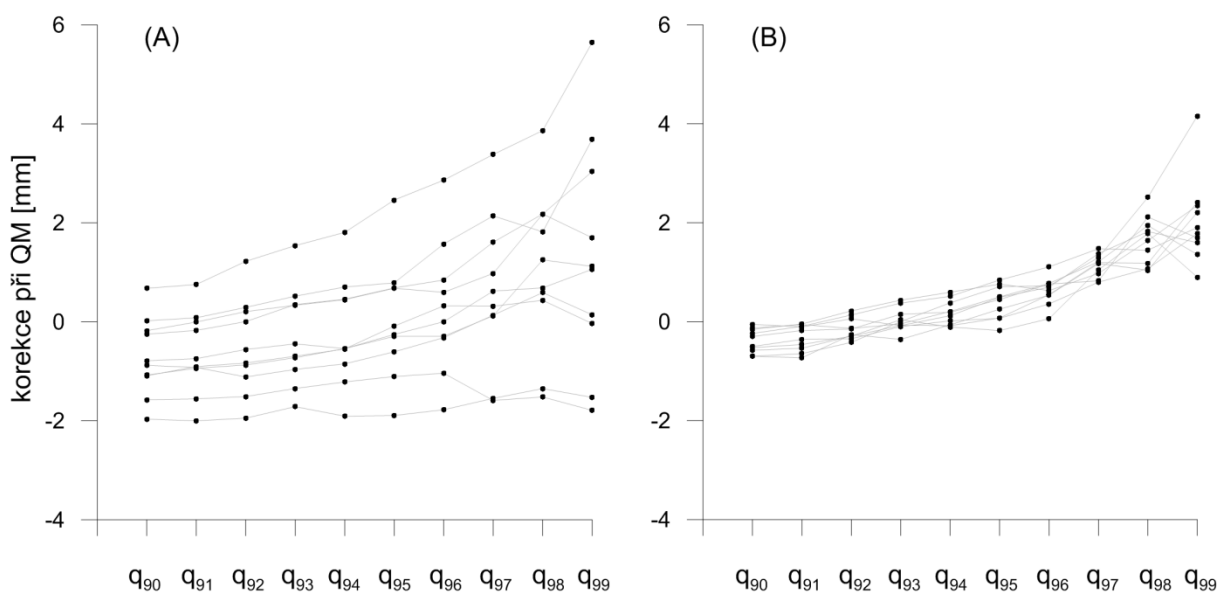
Dále byla zkoumána efektivita metod při korekci rozdělení jednotlivých datových souborů. Pro všechny korigované datové soubory z validačního období byly spočítány kvantily q_{25} , q_{50} , q_{75} a q_{90} . Korigované kvantily byly posléze porovnány s kvantily měřených dat, výsledky tohoto porovnání jsou vyneseny opět ve formě střední absolutní chyby pro jednotlivé kvantily pro obě korekční metody na obrázku 7.10.



Obr. 7.10 Střední absolutní chyby (MAE) jednotlivých kvantilů pro modelová data (černé sloupce), data korigovaná metodou PCC+QM (šedé sloupce) a data korigovaná metodou QM (bílé sloupce).

Výsledky ukazují, že v případě korekce individuálních rozdělání jsou obě prošetřované metody srovnatelně efektivní a obě výrazně snižují modelovou chybu. Podrobnější pohled na obrázek 7.10 ukazuje, že v případě kvantilů q_{25} a q_{50} je mírně úspěšnější metodou kvantilové mapování. To může být zapříčiněno tím, že chyba dosažená metodou QM je způsobena pouze nestacionaritou jednotlivých rozdělání. Naproti tomu chyba kombinované metody PCC+QM je ovlivněna dvěma zdroji nejistoty, protože ke změnám v jednotlivých rozděláních se přidává také změna závislostní struktury, která zde nastává minimálně z důvodu přirozené klimatické variability. Oba tyto zdroje nestacionarity se mohou promítnout do efektivity metody při korekci jednotlivých rozdělání. Obecně lze konstatovat, že PCC+QM je náchylnější k nestacionárním podmínkám než klasické QM, což je dáno vyšší složitostí a komplexností kombinované metody. Tento fakt je zřejmě obecně platný a souvisí také s obsahem kapitoly 6, kde motivací k odvození nových metod lineární korekce bylo právě zjednodušení korekční funkce při snaze o potlačení vlivu nestacionárních podmínek.

Z obrázku 7.10 také plyne, že pro kvantil q_{90} je menší chyby dosaženo naopak použitím metody PCC+QM. Tento výsledek byl podnětem pro další analýzu. Kvantilové mapování je obecně spojeno s problémem při korekci vysokých kvantilů, což je dáno vzácností vysokých srážkových úhrnů. Přesná funkce QM může v oblasti vysokých kvantilů vykazovat nestabilní chování, projevující se náhlými změnami směru a velikosti korekce pro jednotlivé prostorové body. Následkem toho mohou být prostorové vztahy pro vysoké kvantily v korigovaných datech zcela nekonzistentní s měřenými daty. Obrázek 7.11 ukazuje korekci (posun v rámci korekce) vysokých kvantilů pro (A) čisté QM a (B) PCC+QM, hodnoty kvantilů byly odečteny z histogramů použitých při kvantilovém mapování.



Obr. 7.11 Korekce vysokých kvantilů pro (A) metodu QM a (B) metodu PCC+QM. Kvantily z jednotlivých povodí jsou spojeny šedými úsečkami.

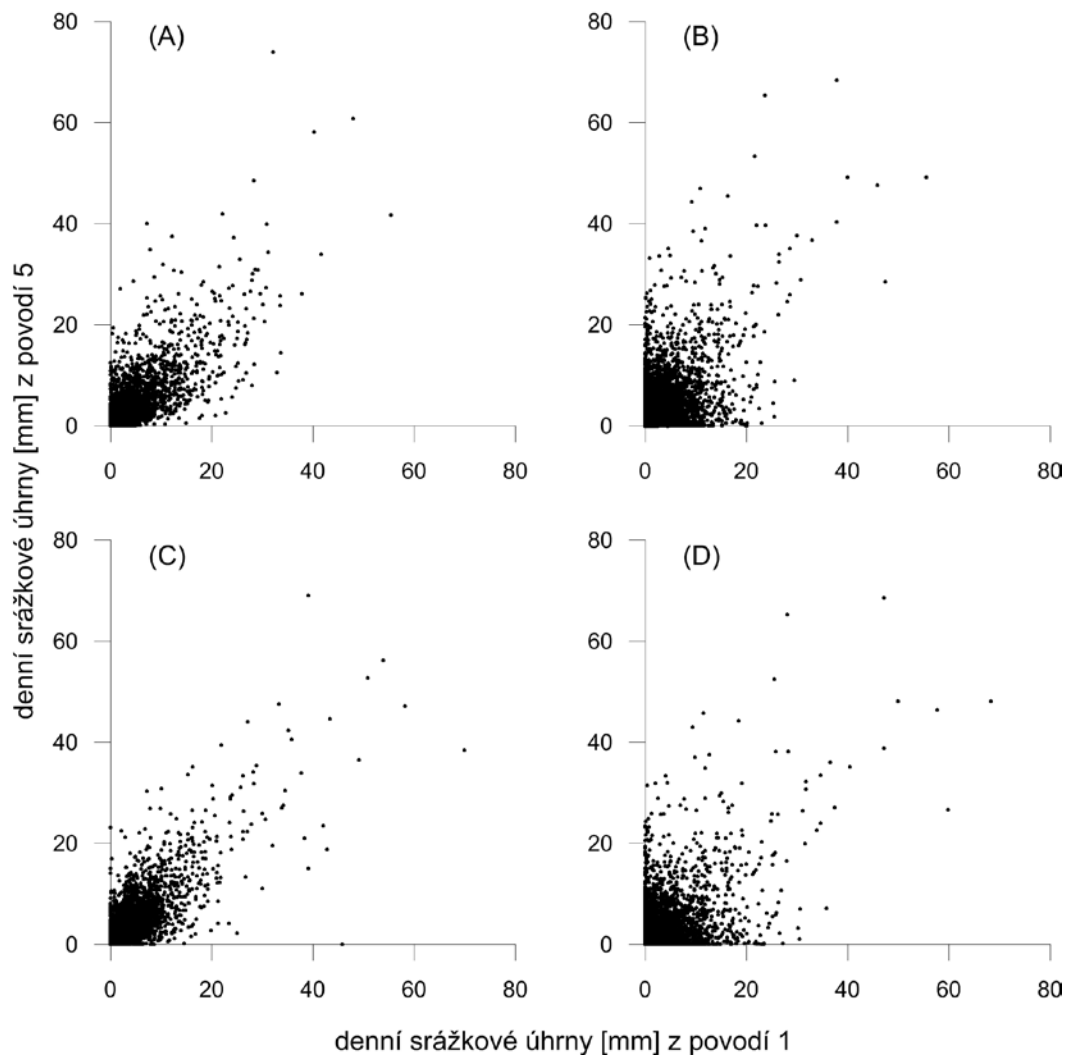
Obrázek ukazuje, že vložení procedury PCC před kvantilové mapování má na korekci vysokých kvantilů stabilizační efekt. Korekce v jednotlivých povodích má pro jednotlivé kvantily v případě PCC+QM mnohem menší rozptyl než v případě samotného QM. Dalším krokem pro posouzení kvality vysokých kvantilů byla analýza jejich korelací. Z validačních dat byly separovány dny, kdy na alespoň jednom povodí nastala srážková událost rovnající se

alespoň kvantilu q_{90} z příslušného povodí. Z těchto dat byly spočítány korelační matice pro měřená a korigovaná data. Celková chyba prostorových korelací byla spočítána jako

$$e = \|\mathbf{C}_O - \mathbf{C}_T\| \quad (7.10)$$

kde \mathbf{C}_O a \mathbf{C}_T jsou korelační matice měřených a korigovaných dat, $\|\cdot\|$ je Frobeniova maticová norma. Hodnota chyby (7.10) pro čisté QM byla 1.275, pro proceduru PCC+QM byla 0.735. To opět naznačuje, že prostorové vztahy vysokých srážek jsou po aplikaci PCC blíže realitě než v případě použití samotného QM. Tyto výsledky ovšem nic nemění na skutečnosti, že QM je stále nedílnou součástí navrhované procedury, tudíž kvalita výsledných extrémů je ovlivnitelná konkrétní implementací QM (histogramy, gama rozdělení, lineární funkce...) a je stále citlivá na výskyt extrémních událostí v kalibračním období.

Efekt PCC na výsledek korekce je jasně viditelný při detailní analýze dat z povodí 5. Modelové korelace a kovariance mezi tímto a ostatními povodími jsou silně ovlivněny chybou modelu, což je patrné z obrázků 7.5 a 7.6. Obrázek 7.12 ukazuje validační data z povodí 5 vynesena proti datům z povodí 1 pro (A) měřená data, (B) modelová data, (C) data korigovaná metodou PCC+QM a (D) data korigovaná metodou QM.



Obr. 7.12 Validační data z povodí 1 vynesena proti datům z povodí 5: (A) měřená data, (B) modelová data, (C) data korigovaná metodou PCC+QM a (D) data korigovaná metodou QM.

Efekt korekce závislostní struktury je jasně patrný v rozdílu ve tvaru datových souborů. Zatímco data korigovaná metodou QM zachovávají „trojúhelníkový“ tvar modelových dat, procedura PCC mění tvar datového souboru, který je po korekci výrazně podobnější měřeným datům.

8 Souhrn a závěry

V první části disertační práce byla provedena literární rešerše problematiky globálního a regionálního klimatického modelování. Byly představeny principy globálních modelů a všechny jejich hlavní části, v rámci problematiky vnořených regionálních modelů bylo poukázáno na některé problematické aspekty tohoto konceptu. Dále byla provedena rešerše nejpoužívanějších metod korekce systematických chyb výstupů z regionálních modelů, která se podrobněji soustředila na nejpoužívanější (tzv. přímé) metody korekce.

Prvním bodem praktické části disertační práce byla demonstrace lineární povahy korekce a podobnosti jednotlivých korekčních metod. Byl proveden teoretický rozbor založený na výběru rozdělení pravděpodobnosti při kvantilovém mapování a teoretické závěry byly ověřeny na reálných měřených a modelových datech. Tyto výsledky se staly podkladem pro další studie.

Druhým bodem praktické části bylo odvození nových metod korekce založených na lineární konverzi mezi dvěma gama rozděleními. Motivací pro odvození lineárních metod bylo přesvědčení, že při nestacionárních podmínkách je užitečnější vystihnout lineární trend korekce, než odvodit funkci, která má schopnost přesně korigovat data z kalibračního období, ale je příliš citlivá na případné změny. Bylo ověřeno, že při určitém stupni nestacionarity mezi kalibrací a aplikací jsou nově odvozené metody efektivnější než kvantilové mapování. Bylo též poukázáno na to, že pokud je nestacionarita snížena, nebo pokud se výchozí datové soubory odchylují od gama rozdělení (např. při interpolacích dat na pravidelnou prostorovou mřížku), dojde ke snížení efektivity nově odvozených metod.

Třetím bodem praktické části bylo odvození procedury, která koriguje závislostní strukturu ve vícerozměrném datovém souboru. Metoda je založena na konverzi mezi alternativními souřadnými systémy, které jsou odvozeny z hlavních komponent měřených a modelových dat. Bylo ověřeno, že metoda koriguje kompletní korelační a kovarianční matice modelových dat. Analýzou kvantilů bylo posléze zjištěno, že nově odvozená procedura nedostatečně koriguje jednotlivá pravděpodobnostní rozdělení. Proto byla nově odvozená procedura vložena jako pre-processing před klasickou korekční metodou (v tomto případě kvantilové mapování). Toto spojení metod využívá poznatku o lineární povaze korekce, protože úprava jednotlivých rozdělení již výrazně neovlivní vazby mezi datovými soubory, korigované novou metodou.

Nově odvozenou vícerozměrnou metodu lze teoreticky použít pro korekci závislostní struktury jakéhokoliv datového souboru, neboť konverze mezi souřadnými systémy je omezena pouze hypotetickými způsoby, které v reálných přírodních datech nemohou nastat (např. při stoprocentní korelaci mezi některými z dílčích souborů by modelová a měřená data mohla mít jiný počet dimenzí atd.). Metodu lze např. použít pro smíšená data, obsahující teplotní i srážkové datové soubory, popř. i jiné meteorologické veličiny, např. vlhkost vzduchu atd.

Otázkou k prošetření je stabilita závislostní struktury při dlouhých klimatických projekcích. Tato otázka byla již autorem práce částečně zkoumána a podle prvních výsledků se zdá, že korelace a kovariance se z hlediska stability chovají odlišně – zatímco kovarianční struktura je ovlivnitelná klimatickým signálem a odlehlými hodnotami, korelační struktura se zdá být mnohem stabilnější. Bylo by tedy zřejmě užitečné postavit odvození alternativních souřadnic spíše na korelačních maticích, nicméně potvrzení těchto dílčích výsledků vyžaduje další výzkum.

Je nutné zdůraznit, že validace metod s použitím měřených a modelových dat sloužila v rámci disertační práce pouze k analýze metod a k demonstraci jejich efektu a z výsledků nebyly činěny žádné klimatologické závěry. Při skutečném použití v dopadových studiích by bylo nutno brát v úvahu některé další problematické aspekty korekcí. Např. Maraun (2013) ukázal, že čistě deterministické metody korekce mohou vést ke sporným výsledkům, pokud jsou použity jako downscaling z hrubšího měřítka modelových výstupů na bodové měřítka meteorologických stanic. To je dáno tím, že modelové simulace neobsahují informaci o prostorové a časové variabilitě na podrobnějších měřících. Korigované časové řady uvnitř jednoho modelového grid-boxu tudíž uniformně adoptují časovou variabilitu modelových dat, což může vést k vážným důsledkům, např. k nadhodnocení rizika povodní (tzv. „inflation problems“). Maraun (2013) v takovém případě doporučuje doplnit výslednou korekci šumem pomocí stochastického generátoru.

Závěrem lze konstatovat, že disertační práce splnila vytčené cíle. Byly odvozeny metody korekce odolnější vůči vlivu nestacionarity a byla odvozena procedura korigující závislostní struktury. Nicméně v oblasti korekcí modelových simulací zůstává k vyřešení mnoho problematických bodů. V korigovaných datech například zůstávají značné reziduální chyby v jiných časových měřících, než ve kterých byly korekce prováděny (např. Haerter et al., 2011). Korekce tedy v budoucnu zřejmě zůstanou nutnou, ale značně problematickou součástí studií hodnotících vliv dopadu klimatické změny.

Seznam literatury

Addor N, Rohrer M, Furrer R, Seibert J. 2016. Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment. *Journal of Geophysical Research – Atmospheres*. 121(5):2075-2089. doi:10.1002/2015JD024040.

Anděl, J. (2007): *Statistické metody*. Matfyzpress, Praha.

Anisimov, O.A. (2007): Potential feedback of thawing permafrost to the global climate system through methane emission. *Environmental Research Letters*, 2(4): 045016. doi: 10.1088/1748-9326/2/4/045016.

Antic, S., Laprise, R., Denis, B., de Elia, R. (2004): Testing the downscaling ability of a one-way nested regional climate model in regions of complex topography. *Climate Dynamics*, 23(5), 473-493. doi: 10.1007/s00382-004-0438-5.

Arnell, N.W., Hudson, D.A., Jones, R.G. (2003) Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa. *Journal of Geophysical Research* 108(D16), 4519. doi:10.1029/2002JD002782.

Arora, V.K., Boer, G.J. (1999): A variable velocity flow routing algorithm for GCMs. *Journal of Geophysical Research – Atmospheres*. 104 (D24): 30965–30979. doi:10.1029/1999JD900905

Bader, D.C., Covey, C., Gutowski, W.J., Held, I.M., Kunkel, K.E., Miller, R.L., Tokmakian, R.T., Zhang, M.H. (2008): *Climate Models: An Assessment of Strengths and Limitations*. US Department of Energy Publications. Paper 8.

Baigorria, G.A., Jones, J.W., Shin, D., Mishra, A., O'Brien, J.J. (2007): Assessing uncertainties in crop model simulations using daily bias-corrected Regional Circulation Model outputs. *Climate Research*, 34:211-222. doi: 10.3354/cr00703.

Berg, P., Feldmann, H., Paintz, H.J. (2012): Bias correction of high resolution regional climate model data. *Journal of Hydrology*, 448-449:80-92. doi: 10.1016/j.jhydrol.2012.04.026.

Block, P.J., Filho, F.A.S., Sun, L., Kwon, H. (2009): A streamflow forecasting framework using multiple climate and hydrological models. *Journal of the American Water Resources Association*, 45(4):828-843. doi: 10.1111/j.1752-1688.2009.00327.x.

Boe, J., Terray, L., Habets, F., Martin, E. (2007): Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, 27(12): 1643-1655. doi: 10.1002/joc.1602.

Bretherton, C.S., Smith, C, Wallace, J.M. (1992): An intercomparison of methods for finding coupled patterns in climate data. *Journal of Climate*, 5(6): 541-560. doi: 10.1175/1520-0442(1992)005<0541:AIOMFF>2.0.CO;2.

Castellvi, F., Mormeneo, I., Perez, P.J. (2004): Generation of daily amounts of precipitation from standard climatic data: a case study for Argentina. *Journal of Hydrology*, 289(1-4): 286-302. doi: 10.1016/j.jhydrol.2003.11.027.

Collins, W.D., Bitz, C.M., Blackmon, M.L., Bonan, G.B., Bretherton, C.S., Carton, J.A., Chang, P., Doney, S.C., Hack, J.J., Henderson, T.B., Kiehl, J.T., Large, W.G., McKenna, D.S., Santer, B.D., Smith, R.D. (2006): The Community Climate System Model version 3 (CCSM3). *Journal of Climate*, 19(11): 2122–2143. doi: 10.1175/JCLI3761.1.

Curry, J.A., Schramm, J.L., Ebert, E.E. (1995): Sea ice-albedo climate feedback mechanism. *Journal of Climate*, 8(2): 240–247. doi:10.1175/1520-0442(1995)008<0240:SIACFM>2.0.CO;2

Delworth, T.L., Broccoli, A.J., Rosati, A., Stouffer, R.J., Balaji, V., Beesley, J.A., Cooke, W.F., Dixon, K.W., Dunne, J., Dunne, K.A., Durachta, J.W., Findell, K.L., Ginoux, P., Gnanadesikan, A., Gordon, C.T., Griffies, S.M., Gudgel, R., Harrison, M.J., Held, I.M., Hemler, R.S., Horowitz, L.W., Klein, S.A., Knutson, T.R., Kushner, P.J., Langenhorst, A.R., Lee, H.C., Lin, S.J., Lu, J., Malyshev, S.L., Milly, P.C.D., Ramaswamy, V., Russell, J., Schwarzkopf, M.D., Shevliakova, E., Sirutis, J.J., Spelman, M.J., Stern, W.F., Winton, M., Wittenberg, A.T., Wyman, B., Zeng, F., Zhang, R. (2006): GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *Journal of Climate*, 19(5): 643-674. doi: 10.1175/JCLI3629.1.

Denis, B., Cote, J., Laprise, R. (2001): Spectral decomposition of two-dimensional atmospheric fields on limited-area domains using the discrete cosine transform (DCT). *Monthly Weather Review*, 130(7): 1812-1829. doi: 10.1175/1520-0493(2002)130<1812:SDOTDA>2.0.CO;2.

Denis, B., Laprise, R., Caya, D., Cote, J. (2002): Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment. *Climate Dynamics*, 18(8): 627-646. doi: 10.1007/s00382-001-0210-0.

Déqué, M. (2007): Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global and Planetary Change*, 57:16-26. doi: 10.1016/j.gloplacha.2006.11.030

Déqué, M., Drevet, C., Braun, A., Cariolle, D. (1994): The ARPEGE/ifs atmosphere model – a contribution to the French community climate modeling. *Climate Dynamics*, 10(4-5): 249-266.

Dickinson, R.E., Errico, R.M., Giorgi, F., Bates, G.T. (1989): A regional climate model for the western United States. *Climatic Change*, 15(3): 383-422.

Ducharne, A., Golaz, C., Leblois, E., Laval, K., Polcher, J., Ledoux, E., Marsily, G. (2003): Development of a high resolution runoff routing model, calibration, and application to assess runoff from the LMD GCM. *Journal of Hydrology*, 280(1-4): 207–228. doi:10.1016/S0022-1694(03)00230-0.

Eby, M., Weaver, A.J., Alexander, K., Zickfeld, K., Abe-Ouchi, A., Cimadoribus, A.A., Cresspin, E., Drijfhout, S.S., Edwards, N.R., Eliseev, A.V., Feulner, G., Fichet, T., Forest, C.E., Goosse, H., Holden, P.B., Joos, F., Kawamiya, M., Kicklighter, D., Kienert, H., Matsumoto, K., Mokhov, I.I., Monier, E., Olsen, S.M., Pedersen, J.O.P., Perrette, M., Philippon-Berthier, G., Ridgwell, A., Schlosser, A., von Deimling, T.S., Shaffer, G., Smith, R.S., Spahni, R., Sokolov, A.P., Steinacher, M., Tachiiri, K., Tokos, K., Yoshimori, M., Zeng, N., Zhao, F. (2013): Historical and idealized

climate model experiments: an intercomparison of Earth system models of intermediate complexity. *Climate of the Past*, 9(3): 1111-1140. doi: 10.5194/cp-9-1111-2013.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., Liebert, J. (2012): "Should we apply bias correction to global and regional climate model data?" *Hydrology and Earth System Sciences*, 16(9): 3391-3404. Doi: 10.5194/hess-16-3391-2012.

Feser, F., Rockel, B., von Storch, H., Winterfeldt, J., Zahn, M. (2011): Regional climate models add value to global model data – a review and selected examples. *Bulletin of the American Meteorological Society*, 92(9): 1181-1192. doi: 10.1175/2011BAMS3061.1.

Foley, A.M. (2010): Uncertainty in regional climate modelling: A review. *Progress in Physical Geography*, 34(5): 647-670. doi:10.1177/0309133310375654.

Giorgi, F. (1990): Simulation of regional climate using a limited area model nested in a general circulation model. *Journal of Climate*, 3(9): 941-963. doi:10.1175/1520-0442(1990)003<0941:SORCUA>2.0.CO;2.

Giorgi, F., Jones, C., Ghassem, R. (2009): Addressing climate information needs at the regional level: the CORDEX framework. *World Meteorological Organization (WMO) Bulletin* 58(3): 175-183.

Giorgi, F., Mearns, L.O. (2002): Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method. *Journal of Climate*, 15(10): 1141-1158. doi:10.1175/1520-0442(2002)015<1141:COAURA>2.0.CO;2.

Giorgi, F., Mearns, L.O. (2003): Probability of regional climate change based on the Reliability Ensemble Averaging (REA) method. *Geophysical Research Letters*, 30(12), 1629. doi:10.1029/2003GL017130.

Graham, L.P., Hagemann, S., Jaun, S., Beniston, M. (2007): On interpreting hydrological change from regional climate models. *Climatic Change*, 81: 97-122. doi: 10.1007/s10584-006-9217-0.

Haerter, J., Hagemann, S., Moseley, C., Piani, C. (2011): Climate model bias correction and the role of timescales. *Hydrology and Earth System Sciences* 15(3): 1065–1079. doi: 10.5194/hess-15-1065-2011.

Hay, L.E., Clark, M.P., Pagowski, M., Leavesley, G.H., Gutowski, W.J. (2006): One-way coupling of an atmospheric and a hydrologic model in Colorado. *Journal of Hydrometeorology*, 7(4): 569–589. doi:10.1175/JHM512.1.

Hay, L.E., Clark, M.P., Wilby, R.L., Gutowski Jr. W.J., Leavesley, G.H., Pan, Z., Arritt, R.W., Takle, E.S. (2002): Use of regional climate model output for hydrologic simulation. *Journal of Hydrometeorology*, 3:571-590.

Hoffmann, H., Rath, T. (2012): Meteorologically consistent bias correction of climate time series for agricultural models. *Theoretical and Applied Climatology* 110(1-2): 129-141. doi:10.1007/s00704-012-0618-x.

Houghton, J. *Globální oteplování* (1998): Academia, Praha.

Chen, J., Brissette, F.P., Lucas-Picher, P. (2015): Assessing the limits of bias-correcting climate model outputs for climate change impact studies. *Journal of Geophysical Research – Atmospheres* 120(3): 1123-1136. doi: 10.1002/2014JD022635.

Chen, C., Haerter, J.O., Hagemann, S., Piani, C. (2011): On the contribution of statistical bias correction to the uncertainty in the projected hydrological cycle. *Geophysical Research Letters*, 38(L20403). doi: 10.1029/2011GL049318.

Choi, S.C., Wette, R. (1969): Maximum likelihood estimation of the parameters of the gamma distribution and their bias. *Technometrics* 11(4): 683-690. doi:10.2307/1266892.

Im, E.S., Park, E.H., Kwon, W.T., Giorgi, F. (2006): Present climate simulation over Korea with a regional climate model using a one-way double-nested system. *Theoretical and Applied Climatology*, 86(1-4): 187-200. doi: 10.1007/s00704-005-0215-3.

Ines, A.W.M, Hansen, J.W. (2006): Bias correction of daily GCM rainfall for crop simulation studies. *Agricultural and Forest Meteorology*, 138:44-53. doi: 10.1016/j.agrformet.2006.03.009.

Jacob, D. (2001) A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. *Meteorology and Atmospheric Physics*, 77:61–73. doi:10.1007/s007030170017.

Jacob, D., Podzun, R. (1997): Sensitivity studies with the regional climate model REMO. *Meteorology and Atmospheric Physics*, 63(1-2): 119-129. doi: 10.1007/BF01025368.

Johns, T.C., Durman, C.F., Banks, H.T., Roberts, M.J., McLaren, A.J., Ridley, J.K., Senior, C.A., Williams, K.D., Jones, A., Rickard, G.J., Cusack, S., Ingram, W.J., Crucifix, M., Sexton, D.M.H., Joshi, M.M., Dong, B.W., Spencer, H., Hill, R.S.R., Gregory, J.M., Keen, A.B., Pardaens, A.K., Lowe, J.A., Bodas-Salcedo, A., Stark, S., Searl, Y. (2006): The new Hadley Centre Climate Model (HadGEM1): Evaluation of coupled simulations. *Journal of Climate*, 19(7): 1327-1353. doi: 10.1175/JCLI3712.1.

Jolliffe, I.T. (2002): *Principal Component Analysis*, second edition. Springer Series in Statistics, New York. Springer-Verlag, New York.

Jones, R.G., Murphy, J.M., Noguer, M. (1995): Simulation of climate change over Europe using a nested regional-climate model. I. Assessment of present climate, including sensitivity to location of lateral boundaries. *Quarterly Journal of the Royal Meteorological Society*, 121(526): 1413-1450. doi:10.1002/qj.49712152610.

Kalvová, J., Kašpárek, L., Janouš, D., Žalud, Z., Kazmarová H. (2002): *Národní klimatický program ČR svazek 32: Scénáře změny klimatu na území České republiky a odhady dopadů*

klimatické změny na hydrologický režim, sektor zemědělství, sektor lesního hospodářství a na lidské zdraví v ČR. ČHMÚ.

Leander, R., Buishand, T.A. (2007): Resampling of regional climate model output for the simulation of extreme river flows. *Journal of Hydrology*, 332:487-496. doi: 10.1016/j.jhydrol.2006.08.006.

Lee, H.C., Rosati, A., Spelman, M.J. (2006): Barotropic tidal mixing effects in a coupled climate model: Ocean conditions in the northern Atlantic. *Ocean Modelling*, 11(3-4): 464–477. doi:10.1016/j.ocemod.2005.03.003.

Lenderink, G., Buishand, A., van Deursen, W. (2007) Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. *Hydrology and Earth System Sciences*, 11(3): 1143-1159.

Leung, L.R., Mearns, L.O., Giorgi, F., Wilby, R.L. (2003): Regional climate research - Needs and opportunities. *Bulletin of the American Meteorological Society*, 84(1): 89-95. doi:10.1175/BAMS-84-1-89.

Liston, G.E., (2004): Representing subgrid snow cover heterogeneities in regional and global models. *Journal of Climate*, 17(6): 1381–1397. doi:10.1175/1520-0442(2004)017<1381:RSSCHI>2.0.CO;2.

Liu, M., Rajagopalan, K., Chung, S.H., Jiang, X., Harrison, J., Nergui, T., Guenther, A., Miller, C., Reyes, J., Tague, C., Choate, J., Salathe, E.P., Stoeckle, C.O., Adam, J.C. (2014): What is the importance of climate model bias when projecting the impacts of climate change on land surface processes? *Biogeosciences* 11(10): 2601-2622. doi:10.5194/bg-11-2601-2014.

Lorenz, P., Jacob, D. (2005): Influence of regional scale information on the global circulation: A two-way nesting climate simulation. *Geophysical Research Letters*, 32(18): L18706. doi: 10.1029/2005GL023351.

Maraun, D. (2012): Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation. *Geophysical Research Letters*, 39: L06706. doi:10.1029/2012GL051210.

Maraun, D. (2013): Bias correction, quantile mapping, and downscaling: revisiting the inflation issue. *Journal of Climate* 26(6): 2137-2143. doi: 10.1175/JCLI-D-12-00821.1.

Mitchell, T.D., Hulme, M. (1999): Predicting regional climate change: living with uncertainty. *Progress in Physical Geography*, 23(1): 57-58. doi: 10.1191/030913399672023346.

Mo, K.C., Schemm, J.K., Juang, H.M.H., Higgins, R.W., Song, Y.C. (2005): Impact of model resolution on the prediction of summer precipitation over the United States and Mexico. *Journal of Climate* 18, 3910-3927. doi: 10.1175/JCLI3513.1.

Monjo, R., Caselles, V., Chust, G. (2012): Alternative model for precipitation probability distribution: application to Spain. *Climate Research*, 51(1): 23-33. doi: 10.3354/cr01055.

- Monjo, R., Chust, G., Caselles, V. (2014): Probabilistic correction of RCM precipitation in the Basque Country (Northern Spain). *Theoretical and Applied Climatology*, 117:317-329. doi: 10.1007/s00704-013-1008-8.
- Moore, K., Pierson, D., Pettersson, K., Schneiderman, E., Samuelsson, P. (2008): Effects of warmer world scenarios on hydrologic inputs to Lake Malaren, Sweden and implications for nutrient loads. *Hydrobiologia*, 599: 191-199. doi: 10.1007/s10750-007-9197-8.
- Nakicenovic, N., Swart, R. (2000): *Special report on emissions scenarios*. Cambridge University Press, Cambridge, UK.
- Pan, Z., Christensen, J.H., Arritt, R.W., Gutowski, W.J., Takle, E.S., Otieno, F. (2001): Evaluation of uncertainties in regional climate change simulations. *Journal of Geophysical Research – Atmospheres*, 106(D16): 17735-17751. doi:10.1029/2001JD900193.
- Piani, C., Haerter, J.O. (2012): Two dimensional bias correction of temperature and precipitation copulas in climate models. *Geophysical Research Letters* 39: L20401. doi:10.1029/2012GL053839.
- Piani, C., Haerter, J.O., Coppola, E. (2010): Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology*, 99:187-192. doi: 10.1007/s00704-009-0134-9.
- Piani, C., Weedon, G.P., Best, M., Gomes, S.M., Viterbo, P., Hagemann, S., Haerter, J.O. (2010): Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *Journal of Hydrology* 395: 199-215. doi: 10.1016/j.jhydrol.2010.10.024.
- Raisanen, J., Palmer, T.N. (2001): A probability and decision-model analysis of a multimodel ensemble of climate change simulations. *Journal of Climate*, 14(15): 3212-3226. doi: 10.1175/1520-0442(2001)014<3212:APADMA>2.0.CO;2.
- Rajagopalan, B., Lall, U., Tarboton, D.G. (1997): Evaluation of kernel density estimation methods for daily precipitation resampling. *Stochastic hydrology and hydraulics*, 11(6):523-547. doi: 10.1007/BF02428432.
- Rignot, E., Kanagaratnam, P. (2006): Changes in the velocity structure of the Greenland ice sheet. *Science*, 311, 986–990. doi: 10.1126/science.1121381.
- Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., Tompkins, A. (2003): *The atmospheric general circulation model ECHAM 5. PART I: Model description*. MPI Report 349, Max Planck Institute for Meteorology, Hamburg, Germany.
- Rojas, R., Feyen, L., Dosio, A., Bavera, D. (2011): Improving pan-European hydrological simulations of extreme events through statistical bias correction of RCM-driven climate. *Hydrology and Earth System Sciences*, 15: 2599-2620. doi: 10.5194/hess-15-2599-2011.

- Rossow, W.B., Schiffer, R.A. (1991): ISCCP cloud data product. *Bulletin of the American Meteorological Society*, 72(1): 2-20. doi: 10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2.
- Rummukainen, M. (2010): State-of-the-art with regional climate models. *Wiley Interdisciplinary Reviews: Climate Change*, 1:82-96. doi: 10.1002/wcc.8.
- Seneviratne, S.I., Lüthi, D., Litschi, M., Schär, C. (2006): Land-atmosphere coupling and climate change in Europe. *Nature*, 443(7108): 205–209. doi:10.1038/nature05095.
- Shabalova, M.V., van Deursen, W.P.A., Buishand, T.A. (2003): Assessing future discharge of the river Rhine using regional climate model integrations and a hydrological model. *Climate Research*, 23(3): 233-246. doi: 10.3354/cr023233.
- Shoji, T., Kitaura, H. (2006): Statistical and geostatistical analysis of rainfall in central Japan. *Computer and Geosciences*, 32(8): 1007-1024. doi: 10.1016/j.cageo.2004.12.012.
- Schmidli, J., Frei, C., Vidale, P.L. (2006): Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology*, 26:679-689. doi: 10.1002/joc.1287.
- Schoof, J.T. (2008): Application of the multivariate spectral weather generator to the contiguous United States. *Agricultural and Forest Meteorology*, 148(3): 517-521. doi: 10.1016/j.agrformet.2007.10.003.
- Schopf, P., Gregg, M., Ferrari, R., Haidvogel, D., Hallberg, R., Large, W., Ledwell, J., Marshall, J., McWilliams, J., Schmitt, R., Skillingstad, E., Speer, K., Winters, K. (2002): Coupling Process and Model Studies of Ocean Mixing to Improve Climate Models - A Pilot Climate Process Modeling and Science Team. Special Report to the U.S. CLIVAR Scientific Steering Committee. Available from the U.S. CLIVAR Office, Washington, DC 20006, 35pp.
- Silverman, B.W. (1986): Density estimation for statistics and data analysis. Chapman and Hall, New York.
- SMIC, (eds). (1971): Inadvertent Climate Modification: Report of the Study of Man's Impact on Climate, Massachusetts Institute of Technology Press, Cambridge.
- Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (eds). (2007): *Climate Change 2007: The Physical Science Basis Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Staniforth, A. (1997): Regional modeling: a theoretical discussion. *Meteorology and Atmospheric Physics*, 63(1-2): 15-29. doi: 10.1007/BF01025361.
- Stocker, T.F., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (eds.) (2013): *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

- Štěpánek, P., Zahradníček, P., Huth, R. (2011): Interpolation techniques used for data quality control and calculation of technical series: an example of a Central European daily time series. *Idojaras*, 115(1-2): 87-98.
- Teng, J., Potter, N.J., Chiew, F.H.S., Zhang, L., Wang, B., Vaze, J., Evans, J.P. (2015): How does bias correction of regional climate model precipitation affect modelled runoff? *Hydrology and Earth System Sciences* 19(2): 711-728. doi:10.5194/hess-19-711-2015.
- Teutschbein, C., Seibert, J. (2012): Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, 456-457: 12-29. doi: 10.1016/j.jhydrol.2012.05.052.
- Teutschbein, C., Seibert, J. (2013): Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? *Hydrology and Earth System Sciences*, 17(12): 5061-5077. doi: 10.5194/hess-17-5061-2013.
- Themessl, M.J., Gobiet, A., Heinrich, G. (2012): Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change*, 112: 449-468. doi: 10.1007/s10584-011-0224-4.
- Themessl, M.J., Gobiet, A., Leuprecht, A. (2011): Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology*, 31: 1530-1544. doi: 10.1002/joc.2168.
- Trenberth, K.E., Dai, A., Rasmussen, R.M., Parsons, D.B. (2003): The changing character of precipitation. *Bulletin of the American Meteorological Society*, 84(9): 1205-1217. doi:10.1175/BAMS-84-9-1205.
- Uppala, S.M., Kallberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.F., Morcrette, J.J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., Woollen, J. (2005): The ERA-40 reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612): 2961–3012. doi:10.1256/ qj.04.176.
- Vlček, O., Huth, R. (2009): Is daily precipitation Gamma-distributed? Adverse effects of an incorrect use of the Kolmogorov-Smirnov test. *Atmospheric Research*, 93:759-766. doi: 10.1016/j.atmosres.2009.03.005.
- von Storch, H., Langenberg, H., Feser, F. (2000): A spectral nudging technique for dynamical downscaling purposes. *Monthly Weather Review*, 128(10): 3664-3673. doi: 10.1175/1520-0493(2000)128<3664:ASNTFD>2.0.CO;2
- Vuuren van, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S.J., Rose, S.K. (2011): The representative concentration pathways: an overview. *Climatic Change*, 109(1-2): 5-31. doi: 10.1007/s10584-011-0148-z.

- Wang, J.L., Kotamarthi, V.R. (2014): Downscaling with a nested regional climate model in near-surface fields over the contiguous United States. *Journal of Geophysical Research – Atmospheres*, 119(14): 8778-8797. doi:10.1002/2014JD021696.
- Wilby, R.L. (1997): Non-stationarity in daily precipitation series: implications for GCM downscaling using atmospheric circulation indices. *International Journal of Climatology*, 17: 439-454. doi:10.1002/(SICI)1097-0088(19970330)17:4<439::AID-JOC145>3.0.CO;2-U.
- Wilby, R.L., Wigley, T.M.L. (1997): Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography*, 21(4): 530-548. doi: 10.1177/030913339702100403.
- Wilby, R.L., Wigley, T.M.L. (2000): Precipitation predictors for downscaling: Observed and general circulation model relationships. *International Journal of Climatology*, 20(6): 641-661. doi: 10.1002/(SICI)1097-0088(200005)20:6<641::AID-JOC501>3.0.CO;2-1.
- Wilcke, R.A.I., Mendlik, T., Gobiet, A. (2013): Multi-variable error correction of regional climate models. *Climatic Change*, 120:871-887. doi: 10.1007/s10584-013-0845-x.
- Wilks, D.S. (2011): *Statistical methods in the atmospheric sciences*, 3rd edn. Academic, Amsterdam.
- Woolhiser, D.A., Roldan, J. (1982): Stochastic daily precipitation models 2. A Comparison of distributions of amounts. *Water Resources Research*, 18(5): 1461-1468. doi: 10.1029/WR018i005p01461.
- Zorita, E., von Storch, H. (1999): The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate*, 12(8): 2474-2489. doi: 10.1175/1520-0442(1999)012<2474:TAMAAS>2.0.CO;2.

Seznam obrázků

1.1 Schématická ilustrace výpočetní sítě modelu cirkulace atmosféry	10
1.2 Vývoj rozlišení globálních klimatických modelů.....	16
2.1 Mechanismus obousměrného doplňování.....	22
2.2 Prostorové rozlišení dosažené globálním a regionálním modelem.....	22
3.1 Efekt „mrhnutí“ – histogramy ze stanice České Budějovice.....	26
3.2 Schéma statistické předpovědi počasí	27
3.3 Rozdíly mezi perfect prognosis a model output statistics	27
3.4 Princip kvantilového mapování.....	35
3.5 Ilustrace jádrového odhadu hustoty	40
4.1 Mapa povodí Malše a použitých meteorologických stanic.....	44
4.2 Mapa vybraných povodí.....	45
5.1 Porovnání histogramů a exponenciálních hustot	48
5.2 Porovnání funkcí QM z histogramů a exponenciálních hustot	48
5.3 Porovnání různých parametrických odhadů hustoty	49
5.4 Porovnání funkce QM odvozených z různých parametrických hustot	50
5.5 Porovnání mocninných korekčních funkcí	52
6.1 Průběh autokorelačních a autokovariančních funkcí.....	60
6.2 Průběh různých vyjádření účelových funkcí.....	62
6.3 Precizní detekce nestacionarity a její vliv na QM.....	66
6.4 Detekce nestacionarity vyhovující kvantilovému mapování.....	67
7.1 Demonstrace souřadné soustavy založené na hlavních komponentách.....	69
7.2 Demonstrace korekce závislostní struktury	70
7.3 Demonstrace problému s orientacemi vlastních vektorů.....	71
7.4 Korekce syntetických dat – demonstrace rozdílu mezi PCC a QM.....	73
7.5 Korelační matice pro měřená, modelová a korigovaná data	74
7.6 Kovarianční matice pro měřená, modelová a korigovaná data	75
7.7 Korelační matice pro data z validačního období.....	78
7.8 Kovarianční matice pro data z validačního období	79
7.9 Střední absolutní chyby MAE korelačních koeficientů.....	80
7.10 Střední absolutní chyby jednotlivých kvantilů	80
7.11 Korekce vysokých kvantilů	81
7.12 Demonstrace efektu PCC: data z povodí 1 a 5	82

Seznam tabulek

1.1 Přehled vybraných globálních klimatických modelů.....	17
4.1 Přehled meteorologických stanic z povodí Malše.....	43
4.2 Přehled vybraných povodí.....	45
6.1 Přehled kritérií pro zhodnocení chyby korekčních metod	57
6.2 Chyby korekčních metod – korekce modelu REMO.....	58
6.3 Různé formulace účelové funkce	61
6.4 Chyby korekčních metod – korekce modelu ALADIN53.....	63
6.5 Hodnoty nezachycené nestacionarity	64

Seznam publikací

Články v časopisech s impakt faktorem:

Hnilica, J., Puš, V. (2013): Linear methods for the statistical transformation of daily precipitation sums from regional climate models. *Theoretical and Applied Climatology*, 111(1-2): 29-36. doi: 10.1007/s00704-012-0638-6.

Hnilica, J., Hanel, M., Puš, V. Multisite bias correction of precipitation data from regional climate models. *International Journal of Climatology*. *Accepted*.

Příspěvky ve sborníku:

Němečková, S., Hnilica, J. (2010): Verification of outputs from regional model REMO for impact studies in the basins in the southern Bohemia. In *Conference Abstracts*. Kyjev: Ukrainian Hydrometeorological Institute: 58-59. ISBN 978-966-521-567-7.

Hnilica, J., Šípek, (2011): V. Statistická transformace srážkových dat z regionálního klimatického modelu na specifické podmínky povodí Malše. In *Hydrologie malého povodí 2011*: 137-142. Ústav pro hydrodynamiku AVČR, 2011, Praha. ISBN 978-80-02-02290-9.

Hnilica, J. (2012): Statistical correction of daily precipitation data from the climate models. *Geophysical Research Abstracts* 14: 4780. ISSN 1607-7962.

Hnilica, J. (2013): An optimal separation of the year to periods with different number of precipitation days. *Geophysical Research Abstracts* 15: 4724. ISSN 1607-7962.

Hnilica, J., Puš, V. (2014): Kernel density estimates used in stochastic precipitation generator. In *Hydrologie malého povodí 2014*: 120-124. Ústav pro hydrodynamiku AVČR, 2014, Praha. ISBN 978-80-02-02525-2.