

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DISERTAČNÍ PRÁCE

2015

Lukáš Pop

Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

DISERTAČNÍ PRÁCE



Lukáš Pop

Analýza nárazů větru na území České republiky

Katedra fyziky atmosféry

Vedoucí disertační práce: doc. RNDr. Zbyněk Sokol, CSc.

Studijní program: Fyzika

Studijní obor: F 8 Meteorologie a klimatologie

Praha

rok

2015

Děkuji všem, kteří mi pomohli při zpracování této práce. Vedoucímu disertační práce doc. RNDr. Zbyňku Sokolovi, CSc. děkuji za vedení a odbornou pomoc. Konzultantům RNDr. Josefu Šteklovi, CSc. a RNDr. Jaroslavu Svobodovi, CSc. a členům Rady doktorského studijního oboru 4f8 děkuji za cenné připomínky ke struktuře práce. Kolegovi RNDr. Janu Kyselému, Ph. D. děkuji za odborné konzultace a kolegům Mgr. Davidu Hanslianovi, Ph. D. a Mgr. Jiřímu Hoškovi, Ph. D. děkuji za konzultace a poskytnutí podkladů ohledně větrné mapy. Dále děkuji Českému hydrometeorologickému ústavu za poskytnutá data.

Výsledky práce byly získány za podpory projektů GAAV KJB300420905 a GA ČR. GAP504/10/0843.

Prohlašuji, že jsem tuto disertační práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

V Praze dne 7. 5. 2015

podpis

Název práce: Analýza nárazů větru na území České republiky

Autor: Lukáš Pop

Katedra: Katedra fyziky atmosféry

Školící pracoviště: Ústav fyziky atmosféry AV ČR, v. v. i.

Vedoucí disertační práce: doc. RNDr. Zbyněk Sokol, CSc., Ústav fyziky atmosféry AV ČR, v. v. i.

Abstrakt:

Disertační práce se zabývá analýzou extrémních nárazů větru na území České republiky. První část práce shrnuje poznatky o metodách měření rychlosti a směru větru především s ohledem na vyhodnocení maximálních nárazů větru. Její součástí je i zpracování vysokofrekvenčního měření rychlosti větru 3-D sonickým anemometrem na stanici Kopisty a homogenizace řad nejvyšších denních nárazů větru na 29 českých stanicích. Pro řady nárazů větru naměřené na stanicích byly stanoveny základní statistické charakteristiky. Další část práce popisuje statistickou teorii extrémních hodnot a diskutuje její aplikovatelnost na extrémní hodnoty nárazů větru. Byly odvozeny některé teoretické poznatky a provedeny četné numerické experimenty s cílem určit vlastnosti používaných metod analýzy extrémů s ohledem na jejich využitelnost při analýze extrémních nárazů větru. V závěrečné části práce byly výsledky této analýzy metod využity při zpracování souboru staničních měření a při odvození výsledného modelu popisujícího závislost extrémního větrného klimatu na větrném klimatu. Aplikací tohoto modelu na větrnou mapu vytvořenou dříve na Ústavu fyziky atmosféry byla získána mapa extrémního větru na území ČR včetně odhadu její přesnosti. Tato mapa byla porovnána s mapami extrémního větru vytvořenými jinými autory.

Klíčová slova:

Nárazy větru, analýza extrémních hodnot, mapa extrémního větru, doba opakování

Title: Analysis of wind gusts over the area of the Czech Republic

Author: Lukáš Pop

Department: Department of Atmospheric Physics

Workplace: Institute of Atmospheric Physics ASCR, v. v. i.

Supervisor: doc. RNDr. Zbyněk Sokol, CSc., Institute of Atmospheric Physics CAS

Abstract:

The Ph.D. thesis deals with extreme wind gust analysis over the area of the Czech Republic. The first part of the thesis deals with processing of wind measurements, in particular maximum wind gusts measurements. Analysis of high-frequency wind measurement using 3-D sonic anemometer on the Kopisty station is included. Homogenization of the highest daily wind gusts was performed. Descriptive statistical analysis of measured wind gust values was performed. The following part of the thesis describes statistical theory of extreme values and discusses its applicability to wind gust data. Some theoretical findings were obtained. Numerous numerical experiments were performed focused on evaluation of proposed method. In the last part of the thesis station measurements were processed using the proposed methods and a model of

dependence between extreme and mean wind climate was derived. The model was applied to the map of mean wind climate calculated earlier on the Institute of Atmospheric Physics and thus a map of extreme wind climate was obtained. The accuracy of this map was estimated. The map was compared with other maps of extreme wind calculated by other authors earlier.

Keywords: Wind gusts, extreme value analysis, extreme wind map, return period

Seznam zkratek

AIM	Automatický imisní monitoring
bias	Systematická chyba
BM	Bloková maxima, metoda výběru extrémních dat
CAPE	Index dostupné konvektivní potenciální energie (Convective Available Potential Energy)
cdf	Distribuční funkce (Cumulative Distribution Function)
CLIDATA	Systém určený k uchování klimatologických dat v současnosti používaný na ČHMÚ
ČHMÚ	Český hydrometeorologický ústav
ČR	Česká republika
EVK	Extrémní větrné klima
GEV	Zobecněné rozdělení extrémních hodnot (General Extreme Value)
GPD	Zobecněné Pareto rozdělení (Generalized Pareto Distribution)
GF	Podíl mezi rychlostí nejvyššího nárazu a nejvyšší rychlostí s danou dobou průměrování (Gust Factor)
GWC	Zobecněné větrné klima (Generalized Wind Climate)
ML	Maximální věrohodnost, případně metoda maximální věrohodnosti (Maximum Likelihood)
pdf	hustota pravděpodobnosti (Probability Density Function)
POT	Peak-over-Threshold, tedy metoda uvažující vzájemně nezávislé naměřené hodnoty přesahující určitou hranici
RMSE	Střední kvadratická chyba (Root Mean Square Error)
STD	Směrodatná odchylka (STandard Deviation)
SYNOP	Běžně používaný formát meteorologických dat, označují se tak i zprávy a databáze s daty v tomto formátu
ÚFA	Ústav fyziky atmosféry
VK	Větrné klima
WAsP	Linearizovaný model proudění vyvinutý v Risø (Wind Atlas Analysis and Application Program)
WMO	Světová meteorologická organizace (World Meteorological Organization)

Obsah

1. Úvod.....	1
1.1 Cíl práce.....	3
1.2 Struktura práce	3
2 Měření větru v ČR.....	6
2.1 Měření větru a jeho nárazů na území ČR	6
2.1.1 Přístroje používané k měření větru, jejich chyby a vlastnosti	6
2.1.2 Přehled dostupných měření větru v ČR	8
2.1.3 Vliv různých přístrojů a algoritmů zpracování na naměřené maximální nárazy větru	9
2.1.4 Vliv různé doby průměrování a nespojitého měření na nejvyšší naměřené průměrné rychlosti větru	10
2.1.5 Vliv umístění měření větru na naměřené rychlosti větru.....	11
2.2 Vyhodnocení rozdílu mezi různými způsoby výpočtu nejvyšších denních nárazů větru na základě vysokofrekvenčního měření 3-D sonickým anemometrem ze stanice Kopisty.....	13
2.2.1 Popis experimentu	13
2.2.2 Zjištění a odstranění hrubých chyb	14
2.2.3 Dosažené výsledky	15
2.3 Homogenizace denních nárazů větru	16
2.3.1 Úvod.....	17
2.3.2 Problém homogenizace větru a volby referenčních dat.....	17
2.3.3 Popis a provedení použité metody	18
2.3.4 Korekční faktory na stanici Brno-Tuřany.....	22
2.3.5 Verifikace výsledků a volba parametrů Gaussových filtrů ve vzorci 2.3.1	23
2.4 Základní statistické vlastnosti nárazů větru na území ČR	25

2.4.1 Denní a roční chod větru a nárazů větru na území ČR v dřívějších pracích	25
2.4.2 Kvantily a maximální hodnoty nejvyšších denních nárazů větru	26
2.4.3 Roční chod nejvyšších denních rychlostí větru.....	28
2.4.4 Denní chod vysokých nárazů rychlosti větru.....	32
3 Analýza statistických metod.....	43
3.1 Definice a odhad doby návratu z empirické distribuční funkce	43
3.1.1 Definice doby návratu a souvislost s distribuční funkcí.....	43
3.1.2 Střední hodnota doby návratu naměřených hodnot.....	44
3.1.3 Nevýhody použití empirické distribuční funkce pro výpočet extrémních hodnot větru	45
3.2 Teorie extrémních hodnot.....	45
3.2.1 Úvod	45
3.2.2 První a druhý teorém extrémních hodnot	46
3.2.3 Problém domén atrakce	47
3.3 Použití GEV rozdělení k odhadu extrémních hodnot	47
3.3.1 Koeficienty a_n a b_n v prvním teorému extrémních hodnot	47
3.3.2 Vlastnosti umocněné distribuční funkce GEV rozdělení a jejich důsledky	48
3.3.3 Diskuze k prvnímu teorému extrémních hodnot	49
3.3.4 Srovnání teorie extrémních hodnot a centrální limitní věty	50
3.3.5 Odůvodnění a limity aproximace extrémních dat.....	51
3.3.6 Věta o průměrné době návratu určené proložení GEV rozdělení náhodnými hodnotami	52
3.4 Metody výběru extrémních hodnot z naměřených dat	53
3.4.1 Porovnání metody blokových maxim a metody Peak-over-Threshold ...	53
3.4.2 Metoda individuálních bouří.....	54

3.4.3 Separace extrémních rychlostí větru	54
3.5 Rozdíl mezi různými definicemi doby návratu.....	57
3.5.1 Úvod.....	57
3.5.2 Rozdíl mezi dvěma definicemi doby návratu.....	57
3.5.3 Aplikace na Gumbelovo rozdělení.....	60
3.6 Vliv tvaru růžice rychlostí větru na vztah mezi průměrnou a extrémní rychlostí větru.....	62
3.6.1 Způsob spojení průměrných a extrémních rychlostí větru z jednotlivých sektorů	63
3.6.2 Příklad spojení průměrných a extrémních rychlostí větru	63
3.7 Aproximace rozdělení naměřených dat pomocí GEV a Gumbelova rozdělení	67
3.7.1 Metoda maximální věrohodnosti.....	67
3.7.2 Metoda momentů.....	68
3.7.3 L-momenty	68
3.7.4 Odhad parametrů Gumbelova a GEV rozdělení metodou L-momentů	69
3.8 Analýza aproximace extrémních dat GEV a Gumbelovým rozdělením metodami L-momentů a ML na uměle vygenerovaných datech	70
3.8.1 Definice a výpočet L-ekvivalentního rozdělení.....	71
3.8.2 Změny charakteristik L-ekvivalentních GEV rozdělení s parametrem ξ .	72
3.8.3 Příprava umělých dat.....	76
3.8.4 Aproximovaná rozdělení a označení výsledků.....	77
3.8.5 Výsledky aproximací	77
3.8.6 Shrnutí výsledků aproximací	87
3.9 Analýza aproximace extrémních dat GEV a Gumbelovým rozdělením metodou L-momentů a ML na naměřených datech.....	88
3.9.1 Úvod.....	88

3.9.2 Data	89
3.9.3 Výsledky a jejich diskuse	90
3.9.4 Závěr	96
4 Výpočet mapy extrémního větru v ČR.....	98
4.1 Vztah mezi větrným klimatem a extrémním větrným klimatem na území ČR	98
4.1.1 Vztah podle normy IEC 61400-1.....	98
4.1.2 Použitá data a jejich příprava	99
4.1.3 Model pro závislost EVK na VK	100
4.1.4 Odhad chyb modelu	103
4.1.5 Kombinace výsledků regresního modelu s odhady ze staničních měření	104
4.1.6 Odhad u_{50} vypočtený z větrného klimatu	106
4.1.7 Odhad chyby u_{50}	108
4.1.8 Testování vhodnosti použití Gumbelova rozdělení.....	110
4.2 Mapa extrémního větru na území ČR	117
4.2.1 Použitá mapa VK	117
4.2.2 Výpočet mapy extrémního větru	119
4.2.3 Porovnání s výsledky jiných autorů	128
4.2.4 Odhad chyby vytvořené mapy u_{50}	133
5. Shrnutí výsledků a závěr	135
5.1 Hlavní výsledky práce.....	135
5.2 Výsledky dosažené v kapitolách 2 a 3	135
5.3 Výhled dalšího výzkumu.....	137
Seznam použité literatury.....	138
Přílohy	147
Příloha A: Problém domén atrakce.....	147

A.1 Nutná a postačující kritéria <i>konvergence</i>	147
A.2 Postačující kritéria konvergence (Von Misesův teorém)	148
Příloha B: Důkaz věty o průměrné době návratu určené proložením GEV rozdělení náhodnými hodnotami (sekce 3.3.5)	149
Příloha C: Vliv hodnot parametrů Gumbelova a GEV rozdělení na odhad hodnoty se zadanou dobou opakování.....	151
C.1 Gumbelovo rozdělení.....	151
C.2 GEV rozdělení	152
Příloha D: Odhad parametrů tvaru GEV rozdělení Bayesovským přístupem	153

1. Úvod

Předložená disertační práce se zabývá analýzou extrémních rychlostí větru na území ČR. Za extrémní rychlosti větru se považují hodnoty rychlosti větru, které jsou překročeny jen zřídka, což ve většině případů znamená jednou za několik desítek let na konkrétní lokalitě. V literatuře se rozlišuje extrémní rychlost větru a extrémní rychlost nárazu větru. Obě dvě veličiny lze ale považovat za průměr okamžité rychlosti větru za určitý časový interval. V případě rychlostí nárazů větru se podle WMO většinou uvažuje délka časového intervalu 3 s, zatímco v případě rychlostí větru 10 nebo 15 minut. Extrémní rychlosti nárazů větru lze tedy považovat za extrémní rychlosti větru s velmi krátkou dobou průměrování, a proto v této práci pod pojem extrémní rychlost větru zahrnujeme i extrémní rychlosti nárazů větru. Vzhledem k tomu, že data o rychlosti nárazů větru jsou výrazně dostupnější především pro starší období, jsou k výpočtům s reálnými daty v této práci použity hodnoty rychlosti nárazu větru. Avšak většina teoretických poznatků je platná i pro extrémní rychlost větru se standardní dobou průměrování, případně i pro extrémní hodnoty jiných veličin.

V textu práce se v zájmu stručnosti používají následující termíny, které nejsou běžné nebo jejichž užití v odborné literatuře není jednotné. Pod termínem „měření větru“ se v celé této práci rozumí měření jeho rychlosti a směru. Termín „větrné klima“ (VK) vyjadřuje celkové statistické charakteristiky rychlosti a směru větru, dané např. jeho průměrnou rychlostí nebo podrobněji četností a hustotou pravděpodobnosti rychlosti větru v jednotlivých sektorech směru větru. Obdobně termín „extrémní větrné klima“ (EVK) znamená obdobné charakteristiky extrémního větru. Termínem „větrná mapa“ se rozumí mapa větrného klimatu a termínem „extrémní větrná mapa“ mapa extrémního větrného klimatu. Pod termínem „maximální náraz“ se rozumí náraz dosahující maximální rychlosti. Pod termínem „korelační koeficient“ se rozumí Pearsonův korelační koeficient a pod termínem „korelace“ lineární korelace. Pod termínem „aproximace rozdělení“ se rozumí aproximace jeho distribuční funkce. Pod termínem „exponovanost lokality“ se rozumí zesílení proudění vzduchu v důsledku místních podmínek. Formulací „data jsou dána rozdělením X“ se rozumí, že data jsou vzájemně nezávislá náhodná čísla pocházející z pravděpodobnostního rozdělení X.

Extrémní rychlosti větru způsobují velké hospodářské škody, jak jsme se mohli přesvědčit např. při ničivé bouři Kyrill v lednu 2007. Tyto škody lze omezit preventivními opatřeními, jako je použití odolnějších konstrukcí a přizpůsobení využívání krajiny, což představuje nezanedbatelné náklady. Pro návrh ideálního provedení těchto preventivních opatření je nutné znát EVK v dané lokalitě. To představuje komplikovaný problém, přičemž jedním z hlavních důvodů je velký počet možných zdrojů chyb, se kterými je spojeno určování EVK. Tyto zdroje jsou:

1. Při statistické analýze je vždy k dispozici pouze konečný náhodný výběr hodnot, tzn. délka měření rychlosti a směru větru je omezená.
2. Omezený počet stupňů volnosti (typ a parametry distribuční funkce) rozdělení použitého k aproximaci naměřených hodnot extrémní rychlosti větru.
3. Chyby a nehomogenity měření.
4. Změny klimatu.
5. Nepřítomnost in situ měření. Dlouhodobá měření rychlosti a směru větru se provádí na omezeném počtu míst. Vzhledem k tomu, že rychlost a směr větru vykazují velkou prostorovou proměnlivost, může interpolace naměřených dat vést ke značným chybám.
6. Neurčitost daná různými definicemi doby návratu.

V této práci je ukázán postup eliminace neurčitosti dané různými definicemi doby návratu. Ostatní chyby zcela eliminovat nelze, ale lze je omezit volbou vhodné metody pro odhad EVK. Při volbě vhodné metody je potřeba postupovat obezřetně, protože se může stát, že snížením chyby jednoho typu dojde naopak k navýšení chyby jiné. Příkladem je chyba v důsledku omezeného výběru a chyba v důsledku omezeného počtu stupňů volnosti u rozdělení extrémních hodnot. Chybu výběru lze eliminovat tím, že výsledné rozdělení extrémních rychlostí větru nebude záviset na naměřených hodnotách a bude mít tedy nulový počet stupňů volnosti. Například můžeme prohlásit, že rozdělení ročních maxim nárazu větru je dáno Gumbelovým rozdělením s parametrem polohy 25 a parametrem tvaru 3 nezávisle na lokalitě. Chyba výběru je v tomto případě nulová, zatímco chyba daná omezením stupňů volnosti rozdělení extrémních hodnot bude v tomto případě vysoká. Naopak chybu omezeného počtu stupňů volnosti rozdělení lze eliminovat aproximací rozdělení použitím empirické distribuční funkce naměřených extrémních hodnot. Tato metoda umožňuje libovolný tvar rozdělení extrémních hodnot, její počet stupňů volnosti je tedy neomezený,

ale naopak chyba výběru je u ní značně vysoká. Při výběru konkrétní metody určení rozdělení extrémních hodnot rychlostí větru je tedy nutno vždy posoudit více možných zdrojů chyb, jejichž odhad je často zatížen značnou dávkou subjektivity.

Typická doba návratu extrémních rychlostí větru používaná v praktických aplikacích je 50 let. Pro takovou délku návratu není možné přímo ověřit správnost odhadu extrémní rychlosti větru, protože nejsou k dispozici potřebná data. Nejdelší kvalitní měření větru z území ČR jsou dlouhá okolo 50 let a připadá na ně tedy v průměru přibližně jedna událost překračující hodnotu 50leté rychlosti větru, což je příliš málo pro provedení statistických testů. Tento problém lze částečně obejít metodami tzv. regionální analýzy často využívané v případě srážek, kdy se k ověření a odhadům parametrů rozdělení extrémních hodnot používají současně data z více stanic, čímž se celkový počet dat podstatně zvýší. Tyto metody ale zavádějí dodatečné předpoklady o nezávislosti dat na jednotlivých stanicích a společných rysech rozdělení extrémních rychlostí větru na všech těchto stanicích, a to je zřejmě důvodem, že jejich použití není obvyklé v případě analýzy extrémních rychlostí větru.

1.1 Cíl práce

Cílem práce je vytvoření mapy EVK na území ČR a odhad její chyby. K dosažení tohoto cíle je provedena analýza měření nárazů větru a analýza metod jejich statistického zpracování.

K tomuto účelu je nutné zkoumat různé zdroje chyb, viz výše, spojené jak s měřením nárazů větru, tak se statistickou analýzou extrémních hodnot. Tyto chyby závisí na zvolené metodě. Kapitola 2 se proto zabývá analýzou měření nárazů větru, a kapitola 3 analýzou jejich statistického zpracování různými metodami. Na základě analýzy provedené v těchto dvou kapitolách je v kapitole 4 navržena a použita konkrétní metoda vytvoření mapy extrémních nárazů větru na území ČR.

1.2 Struktura práce

Práce se skládá z 5 kapitol. Kromě toho obsahuje český a anglický abstrakt, seznam zkratek a seznam literatury.

Kapitolu 1 tvoří úvod, ve kterém jsou definovány cíle práce. Odborná část práce je obsahem kapitol 2, 3 a 4, závěr práce je uveden jako kapitola 5.

Kapitola 2 se zabývá měřením rychlosti a směru větru a skládá se ze čtyř podkapitol. Podkapitola 2.1 představuje rešerši poznatků o měření větru, které souvisejí s odhadem EVK na území ČR. Podkapitola 2.2 ukazuje a kvantifikuje na konkrétních datech nejistoty a neurčitosti spojené s měřením maximálních nárazů větru. V podkapitole 2.3 je popsána homogenizace řad nejvyšších denních nárazů větru na 29 českých stanicích. Podkapitola 2.4 zkoumá základní statistické charakteristiky rychlostí nárazů větru naměřených na stanicích v ČR.

Třetí kapitola se zabývá statistickou analýzou extrémních hodnot, především rychlosti větru. Podkapitola 3.1 uvádí definici doby návratu a zabývá se jejím výpočtem pomocí empirické distribuční funkce. V podkapitole 3.2 jsou popsány základy teorie extrémních hodnot. V podkapitole 3.3 je diskutováno použití GEV rozdělení pro aproximaci naměřených hodnot. V podkapitole 3.4 jsou popsány metody výběru extrémních hodnot z naměřených dat. V podkapitole 3.5 je kvantifikován rozdíl v odlišných definicích doby návratu a jeho korekce. V podkapitole 3.6 je ukázán a diskutován odlišný vliv výpočtu celkového EVK z EVK pro jednotlivé sektory směru větru oproti výpočtu celkového VK z VK pro jednotlivé sektory. Podkapitola 3.7 popisuje metody aproximace rozdělení extrémních hodnot z naměřených dat, především nejvíce používané metody L-momentů a maximální věrohodnosti. Podkapitola 3.8 obsahuje rozsáhlou analýzu výsledků těchto metod v případě aproximace umělých dat vygenerovaných metodou Monte Carlo Gumbelovým a GEV rozdělením a podkapitola 3.9 obdobnou analýzu v případě dat vygenerovaných metodou bootstrap z homogenizovaných maximálních ročních nárazů větru na profesionálních stanicích ČHMÚ.

Čtvrtá kapitola využívá poznatky získané v přechozích dvou kapitolách. V podkapitole 4.1 je na základě analýzy v kapitole 3 zvolena a aplikována vhodná metodika určení EVK z naměřených maximálních denních nárazů na profesionálních stanicích ČHMÚ. Na těchto stanicích je spočítáno i VK pomocí patnáctiminutových průměrných rychlostí větru z databáze CLIDATA. Je odvozen model pro závislost EVK na VK, který je verifikován a porovnán se vztahem z normy IEC 61400-1. Součástí této kapitoly je i diskuze správné váhy chvostu použitého rozdělení ročních maxim rychlosti větru. V podkapitole 4.2 je vytvořena mapa extrémního větru aplikací vztahu mezi VK a EVK odvozeného v podkapitole 4.1. Tato

mapa je srovnána s mapami vytvořenými jinými autory a je diskutována a odhadnuta její chyba.

Kapitola 5 shrnuje a hodnotí výsledky dosažené v této práci a diskutuje možnosti navazujícího výzkumu.

2 Měření větru v ČR

2.1 Měření větru a jeho nárazů na území ČR

Tato kapitola stručně shrnuje otázky spojené s měřením větru a jeho nárazů na území ČR v období, ze kterého pocházejí naměřená data použitá v této práci. Tyto znalosti jsou důležité pro zpracování naměřených dat, protože měření větru je zatíženo řadou komplikací. Požadavky na měření větru jsou dány Světovou meteorologickou organizací (WMO) (WMO, 2010).

Dopad některých níže popsaných vlivů na naměřené hodnoty byl zkoumán s využitím vysokofrekvenčního měření 3D sonickým anemometrem na stanici Kopisty. Získané poznatky jsou uvedeny v podkapitole 2.2.

2.1.1 Přístroje používané k měření větru, jejich chyby a vlastnosti

Přístroje měřící vítr, respektive rychlost větru, se obecně nazývají anemometry. Zde se budeme zabývat jen typy relevantními pro naměřená data využitá v této práci, přestože různých typů existuje velké množství (ČMES, 2015).

Nejpoužívanějším typem přístroje pro měření rychlosti větru je miskový anemometr, který bývá většinou doplněn větrnou směrovkou pro měření směru větru (Řepka, 2011). Měření spočívá v určení rychlosti otáčení misek, přičemž se předpokládá, že rychlost větru je lineární funkcí této rychlosti otáčení (Kristersen, 1993). Parametry této funkce se nastavují kalibrací ve větrném tunelu.

Část chyb těchto přístrojů je způsobená odlišnými podmínkami proudění na stanici ve srovnání s větrným tunelem. Proudění ve větrném tunelu je stabilní a horizontální. Protože miskový anemometr reaguje pomaleji na poklesy než na vzestupy rychlosti větru, nastává vlivem fluktuací rychlosti větru nadhodnocení naměřené rychlosti větru nazývané overspeeding (Kristensen, 1998; Bush et Kristensen, 1976). Tento efekt roste se setrvačností přístroje. Vertikální složky rychlosti větru způsobují další chybu naměřené rychlosti, přičemž různé anemometry se chovají odlišně (Bégin-Drolet et al., 2011; Papadopoulos et al., 2001).

Protože je miskový anemometr mechanický přístroj, dochází postupem času k jeho mechanickému opotřebení, především ložisek. Proto se doporučuje provádět pravidelně jeho kalibraci a případnou výměnu. Stárnutím miskových anemometrů se podrobně zabýval článek Pindado et al. (2012), podle něhož dochází ke změnám vlastností

přístrojů dokonce i když nejsou používány. Stárnutí ložisek se ale více projevuje například zvýšením podílu bezvětří v důsledku zvýšení prahové rychlosti přístroje než při extrémních rychlostech.

Dalším typem anemometrů jsou anemometry akustické. Ty se dříve používaly především pro výzkum turbulence v mezní vrstvě, k čemuž jsou vhodnější než miskové anemometry, u kterých vadí jejich setrvačnost (Wieser et al., 2001). Pro operativní měření rychlosti větru začaly být běžně používány až v posledních letech (Řepka, 2011), protože dříve byly velmi drahé. Dalšími výhodami akustických anemometrů proti miskovým jsou kromě zmíněné absence setrvačnosti a s tím spojeného možného vysokého časové rozlišení také možnost snímání všech tří složek vektoru rychlosti větru, absence prahové rychlosti přístroje a absence mechanického opotřebení, protože tento typ anemometru neobsahuje žádné pohyblivé součásti. Nevýhodou je ovlivnění proudění vzduchu konstrukcí přístroje (Dellwik et al., 2010). Toto ovlivnění je ve výstupních datech odstraněno proprietárním algoritmem. Tento algoritmus se nastavuje opět na základě měření ve větrném tunelu, a proto může být měření na stanici zatíženo chybou způsobenou odlišnou mírou turbulence na stanici ve srovnání s větrným tunelem a vertikálními složkami proudění obdobně jako v případě miskových anemometrů (Wieser et al., 2001). Dále tyto anemometry poskytují chybné hodnoty, pokud se na vysílačích usadí kapičky vody (Pahlow et Parlange, 2001). Podle zkušeností s měřeními sonickými anemometry různých výrobců na Ústavu fyziky atmosféry jsou tato období velmi zřetelná a nezbyvá, než je vymazat. Chování jiných typů sonických anemometrů však může být jiné.

Dalším typem anemometru jsou anemometry založené na principu tzv. Pitotovy trubice. Ta měří rozdíl statického a dynamického tlaku, jež závisí na rychlosti větru. Její fyzikální princip je popsán v Kristensen (1993). Pitotova trubice byla součástí tzv. Univerzálního anemografu (Kozumplík, 1950), kde sloužila k měření okamžité rychlosti větru a jeho nárazů vzhledem ke své nízké setrvačnosti. Pro správnou funkci přístroje je nutné zajistit neustálé směřování ústí trubice proti směru větru.

Měřicí přístroje a především miskové anemometry mohou být mechanicky poškozeny například nárazem ptáka, při vichřici apod. Pokud je ovlivnění velké a zřejmé, je nutno data z dalšího zpracování vyloučit. Menší ovlivnění se obvykle nezjistí a jsou opravena až při příští kalibraci.

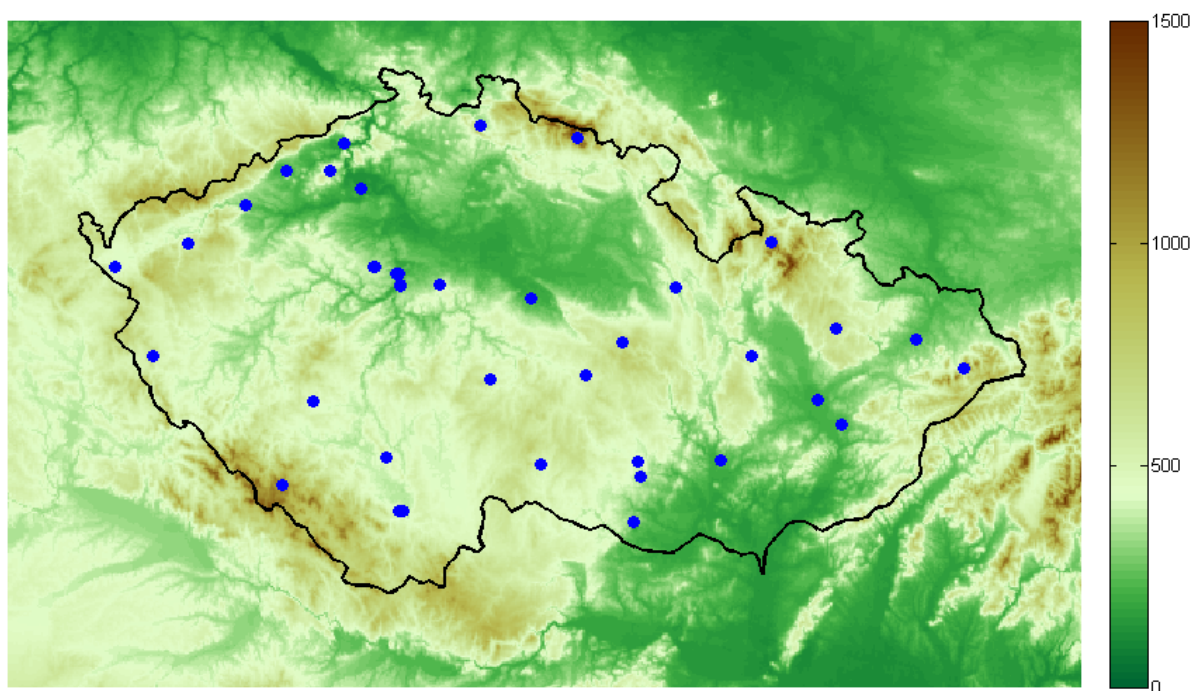
U starších dat se občas vyskytují chyby při zápisu hodnot. Někdy jsou tyto chyby patrné na první pohled, kdy je uváděná hodnota nesmyslně vysoká. Tyto hodnoty je obvykle nutno vyřadit. V některých případech je ale zřejmé, že chyba souvisí s manuálním zapisováním, kdy došlo k posunu desetinné čárky nebo k přehození číslic. V takových případech lze chybu opravit.

2.1.2 Přehled dostupných měření větru v ČR

Přehled dostupných měření větru v ČR je popsán v publikaci Řepky (2011), ze které zde vycházíme. V této publikaci se stanice měřící vítr dělí na následující skupiny:

1. profesionální stanice,
2. dobrovolnické automatizované stanice,
3. stanice s anemografem,
4. stanice AIM,
5. manuální klimatologické stanice.

Za jednoznačně nejkvalitnější jsou považována měření z profesionálních stanic, proto je využíváme v této práci. Mapa současného rozložení profesionálních stanic je na obr. 2.1.1.



Obr 2.1.1: Mapa profesionálních meteorologických stanic v ČR spolu s orografií. Nadmořská výška je v m.

V této práci zpracováváme data z období 1961-2009. K měření větru na těchto stanicích se v tomto období přibližně do přelomu století používal tzv. Univerzální anemograf, který se skládal z miskového anemometru pro určení průměrné rychlosti větru, větrné směrovky a Pitotovy trubice pro záznam okamžité rychlosti větru. Nejvyšší denní nárazy se odečítaly z tohoto záznamu. Univerzální anemografy byly poté v rámci tzv. automatizace nahrazeny měřením průměrné rychlosti i rychlosti nárazů větru miskovými anemometry Vaisala WAV 151, případně WAV 251, který je mechanicky stejný, ale má výkonnější vytápění. Údaje o modernizaci stanic v ČR jsou popsány Řepkou (2011) a technické specifikace přístrojů jsou dostupné na <http://www.vaisala.com>. V poslední době se na mnoha profesionálních stanicích přešlo na měření sonickými anemometry WAS 425 AH (technickou dokumentaci lze nalézt na <http://www.vaisala.com>). Hodnota nejvyššího nárazu se v případě těchto anemometrů získá softwarovým zpracováním naměřených hodnot, a je tedy závislá na nastavení tohoto algoritmu.

V této práci využíváme dva typy dat.

1. Data o průměrné rychlosti a směru větru v 15minutových intervalech z miskových anemometrů Vaisala, která jsou dostupná přibližně od roku 2004 (Řepka, 2011).
2. Data o času, rychlosti a směru nejvyššího denního nárazu větru, která jsou dostupná od roku 1961. Tato data byla nejprve odečítána ze záznamu vytvořeného Pitotovou trubicí v Univerzálním anemografu, a poté z miskových anemometrů Vaisala (Řepka, 2011).

2.1.3 Vliv různých přístrojů a algoritmů zpracování na naměřené maximální nárazy větru

Nejvyšší náraz větru by měl podle WMO odpovídat plovoucímu intervalu průměrování 3 s (WMO, 2010). Meteorologický slovník (ČMES, 2015) udává pro meteorologické potřeby dobu trvání nárazu větru jako 1 až 20 s. Ve skutečnosti ale zpravidla naměřený maximální náraz závisí na setrvačnosti přístroje a odpovídající interval průměrování bývá nižší než 3 s požadované WMO (Holmes et al., 2014). To platí

zejména v případě měření pomocí Pitotovy trubice. Podle Miller et al. (2013) vedou měření přístroji s Pitotovou trubicí k o 6 až 19 % vyšším naměřeným nárazům než odpovídá normě WMO, což dobře odpovídá výsledkům podkapitoly 2.3. V Miller et al. (2013) a Ginger et al. (2011) je navíc ukázáno, že odezva Pitotovy trubice na změny rychlosti větru vykazuje rezonanční frekvenci v řádu jednotek Hz. Další nehomogenita může nastat i při výměně zmíněných miskových anemometrů Vaisala za sonické anemometry WAS 425 AH.

Výměny přístrojů na stanicích jsou finančně nákladné, a proto nejsou nové přístroje na stanice nasazovány plošně. Vyhodnocení vlivu změn přístrojů dále komplikuje fakt, že termíny výměn přístrojů jsou často neznámé nebo obtížně dohledatelné.

Významnou roli při výpočtu maximálních nárazů větru hraje způsob skalárního nebo vektorového průměrování rychlosti větru. V případě 3D sonických anemometrů existují i různé algoritmy zahrnutí vertikální složky rychlosti větru. Je zde třeba upozornit, že výstupem měření pro uživatele nejsou přímo naměřená surová data, ale data zpracovaná softwarem, jehož algoritmy zpravidla nejsou explicitně známy.

Podhodnocení naměřených nejvyšších nárazů větru nastane při výpočtu nejvyššího nárazu jako maxima z pevně daných po sobě následujících intervalů namísto plovoucích, které jsou požadované WMO.

2.1.4 Vliv různé doby průměrování a nespojitého měření na nejvyšší naměřené průměrné rychlosti větru

Alternativou k vyhodnocení nejvyšších nárazů větru je vyhodnocení nejvyšších průměrných rychlostí za podstatně delší dobu (obvykle 10 minut až 1 hodina). U vyhodnocení těchto dat ale mohou nastat problémy při změnách použité délky průměrování, případně při porovnání extrémních hodnot mezi stanicemi používajícími různé doby průměrování. Problematická je též situace, kdy data nebyla zaznamenávána spojitě, což je případ měření SYNOP, kde se zaznamenává desetiminutový průměr rychlosti větru jednou za hodinu. Dříve se zaznamenával dokonce jednou za tři hodiny. Tyto faktory nezpůsobí systematickou chybu v odhadu průměrných rychlostí větru za dlouhé období, ale jsou důležité například při určení maximálních ročních hodnot.

Korekce nejvyšších naměřených rychlostí na různé doby průměrování je řešena v Harper et al. (2008), kde se vychází ze spektrální analýzy rychlosti větru. Tato práce se

sice zabývá tropickými cyklony, ale navržený přístup je aplikovatelný i na extrémní rychlosti větru v mírných zeměpisných šířkách. V této práci lze najít i odhad poměru mezi nejvyšší desetiminutovou rychlostí a nejvyšším nárazem (doba průměrování 3 s), který roste se složitější orografií v okolí stanice. Podle našich výsledků tento podíl také klesá s rostoucí výškou nad zemským povrchem (Pop, 2007). Obě tyto závislosti jsou logické, protože rozdíl mezi nejvyšší rychlostí s různou dobou opakování je způsoben turbulencí a ta je mimo termických příčin generována interakcí proudění se zemským povrchem.

Jiný přístup založený na aproximaci časových řad rychlosti větru Markovovými řetězci prvního řádu je popsán v Larsén et Mann (2006). Tento přístup umožňuje korekci jak na změnu délky průměrování, tak na nespojitě měření. Přístupy v obou pracích ale vyžadují nastavení specifických parametrů pro každou stanici, což je nutno provést buď z dostupného dostatečně dlouhého spojitého měření o vysoké vzorkovací frekvenci, nebo odhadem na základě posouzení charakteru okolí stanice.

Podíl mezi rychlostí nejvyššího nárazu a nejvyšší rychlostí s danou dobou průměrování se nazývá anglicky gust factor (GF). Korekce nejvyšších rychlostí pro různé doby opakování vychází z podílu GF pro tyto doby. Jde o analogický problém výše popsanému jen s jiným názvoslovím. Statistické zpracování GF pro různé konkrétní případy extrémního větru lze pro Singapur nalézt v publikaci Choi et Hidayat (2002) a pro měření v jihovýchodní Anglii v práci Deaves (1993). Vztahem mezi desetiminutovými nejvyššími rychlostmi a nejvyššími nárazy pro území České republiky se zabýval Král (2007), kterému vychází vyšší poměry ve srovnání s jinými pracemi. Tento poznatek lze vysvětlit dvěma faktory: (i) Králova studie vychází z měření z let 1981 - 1999, která se prováděla téměř výhradně Univerzálními anemometry, které mají malou setrvačnost a naměřené nárazy jsou tak vyšší než nárazy naměřené novějšími přístroji a než odpovídá požadavkům WMO, (ii) desetiminutové průměry jsou z databáze SYNOP, takže obsahují vždy jen posledních 10 minut z každé hodiny, a nebyla provedena příslušná korekce (Larsén et Mann, 2006).

2.1.5 Vliv umístění měření větru na naměřené rychlosti větru

Další skutečností, kterou je nutné při vyhodnocování měření větru uvažovat, je jeho umístění. To je v případě větru důležitější než u jiných meteorologických veličin vzhledem k jeho velké proměnlivosti na malých prostorových škálách a jeho silnému

ovlivnění i poměrně malými změnami v okolí stanice. I zdánlivě malá změna v okolí může způsobit nehomogenitu naměřených řad. Vliv umístění měření větru lze zkoumat na více úrovních.

Za prvé jde o vliv samotné meteorologické stanice. Standardní meteorologická měření větru se provádějí ve výšce 10 m nad zemí (WMO, 2010). Tato měření jsou často umístěna na střeše budovy, nad kterou dochází ke zvýšení rychlosti větru v důsledku obtékání této budovy. Zjednodušený výpočet tohoto zvýšení rychlosti modelem WAsP (jeho popis je dostupný na <http://www.wasp.dk>) je navržen v publikaci Landberg (2000). Jako příklad je zde uveden výpočet pro stanici Qantara v Egyptě, která má horizontální rozměry 14,3 x 13,4 m a výšku 4,0 m, a je na ní umístěn 9 m vysoký měřicí stožár. V tomto případě model dává zvýšení rychlostí větru v bodě měření o znatelných 9 %. Použití stejné metodiky na typizovanou profesionální meteorologickou stanici v České republice vede ke zvýšení průměrné rychlosti větru měřené anemometry umístěnými nad těmito stanicemi zhruba o 15 až 20 % (Hanslian, 2014).

Za druhé jde o vliv překážek v okolí místa měření. Podle WMO (2010) se jako absolutní minimum vyžaduje, aby vzdálenost mezi překážkou a měřením byla nejméně desetinásobek výšky této překážky, přičemž ideální je podstatně vyšší vzdálenost. Konkrétní vliv překážek na proudění byl studován s využitím matematických modelů. Zjednodušený model pro odhad vlivu 2D (s nekonečnou délkou) překážky na proudění vzduchu je využit v modelu WAsP a byl navržen v Perera (1981). V Taylor et Salmon (1993) jsou diskutovány 4 různé modely zahrnující i 3D překážky. Podle nich je vliv 3D překážek znatelně menší než vliv odpovídajících nekonečných 2D překážek.

Za třetí jde o vliv parametru drsnosti povrchu v okolí místa měření. WMO doporučuje parametr drsnosti povrchu v okolí stanice rovný 0,03 m (WMO, 2010). Používá přitom Davenportovu klasifikaci drsností povrchů převzatou z Wieringa (1980) a podle ní tato hodnota odpovídá trávníku. Ale typická hodnota drsnosti povrchu v ČR je odhadem spíše 0,2 m. Časté jsou i vyšší hodnoty (zmíněná klasifikace přiřazuje lesům parametr drsnosti 1 m a centřum měst vyšší než 2 m). Pro účely modelování jsou naopak žádoucí měření větru co nejméně ovlivněná zemským povrchem, tedy umístěná výše nad zemí a v místech s co nejmenším parametrem drsnosti. WMO (WMO, 2010) doporučuje zvýšit výšku měření v případě vyšší okolní drsnosti povrchu (na 13 m v případě drsnosti 0,1 m a 19 m v případě drsnosti 0,5 m), což je ale problematické

v případě, kdy typická drsnost závisí na směru od místa měření. Určení parametru drsnosti podle výše zmíněné klasifikace je také značně subjektivní. Kvantifikace vlivu lokální topografie je popsána v Walmsley et al. (1990) a je také součástí modelu WASP.

Za čtvrté jde o vliv orografie v okolí stanice. Při interpolaci naměřených dat je vhodné také zohlednit fakt, že ani orografie v okolí stanice nemusí být typická. Například vrchol Milešovky určitě nepředstavuje typickou lokalitu v ČR o nadmořské výšce 837 m n. m. Jiným příkladem je měření větru v horském údolí (např. Pec pod Sněžkou). Rychlost a směr větru v údolí se zásadně liší od rychlosti a směru větru v otevřené krajině v blízké lokalitě se stejnou nadmořskou výškou a naopak.

S umístěním měření úzce souvisí pojem jeho reprezentativnosti. Podle meteorologického slovníku (ČMES, 2015) je „reprezentativní meteorologické pozorování, při němž jsou dodržovány předepsané postupy, např. výška sensoru nad zemí, a jehož výsledky mají platnost pro širší okolí místa pozorování. Velikost tohoto okolí závisí na prostorové proměnlivosti daného met. prvku, na charakteru terénu a na účelu pozorování.“ Je zřejmé, že v případě větru bude reprezentativnost měření důležitější než například u srážek z důvodů jeho velké prostorové proměnlivosti (Wieringa, 1996).

Z výše uvedeného vyplývá, že při zpracování měření větru je vhodné mít k dispozici fotografickou dokumentaci lokálních podmínek, jak je doporučeno v WMO (2010) a v článku Wieringa (1983), a provést kvantifikaci nežádoucích okolních vlivů buď kvalifikovaným odhadem nebo pomocí numerického modelu proudění. Numerický model proudění je také vhodný pro extrapolaci výsledků ze stanice do jiné lokality.

2.2 Vyhodnocení rozdílu mezi různými způsoby výpočtu nejvyšších denních nárazů větru na základě vysokofrekvenčního měření 3-D sonickým anemometrem ze stanice Kopisty

2.2.1 Popis experimentu

V této kapitole budou kvantifikovány vlivy různých neurčitostí definice nejvyššího nárazu zmíněných v podkapitole 2.1. K tomuto účelu byla použita data o rychlosti a směru větru z přístroje Gill WindMaster 3D Anemometer (parametry, technickou

dokumentaci a návod k použití lze nalézt na webu <http://www.gillinstruments.com>) umístěného nad střechou budovy meteorologické stanice v Kopistech.

Použili jsme naměřená data z období 24. 8. 2012 až 1. 4. 2014, která se sestávala ze složek rychlostí větru ve všech třech dimenzích s frekvencí 20 Hz, tj. s intervalem měření 0,05 s. Na těchto datech byly aplikovány různé metody pro výpočet nejvyšších denních nárazů větru. Z dat jsme vyřadili všechny dny, kdy nastal jakýkoli výpadek měření. Celkem jsme zpracovávali data ze 493 dní bez výpadků měření.

Jako základní metodu pro výpočet nejvyšších denních nárazů větru jsme použili metodu založenou na využití pouze horizontální složky rychlosti větru. Denní náraz se počítal jako maximum průměrných rychlostí z 3 s dlouhých plovoucích intervalů, což odpovídá požadavkům WMO. V celé této podkapitole jsme používali průměrování rychlosti větru uvažované jako skalár. Další testované metody vznikly následujícími modifikacemi základní metody:

- (i) výpočet maximální rychlosti se provedl z 3D hodnot, tj. včetně vertikální složky,
- (ii) při výpočtu průměrů se vycházelo z pevně daných po sobě následujících 3 s intervalů namísto plovoucích,
- (iii) průměry se počítaly z časových intervalů o jiné délce než 3 s.

2.2.2 Zjištění a odstranění hrubých chyb

V datech se vyskytla velmi vysoká hodnota nejvyššího nárazu větru (vypočítaná základní metodou) pro den 22. 2. 2013 (cca 56 m/s). Po prohlédnutí původních 20 Hz dat bylo zcela zřejmé, že se jedná o důsledek chyby, kdy v jednom termínu měření byla v jednom ze směrů zapsána hodnota rychlosti větru 3000 m/s. Jiné hodnoty extrémních rychlostí větru již nápadné nebyly, avšak při kontrole jednotlivých termínových měření s frekvencí 20 Hz se našly další termíny, ve kterých byla hodnota 3D rychlosti vyšší než 40 m/s. Bylo jich celkem 5, a po prohlédnutí časových řad v těchto termínech bylo zřejmé, že se ve všech těchto případech opět jedná o hrubé chyby v jednom termínu v jedné ze složek rychlosti větru.

Při frekvenci 20 Hz znamená chyba 45 m/s v jednom termínu v původních datech, která je na první pohled zřejmá, poměrně malou chybu 0,75 m/s v nárazech (třísekundových průměrech). Zpracování dat ukázalo, že tyto chyby se objevují jednotlivě, a proto je lze odstranit a nahradit interpolovanou hodnotou ze sousedních

termínů. Testy prokázaly, že chyba interpolace originálních dat je v naprosté většině případů pod 1 m/s a výsledná chyba v 3 s nárazech je zanedbatelná.

Pokud nastane hrubá chyba o velikosti 5 m/s v řadě patnáctiminutových, hodinových nebo dokonce denních nárazů, pak není zpravidla v datové řadě zřejmá, a tedy je obtížné ji odstranit. Odhalení a odstranění této chyby je však snadné pokud jsou k dispozici originální 20 Hz data ze sonického anemometru, protože takto velké chybě odpovídá chyba rychlosti 300 m/s v originálních datech.

Proto je velmi důležité uchovávat původní vysokofrekvenční data, která jsou dostupná u moderních sonických přístrojů. Současná kapacita výpočetní techniky umožňuje jejich uchovávání a zpracovávání, a lze očekávat, že se její možnosti budou dále zvyšovat. Tyto hrubé chyby nejsou významné v případě analýzy VK, ale mohou mít zásadní význam při analýze EVK.

V naší 20 Hz řadě jsme zjištěné chybné hodnoty nahradili lineární interpolací sousedních hodnot. Výpočty denních nárazů základní metodou a jejími třemi modifikacemi jsme pak provedli znovu s těmito upravenými hodnotami.

2.2.3 Dosážené výsledky

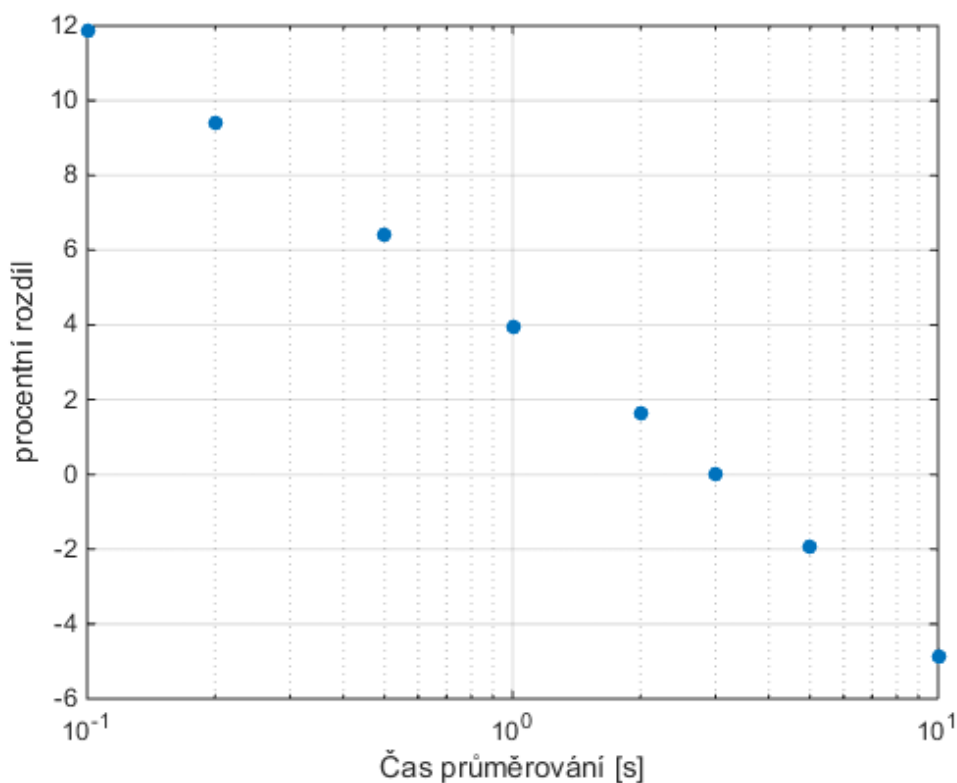
S opravenými daty byly spočítány průměrné rozdíly rychlostí nejvyšších denních nárazů větru pro modifikace (i), (ii) a (iii) a srovnány se základní metodou ve tvaru procentních rozdílů mezi průměrnými hodnotami nejvyšších denních nárazů, jak je u podobných studií obvyklé (Durst, 1960; Deaves, 1993; Choi et Hidayat, 2002; Kwon et Kareem, 2014). V případě (iii) byly použity délky průměrování 0,1; 0,2; 0,5; 1; 2; 5 a 10 s. Zmíněné články se zabývají vlivem délky průměrování na naměřené nejvyšší rychlosti větru, ale tyto délky jsou v nich podstatně delší než odpovídá případu (iii) – až 3600 s. Procentní rozdíly mezi nejvyššími nárazy s délkou průměrování 1 s a 3 s podle tří různých metodik podle Kwon et Kareem (2014) jsou 3,3 %, 5 % a 5,7 %.

Spočtené rozdíly pro uvedené tři modifikace jsou v tabulce 2.2.1. Nejvýraznější z vyzkoušených modifikací je vliv doby průměrování. Rozdíl mezi dobami průměrování 1 s a 3 s je v souladu s výsledky uvedenými v Kwon et Kareem (2014). Na obr. 2.2.1 je graf závislosti procentního rozdílu nejvyššího rozdílu na délce průměrování. Tato závislost se dá dobře aproximovat logaritmickou funkcí.

Spočtené rozdíly odpovídají měření na stanici v Kopistech. Z obdobné logiky jako uvedené v Larsén et al. (2006) lze očekávat, že se budou zvyšovat s vyšší mírou turbulence v místě měření.

případ	(i)	(ii)	0,1 s	0,2 s	0,5 s	1 s	2 s	5 s	10 s
rozdíl	0,4	-1,3	11,9	9,4	6,4	3,9	1,6	-1,9	-4,9

Tabulka 2.2.1: Procentní rozdíly mezi průměrnými nejvyššími denními nárazy spočítanými různými způsoby na stanici Kopisty oproti základní metodě: (i) odpovídá uvažování i vertikální složky rychlosti a (ii) používá pevné intervaly průměrování místo plovoucích, další sloupce odpovídají jiným dobám průměrování než jsou 3 s požadované WMO.



Obr. 2.2.1: Graf závislosti procentního rozdílu mezi průměrnými hodnotami maximálních denních nárazů na stanici Kopisty spočítané s použitím různě dlouhých intervalů průměrování.

2.3 Homogenizace denních nárazů větru

2.3.1 Úvod

Homogenní klimatická časová řada je definována jako řada, jejíž variabilita je způsobena pouze variabilitou klimatu. Dlouhodobě homogenní řady měření klimatických prvků jsou však vzácností, neboť velká většina těchto řad je ovlivněna též neklimatologickými faktory (Aguilar et al., 2003). K narušení homogenity dat může dojít především z následujících důvodů (Costa et Soares, 2009):

- a) změna místa měření,
- b) výměna měřicího přístroje,
- c) změny v okolí místa měření,
- d) stárnutí a poruchy přístroje,
- e) změny metod pozorování a zpracování dat.

Proces eliminace těchto nežádoucích nehomogenit se nazývá homogenizace. Jde o komplikovanou úlohu, pro kterou neexistuje univerzální metoda. Obvykle se skládá z následujících kroků (Aguilar et al., 2003):

- (i) analýza metadat a kontrola zjevných chyb měření,
- (ii) vytvoření referenční časové řady,
- (iii) detekce bodů zlomu a
- (iv) oprava naměřených hodnot.

Metadaty se v této souvislosti rozumí především informace o způsobu měření meteorologických dat a jeho změnách, zvláště o změnách měřicích přístrojů a jejich umístění.

Odborná literatura se nejvíce zabývá homogenizací řad teploty – např. Della-Marta et Wanner (2006), Camuffo (2002) nebo srážek – např. Alexandersson (1986), Rhoades et Salinger (1993), Guentchev et al. (2010), Wang et al. (2010). Existují i práce zabývající se vlhkostí a atmosférickým tlakem, např. Domonkos et Štěpánek (2009), Begert et al. (2008), Sugahara et al. (2012), Costa et Soares (2009).

2.3.2 Problém homogenizace větru a volby referenčních dat

Analýza homogenity dat i jejich homogenizace jsou v případě měření větru obtížnější než například pro teplotu. Výsledky měření větru jsou velmi citlivé na kvalitu měřicího přístroje a bezprostřední okolí místa měření (Wieringa, 1996), které se obojí mohou v čase skokově či pozvolna měnit. V případě změn v bezprostředním okolí místa měření nehomogenity navíc silně závisejí na směru větru. Ve výsledku je míra

nehomogenity větroměrných dat často značná, přičemž konkrétní nehomogenity je často obtížné až nemožné identifikovat a odstranit. To je nejspíše důvodem, proč se v literatuře práce týkající se homogenizace řad větru téměř nevyskytují (jedinou dohledanou prací na toto téma je níže diskutovaná studie (Azorin-Molina et al., 2014)).

Další komplikací je volba referenční řady. Obvyklým přístupem je využití řad měření daného prvku na okolních stanicích (Štěpánek et al., 2009). Ten má ale dvě rizika: jednak i řady z okolních stanic mohou být postiženy nehomogenitami a jednak přirozené chování daného prvku může být na okolních stanicích odlišné od místa měření homogenizované řady. V případě měření větru se oba tyto faktory uplatňují ve velké míře, proto je použití tohoto přístupu pro homogenizaci řad větru problematické.

V poslední době se proto k homogenizaci meteorologických dat často využívají reanalýzy získané aplikací numerických modelů atmosféry (Kalnay et al., 1996). Výhoda tohoto postupu spočívá ve skutečnosti, že reanalýzy poskytují poměrně dlouhé řady dat, které lze alespoň v hrubém přiblížení považovat za homogenní (Sterl, 2004). Pro získání věrnějších podmínek lze data reanalýz dále dále použít jako vstupní data simulací mezoměřítkovými modely v jemnějším rozlišení. Takto postupovali například Azorin-Molina et al. (2014), kteří použili mezoměřítkový model MM5 (Grell et al., 1994) řízený reanalýzou ERA 40 (Uppala et al., 2005). Nevýhodou tohoto přístupu je značná pracnost a výpočetní náročnost.

Homogenita samotných reanalýz je diskutována ve člancích Sterl (2004) a Gruber et Haimberger (2008). V nich je ukázáno, že v námi sledovaném období 1961 až 2009 nejsou reanalýzy NCEP-NCAR (Kalnay et al., 1996) i ERA-40 (Uppala et al., 2005) zcela homogenní. K nehomogenitě dochází zejména kolem roku 1980, kdy se do reanalýz začala zahrnovat data ze satelitů. Přesto se reanalýzy považují za podstatně homogennější než přízemní měření větru.

2.3.3 Popis a provedení použité metody

V této práci byla použita metoda homogenizace denních nárazů větru s využitím reanalýz NCEP/NCAR (Kalnay et al., 1996). Použitá metoda vychází ze dvou předpokladů. Za prvé se předpokládá, že existuje poměrný vztah mezi průměrem (za delší období, řádově několika let) z hodnot nejvyšších denních nárazů větru z daného směru a průměrem rychlostí větru v reanalýze v odpovídajících termínech. Druhým předpokladem je, že tento vztah se nebude v čase měnit. Tento předpoklad by mohl být

porušen, pokud by se např. vlivem klimatických změn změnila průměrná stabilita mezní vrstvy atmosféry. Riziko spojené s nenaplněním těchto předpokladů ale považujeme pro účel této práce za méně závažné než riziko nehomogenit přízemních měření rychlosti větru.

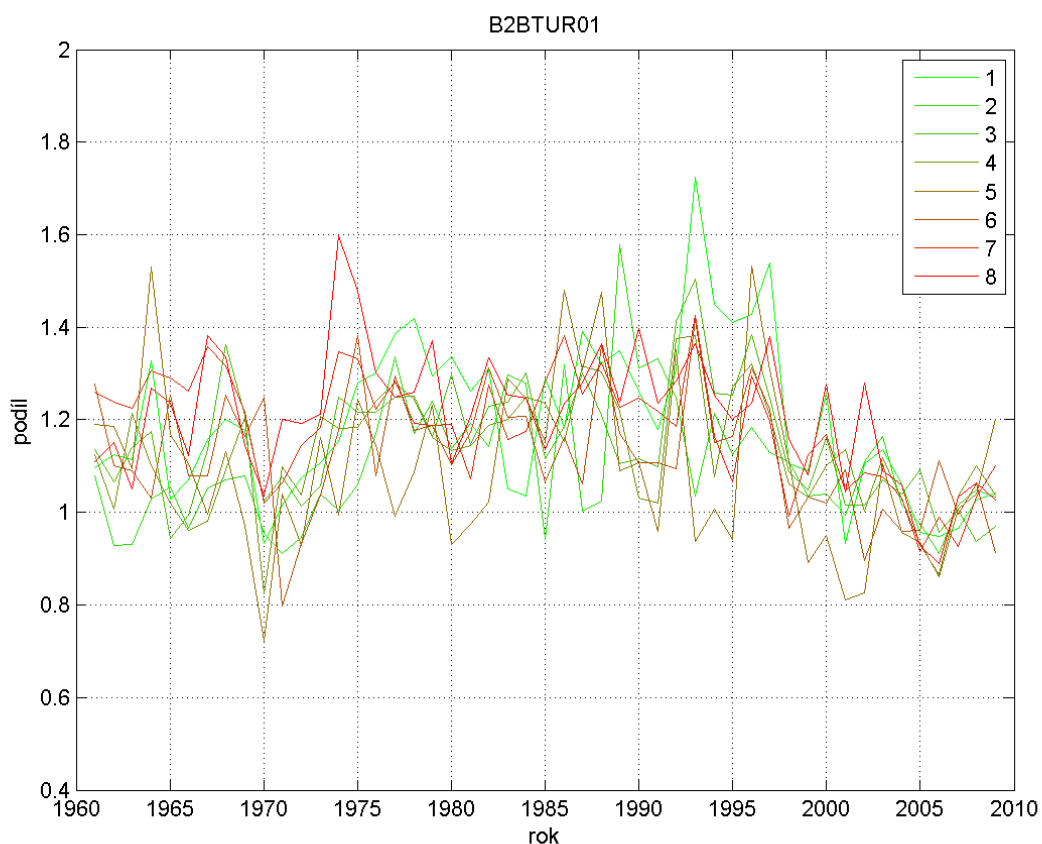
Prvním krokem použité metody je rozdělení nejvyšších denních nárazů větru do jednotlivých sektorů podle směru. Zvolený počet směrů je limitovaný nutností dostatečného počtu případů z jednotlivých směrů a strukturou historických dat. Bylo proto zvoleno členění do 8 sektorů. Vzhledem k tomu, že se v minulosti směr větru na stanicích udával v 16 světových směrech (sever, severoseverovýchod, severovýchod atd.), bylo zvoleno poněkud netradiční rozdělení na 8 sektorů tak, aby vždy dva z udávaných 16 směrů byly symetricky umístěné uvnitř daného sektoru. První sektor má tedy hranice 11,25 a 56,25 stupňů, další následují v pravidelných 45 stupňových intervalech. Vliv konkrétní volby hranic sektorů byl minimalizován použitím Gaussova filtru při výpočtu závislosti homogenizační funkce na směru, což bude popsáno níže.

Za „normálové“ období, na které byla data homogenizována, bylo určeno posledních 5 let měření, tedy období 2005 až 2009. Toto období bylo vybráno, protože je pro něj k dispozici nejvíce údajů o měření včetně podrobného měření průměrné rychlosti větru a lze doložit, že kvalita měření je vyšší než v předchozích letech (Řepka, 2011). Použití nejnovějšího období jako normálového doporučují také Aguilar et al. (2003).

K homogenizaci byla použita data o rychlosti a směru větru v nejbližším bodě v síti reanalýz NCEP/NCAR ke stanici. Jako referenční hodnoty byla použita data z hladin 10 m, 925 hPa a 850 hPa a z času reanalýz, který byl nejbližší k času naměření nejvyšší denní rychlosti větru. Pro danou stanici byla použita data z hladiny, která vykazovala pro danou stanici a sektor nejvyšší korelační koeficient s naměřenými rychlostmi. V 116 případech vyšla hladina 10 m, v 21 případech hladina 925 hPa a v 95 případech hladina 850 hPa. Byla testována i hladina 700 hPa, u které však vyšel nejvyšší korelační koeficient pouze v jednom případě, a nebyla proto použita.

Pro každou stanici, rok měření a sektor byl spočítán podíl sum naměřených nejvyšších denních nárazů větru na stanici a sum rychlostí větru v nejlépe korelované hladině reanalýz v odpovídajících termínech a tento podíl byl vydělen průměrnou

hodnotou tohoto podílu za roky 2005 až 2009. Tyto podíly budou dále označovány jako "homogenizační podíly".



Obr. 2.3.1: Časová závislost homogenizačních podílů v jednotlivých sektorech pro stanici Brno-Tuřany před aplikací Gaussova filtru.

Typický obrázek homogenizačních podílů pro jednotlivé sektory a jednotlivé roky je ilustrován na obr. 2.3.1. Z grafu je zřejmá velká meziroční variabilita tohoto podílu. Křivky pro různé sektory se liší, přičemž bližší sektory mají podobnější průběh než vzdálenější, ale v určitých obdobích má homogenizační podíl rostoucí nebo klesající tendenci ve všech sektorech současně. První případ lze vysvětlit změnou, která se týká určitého směru, zatímco druhý změnou v samotném přístroji, např. jeho stárnutím nebo výměnou za nový stejného nebo i jiného typu. Odlišit vliv přirozené variability a nehomogenit je v tomto grafu prakticky nemožné, změny navíc obvykle nejsou podloženy metadaty. Situace v případě ostatních stanic je obdobná.

Existuje ale jedna výjimka. Znatelný zlom lze identifikovat mezi roky 1997 a 1998, kdy došlo k zřetelnému snížení naměřených rychlostí větru oproti reanalýzám. Obdobné

snížení rychlosti se vyskytuje u velké většiny stanic, a časově odpovídá změně přístroje pro určení nejvyšších rychlostí větru z Pitotovy trubice na anemometr Vaisala WAV 151, případně Vaisala WAV 251.

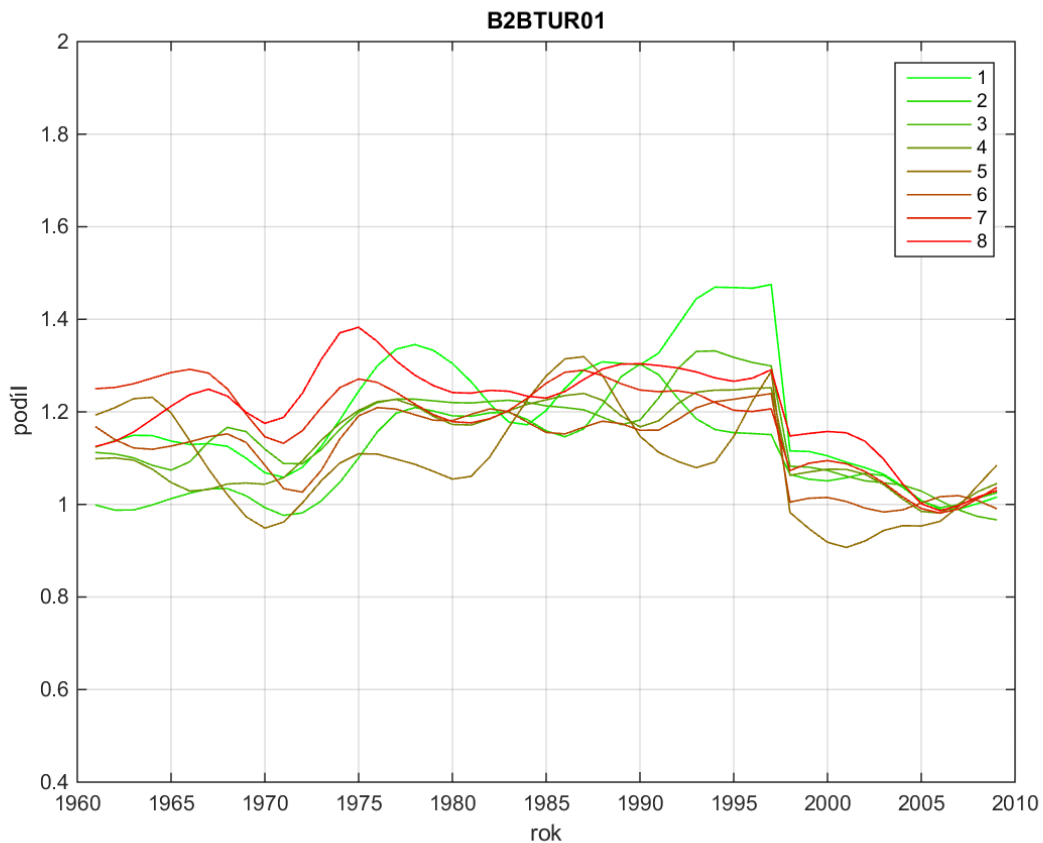
Dalším krokem je odstranění krátkodobých náhodných fluktuací homogenizačních podílů, k čemuž byl použit Gaussův filtr se směrodatnou odchylkou 1,6 roků. Volba této hodnoty je popsána v sekci 2.3.5. V případě, že data byla ovlivněna přechodem na anemometr firmy Vaisala (viz výše), byl tento filtr aplikován zvlášť na data před a po změně přístroje. Data byla považována za ovlivněná změnou přístroje, pokud se průměrný homogenizační podíl v posledních 3 letech před změnou a v prvních 3 letech po změně ve většině sektorů lišil alespoň o 10 %. Výsledek aplikace tohoto časového Gaussova filtru pro stanici Brno-Tuřany je zobrazen na obr. 2.3.2.

Takto spočtené homogenizační podíly vykazují skokové změny na hranicích sektorů. Proto jsme dále aplikovali směrový Gaussův filtr se směrodatnou odchylkou 22°.

Nakonec byly naměřené hodnoty nárazů větru homogenizovány vydělením maximálních denních nárazů větru "korekčním faktorem" F ve tvaru

$$F(T, D) = \frac{GFD(GFT(h(t,d),T,1.6),D,22)}{GFD(\frac{1}{5}\sum_{t=2005}^{2009} h(t,d),D,22)}, \quad (2.3.1)$$

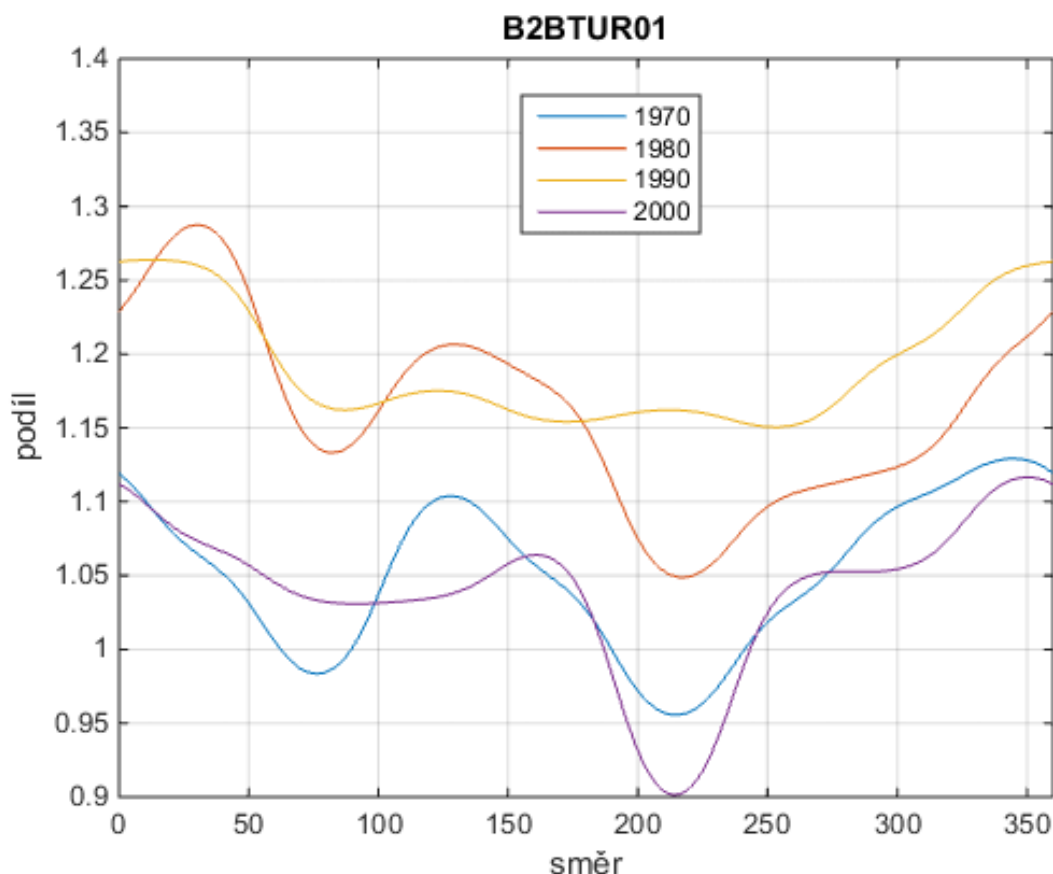
kde T je čas naměřeného nárazu, D je jeho směr, $h(t, d)$ je homogenizační podíl před aplikací Gaussových filtrů (obr. 2.3.1) pro rok t a sektor d , $GFD(x, D, \sigma)$ je hodnota proměnné x pro směr D filtrovaná směrovým Gaussovým filtrem se směrodatnou odchylkou σ a $GFT(x, T, \sigma)$ je hodnota proměnné x pro čas T filtrovaná časovým Gaussovým filtrem se směrodatnou odchylkou σ .



Obr. 2.3.2: Časová závislost homogenizačních podílů v jednotlivých sektorech pro stanici Brno-Tuřany po aplikaci časového Gaussova filtru.

2.3.4 Korekční faktory na stanici Brno-Tuřany

Chování korekčních podílů a jejich časová variabilita je ilustrována pro stanici Brno-Tuřany a roky 1970, 1980, 1990 a 2000 na obr. 2.3.3. Výkyvy těchto podílů směrem dolů lze vysvětlit překážkou, která snižovala v daném roce rychlost větru z daného směru, ale která se nevyskytovala v období 2005 až 2009. Naopak výkyvy směrem nahoru mohou být způsobeny překážkou, která se v daném směru vyskytovala v období 2005 až 2009, ale ne v roce, pro který je daná křivka spočítána. Z grafu se jeví jako pravděpodobné, že v letech 1970, 1980 a 2000 byl přístroj zastíněn ze směru cca 215° oproti referenčnímu období a naopak ve směru cca 0° byl přístroj pravděpodobně více zastíněn v referenčním období. Fakt, že tyto výkyvy (i další) se chovají často velmi podobně pro různé roky, přidává váhu názoru, že jde o skutečné jevy a ne jen o náhodné fluktuační.



Obr. 2.3.3: Výsledná směrová závislost korekčních podílů F pro roky 1970, 1980, 1990 a 2000 pro stanici Brno-Tuřany.

2.3.5 Verifikace výsledků a volba parametrů Gaussových filtrů ve vzorci 2.3.1

K verifikaci použité metody homogenizace byl použit Kolmogorovův-Smirnovův test (Wilks, 2011). Nejprve byly z časové osy pro každou stanici odstraněny dny, pro které nebyl k dispozici nejvyšší denní náraz, a pro každou stanici bylo vybráno 50 nejvyšších naměřených hodnot nejvyššího denního nárazu větru. Pokud by řada byla homogenní, a odhlédneme-li od kolísání klimatu a ročního chodu, mělo by rozdělení časů měření těchto nejvyšších hodnot mít rovnoměrné rozdělení mezi začátkem prvního a koncem posledního dne měření. I když tyto předpoklady nejsou splněny, lze očekávat, že by vhodná metoda homogenizace měla rozdělení časů měření nejvyšších hodnot nárazů větru přiblížit rovnoměrnému rozdělení.

Jsme si vědomi toho, že při vyhodnocování výsledků naše homogenizace využívá i těchto 50 nejvyšších naměřených hodnot pro každou stanici, a ověřuje se tedy na datech, ze kterých vychází. Tuto závislost ale nepovažujeme za zásadní, protože

zmíněných 50 nejvyšších hodnot tvoří na jednotlivých stanicích méně než 0,5 % všech naměřených hodnot, ze kterých metoda homogenizace vychází.

Další komplikací při vyhodnocení je různá délka naměřených řad, takže zmíněných 50 nejvyšších denních nárazů pro jednotlivé stanice představuje různou „extremitu“. Toto by vyřešila modifikace, která by místo pevného počtu 50 nejvyšších hodnot brala počet hodnot přímo úměrný celkovému počtu naměřených hodnot na dané stanici. Protože se ale velmi obtížně porovnávají výsledky Kolmogorova-Smirnovova testu mezi případy s různým počtem dat, ze kterých vychází, byl zvolen původní postup, tedy pevného počtu vybraných nejvyšších případů.

Byly sledovány dvě testovací statistiky, které vyjadřují míru splnění předpokladu, že časy 50 nejvyšších denních nárazů větru pocházejí z rovnoměrného rozdělení. První testovací statistikou je počet stanic, na kterých se nedá na hladině spolehlivosti 5 % vyloučit, že časy pocházejí z rovnoměrného rozdělení. Druhou testovací statistikou je počet stanic, na kterých vyšel Kolmogorovův-Smirnovův test lépe pro homogenizovaná než původní data.

Výsledky těchto testovacích statistik byly nejprve použity k volbě směrodatných odchylek Gaussových filtrů použitých v kapitole 2.3.3 pro odstranění krátkodobých fluktuací homogenizačních podílů a k odstranění jejich skokových změn na hranici sektorů. Tyto směrodatné odchylky byly zvoleny tak, aby zvolené testovací statistiky dávaly nejlepší výsledky. Obě testovací statistiky dávají velmi podobný výsledek a shodují se na optimální hodnotě přibližně 1,6 roku v případě časového Gaussova filtru a 22° v případě směrového Gaussova filtru (přesné hodnoty nebylo možné takto určit, protože obě testovací statistiky jsou diskrétní).

Při aplikaci Kolmogorova-Smirnovova testu a výpočtu hladin spolehlivosti byla předpokládána nezávislost časů nejvyšších denních nárazů větru na jednotlivých stanicích. Získané výsledky testu ukázaly, že pro původní data lze u 22 stanic z 29 na hladině spolehlivosti 5 % vyloučit hypotézu, že časy pocházejí z rovnoměrného rozdělení, což odpovídá zamítnutí nulové hypotézy, že časy pocházejí z rovnoměrného rozdělení na všech stanicích, na hladině spolehlivosti přibližně $8,4 \cdot 10^{-25}$. Po provedení homogenizace lze tuto hypotézu vyloučit u 7 stanic, což odpovídá hladině spolehlivosti $9,1 \cdot 10^{-4}$. Homogenizace tedy výrazně zvýšila počet stanic, u kterých se neprokázala nerovnoměrnost časů nejvyšších denních nárazů větru. Z tohoto hlediska je zřejmé, že

navržená metoda zlepšuje homogenitu dat. Zároveň je však třeba konstatovat, že i po homogenizaci časy extrémů nepocházejí z rovnoměrného rozdělení.

Kolmogorovův-Smirnovův test vyšel u 24 stanic lépe po homogenizaci než v případě původních dat, což vede k zamítnutí nulové hypotézy, že rovnoměrnost časů nejvyšších denních nárazů větru po homogenizaci zůstala stejná, na hladině významnosti $5,5 \cdot 10^{-4}$. Celkově testy potvrdily, že homogenizace zlepšila kvalitu dat.

Sedm stanic, u kterých po navržené homogenizaci na hladině významnosti 5 % lze vyloučit, že časy nejvyšších denních nárazů větru pocházejí z rovnoměrného rozdělení, jsou (v pořadí od nejhorší) Kopisty, Holešov, Lysá hora, Doksany, Karlovy Vary, Svratouch, Ruzyně. Lze předpokládat, že jejich homogenita je i po provedené homogenizaci nejvíce problematická. Nevidíme mezi nimi na první pohled žádnou zřejmou souvislost.

2.4 Základní statistické vlastnosti nárazů větru na území ČR

V této kapitole jsou analyzovány maximální denní nárazy větru, které byly homogenizovány metodou popsanou v podkapitole 2.3. Kromě toho jsou použita data z databáze CLIDATA (Řepka, 2011) obsahující maximální nárazy z 15minutových intervalů. Tato data homogenizována nebyla z důvodu malé délky jejich řad.

2.4.1 Denní a roční chod větru a nárazů větru na území ČR v dřívějších pracích

Ročním chodem počtu dní se silným větrem se zabývají publikace Brázdila et al. (2004 a 2007). Za den se silným větrem je v těchto pracích považován den s nejvyšším nárazem vyšším nebo rovným než 17 m/s a byla zpracována data nejvyšších denních nárazů z let 1961-1990. Data nebyla homogenizována. Roční chod výskytu těchto dní tvoří podle těchto prací jednoduchou křivku s maximem v zimě nebo na začátku jara. Denní chod případů s extrémním větrem v těchto pracích diskutován není. Jak ukážeme později, tato data jsou k určení denního chodu nedostatečná.

Ročním i denním chodem se zabývala zpráva Sokol et al. (2007), ale ta se zabývala průměrnou rychlostí větru, ne jejími extrémy. Také vycházela z jiných zdrojů naměřených dat. Byly použity zprávy SYNOP (ČHMÚ, 2007) z období 1996 až 2006, ve kterých je zaznamenána průměrná rychlost větru za posledních 10 minut každé hodiny.

Dále byla použita data ze stanic Automatického imisního monitoringu (AIM), které měří rychlosti větru kontinuálně, ale kvalitě jejich měření se nevěnuje taková pozornost jako profesionálním stanicím ČHMÚ (Řepka, 2011). Dalším typem dat zpracovaným v této studii jsou stožárová měření, která přinášejí informace i z vyšších hladin než je standardních 10 m a obsahují průměrnou rychlost větru po 10 minutách. Tato měření ale bývají velmi krátká (často kratší než 1 rok), nejsou standardizovaná a trpí častými výpadky. Roční chod v této zprávě vyšel jako jednoduchá vlna s maximem v zimě. V případě denního chodu vyšla plochá závislost v nočních hodinách. V denních hodinách dochází ke zvýšení průměrné rychlosti větru s výjimkou vrcholových stanic, kde dochází naopak ke snížení průměrné rychlosti větru. Zvýšení průměrné rychlosti větru v denních hodinách roste s klesající nadmořskou výškou stanic. Amplituda denního chodu rychlosti větru se výrazně mění v průběhu roku, největší je v létě a nejmenší v zimě, což ukazuje na to, že denní chod je způsoben radiačními procesy. Naproti tomu roční chod je dosti podobný na všech stanicích, proto předpokládáme, že je z velké části způsoben zesílením průměrného tlakového gradientu na území České republiky v zimních měsících (Tolasz et al., 2007).

Většina publikovaných prací se zabývá pouze rychlostí větru. Periodicitou směru větru na stanici Praha-Ruzyně se zabývali Kalvová a Sobišek (1981) a našli dvě periody: 24hodinovou způsobenou denním chodem a 96hodinovou způsobenou změnami synoptické situace.

2.4.2 Kvantily a maximální hodnoty nejvyšších denních nárazů větru

Byly spočteny mediány, kvantily (0,9, 0,99, 0,999) a maximální hodnoty naměřených nejvyšších denních nárazů větru na profesionálních stanicích ČHMÚ, na kterých bylo dostupných minimálně 20 let měření z let 1961 až 2009. Výsledky jsou uvedeny v tabulce 2.4.1.

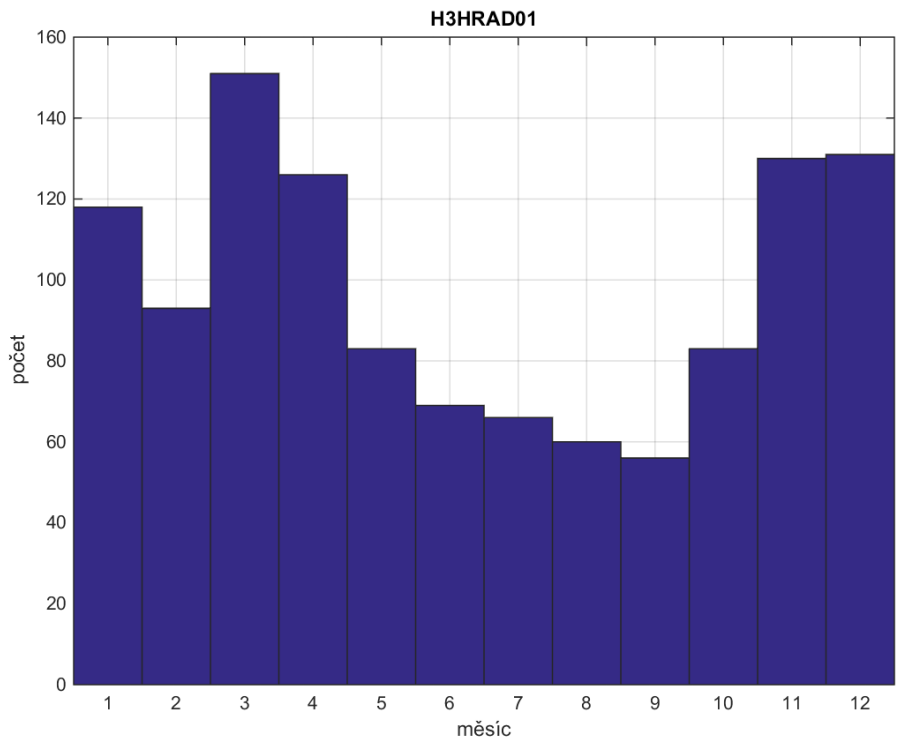
indikativ	medián	Q _{0,90}	Q _{0,99}	Q _{0,999}	maximum
B1HOLE01	9,3	14,6	20,7	26,6	35,5
B2BTUR01	9,6	14,9	21,0	26,7	35,0
B2DUKO01	9,4	14,5	20,5	25,6	29,5
B2KMYS01	9,7	16,0	22,7	29,8	39,6
B2KUCH01	9,7	16,0	23,0	30,3	34,3
B2NEDV01	10,3	16,3	22,4	28,1	31,0
C1CHUR01	9,3	16,1	22,7	27,6	38,4
C1KOCE01	10,2	17,2	25,2	30,8	43,3
H3HRAD01	9,5	15,6	23,3	29,0	41,9
H3SVRA01	13,9	22,6	31,1	38,1	44,7
L2PRIM01	11,1	18,1	25,4	30,7	39,9
L3CHEB01	8,3	14,0	20,4	25,5	34,8
L3KVAL01	8,6	14,3	21,0	26,5	30,7
O1CERV01	10,7	16,4	22,0	27,4	33,0
O1LYSA01	15,3	24,7	33,8	41,4	51,3
O1MOSN01	9,9	15,5	21,9	27,5	33,5
O1PORU01	8,2	12,9	17,6	21,6	27,2
O2LUKA01	10,5	17,6	26,7	33,0	37,3
O3PRER01	8,4	13,8	19,8	26,1	30,6
O7KRNO01	9,8	17,1	22,9	27,4	32,8
P1PKAR01	10,1	16,9	25,0	32,5	36,3
P1PLIB01	9,0	14,9	21,4	27,6	33,4
P1PRUZ01	9,1	15,3	22,0	28,3	34,6
P3PRIB01	10,0	16,1	22,6	29,5	37,3
U1DOKS01	8,1	14,2	21,7	28,9	33,8
U1KATU01	10,3	18,3	26,9	34,2	41,3
U1KOPI01	8,0	13,8	21,6	28,4	32,1
U1MILE01	16,1	26,0	37,0	43,2	47,8
U2LIBC01	9,8	15,4	22,6	28,4	37,7

Tabulka 2.4.1. Seznam českých stanic s alespoň 20letým měřením z let 1961 až 2009. Tabulka obsahuje indikativy stanic, mediány, kvantily (0,9, 0,99, 0,999) a maximální hodnoty naměřených nejvyšších denních nárazů větru. Hodnoty jsou v m/s. Medián má na většině stanic hodnotu kolem 10 m/s a maximum kolem 35 m/s. Hodnoty mediánu a kvantilů dosahují nejvyšších hodnot na stanici Milešovka. Nejvyšší hodnota nárazu větru 51,3 m/s byla naměřena 27. 12. 1990 ve tři hodiny ráno na stanici Lysá hora.

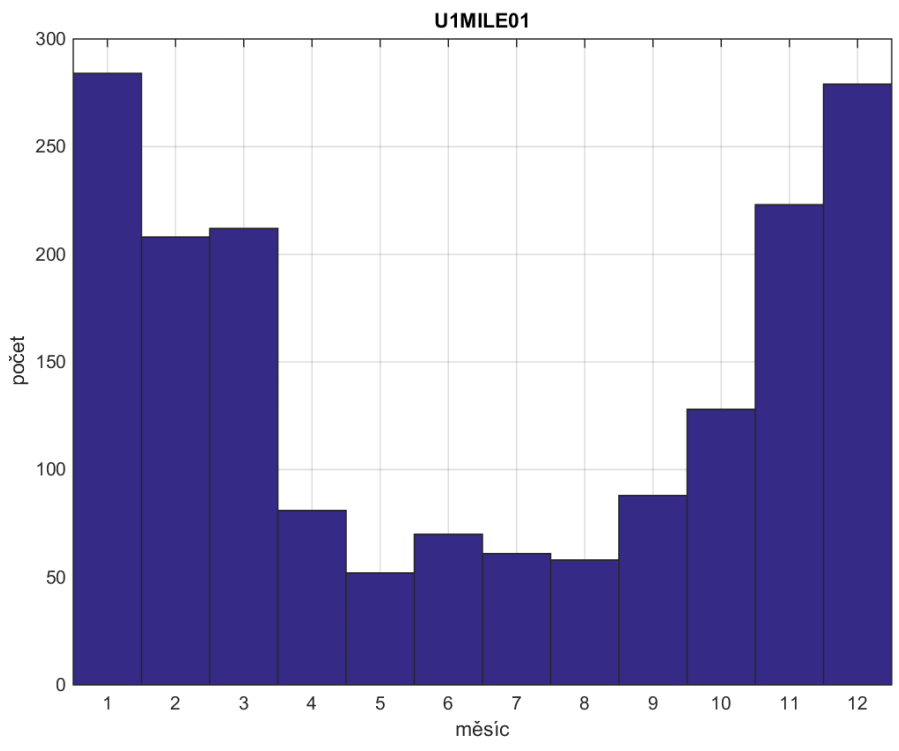
2.4.3 Roční chod nejvyšších denních rychlostí větru

Roční chod byl spočítán z nejvyšších denních nárazů z let 1961 až 2009. Na rozdíl od prací Brázdila et al. (2004 a 2007) bylo použito delší období a data byla homogenizována. Dále byl zvolen jiný způsob výběru dnů se silnými nárazy větru. Místo použití pevné hranice 17 m/s bylo na každé stanici vybráno 10 % dní s nejvyššími naměřenými nárazy větru. Maximální výskyt těchto dnů se vyskytl v zimě. V práci Brázdil et al. (2004) byl mimo jiné zkoumán vliv zvýšení hranice ze 17 m/s na 20 m/s, 25 m/s a 30 m/s a bylo konstatováno, že se velikost ročního chodu se zvyšující se hranicí zvyšuje. Toto zjištění se potvrdilo i na našich datech při omezení počtu zkoumaných dat na 1 % nejvyšších denních nárazů na každé stanici, což způsobilo zvýraznění ročního chodu oproti původnímu výběru 10 % nejvyšších případů (dní). Amplituda ročního chodu v práci Brázdila et al. (2004) vykazuje variabilitu mezi jednotlivými stanicemi, ale není zde patrná nějaká výrazná závislost na nadmořské výšce nebo exponovanosti stanice.

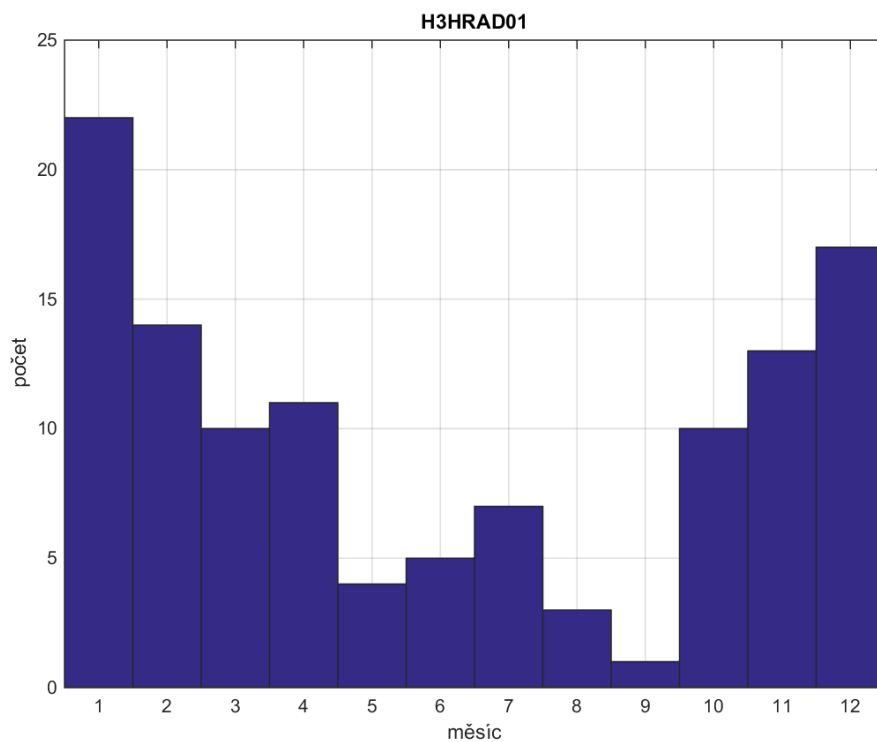
Zde vypočtené roční chody pro nížinné a vrcholové horské stanice se liší od výsledků Brázdila et al. (2004) a jsou zobrazeny na obr. 2.4.1 až 2.4.4 pro stanice Hradec Králové a Milešovka. Hlavní rozdíl je v tom, že na našich datech je amplituda ročního chodu vyšší na horských exponovaných stanicích. Rozdílné výsledky lze vysvětlit výběrem zpracovávaných dat. Při použití pevné hranice rychlosti větru je na těchto stanicích použit vyšší podíl dat, což redukuje velikost ročního chodu na těchto stanicích ve srovnání s přístupem, kdy byl použit pevný podíl nejvyšších dat, vzhledem k výše zmíněné skutečnosti, že velikost ročního chodu se zvětšuje s extremitou vybraných dat.



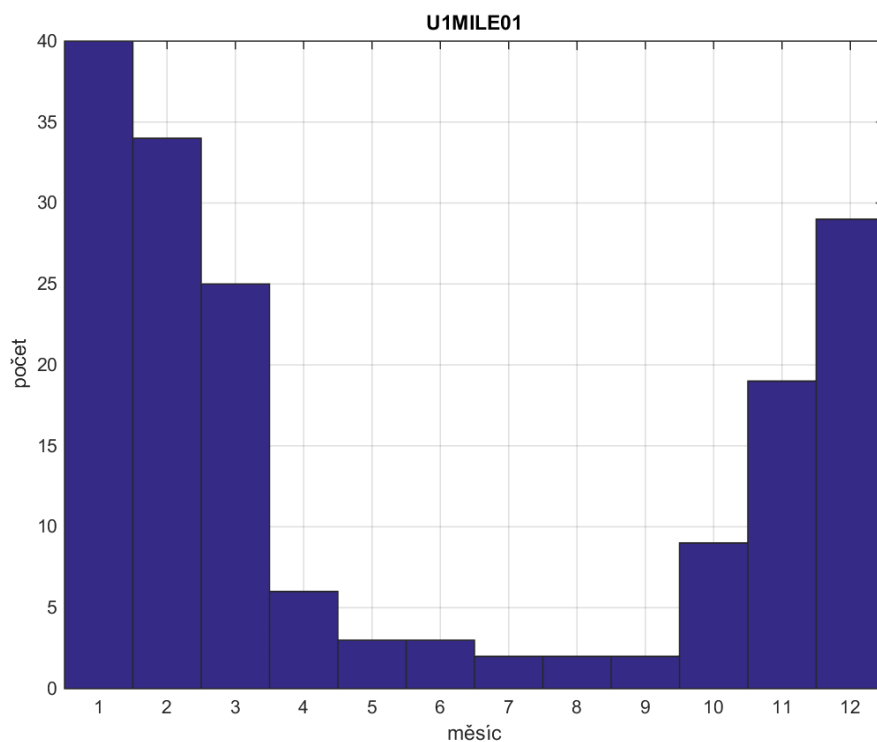
Obr. 2.4.1: Roční chod 10 % nejvyšších denních nárazů na nížinné stanici Hradec Králové.



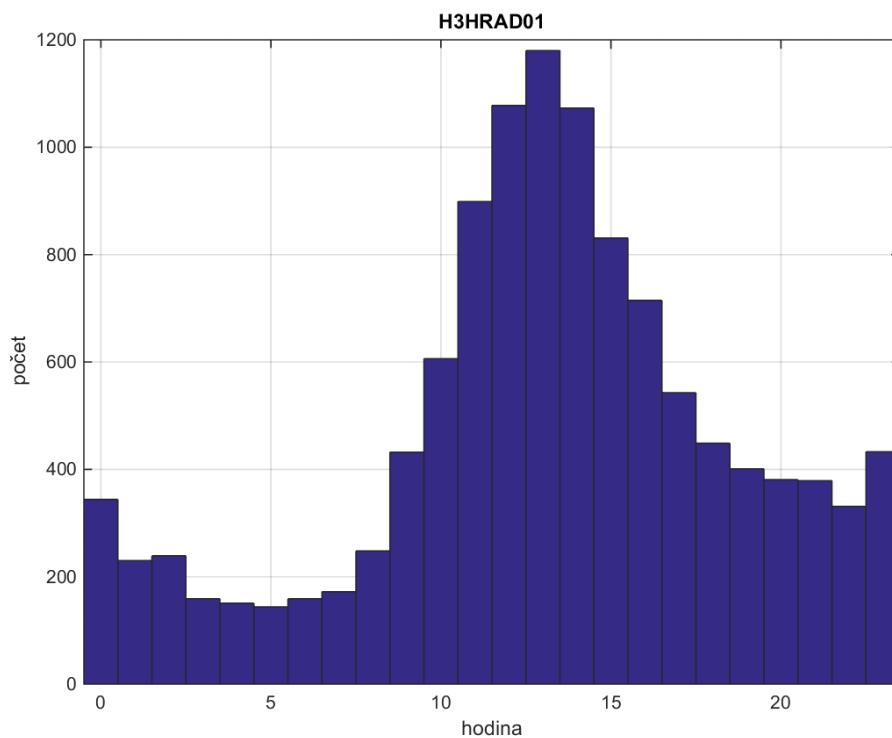
Obr. 2.4.2: Roční chod 10 % nejvyšších denních nárazů na horské stanici Milešovka.



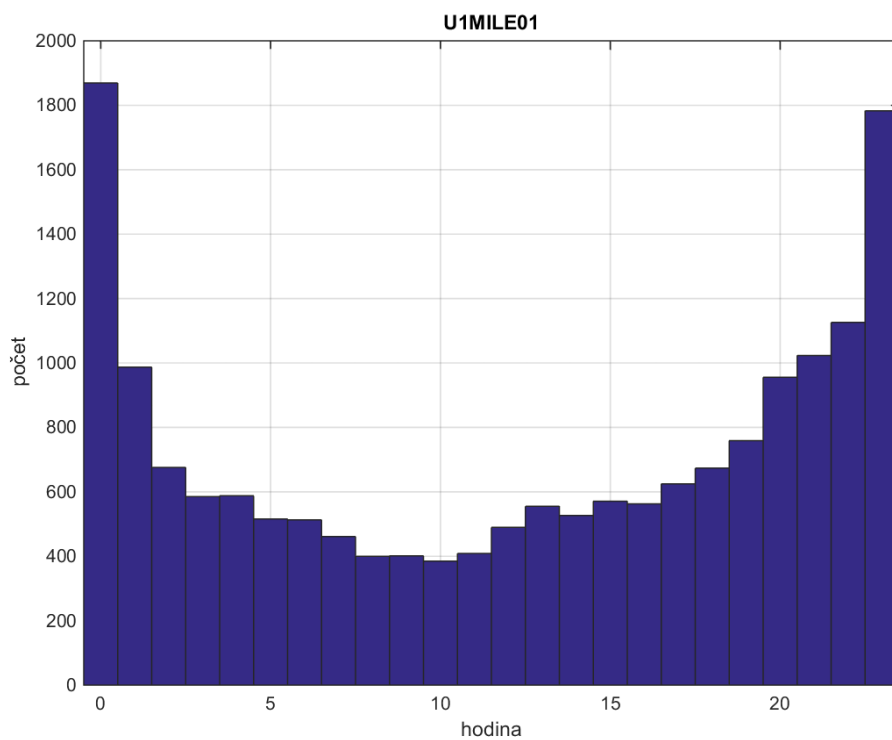
Obr. 2.4.3: Roční chod 1 % nejvyšších denních nárazů na nížinné stanici Hradec Králové.



Obr. 2.4.4: Roční chod 1 % nejvyšších denních nárazů na horské stanici Milešovka.



Obr. 2.4.5: Denní chod časů nejvyšších denních nárazů na nížinné stanici Hradec Králové.



Obr. 2.4.6: Denní chod časů nejvyšších denních nárazů na stanici Milešovka.

2.4.4 Denní chod vysokých nárazů rychlosti větru

Nejvyšší denní nárazy rychlosti větru se ukázaly jako nevhodné pro určení denního chodu událostí se silným větrem. Časy těchto událostí vykazují oproti dennímu chodu zjištěnému v práci Sokol et al. (2007) zvýšenou četnost okolo půlnoci, a to na všech stanicích, přičemž zvýšení je výraznější na horských vrcholových stanicích, jak je patrné z obr. 2.4.5 a 2.4.6.

Toto zvýšení má jednoduché vysvětlení. Rychlost větru je kromě denního chodu dána synoptickou situací, která se často mění v řádu dní, a v důsledku toho mívá rychlost větru až po několik dní vzestupnou nebo naopak sestupnou tendenci. Protože je denní maximum maximálním nárazem větru v období od půlnoci do půlnoci následující, nastane toto maximum v případě vzestupné tendence rychlosti větru pravděpodobně na konci dne a v případě sestupné tendence naopak pravděpodobně na začátku dne. Tím dochází k deformaci denního chodu tak, že se zvyšuje četnost maxim naměřených v okolí hranice mezi dvěma uvažovanými obdobími. Na horských vrcholových stanicích se tento jev projevuje více než na ostatních stanicích, protože proudění na nich se blíží proudění ve volné atmosféře, které je téměř výlučně určeno synoptickou situací (Holton, 2004).

Proto se pro výpočet denního chodu rychlosti větru použily hodnoty maximálních nárazů větru za každých 15 minut z databáze CLIDATA. Tato data byla k dispozici za podstatně kratší dobu než data maximálních denních nárazů a to na většině stanic pouze z období 2004-2009. Byl zkoumán denní chod dvou souborů dat:

1. hodnot nárazů splňujících podmínku, že v předchozích 12 hodinách nebyl zaznamenán vyšší nebo stejný náraz větru a v následujících 12 hodinách nebyl zaznamenán vyšší náraz větru – v_1 ,
2. maximálních hodnot nárazů ve všech naměřených 15 minutových intervalech – v_2 .

První soubor dat, v_1 , byl vybrán, protože by měl přibližně odpovídat maximálním denním nárazům zkoumaným v dřívějších pracích. Soubor v_2 obsahuje mnohem více denních dat a denní chod jejich počtu je z definice přibližně rovnoměrný. Proto je třeba se zaměřit na denní chod výskytu hodnot z některého jejich nejvyššího kvantilu.

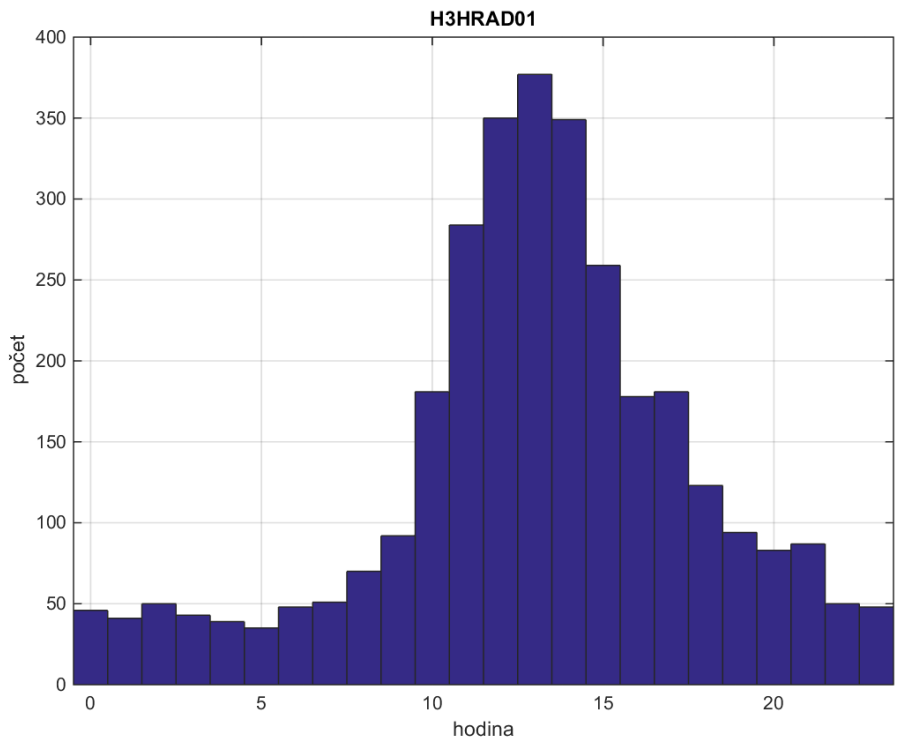
Lze očekávat, že v souborech dat v_1 a v_2 se odlišným způsobem projeví konvektivní a velkoprostorové události se silným větrem. Konvektivní událost se pravděpodobně projeví jedinou hodnotou v souborech dat v_1 i v_2 , zatímco velkoprostorová událost se pravděpodobně projeví jedinou hodnotou v souboru dat v_1 , ale více hodnotami

v souboru v_2 . To způsobí, že v souboru dat v_2 bude ve srovnání se souborem v_1 vyšší podíl hodnot z velkoprostorových událostí oproti konvektivním.

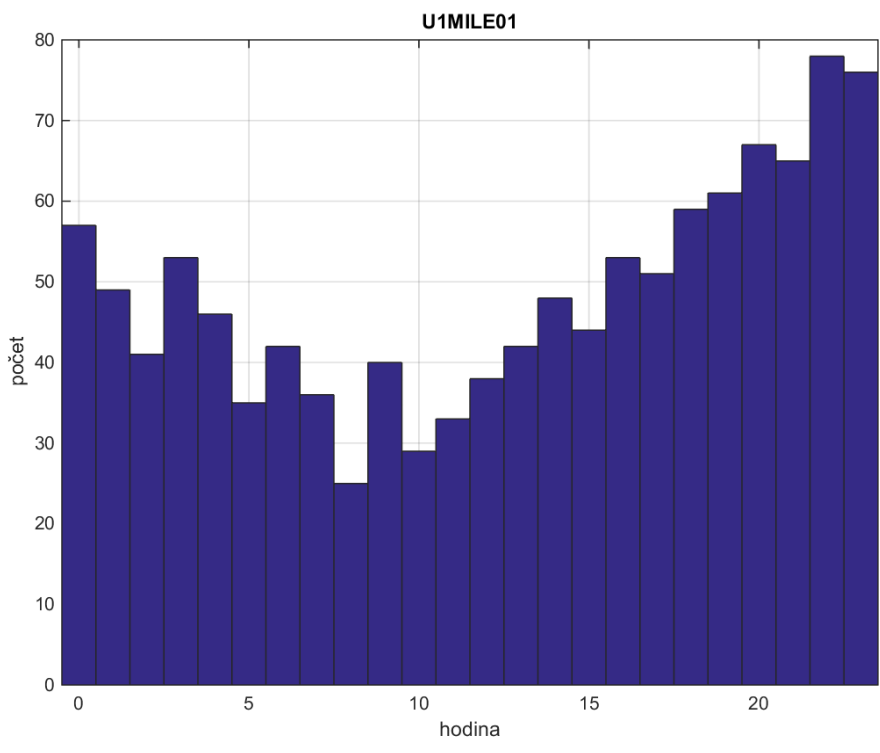
Srovnání souborů dat v_1 a v_2 na stanicích Hradec Králové a Milešovka je na obr. 2.4.7 až 2.4.16. Denní chod je mnohem výraznější na stanici Hradec Králové než na stanici Milešovka. Na stanici Hradec Králové je výrazné maximum denního chodu krátce po poledni a toto maximum je výraznější na datech v_1 než v případě dat v_2 . To souhlasí s výše uvedenou úvahou a odpovídá to skutečnosti, že maximum konvektivní aktivity nastává v odpoledních hodinách. V případě stanice Milešovka nastalo pro data v_1 minimum denního chodu v ranních hodinách a maximum ve večerních hodinách. V ostatních případech je denní chod nevýrazný.

Podle práce Sokol et al. (2007) má denní chod rychlosti větru na všech stanicích kromě horských výrazné maximum po poledni a na horských vrcholových stanicích má naopak v této době nevýrazné minimum. V odpoledních hodinách nastává maximum konvektivní aktivity a tím dochází na níže položených stanicích k dalšímu zesílení výskytu extrémních rychlostí větru, což vysvětluje velmi výrazný denní chod na stanici Hradec Králové. Naproti tomu na vrcholových stanicích proti sobě působí maxima konvektivní aktivity a minima průměrné rychlosti větru. Tím lze vysvětlit nevýrazný denní chod vysokých rychlostí větru na stanici Milešovka.

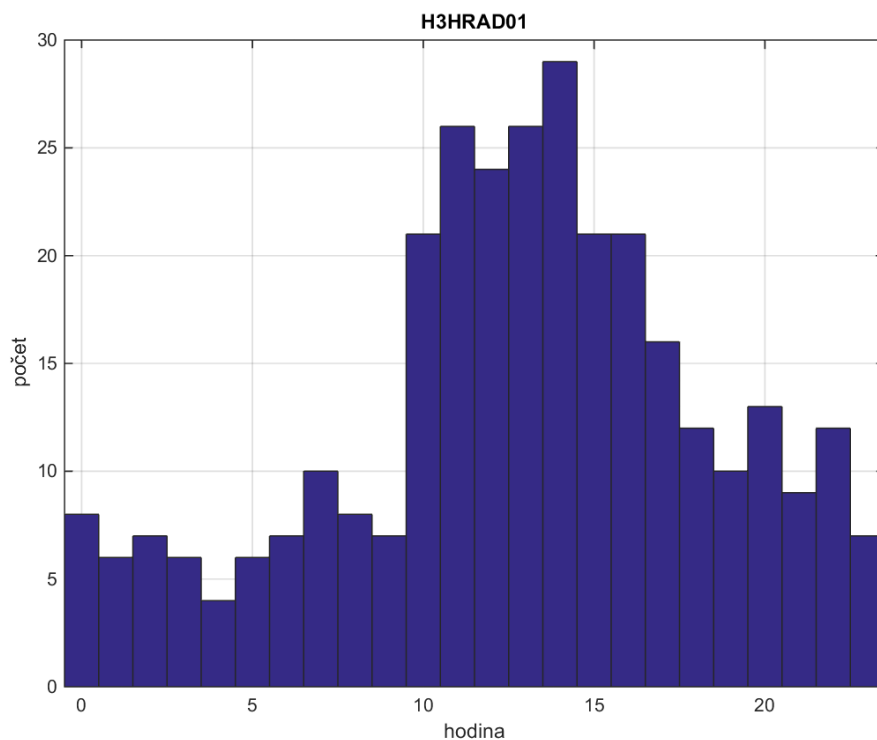
Na rozdíl od ročního chodu se při zvyšování extremity vybraných událostí amplituda denního chodu snižuje. To vyplývá z toho, že tyto události nastávají výrazně častěji v zimě, kdy je denní chod výskytu silných rychlostí větru slabší než v létě. To ilustrují obr. 2.4.17 až 2.4.24.



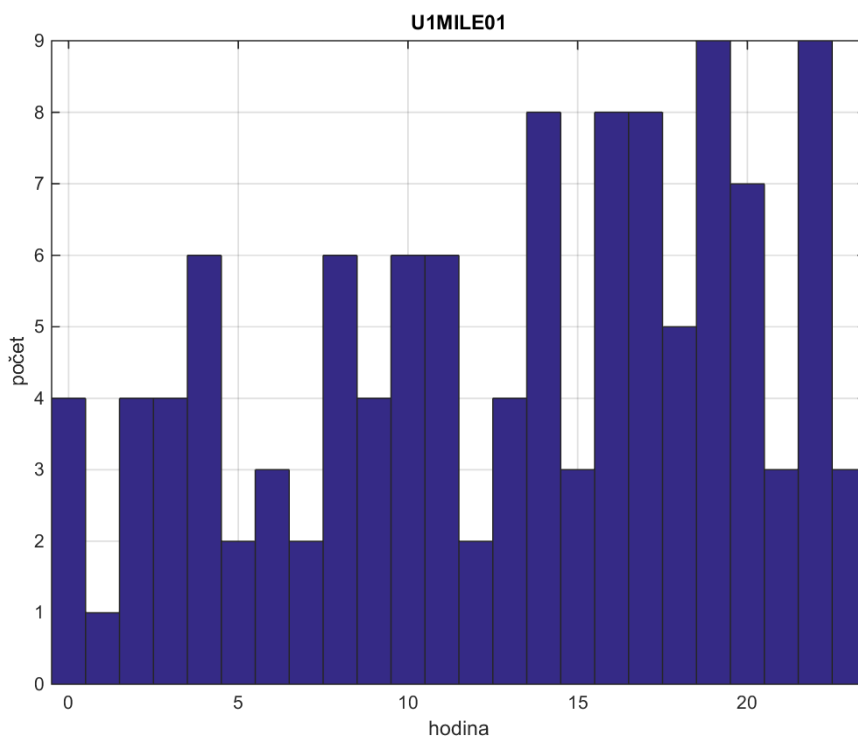
Obr. 2.4.7 Denní chod souboru dat v_I na nížinné stanici Hradec Králové.



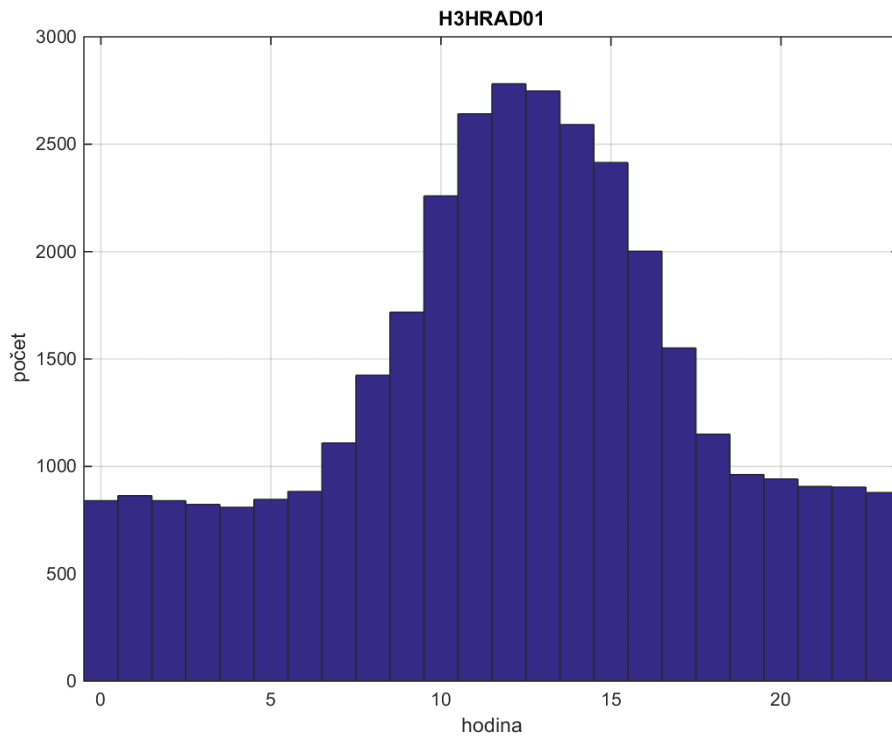
Obr. 2.4.8 Denní chod souboru dat v_I na horské stanici Milešovka.



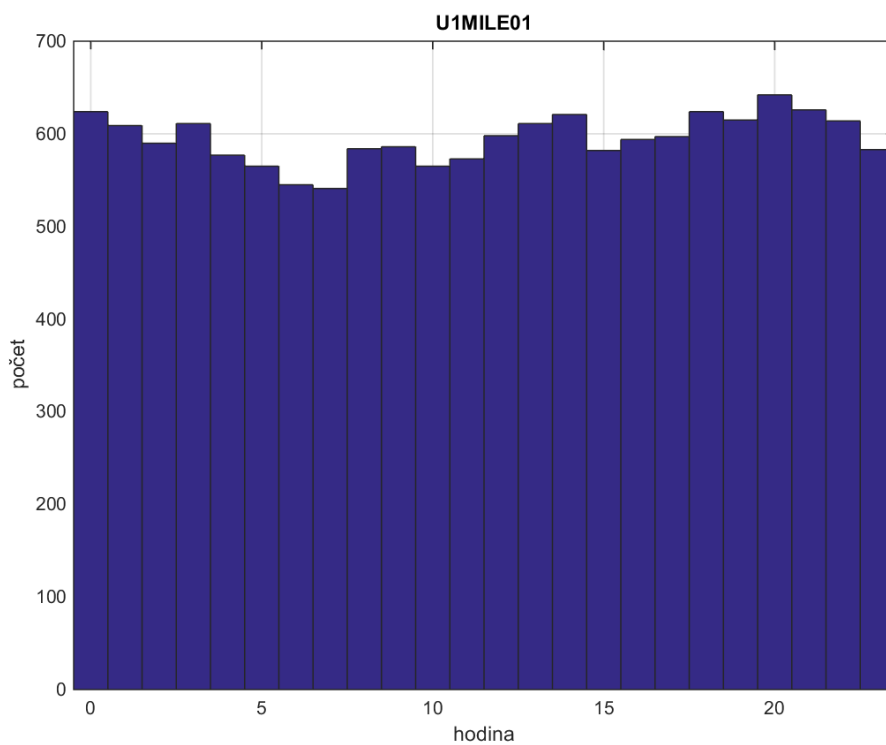
Obr. 2.4.9 Denní chod nejvyšších 10 % ze souboru dat v_1 na nížinné stanici Hradec Králové.



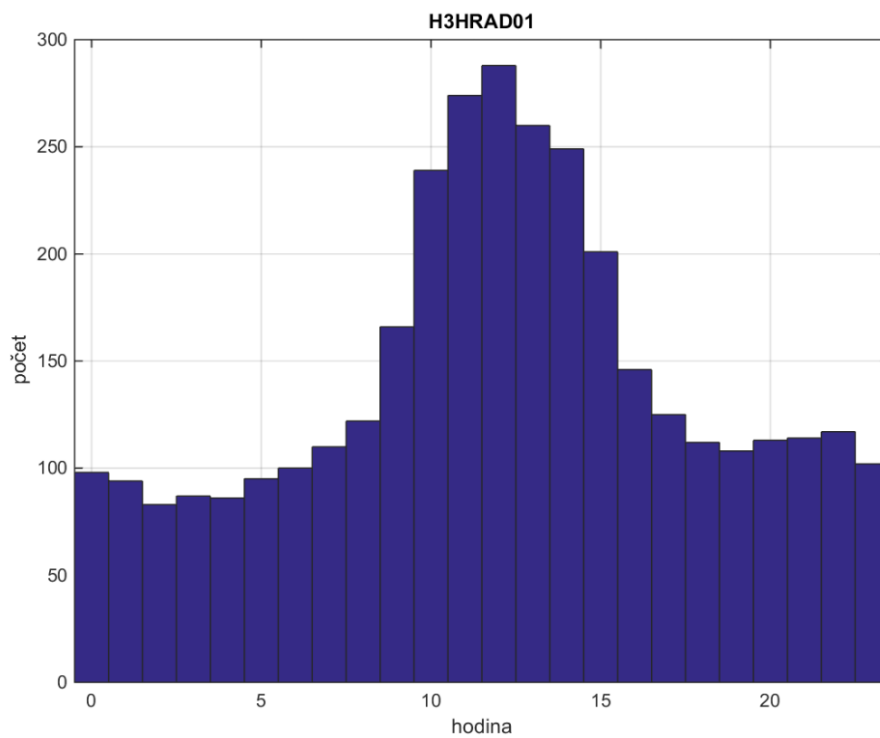
Obr. 2.4.10 Denní chod nejvyšších 10 % ze souboru dat v_1 na horské stanici Milešovka.



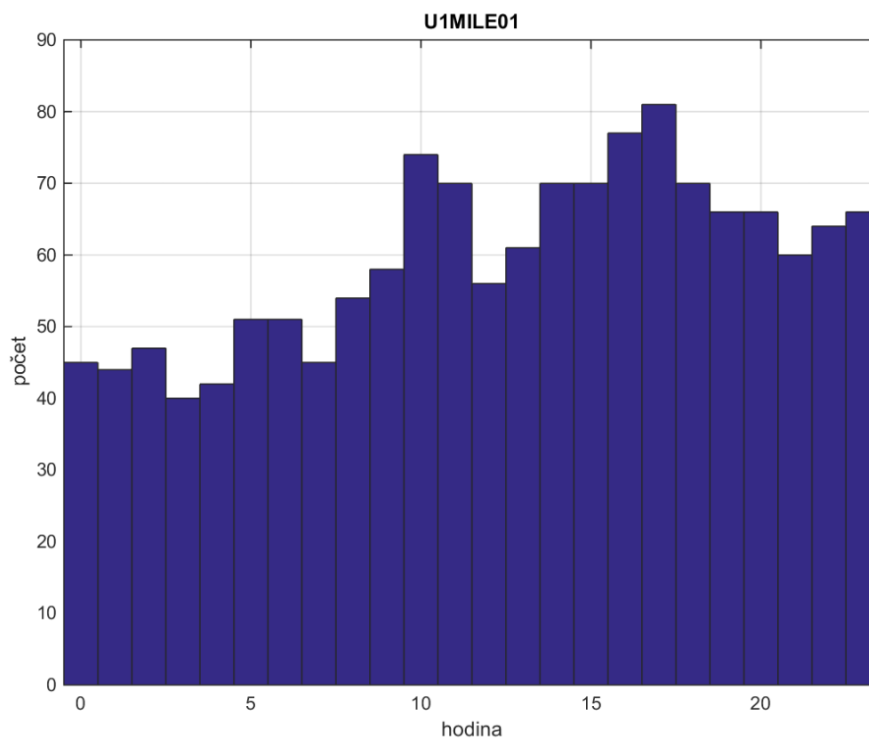
Obr. 2.4.11 Denní chod nejvyšších 10 % ze souboru dat v_2 na nížinné stanici Hradec Králové.



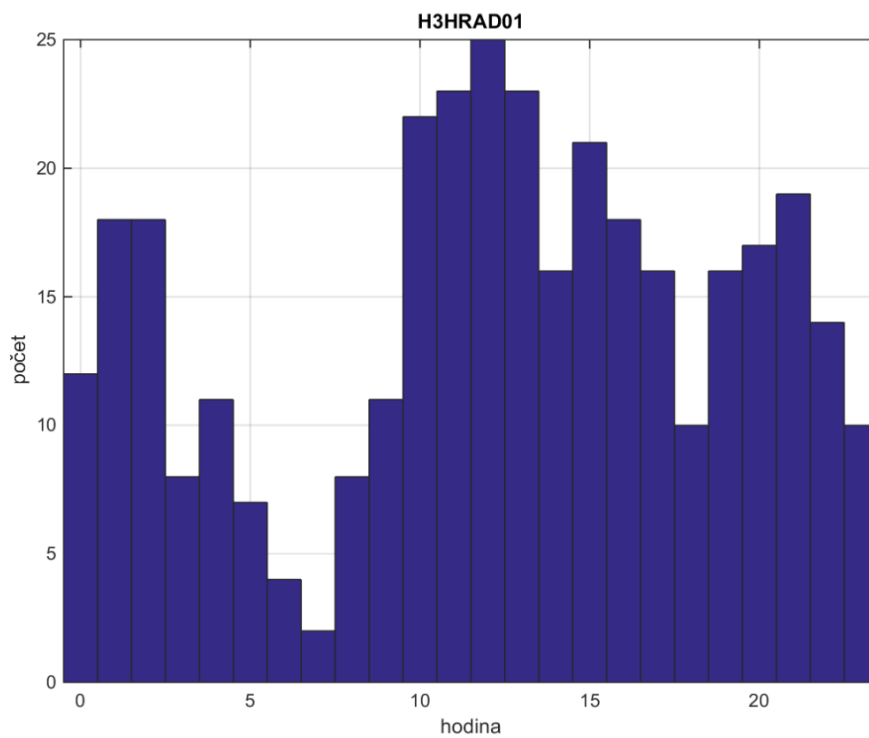
Obr. 2.4.12 Denní chod nejvyšších 10 % ze souboru dat v_2 na horské stanici Milešovka.



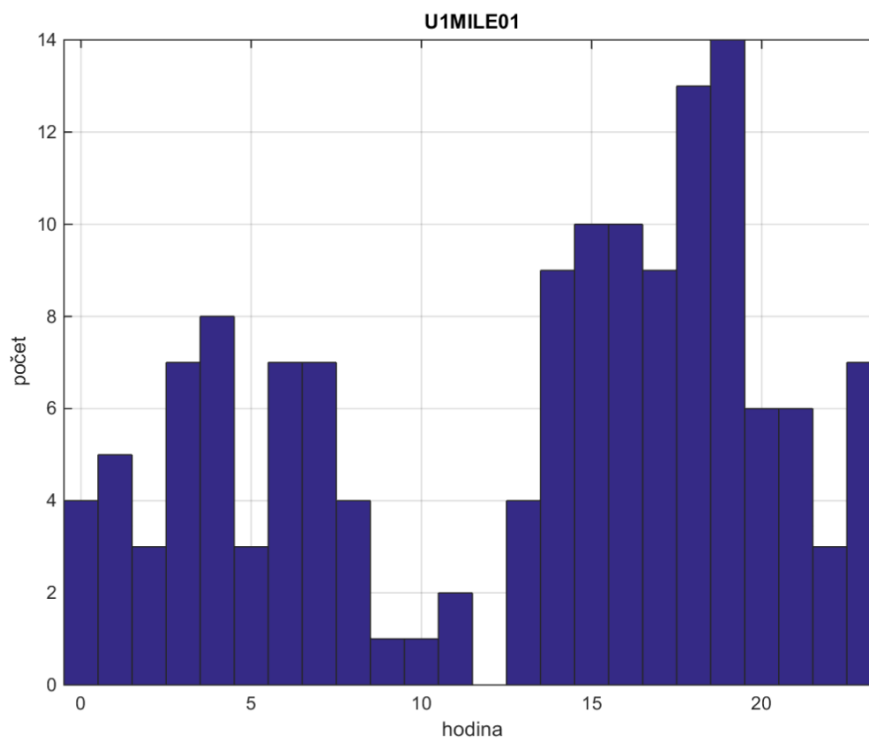
Obr. 2.4.13 Denní chod nejvyššího 1 % ze souboru dat v_2 na nížinné stanici Hradec Králové.



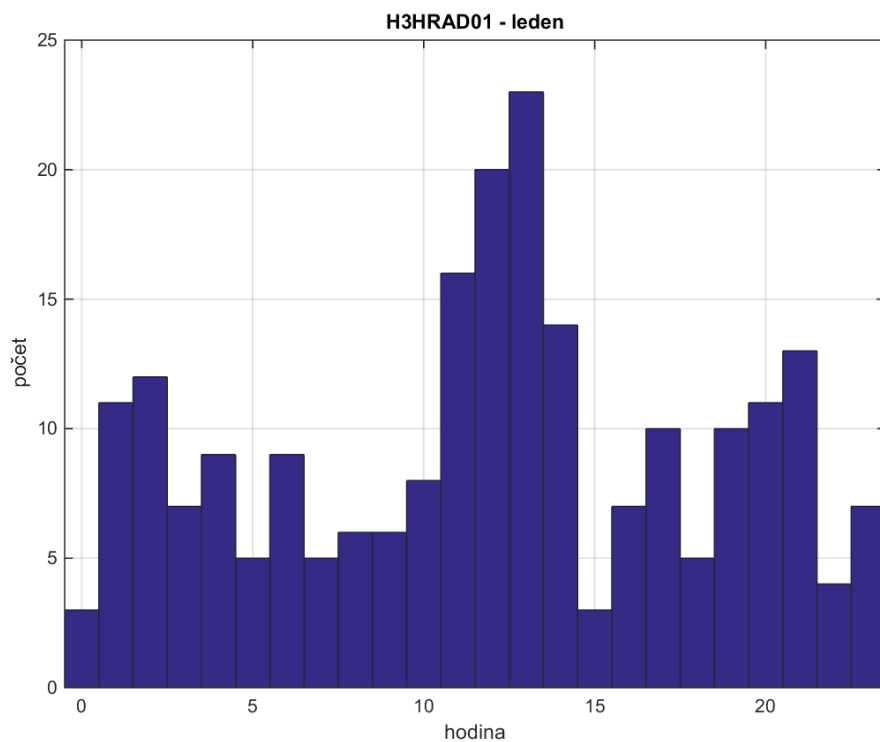
Obr. 2.4.14 Denní chod nejvyššího 1 % ze souboru dat v_2 na horské stanici Milešovka.



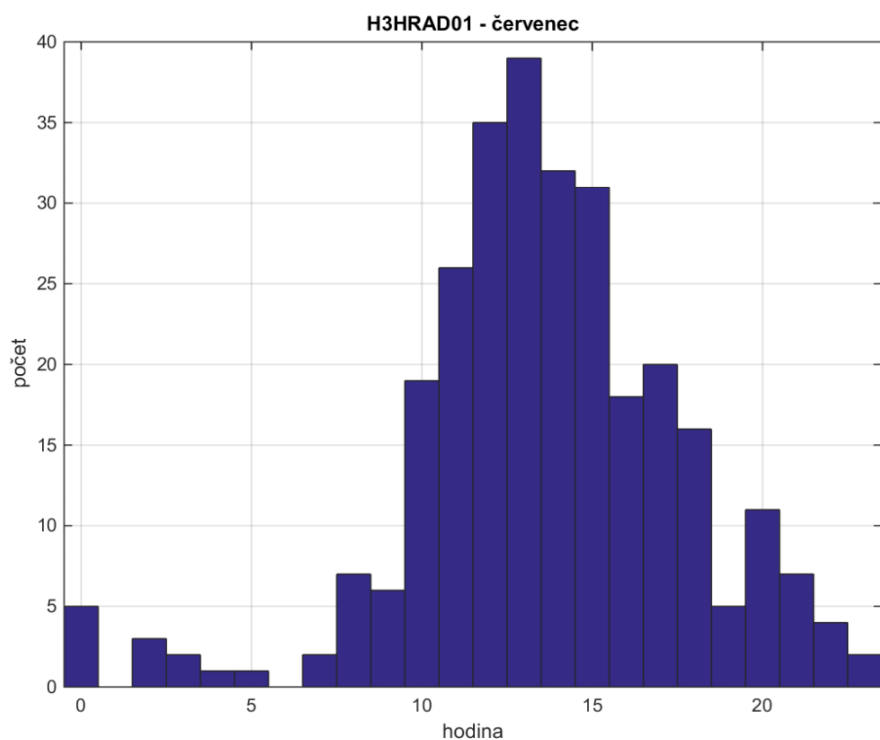
Obr. 2.4.15 Denní chod nejvyššího 0,1 % ze souboru dat v_2 na nížinné stanici Hradec Králové.



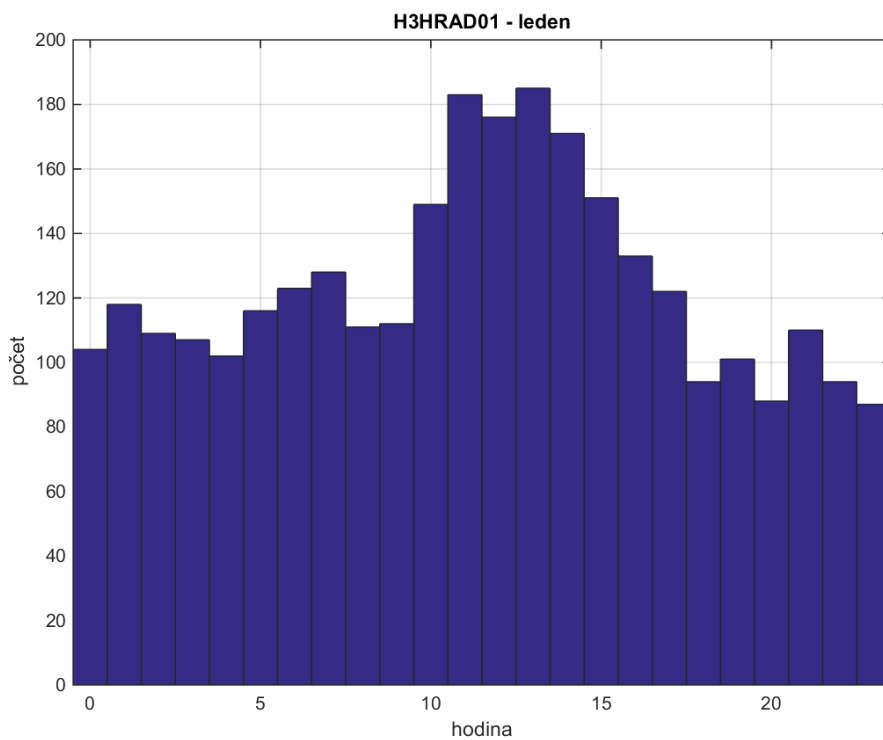
Obr. 2.4.16 Denní chod nejvyššího 0,1 % ze souboru dat v_2 na horské stanici Milešovka.



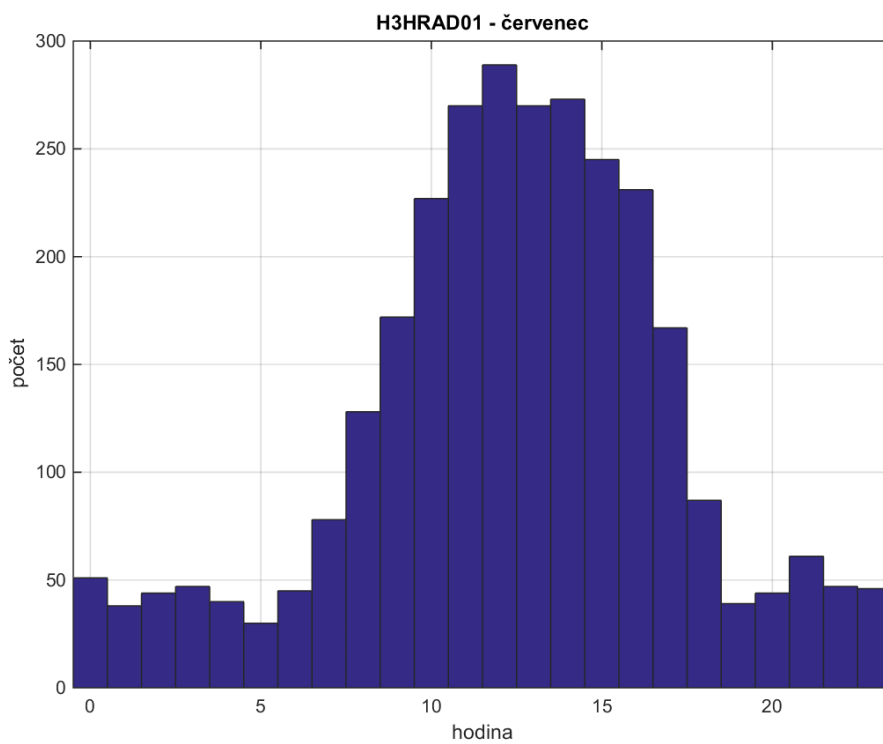
Obr. 2.4.17 Denní chod souboru dat v_1 v měsíci lednu na nížinné stanici Hradec Králové.



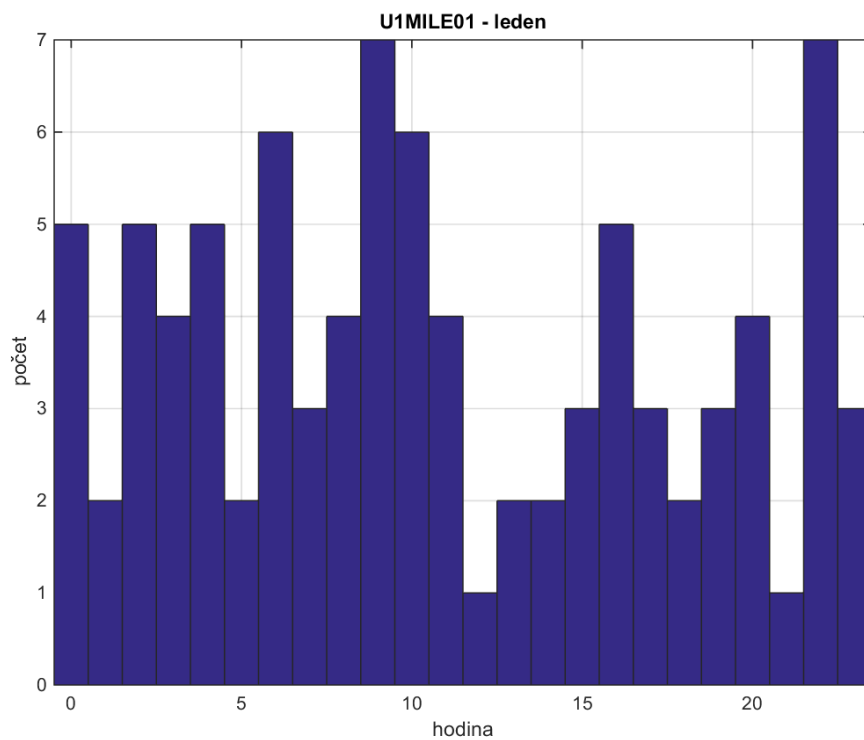
Obr. 2.4.18 Denní chod souboru dat v_1 v měsíci červenci na nížinné stanici Hradec Králové.



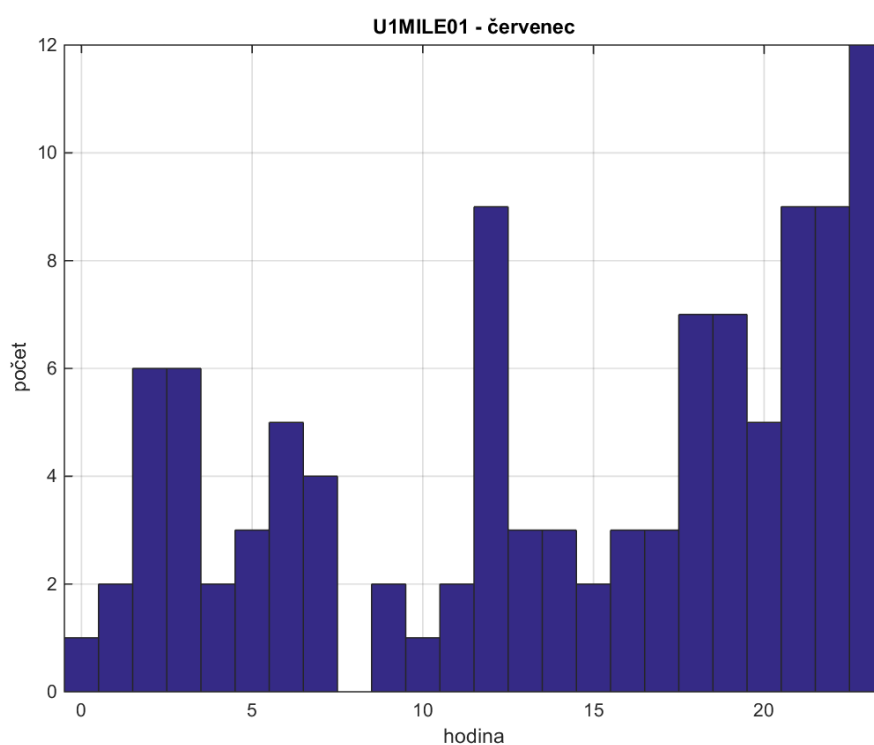
Obr. 2.4.19 Denní chod nejvyšších 10 % ze souboru dat v_2 v měsíci lednu na nížinné stanici Hradec Králové.



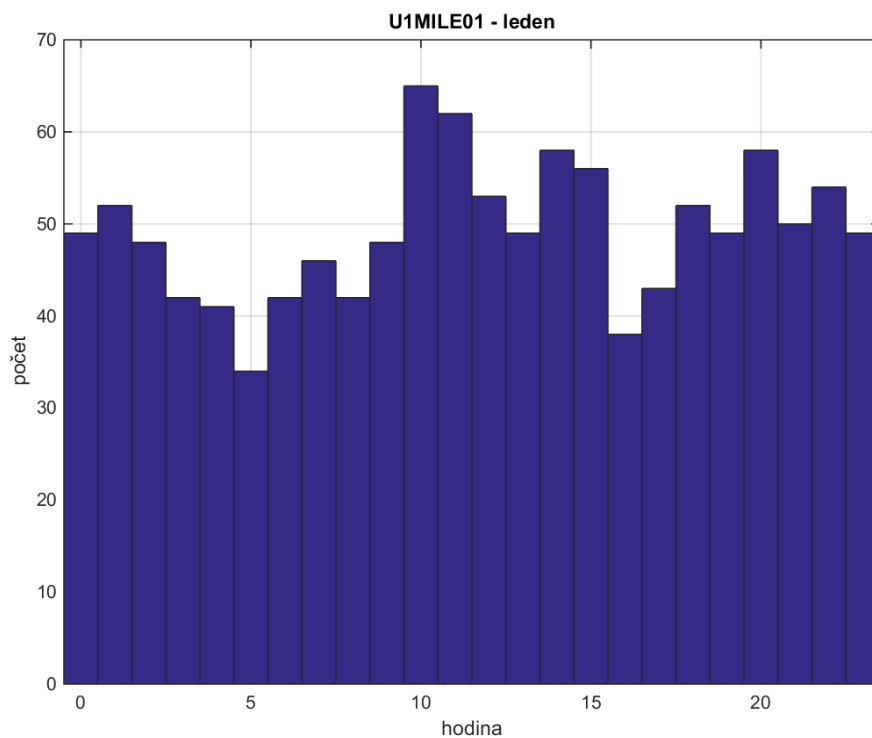
Obr. 2.4.20 Denní chod nejvyšších 10 % ze souboru dat v_2 v měsíci červenci na nížinné stanici Hradec Králové.



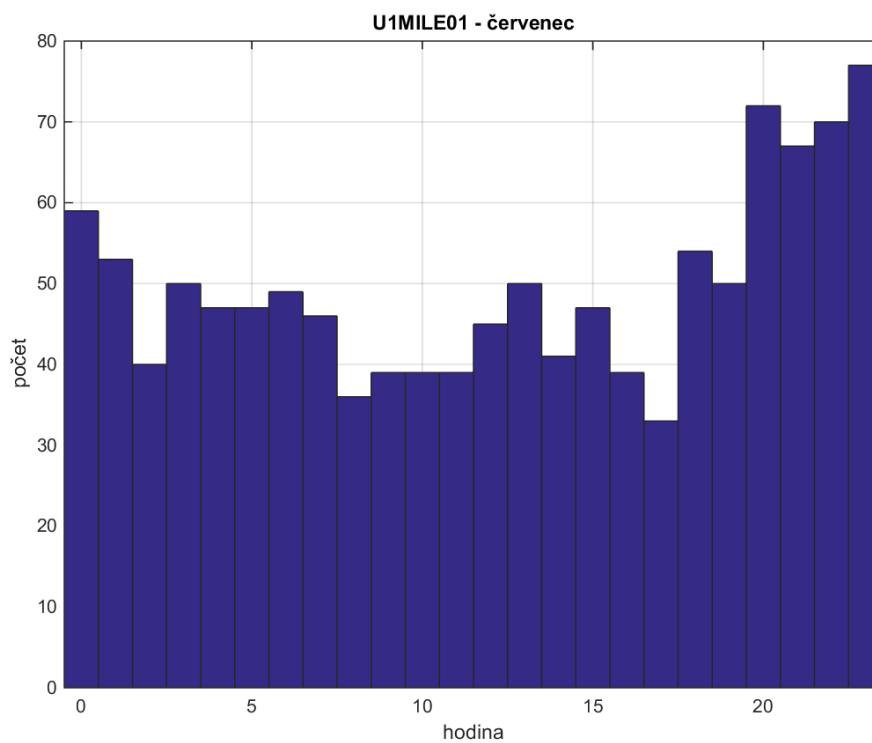
Obr. 2.4.21 Denní chod souboru dat v_I v měsíci lednu na horské stanici Milešovka.



Obr. 2.4.22 Denní chod souboru dat v_I v měsíci červenci na horské stanici Milešovka.



Obr. 2.4.23 Denní chod nejvyšších 10 % ze souboru dat v_2 v měsíci lednu na horské stanici Milešovka.



Obr. 2.4.24 Denní chod nejvyšších 10 % ze souboru dat v_2 v měsíci červenci na horské stanici Milešovka.

3 Analýza statistických metod

3.1 Definice a odhad doby návratu z empirické distribuční funkce

Při analýze extrémních hodnot se velmi často pracuje s termínem doba návratu. Určení doby návratu určité zpravidla vysoké hodnoty rychlosti větru je klíčovou bezpečnostní a ekonomickou otázkou v případě návrhu staveb a jiných inženýrských objektů. Při jejím stanovení se abstrahuje od doby, po kterou byla daná rychlost překročena, ale je uvažována pouze existence jevu překročení zadané rychlosti. Proto se obvykle pracuje pouze s posloupností vzájemně nezávislých maxim rychlosti větru. Tato vzájemně nezávislá maxima rychlosti větru budeme pro zjednodušení v následujícím textu většinou nazývat pouze rychlostí větru. Podmínka nezávislosti není ve skutečnosti zpravidla zcela splněna, ale uvažování závislosti by velmi komplikovalo teoretická odvození.

3.1.1 Definice doby návratu a souvislost s distribuční funkcí

Mějme n vzájemně nezávislých lokálních maxim rychlosti větru za dobu T_0 . Tyto hodnoty seřadíme podle velikosti a označíme je x_1, x_2, \dots, x_n . Zde nastává terminologický problém, protože v teoretické literatuře se obvykle hodnoty řadí od nejnižšího k nejvyššímu, zatímco v konkrétních případech je praktičtější řadit hodnoty obráceně, protože je mnohem názornější mluvit o 2. nejvyšší rychlosti větru v daném období než o 1354. nejnižší. Ve druhé variantě navíc pořadí nejvyšších hodnot silně závisí na metodě výběru případů. Abychom předešli nedorozumění, budeme za číslo označující pořadí přidávat znaménko “-”, znamenající vzestupné seřazení, nebo “+”, znamenající sestupné seřazení.

Doba návratu $T(x)$ pro rychlost větru x se definuje jako průměrná doba mezi dvěma po sobě následujícími stavy, kdy rychlost větru byla minimálně x . To je ekvivalentní době, za kterou nastane v průměru jedna událost s rychlostí větru vyšší nebo rovnou x . Za dobu T_0 připadající na jednu hodnotu rychlosti větru nastane průměrně $1 - F(x)$ takovýchto událostí, kde F je distribuční funkce rychlosti větru. Z toho plyne vzorec

$$T(x) = \frac{T_0}{(1-F(x))}. \quad (3.1.1)$$

Pro inverzní úlohu, kdy je cílem určit hodnotu rychlosti x s danou dobou opakování T , dostaneme úpravou vztahu (3.1.1)

$$x(T) = Q\left(1 - \frac{T_0}{T}\right), \quad (3.1.2)$$

kde Q je inverzní funkce k distribuční funkci F , tzv. kvantilová funkce (Wilks, 2011). Pro nevychýlený odhad distribuční funkce z naměřených hodnot platí vztah (Gumbel, 1958)

$$\tilde{F}(x_{i-}) = \frac{i}{n+1}, \quad (3.1.3)$$

kde n je počet naměřených hodnot. Vlnovkou nad F je zde značen výběrový odhad. Tento odhad je nevychýleným odhadem distribuční funkce. Nevychýlenost již ale neplatí pro nelineární funkce distribuční funkce jako je například doba návratu daná vztahem (3.1.1). Tato problematika je diskutována v článku Cook (2011).

V této práci budeme používat odhad (3.1.3), protože jde o nevychýlený odhad pravděpodobnosti překročení, který je jako doplněk hodnoty distribuční funkce do hodnoty 1 její lineární funkcí. Nevychýlené odhady doby opakování mohou vést k zavádějícím výsledkům, což ukážeme na jednoduchém příkladu v sekci 3.1.2 založeném na práci Cooka (2011). Této problematice se věnuje také sekce 3.3.6.

3.1.2 Střední hodnota doby návratu naměřených hodnot

Střední hodnota funkce f distribuční funkce F pro i -tou nejnižší hodnotu x_i z náhodného výběru o velikosti n se spočítá vztahem (Cook, 2011)

$$E(f(F, i, n)) = \int_0^1 f(F) \varrho(F, i, n) dF, \quad (3.1.4)$$

kde E je označení pro střední hodnotu a $\varrho(F, i, n)$ je hustota pravděpodobnosti odhadu distribuční funkce hodnoty x_i . Dosazením $f(F) = F$ a úpravou dostaneme vztah (3.1.3) pro nevychýlený odhad distribuční funkce.

Pokud známe pouze jednu naměřenou hodnotu, a interval měření je pro jednoduchost jednotkový, pak pomocí (3.1.4) vypočteme střední hodnotu doby návratu podle vztahu:

$$f(F) = T(f) = \frac{1}{1-T} \quad (3.1.5)$$

a

$$\varrho(F, i, n) = \varrho(F, 1, 1) = 1; F \in \langle 0, 1 \rangle, \quad (3.1.6)$$

protože distribuční funkce jediného náhodného měření má rovnoměrné rozdělení. Dosazením a jednoduchým výpočtem dostaneme výsledek

$$E(T, 1, 1) = \int_0^1 \frac{dF}{1-F} = \infty, \quad (3.1.7)$$

a tedy průměrná hodnota doby návratu hodnoty dané jedním měřením je nekonečná. Pro srovnání nevychýlený odhad pravděpodobnosti překročení podle vzorce (3.1.3) dává hodnotu $\frac{1}{2}$, které odpovídá doba návratu 2.

3.1.3 Nevýhody použití empirické distribuční funkce pro výpočet extrémních hodnot větru

Výhodou použití empirické distribuční funkce pro výpočet extrémních hodnot je, že neklade žádné dodatečné předpoklady na jejich statistickém rozdělení. To ale způsobuje velkou výběrovou chybu této metody. Tuto chybu lze redukovat použitím jiných metod, které zavádějí a využívají určité předpoklady o rozdělení extrémních hodnot (Harris, 1996). Analýza těchto metod je provedena v dalších podkapitolách kapitoly 3.

Tyto metody obvykle řeší i další nevýhodu použití empirické distribuční funkce podle vzorce (3.1.3), kterou představuje skutečnost, že je definována pouze pro diskrétní hodnoty odpovídající naměřeným datům. Použitím vzorců (3.1.1) a (3.1.3) zjistíme, že v případě padesátileté řady ročních maxim odpovídají dvěma nejvyšším naměřeným hodnotám doby návratu, které odpovídají nevychýleným odhadům distribuční funkce 51 a 25,5 let.

3.2 Teorie extrémních hodnot

3.2.1 Úvod

Teorie extrémních hodnot představuje rozsáhlou problematiku. Tato podkapitola představuje její shrnutí a vychází z publikací Beirlant et al. (2004), Reiss et Thomass (2001), Coles et al. (2001) a Kotz et Nadarajah (2000). Aplikovatelností této teorie na určení extrémního větrného klimatu se věnuje podkapitola 3.3.

V této práci používáme formulaci teorie extrémních hodnot pro maxima. Formulace pro minima je analogická.

Základem této teorie je řešení následujících dvou otázek:

- (i) jaké jsou možné tvary limitního rozdělení extrémních hodnot,
- (ii) jaké jsou předpoklady, za kterých extrémní hodnoty k těmto limitním rozdělením konvergují?

První otázku řeší první a druhý teorém extrémních hodnot, druhá otázka je problém domén atrakce.

V teorii extrémních hodnot se pracuje se dvěma soubory dat. Prvním jsou původní data, jejichž hodnoty označme jako X_i , kde index i je pořadí dané hodnoty, a předpokládá se, že X_i jsou vzájemně nezávislé náhodné hodnoty pocházejícími ze stejného rozdělení. Druhým souborem jsou extrémní data, která jsou určitým výběrem z původních dat. Tento výběr se obvykle provádí dvěma způsoby, které vedou k různým typům (souborům) extrémních dat. Prvním typem jsou bloková maxima (BM) obvykle označovaná jako M_n , definovaná jako maxima ze vzájemně výlučných souborů původních dat o velikost n prvků. Druhým typem jsou hodnoty překračující zadanou mez. Tento výběr se anglicky nazývá Peak-over-Threshold (POT).

3.2.2 První a druhý teorém extrémních hodnot

První teorém extrémních hodnot (Coles et al., 2001) zní: *Necht' M_n jsou bloková maxima původních dat. Pokud existují posloupnosti náhodných čísel $a_n > 0$, b_n takové, že*

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq z\right) = G(z), \quad (3.2.1)$$

kde P značí pravděpodobnost a G je nedegenerovaná distribuční funkce, pak se tato distribuční funkce dá vyjádřit ve tvaru

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \quad (3.2.2)$$

která je definovaná pro hodnoty z takové, že $1 + \xi(z-\mu)/\sigma > 0$, kde $\sigma > 0$ a μ a ξ mohou být libovolná reálná čísla. Toto rozdělení se označuje jako tzv. zobecněné rozdělení extrémních hodnot neboli GEV (General Extreme Value) rozdělení. Pro případ $\xi = 0$ je nutno použít limitní vztahu (3.2.2) ve tvaru

$$G(z) = \exp\left[-\exp\left\{-\left(\frac{z-\mu}{\sigma}\right)\right\}\right]. \quad (3.2.3)$$

Rozdělení s touto distribuční funkcí se nazývá Gumbelovo rozdělení. GEV rozdělení s $\xi > 0$ se nazývá Fréchetovo rozdělení, pokud je $\xi < 0$, pak hovoříme o Weibullově rozdělení. Parametr ξ se nazývá parametr tvaru. V dalším textu se pod konvergencí

blokových maxim ke GEV (případně Gumbelovu) rozdělení rozumí konvergence ve smyslu vzorců (3.2.1) a (3.2.2).

Druhý teorém extrémních hodnot zhruba tvrdí, že rozdělení extrémních hodnot vybraných metodou POT konverguje se zvyšující se hodnotou meze k tzv. Generalizovanému Pareto rozdělení (GPD – Generalized Pareto Distribution) právě tehdy, když pro bloková maxima vybraná ze stejných původních hodnot platí konvergence ke GEV rozdělení. Tento teorém zároveň udává vztah mezi hodnotami parametrů těchto limitních GEV a GPD rozdělení. Přesné znění tohoto teorému zde není uvedeno vzhledem k faktu, že se dále v této práci pracuje s blokovými maximy. Lze jej nalézt např. v práci Coles et al. (2001).

3.2.3 Problém domén atrakce

Tento problém se zabývá odpovědí na otázku, jaké jsou požadavky na rozdělení původních hodnot, aby bloková maxima těchto hodnot konvergovala podle 1. teorému extrémních hodnot. Navíc řeší i souvislost parametru ξ limitního rozdělení těchto blokových maxim s rozdělením původních hodnot. Pro další diskuzi v následující podkapitole 3.3 je podstatné, že pro konvergenci rozdělení blokových maxim ke GEV rozdělení je nutnou podmínkou, aby se pravý chvost rozdělení limitně choval určitým poměrně striktně daným způsobem. Přesná formulace je obsahem přílohy A.

3.3 Použití GEV rozdělení k odhadu extrémních hodnot

Použití GEV rozdělení k aproximaci rozdělení extrémních hodnot se často teoreticky zdůvodňuje teorií extrémních hodnot popsanou v kapitole 3.2, ale tato teorie je mnohem komplikovanější, než tvrzení, že bloková maxima konvergují ke GEV rozdělení. V této kapitole budou na pozadí teorie extrémních hodnot diskutovány možné problémy vyplývající z použití GEV rozdělení k aproximaci extrémních hodnot.

3.3.1 Koeficienty a_n a b_n v prvním teorému extrémních hodnot

V prvním teorému extrémních hodnot se vyskytují koeficienty a_n a b_n . Pro praktické aplikace se ale využívá faktu, že ze vzorců (3.2.2) a (3.2.3) plyne pro distribuční funkci GEV rozdělení vztah

$$G(z, \mu, \sigma, \xi) = G(az + b, \mu^*, \sigma^*, \xi); a > 0. \quad (3.3.1)$$

Pokud tedy z má GEV rozdělení, potom lineární funkce z bude mít také GEV rozdělení s jinými parametry μ a σ , ale stejným parametrem ξ . Z toho plyne, že pro velké délky bloků n budou za předpokladu konvergence podle 1. teorému extrémních hodnot samotné distribuční funkce blokových maxim přibližně dány GEV rozdělením a koeficienty a_n, b_n vyskytující se v 1. teorému extrémních hodnot je možné při aproximaci těchto blokových maxim GEV rozdělením ignorovat (Coles et al., 2001).

3.3.2 Vlastnosti umocněné distribuční funkce GEV rozdělení a jejich důsledky

Tyto vlastnosti vyplývají z důkazu 1. teorému extrémních hodnot. Odvození 1. teorému extrémních hodnot je popsáno v Kotz et al. (2000) a zde je ve zkrácené podobě reprodukováno. Základní myšlenka vychází ze vzorce

$$G_n(z) = (F(z))^n, \quad (3.3.2)$$

kde je F je distribuční funkce původních hodnot a G_n distribuční funkce blokových maxim vybraných vždy z n těchto původních hodnot. Tento vzorec vyplývá ze vzorce pro pravděpodobnost průniku nezávislých náhodných jevů. Základní myšlenkou odvození 1. teorému extrémních hodnot je otázka, k jaké distribuční funkci bude $G_n(z)$ konvergovat pro $n \rightarrow \infty$. Zřejmě $\lim_{n \rightarrow 0} G_n(z) = 0$ pro $F(z) < 1$ a $\lim_{n \rightarrow \infty} G_n(z) = 1$ pro $F(z) = 1$, což je triviální výsledek. Důkaz 1. teorému extrémních hodnot spočívá v důkazu, že distribuční funkce veličiny $(M_n - b_n)/a_n$, $a_n > 0$, kde a_n a b_n závisejí na n , ale ne na z , může konvergovat k nedegenerované distribuční funkci G pouze tehdy, pokud pro všechna přirozená čísla n, N platí vzorec

$$G^N(a_n z + b_n) = G(a_{n*N} z + b_{n*N}), \quad (3.3.3)$$

a že obecným řešením této rovnice je distribuční funkce GEV rozdělení. To znamená, že umocněním distribuční funkce GEV rozdělení dostaneme stejnou funkci pouze s lineárně transformovanou nezávisle proměnnou.

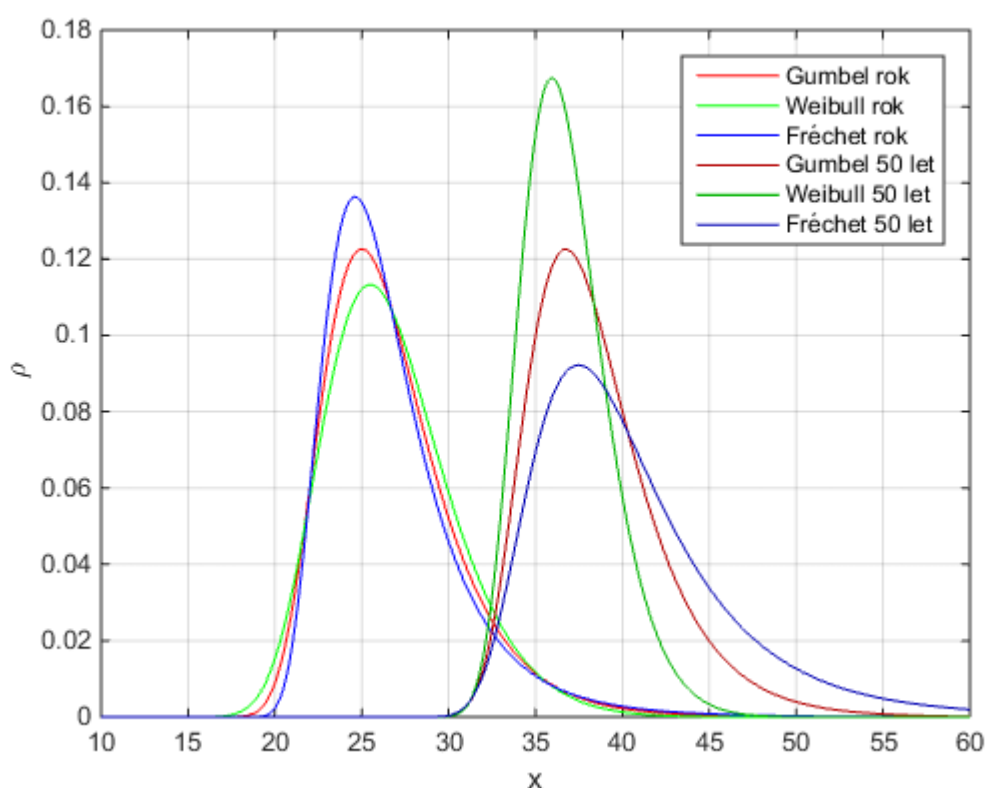
Z těchto úvah vyplývá, že pokud předpokládáme, že distribuční funkce ročních maxim nějaké veličiny má GEV rozdělení, pak bude mít distribuční funkce padesátiletých maxim této veličiny také GEV rozdělení se stejným parametrem ξ . To je ilustrováno na obr. 3.2.1, na kterém jsou srovnána roční a padesátiletá maxima veličin se třemi různými GEV rozděleními ročních maxim. Parametry těchto rozdělení byly vybrány tak, že hodnoty parametru ξ jsou 0, -0,1 a +0,1, a hodnoty ostatních dvou parametrů jsou zvoleny tak, aby rozdělení ročních hodnot byly co nejbližší. Přesná definice této blízkosti je v sekci 3.8.1. Pro Gumbelovo rozdělení bude rozdělení maxim za n let pouze

posunutým rozdělením ročních maxim, přičemž je snadné spočítat velikost posunu, který bude $\sigma * \ln(n)$, protože pro distribuční funkci G Gumbelova rozdělení platí

$$G^n(x, \mu, \sigma) = \left(\exp \left(- \exp \left(- \frac{x - \mu}{\sigma} \right) \right) \right)^n = \exp \left(- n * \exp \left(- \frac{x - \mu}{\sigma} \right) \right) =$$

$$= \exp \left(- \exp \left(- \frac{x - (\mu + \sigma * \ln(n))}{\sigma} \right) \right) = G(x, \mu + \sigma * \ln(n), \sigma). \quad (3.3.4).$$

V případě GEV rozdělení s nenulovým parametrem ξ dojde při umocnění distribuční funkce kromě posunu také ke změně parametru σ , jak je ukázáno v Beirlant et al. (2004) a ilustrováno na obr. 3.2.1.



Obr 3.2.1: Srovnání rozdělání hustot pravděpodobností ρ rozdělání maxim určité veličiny za 1 rok a 50 let za předpokladu, že její roční maxima jsou dána GEV rozděláním.

3.3.3 Diskuze k prvnímu teorému extrémních hodnot

Problém, který řeší 1. teorém extrémních hodnot, lze definovat jako speciální případ širšího problému nalezení možných tvarů nedegenerované distribuční funkce G , pro kterou platí

$$f(x, n) = G^{-1}((G^n(x))^{1/n}). \quad (3.3.5)$$

kde $f(x, n)$ má určitý tvar. V případě 1. teorému extrémních hodnot je

$$f(x, n) = a(n)x + b(n), \quad (3.3.6)$$

a jeho řešením je distribuční funkce GEV rozdělení. Z předchozí sekce 3.3.2 plyne, že pro

$$f(x, n) = x \quad (3.3.7)$$

nemá problém řešení, a jak je dokázáno např. v Beirlant et al. (2004), pro

$$f(x, n) = x + b(n), \quad (3.3.8)$$

je řešením Gumbelovo rozdělení.

Při podmínce (3.3.8) by tedy „obecným“ rozdělením extrémních hodnot nebylo GEV, ale Gumbelovo rozdělení. Jiné parametrické tvary funkce $f(x, n)$ vedou k jiným „obecným“ rozdělením extrémních hodnot. Z této úvahy plyne, že výsledné GEV rozdělení v 1. teorému extrémních hodnot je důsledek použití konkrétní podmínky (3.3.6).

3.3.4 Srovnání teorie extrémních hodnot a centrální limitní věty

V této sekci uvažujeme teorii extrémních hodnot ve formulaci pro bloková maxima. Tato teorie bývá často přirovnávána k známější centrální limitní větě říkající, že rozdělení součtu náhodných hodnot konverguje k normálnímu rozdělení nezávisle na původním rozdělení. Stejně tak se často tvrdí, že podle teorie extrémních hodnot konverguje rozdělení maxim náhodných hodnot ke GEV rozdělení nezávisle na původním rozdělení. Ve skutečnosti je to podstatně složitější.

Existuje mnoho tvarů centrální limitní věty podrobně popsanych a dokázaných např. v Chung (2000). Zde použijeme Lévyho-Lindebergův tvar, který je asi nejznámější.

Nechť náhodná veličina X je součtem n vzájemně nezávislých náhodných proměnných X_1, X_2, \dots, X_n se shodným rozdělením libovolného typu, s konečnou střední hodnotou $E(X_i) = \mu$ a konečným nenulovým rozptylem $D(X_i) = \sigma^2$. Pak pro normovanou náhodnou veličinu

$$U = \frac{X - n\mu}{\sqrt{n\sigma^2}}, \quad (3.3.9)$$

platí vztah

$$\lim_{n \rightarrow \infty} P(U < u) = \phi(u) \quad (3.3.10)$$

pro $-\infty < u < \infty$, kde P značí pravděpodobnost a $\phi(u)$ je distribuční funkce normovaného normálního rozdělení. Veličina U má tedy asymptoticky normované normální rozdělení.

Tvrzení teorie extrémních hodnot má na první pohled podobnou strukturu, a to že distribuční funkce náhodné veličiny definované určitým způsobem za jistých podmínek

konverguje k určitému tvaru. Zásadní rozdíl mezi těmito podmínkami je v definici těchto „jistých podmínek“. V uvedeném tvaru centrální limitní věty jsou dosti obecné. Stačí, aby původní rozdělení nemělo příliš těžké chvosty tak, aby jeho střední hodnota a rozptyl byly konečné. Naopak v teorii extrémních hodnot se pro konvergenci vyžaduje, aby bylo možné pravý chvost rozdělení původních hodnot limitně aproximovat funkcí určitého tvaru (sekce 3.2.3 a příloha A), což je podstatně silnější a v případě naměřených dat těžko ověřitelný předpoklad.

3.3.5 Odůvodnění a limity aproximace extrémních dat

Nemožnost ověření podmínek konvergence blokových maxim ke GEV rozdělení vede k otázce, zda má vůbec smysl pro účely aproximace vybírat ze všech naměřených hodnot hodnoty extrémní. Tuto konvergenci lze snadno posoudit, pokud předpokládáme, že rozdělení původních dat má určitý tvar. Uvažujme např., že je dáno Weibullovým rozdělením, kterým se často aproximuje rozdělení rychlosti větru (Troen et Petersen, 1993). (Weibullovo rozdělení se pro účely rozdělení rychlostí větru používá v jiném tvaru než v podkapitole 3.2, který vznikne otočením a posunutím jeho hustoty pravděpodobnosti tak, aby jeho levý krajní bod byl 0 m/s.) Potom jeho bloková maxima budou konvergovat ke Gumbelovu rozdělení (Beirlant et al., 2004). Ze vzorce (3.3.2) ovšem plyne, že rozdělení blokových maxim je jednoznačně určeno rozdělením původních hodnot a naopak. Určení parametrů limitního Gumbelova rozdělení BM a parametrů Weibullova rozdělení původních hodnot tedy představuje ekvivalentní úlohu. Určení parametrů rozdělení původních dat bude nejpřesnější, pokud se k němu využijí všechna naměřená data. Tato úvaha ukazuje, že výběr extrémních dat a jejich následná aproximace je v případě znalosti tvaru rozdělení původních hodnot kontraproduktivní.

Přesto je tento postup v případě analýzy extrémních rychlostí větru obvyklý (Palutikof et al., 1999). Odůvodnit se dá v případě, kdy předpokládáme, že rozdělení původních hodnot neznáme, ale že události s obdobnou rychlostí větru vykazují podobné statistické chování a má smysl je tedy aproximovat samostatně. To v případě analýzy extrémních rychlostí větru dává intuitivní smysl v případě, kdy si budou jednotlivé případy s vysokou rychlostí větru fyzikálně podobné, což je zřejmě případ přechodů zimních front nebo konvektivních bouří. Z této úvahy dále vyplývá, že touto metodou nelze očekávat dobrou předpověď extrémních rychlostí větru způsobených

meteorologickými jevy, které jsou tak vzácné, že se v naměřených datech nevyskytly, např. tornády. Extrémní rychlosti větru při těchto jevech je nutno analyzovat odděleně.

3.3.6 Věta o průměrné době návratu určené proložením GEV rozdělení náhodnými hodnotami

V příloze B je ukázáno, že průměrná doba návratu určená proložením GEV rozdělení náhodnými hodnotami je za velmi obecných podmínek nekonečná. Průměrná doba návratu T náhodné veličiny s distribuční funkcí F a dobou mezi jednotlivými měřeními T_0 je dána vztahem (podkapitola 3.1)

$$T(x) = \frac{T_0}{1-F(x)}. \quad (3.3.11)$$

Z tohoto vztahu platí, že doba opakování je nekonečná pro takové hodnoty x , že $F(x) = 1$. Pokud je distribuční funkce veličiny x dána GEV rozdělením, čili platí

$$F(x) = \exp\left\{-\left[1 + \xi \left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \quad (3.3.12)$$

(případ $\xi = 0$ je dán limitou tohoto rozdělení) je $F(x) = 1$ a tedy $T(x) = \infty$ pro $x \geq \mu - \sigma/\xi$ a $\xi < 0$. Tedy pro jednu třídu GEV rozdělení (Weibullovo rozdělení) je doba opakování nekonečná pro hodnoty x vyšší než určitá hranice. Následující věta ukazuje, že tento fakt zásadně ovlivňuje průměrnou dobu návratu počítanou z náhodného výběru hodnot proložením GEV rozdělení těmito daty.

Věta: Necht' náhodná veličina x má hustotu pravděpodobnosti $\rho(x)$ takovou, že

$$\rho(x) > 0 \quad \forall x \in (x_{inf}, x_{sup}), \quad (3.3.13)$$

kde x_{inf} , x_{sup} jsou z R^ , tedy buď reálná čísla, nebo $\pm\infty$, a $x_{inf} < x_{sup}$ a současně je $\rho(x)$ na intervalu (x_{inf}, x_{sup}) spojitá. Dále necht' n je přirozené číslo a G_n je metoda přiřazující distribuční funkci GEV rozdělení naměřeným datům sestávající se z funkcí $R^n \rightarrow R$ μ_n , σ_n a ξ_n . Současně předpokládejme, že existuje taková n -tice reálných čísel $X_1, X_2, \dots, X_n = \mathbf{X}$, která nejsou všechna stejná taková, že $\xi_n(\mathbf{X}) < 0$ a μ_n , σ_n a ξ_n jsou v bodě \mathbf{X} spojitě, G_n nezávisí na pořadí proměnných a škáluje a posouvá se stejně jako data, tedy platí:*

$$G_n(x, \mathbf{X}) = G_n(ax + b, a\mathbf{X} + b) \quad \forall a, b \in R, a \neq 0, \quad (3.3.14)$$

kde x je hodnota, pro kterou se distribuční funkce počítá, a \mathbf{X} je n -tice hodnot, kterými se GEV rozdělení prokládá. Dobu opakování určíme podle vztahu (3.3.11), ve kterém zvolíme $T_0 = 1$, přičemž definujeme podle obvyklé praxe dobu opakování jako $+\infty$, pokud je hodnota distribuční funkce rovna 1. Potom průměrná doba opakování určená z n náhodných

hodnot s hustotou pravděpodobnosti $\rho(x)$ metodou proložení GEV rozdělení G_n je nekonečná pro všechna $y > x_{inf}$.

Důkaz této věty je příloze B.

Pokud předpokládáme, že rozdělení, ze kterého jsou generovány náhodné hodnoty, je skutečně GEV, tak potom podle této věty platí, že průměrná doba opakování je nekonečná pro všechna reálná čísla v případě, že toto rozdělení je Weibullovo nebo Gumbelovo. Pokud se jedná o Fréchetovo rozdělení s parametry μ , σ a ξ , je pak nekonečná pro všechny hodnoty vyšší než $\mu - \sigma/\xi$. Tento výsledek je ale pro praktické aplikace nepoužitelný. Proto považujeme za smysluplnější usilovat o nevyčýlený odhad například pravděpodobnosti překročení v daném období než doby opakování. Obdobná věta platí i pro GPD rozdělení, protože má pro záporný parametr tvaru také omezený pravý chvost, proto problém nejde obejít použitím dat vybraných metodou Peak-over-Threshold a jejich aproximací GPD rozdělením.

3.4 Metody výběru extrémních hodnot z naměřených dat

Obvyklý způsob analýzy extrémních hodnot spočívá v aproximaci naměřených extrémních hodnot teoretickým rozdělením určitého tvaru. Tato kapitola se zabývá otázkou, jakým způsobem tato extrémní data vybrat.

3.4.1 Porovnání metody blokových maxim a metody Peak-over-Threshold

Toto srovnání vychází z diskuze uvedené v článku Palutikov et al. (1999). První teorém extrémních hodnot uvažuje tzv. bloková maxima, tedy maxima z pevně daného počtu po sobě následujících naměřených hodnot, zatímco druhý teorém extrémních hodnot uvažuje hodnoty vybrané metodou Peak-over-Threshold (POT), tedy všechny vzájemně nezávislé hodnoty přesahující určitou hranici. Jak bude ukázáno v podkapitole 3.5, je použití těchto dvou přístupů spojeno s odlišnými hodnotami doby návratu, mezi kterými je systematický rozdíl, jenž se pro velké doby opakování přibližně rovná jedné polovině délky používaných bloků v metodě BM.

Za výhodu metody POT se považuje efektivnější výběr extrémních událostí. Průměrná hodnota výběru n nejvyšších hodnot vybraných metodou POT je obvykle vyšší

než průměrná hodnota stejného počtu hodnot vybraných metodou blokových maxim. V případě analýzy extrémních rychlostí větru je často tento rozdíl významný.

Výhodou metody blokových maxim oproti metodě Peak-over-Threshold je menší počet parametrů, které je nutno zvolit, a tedy jednodušší aplikace, což je důvod, proč se při analýze extrémních rychlostí větru používá častěji. Při aplikaci metody blokových maxim se prakticky vždy používá délka bloku 1 rok. Naproti tomu v případě metody Peak-over-Threshold neexistuje všeobecně přijímaná metoda volby hranice (Palutikof et al., 1999). Navíc je nutno zajistit, aby vybrané hodnoty byly nezávislé. To se zpravidla řeší tak, že se vyžaduje minimální časový rozdíl mezi dvěma extrémními hodnotami. V evropských podmínkách se obvykle používá interval 48 hodin (Palutikof et al., 1999).

3.4.2 Metoda individuálních bouří

Alternativní přístup pro výběr extrémních událostí představuje tzv. metoda individuálních bouří zavedená v článku Cooka (1982). Je podrobně diskutována v článku Harrise (1999), kde jsou navrženy i některé její úpravy. Tato metoda ale vyžaduje navíc subjektivní volbu několika parametrů. Nejprve se řada rychlostí větru zhladí, aby se odstranily krátkodobé oscilace. Toto shlazení se použije jen pro oddělení jednotlivých událostí a hodnoty maximálních rychlostí se odečítají z původní neshlazené řady. Potom se zvolí hranice rychlosti větru, která odděluje jednotlivé bouře. Cook původně použil hranici 5 m/s, což vedlo k počtu cca 100 bouří za rok, ale pro odhad rychlosti větru s dobou návratu 50 let doporučil tuto hranici zvýšit tak, aby bylo identifikováno v průměru 10 bouří za rok. Maximální dosažené rychlosti v těchto jednotlivých bouřích se pak aproximují Gumbelovým rozdělení. Cook doporučuje toto rozdělení z důvodu rychlejší konvergence prokládat druhými mocninami vybraných rychlostí.

3.4.3 Separace extrémních rychlostí větru

Analýza extrémních rychlostí větru se obvykle provádí pro všechna data najednou jako např. v Abild et al. (1992) a Perrin et al. (2006). Někdy se ale naměřená data rozdělují do různých kategorií a analýza se provádí v kategoriích nezávisle na sobě. Motivací pro toto dělení je předpoklad, že charakteristiky dat z jedné kategorie budou navzájem podobné a aproximace jejich rozdělení pomocí vhodného teoretického rozdělení bude přesnější než aproximace všech dat najednou jedním rozdělením. Rozdělení na kategorie extrémních událostí může být užitečné i z dalších důvodů, které

budou dále zmíněny. Nevýhodou dělení je skutečnost, že se k aproximaci extrémního rozdělení používá více hodnot, které jsou v průměru nižší, a tedy méně extrémní, než v případě analýzy všech hodnot najednou.

Pokud při přiřazování do kategorií budou vždy „celé události“ s extrémním větrem přiřazeny do jedné kategorie (tj. nedojde např. k tomu, že přechod fronty s extrémními rychlostmi větru trvající několik hodin bude rozdělen do dvou kategorií), budou hodnoty vybrané z různých kategorií vzájemně nezávislé. Distribuční funkce extrémních rychlostí větru nezávisle na kategorii bude mít tvar

$$F(v) = \prod_{i=1}^n F_i(v), \quad (3.4.1)$$

kde $F_i(v)$ jsou distribuční funkce extrémních rychlostí větru v jednotlivých kategoriích a $F(v)$ je celková distribuční funkce. Číslo n představuje počet kategorií. Inverzní funkci ke vzorci (3.4.1) vyjadřující závislost rychlosti větru na distribuční funkci extrémních rychlostí větru, a tedy i na době návratu podle (3.1.1), nelze analyticky vyjádřit. Ale vzhledem k tomu, že je monotónní, lze pro výpočet extrémní rychlosti s určitou dobou návratu použít jednoduchou numerickou metodu, např. metodu půlení intervalů. Skládání podle vzorce (3.4.1) má odlišné vlastnosti než skládání průměrných hodnot, což bude na příkladu separace naměřených hodnot rychlosti větru podle směru dokumentováno v podkapitole 3.6.

Pro rozdělování dat do kategorií se používají různé metody. Jednou z nich je kategorizace podle směru větru provedená např. v Moriarty et Templeton (1983) a Žuraňski et Jašpińska (1996). Toto rozdělení do jednotlivých sektorů se odůvodňuje možností odstranění vlivu lokální orografie a překážek, případně i možností interpolace extrémních rychlostí větru do plochy. Ta není v případě nerozlišení směru větru extrémních událostí možná, protože lokální změny rychlosti větru silně závisejí na jeho směru. To je i důvodem využití kategorizace podle sektorů v této práci (kapitola 4). Směr extrémních rychlostí větru je důležitý také pro aplikace ve stavebnictví a v územním plánování. Směr se obvykle dělí do určitého počtu stejně velkých sektorů, ale možné je i rozdělení do nestejně velkých sektorů. Ty jsou v dané lokalitě stanoveny tak, aby „chování“ větru v jednotlivých sektorech bylo co nejpodobnější. Problémem je nalezení vhodné objektivní definice tohoto „chování“. Jiný přístup je použit v práci Coles et Walshaw (1994), kde se předpokládá, že parametry extrémního rozdělení větru jsou harmonickými funkcemi jeho směru.

Další možností kategorizace je rozdělení podle druhu situace způsobující extrémní událost. V našich zeměpisných šířkách je obvyklé dělení do dvou kategorií zhruba odpovídajících konvektivním událostem a událostem způsobených silným synoptickým tlakovým gradientem. Tato kategorizace byla prosazována např. v článcích Twisdale et Vickery (1992) a Lombardo et al. (2009). Problémem bývá odlišení těchto kategorií, které není tak jednoduché jako v případě směru větru. Použití pozorování na meteorologických stanicích je z řady důvodů problematické (například jde o vliv subjektivity pozorování na zaznamenaná data a dostupnosti). Subjektivní je v tomto případě i volba hranice mezi kategoriemi. Proto je obvyklá aplikace indexu dostupné konvektivní potenciální energie CAPE (Pešice et al., 2003), který je součástí různých reanalýz (NCEP/NCAR, ERA Interim). Vazba mezi tímto indexem a výskytem bouřek ale není příliš těsná a různé práce navrhují různé hraniční hodnoty (Pešice et al., 2003; Siedlecki, 2009; Chaudhuri et Middey, 2012). Tato kategorizace je výhodná při uvažování vlivu klimatických změn. Předpokládá se, že klimatické změny způsobí zeslabení větrů způsobených silným synoptickým gradientem (Francis et Vavrus, 2012), což je v souladu s článkem Brázdil et al. (2009), podle kterého vykazuje většina stanic v ČR sestupný trend rychlosti větru. Naopak v případě extrémních větrů způsobených silnou konvekcí se očekává jejich zesílení (Diffenbaugh et al, 2013). Možnosti zahrnutí vlivu klimatických změn do odhadu extrémních hodnot jsou popsány v AghaKouchak et al. (2012).

Další používaný způsob kategorizace je podle ročního období. Takové rozdělení událostí s extrémní rychlostí větru může být důležité například vzhledem ke škodám na vegetačním pokryvu. Podnebí v České republice má výrazný roční chod. Častým způsobem vypořádání se s tímto jevem je odstranění ročního chodu rychlosti větru a jejího rozptylu transformací, jejíž koeficienty závisejí na roční době. Problémem tohoto přístupu je, že extrémní případy se mohou řídit jiným ročním chodem než „průměrné“ případy (Davison et Smith, 1990). Alternativní postup spočívá v separaci událostí s neodstraněným ročním chodem do jednotlivých sezón, které se ale provádí zřídka (Cook, 1983).

V této práci bude pro analýzu naměřených hodnot aplikována separace podle sektorů směru větru, protože je nutná pro použitý způsob výpočtu mapy extrémního větru.

3.5 Rozdíl mezi různými definicemi doby návratu

3.5.1 Úvod

Doba návratu $T(x)$ pro rychlost větru x se definuje jako průměrná doba mezi dvěma po sobě následujícími událostmi s rychlostí větru vyšší než x (podkapitola 3.1). Výsledek vyplývající z této definice však závisí na tom, zda v ní uvažujeme všechny rychlosti větru (původní data) nebo pouze jejich bloková maxima. V této kapitole bude ukázáno, že mezi těmito dvěma přístupy existuje systematický rozdíl, který bude vyčíslen.

V této podkapitole budeme používat následující definice: (i) doba návratu $T(x)$ je průměrný čas mezi dvěma událostmi s rychlostí větru vyšší než x a (ii) doba návratu blokových maxim $T_b(n, x)$ je průměrný čas mezi dvěma událostmi s rychlostí větru vyšší než x a které jsou zároveň maximy ve svých blocích o délce n . Je zřejmé, že T_b závisí na délce bloku n .

3.5.2 Rozdíl mezi dvěma definicemi doby návratu

Uvažujme náhodnou proměnnou x s distribuční funkcí $F(x)$. Potom dobu návratu $T(x_0)$ hodnoty x_0 , která je definována jako průměrný čas mezi po sobě následujícími událostmi $x; x > x_0$, lze vyjádřit pomocí distribuční funkce F (podkapitola 3.1) vztahem

$$T(x) = \frac{T_0}{1-F(x)}, \quad (3.5.1)$$

kde T_0 je doba připadající na jedno měření. Pro jednoduchost budeme v následujícím textu předpokládat, že $T_0 = 1$.

V případě metody POT se výpočet uskutečňuje ve dvou krocích. Nejprve je pro hodnoty vyšší než mez u metody POT spočítána podmíněná distribuční funkce $H_p(u, x)$

$$H_p(u, x) = \frac{F(x)-F(u)}{1-F(u)} \quad (3.5.2)$$

a pak následuje výpočet doby návratu $T_p(u, v)$ pomocí (3.5.1)

$$T_p(u, x) = \frac{T(u)}{1-H_p(u, x)}. \quad (3.5.3)$$

Dolní index p ve vzorcích indikuje použití metody POT. Je snadné ukázat, že $T(x) = T_p(u, x)$ pro všechny hodnoty u takové, že $F(u) < 1$, a tedy že metoda POT vede ke stejným dobám návratu jako původní data:

$$T_p(u, x) = \frac{T(u)}{1-H_p(u, x)} = \frac{1}{1-F(x)} = T(x) \text{ for } x \geq u. \quad (3.5.4)$$

Při použití metody blokových maxim (BM) je situace komplikovanější. V tomto případě odpovídá jeden vybraný extrémní případ časovému intervalu $T_{b0}(n) = n$, kde dolní index b značí, že je používána metoda BM. Distribuční funkci G blokových maxim M_n lze vyjádřit pomocí distribuční funkce F (vztah 3.3.2)

$$G(n, x) = \{F(x)\}^n \quad (3.5.5)$$

a následně doba návratu $T_b(n, x)$ je

$$T_b(n, x) = \frac{T_{b0}(n)}{1-G(n, x)} = \frac{n}{1-\{F(x)\}^n}. \quad (3.5.6)$$

Ze srovnání vzorců (3.5.4) a (3.5.6) vyplývá, že se doby návratu získané metodami BM a POT liší. Důvodem je to, že $T_b(n, x)$ musí splňovat stejnou podmínku na překročení dané hodnoty jako $T_p(n, x)$, ale daná událost současně musí být maximem ve svém bloku. Je zřejmé, že pro jakékoli x_0 takové, že $F(x_0) < 1$, existuje nenulová pravděpodobnost, že dvě nebo více událostí x takových, že $x \geq x_0$, může nastat v jednom bloku. Pro x splňující $F(x) < 1$ a $n \geq 2$ tak platí následující nerovnost:

$$T_b(n, x) - T(x) > 0 \quad . \quad (3.5.7)$$

Limitu rozdílu na levé straně (3.5.7) pro $T(x) \rightarrow \infty$ lze odhadnout. Pokud označíme pravděpodobnost, že $x \geq x_0$, jako $t(x)$

$$t(x) = 1 - F(x) = \frac{1}{T(x)} \quad , \quad (3.5.8)$$

potom

$$T_b(n, x) - T(x) = \frac{n}{1-(F(x))^n} - \frac{1}{1-F(x)} = \frac{n}{1-(1-t(x))^n} - \frac{1}{t(x)} \quad (3.5.9)$$

a aplikací limity $T(x) \rightarrow \infty$ získáme

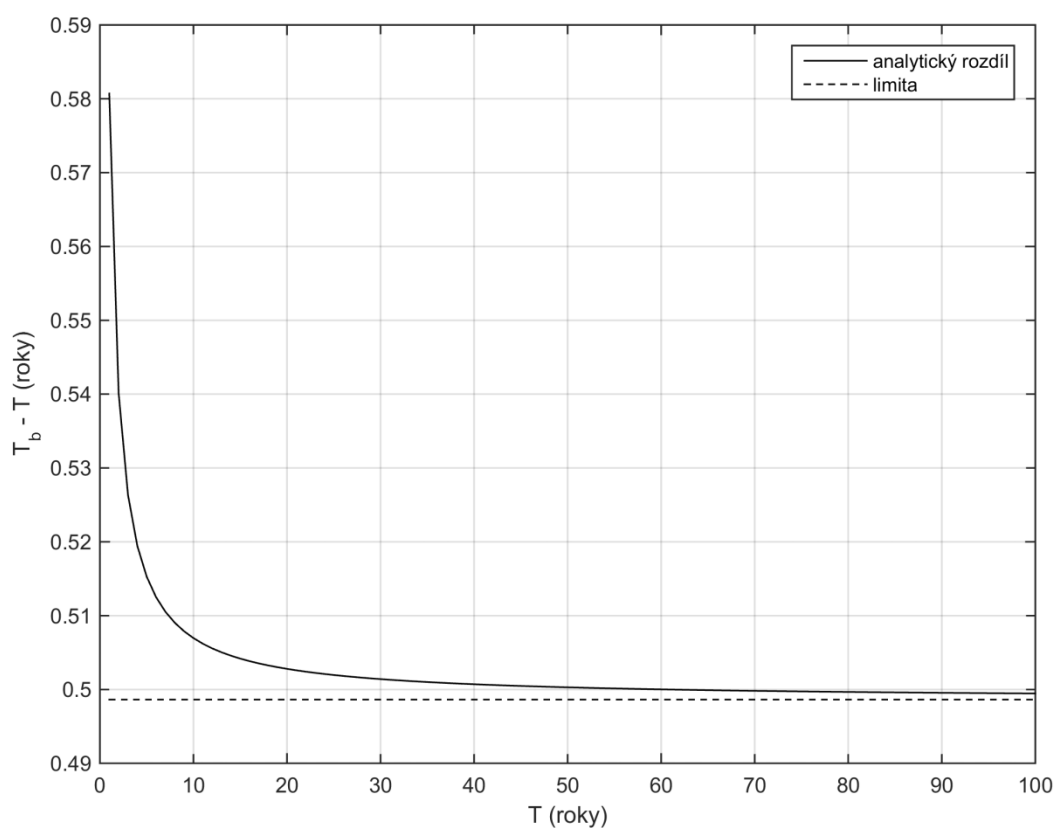
$$\begin{aligned} \lim_{T(x) \rightarrow \infty} \{T_b(n, x) - T(x)\} &= \lim_{t(x) \rightarrow 0} \{T_b(n, x) - T(x)\} = \\ \lim_{t(x) \rightarrow 0} \left(\frac{n}{1 - \left[1 - n * t(x) + \frac{n * (n-1) * t(x)^2}{2} + o\{t(x)^2\} \right]} - \frac{1}{t(x)} \right) &= \frac{n-1}{2}. \end{aligned} \quad (3.5.10)$$

Vzhledem k tomu, že jsme položili $T_0 = 1$, je výsledek v jednotkách intervalu měření původních dat. Získané výsledky lze shrnout takto. Zatímco výběr extrémních dat metodou POT vede ke stejným dobám návratu jako doba návratu původních dat, výběr metodou BM vede k systematicky vyšší hodnotě doby návratu, přičemž rozdíl je pro dlouhé doby návratu přibližně roven jedné polovině délky jednoho bloku. V případě použití ročních maxim je tedy tento rozdíl přibližně roven polovině roku.

Rovnici (3.5.10) můžeme použít k porovnání dob návratu získaných metodou BM s různými délkami bloků n_1 a n_2

$$\lim_{T(x) \rightarrow \infty} \{T_b(n_1, x) - T_b(n_2, x)\} = \frac{n_1-1}{2} - \frac{n_2-1}{2} = \frac{n_1-n_2}{2} . \quad (3.5.11)$$

Tento výsledek znamená, že rozdíl v dobách návratu se bude pro dlouhé doby návratu blížit polovině rozdílu délek bloků. Rychlost konvergence na příkladu, kdy původními daty jsou denní hodnoty a blokovými maximy roční maxima, je ukázána na obr. 3.5.1. Doby návratu jsou vyjádřeny v letech. Konvergence k limitní hodnotě je s rostoucí dobou návratu T velmi rychlá.



Obr. 3.5.1: Konvergence rozdílu dob návratu počítaných z denních hodnot a ročních maxim (vertikální osa). T_b a T jsou po řadě doby návratu počítané metodou BM a s použitím všech dat. Přerušovaná čára představuje limitní hodnotu danou vzorcem (3.5.10).

Rovnice (3.5.1) a (3.5.6) lze použít ke korekci dob návratu získaných metodou BM tak, aby se odstranil jejich systematický rozdíl od dob návratu původních dat. Z těchto rovnic lze vyjádřit $T(x)$ jako funkci $T_b(n, x)$:

$$T(x) = \frac{1}{1 - n \sqrt{\frac{T_b(x) - n}{T_b(x)}}} . \quad (3.5.12)$$

3.5.3 Aplikace na Gumbelovo rozdělení

Distribuční funkce Gumbelova rozdělení (Kapitola 3.2) má tvar

$$F(x) = \exp \left[- \exp \left\{ \frac{-(x-\mu)}{\sigma} \right\} \right], \quad (3.5.13)$$

kde μ a σ jsou jeho parametry. Toto rozdělení se často používá pro aproximaci ročních maxim rychlosti větru a je použito v kapitole 4. Pokud jsou roční maxima dána tímto rozdělením, pak z této rovnice a z rovnice (3.5.6) plyne pro hodnoty x_b s dobou návratu T_b vztah

$$x_b(T_b) = \mu + \ln(T_b) * \sigma. \quad (3.5.14)$$

Pokud jsou doby návratu počítány ze všech hodnot, lze hodnoty $x(T)$ s dobou návratu T vyjádřit ve formě

$$x(T) = x_b\{T_b(T)\} = \mu + \ln\{T_b(T)\} * \sigma \quad (3.5.15)$$

a $T_b(T)$ lze vyjádřit pomocí 3.5.12 jako

$$T_b(T) = \frac{1}{1 - \left(\frac{n * T - 1}{n * T}\right)^n}, \quad (3.5.16)$$

kde $n = 365$ je počet dní v roce. Systematický rozdíl v dobách návratu b

$$b(T) = x_b\{T_b(T)\} - x_b(T), \quad (3.5.17)$$

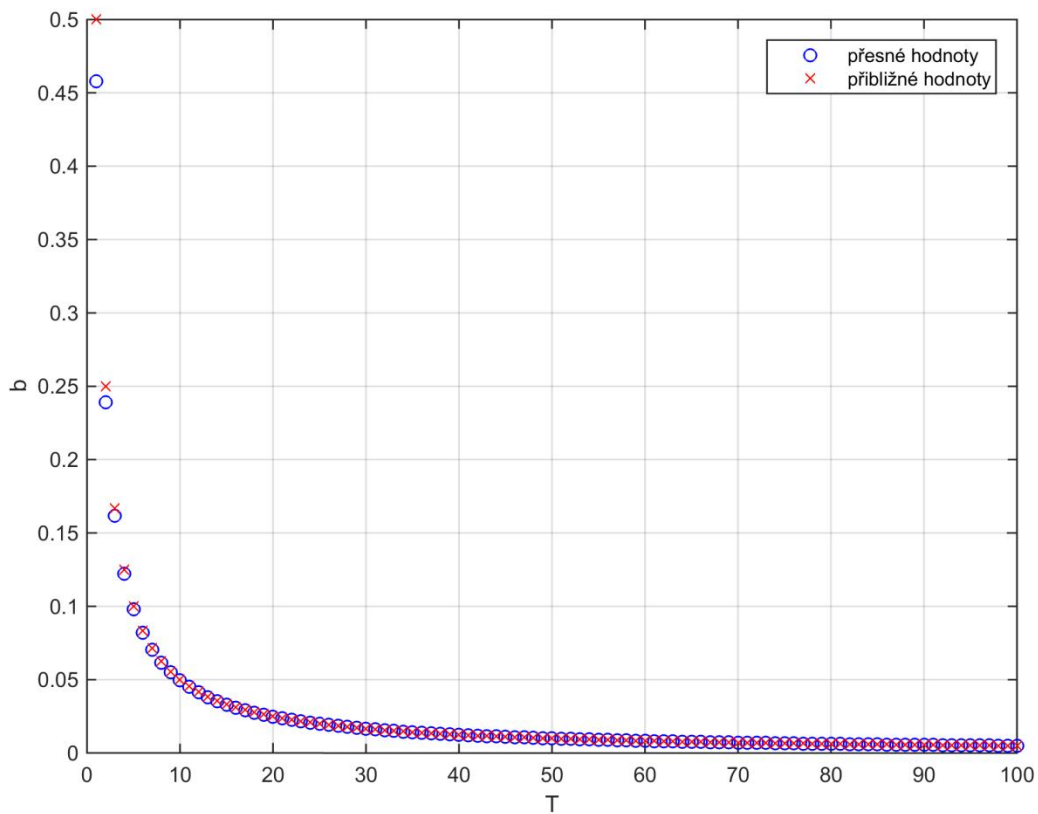
lze vyjádřit pomocí vzorce (3.5.15) jako

$$b(T) = \sigma * [\ln\{T_b(T)\} - \ln(T)]. \quad (3.5.18)$$

Pokud $T \gg 1$, potom podle (3.5.10) $T_b(T) - T \approx \frac{1}{2}$ a $\ln(T+1/2) \approx \ln(T) + 1/2T$, z čehož dostaneme následující odhad:

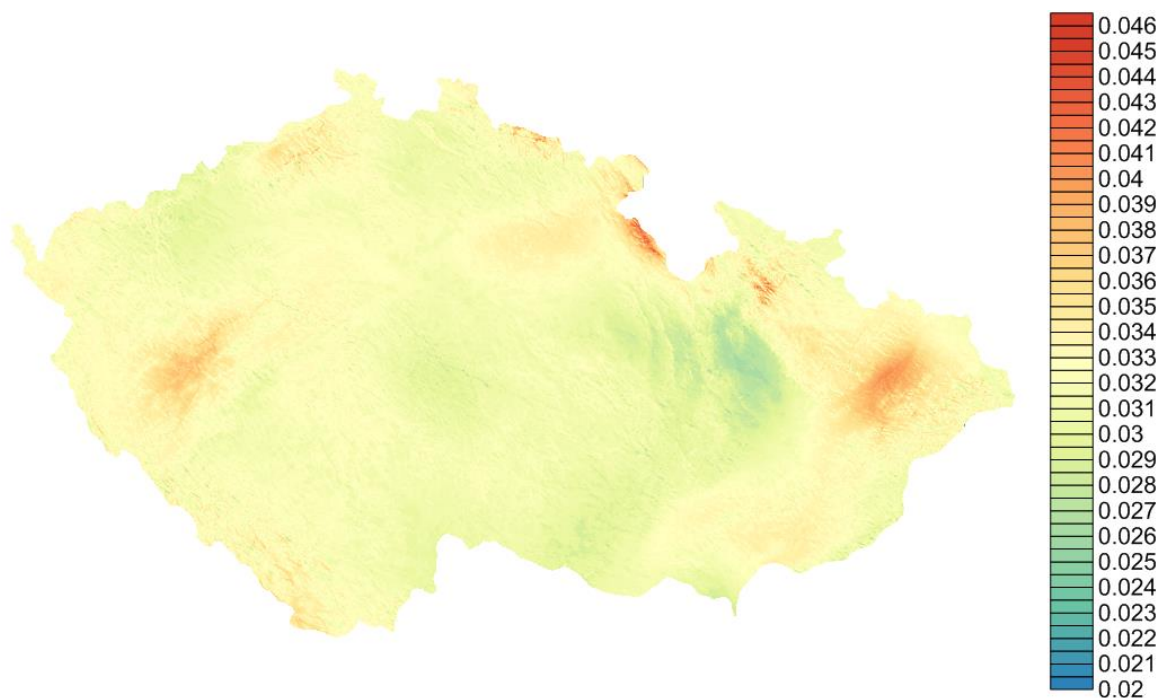
$$b(T) \approx \frac{\sigma}{2T}, \quad (3.5.19)$$

který říká, že pro vysoké hodnoty doby návratu T je hodnota systematického rozdílu $b(T)$ přímo úměrná parametru σ Gumbelova rozdělení a nepřímo úměrná době návratu T . To znamená, že význam tohoto systematického rozdílu klesá s rostoucí dobou návratu. Obr. 3.5.2 ukazuje, že aproximace (3.5.19) je velmi přesná pro doby návratu delší než 5 let.



Obr. 3.5.2: Porovnání přesnosti odhadu $b(T)$ podle rovnice (3.5.19). Srovnání je provedeno pro roční maxima pocházející z Gumbelova rozdělení s $\sigma = 1$ (rovnice 3.5.13).

Praktický dopad systematického rozdílu extrémních rychlostí s danou dobou návratu není velký a je ilustrován na obr. 3.5.3, kde je zobrazen odhad tohoto systematického rozdílu pro nárazy větru ze sektoru č. 7 (přibližně severozápad) pro území ČR pro dobu návratu 10 let. Hodnoty byly získány pomocí rovnice (3.5.18). Výpočet mapy hodnot parametru σ Gumbelova rozdělení je popsán v podkapitole 4.2. Systematický rozdíl je v průměru přibližně 0,30 m/s pro dobu návratu 20 let (obr. 3.5.3), což odpovídá přibližně 0,5 %.



Obr. 3.5.3: Odhad systematického rozdílu rychlosti nárazu větru s dobou návratu 10 let ze sektoru č. 7 způsobeného různými definicemi doby návratu podle (3.5.1) a (3.5.6).

3.6 Vliv tvaru růžice rychlostí větru na vztah mezi průměrnou a extrémní rychlostí větru

Jednou z nejjednodušších metod odhadu extrémních rychlostí větru je prosté vynásobení průměrné rychlosti větru empirickou konstantou, což je přístup použitý například v normě IEC 61400-1 (IEC, 2005, dále jen jako norma), a diskutovaný a verifikovaný v podkapitole 4.1. Výhodou tohoto postupu je jeho robustnost. Takto lze snadno získat plošnou informaci z map průměrné rychlosti větru, které často vznikají za účelem odhadu potenciálu daného území z hlediska větrné energetiky – viz Landberg et al. (2003) a Ayotte (2008). Pro naši republiku lze mapu průměrných rychlostí větru ve standardní výšce 10 m a popis jejího vzniku nalézt, např. v článkách Hanslian et al. (2013a) a Hanslian et al. (2013b), technické detaily pak v článkách Hanslian et al. (2012) a Svoboda et al. (2012). Obdobné mapy lze nalézt i pro jiné země, hlavně s rozvinutou větrnou energetikou, např. Fueyo et al. (2010) a Costa et al. (2006). Problematickými vlastnostmi těchto map je, že různí autoři používají různou metodiku výpočtů a že velké množství studií je neveřejných z důvodu jejich komerční využitelnosti ve větrné energetice.

Jednou z nevýhod výše uvedené metody odhadu extrémních rychlostí větru je to, že zanedbává vliv tvaru větrné růžice v lokalitě, což může vést k velkým chybám. V této kapitole ukážeme, jaký vliv má složení průměrných a extrémních rychlostí větru z jednotlivých sektorů na celkový podíl mezi průměrnou a extrémní rychlostí větru za předpokladu, že extrémní rychlosti v jednotlivých sektorech mají Gumbelovo rozdělení. Získané poznatky jsou využity v podkapitole 4.1.

3.6.1 Způsob spojení průměrných a extrémních rychlostí větru z jednotlivých sektorů

Průměrná rychlost větru v se při znalosti větrné růžice vypočte vztahem

$$v = \sum_{i=1}^n f_i v_i, \quad (3.6.1)$$

kde n je počet směrových sektorů, f_i jsou četnosti výskytu větru z jednotlivých sektorů a v_i průměrné rychlosti větru z těchto sektorů. Průměrná rychlost je tedy významně ovlivněna průměrnými rychlostmi ze všech sektorů, u kterých je četnost výskytu větru nezanedbatelná.

Postup spojení extrémních rychlostí větru z jednotlivých sektorů do jediné celkové hodnoty je popsán v podkapitole 3.4 a vede ke vzorci

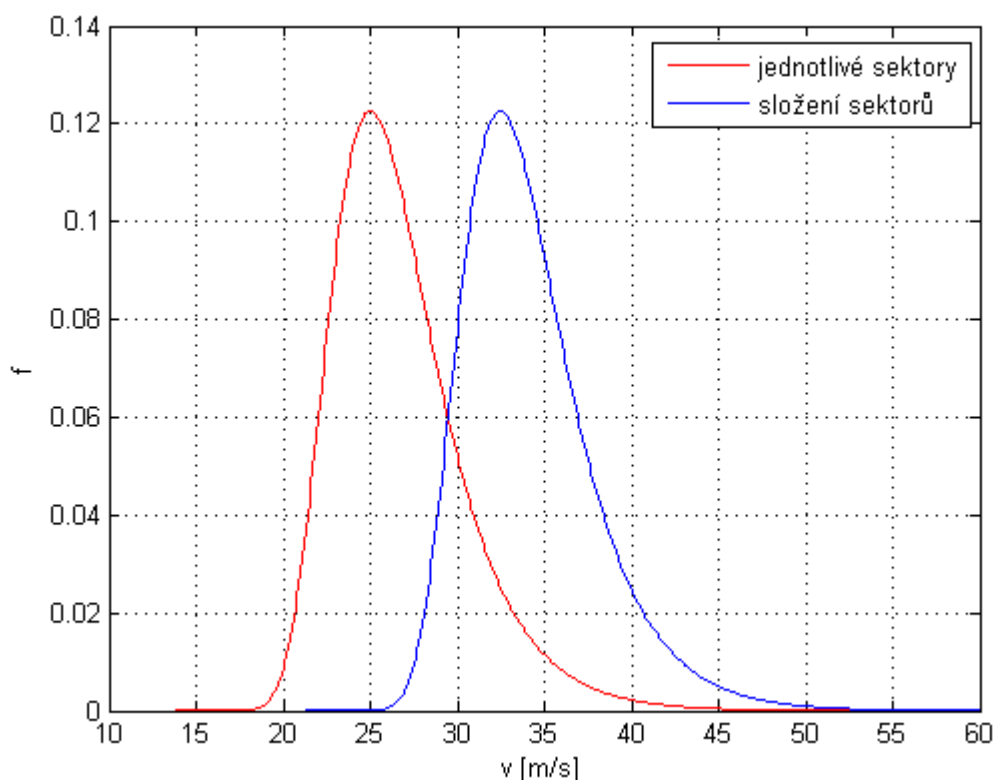
$$F(x) = \prod_{i=1}^n F_i(x), \quad (3.6.2)$$

kde $F_i(x)$ jsou distribuční funkce extrémních rychlostí větru z jednotlivých sektorů, $F(x)$ je celková distribuční funkce a n je počet sektorů. Celková rychlost větru s dobou návratu n let musí být z definice vyšší nebo rovna rychlosti s dobou návratu n let z jednotlivých sektorů. Při této hodnotě rychlosti budou ale distribuční funkce extrémních rychlostí větru z málo větrných sektorů velmi blízké hodnotě 1, a proto bude vliv těchto sektorů na celkovou extrémní rychlost minimální, což bude ukázáno na následujícím příkladu.

3.6.2 Příklad spojení průměrných a extrémních rychlostí větru

Předpokládejme, že větrná růžice je tvořena 12 stejně velkými sektory, přičemž četnost větru ze všech sektorů f_i je stejná, tj. $1/12$. Ve všech sektorech předpokládáme stejný konstantní poměr $w(T)$ mezi rychlostí $v_i(T)$ s dobou návratu T ze sektoru i a průměrnou rychlostí v_i z tohoto sektoru. Dále předpokládejme, že rozdělení extrémních rychlostí větru z jednotlivých směrů je dáno Gumbelovým rozdělením. Porovnáme dva případy:

(i) Uvažujeme izotropní větrnou růžici, což znamená, že v_i je ve všech sektorech stejná. Potom budou stejné i všechny distribuční funkce Gumbelova rozdělení extrémních rychlostí větru z jednotlivých sektorů. Předpokládejme, že tato Gumbelova rozdělení mají parametry $\mu_0 = 25 \text{ m/s}$ a $\sigma_0 = 3 \text{ m/s}$. V tomto případě lze celkovou distribuční funkci podle vzorce (3.6.2) spočítat přímo, protože n -tou mocninou Gumbelova rozdělení s parametry μ a σ je Gumbelovo rozdělení posunuté o $\sigma \cdot \ln(n)$ doprava (vzorec 3.3.4). Pro uvažované hodnoty parametrů a počet sektorů je velikost tohoto posunu $7,45 \text{ m/s}$. Výsledné pravděpodobnostní rozdělení ročních maximálních rychlostí větru z jednotlivých sektorů a celkově jsou zobrazeny na obr. 3.6.1. Nejvyšší rychlost větru s dobou opakování 50 let u_{50} má v jednotlivých sektorech hodnotu $36,71 \text{ m/s}$ a odpovídající celková hodnota pro všechny sektory dohromady činí $44,16 \text{ m/s}$.



Obr. 3.6.1: Hustoty pravděpodobnosti f nejvyšších ročních rychlostí větru z jednotlivých sektorů a pro všechny směry dohromady v případě izotropní větrné růžice.

(ii) Uvažujeme modifikaci izotropní větrné růžice spočívající v tom, že rychlost větru je v jednom ze sektorů zesílena vynásobením konstantou k_1 a v ostatních

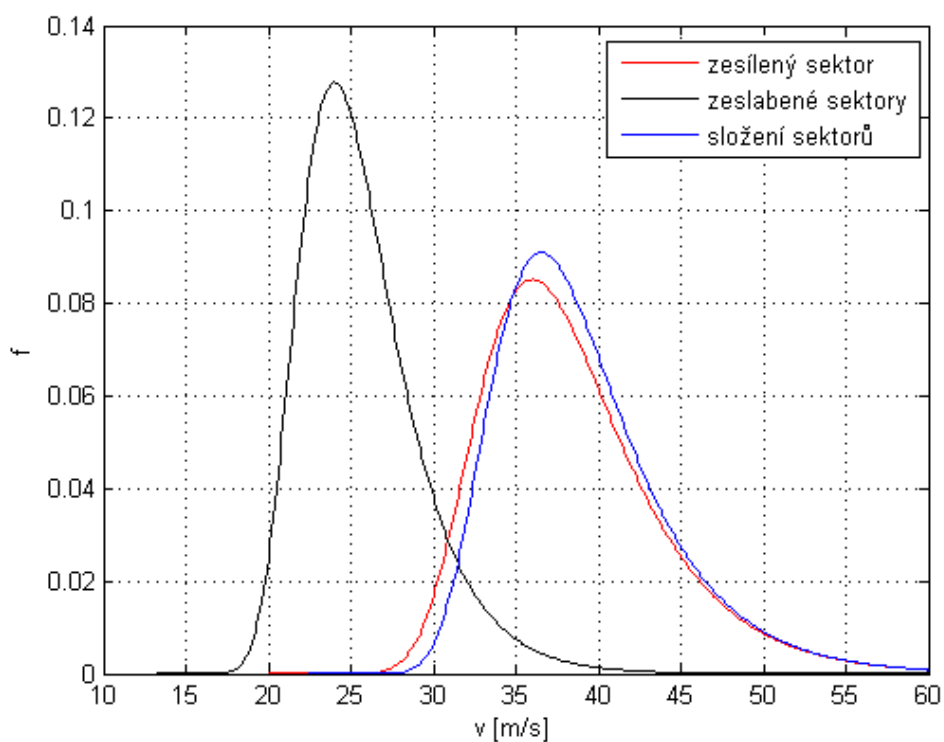
sektorech naopak zeslabena vynásobením konstantou k_2 tak, aby průměrná rychlost zůstala stejná. Tomu podle uvedených předpokladů odpovídá i vynásobení parametrů μ a σ Gumbelova rozdělení extrémních rychlostí větru těmito konstantami. V našem případě zvolíme hodnoty konstant k_1 a k_2 tak, aby se průměrná rychlost větru ze zesíleného sektoru rovnala třem polovinám průměrné rychlosti ze zeslabených sektorů, což odpovídá následující soustavě rovnic:

$$k_1 = 3/2k_2 \quad (3.6.3)$$

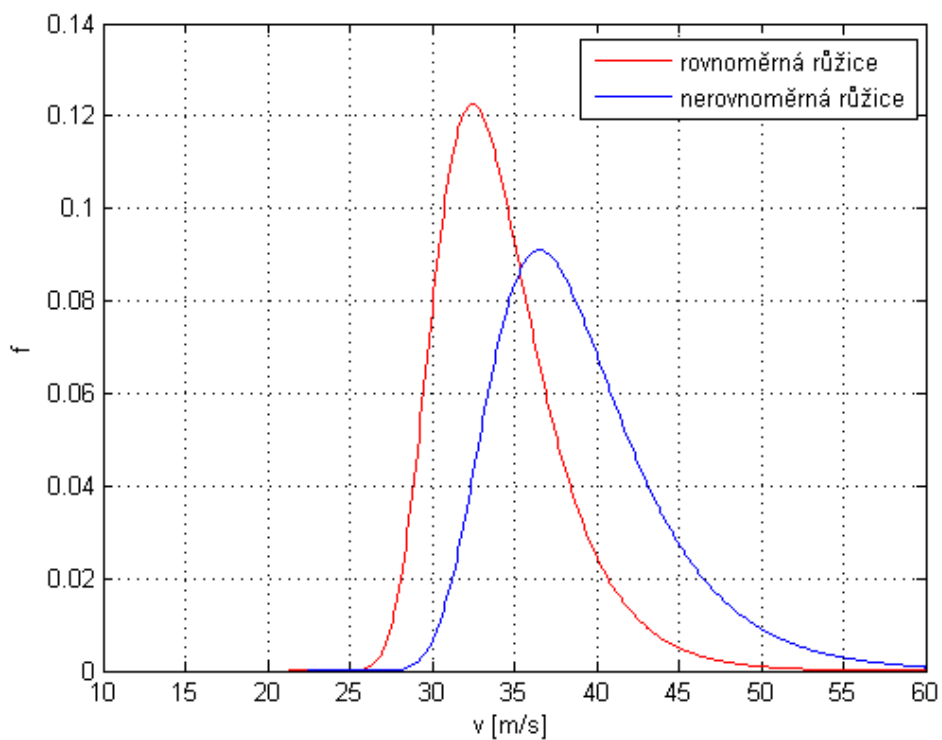
$$\frac{1}{12}k_1 + \frac{11}{12}k_2 = 1. \quad (3.6.4)$$

Řešením této soustavy jsou hodnoty $k_1 = 36/25$ a $k_2 = 24/25$. Odpovídající parametry Gumbelova rozdělení mají hodnotu $\mu_1 = 36$ m/s, $\sigma_1 = 4,32$ m/s, $\mu_2 = 24$ m/s a $\sigma_2 = 2,88$ m/s. Výsledné pravděpodobnostní rozdělení ročních maximálních rychlostí větru z jednotlivých sektorů a celkové rozdělení je zobrazeno na obr. 3.6.2. Z něho je patrné, že hlavně pro dlouhé doby opakování se celkové rozdělení nejvyšších ročních rychlostí větru blíží rozdělení nejvyšších ročních rychlostí větru ze zesíleného sektoru. u_{50} má v jednotlivých zeslabených sektorech hodnotu 35,24 m/s, v zesíleném sektoru 52,86 m/s a celkově nezávisle na směru 52,96 m/s. Je tedy o pouhých 0,1 m/s vyšší než u_{50} ze zesíleného sektoru a o významných 8,80 m/s vyšší než u_{50} z případu se stejnou průměrnou rychlostí větru s izotropní růžicí.

Porovnání rozdělení pravděpodobnosti nejvyšších ročních rychlostí větru v těchto dvou případech se stejnou průměrnou rychlostí větru, ale s rozdílným tvarem větrné růžice je na obr. 3.6.3. Výpočet tohoto modelového případu ukazuje, že ve srovnání s odhadem extrémních rychlostí větru pouze na základě průměrné rychlosti obdobně jako v normě je možno dosáhnout podstatného zlepšení odhadu zohledněním tvaru větrné růžice.



Obr. 3.6.2: Hustoty pravděpodobností f nejvyšších ročních rychlostí větru z jednotlivých sektorů a celkové složené rychlosti v případě nerovnoměrné růžice větru.



Obr. 3.6.3: Srovnání celkových hustot pravděpodobnosti f nejvyšších ročních rychlostí větru v případě rovnoměrné a nerovnoměrné růžice větru (viz text).

3.7 Aproximace rozdělení naměřených dat pomocí GEV a Gumbelova rozdělení

Tato kapitola popisuje dvě často používané metody aproximace rozdělení naměřených hodnot zvoleným teoretickým rozdělením a jejich aplikace na GEV a Gumbelovo rozdělení. Jedná se o metody maximální věrohodnosti (ML) a L-momentů. V podkapitolách 3.8 a 3.9 budou tyto metody porovnány a jejich výsledky analyzovány na uměle vygenerovaných a naměřených datech.

3.7.1 Metoda maximální věrohodnosti

Metoda maximální věrohodnosti, popsána např. ve Wilks (2011), spočívá v maximalizaci tzv. věrohodnostní funkce, která vyjadřuje míru shody naměřených dat s hustotou pravděpodobnosti teoretického rozdělení. Pro jedno měření je tato funkce identická hustotě pravděpodobnosti teoretického rozdělení pro hodnotu tohoto měření, přičemž hodnota měření se bere jako fixní a proměnnými jsou parametry zvoleného teoretického rozdělení.

Věrohodnostní funkce pro více bodů má tvar součinu věrohodnostních funkcí pro jednotlivé body, což odpovídá faktu, že pravděpodobnost průniku více nezávislých jevů je přímo úměrná součinu jejich jednotlivých pravděpodobností.

$$\Lambda(\mathbf{p}) = \prod_{i=1}^n \varrho(x_i, \mathbf{p}), \quad (3.7.1)$$

kde x_i jsou jednotlivá měření, \mathbf{p} jsou parametry rozdělení a ϱ hustota pravděpodobnosti rozdělení. Maximalizaci $\Lambda(\mathbf{p})$ lze nahradit maximalizací funkce

$$L(\mathbf{p}) = \ln(\Lambda(\mathbf{p})) = \sum_{i=1}^n \ln(\varrho(x_i, \mathbf{p})). \quad (3.7.2)$$

Pro některá rozdělení, např. Gaussovo, existuje analytické řešení tohoto problému, ale většinou je nutno maximalizaci řešit numericky iterativními metodami. Pro numerické řešení těchto úloh zpravidla existují osvědčené metody, které jsou součástí většiny statistických softwarů.

Zřejmě platí, že $L(\mathbf{p}) = -\infty$ pokud $\varrho(x_i, \mathbf{p}) = 0$ pro nějakou naměřenou hodnotu x_i , a proto bude pro odhad metodou maximální věrohodnosti platit, že hustota pravděpodobnosti odhadnutého rozdělení je ve všech naměřených hodnotách kladná. Důsledkem pro odhad parametrů GEV rozdělení je, že žádná z naměřených hodnot

nebude ležet za pravým krajním bodem vypočteného rozdělení, a proto odhadnutá doba opakování bude konečná pro všechny naměřené hodnoty, což není zaručeno v případě aproximace parametrů GEV rozdělení metodou L-momentů.

3.7.2 Metoda momentů

Metoda momentů (Wilks, 2011) počítá parametry rozdělení tak, že vyžaduje shodu prvních n statistických momentů naměřeného vzorku a zvoleného rozdělení, kde n je počet parametrů daného rozdělení. V případě Gumbelova rozdělení $n = 2$ a v případě GEV rozdělení $n = 3$. Takto získáme soustavu n rovnic pro n neznámých hodnot parametrů, která se analyticky nebo numericky vyřeší.

Nevýhodou této metody je její malá robustnost plynoucí z faktu, že při výpočtu n -tého momentu se naměřené hodnoty umocňují na n . Kromě toho v případě GEV jsou momenty vyšší nebo rovné než $1/\xi$ (Reiss at Thomas, 2001) nekonečné, což omezuje výslednou hodnotu parametru ξ na hodnoty nižší než $1/3$. Proto se častěji používají odhady založené na aplikaci L-momentů, které z velké části uvedené nedostatky odstraňují (viz níže).

3.7.3 L-momenty

V odborné literatuře se používají dva ekvivalentní přístupy při definování „alternativních“ momentů, které nahrazují standardní statistické momenty. První, popsáný v Greenwood at al. (1979), využívá speciální případy pravděpodobnostně vážených momentů ve tvaru $M_{1,0,s}$ a $M_{1,r,0}$, kde pravděpodobnostně vážené momenty $M_{p,r,s}$ se definují vztahem

$$M_{p,r,s} = E(x^p (F(x))^r (1 - F(x))^s) \quad (3.7.3)$$

kde F značí distribuční funkci a E střední hodnotu. V této práci je použita metoda L-momentů, kterou zavedl Hosking (1990), a která je ekvivalentní výše uvedenému postupu. Výhodou L-momentů je fakt, že jsou lépe interpretovatelné a představují robustní alternativu klasických momentů.

L-momenty se definují následovně. Mějme hodnoty seřazeného náhodného výběru hodnot o velikosti n proměnné X , které označíme jako $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$. Potom L-moment r -tého řádu proměnné X definujeme jako

$$\lambda_r \equiv r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r-k:r}). \quad (3.7.4)$$

E opět značí střední hodnotu. První L-moment je zřejmě stejný jako první statistický moment. L-momenty všech řádů existují, pokud je průměr rozdělení konečný (David et Nagaraja, 2003), tedy v případě GEV rozdělení pokud je parametr $\xi < 1$, na rozdíl od klasických momentů, které jsou všechny definovány pouze pro hodnoty $\xi \leq 0$ (Reiss et Thomas, 2001).

Nevychýlený odhad L-momentu řádu k , l_k , dostaneme nahrazením střední hodnoty ve vzorci (3.7.4) průměrem přes všechny seřazené kombinace bez opakování k hodnot vybraných z náhodného vzorku x_i o velikosti n (Hosking, 1990), což je vyjádřeno vzorcem

$$l_r = \binom{n}{r}^{-1} \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} x_{i_{r-k:n}}. \quad (3.7.5)$$

Místo L-momentů třetího a vyššího řádu se obvykle používají tzv. normované L-momenty τ_r definované vztahem

$$\tau_r \equiv \frac{\lambda_r}{\lambda_2}, \quad r = 3, 4, \dots \quad (3.7.6)$$

První L-momenty Gumbelova a GEV rozdělení jsou uvedeny v Hosking (1990). V případě Gumbelova rozdělení jsou dány vztahy

$$\lambda_1 = \mu + \gamma\sigma, \quad (3.7.7)$$

$$\lambda_2 = \sigma \ln(2), \quad (3.7.8)$$

kde γ značí Eulerovu-Mascheroniho konstantu s přibližnou hodnotou 0,5772.

V případě GEV rozdělení se jedná o vztahy

$$\lambda_1 = \mu - \sigma\{1 - \Gamma(1 - \xi)\}/\xi, \quad (3.7.9)$$

$$\lambda_2 = \sigma(2^\xi - 1)\Gamma(1 + \xi)/\xi, \quad (3.7.10)$$

$$\tau_3 = \frac{2(1-3^\xi)}{(1-2^\xi)} - 3, \quad (3.7.11)$$

kde Γ značí Gama funkci.

3.7.4 Odhad parametrů Gumbelova a GEV rozdělení metodou L-momentů

K odhadu parametrů teoretického rozdělení metodou L-momentů se používá analogická metoda jako v případě klasické metody momentů, tj. sestaví se n rovnic pro n neznámých vyjadřujících rovnost prvních L-momentů předpokládaného rozdělení s odhady L-momentů z daného výběru dat. Tím se získá soustava n rovnic pro n neznámých, kde n je počet parametrů zvoleného rozdělení. Řešení této soustavy rovnic pro nejčastěji používaná rozdělení je uvedeno v Hosking (1990).

V případě Gumbelova rozdělení mají tyto soustavy rovnic řešení

$$\sigma = l_2/\ln(2), \quad (3.7.12)$$

$$\mu = l_1 - \gamma\sigma. \quad (3.7.13)$$

V případě GEV rozdělení jsou řešením vztahy

$$z = \frac{2}{3+\tau_3} - \ln(2)/\ln(3), \quad (3.7.14)$$

$$\xi \cong -7.8590z - 2.9554z^2, \quad (3.7.15)$$

$$\sigma = \frac{l_2\xi}{(2^\xi-1)\Gamma(1-\xi)}, \quad (3.7.16)$$

$$\mu = l_1 + \frac{\sigma\{1-\Gamma(1-\xi)\}}{\xi}. \quad (3.7.17)$$

3.8 Analýza aproximace extrémních dat GEV a Gumbelovým rozdělením metodami L-momentů a ML na uměle vygenerovaných datech

Současné možnosti výpočetní techniky umožňují uměle vygenerovat a zpracovat velká množství dat. Této skutečnosti bylo využito pro porovnání metod L-momentů a ML, které byly použity k aproximaci rozdělení uměle generovaných dat pomocí GEV a Gumbelova rozdělení. Použití umělých dat má výhodu znalosti některých jejich vlastností, které v případě naměřených dat nejsou přesně známy, protože není známa distribuční funkce rozdělení, ze kterého naměřená data pocházejí. Výsledky této podkapitoly se využijí jednak pro srovnání s výsledky podkapitoly 3.9, jednak spolu s dalšími analýzami pro volbu vhodné metody výpočtu větrné mapy v kapitole 4 a odhad její chyby.

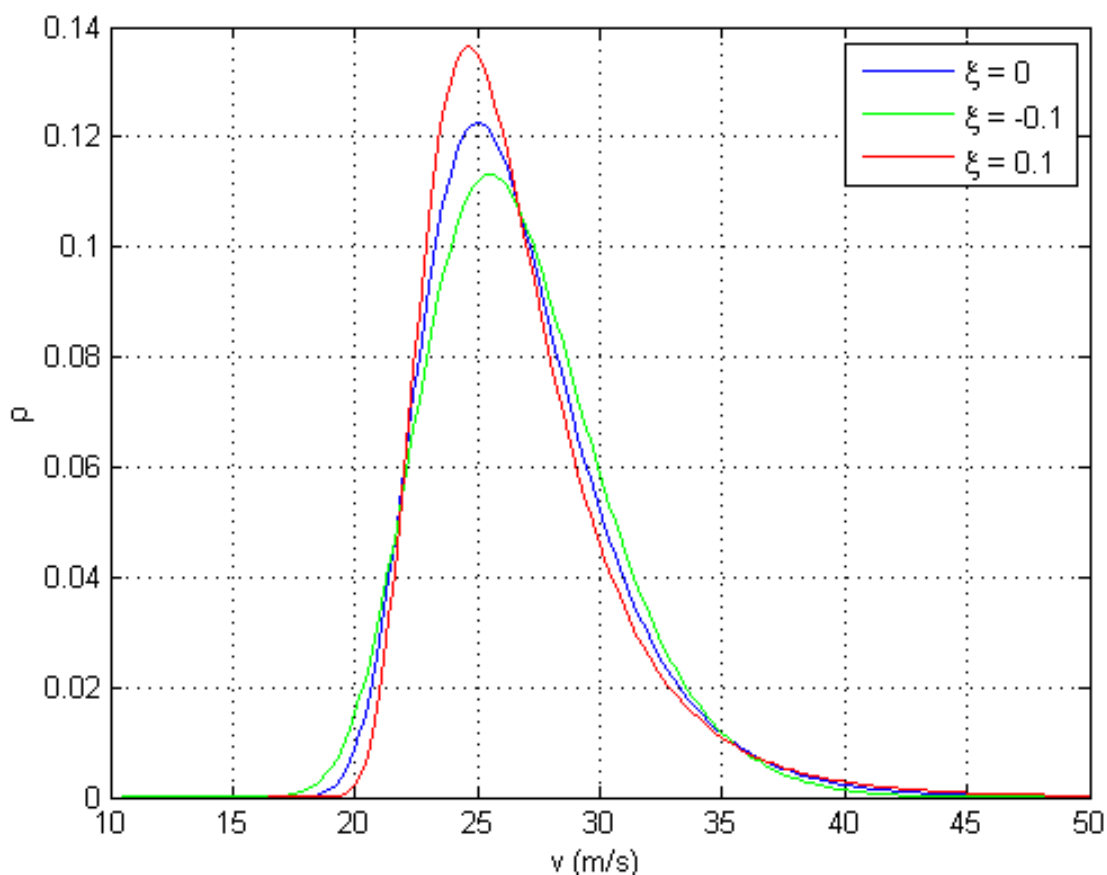
Z důvodu názornosti jsou následující testování a výpočty provedeny na datech, jejichž vlastnosti byly zvoleny tak, aby odpovídaly typickému rozdělení extrémních rychlostí větru v ČR. Většina získaných poznatků však není omezena na extrémní vítr, ale platí i pro extrémny jiných veličin. Umělá data byla generována na základě Gumbelova rozdělení s parametry $\mu = 25$ a $\sigma = 3$. Takto zvolené hodnoty parametrů nejsou na újmu obecnosti, protože Gumbelovo rozdělení s jinými parametry získáme z tohoto základního rozdělení pouze lineární transformací nezávisle proměnné.

Pro určení vlivu odhadu parametru tvaru GEV rozdělení na odhady dalších veličin se ukázalo nutné definovat tzv. L-ekvivalentní GEV rozdělení.

3.8.1 Definice a výpočet L-ekvivalentního rozdělení

V této sekci chápeme Gumbelovo rozdělení jako speciální případ GEV rozdělení. Pro vyšetřování vlivu parametru ξ GEV rozdělení na různé metody jsme k danému (původnímu) GEV rozdělení definovali jemu nejbližší GEV rozdělení s jiným parametrem ξ jako GEV rozdělení, které má tento parametr tvaru a stejné hodnoty prvních dvou L-momentů jako původní rozdělení. Toto rozdělení jsme nazvali jako L-ekvivalentní původnímu rozdělení. Jako L-ekvivalentní také budeme označovat množinu GEV rozdělení se stejnými parametry prvních dvou L-momentů.

Určení L-ekvivalentního GEV rozdělení s daným parametrem ξ rozdělení k původnímu GEV rozdělení spočívá ve výpočtu jeho parametrů μ a σ . To se provede tak, že spočítáme první dva L-momenty původního GEV rozdělení vzorci (3.7.7) a (3.7.8) ($\xi = 0$) nebo (3.7.9) a (3.7.10). Parametry μ a σ výsledného GEV rozdělení pro tyto spočtené hodnoty L-momentů jsou dány rovnicemi (3.7.12) a (3.7.13) ($\xi = 0$) nebo (3.7.16) a (3.7.17).

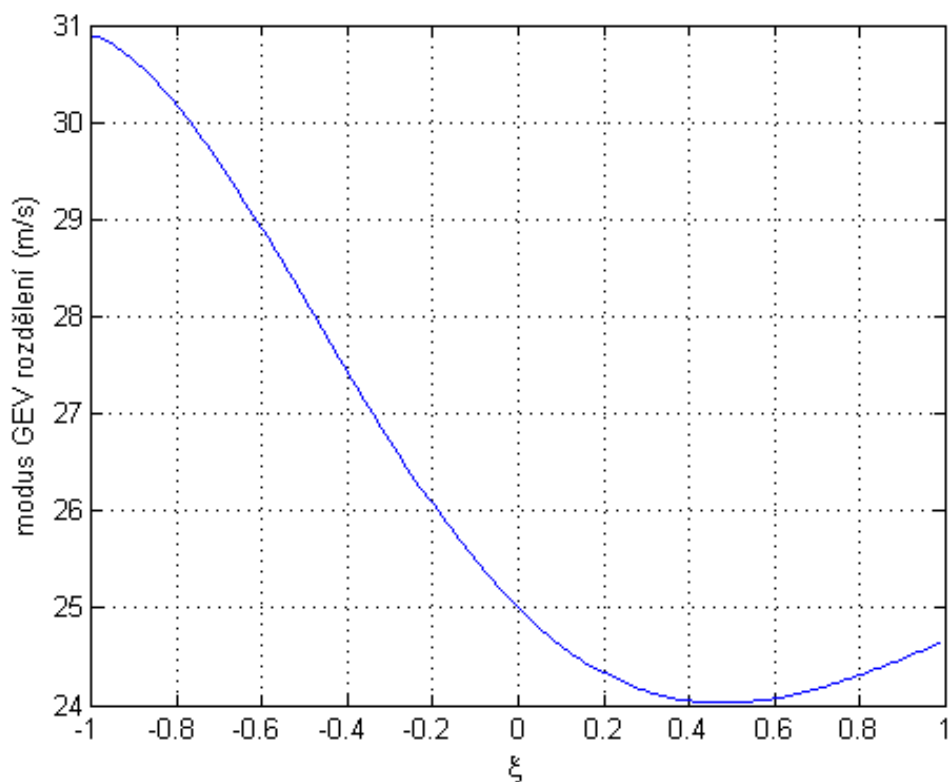


Obr. 3.8.1: Porovnání L-ekvivalentních GEV rozdělení.

Výše uvedené ilustruje obr. 3.8.1., kde byl testován vliv parametru ξ na různé charakteristiky L-ekvivalentních GEV rozdělání. Za základní rozdělání jsme zvolili Gumbelovo rozdělání s $\mu = 25$ m/s a $\sigma = 3$ m/s a k němu vypočítali L-ekvivalentní GEV rozdělání pro hodnoty $\xi = 0,1$ a $\xi = -0,1$. Na obr. 3.8.1 jsou navzájem porovnány hustoty pravděpodobnosti těchto tří rozdělání.

3.8.2 Změny charakteristik L-ekvivalentních GEV rozdělání s parametrem ξ

Byl testován vliv parametru ξ na charakteristiky L-ekvivalentních GEV rozdělání. Tyto závislosti jsou v následujícím textu zobrazeny ve formě grafů pro třídu L-ekvivalentních rozdělání s rozděláními na obr. 3.8.1. Pro L-ekvivalentní rozdělání na obr. 3.8.1 platí, že se modus snižuje s rostoucím parametrem ξ , ale graf na obr. 3.8.2 ukazuje, že modus dosahuje minima přibližně pro $\xi = 0,48$ a potom narůstá.

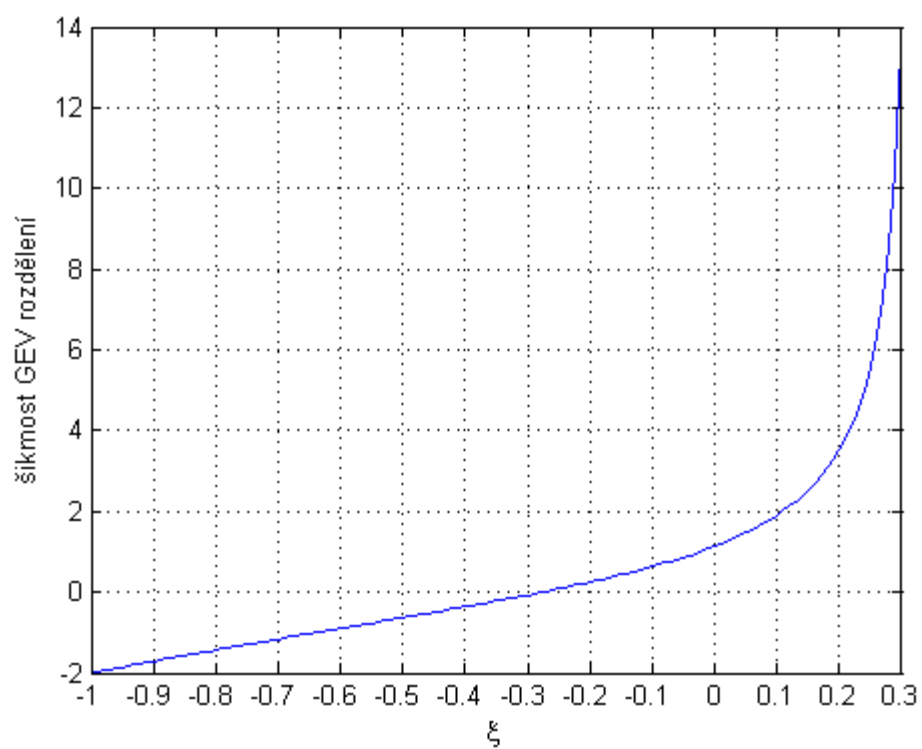


Obr. 3.8.2: Závislost modu GEV rozdělání L-ekvivalentních Gumbelovu rozdělání s parametry $\mu = 25$ m/s a $\sigma = 3$ m/s na parametru ξ .

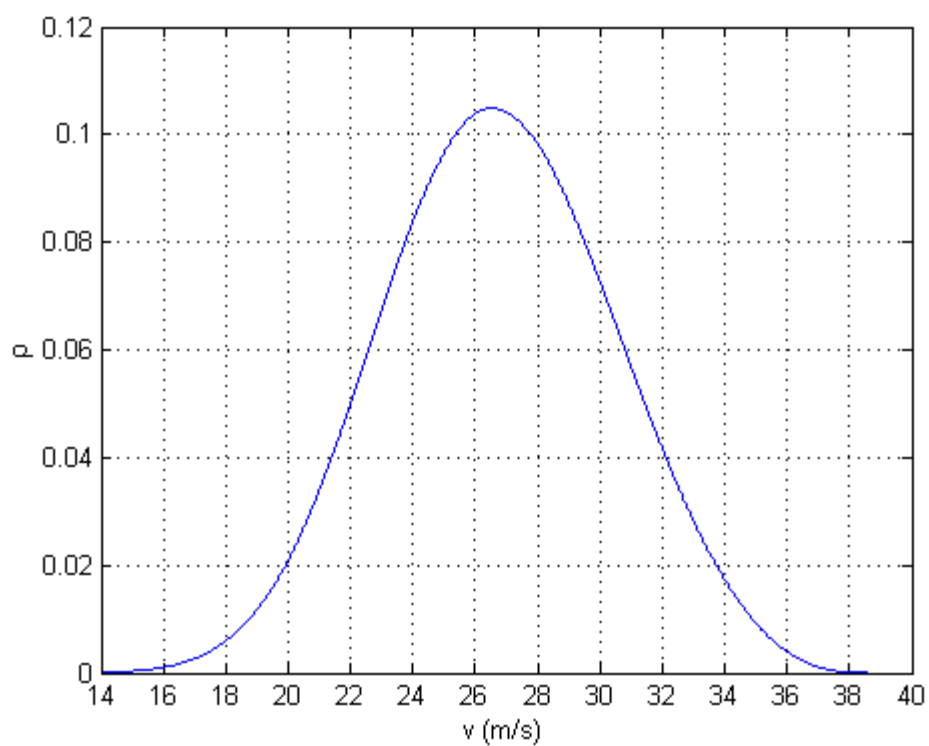
Graf závislosti šikmosti na ξ je zobrazen na obr. 3.8.3. Tato závislost je rostoucí a pro ξ blíží se k 1/3 jde do nekonečna. Nulová šikmost odpovídá hodnotě ξ přibližně -0,2776. Hustota pravděpodobnosti tohoto GEV rozdělení s nulovou šikmostí, které je L-ekvivalentní s rozdělením z obr. 3.8.1, je zobrazena na obr. 3.8.4. Toto rozdělení je velmi podobné normálnímu rozdělení.

Pro výpočet modu a šikmosti GEV rozdělení byly využity vztahy uvedené v monografii Reiss et Thomas (2001).

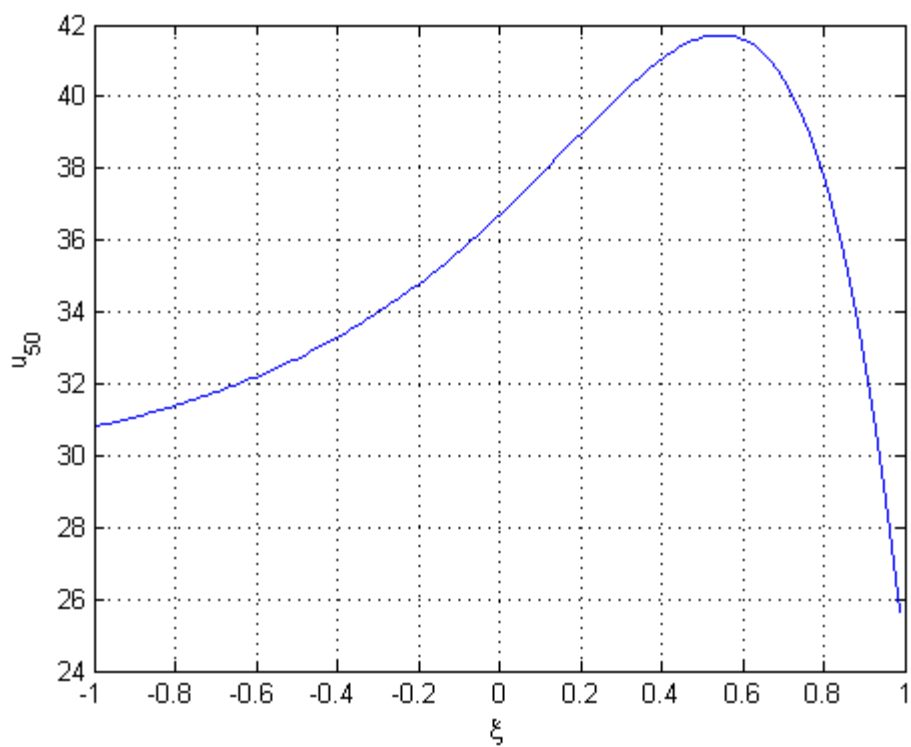
Na obr. 3.8.5 je závislost rychlosti větru s dobou návratu 50 let u_{50} na ξ GEV rozdělení L-ekvivalentním rozdělením z obr. 3.8.1. Podle očekávání se u_{50} na velké části grafu zvyšuje se zvyšující se hodnotou ξ , a tedy s těžknoucím pravým chvostem rozdělení. Toto platí jen přibližně do hodnoty 0,55 a poté dochází k prudkému poklesu. Pro vysvětlení tohoto jevu jsme porovnali hustoty pravděpodobnosti našeho výchozího Gumbelova rozdělení a L-ekvivalentního GEV rozdělení s hodnotou parametru $\xi = 0,9$. Hustoty jsou zobrazeny na obr. 3.8.6 a na výřezu 3.8.7. Z nich je patrné, že hustota pravděpodobnosti GEV rozdělení s hodnotou $\xi = 0,9$ ve srovnání s rozdělením s $\xi = 0$ klesá k nule podstatně rychleji hned za módem rozdělení, který je v obou případech v blízkosti hodnoty 25 m/s. Při rychlostech větru nad přibližně 45 m/s se ale situace obrátí a se zvyšující se hodnotou rychlosti se rozdíl stále zvětšuje. V této části rozdělení je příliš málo dat, než aby výrazněji ovlivnily u_{50} , ale jejich vliv stačí na to, aby byly hodnoty např. druhého a vyšších klasických momentů nekonečné.



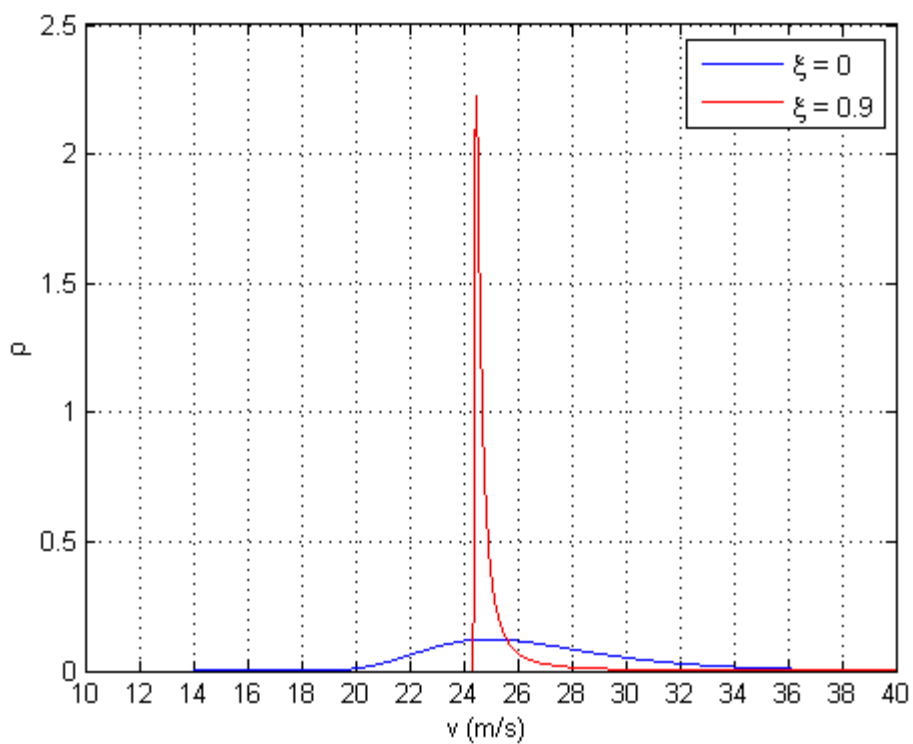
Obr. 3.8.3: Závislost šikmosti GEV rozdělení na parametru ξ .



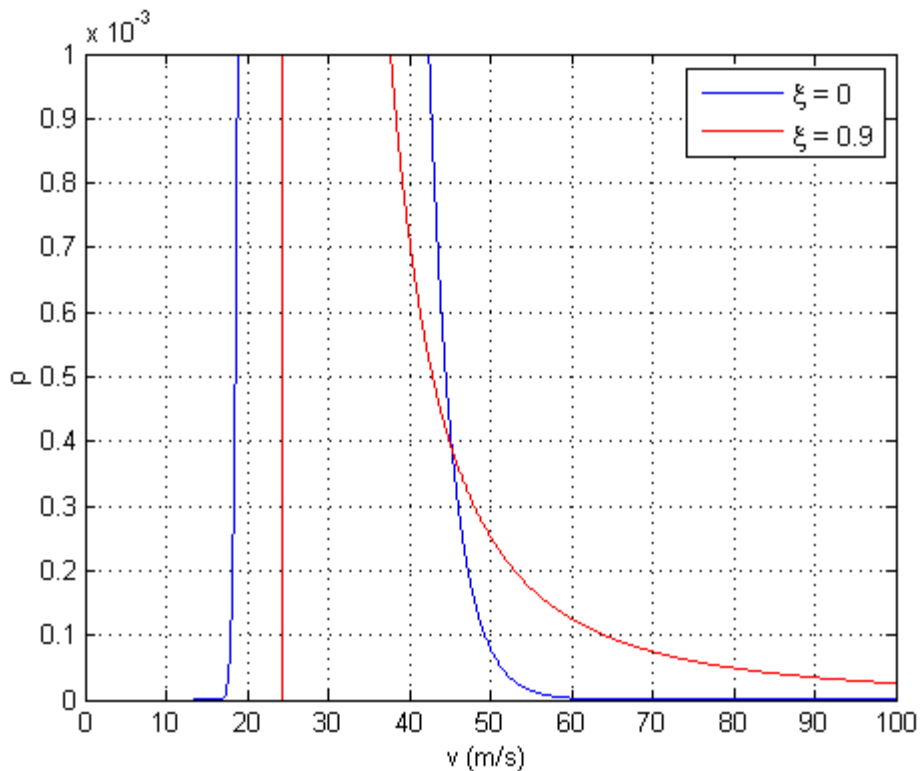
Obr. 3.8.4: Hustota pravděpodobnosti ρ GEV rozdělení s nulovou šikmostí.



Obr. 3.8.5: Závislost u_{50} na parametru ξ pro L-ekvivalentní rozdělení z obr. 3.8.1.



Obr. 3.8.6: Porovnání hustot pravděpodobnosti ρ pro L-ekvivalentních GEV rozdělení s hodnotami $\xi = 0$ a $\xi = 0,9$.



3.8.7: Totéž jako obr. 3.8.6 pro jiné měřítko.

3.8.3 Příprava umělých dat

Výchozím rozdělením pro přípravu umělých dat bylo Gumbelovo rozdělení s parametry $\mu = 25$ m/s a $\sigma = 3$ m/s. Z tohoto rozdělení byly vygenerovány řady o padesáti prvcích, které představovaly padesátileté řady ročních maxim rychlosti větru.

Od tohoto základního postupu přípravy dat byly odvozeny dva typy modifikací. První z nich spočívá v použití L-ekvivalentních rozdělení k tomuto základnímu rozdělení s parametry $\xi = -0,1$ a $0,1$. Hustoty pravděpodobnosti těchto rozdělení a výchozího jsou na obr. 3.8.1 a jejich parametry spolu s hodnotami s dobou návratu 50 let u_{50} v tabulce 3.8.1. V této podkapitole se dobou návratu rozumí doba návratu ročních maxim. Pro každé z těchto rozdělení byl vygenerován milion padesátiletých řad.

Druhý typ modifikace spočíval v simulaci chybných měření přičtením hodnoty 3 m/s nebo 6 m/s k nejvyšší simulované hodnotě rychlosti větru v každé padesátileté řadě vygenerované Gumbelovým rozdělením. Takto velká chyba je například v řadě nejvyšších denních nárazů větru velmi těžko rozpoznatelná.

ξ	μ (m/s)	σ (m/s)	u_{50} (m/s)
-0,1	25,144	3,264	35,69
0	25,000	3,000	36,71
0,1	24,871	2,711	37,81

Tabulka 3.8.1: Parametry a u_{50} L-ekvivalentních rozdělení z obr. 3.8.1.

3.8.4 Aproximovaná rozdělení a označení výsledků

Pro aproximaci rozdělení vygenerovaných dat byly použity tři typy teoretických rozdělení: (i) Gumbelovo rozdělení, (ii) GEV rozdělení, (iii) GEV rozdělení, u kterého jsme předpokládali, že známe správnou hodnotu parametru ξ . K aproximacím byla použita metoda L-momentů a ML.

Každý výsledek byl označen ve tvaru A_B_C_D, kde význam a možné hodnoty jednotlivých částí jsou následující:

A: Metoda aproximace, lm = L-momenty, ml = ML

B: Parametr tvaru rozdělení použitého pro přípravu dat (dále generující rozdělení),
-01 = -0,1, 00 = 0, 01 = 1

C: Chyba přidaná k nejvyšší hodnotě v řadě, 0 = bez přidané chyby, 3 = 3 m/s,
6 = 6 m/s

D: Rozdělení použité k aproximaci hodnot (dále aproximované rozdělení), GU = Gumbelovo, GEVV = GEV, GEVS = GEV s pevně nastaveným ξ .

Označení XX u kterékoli z částí A, B, C, D znamená, že tato část může nabývat jakoukoli hodnotu, pro kterou byla provedena aproximace. Každou unikátní zkoumanou kombinaci A_B_C_D budeme označovat jako jeden případ.

3.8.5 Výsledky aproximací

Provedením popsané aproximace bylo pro každou z možností XX_XX_XX_XX získáno milion odhadů parametrů μ a σ a, v případě aproximace GEV rozdělení, i ξ , a z těchto hodnot byl vztahem 3.1.2 získán milion odhadů u_{50} . Z těchto odhadů byl spočítán průměr a medián, v případě u_{50} ve formě systematické chyby (označeno na obrázcích bias), směrodatná odchylka a histogramy. Byly také spočítány korelace odhadů jednotlivých parametrů mezi odpovídajícími soubory lm_XX_XX_XX a

ml_XX_XX_XX. Histogramy jsou z důvodu stručnosti v této práci uvedeny jen pro případy XX_00_0_GU a XX_00_0_GEVV (obr. 3.8.8 až 3.8.14). Ostatní výsledky jsou uvedeny ve formě tabulek (tabulka 3.8.2 až 3.8.8). Chyby hodnot uvedených v tabulkách jsou vzhledem k počtu vygenerovaných dat zanedbatelné.

soubor výsledků	průměr (m/s)	medián (m/s)	STD (m/s)
lm_00_0_GU	25,00	24,99	0,49
ml_00_0_GU	25,02	25,01	0,49
lm_00_0_GEVV	25,01	25,00	0,48
ml_00_0_GEVV	25,04	25,03	0,49
lm_-01_0_GU	25,00	24,99	0,48
ml_-01_0_GU	25,00	24,99	0,48
lm_-01_0_GEVV	25,14	25,14	0,48
ml_-01_0_GEVV	25,16	25,16	0,48
lm_-01_0_GEVV	25,16	25,15	0,52
ml_-01_0_GEVV	25,19	25,18	0,52
lm_01_0_GU	25,00	24,99	0,42
ml_01_0_GU	25,04	25,03	0,42
lm_01_0_GEVV	24,87	24,86	0,41
ml_01_0_GEVV	24,89	24,88	0,41
lm_01_0_GEVV	24,89	24,87	0,44
ml_01_0_GEVV	24,90	24,89	0,44
lm_00_3_GU	25,01	25,00	0,45
ml_00_3_GU	25,04	25,03	0,45
lm_00_3_GEVV	24,97	24,96	0,48
ml_00_3_GEVV	24,99	24,97	0,48
lm_00_6_GU	25,02	25,01	0,45
ml_00_6_GU	25,06	25,04	0,45
lm_00_6_GEVV	24,93	24,92	0,48
ml_00_6_GEVV	24,95	24,93	0,47

Tabulka 3.8.2: Výsledky odhadů parametru μ .

soubor výsledků	průměr (m/s)	medián (m/s)	STD (m/s)
lm_00_0_GU	3,000	2,981	0,383
ml_00_0_GU	2,954	2,945	0,331
lm_00_0_GEVV	2,993	2,983	0,372
ml_00_0_GEVV	2,938	2,931	0,352
lm_-01_0_GU	3,000	2,992	0,331
ml_-01_0_GU	3,130	3,124	0,338
lm_-01_0_GEVS	3,264	3,255	0,360
ml_-01_0_GEVS	3,212	3,206	0,333
lm_-01_0_GEVV	3,256	3,248	0,385
ml_-01_0_GEVV	3,201	3,196	0,371
lm_01_0_GU	2,999	2,962	0,463
ml_01_0_GU	2,790	2,773	0,350
lm_01_0_GEVS	2,711	2,677	0,419
ml_01_0_GEVS	2,669	2,657	0,322
lm_01_0_GEVV	2,706	2,691	0,358
ml_01_0_GEVV	2,652	2,642	0,332
lm_00_3_GU	3,087	3,068	0,383
ml_00_3_GU	2,993	2,984	0,331
lm_00_3_GEVV	2,977	2,966	0,373
ml_00_3_GEVV	2,934	2,927	0,349
lm_00_6_GU	3,173	3,154	0,383
ml_00_6_GU	3,030	3,022	0,331
lm_00_6_GEVV	2,960	2,949	0,374
ml_00_6_GEVV	2,935	2,928	0,348

Tabulka 3.8.3: Výsledky odhadů parametru σ .

soubor výsledků	průměr	medián	STD
lm_00_0_GEVV	-0,009	-0,006	0,110
ml_00_0_GEVV	-0,009	-0,007	0,119
lm_-01_0_GEVV	-0,106	-0,103	0,106
ml_-01_0_GEVV	-0,114	-0,111	0,113
lm_01_0_GEVV	0,087	0,089	0,117
ml_01_0_GEVV	0,095	0,096	0,125
lm_00_3_GEVV	0,028	0,031	0,103
ml_00_3_GEVV	0,035	0,031	0,098
lm_00_6_GEVV	0,063	0,065	0,097
ml_00_6_GEVV	0,069	0,062	0,086

Tabulka 3.8.4: Výsledky odhadů parametru ξ .

soubor výsledků	bias průměru (m/s)	bias mediánu (m/s)	STD (m/s)
lm_00_0_GU	0,00	-0,08	1,66
ml_00_0_GU	-0,16	-0,21	1,49
lm_00_0_GEVV	0,02	-0,16	2,35
ml_00_0_GEVV	-0,11	-0,36	2,52
lm_-01_0_GU	1,02	0,98	1,40
ml_-01_0_GU	1,52	1,48	1,45
lm_-01_0_GEVS	0,00	-0,03	1,30
ml_-01_0_GEVS	-0,15	-0,17	1,24
lm_-01_0_GEVV	0,03	-0,06	1,80
ml_-01_0_GEVV	-0,23	-0,34	1,83
lm_01_0_GU	-1,11	-1,26	2,00
ml_01_0_GU	-1,88	-1,96	1,61
lm_01_0_GEVS	0,00	-0,17	2,17
ml_01_0_GEVS	-0,18	-0,25	1,76
lm_01_0_GEVV	-0,01	-0,34	3,07
ml_01_0_GEVV	0,08	-0,36	3,40
lm_00_3_GU	0,35	0,27	1,66
ml_00_3_GU	0,01	-0,04	1,49
lm_00_3_GEVV	0,76	0,59	2,36
ml_00_3_GEVV	0,75	0,49	2,42
lm_00_6_GU	0,70	0,62	1,66
ml_00_6_GU	0,17	0,12	1,49
lm_00_6_GEVV	1,50	1,34	2,36
ml_00_6_GEVV	1,54	1,27	2,35

Tabulka 3.8.5: Výsledky odhadů u_{50} .

dvojice parametrů	μ a σ	μ a u_{50}	σ a u_{50}
soubor výsledků			
lm_00_0_GU	0,232	0,480	0,965
ml_00_0_GU	0,305	0,565	0,958
lm_-01_0_GU	0,059	0,397	0,940
ml_-01_0_GU	0,101	0,423	0,944
lm_-01_0_GEVs	0,092	0,454	0,929
ml_-01_0_GEVs	0,132	0,504	0,923
lm_01_0_GU	0,367	0,541	0,981
ml_01_0_GU	0,478	0,666	0,973
lm_01_0_GEVs	0,325	0,489	0,984
ml_01_0_GEVs	0,448	0,625	0,978
lm_00_3_GU	0,232	0,480	0,965
ml_00_3_GU	0,305	0,565	0,958
lm_00_6_GU	0,232	0,480	0,965
ml_00_6_GU	0,305	0,565	0,958

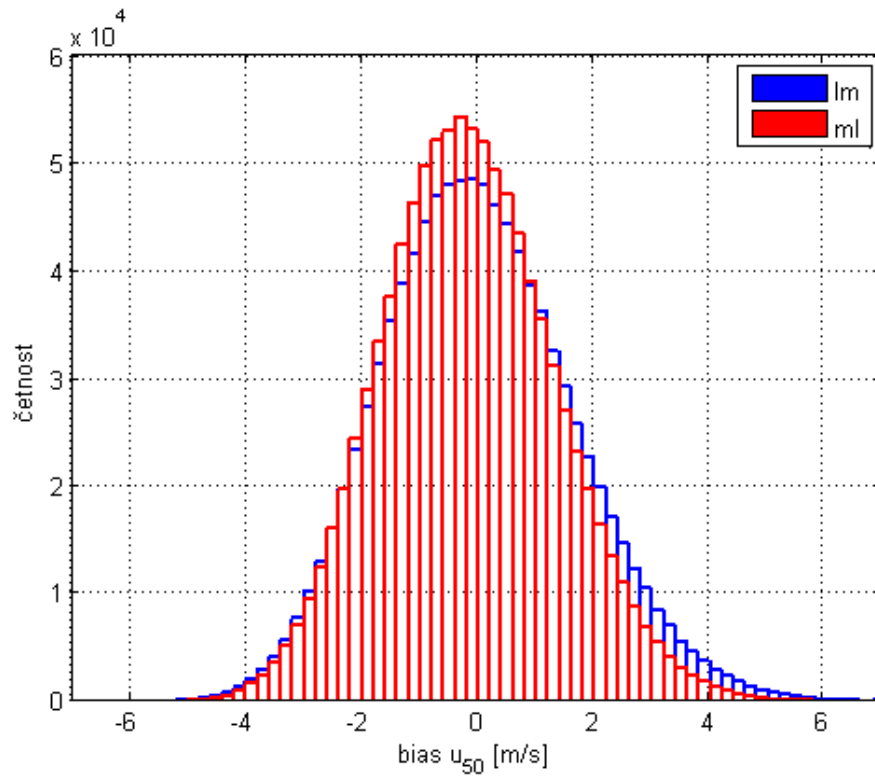
Tabulka 3.8.6: Korelace mezi odhady μ , σ a u_{50} pro aproximace Gumbelovým rozdělení a GEV rozdělením s pevným parametrem tvaru.

dvojice parametrů	μ a σ	μ a ξ	σ a ξ	μ a u_{50}	σ a u_{50}	ξ a u_{50}
soubor výsledků						
lm_00_0_GEVV	0,378	-0,362	-0,360	0,058	0,306	0,744
ml_00_0_GEVV	0,394	-0,382	-0,337	0,008	0,259	0,785
lm_-01_0_GEVV	0,260	-0,377	-0,423	0,049	0,284	0,701
ml_-01_0_GEVV	0,283	-0,394	-0,428	0,023	0,230	0,731
lm_01_0_GEVV	0,484	-0,351	-0,307	0,057	0,304	0,783
ml_01_0_GEVV	0,491	-0,367	-0,243	0,000	0,289	0,823
lm_00_3_GEVV	0,377	-0,357	-0,409	0,069	0,293	0,722
ml_00_3_GEVV	0,379	-0,341	-0,366	0,075	0,304	0,744
lm_00_6_GEVV	0,374	-0,352	-0,451	0,081	0,285	0,698
ml_00_6_GEVV	0,369	-0,316	-0,401	0,120	0,332	0,702

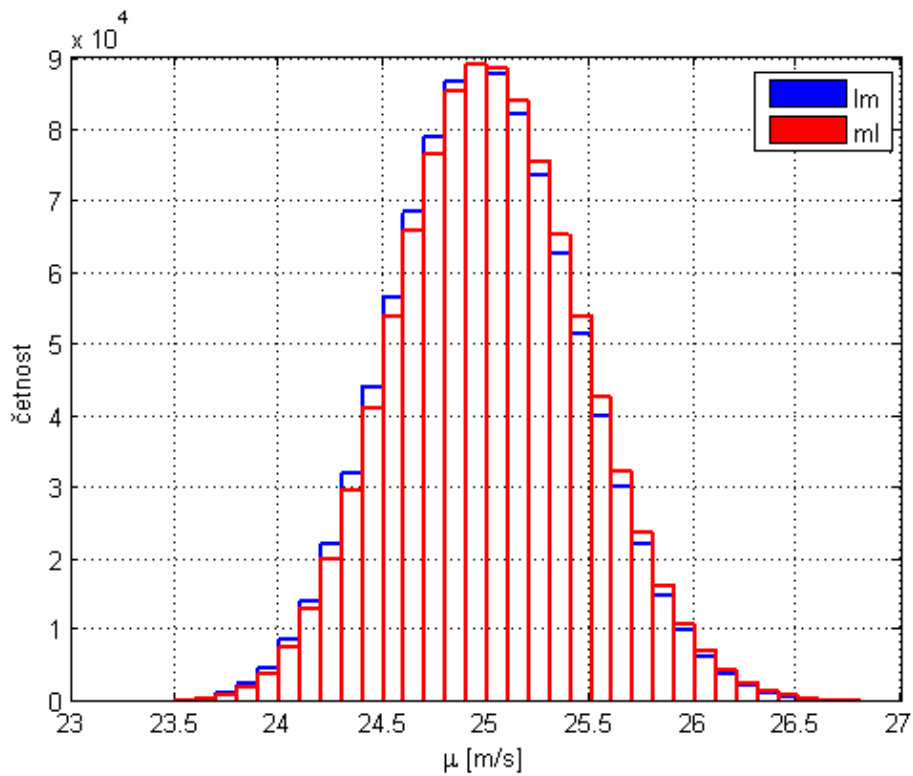
Tabulka 3.8.7: Korelace mezi odhady μ , σ , ξ a u_{50} pro aproximace GEV rozdělením.

soubor výsledků	μ	σ	ξ	u_{50}
XX_00_0_GU	0,998	0,878	-	0,910
XX_00_0_GEVV	0,993	0,950	0,931	0,966
XX_-01_0_GU	0,997	0,826	-	0,868
XX_-01_0_GEVV	0,996	0,941	-	0,962
XX_-01_0_GEVV	0,992	0,965	0,932	0,965
XX_01_0_GU	0,998	0,922	-	0,939
XX_01_0_GEVV	0,999	0,784	-	0,825
XX_01_0_GEVV	0,989	0,930	0,903	0,947
XX_00_3_GU	0,998	0,880	-	0,912
XX_00_3_GEVV	0,992	0,940	0,902	0,958
XX_00_6_GU	0,998	0,883	-	0,913
XX_00_6_GEVV	0,990	0,931	0,870	0,948

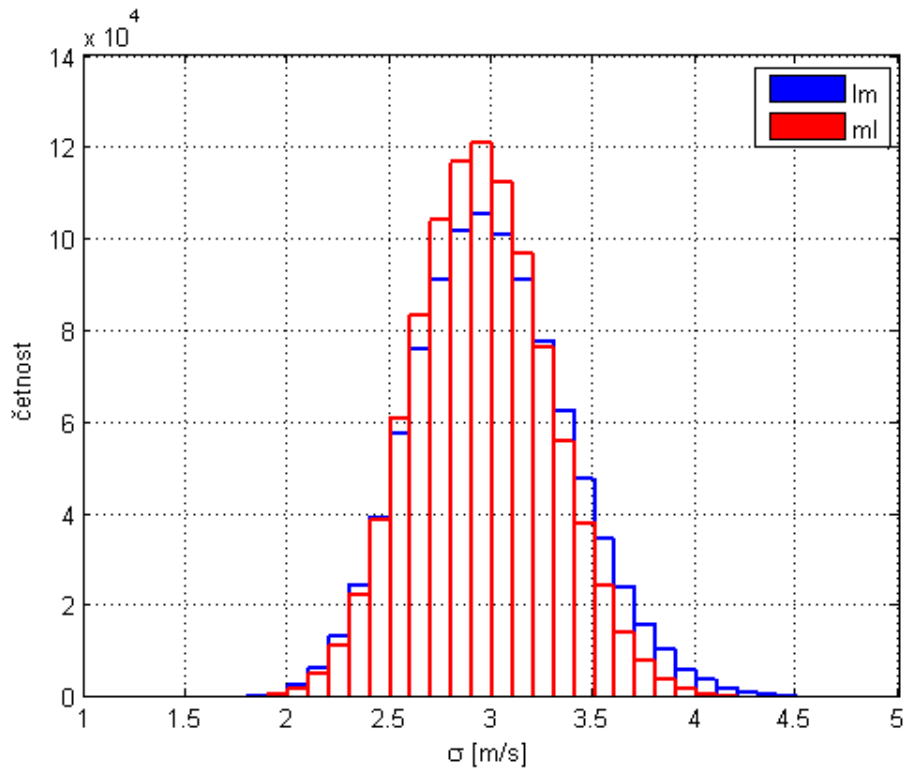
Tabulka 3.8.8: Korelace odhadů mezi metodami L-momentů a ML.



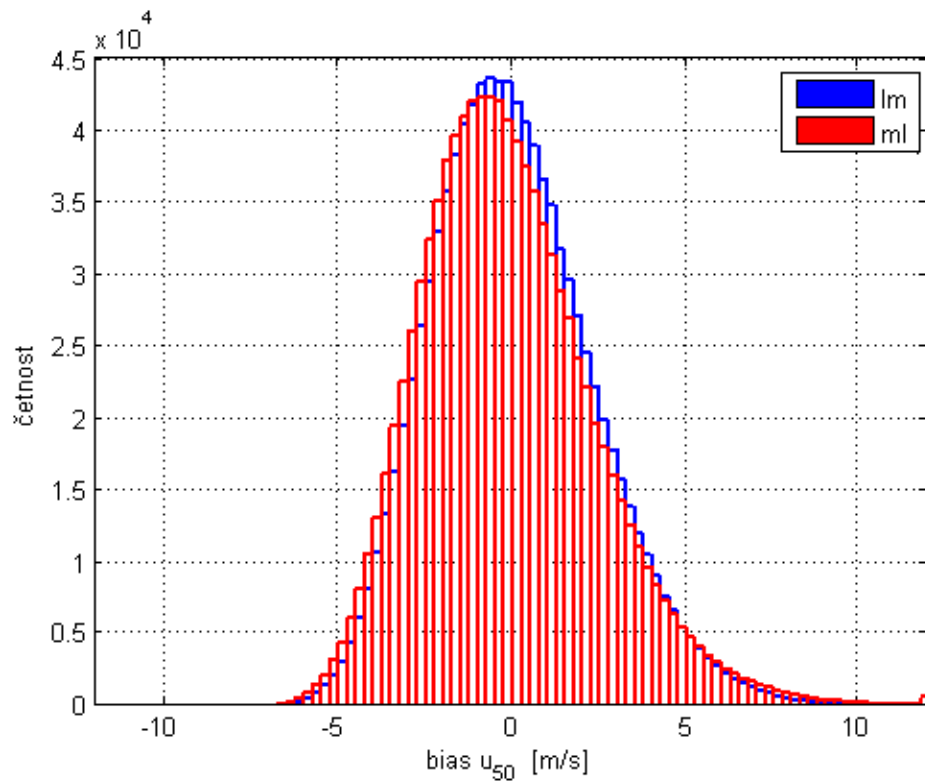
Obr. 3.8.8: Histogramy systematické chyby odhadu u_{50} případů XX_00_0_GU.



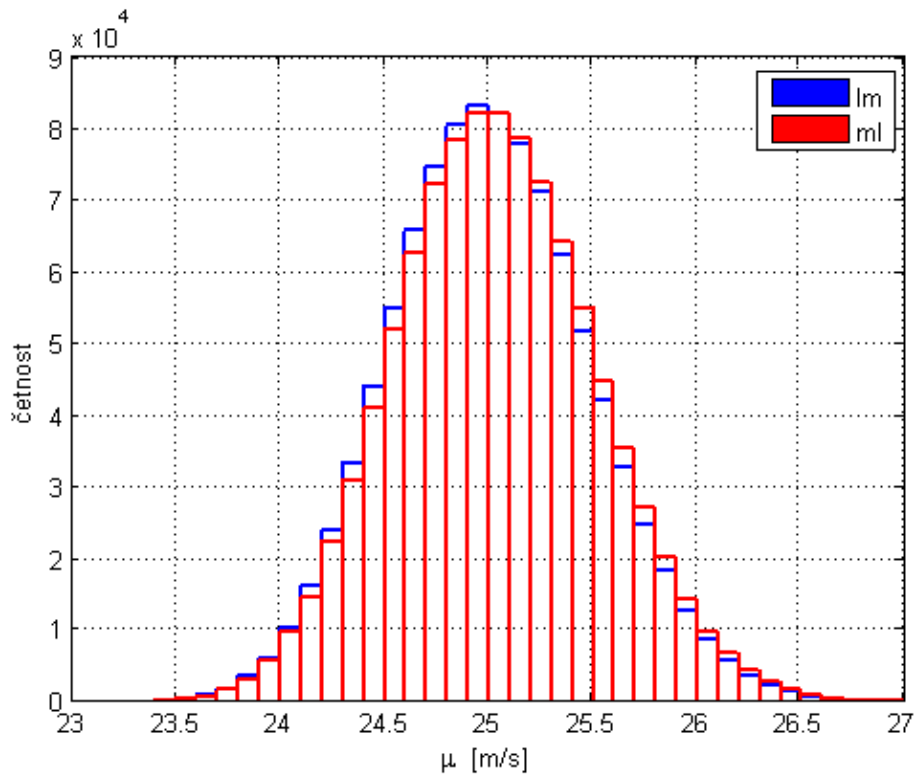
Obr. 3.8.9: Histogramy odhadu parametru μ případů XX_00_0_GU.



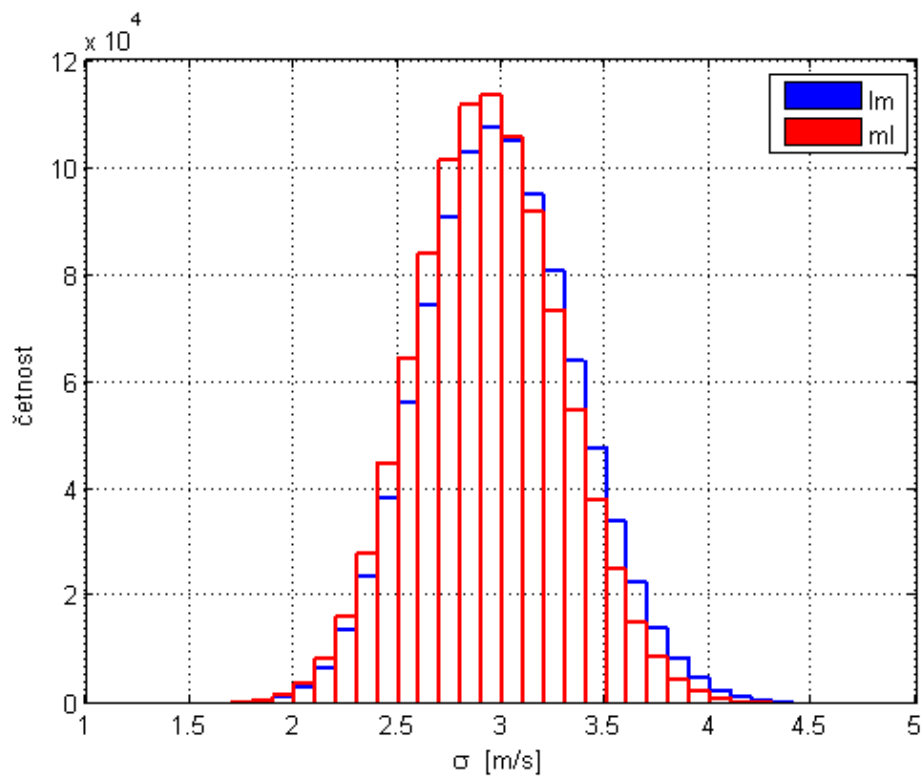
Obr. 3.8.10: Histogramy odhadu parametru σ případů XX_00_0_GU.



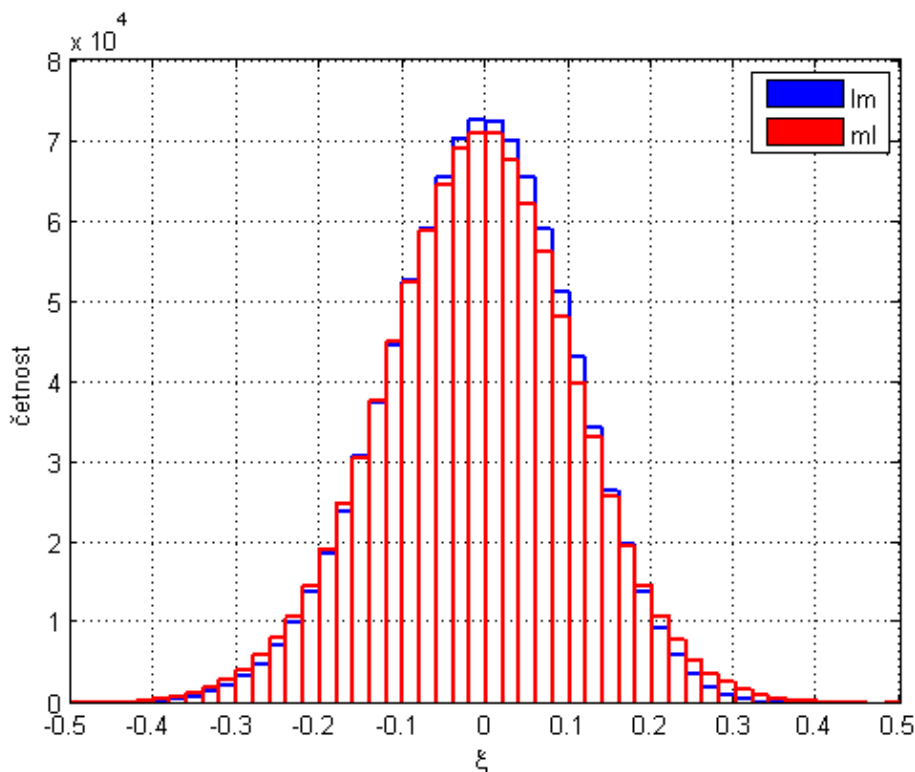
Obr. 3.8.11: Histogramy systematické chyby odhadu u_{50} případů XX_00_0_GEVV.



Obr. 3.8.12: Histogramy odhadu parametru μ případů XX_00_0_GEVV.



3.8.13: Histogramy odhadu parametru σ případů XX_00_0_GEVV.



Obr. 3.8.14: Histogramy odhadu parametru ξ případů XX_00_0_GEVV.

3.8.6 Shrnutí výsledků aproximací

V této sekci budou shrnuty a diskutovány pouze nejdůležitější závěry z provedených testovacích výpočtů. Doplňující informace lze nalézt v přílohách C a D.

Odhad u_{50} je v případě dvouparametrických metod, Gumbelova rozdělení a GEV rozdělení s pevným parametrem tvaru, dán především odhadem parametru σ , zatímco v případě tříparametrického GEV rozdělení odhadem parametru tvaru ξ . V příloze C je ukázáno, že tento vliv narůstá s rostoucí dobou návratu.

Mezi výsledky získanými metodou L-momentů a ML nejsou významné rozdíly. Při použití dvouparametrických metod se správným parametrem ξ GEV rozdělení (včetně Gumbelova rozdělení jako speciálního případu GEV rozdělení) i v případě použití tříparametrického GEV rozdělení vede metoda ML k záporné systematické chybě parametru σ , což se projeví i v záporné systematické chybě odhadu rychlosti větru s dobou opakování 50 let.

Při aplikaci Gumbelova rozdělení pro aproximaci rozdělení dat generovaných rozdělením s nenulovým parametrem ξ vede metoda ML k větší systematické chybě odhadu u_{50} než metoda L-momentů.

V případě aplikace Gumbelova rozdělení vykazuje metoda ML podstatně menší citlivost na přidání chybné vysoké hodnoty do souboru ročních maxim než metoda L-momentů. Naopak při aplikaci GEV rozdělení s třemi volnými parametry je metoda ML k této chybě citlivější.

Použití GEV rozdělení vykazuje vyšší směrodatnou odchylku určení u_{50} ve srovnání s Gumbelovým rozdělením nebo GEV rozdělením s pevným parametrem tvaru. Dále vykazuje podstatně vyšší citlivost odhadu na přidání chybné vysoké hodnoty do souboru dat. Jedinou výhodou použití GEV rozdělení se třemi volnými parametry oproti dvouparametrickým metodám tak představuje eliminace systematické chyby odhadu extrémních rychlostí větru v případě aproximace dvouparametrickým rozdělením s jiným parametrem ξ než mají data.

Odhady jednotlivých parametrů Gumbelova a GEV rozdělení nelze považovat za nezávislé. Významem odhadu jednotlivých parametrů pro odhad rychlosti s danou dobou návratu se zabývá příloha C.

Pomocí výsledků z této kapitoly je v příloze D navržen způsob odhadu parametru ξ GEV rozdělení založený na Bayesovském přístupu.

3.9 Analýza aproximace extrémních dat GEV a Gumbelovým rozdělením metodou L-momentů a ML na naměřených datech

3.9.1 Úvod

Tato kapitola se zabývá analýzou aproximace naměřených maximálních rychlostí větru pomocí GEV a Gumbelova rozdělení se zaměřením na odhad rychlostí s dobou opakování 50 let (u_{50}). Aproximované řady jsou vygenerovány metodou bootstrap (Wilks, 2011). Naměřená data nemusí pocházet z GEV nebo Gumbelova rozdělení, jak plyne z podkapitoly 3.2. Přesto rozdělení GEV a Gumbelovo použijeme, protože je to u tohoto typu dat obvyklé (An et Pandey, 2005). Stejně jako v podkapitole 3.8 byly pro odhad parametrů rozdělení použity metody L-momentů a ML.

Cílem této kapitoly je odpovědět na následující otázky:

1. Jsou statistické charakteristiky odhadů parametrů GEV a Gumbelova rozdělení (rozptyly odhadu jednotlivých parametrů, korelace mezi odhady vypočtenými metodou L-momentů a ML a korelace mezi odhady jednotlivých parametrů (včetně u_{50})) obdobné jako v případě umělých dat?
2. Projeví se přidání chybné hodnoty do časové řady stejně jako v případě uměle vygenerovaných dat?

3.9.2 Data

Byla použita data nejvyšších denních nárazů větru z let 1961-2009 využitá již v podkapitole 2.3. Data byla z 29 stanic, na kterých bylo k dispozici alespoň 20 let úplných měření bez výpadků. Neúplné roky byly ze zpracování vyřazeny. Na data byla aplikována metoda homogenizace popsaná v podkapitole 2.3 a data byla rozdělena do 8 sektorů stejných jako v uvedené podkapitole. Z let, za která bylo na dané stanici k dispozici měření nejvyšších denních nárazů větru, byla pro každou stanici a sektor vybrána roční maxima. Z těchto hodnot bylo pro každou kombinaci stanice a sektoru metodou bootstrap vygenerováno 10 000 souborů řad o padesáti prvcích, které byly interpretovány jako 50 let dlouhé řady ročních maxim.

Pro aproximaci metodou ML byl použit software Matlab a jeho procedura `gevfit`. Hledání optimálních hodnot parametrů pomocí kritéria maximální věrohodnosti představuje numerické řešení minimalizace nelineárního funkcionálu. I když použité algoritmy ve velké většině případů konvergují, obecně tuto konvergenci nelze zajistit (Mantel et Mayers, 1971). Kromě toho může algoritmus konvergovat k lokálnímu minimu, a tedy k zcela jiným hodnotám parametrů než odpovídají globálnímu minimu. Již při výpočtech s uměle generovanými daty existovaly případy, kdy metoda nekonvergovala nebo dala zjevně nesmyslné hodnoty parametrů. V případě naměřených dat k těmto chybným odhadům docházelo výrazně častěji. V případě umělých dat se problém podařilo vyřešit individuálně nastavenými parametry minimalizačního procesu. U naměřených dat tento postup nebyl vždy úspěšný. Tyto případy identifikovala samotná funkce `gevfit`, která varovala, že výpočet probíhá nestandardně.

Ukázalo se, že problémy s konvergencí nastávají především u stanic s kratšími dobami pozorování a to tím, že v generovaných řadách ročních maxim se často opakují stejné hodnoty. S obdobnými problémy se setkal M. Kašpar při zpracování ročních maxim teploty (osobní komunikace). Proto bylo ke každé hodnotě v řadách

vygenerovaných metodou bootstrap přidáno náhodné číslo s normálním rozdělením s nulovou střední hodnotou a směrodatnou odchylkou 0,2 m/s. Protože je tato směrodatná odchylka podstatně menší než chyby měření, nemůže tato změna dat viditelně ovlivnit rozdělení ročních maxim na jednotlivých stanicích. Aby nedošlo k nechtěným ovlivněním, byly tyto upravené hodnoty použity i pro metody odhadu, při kterých nebyly problémy s konvergencí (aproximace Gumbelovým a GEV rozdělením pomocí L-momentů a Gumbelovým rozdělením pomocí metody ML). Popsaná úprava skutečně konvergenci odhadů parametrů GEV rozdělení metodou ML podstatně zlepšila. Problémy se po úpravě vyskytly v méně než půl promile případů.

I tento velmi malý počet případů by však mohl ovlivnit výsledné statistiky, pokud by tyto odhady parametrů GEV rozdělení byly zásadně chybné. Protože jejich vyloučení by mohlo ovlivnit systematickou chybu výsledků, byly tyto případy nahrazeny odhady metodou L-momentů.

3.9.3 Výsledky a jejich diskuse

V této podkapitole používáme označení metod aproximace obdobné označení zavedenému v podkapitole 3.8:

gu_lm pro aproximaci Gumbelovým rozdělením metodou L-momentů,

gu_ml pro aproximaci Gumbelovým rozdělením metodou ML,

gev_lm pro aproximaci GEV rozdělením metodou L-momentů,

gev_ml pro aproximaci GEV rozdělením metodou ML.

Pro orientační srovnání jednotlivých metod byly vyhledány kombinace stanice a sektoru s nejnižší a nejvyšší hodnotou průměru a mediánu odhadu u_{50} pro jednotlivé metody výpočtu. Výsledky jsou zobrazeny v tabulce 3.9.1 (průměr) a 3.9.2 (medián).

Tyto tabulky ukazují, že hodnoty maximálních a minimálních průměrů a mediánů se příliš neliší, což ukazuje na značnou míru symetrie odhadu u_{50} . Zajímavá je skutečnost, že aproximace Gumbelovým rozdělením vede k vyšším maximálním průměrných i mediánovým odhadům extrémní rychlosti větru než interpolace GEV rozdělením zvláště v případě metody ML. To dokládá tabulka 3.9.3, která obsahuje stanici a sektory s maximální průměrnou směrodatnou odchylkou odhadů u_{50} . Z ní je zřejmé, že na rozdíl od uměle vygenerovaných dat (podkapitola 3.8) dává metoda ML na naměřených datech v některých případech podstatně větší směrodatné odchylky

odhadů extrémních rychlostí větru než metoda L-momentů. To platí jak pro aproximaci Gumbelovým, tak především GEV rozdělením.

metoda	min stanice	min sektor	min hodnota	max stanice	max sektor	Max hodnota
gu_lm	O1PORU01	1	15,02	U1MILE01	7	50,40
gu_ml	O1PORU01	1	14,63	U1MILE01	7	55,31
gev_lm	U1KOPI01	3	14,12	U1MILE01	6	47,82
gev_ml	O1PORU01	1	14,63	O1LYSA01	6	48,25

Tabulka 3.9.1: Minimální a maximální průměry odhadů u_{50} v m/s.

metoda	min stanice	min sektor	min hodnota	max stanice	max sektor	Max hodnota
gu_lm	O1PORU01	1	15,01	U1MILE01	7	50,40
gu_ml	O1PORU01	1	14,60	U1MILE01	7	55,67
gev_lm	U1KOPI01	3	14,12	U1MILE01	6	47,86
gev_ml	O1PORU01	1	14,61	U1MILE01	6	47,21

Tabulka 3.9.2: Minimální a maximální mediány odhadů u_{50} v m/s.

metoda	stanice	sektor	hodnota
gu_lm	L3CHEB01	2	2,80
gu_ml	O7KRNO01	5	4,00
gev_lm	H3HRAD01	2	4,63
gev_ml	O1LYSA01	6	14,86

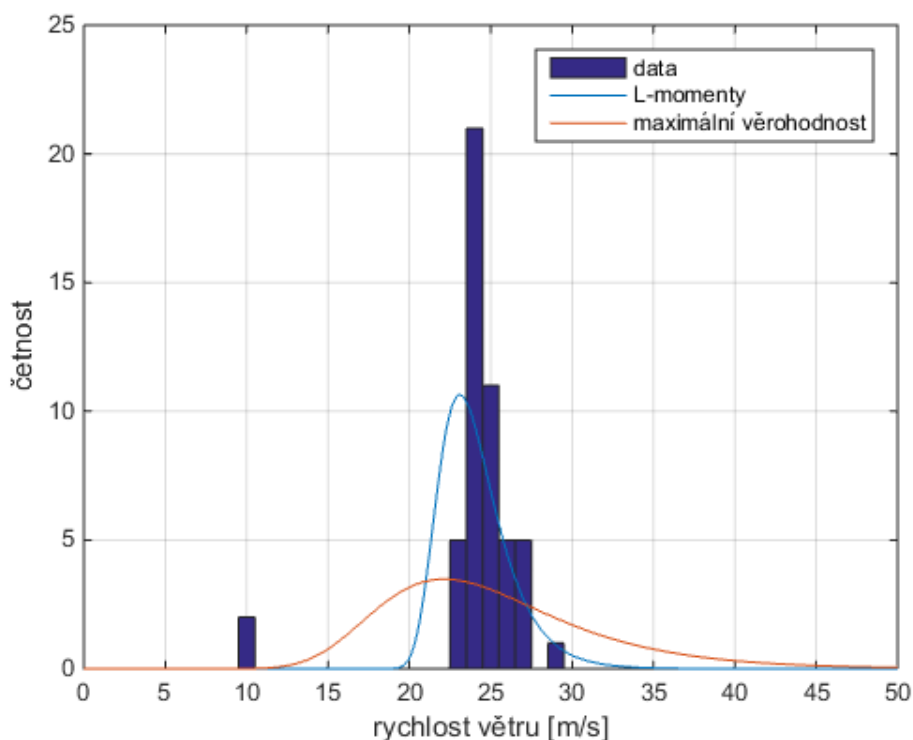
Tabulka 3.9.3: Maximální průměrné směrodatné odchytky odhadů u_{50} v m/s.

Tyto výsledky ukazují, že mezi metodami L-momentů a ML jsou větší rozdíly než v případě uměle vygenerovaných dat. Pro objasnění byl pro Gumbelovo i GEV rozdělení vybrán případ s největším rozdílem R mezi těmito metodami, přičemž R byl definován vzorcem:

$$R = \max\left(\frac{u_{50}^{lm}}{u_{50}^{ml}}, \frac{u_{50}^{ml}}{u_{50}^{lm}}\right), \quad (3.9.1)$$

kde u_{50}^{lm} je nejvyšší rychlost větru s dobou návratu 50 let určená aproximací metodou L-momentů a u_{50}^{ml} je nejvyšší rychlost větru s dobou návratu 50 let určená aproximací metodou ML.

Pro Gumbelovo rozdělení byla dosažena nejvyšší hodnota $R = 1,43$ v jednom souboru dat z 5. sektoru stanice Krnov. Histogram tohoto případu spolu s hustotami pravděpodobnosti aproximovaných rozdělení je zobrazen na obr. 3.9.1. Odhady parametru μ jsou 23,10 m/s v případě metody L-momentů a 22,14 m/s v případě metody ML, parametru σ 1,73 m/s v případě metody L-momentů a 5,29 m/s v případě metody ML. Zatímco odhady μ jsou podobné, odhady σ se výrazně liší. Hodnoty u_{50} odpovídající těmto odhadům parametrů jsou 29,85 m/s v případě metody L-momentů a 42,76 m/s v případě metody ML.



Obr. 3.9.1: Histogram případu s nejvyšším rozdílem R mezi Gumbelovým rozdělením vypočteným metodami L-momentů a ML spolu s hustotami pravděpodobnosti Gumbelova rozdělení aproximovaného těmito dvěma metodami.

Z grafu je zřejmé, že problém způsobila odlehlá nízká hodnota cca 10 m/s. Takto odlehlé nízké hodnoty mohou být v souboru dat způsobeny malou četností dat v určitém sektoru. S touto odlehlou hodnotou si mnohem lépe poradila metoda L-momentů než metoda ML. Zde připomeňme, že na umělých datech metoda L-momentů dávala horší výsledky, když se přidala naopak vysoká hodnota. To vysvětluje podstatně těžší pravý chvost Gumbelova rozdělení oproti jeho levému chvostu – viz tabulka 3.9.4. Tato tabulka obsahuje desítkové logaritmy hustoty pravděpodobnosti (pdf) Gumbelova rozdělení s parametry $\mu = 0$ a $\sigma = 1$, které jsou přímo úměrné penalizaci metody ML (podkapitola 3.7). V případě hodnot $\mu = 25$ m/s a $\sigma = 3$ m/s, použitých v kapitole 3.8, by hodnotě $x = -5$ odpovídala rychlost větru 10 m/s a hodnotě $x = 5$ rychlost větru 40 m/s. Hodnoty v tabulce ukazují, že metoda ML je v případě aproximace Gumbelova rozdělení mnohem citlivější na odlehlé nízké hodnoty než na odlehlé vysoké hodnoty.

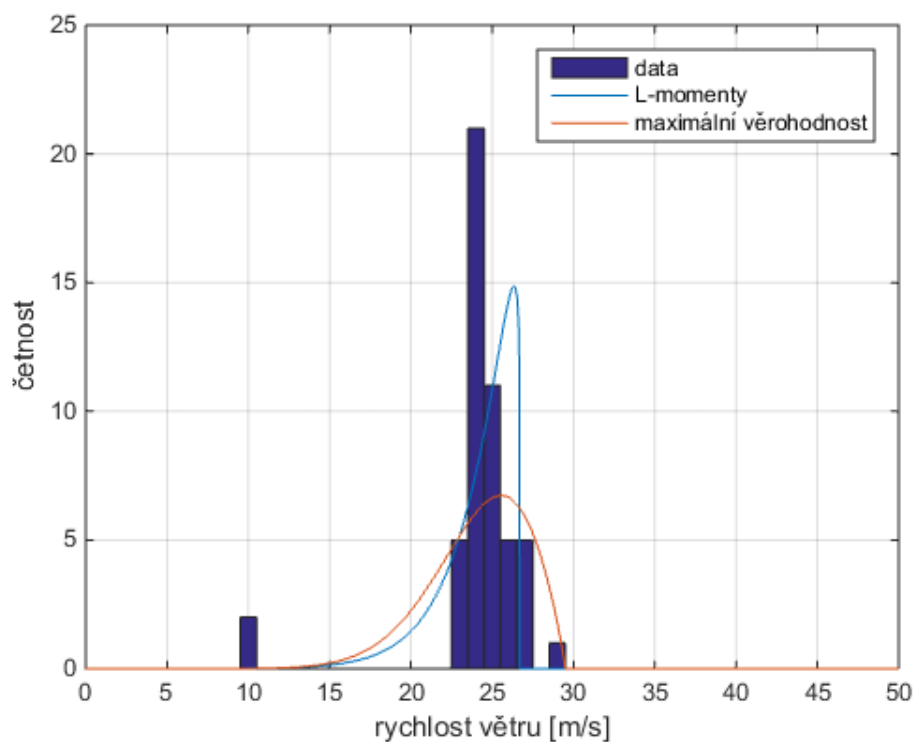
x	-5	-4	-3	-2	-1
log(pdf)	-62,28	-21,97	-7,42	-2,34	-0,75

0	1	2	3	4	5
-0,43	-0,59	-0,92	-1,32	-1,75	-2,17

Tabulka č. 3.9.4: Desítkové logaritmy pdf (x) Gumbelova rozdělení s parametry $\mu = 0$ a $\sigma = 1$.

Pro srovnání byla pro tento případ provedena i interpolace GEV rozdělením, která je zobrazena na obr. 3.9.2. Odhady parametrů μ , σ a ξ jsou při aproximaci metodou L-momentů 24,00 m/s, 2,42 m/s a -0,900 a při aproximaci metodou ML 23,43 m/s, 3,30 m/s a -0,545. Odlehlá hodnota vychýlila odhad parametru ξ do hlubokých záporných hodnot, a to podstatně více u metody L-momentů než u metody ML. Tím způsobila omezení pravého chvostu rozdělení nereálně nízkými hodnotami. V případě metody ML

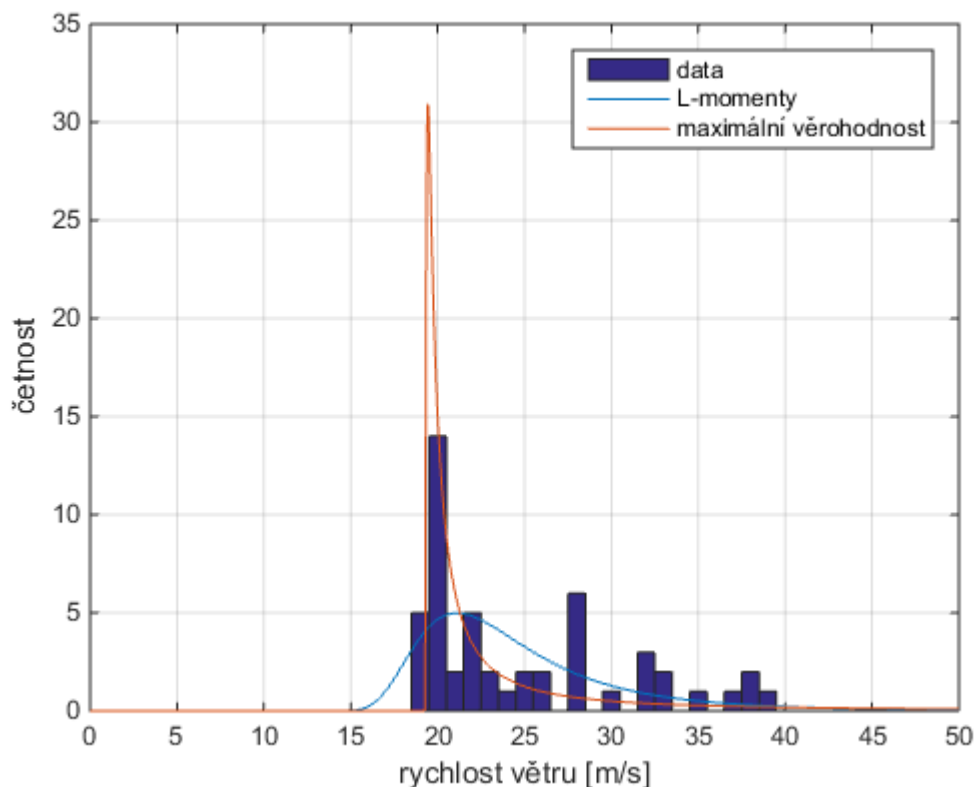
aproximace s těmito parametry vylučuje výskyt rychlostí větru vyšších než 29,5 m/s. V případě metody L-momentů je tato hranice dokonce 26,7 m/s. Je tedy dokonce znatelně nižší než nejvyšší hodnota rychlosti větru v daném souboru dat. V těchto případech s výrazně odlehlými hodnotami je vhodné používat aproximaci Gumbelovým rozdělením metodou L-momentů.



Obr. 3.9.2: Histogram případu s nejvyšším rozdílem R mezi Gumbelovým rozdělením aproximovaným metodami L-momentů a ML spolu s hustotami pravděpodobnosti GEV rozdělení aproximovaného těmito dvěma metodami.

V případě aproximace GEV rozdělením dosáhl rozdíl R nejvyšší hodnoty 11,53 pro jeden soubor dat ze 4. sektoru stanice Milešovka. Histogram tohoto případu spolu s hustotami pravděpodobnosti GEV rozdělení aproximovaných pomocí obou metod je zobrazen na obr. 3.9.3. Odhady parametrů μ , σ a ξ jsou při aplikaci metody L-momentů 21,78m/s, 3,76 m/s a 0,199 a při aplikaci metody ML 20,15 m/s, 1,46 m/s a 1,612. Odhady u_{50} jsou 43,96 m/s v případě metody L-momentů a 506,63 m/s v případě metody ML.

Odhad u_{50} metodou ML je zjevně chybný, což platí o celé aproximované křivce (obr. 3.9.3). Křivka aproximovaná metodou L-momentů vychází v tomto případě rozumně včetně odhadu u_{50} .



Obr. 3.9.3: Histogram případu s nejvyšším rozdílem R mezi GEV rozdělením aproximovaným metodami L-momentů a ML spolu s hustotami pravděpodobnosti GEV rozdělení aproximovaného těmito dvěma metodami.

Histogram rychlostí větru v této datové řadě zjevně neodpovídá předpokládanému GEV rozdělení, avšak má tvar, který u naměřených dat není neobvyklý. Při bližším pohledu si lze všimnout dvou jeho charakteristik:

1. většího počtu hodnot kolem rychlosti 20 m/s,
2. poměrně vysokého počtu hodnot na jeho pravém okraji.

GEV rozdělení aproximované metodou ML lze těmito dvěma charakteristikami vysvětlit, protože má výrazné maximum přibližně na hodnotě 20 m/s a těžký pravý chvost. Metoda ML nezohledňuje, že má toto rozdělení extrémní vlastnosti, jako např. nekonečný průměr a hodnoty dalších momentů (podkapitola 3.7). Naproti tomu metoda

L-momentů zaručuje, že průměr rozdělení bude stejný jako průměr prokládaných dat a hodnoty všech L- momentů rozdělení budou konečné (podkapitola 3.7).

Aproximace metodou ML je v tomto případě velmi citlivá na naměřené hodnoty v blízkosti jeho levého krajního bodu. Hustota GEV rozdělení aproximovaného metodou ML v blízkosti jeho levého krajního bodu extrémně rychle narůstá od nuly na poměrně vysokou hodnotu, což odpovídá nekonečné změně penalizace metody ML (podkapitola 3.7). Tento efekt byl ilustrován na odhadu u_{50} aproximací GEV rozdělením stejných dat jako na obr. 3.9.3 oběma metodami, pouze nejnižší rychlost větru v souboru byla snížena o 1 m/s z 19,32 m/s na 18,32 m/s. Tato změna způsobila v případě metody L-momentů snížení odhadu u_{50} o 0,11 m/s na 43,85 m/s zatímco v případě metody maximální věrohodnosti snížení odhadu u_{50} o 432,80 m/s na 73,83 m/s. To znamená, že malá změna ve vstupních datech může v určitých případech u metody ML způsobit obrovskou změnu ve výsledku, což se nestává u metody L-momentů.

Obdobně jako v podkapitole 3.8 byla pro metodu L-momentů spočítána citlivost odhadu u_{50} na zvýšení maximální hodnoty v souboru o 3 m/s a 6 m/s. Při aproximaci metodou Gumbelovým rozdělením metodou L-momentů došlo při zvýšení nejvyšší hodnoty v souboru dat o 3 m/s ke zvýšení odhadu u_{50} v průměru o 0,35 m/s. Při zvýšení nejvyšší hodnoty v souboru dat o 6 m/s došlo ke zvýšení odhadu u_{50} o 0,70 m. Tyto hodnoty jsou naprosto stejné, jako v případě uměle vytvořených dat (podkapitola 3.8).

3.9.4 Závěr

Analýza ukázala, že aproximace naměřených dat Gumbelovým a zejména GEV rozdělením je mnohem méně robustní než aproximace uměle vygenerovaných dat. Robustností zde máme na mysli malou citlivost metody na změny ve vstupních datech a její využitelnost i na data, která nesplňují teoretické předpoklady.

Zdaleka nejrobustnější ze zkoumaných metod aproximace extrémních dat se na naměřených datech ukázala aproximace Gumbelovým rozdělením metodou L-momentů. Proto byla tato metoda aplikována v kapitole 4.

Výsledky této analýzy zeslabují výsledky některých metod analýzy extrémních dat, které předpokládají, že tvar rozdělení extrémních hodnot je známý a neznámé jsou pouze jeho parametry. Např. metody testování, zda hodnoty spíše pocházejí z Gumbelova rozdělení nebo z GEV rozdělení s nenulovým parametrem tvaru popsané v Hosking (1984) apriori předpokládají, že hodnoty pocházejí z GEV rozdělení. Navíc

tyto metody často vycházejí z odhadů parametrů pomocí metody ML, která se v případě reálných dat ukázala jako málo robustní. Výsledky jejich mechanické aplikace na reálných datech proto mohou být zavádějící.

4 Výpočet mapy extrémního větru v ČR

4.1 Vztah mezi větrným klimatem a extrémním větrným klimatem na území ČR

V této podkapitole je na základě měření větru na profesionálních stanicích ČHMÚ odvozen vztah mezi extrémním větrným klimatem (EVK) a větrným klimatem (VK) na území ČR. Znalost tohoto vztahu lze využít pro určení EVK, protože VK se ve srovnání s EVK zabývá podstatně větší množství prací a lze ho odhadnout z kratšího období měření. Například Rogers et al. (2005) uvádí, že dobré odhady lze získat při prodloužení měření větru vhodnými metodami MCP (measure-correlate-predict) i z měření trvajících necelý rok. Navíc je na Ústavu fyziky atmosféry již k dispozici tzv. větrná mapa, tedy plošná informace o VK na území České republiky (Hanslian et. al, 2013a a 2013b; Hanslian et Hošek, 2015), na kterou lze takový vztah aplikovat (kapitola 4.2).

4.1.1 Vztah podle normy IEC 61400-1

Jednoduchý vztah mezi průměrnou rychlostí větru u a maximální rychlostí větru s dobou opakování 50 let je popsán v normě IEC 61400-1 (IEC, 2005), dále jen norma. Ta odhaduje desetiminutový průměr rychlosti větru s dobou opakování 50 let jako pětinašobek u a rychlost nárazu větru s dobou opakování 50 let jako sedminásobek. Předpokládá se, že tyto odhady platí ve výšce nad zemí odpovídající výšce osy turbíny větrné elektrárny, a že závislost průměrné rychlosti větru na výšce nad zemí je dána mocninnou funkcí této výšky s exponentem 0,2. Budeme předpokládat, že uvedené podíly platí ve výšce 100 m nad zemí, což je typická výška os současných větrných turbín (viz webové stránky České společnosti pro větrnou energii). Závislost průměrné desetiminutové rychlosti větru i rychlosti nárazu s dobou opakování 50 let na výšce nad zemí je dána mocninnou funkcí této výšky s exponentem 0,11. Z těchto vztahů plyne, že, ve standardní výšce meteorologických měření 10 m nad zemí je podíl mezi nejvyšší 10 minutovou rychlostí větru s dobou opakování 50 let a u 6,15 a podíl mezi nejvyšším nárazem větru s dobou opakování 50 let a u je 8,61.

V těchto vztazích je ale zanedbáno několik skutečností:

1. Uvedené podíly představují odhad pro typické lokality s větrnými elektrárnami. Typické podmínky v České republice však mohou být odlišné, a jednotlivé lokality se navíc liší podle orografických podmínek.
2. Vztahy předpokládají, že podíl mezi nejvyšší průměrnou rychlostí s dobou opakování 50 let a nejvyšším nárazem s touto dobou opakování se nebude měnit s výškou nad zemí, ale práce Choi et Hidayat (2002), Deaves (1993) a Popa (2007) naznačují, že tento podíl by se měl s výškou nad zemí snižovat.
3. Vztahy nereflektují odlišné skládání průměrné a extrémní rychlosti větru popsané v kapitole 3.6.

4.1.2 Použitá data a jejich příprava

Vzhledem k výše uvedeným nedostatkům vztahu mezi VK a EVK podle normy byl odvozen nový vztah, který se snaží zohlednit tyto připomínky. EVK bylo odvozeno z hodnot nejvyšších denních nárazů na profesionálních stanicích v ČR z let 1961 až 2009, přičemž byly použity stanice, na kterých jsou k dispozici výsledky měření nárazů větru po dobu alespoň 20 let. VK bylo spočítáno z 15 minutových průměrných rychlostí z databáze CLIDATA, z období od počátku dostupnosti těchto dat na dané stanici do konce roku 2008, což na většině stanic činilo cca 5 let. Podrobnější popis těchto dat lze nalézt v práci Řepka (2011). Oba typy dat současně byly dostupné na celkem 28 profesionálních stanicích ČHMÚ. Tyto stanice byly použity k určení hledaného vztahu. Protože se používají naměřené extrémní hodnoty z více stanic současně, lze tento postup považovat za variantu regionálního modelování extrémních hodnot běžně používaného v případě srážek (Kyselý et al., 2011).

Hodnoty nejvyšších denních nárazů byly homogenizovány postupem uvedeným v podkapitole 2.3. Hodnoty průměrných rychlostí větru homogenizované nebyly vzhledem ke krátké délce měření.

Data byla rozdělena na 8 sektorů stejných jako v podkapitole 2.3, tj. hranice prvního sektoru je mezi $11,25^\circ$ a $56,25^\circ$ a další sektory jsou posunuté o násobky 45° ve směru hodinových ručiček. Průměrné rychlosti z každého sektoru byly aproximovány Weibullovým rozdělením metodou popsanou v Troen et Petersen (1989). Tato metoda byla použita, protože byla použita i při konstrukci větrné mapy ČR (Hanslian et al., 2013a a 2013b), která je využita v podkapitole 4.2. Hodnoty nejvyšších ročních nárazů

byly aproximovány Gumbelovým rozdělením metodou L-momentů. Netradiční volba sektorů byla zdůvodněna v podkapitole 2.3.

4.1.3 Model pro závislost EVK na VK

Výpočet EVK je založen na aplikaci 16 modelů vícenásobné regrese (2 prediktandy (μ , σ), 8 sektorů) ve tvaru

$$\hat{y}_i = \sum_{j=1}^m a_{yij} x_{ij} + b_{y,i}, y = \mu, \sigma \quad (4.1.1)$$

kde i je index sektoru, j je index prediktoru, \hat{y}_i je odhad parametru y Gumbelova rozdělení ročních maxim nárazů větru v sektoru i , x_{ij} je hodnota prediktoru x_j v sektoru i , a_{yij} a b_{yi} jsou parametry lineární regrese a m je počet prediktorů.

Potenciální prediktory byly odvozeny a vybírány z následujících veličin:

1. Průměrná rychlost větru v daném sektoru u .
2. Parametr tvaru Weibullova rozdělení rychlostí větru v daném sektoru k .
3. Četnost výskytu větru z daného sektoru f .
4. Nadmořská výška stanice h .
5. Zesílení v daném sektoru z , definované jako podíl průměrné rychlosti v daném sektoru a celkové průměrné rychlosti větru.

Z těchto veličin bylo vytvořeno 5 dalších prediktorů mocninnou transformací (Wilks, 2011)

$$\begin{aligned} x^* &= x^\lambda, \lambda > 0 \\ x^* &= \ln(x), \lambda = 0 \\ x^* &= -(x^\lambda), \lambda < 0, \end{aligned} \quad (4.1.2)$$

kde x je původní veličina a x^* transformovaná veličina s parametrem λ zvoleným tak, aby rozdělení proměnné x^* bylo symetrické. Výše uvedené prediktory byly aplikovány jednak jako průměrné hodnoty přes všechny směry, tak i odděleně po směrech.

Výběr prediktorů a výpočet modelových parametrů probíhal zvláště pro parametry μ a σ , které budou označovány dále jako y .

Byl zvolen empirický postup konstrukce modelu, který se liší od aplikace standardní vícenásobné lineární regrese metodou nejmenších čtverců. Motivací pro zvolený postup byla snaha při výběru prediktorů zajistit, aby model pro každý sektor obsahoval stejné prediktory, a odpovědět na otázku, zda jsou parametry regrese

sektorově závislé. I když výsledné odhady modely nejsou optimální ve smyslu regrese, užitý postup výběru prediktorů umožňuje jednoduše získat poznatky o směrové závislosti prediktandů na prediktorech, které nejsou při aplikaci standardní regrese snadno dostupné.

Výpočet probíhal následujícím postupem:

1. Pro každou veličinu (prediktor) byly vypočteny absolutní členy $b_{y,i}$ v rovnici (4.1.1) jako průměrné hodnoty y v každém ze sektorů. Bylo spočteno reziduum r jako odchylka y od těchto absolutních členů.
2. Byly spočteny t-statistiky r se všemi dosud nepoužitými prediktory pro všechny sektory a sečteny jejich absolutní hodnoty pro každý prediktor pro všechny sektory. Byl vybrán prediktor p s nejvyšší hodnotou tohoto součtu.
3. Metodou bootstrap byla vypočtena kritická hodnota součtu absolutních hodnot t-statistik z kroku 2 na hladině významnosti 0,95 při nulové hypotéze, že r nezávisí na p . Pokud nešlo nulovou hypotézu na této hladině významnosti zamítnout, byl výpočet ukončen.
4. Byly vypočteny lineární parametry lineární regrese pro závislost r na p pro každý sektor i a_{yip} .
5. Pro všechny sektory byly spočítány následující statistiky Δ_i určující, nakolik se odhad lineárního koeficientu regrese závislosti z na prediktoru x_k počítaný z dat v sektoru i liší od odhadu tohoto koeficientu počítaného z dat ze všech ostatních sektorů dohromady

$$\Delta_i = \frac{\hat{a}_i - \hat{a}_{\sim i}}{\sqrt{s^2(\hat{a}_i) + s^2(\hat{a}_{\sim i})}} \quad (4.1.3)$$

Kde \hat{a}_i a $\hat{a}_{\sim i}$ jsou odhady koeficientu lineární regrese pro závislost r na prediktoru p počítané z dat v sektoru i a z dat ze všech ostatních sektorů dohromady a $s(\hat{a}_i)$ a $s(\hat{a}_{\sim i})$ odhady jejich chyb. Byl spočítán součet absolutních hodnot těchto statistik S_i . V Gujarati (2004) je dokázáno, že za předpokladů platnosti lineárního regresního modelu má statistika Δ_i limitně pro velký počet kalibračních dat normované normální rozdělení při platnosti nulové hypotézy, že lineární parametr regresního vztahu pro sektor i je stejný jako ve všech ostatních sektorech uvažovaných dohromady. Zde bylo z tohoto tvrzení využito, že tato statistika limitně nezávisí na hodnotě koeficientu a . Z toho plyne, že na tomto koeficientu limitně nezávisí ani statistika S_i . To

umožňuje odhadnout kritické hodnoty S_i metodou Monte Carlo (Robert et Casella, 2004) generovanou ze vzájemně nezávislých dat při platnosti nulové hypotézy, že hodnota parametru a je ve všech sektorech stejná, přičemž tato data stačí vygenerovat pro případ $a = 0$. Pokud tuto nulovou hypotézu nebylo možno zamítnout na hladině významnosti 0,98, byly parametry lineární regrese a_{yip} nahrazeny jejich průměrnou hodnotou přes všechny sektory i .

6. Byly upraveny odhady absolutních členů modelu (4.1.1) $b_{y,i}$ tak, aby průměrná hodnota reziduí v každém se sektorů byla 0 a do modelů (4.1.1) byly přidány členy $a_{yip}x_{ip}$, kde se za a_{yip} použily hodnoty vypočtené v krocích 5 a 6.
7. Byla vypočtena nová rezidua r
8. Ze seznamu prediktorů byl odstraněn vybraný prediktor p spolu s prediktorem s ním spojeným pomocí rovnic (4.1.2).
9. Pokračovalo se krokem 2.

Nakonec bylo u výsledných odhadů absolutních koeficientů b stejně jako u lineárních koeficientů kritériem založeným na statistice (4.1.3) rozhodnuto, zda závisejí na sektoru nebo ne.

Výsledné modely lze zapsat ve tvaru

$$\begin{aligned}\mu_i(u_i, f_i, i) &= a_{\mu,u}(i) * u_i + a_{\mu,f} * f_i + b_{\mu}(i), \\ \sigma_i(u_i, f_i^*, i) &= a_{\sigma,u} * u_i + a_{\sigma,f^*} * f_i^* + b_{\sigma}(i),\end{aligned}\tag{4.1.4}$$

kde index i v argumentu značí závislost na sektoru i . Tento model se sestává z 16 dílčích lineárních modelů. f^* je četnost f transformovaná vzorcem (4.1.2) s $\lambda = 0,1440$. Odhady koeficientů jsou uvedeny v tabulce 4.1.1.

Ukázalo se, že oba parametry Gumbelova rozdělení zdaleka nejvíce závisejí na průměrné rychlosti větru. Jako jediný další prediktor zahrnutý do modelu vyšla četnost výskytu větru z daného sektoru f a v případě parametru σ opět tato četnost transformovaná vztahem (4.1.2). Ze znamének koeficientů plyne, že oba parametry Gumbelova rozdělení ročních maxim rostou s průměrnou rychlostí větru. S rostoucí četností větru (při stejné průměrné rychlosti větru) z daného sektoru roste parametr μ , ale klesá parametr σ , což znamená, že se rozdělení ročních maxim posunuje doprava a současně klesá jeho rozptyl.

	1	2	3	4	5	6	7	8
$a_{\mu,u}$	1,629	1,812	2,058	2,516	2,527	2,100	2,268	2,218
$a_{\mu,f}$	8,183	8,183	8,183	8,183	8,183	8,183	8,183	8,183
b_{μ}	8,947	8,801	8,419	8,200	9,373	12,712	12,543	10,078
$a_{\sigma,u}$	0,1090	0,1090	0,1090	0,1090	0,1090	0,1090	0,1090	0,1090
a_{σ,f^*}	-5,0153	-5,0153	-5,0153	-5,0153	-5,0153	-5,0153	-5,0153	-5,0153
b_{σ}	5,5220	5,8188	5,9506	6,0582	6,7530	6,6348	6,6780	6,1812

Tabulka 4.1.1: Odhad parametrů modelu 4.1.4. Sloupce odpovídají jednotlivým sektorům.

4.1.4 Odhad chyb modelu

Odhad chyby parametrů Gumbelova rozdělení podle vzorců (4.1.4) komplikují dvě skutečnosti.

1. Oba odhady (4.1.4) pro μ i σ se skládají z 8 lineárních regresních modelů, ale koeficienty v těchto modelech nebyly vypočítány nezávisle.
2. Samotné „správné“ hodnoty parametrů Gumbelova rozdělení, ze kterých byla chyba vypočtena, jsou odhadem z omezeného počtu ročních maxim.

Proto byl k odhadu chyb použit částečně empirický postup s aplikací bootstrap techniky.

Odhad chyby lineární regrese prediktandu y počítané metodou nejmenších čtverců, \hat{s}_y , lze spočítat podle vzorce (Wilks, 2011):

$$\hat{s}_y^2 = \frac{1}{n-m} \sum_{i=1}^n (\hat{y}_i - y_i)^2, \quad (4.1.5)$$

kde n je celkový počet případů použitých pro výpočet parametrů regrese, m počet parametrů regresního modelu a \hat{y}_j jsou regresní odhady y . Tento vzorec byl použit pro výpočet chyby odhadu μ a σ podle (4.1.4), i když se nejedná o klasickou lineární regresi počítanou metodou nejmenších čtverců. Za n byl dosazen celkový počet případů, ze kterých byly parametry vypočteny – $8 \times 28 = 224$, a za m počet nezávislých koeficientů vždy v 8 regresích pro μ ($m = 17$) a σ ($m = 10$).

Zřejmě platí vztah

$$\hat{y}_m - \hat{y}_b = (\hat{y}_m - y) + (y - \hat{y}_b), \quad (4.1.6)$$

kde y je hodnota parametru Gumbelova rozdělení ročních maxim (μ nebo σ) počítaného metodou L-momentů při přesné znalosti prvních dvou L-momentů rozdělení těchto ročních maxim, \widehat{y}_m je jeho odhad z větrného klimatu vzorci (4.1.4) a \widehat{y}_b jeho odhad z omezeného počtu ročních maxim. Označme si

$$s_{mb}^2 \equiv E((\widehat{y}_m - \widehat{y}_b)^2), s_m^2 \equiv E((\widehat{y}_m - y)^2), s_b^2 \equiv E((\widehat{y}_b - y)^2), \quad (4.1.7)$$

kde E značí střední hodnotu. Potom za předpokladu vzájemné nezávislosti odchylek \widehat{y}_m a \widehat{y}_b od y ze vztahu (4.1.6) platí:

$$s_m^2 = s_{mb}^2 - s_b^2. \quad (4.1.8)$$

Odhad s_{mb} byl vypočten rovnicí (4.1.5) a s_b byl odhadnut pomocí bootstrap analýzy obdobně jako v podkapitole 3.9. Délka souborů však v tomto případě nebyla rovná 50, ale skutečnému počtu ročních maxim na dané stanici. s_b bylo vypočteno jako průměr rozptylu odhadů parametrů Gumbelova rozdělení přes všechny stanice a sektory. Pro parametr μ vyšlo 0,566 m/s a pro parametr σ 0,386 m/s.

Odhady $s_{m,\mu}$ a $s_{m,\sigma}$ spočítané vztahem (4.1.5) jsou rovny 1,977 m/s a 0,643 m/s (údaj navíc v indexu označuje veličinu, k níž se odhad vztahuje). Dosazením do (4.1.8) dostaneme $s_{m,\mu} = 1,895$ m/s a $s_{m,\sigma} = 0,514$ m/s.

Pro zhodnocení výsledků modelu (4.1.4) byly pro porovnání spočteny odhady STD S_μ a S_σ . Tyto odhady byly zjištěny tak, že byly vypočteny odhady rozptylů μ a σ odhadnutých z naměřených ročních maxim. Ty byly opět upraveny o směrodatnou odchylku těchto odhadů z naměřených ročních maxim pomocí vztahu (4.1.8). Výsledkem je $S_\mu = 4,974$ m/s a $S_\sigma = 0,688$ m/s. To znamená, že model (4.1.4) dává podstatně lepší odhad parametru μ než σ oproti jejich průměrné hodnotě přes všechny stanice. Redukuje rozptyl parametru μ o 88,5 %; rozptyl parametru σ jen o 44,2 %.

4.1.5 Kombinace výsledků regresního modelu s odhady ze staničních měření

Odhady parametrů Gumbelova rozdělení modelem (4.1.4) se dají zkombinovat s odhady z ročních maxim metodou založenou na metodě optimální interpolace (Barth et al., 2008). Výsledný odhad byl hledán ve tvaru lineární kombinace těchto odhadů:

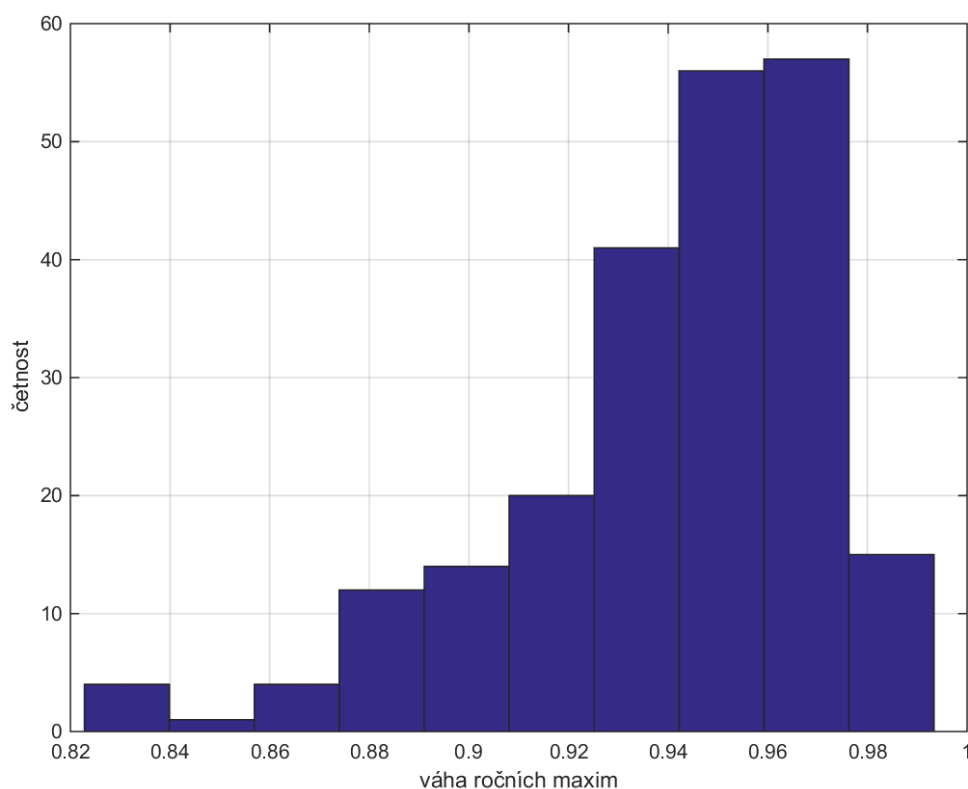
$$\widehat{y} = \omega_b * \widehat{y}_b + \omega_m * \widehat{y}_m, \quad (4.1.9)$$

kde ω_b a ω_m jsou konstanty, \widehat{y} hledaný odhad a \widehat{y}_b a \widehat{y}_m jsou odhady z ročních maxim a větrného klimatu. Váhy ω_b a ω_m byly určeny z požadavku, aby byl odhad \widehat{y} nevychýlený

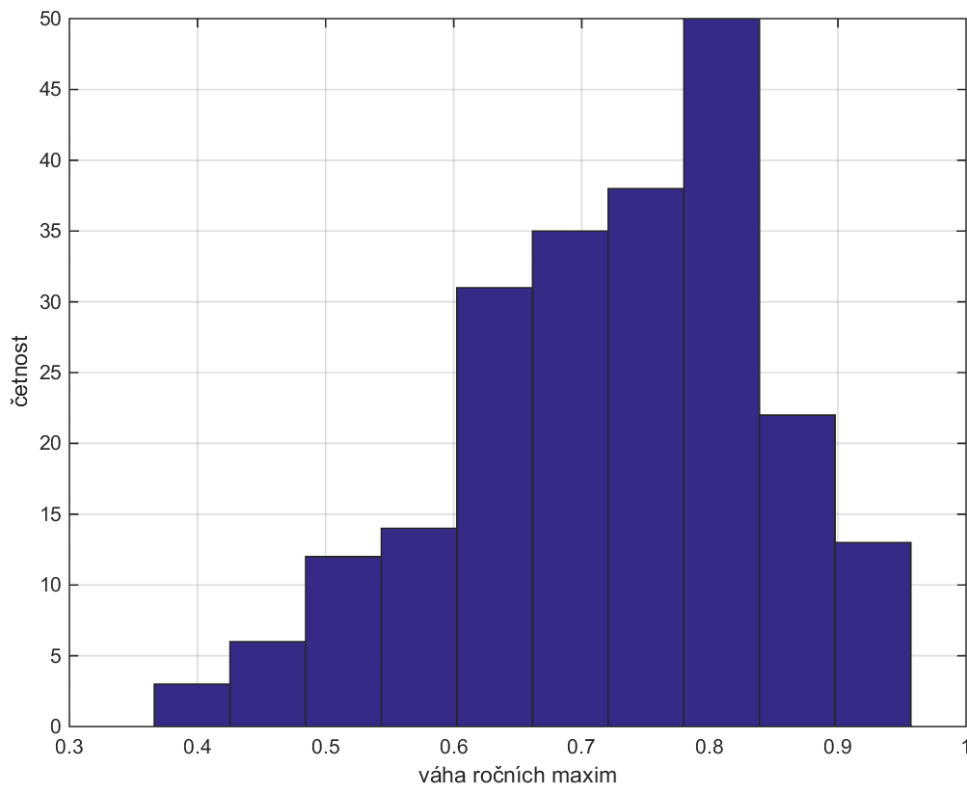
a jeho RMSE minimální. To za předpokladu, že odhady \widehat{y}_b a \widehat{y}_m jsou nevyčýlené a vzájemně nezávislé vede k řešení

$$\widehat{y} = \frac{s_m^2 \widehat{y}_b + s_b^2 \widehat{y}_m}{s_m^2 + s_b^2}, \quad (4.1.10)$$

kde s_b a s_m jsou RMSE odhadů \widehat{y}_b a \widehat{y}_m . s_b bylo odhadnuto metodou bootstrap z naměřených ročních maximálních nárazů větru (zvláště pro každou kombinaci stanice a sektoru). s_m byla vypočtena v sekci 4.1.4. Histogramy vah, které ve vzorci (4.1.10) přísluší odhadům získaných z ročních maxim pro jednotlivé zkoumané stanice a sektory, ze kterých vycházel model (4.1.4), jsou zobrazeny na obr. 4.1.1 a 4.1.2. Z histogramů je zřejmé, že na těchto stanicích je důležitější odhad z naměřených ročních maxim než z průměrného větrného klimatu. To platí především v případě parametru μ Gumbelova rozdělení.



Obr. 4.1.1: Histogram vah odpovídajících odhadům parametru μ Gumbelova rozdělení vypočteného z ročních maxim (viz vztah 4.1.10).



Obr. 4.1.2: Histogram vah odpovídajících odhadům parametru σ Gumbelova rozdělení vypočteného z ročních maxim (viz vztah 4.1.10).

4.1.6 Odhad u_{50} vypočtený z větrného klimatu

Odhad nárazu větru s dobou návratu n let, u_n , byl proveden z VK následujícím postupem:

1. Pokud je uvažována doba návratu všech nárazů, a ne jen jejich ročních maxim, provede se korekce n podle vzorce (3.5.16) (v kapitole 4 byla tato korekce použita).
2. Provede se výpočet odhadů parametrů Gumbelova rozdělení v jednotlivých sektorech podle modelu (4.1.4).
3. Výsledky z kroků 1 a 2 se využijí k odhadu u_n podle vzorce (3.4.1).

V případě znalosti ročních maxim je místo kroku 2 možno použít kombinované odhady ze sekce 4.1.5. Výsledky odhadů u_{50} z ročních maxim, $\widehat{u_{50,b}}$, průměrného klimatu, $\widehat{u_{50,m}}$, a jejich kombinace, $\widehat{u_{50,c}}$, jsou v tabulce 4.1.2.

indikativ	$\widehat{u}_{50,b}$	$\widehat{u}_{50,m}$	$\widehat{u}_{50,c}$	indikativ	$\widehat{u}_{50,b}$	$\widehat{u}_{50,m}$	$\widehat{u}_{50,c}$
B1HOLE01	36,92	34,77	36,52	O1LYSA01	51,92	50,69	51,24
B2BTUR01	34,48	38,39	34,99	O1MOSN01	37,88	37,65	37,01
B2DUKO01	33,22	39,79	33,95	O1PORU01	28,59	33,43	29,03
B2KMYS01	37,03	36,45	36,94	O2LUKA01	43,31	40,39	42,56
B2KUCH01	40,38	39,41	39,79	O3PRER01	35,05	35,13	35,10
B2NEDV01	35,64	41,14	36,22	P1PKAR01	44,00	38,44	42,26
C1CHUR01	38,19	37,05	37,22	P1PLIB01	36,44	34,61	36,19
C1KOCE01	42,76	38,74	41,07	P1PRUZ01	36,13	39,72	36,58
H3HRAD01	40,77	36,79	38,11	P3PRIB01	38,61	39,39	38,31
H3SVRA01	48,67	45,75	47,62	U1DOKS01	38,39	36,06	37,03
L2PRIM01	41,40	39,11	40,41	U1KATU01	45,27	36,84	43,05
L3CHEB01	34,03	35,62	33,76	U1KOPI01	36,32	35,06	36,24
L3KVAL01	35,02	35,85	34,99	U1MILE01	54,35	52,99	53,94
O1CERV01	36,65	38,94	36,68	U2LIBC01	38,70	37,47	38,14

Tabulka 4.1.2: Srovnání odhadů u_{50} proložením ročních maxim metodou L-momentů, $\widehat{u}_{50,b}$, odhadem z větrného klimatu vzorci (4.1.4), $\widehat{u}_{50,m}$, a jejich kombinací podle (4.1.10), $\widehat{u}_{50,c}$. Hodnoty jsou v m/s.

Rozdíly mezi odhady $\widehat{u}_{50,b}$ a $\widehat{u}_{50,m}$ nevykazují žádnou jednoznačnou souvislost. Lze předpokládat, že stanice s měřením silně ovlivněným střechami, které způsobují silnou turbulenci – viz Landberg (2000) (Karlovy, Tušimice) mají nižší odhady extrémních rychlostí podle modelu (4.1.4), než dávají odhady přímo z ročních maxim, a otevřené stanice (Holešov, Dukovany, Ruzyně, ale ne Mošnov) mají tyto odhady vyšší, ale toto tvrzení není při daném malém počtu zkoumaných stanic dostatečně průkazné. Vrcholové stanice Lysá hora a Milešovka jsou sice také velmi otevřené, ale lze předpokládat, že tato jejich otevřenost je již zahrnuta v celkové vysoké rychlosti větru na nich, a proto i přímo ve vstupních hodnotách modelu (4.1.4).

Odhady parametrů Gumbelova rozdělení v jednotlivých sektorech pomocí modelu (4.1.4) jsou nevychýlené a nevychýlené tedy budou i odhady rychlostí se zadanou dobou návratu v jednotlivých sektorech, které jsou jejich lineární kombinací. Průměrná

hodnota $\widehat{u}_{50,m}$ je však o 0,52 m/s a průměrná hodnota $\widehat{u}_{50,c}$ o 0,54 m/s nižší než průměrná hodnota $\widehat{u}_{50,b}$. To se dá vysvětlit tím, že model (4.1.4) snižuje rozptyl odhadů parametrů Gumbelova rozdělení v jednotlivých sektorech a tím snižuje pravděpodobnost velmi vysokých i nízkých rychlostí s touto dobou návratu v jednotlivých sektorech. Při jejich složení podle podkapitoly 3.6 mají na výsledný odhad u_{50} největší vliv sektory s nejvyššími extrémními rychlostmi, a proto vede redukce rozptylu odhadů parametrů Gumbelova rozdělení v jednotlivých sektorech ke zjištěnému systematickému podhodnocení odhadu u_{50} . Na druhé straně budou samotné odhady $\widehat{u}_{50,b}$ pravděpodobně vykazovat kladnou systematickou chybu, která je obdobným způsobem způsobena zvýšením rozptylu odhadů parametrů Gumbelova rozdělení v jednotlivých sektorech v důsledku chyby výběru.

4.1.7 Odhad chyby u_{50}

Pro odhad RMSE výpočtu u_{50} nelze použít vzorec (4.1.5), protože odhady parametrů modelu (4.1.4) se uskutečnily podstatně jinou metodou než minimalizací RMSE odhadu u_{50} , což je předpokladem použití tohoto vzorce. K odhadu chyby byla použita metoda křížové validace, při které byla u_{50} pro každou ze stanic počítána vždy z odhadů VK a EVK na ostatních stanicích. Tato metoda odhad chyby mírně přeceňuje, protože počítá regresní koeficienty z menšího počtu stanic. Stejně jako v případě odhadů chyb parametrů Gumbelova rozdělení je nutno vzít v úvahu, že hodnoty u_{50} počítané z naměřených ročních maxim jsou také odhady a využívají vzorec (4.1.8), do kterého se za s_b dosadila hodnota 1,09 m/s získaná stejným způsobem jako v případě parametrů Gumbelova rozdělení. Odhady u_{50} získané z naměřených ročních maxim $\widehat{u}_{50,b}$ křížovou validací modelu (4.1.4) $\widehat{u}_{50,m}$ a podle normy $\widehat{u}_{50,n}$ jsou porovnány v tabulce 4.1.3.

RMSE $_{Su_{50,m}}$ odhadů $\widehat{u}_{50,m}$ odhadnutá z této tabulky při použití vztahu (4.1.8) je 3,19 m/s; RMSE $_{Su_{50,n}}$ odhadů $\widehat{u}_{50,m}$ odhadnutá stejným způsobem je 12,22 m/s. Systematická chyba odhadů $\widehat{u}_{50,m}$ je -0,52 m/s a systematická chyba odhadů $\widehat{u}_{50,n}$ je -8,95 m/s.

$\widehat{u}_{50,n}$ je vyšší než $\widehat{u}_{50,b}$ pouze na vrcholových stanicích Nedvězí, Svratouch, Lysá hora a Milešovka, na poslední jmenované stanici výrazně. Z toho usuzujeme, že tato norma odpovídá typickým podmínkám na vyvýšených hřebenech. V případě výrazných

izolovaných vrcholů odhady u_{50} přeceňuje a v ostatních případech, které tvoří velkou většinu území, naopak podceňuje.

Indikativ	$\widehat{u}_{50,b}$	$\widehat{u}_{50,m}$	$\widehat{u}_{50,n}$	indikativ	$\widehat{u}_{50,b}$	$\widehat{u}_{50,m}$	$\widehat{u}_{50,n}$
B1HOLE01	36,92	34,86	18,18	O1LYSA01	51,92	50,22	56,46
B2BTUR01	34,48	38,46	28,76	O1MOSN01	37,88	37,80	29,94
B2DUKO01	33,22	39,88	32,95	O1PORU01	28,59	33,48	17,42
B2KMYS01	37,03	36,49	25,65	O2LUKA01	43,31	40,45	33,71
B2KUCH01	40,38	39,51	28,74	O3PRER01	35,05	35,20	21,54
B2NEDV01	35,64	41,26	39,60	P1PKAR01	44,00	38,46	26,02
C1CHUR01	38,19	37,00	25,25	P1PLIB01	36,44	34,66	19,64
C1KOCE01	42,76	38,67	30,90	P1PRUZ01	36,13	39,76	32,69
H3HRAD01	40,77	36,82	26,08	P3PRIB01	38,61	39,48	33,08
H3SVRA01	48,67	45,48	51,09	U1DOKS01	38,39	36,06	20,10
L2PRIM01	41,40	39,04	36,13	U1KATU01	45,27	36,90	25,17
L3CHEB01	34,03	35,64	21,00	U1KOPI01	36,32	35,14	17,79
L3KVAL01	35,02	35,85	23,64	U1MILE01	54,35	52,57	71,11
O1CERV01	36,65	39,00	31,00	U2LIBC01	38,70	37,50	25,85

Tabulka 4.1.3: Srovnání odhadů u_{50} z naměřených ročních maxim $\widehat{u}_{50,b}$, z průměrného větrného klimatu pomocí křížové validace $\widehat{u}_{50,m}$ a podle normy $\widehat{u}_{50,n}$. Hodnoty jsou v m/s .

Ze srovnání tabulek 4.1.2 a 4.1.3 plyne, že se odhady u_{50} vypuštěním jedné stanice z odhadu parametrů změnilo jen minimálně. Rozdíl je na většině stanic menší než 0,1 m/s . Výrazně vyšší je na vrcholových stanicích Milešovka a Lysá hora, kde se blíží 0,5 m/s . Tento vyšší rozdíl se dá vysvětlit tím, že hodnoty průměrné rychlosti jsou na těchto stanicích odlehle od ostatních stanic, a proto silně ovlivňují výpočet parametrů modelu (4.1.4).

Celkově vycházejí regresní odhady podle vztahů (4.1.4) mnohem lépe než norma. Navíc oproti ní přidávají informaci o směrové závislosti odhadu u_{50} .

4.1.8 Testování vhodnosti použití Gumbelova rozdělení

Dále bylo zkoumáno, nakolik je Gumbelovo rozdělení vhodným rozdělením pro maximální roční nárazy větru. Kolmogorovův-Smirnovův, případně Cramér-von Misesův test (oba popsány v Darling, 1957) se pro tyto účely nepovažují za příliš vhodné, protože jsou nejcitlivější na odchylky kolem středu pravděpodobnostního rozdělení. Za vhodnější se považuje Anderson-Darlingův test (Hanel et al., 2009), který je speciálním případem třídy testovacích statistik T zavedených v Anderson et Darling (1952) ve tvaru

$$T = n \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 \omega(x) dF(x), \quad (4.1.11)$$

kde $F(x)$ je testovaná distribuční funkce a $F_n(x)$ empirická distribuční funkce, n je počet hodnot výběru a $\omega(x)$ váha. Speciálním případem je Cramér-von Misesův test, který získáme zavedením váhy $\omega(x) = 1$. Testovací statistika Anderson-Darlingova testu se liší zavedením váhy kladoucí větší důraz na chvosty rozdělení. Tato testovací statistika má tvar:

$$A = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x) * (1 - F(x))} dF(x). \quad (4.1.12)$$

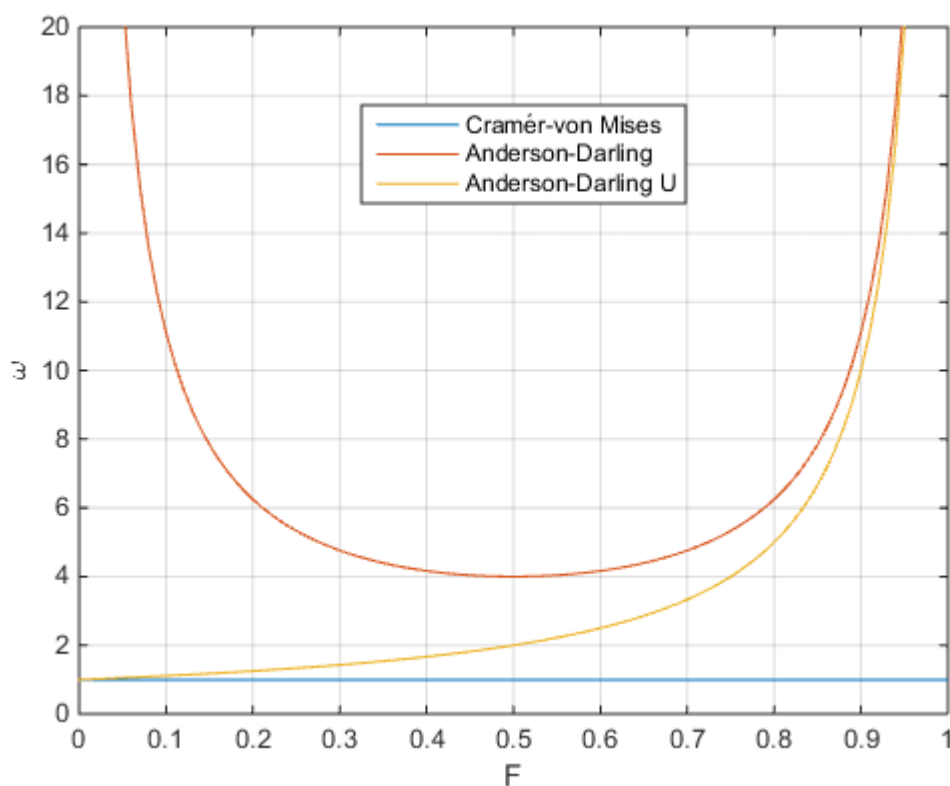
Pokud se parametry testovaného rozdělení odhadují z výběrových hodnot, závisí kritické hodnoty Andersonovy-Darlingovy statistiky na druhu rozdělení, a počítají se Monte-Carlo metodou (Lewis, 1961). To lze přímo provést pro Gumbelovo rozdělení, ale ne pro GEV rozdělení s neznámým parametrem tvaru, protože v tomto případě není možné vygenerovat potřebné řady Monte-Carlo metodou. Toto omezení platí i pro ostatní testy dobré shody.

Vzorec (4.1.12) klade vyšší důraz jak na levý, tak na pravý chvost rozdělení, ale při zkoumání extrémních hodnot považujeme obvykle za důležitý pouze jeden z nich. Proto byly zavedeny modifikace tohoto vzorce zdůrazňující vliv pouze jednoho z chvostů rozdělení (Ahmad et al., 1988). Pro tuto práci je důležitá modifikace zdůrazňující pravý chvost rozdělení, jejíž testovací statistika AU je vyjádřena vztahem

$$AU = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{1 - F(x)} dF(x). \quad (4.1.13)$$

Průběh závislosti vah ω ze vzorce (4.1.11) na distribuční funkci F je zobrazen na obr. 4.1.3. Z něj je zřejmé, že Anderson-Darlingův test skutečně oproti Cramér-von Misesovu testu zdůrazňuje vliv obou chvostů rozdělení. Modifikace Anderson-Darlingova testu podle vzorce (4.1.13) zdůrazňuje vliv pouze pravého chvostu rozdělení, a to ještě výrazněji, než původní Anderson-Darlingův test, protože váhy v pravém

chvostu jsou v případě těchto chvostů prakticky totožné, ale váha v ostatních částech grafu je u modifikované verze podstatně nižší.



Obr. 4.1.3: Průběh závislosti vah ω ze vzorce (4.1.11) na distribuční funkci F pro Cramér-von Misesův test, Anderson-Darlingův test a jeho modifikaci podle vztahu (4.1.13).

Pro testování hypotézy že roční maxima pro jednotlivé kombinace stanice a sektoru pocházejí z Gumbelova rozdělení byla použita testovací statistika AU , protože klade z výše uvedených možností nejvyšší váhu na pravý chvost rozdělení. Byl proveden výpočet kritických hodnot této testovací statistiky pro Gumbelovo rozdělení při odhadu obou jeho parametrů metodou L-momentů a pro počet případů $n = 10$ až 100 , přičemž pro každé n bylo provedeno 10^5 náhodných simulací. Vypočítané kritické hodnoty se využily pro testování naměřených ročních maxim.

Výsledky tohoto testování jsou uvedeny v tabulce 4.1.4. V prvním řádku jsou hladiny spolehlivosti, na kterých se zkoumala nulová hypotéza zvláště pro jednotlivé případy. Tyto hladiny spolehlivosti se dále nazývají primárními hladinami spolehlivosti. Ve druhém řádku je počet případů, ve kterých došlo k zamítnutí nulové hypotézy pro danou primární hladinu spolehlivosti. Ve třetím řádku je průměrný počet těchto

zamítnutých případů za předpokladu splnění nulové hypotézy ve všech případech. V posledním řádku je hladina spolehlivosti zamítnutí hypotézy, že skutečný počet překročení není pro danou primární hladinu významnosti vyšší než jeho teoretická střední hodnota. Tuto hladinu významnosti nazveme sekundární hladinou významnosti. Pro primární hladiny významnosti 0,95 a 0,99 je skutečný počet překročení téměř rovný své teoretické střední hodnotě. Pro nižší primární hladiny významnosti je skutečný počet překročení na hladině významnosti 0,95 statisticky významně vyšší než jeho teoretická hodnota. Vlastní číselné rozdíly nejsou výrazné, což se nevylučuje se statistickou významností vzhledem k poměrně velkému počtu 224 případů.

Podle tohoto testu lze konstatovat, že se podařilo poměrně jednoznačně zamítnout hypotézu, že data ve všech jednotlivých kombinacích stanice a sektoru pocházejí z Gumbelova rozdělení. Pozorované rozdíly se však nezdají příliš výrazné.

p primární	0,5	0,75	0,9	0,95	0,99
počet překročení	126	70	33	11	2
teoretický počet	112	56	22,4	11,2	2,24
p sekundární	0,9645	0,9796	0,9842	0	0

Tabulka 4.1.4: Výsledky testování, zda naměřená roční maxima pocházejí z Gumbelova rozdělení modifikovaným Anderson-Darlingovým testem. Podrobnosti jsou uvedeny v textu.

Pro analýzu extrémních hodnot je zásadní otázka, zda mají naměřené hodnoty lehčí nebo těžší pravé chvosty než odpovídá Gumbelovu rozdělení. Na tuto otázku předchozí test nedává odpověď.

Proto byly navrženy tři další testy, které srovnávají váhu pravého chvostu naměřených dat a Gumbelova rozdělení. Pokud by podle všech těchto testů vyšlo, že mají naměřená data buď těžší nebo lehčí pravý chvost než Gumbelovo rozdělení, bylo by možné najít vhodnou hodnotu parametru ξ GEV rozdělení, která by tuto odchylku kompenzovala. V tom případě by se nahradila aproximovaná Gumbelova rozdělení v jednotlivých kombinacích stanice a sektoru L-ekvivalentními GEV rozděleními s tímto parametrem ξ .

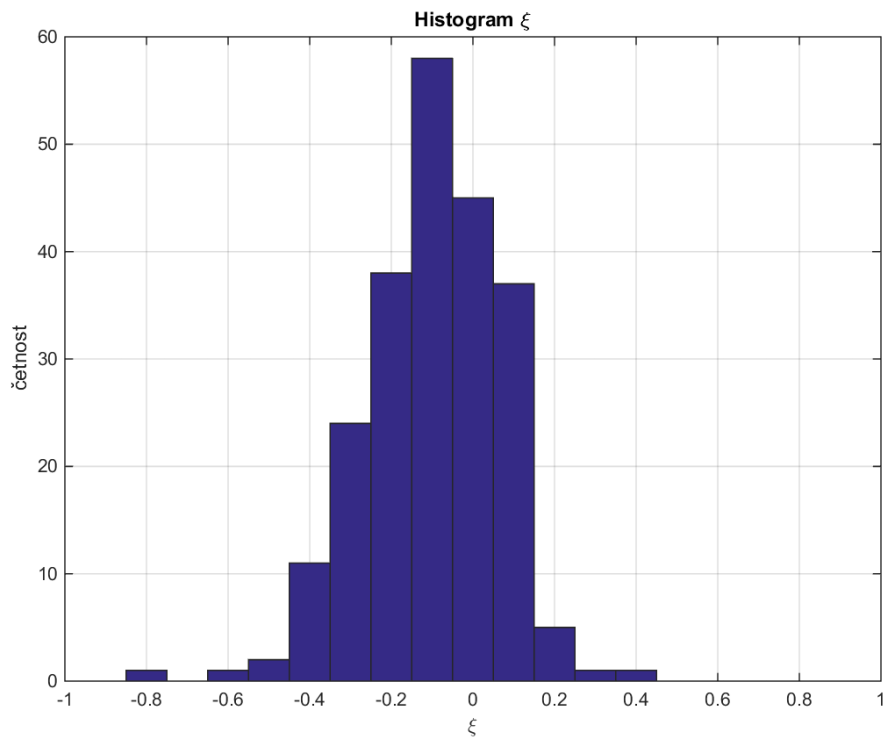
Tyto tři testy spočívaly ve sledování těchto charakteristik:

1. Podílu odhadů parametru ξ menších než 0 při aproximaci GEV rozdělením metodou L-momentů.
2. Počtu n_2 překročení odhadů u_{50} ve všech kombinacích stanice a sektoru.
3. Počtu n_3 kombinací stanice a sektoru, na kterých byl překročen odhad u_{50} .

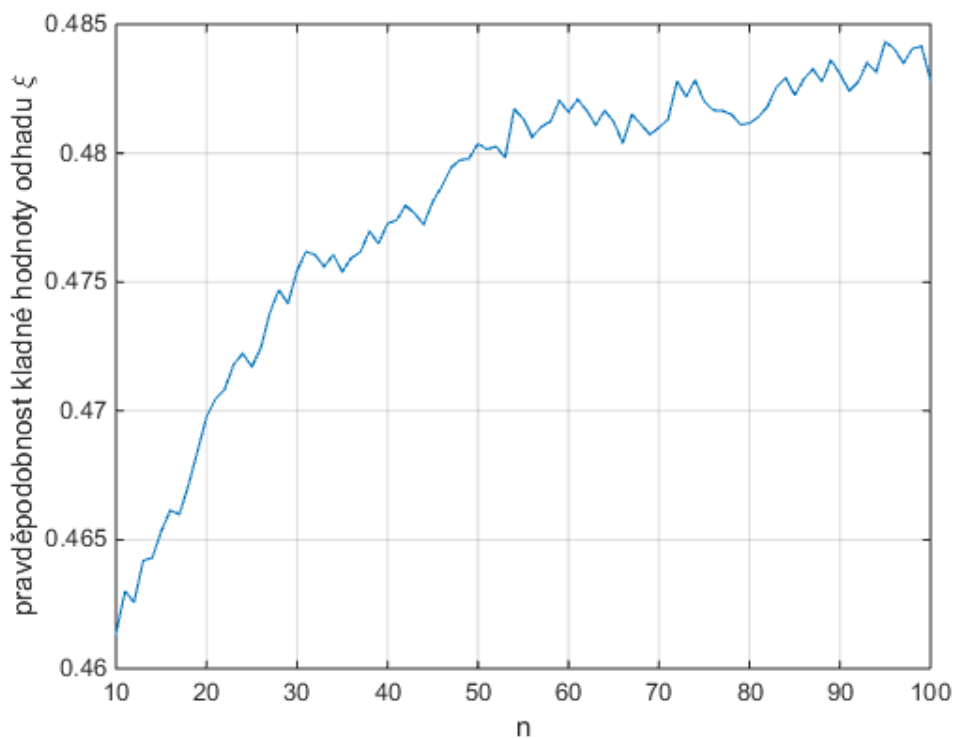
Histogram odhadu parametru tvaru ξ GEV rozdělení metodou L-momentů pro všechny kombinace stanice a sektoru je zobrazen na obr. 4.1.4. Většina odhadnutých hodnot je záporná, průměr má hodnotu -0,097 a medián -0,087. Celkový počet kladných odhadů parametru ξ je 68 z 224, což je méně než 1/3. Metodou Monte Carlo byla spočtena pravděpodobnost odhadu kladné hodnoty parametru ξ z výběru n hodnot pocházejících z Gumbelova rozdělení. Pro každou hodnotu n od 10 do 100 bylo provedeno 10^5 simulací. Výsledek je na obr. 4.1.5. Výsledná pravděpodobnost je sice menší než 1/2, ale podstatně vyšší než empirický výsledek $68/224 = 0,304$. Pro zkoumaný rozsah n mezi 20 a 49 (rozsah počtu ročních maxim na stanicích) má křivka z obr. 4.1.5 minimum 0,470. Z tohoto minima můžeme spočítat horní hranici hladiny významnosti, že hodnoty pocházejí z Gumbelova rozdělení. Tato horní hranice má hodnotu $5,68 \times 10^{-7}$. Nulová hypotéza, že hodnoty pocházejí z Gumbelova rozdělení, je tedy zamítnuta na podstatně vyšší hladině spolehlivosti, než vyplývá z modifikovaného Anderson-Darlingova testu podle vzorce (4.1.13). Podle tohoto testu mají naměřená data lehčí chvost, než odpovídá Gumbelovu rozdělení.

Pro zjištění kritických hodnot n_2 a n_3 byla opět použita metoda Monte Carlo. Bylo vygenerováno 10^4 krát 224 řad umělých ročních maxim s Gumbelovým rozdělením s délkami odpovídajícími délkám měření na jednotlivých stanicích. Z těchto hodnot byly vypočteny medián a kvantily odpovídající hladině významnosti 0,05 pro doby opakování od 20 do 100 let a vyneseny do grafu společně se skutečnými hodnotami.

Graf pro n_2 je na obr. 4.1.6 a graf pro n_3 je na obr. 4.1.7. Zde je n_2 statisticky významně vyšší než v případě Gumbelova rozdělení, což naznačuje těžší pravý chvost rozdělení oproti Gumbelovu rozdělení.



Obr. 4.1.4: Histogram odhadu parametru tvaru ξ GEV rozdělení metodou L-momentů pro všechny kombinace stanice a sektoru.



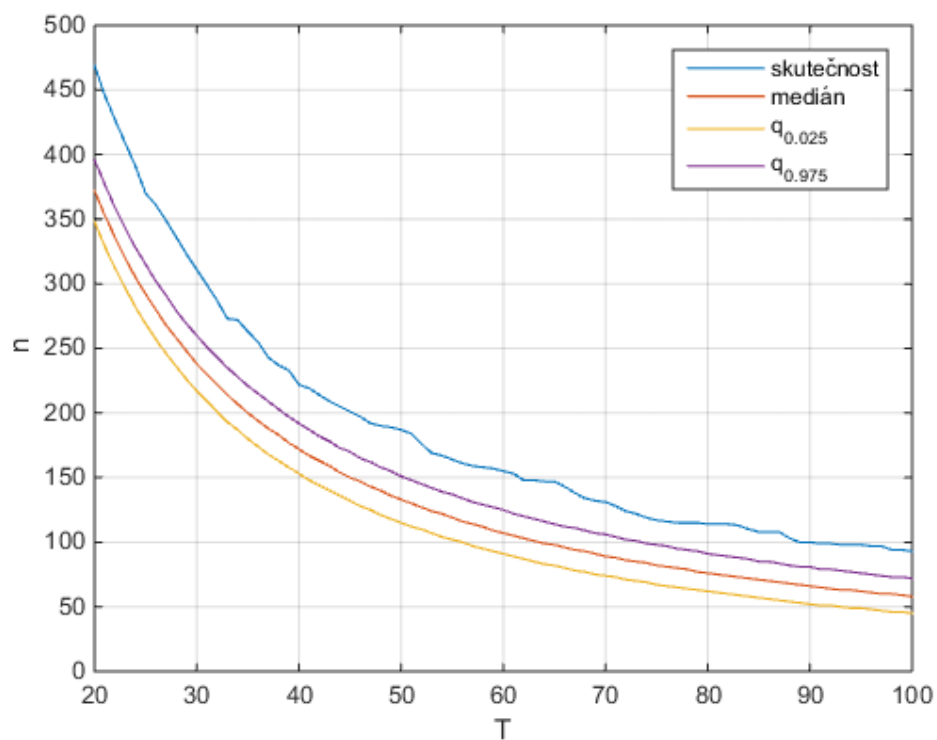
Obr. 4.1.5: Pravděpodobnost odhadu kladné hodnoty ξ při aproximaci výběru n hodnot pocházejících z Gumbelova rozdělení GEV rozdělením metodou L-momentů

Naopak n_3 je pro doby návratu nižší než 50 let statisticky významně nižší než u Gumbelova rozdělení, přičemž pokles počtu těchto řad s rostoucí délkou návratu je podstatně pozvolnější, než odpovídá nulové hypotéze. Srovnáním těchto dvou grafů (obr. 4.1.4 a 4.1.5) lze říci, že pokud dojde pro danou řadu k překročení rychlosti s určitou dobou návratu, je pravděpodobné, že k tomuto překročení dojde ve více letech než odpovídá Gumbelovu rozdělení.

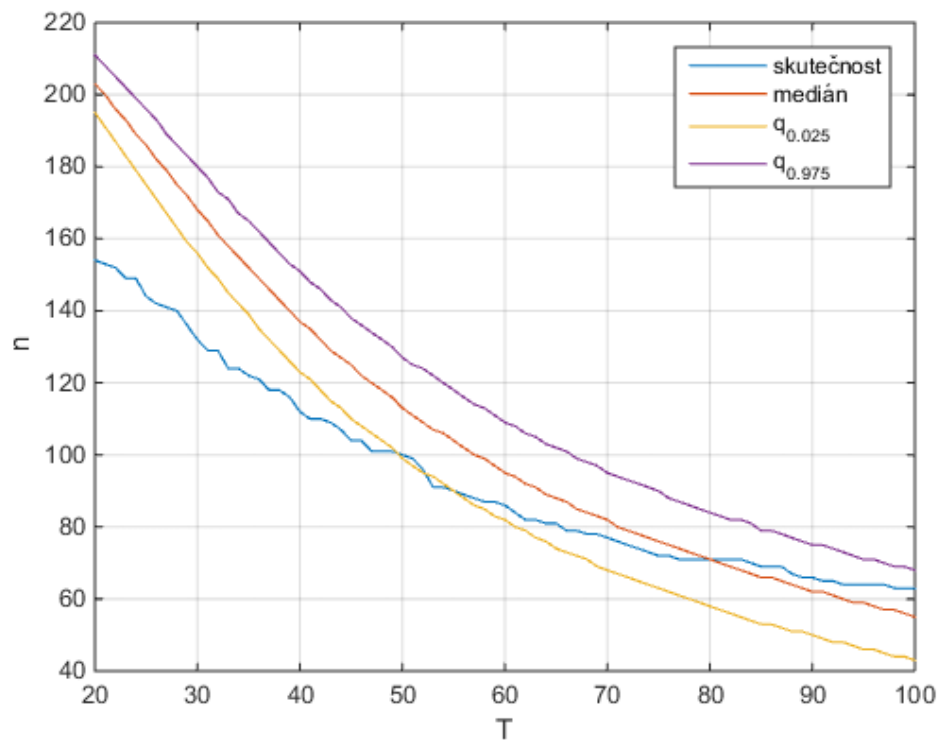
Křivky odpovídající naměřeným datům na obr. 4.1.6 a 4.1.7 se od mediánových křivek Gumbelova rozdělení podstatně liší, ale doby opakování, které odpovídají celkovému počtu překročení nebo počtu řad s překročením v případě mediánové křivky nulové hypotézy a reálných dat, se liší pouze o desítky procent a nikoli násobně.

Další zkoumanou otázkou bylo, zda parametr ξ GEV rozdělení vypočtený aproximací řad ročních maxim rychlosti větru pro jednotlivé kombinace stanice a sektoru koreluje s prediktory parametrů Gumbelova rozdělení uvedenými v sekci 4.1.3. Nejvyšší korelace vyšla stejně jako v případě parametrů Gumbelova rozdělení pro průměrnou rychlost u . Bodový graf mezi těmito odhady ξ a u je na obr. 4.1.8. Korelační koeficient vyšel velmi těsně statisticky významně nenulový na hladině významnosti 0,05. Ale tento výpočet vychází z předpokladu, že obě veličiny mají normální rozdělení. Z bodového grafu je ale vidět, že existuje několik odlehlých hodnot průměrné rychlosti větru a tyto hodnoty mají nízké hodnoty odhadu parametru ξ , čímž zásadně ovlivňují výsledky korelace. To snižuje vypovídací schopnost použité testovací metody. Navíc zjištěná regresní přímka vysvětluje méně než 2 % celkového rozptylu hodnot ξ , takže navržená regrese je prakticky nepoužitelná.

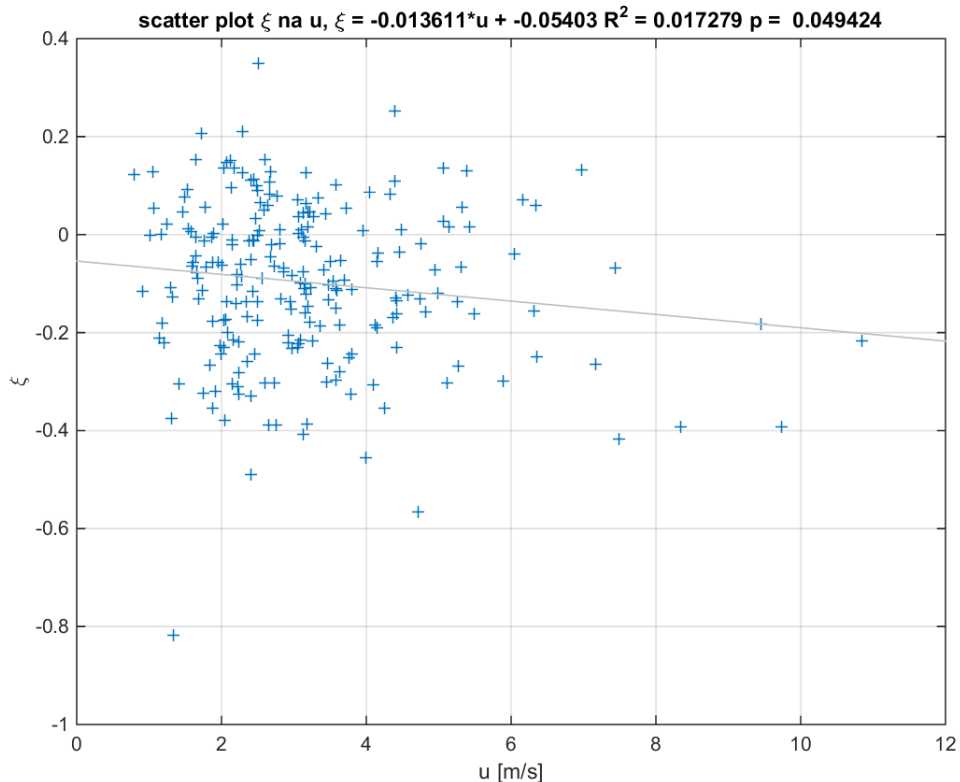
Celkově se dá konstatovat, že i když naměřená data nepocházejí z Gumbelova rozdělení, je toto rozdělení vhodné k jejich aproximaci. Proto bylo použito při výpočtu mapy extrémního větru. Toto rozdělení má nekonečný, ale přitom lehký (ve smyslu existence všech momentů) pravý chvost, čímž zabraňuje velmi odlehlým odhadům u_n . Provedené testy neukázaly, že by tento chvost byl oproti naměřeným datům příliš lehký nebo naopak příliš těžký. Dalším argumentem pro použití Gumbelova rozdělení je, že jeho použití je standardní při výpočtu extrémních rychlostí větru.



Obr. 4.1.6: Statistika n_2 , naměřená data a kvantily Monte Carlo simulace Gumbelova rozdělení v závislosti na době návratu T .



Obr. 4.1.7: Statistika n_3 , naměřená data a kvantily Monte Carlo simulace Gumbelova rozdělení v závislosti na době návratu T .



Obr. 4.1.8: Bodový graf a lineární regrese mezi průměrnou rychlostí větru a parametrem ξ GEV rozdělení maximálních ročních nárazů větru aproximovaného metodou L-momentů.

4.2 Mapa extrémního větru na území ČR

Obsahem této kapitoly je výpočet mapy extrémního větru na území ČR aplikací vztahu mezi VK a EVK odvozeného v podkapitole 4.1 na mapu VK vytvořenou v Ústavu fyziky atmosféry (Hanslian et al., 2012, Hanslian et Hošek, 2015) a její srovnání s obdobnými mapami vytvořenými jinými autory.

4.2.1 Použitá mapa VK

Výpočet použité mapy VK je založen na metodologii větrného atlasu (Troen et Petersen, 1989). Tato metodologie je založena na myšlence, že VK v daném místě se skládá ze zobecněného větrného klimatu (GWC – generalized wind climate), které představuje idealizované VK nad plochým homogenním terénem bez překážek a z místních vlivů, jež GWC modifikují.

Výpočet větrné mapy probíhal ve třech krocích. V prvním kroku bylo z měření na stanicích spočítáno GWC modelem WASP (Mortensen et al., 1993) pro lokalitu každého měření, přičemž se braly v úvahu i překážky. Vstupní měření pocházela z profesionálních stanic ČHMÚ, dobrovolných stanic ČHMÚ, stanic AIM mimo zastavěné oblasti a stožárů (Hanslian et Hošek, 2015). Použita byla data z období 1997 až 2006. GWC bylo spočítáno pro dvanáctisměrné růžice, přičemž GWC v každém sektoru bylo dáno dvěma parametry Weibullova rozdělení a četností výskytu větru v daném sektoru. Ve druhém kroku byla tato GWC prostorově interpolována metodou VAS (Sokol et Štekl, 1995). Metoda VAS byla vyvinuta pro třírozměrnou interpolaci průměrného VK mezi jednotlivými stanicemi, tedy se zohledněním nadmořské výšky. V posledním třetím kroku byl v každém výstupním gridovém bodě spočítán vliv místních podmínek na GWC. V úvahu byly na rozdíl od prvního kroku brány pouze vlivy lokální orografie a parametru drsnosti zemského povrchu, nikoli překážek. Výpočet probíhal pro celé území ČR, které bylo pokryto sítí uzlových bodů s vysokým rozlišením 100 m. Pro popis orografie byl použit digitální model terénu DMU25 a rozložení parametru drsnosti zemského povrchu bylo odvozeno z databáze CORINE (European Environment Agency, 2006), jejíž typy povrchu byly převedeny na parametr drsnosti podle článku Wieringa (1993). Výsledná mapa VK, a tedy i EVK, odpovídá použitým mapám orografie a drsnosti povrchu. Nejsou v ní tedy zohledněny například vlivy těch detailů povrchu, které v těchto mapách nejsou zahrnuty, a místních překážek, například stromů či budov.

Pro výpočet větrné mapy byla využita všechna dostupná měření větru, která prošla kontrolou kvality, což neumožnilo přímou validaci spočítané větrné mapy. Avšak později byla provedena validace této mapy pomocí 15 nezávislých měření větru ze stožárů a jednoho měření z větrné elektrárny uskutečněných po dokončení této mapy (Hanslian et Hošek, 2015). Z této validace vyšel odhad RMSE vypočtené průměrné rychlosti větru ve 100 m nad zemí 0,40 m/s s kladnou systematickou chybou 0,27 m/s, což činí 4,7 % odhadu průměrné rychlosti. Jak plyne z diskuze v uvedeném článku, tato systematická chyba může být z velké části způsobena podceněním efektu urychlení proudění nad střechami při přepočtu rychlostí větru ze staničních měření na GWC. Předpokládáme, že ve výšce 10 m nad povrchem bude RMSE v absolutní hodnotě přibližně stejná. Na jednu stranu lze v této výšce očekávat vyšší relativní chybu, protože proudění ve výšce 10 m nad zemí má vyšší prostorovou variabilitu než proudění ve výšce 100 m nad zemí, na druhou stranu je průměrná rychlost větru nad naším územím ve výšce 10 m nad zemí

zřetelně nižší než ve výšce 100 m nad zemí. Pro zohlednění zjištěné systematické chyby bylo rozhodnuto před dalším použitím vynásobit hodnoty průměrné rychlosti větru větrné mapy faktorem 0,953. Tato korekce vedla ke snížení odhadu RMSE na 0,30 m/s. Je také nutno brát v úvahu, že evaluace byla provedena v lokalitách uvažovaných pro výstavbu větrných elektráren, které nejsou z hlediska celého území ČR typické.

4.2.2 Výpočet mapy extrémního větru

Pro samotný výpočet mapy extrémního větru byly použity postupy popsané v předchozích podkapitolách. Mapy parametrů Gumbelova rozdělení nejvyšších ročních nárazů větru v jednotlivých sektorech byly vypočteny aplikací modelu (4.1.4) na větrnou mapu. Mapy rychlostí maximálního nárazu větru u_n se zadanými dobami opakování n v jednotlivých sektorech a celkově byly vypočteny postupem popsáním v sekci 4.1.6.

Výsledné mapy odhadů u_{50} v jednotlivých sektorech jsou ukázány na obr. 4.2.1 až 4.2.8. Nejvyšší rychlosti nárazů nastávají ze sektorů číslo 5 (191,25° až 236,25°), 6 (236,25° až 281,25°) a 7 (281,25° až 326,25°) tedy ze západního směru. Na obr. 4.2.9 až 4.2.12 jsou mapy odhadů u_{10} , u_{20} , u_{50} a u_{100} nezávisle na sektoru. Na obr. 4.2.13 je pro ilustraci zobrazen detail mapy odhadů u_{50} v Praze a jejím okolí, který ukazuje vysoké prostorové rozlišení vytvořené mapy. Na obr. 4.2.14 je mapa podílu odhadu u_{50} a průměrné rychlosti větru u v daném bodě. Oproti konstantní hodnotě 8,61 uvedené v normě se vypočtený podíl pohybuje v širokém rozmezí cca od 5 do 20, přičemž nejnižší je na exponovaných vrcholech a hřebenech hor a nejvyšší v údolích. Hodnota 8,61 zhruba odpovídá typickým lokalitám s větrnými elektrárnami v ČR, ale v případě jiných lokalit jsou hodnoty výrazně odlišné.

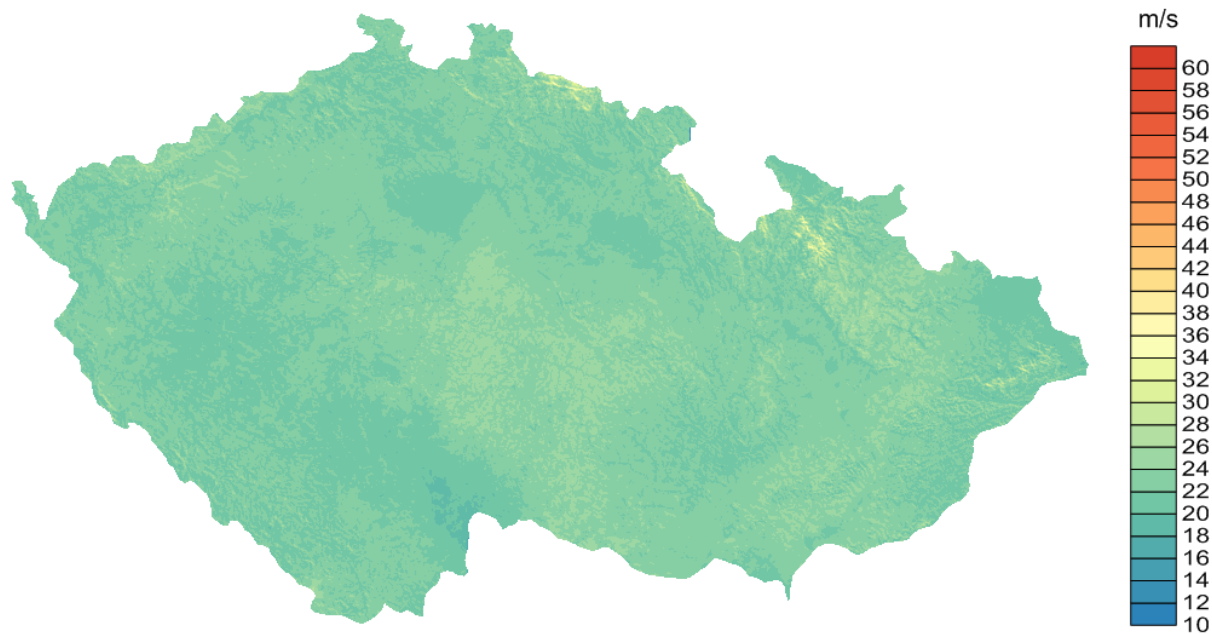
Otázkou je, nakolik je mapa odhadu u_{50} pouze modifikovanou mapou průměrné rychlosti větru z práce Hanslian et al. (2007). Proto byla spočítána lineární regrese mezi průměrnou rychlostí u z této mapy a odhadem u_{50} s využitím všech gridových bodů na území ČR. Mapa podílu odhadu z obr. 4.2.11 a odhadu z této regrese je na obr. 4.2.15. Rovnice této regrese je

$$u_{50} = 2,977 * u + 27,820, \quad (4.2.1)$$

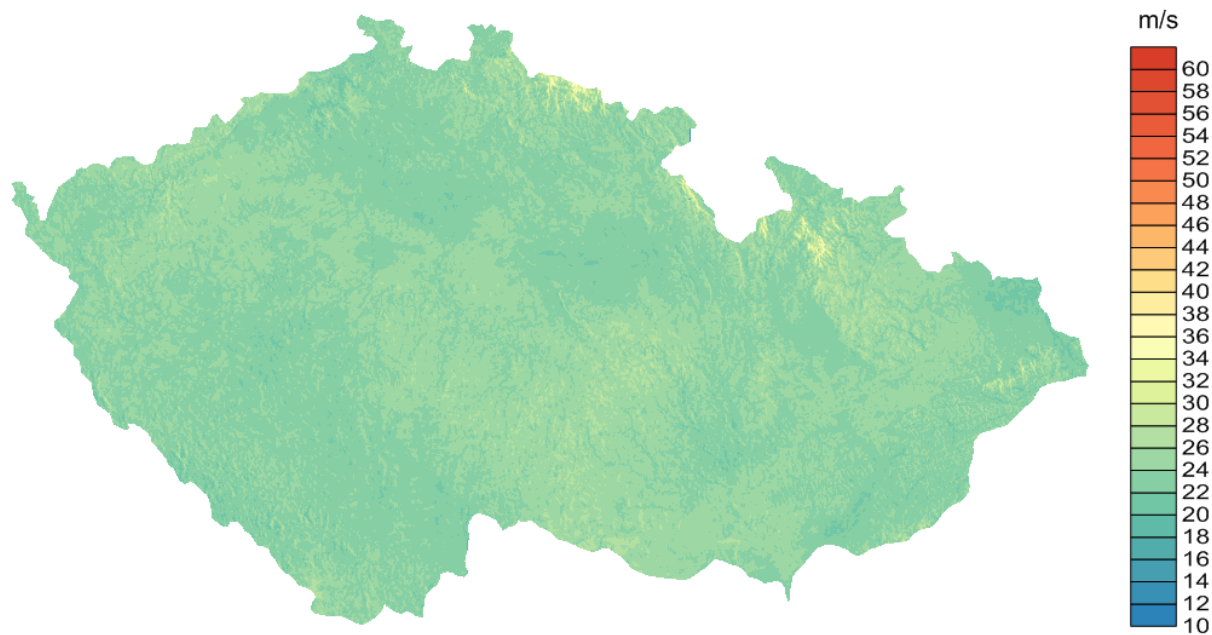
a vysvětluje 95,3 % rozptylu odhadu u_{50} modelu podle 4.1.4. V případě energie proudění, která je přímo úměrná třetí mocnině rychlosti větru (Sathyajith, 2006), však budou relativní rozdíly mezi modely podstatně vyšší. Zjištěný regresní vztah se zásadně liší od vztahu v normě

$$u_{50} = 8,61 * u.$$

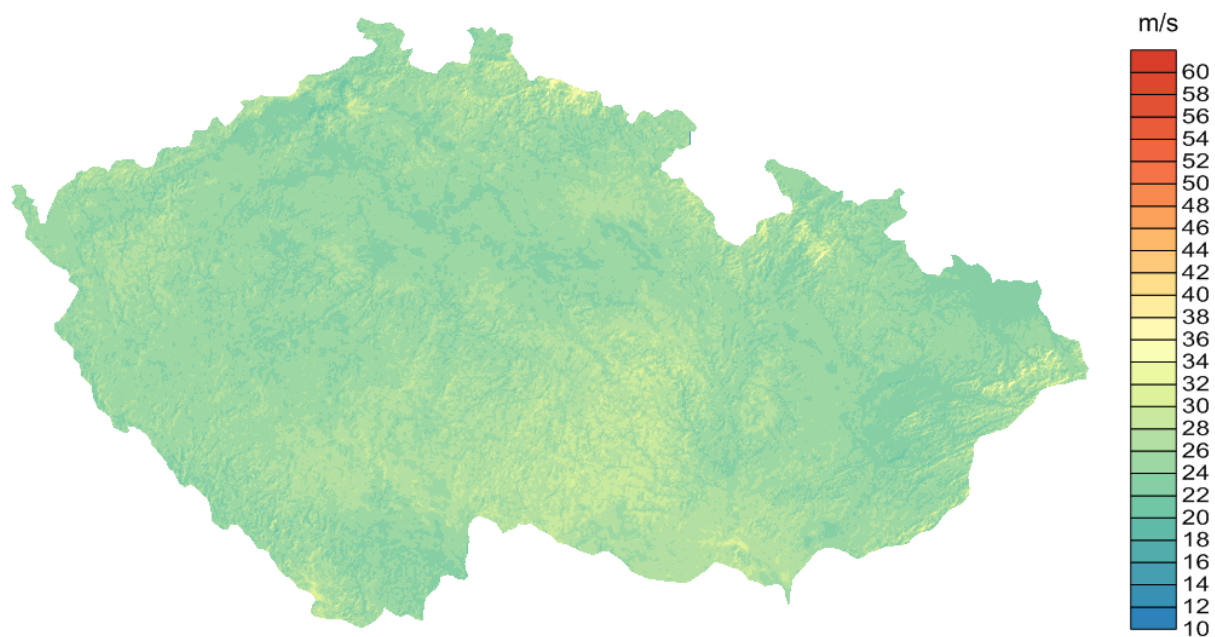
(4.2.2)



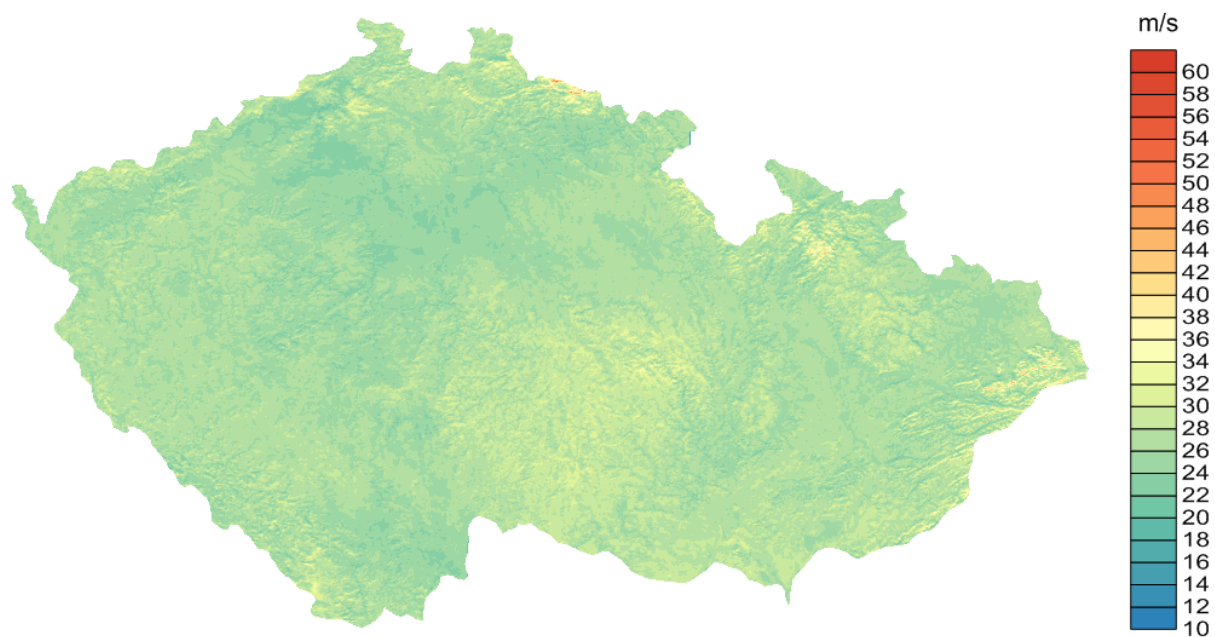
Obr. 4.2.1: Odhad u_{50} ze sektoru č. 1.



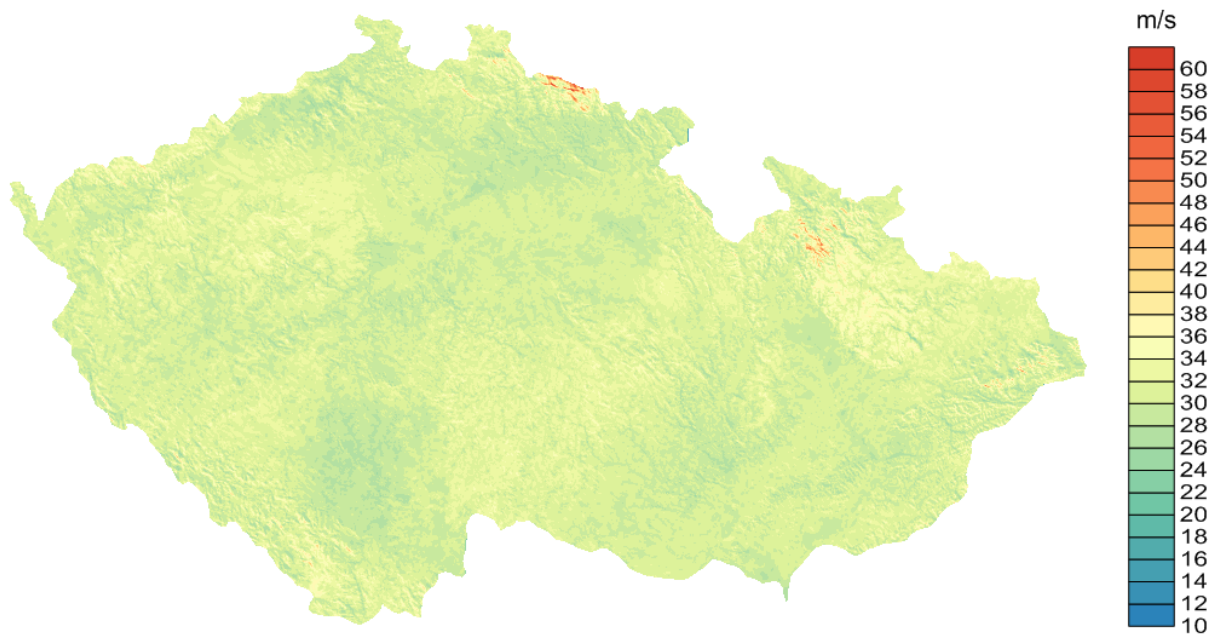
Obr. 4.2.2: Odhad u_{50} ze sektoru č. 2.



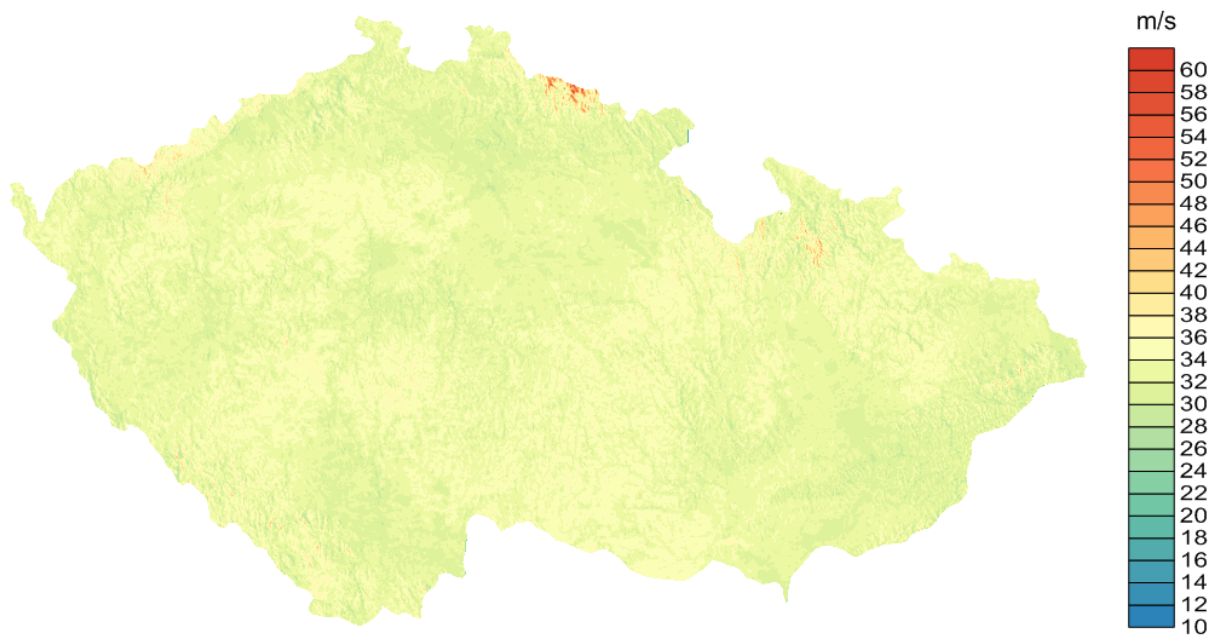
Obr. 4.2.3: Odhad u_{50} ze sektoru č. 3.



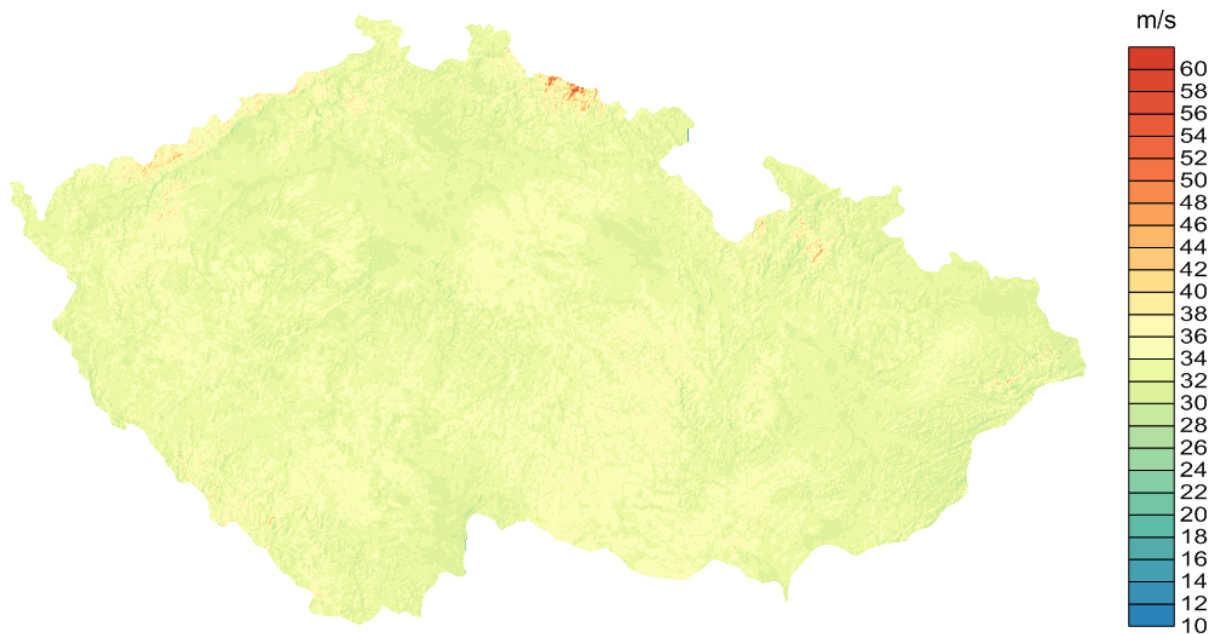
Obr. 4.2.4: Odhad u_{50} ze sektoru č. 4.



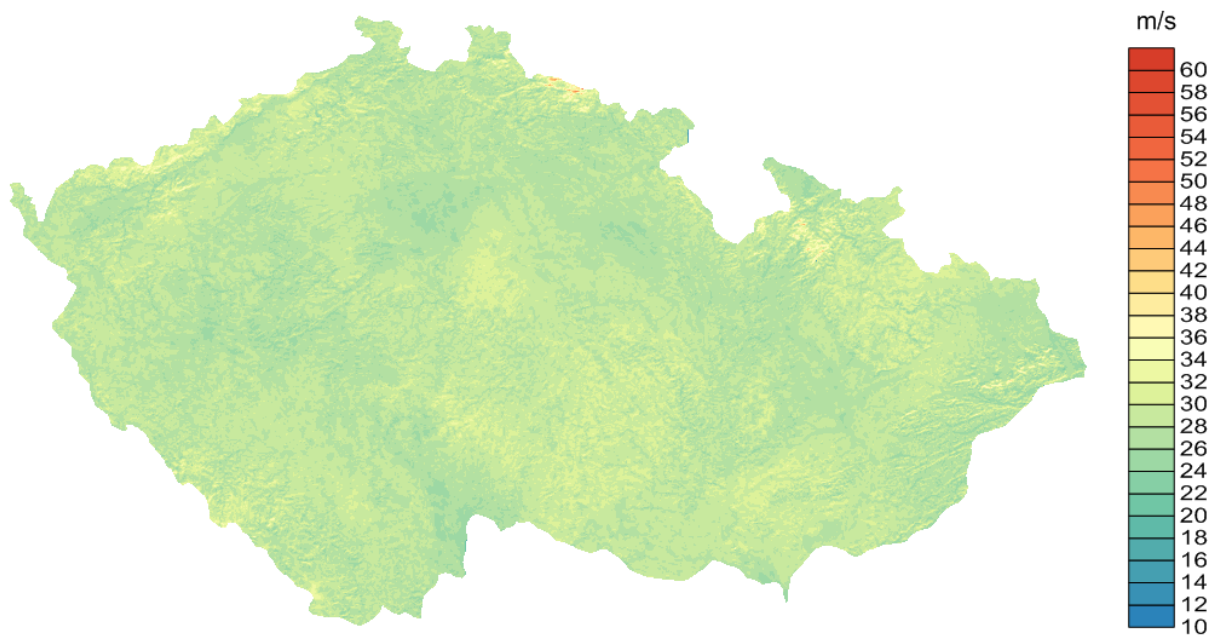
Obr. 4.2.5: Odhad u_{50} ze sektoru č. 5.



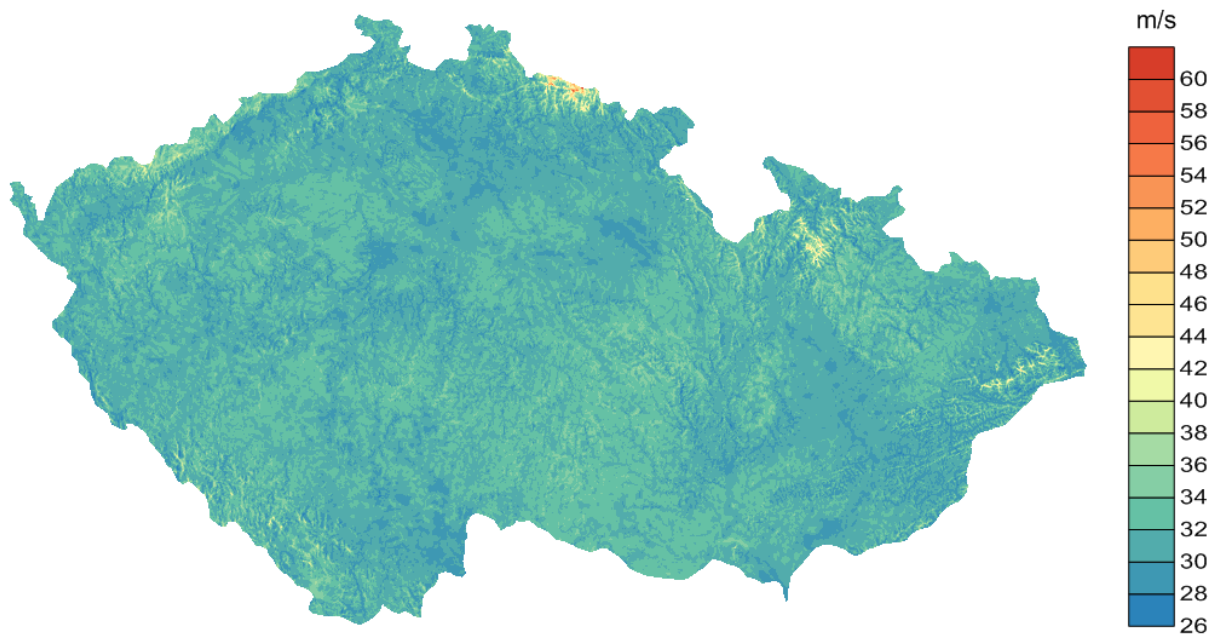
Obr. 4.2.6: Odhad u_{50} ze sektoru č. 6.



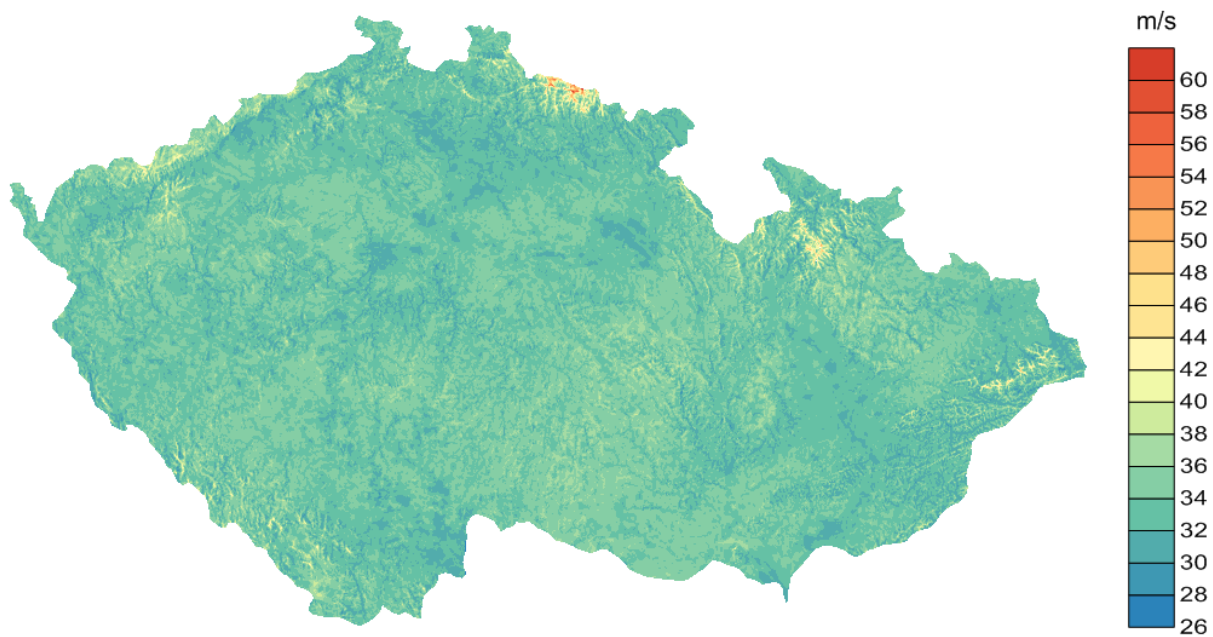
Obr. 4.2.7: Odhad u_{50} ze sektoru č. 7.



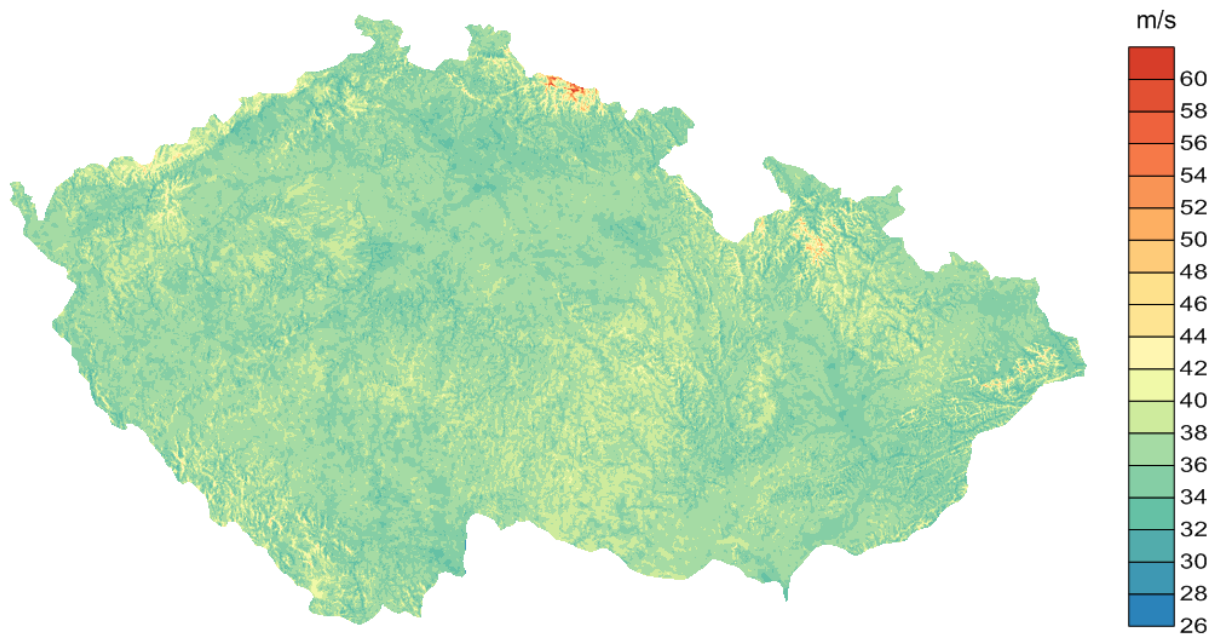
Obr. 4.2.8: Odhad u_{50} ze sektoru č. 8.



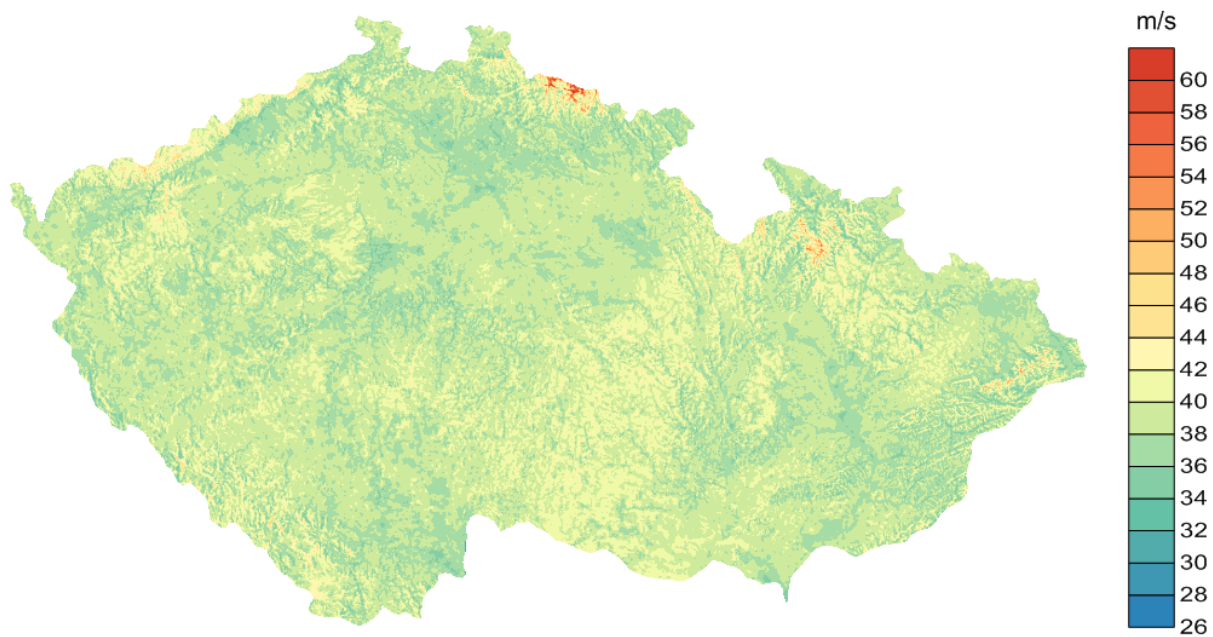
Obr. 4.2.9: Odhad u_{10} .



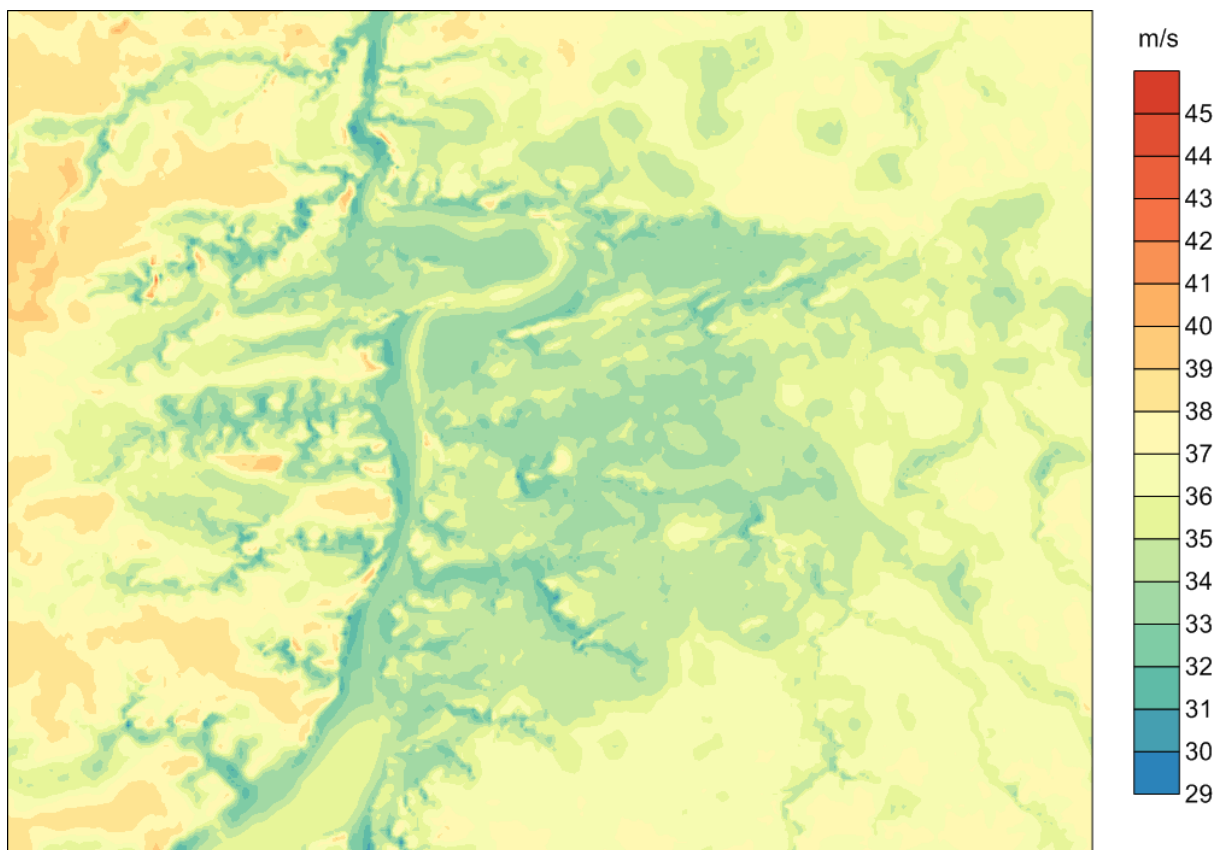
Obr. 4.2.10: Odhad u_{20} .



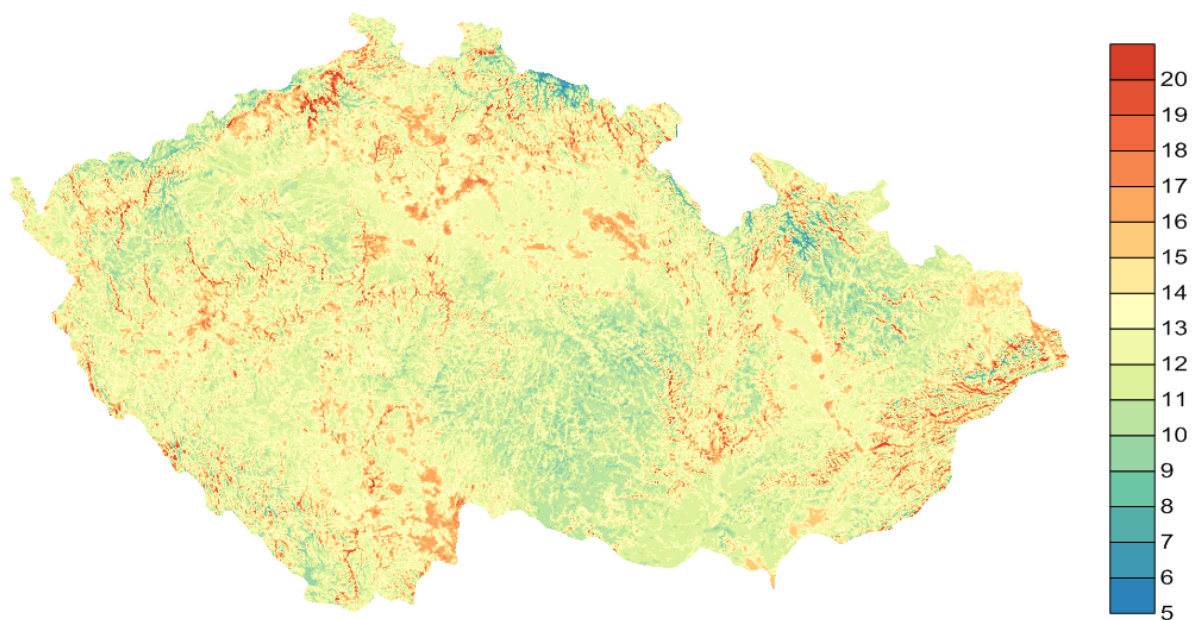
Obr. 4.2.11: Odhad u_{50} .



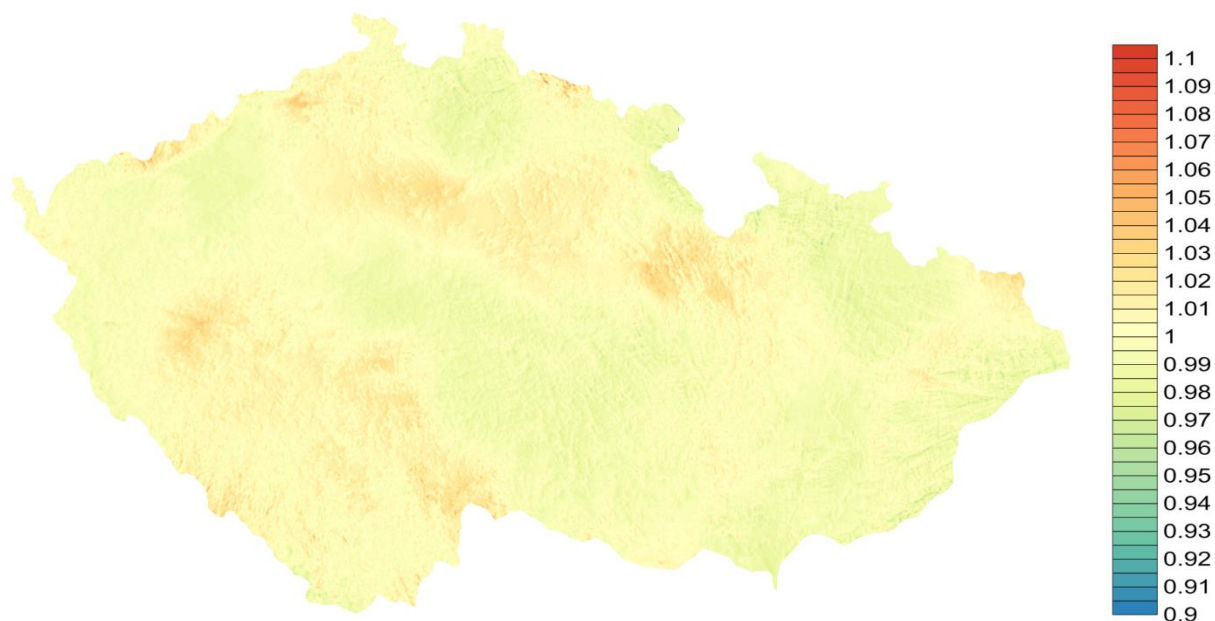
Obr. 4.2.12: Odhad u_{100} .



Obr. 4.2.13: Odhad u_{50} v Praze a jejím okolí.

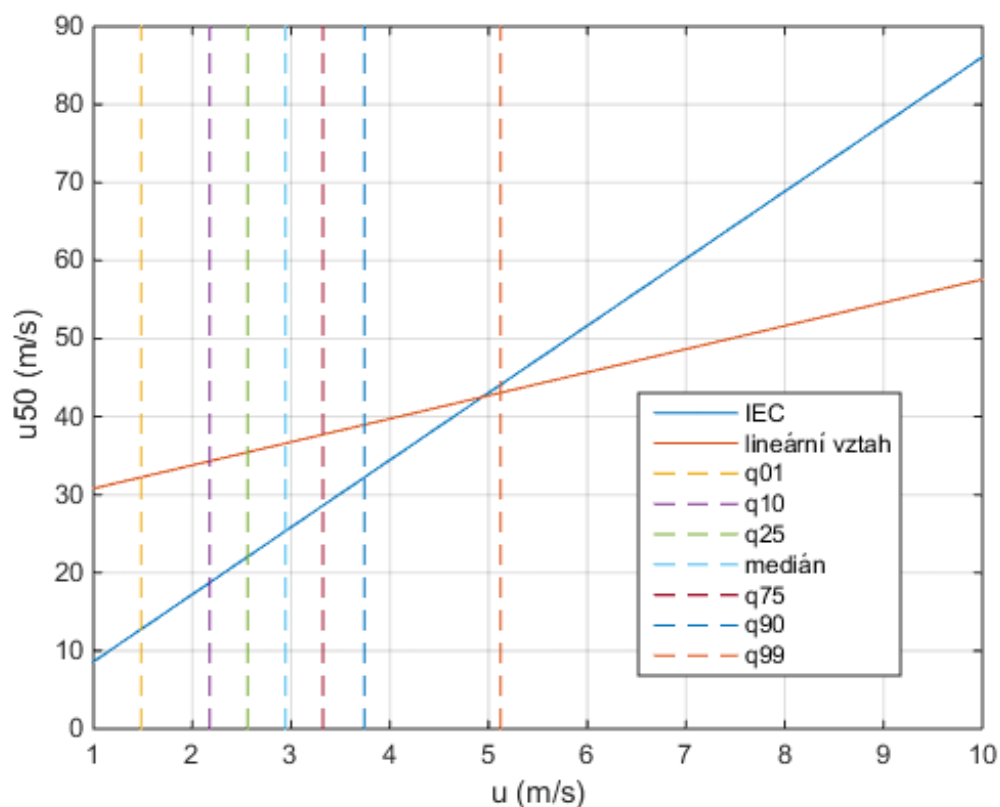


Obr. 4.2.14: Podíl odhadu u_{50} a průměrné rychlosti větru.



Obr. 4.2.15: Podíl odhadu u_{50} podle (4.1.4) a odhadu založeného na zjednodušené regresi pouze na základě průměrné rychlosti podle (4.2.1).

Srovnání vztahů 4.2.1 a 4.2.2 je zobrazeno na grafu na obr. 4.2.16, kde jsou zobrazeny i význačné kvantily průměrné rychlosti větru na území ČR. Pro medián rychlosti je odhad u_{50} podle normy přibližně o 11 m/s vyšší než podle vztahu 4.2.1. Odhad u_{50} podle normy je nižší než námi odvozené hodnoty při průměrných rychlostech větru nižších než necelých 5 m/s, což pokrývá téměř 99 % území ČR. Nicméně v případě extrémně exponovaných lokalit může být odhad u_{50} podle vztahu 4.2.1 naopak až o desítky m/s nižší než podle normy.

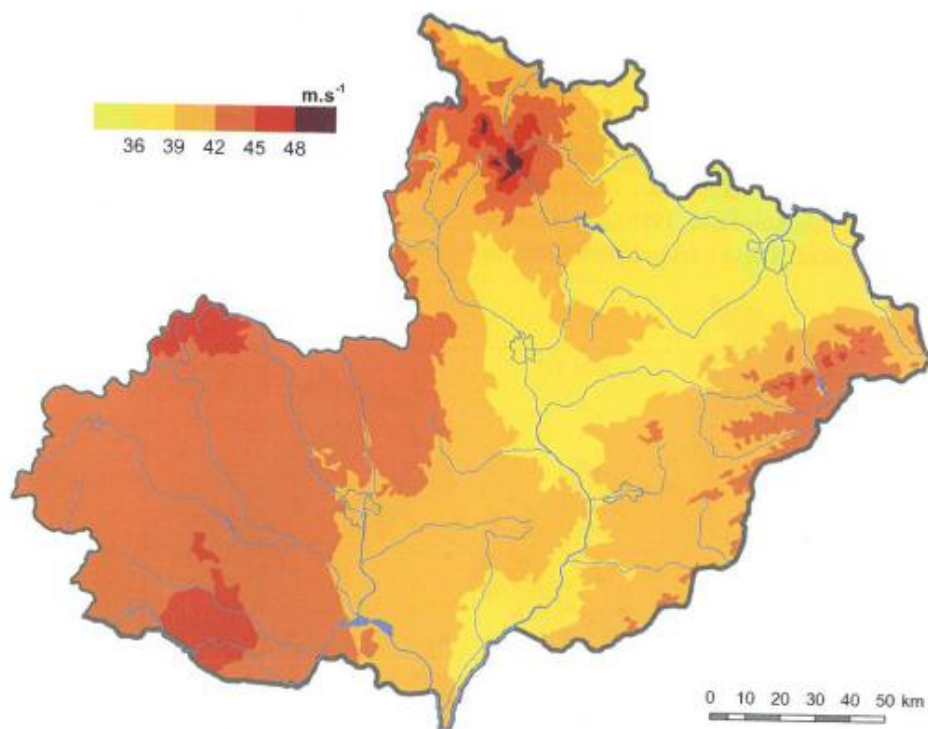


Obr. 4.2.16: Srovnání vztahu mezi průměrnou rychlostí větru a u_{50} podle normy a lineární regresi podle (4.2.1) spolu s kvantily průměrné rychlosti větru na území ČR.

4.2.3 Porovnání s výsledky jiných autorů

Vytvořené mapy extrémního větru byly srovnány s mapami publikovanými v pracích Brázdil et al. (2007) a Návorka (2011).

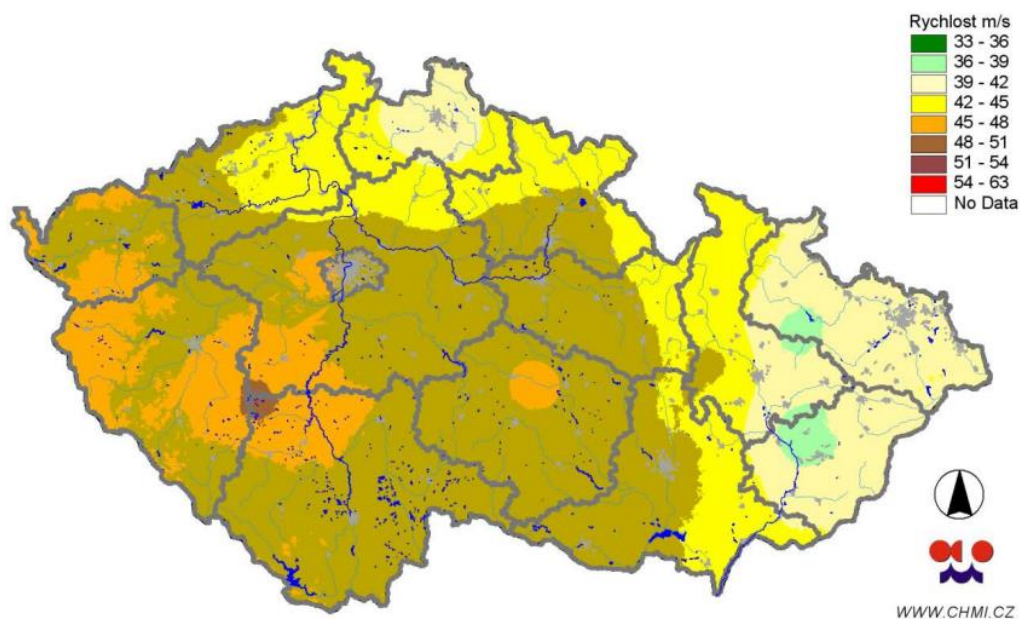
V práci Brázdil et al. (2007) bylo zpracováno pouze území Moravy. Na obr. 4.2.17 je mapa extrémních nárazů větru získaná zpracováním dat z profesionálních stanic ČHMÚ z let 1981 až 2000. Roční maxima byla aproximována Gumbelovým rozdělením a na výsledky této aproximace byla aplikována interpolační metoda založená na určení regresní závislosti mezi hodnotami maximálního nárazu větru a nadmořskou výškou. Na výslednou mapu byl aplikován nízkofrekvenční filtr. Výsledná mapa silně připomíná mapu nadmořské výšky. Nejsou zde, na rozdíl od níže zmíněné mapy na obr. 4.2.18, na první pohled patrná „oka“ kolem jednotlivých stanic snad s výjimkou stanice Kuchařovice. To lze vysvětlit buď odlišnou metodou interpolace, nebo subjektivními opravami, které ale v práci nejsou zmíněny.



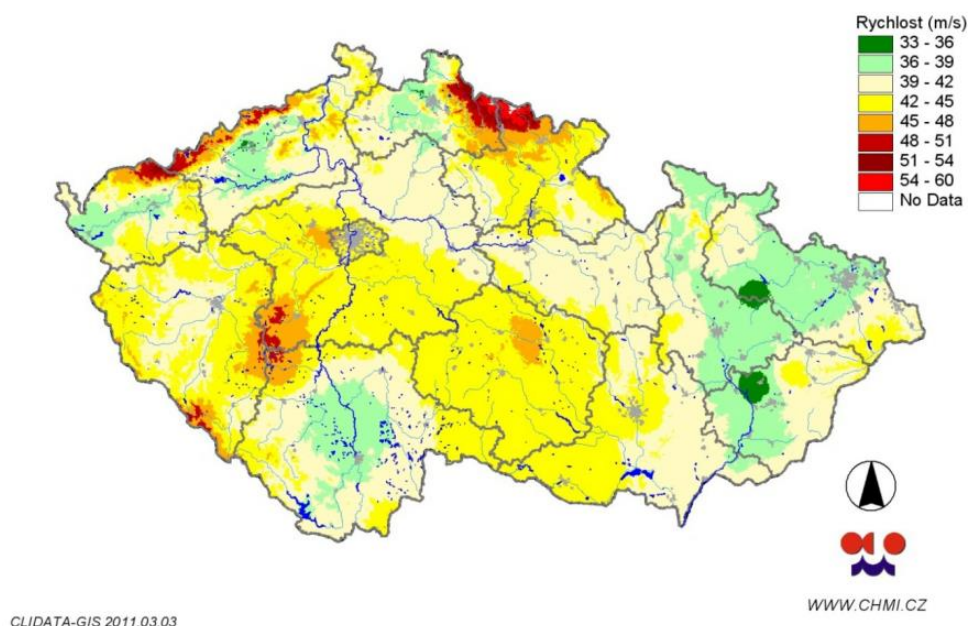
Obr. 4.2.17: Odhad u_{50} podle Brázdila et al. (2007).

V práci Návorka (2011) byla nejprve zpracována data z 16 stanic, které byly vybrány jako „nejkvalitnější“. Většina řad byla z období 1961 až 2010. Pro těchto 16 stanic byla použita aproximace nejvyšších ročních nárazů větru lognormálním rozdělením a pro prostorovou interpolaci byla použita aplikace Clidata-GIS. Výsledná mapa pro u_{100} je na obr. 4.2.18. Je zřejmé na první pohled, že dává naprosto nereálné výsledky. Také došlo pravděpodobně k záměně barev v kategoriích 45 až 48 a 48 až 51 m/s. Za hlavní problém autor práce považoval to, že výběr stanic neobsahuje žádné horské stanice, a proto došlo ke špatnému „zkalibrování“ regresního modelu. Proto byl proveden dopočet hodnot pro 16 dalších převážně vrcholových horských stanic. Práce blíže nespecifikuje, v čem daný „dopočet“ spočíval, ani jak byly tyto doplňkové stanice zvoleny. Maximální roční nárazy z těchto dopočtených stanic byly aproximovány Gumbelovým rozdělením (pro aproximaci Lognormálním i Gumbelovým rozdělením byla v této práci použita metoda maximální věrohodnosti). Výsledek je na obr. 4.2.19 a oproti předchozímu obrázku vypadá podstatně odlišně. Ze stanic byly nakonec odstraněny stanice Červená a Holešov z důvodů „odlišných výsledků od okolních stanic“.

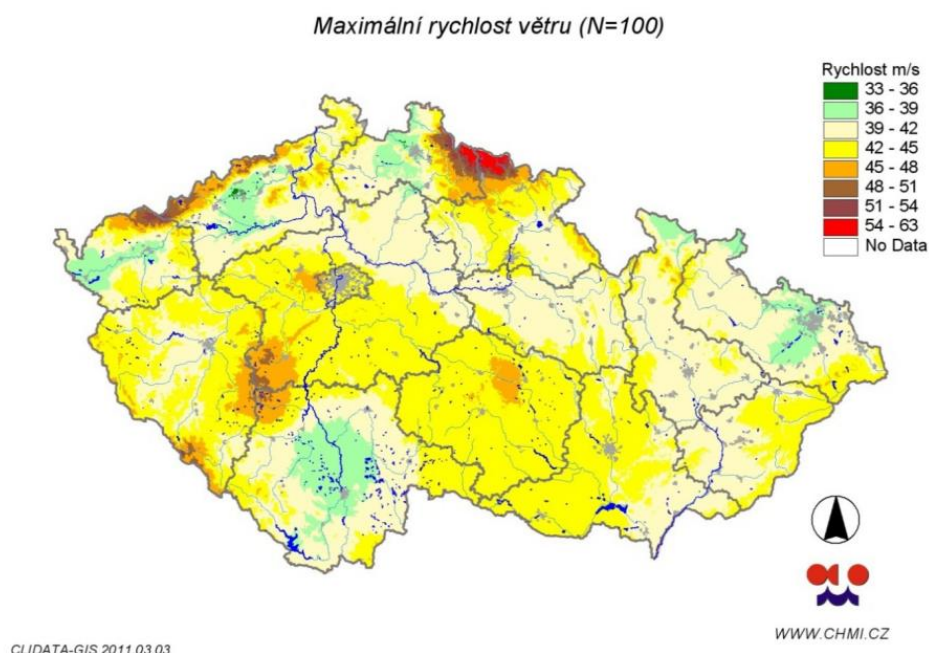
Tím vznikla mapa na obr. 4.2.20, která vypadá oproti původní mapě na obr. 4.2.18 podstatně důvěryhodněji, i když stále vykazuje některé „podezřelé“ rysy. Např. odhadnuté rychlosti extrémního nárazu jsou v oblasti západně od Prahy obdobné jako ve vrcholových partiích Hrubého Jeseníku a Moravskoslezských Beskyd. Oproti mapě z práce Brázdila et al. (2007) se méně shoduje s mapou nadmořské výšky, ale obsahuje více „ok“ kolem použitých stanic.



Obr. 4.2.18: Odhad u_{100} podle Návorka (2011) – původní data.

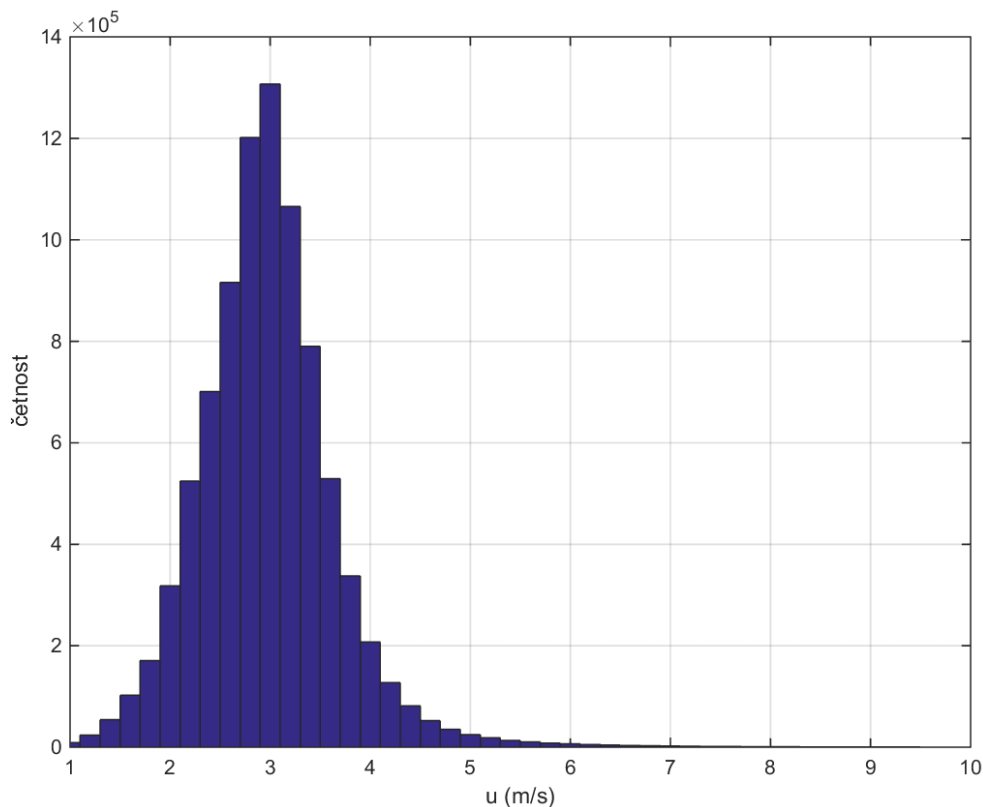


Obr. 4.2.19: Odhad u_{100} podle Návorky (2011) – zahuštěná data.



Obr. 4.2.20: Odhad u_{100} podle Návorky (2011) – zahuštěná data, odstraněny Červená a Holešov.

Problémem takto vytvořených map může být i skutečnost, že stanice s dlouhými řadami měření netvoří reprezentativní vzorek území ČR, ale jsou často umístěny na více exponovaných lokalitách, což může způsobit nadhodnocení odhadnutých rychlostí extrémního větru. Důležitý je také způsob zahrnutí nadmořské výšky v metodě interpolace. Histogram průměrné rychlosti větru na území ČR podle větrné mapy z práce Hanslian et Hošek (2015) je na obr. 4.2.21. Medián má hodnotu 2,95 m/s a průměr 2,98 m/s a rozdělení je téměř symetrické. Existují sice místa s velmi vysokými hodnotami průměrné rychlosti větru, ale jejich celková plocha je velmi malá. 28 stanic použitých pro určení vztahu mezi průměrným a extrémním klimatem má medián průměrné rychlosti větru 3,18 m/s a průměrnou hodnotu 3,52 m/s. Bližší pohled ukazuje, že rozdíl je způsoben několika stanicemi z odlehlého pravého chvostu histogramu. 3 stanice z 28, tedy 10,7 %, mají průměrnou rychlost větru vyšší než 99% kvantil z větrné mapy. Jsou to Milešovka (8,26 m/s), Lysá hora (6,56 m/s) a Svratouch (5,93 m/s).



Obr. 4.2.21: Histogram průměrné rychlosti větru u na území ČR podle větrné mapy z práce Hanslian et Hošek (2015).

Průměrná nadmořská výška použitých stanic je 486 m, její medián 402 m, první kvartil 247 m a třetí kvartil 663 m; průměrná nadmořská výška celého území ČR je 444 m, její medián 434 m, první kvartil 308 m a třetí kvartil 542 m. To znamená, že použité stanice se přednostně nacházejí jak v místech s nízkou, tak i vysokou nadmořskou výškou. Z toho, že mediánová hodnota nadmořské výšky stanic je nižší než mediánová výška území, ale mediánová rychlost je přitom na stanicích vyšší než v případě celého území, lze usuzovat, že stanice jsou umístěny na více exponovaných místech, než je průměr území ČR. Interpolace těchto dat způsobuje kladnou systematickou chybu takto vytvořených map extrémního větru.

Další nevýhodou takto vytvořených map extrémního větru je, že silně závisí na výběru použitých stanic a na metodě interpolace a jejím nastavení, což je zřejmé ze srovnání obr. 4.2.17 a 4.2.18. Tyto volby jsou značně subjektivní. Špatně zvolená metoda

prostorové interpolace se může projevit tím, že výsledná mapa je buď spleť „ok“ kolem naměřených stanic, nebo je příliš závislá na nadmořské výšce.

Navržený způsob výpočtu mapy extrémního větru tyto nedostatky z velké části eliminuje odvozením vztahu mezi VK a EVK a jeho aplikací na mapu VK vytvořenou modelem proudění. Použitý vztah mezi VK a EVK je na rozdíl od verze v normě odvozen přímo pro území ČR. Zlepšením oproti předchozím verzím map je také určení směrové závislosti a podstatně vyšší prostorové rozlišení a menší citlivost na hrubé chyby měření na jednotlivých stanicích.

4.2.4 Odhad chyby vytvořené mapy u_{50}

Pro praktické aplikace mapy extrémního větru je důležitý odhad její chyby. Ta je tvořena několika zdroji, jejichž RMSE se za předpokladu vzájemné nezávislosti kvadraticky sčítají. Přesné určení těchto chyb není možné, proto byl proveden jejich přibližný odhad.

Prvním zdrojem chyby je chyba regrese mezi průměrným a extrémním větrným klimatem podle modelu (4.1.4). Podle kapitoly 4.1 je odhad RMSE určení u_{50} podle tohoto modelu 3,19 m/s.

Druhým zdrojem chyby je vlastní chyba větrné mapy. Pro její přibližné určení byl použit výše uvedený odhad, že 95,3 % rozptylu modelu (4.1.4) se dá vysvětlit vztahem (4.2.1). Z něj a z odhadu chyby určení průměrné rychlosti větru 0,30 m/s podle sekce 4.2.1 plyne, že odhad RMSE způsobený neurčitostí odhadu průměrné rychlosti je 0,89 m/s. Chybu způsobenou chybou určení četnosti větru z jednotlivých směrů lze očekávat podstatně nižší, protože tyto četnosti mají v modelu (4.1.4) výrazně nižší váhu než průměrné rychlosti větru. Proto byla zanedbána.

Dalším zdrojem chyb je chyba způsobená aplikací konkrétního rozdělení extrémních rychlostí. Odhad vycházel z předpokladu, že tato chyba je dána především odlišnou „váhou pravého chvostu“ od Gumbelova rozdělení. Na základě analýzy naměřených dat v podkapitolách 3.9 a 4.1 byl odhad této chyby odvozen z GEV rozdělení s parametrem $\xi = -0,1$ a $\xi = 0,1$ pomocí Monte Carlo simulací provedených v podkapitole 3.8. Výsledná chyba vyšla 1,1 m/s.

Kvadratickým sečtením uvedených tří zdrojů chyb dostaneme odhad RMSE určení u_{50} 3,49 m/s. Odhady druhého (vzhledem k nereprezentativnosti lokalit určených k ověření) a zvláště třetího zdroje chyb jsou zatíženy značnou nejednoznačností, pro

výslednou hodnotu však má určující význam odhad chyby z prvního zdroje, který lze považovat za dostatečně spolehlivý. Získaný odhad platí pro podmínky typické na území ČR. V místech s abnormálně složitými místními vlivy lze očekávat chybu vyšší.

V tomto odhadu není zahrnuta nejistota (chyba) způsobená klimatickými změnami. Dále není zahrnuta chyba způsobená hrubými chybami měření. Vliv hrubých chyb měření však byl eliminován zvoleným postupem, kdy se data ze všech stanic používají při určení parametrů modulu 4.1.4 současně. Současně mohou hrubé chyby měření vést k přecenění odhadu chyby způsobené prvními z uvedených zdrojů.

5. Shrnutí výsledků a závěr

5.1 Hlavní výsledky práce

Za hlavní výsledek této práce lze považovat vztah mezi větrným klimatem (VK) a extrémním větrným klimatem (EVK) odvozený a verifikovaný v podkapitole 4.1 a jeho aplikaci na větrnou mapu vytvořenou dříve v Ústavu fyziky atmosféry, čímž vznikla mapa extrémního větru na území ČR. Vztah mezi VK a EVK byl verifikován křížovou validací na naměřených hodnotách spolu se vztahem uvedeným v normě IEC 61400-1 a z této verifikace vyšel pro území ČR jako podstatně vhodnější. Získaná mapa extrémního větru byla porovnána s obdobnými mapami vytvořenými jinými autory. Přestože přímá verifikace těchto map není možná, lze považovat vytvořenou mapu extrémního mapu za lepší v následujících ohledech:

1. má podstatně vyšší prostorové rozlišení,
2. výrazně eliminuje vliv výběru jednotlivých stanic, jejich místních podmínek a chyb měření,
3. její výpočet je dokumentován velmi podrobně, což umožňuje využít tuto metodiku i pro jiná území a aplikovat ji na větrné mapy spočítané jinými modely proudění a také ji dále vylepšovat,
4. byla odhadnuta její chyba.

Podkapitola 4.1 rovněž obsahuje analýzu vhodnosti použití Gumbelova rozdělení k aproximaci extrémních hodnot rychlosti větru, především v souvislosti s váhou pravého chvostu, a diskutuje možnost případné úpravy metodiky podle výsledků této analýzy. Podle výsledků tohoto zkoumání (a s přihlédnutím k výsledkům kapitol 3.3 a 3.9) lze sice zamítnout hypotézu, že maximální roční rychlosti větru pocházejí z Gumbelova rozdělení, nicméně váha jeho pravého chvostu byla přibližně v souladu s naměřenými hodnotami. Proto byla aproximace naměřených hodnot Gumbelovým rozdělením použita i vzhledem k robustnosti odhadů pomocí tohoto rozdělení a jeho použití v obdobných studiích.

5.2 Výsledky dosažené v kapitolách 2 a 3

V kapitolách 2 a 3 je provedena analýza sloužící k volbě metodiky pro výpočet mapy extrémního větru a pro určení její spolehlivosti. Součástí těchto kapitol je i výběr a

příprava vstupních dat. Určení spolehlivosti výsledků se ukázalo být velmi složité, protože přímé ověření správnosti odhadu extrémních hodnot není možné v důsledku nedostatečné délky měření.

Rešeršní jsou podkapitoly 2.1 o měření větru, 3.1 o době návratu, 3.2 o teorii extrémních hodnot, 3.4 o metodách výběru extrémních hodnot a 3.7 o aproximaci hodnot.

Ostatní podkapitoly obsahují původní poznatky a výsledky studia.

V podkapitole 2.2 bylo na základě měření 3-D sonickým anemometrem kvantifikováno několik systematických chyb měření nejvyšších denních nárazů větru způsobenými různým způsobem zpracování naměřených dat. Tyto výsledky lze použít k opravě nejvyšších denních nárazů větru naměřených různými přístroji, popř. spočítaných různými metodikami.

Podkapitola 2.3 popisuje homogenizaci nejvyšších denních nárazů větru z profesionálních stanic ČHMÚ a její výsledek. Homogenizované hodnoty jsou použity v podkapitolách 2.4, 3.9 a 4.1.

Podkapitola 2.4 se věnuje výpočtu základních statistických vlastností nárazů větru.

V podkapitole 3.3 je provedena diskuze praktického použití GEV rozdělení k aproximaci rozdělení extrémních hodnot.

Podkapitola 3.5 vyčísluje systematickou chybu mezi dvěma různými definicemi doby návratu vycházející ze všech hodnot nebo pouze z blokových (obvykle ročních) maxim, a navrhuje opravu, která byla použita v kapitole 4.1. Vypočítaný rozdíl je ve většině praktických případů poměrně malý, ale je užitečné vědět, že existuje.

V podkapitole 3.6 je na jednoduchém uměle vytvořeném příkladu ilustrován vliv tvaru růžice na vztah mezi VK a EVK. Tento vliv není na rozdíl od vztahu odvozeného v podkapitole 4.1 respektován ve vztahu podle normy IEC 61400-1.

V podkapitolách 3.8 a 3.9 jsou provedeny numerické simulace aproximace dat Gumbelovým a GEV rozdělením metodami L-momentů a maximální věrohodnosti a je provedena analýza jejich výsledků. V podkapitole 3.8 jsou analyzována uměle vytvořená data metodou Monte Carlo. Výhodou tohoto postupu je, že jsou známy správné hodnoty aproximovaného rozdělení. Výsledky této kapitoly byly využity při odhadu chyby extrémní větrné mapy. V podkapitole 3.9 jsou analyzována data získaná metodou bootstrap z naměřených hodnot ročních maxim rychlosti větru. Bylo zjištěno, že analýza

těchto dat vede k některým problémům nepozorovaným v případě uměle vygenerovaných dat. Rovněž rozdíly mezi metodami L-momentů a maximální věrohodnosti jsou v tomto případě podstatně větší než v případě uměle generovaných dat. Jako nejrobustnější se ukázala aproximace Gumbelovým rozdělením metodou L-momentů, která proto byla použita v podkapitole 4.1 s tím, že použití tohoto rozdělení bylo následně ještě diskutováno, jak je uvedeno v předchozí sekci 5.1.

5.3 Výhled dalšího výzkumu

Z provedené analýzy vyplývá velký prostor k dalšímu vývoji metody homogenizace popsané v podkapitole 2.3, pokud se podaří získat data pro její lepší verifikaci. Těmito daty mohou být jednak vysokofrekvenční měření obdobná měřením z podkapitoly 2.2, která byla bohužel příliš krátká pro ověření provedené metody homogenizace, a dále souběžná měření různými anemometry, která bohužel nebyla k dispozici.

Obdobné numerické simulace jako v podkapitolách 3.8 a 3.9 je možno provést i pro extrémní data získaná metodou Peak-over-Threshold, a výsledky porovnat.

Odvození vztahu mezi větrným klimatem a extrémním větrným klimatem uvedené v podkapitole 4.1 je možné relativně snadno provést i pro jiná území než ČR. Použití více dat by mohlo vést k dalšímu zpřesnění vztahu. Také by bylo zajímavé zjistit zda tento vztah vykazuje nějakou významnou regionální variabilitu. Pokud budou k dispozici vhodná data, bude možno obdobným způsobem odvodit i vztah mezi VK a EVK definovaným pomocí jiné charakteristiky, než jsou nárazy větru, např. podle maximálního desetiminutového průměru rychlosti větru. Při použití dat z více stanic také bude možno znovu otevřít otázku nalezení vhodnějšího tvaru rozdělení extrémních hodnot rychlosti větru než je rozdělení Gumbelovo.

Při vydání nových větrných map by mělo být poměrně snadné a přímočaré z těchto map pomocí vztahu z podkapitoly 4.1 spočítat mapy extrémního větru. Tyto mapy by mohly zahrnovat vliv jevů, které model v použité větrné mapě není schopen postihnout, např. termických větrů.

Seznam použité literatury

- Abild, J., Andersen, E. Y. and Rosbjerg, D. (1992). The climate of extreme winds at the Great Belt, Denmark. *Journal of Wind Engineering and Industrial Aerodynamics*, 41(1), 521-532.
- Ahmad, M. I., Sinclair, C. D. and Spurr, B. D. (1988). Assessment of flood frequency models using empirical distribution function statistics. *Water Resources Research*, 24(8), 1323-1328.
- Alexandersson, H. (1986). A homogeneity test applied to precipitation data. *Journal of climatology*, 6(6), 661-675.
- An, Y. and Pandey, M. D. (2007). The r largest order statistics model for extreme wind speed estimation. *Journal of wind engineering and industrial aerodynamics*, 95(3), 165-182.
- Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics*, 193-212.
- AghaKouchak, A., Easterling, D., Hsu, K., Shubert, S. and Sorooshian, S. (2012): Extremes in a changing Climate: Detection, Analysis and Uncertainty, *Springer*, 426 pp.
- Aguilar, E., Auer, I., Brunet, M., Peterson, T. C. and Wieringa, J. (2003). Guidance on metadata and homogenization. *WMO TD, 1186*, 53 pp.
- Ayotte, K. W. (2008). Computational modelling for wind energy assessment. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(10), 1571-1590.
- Azorin-Molina, C., Vicente-Serrano, S. M., McVicar, T. R., Jerez, S., Sanchez-Lorenzo, A., ... and Espírito-Santo, F. (2014). Homogenization and Assessment of Observed Near-Surface Wind Speed Trends over Spain and Portugal, 1961–2011. *Journal of Climate*, 27(10), 3692-3712.
- Barth, A., Joassin, P., Beckers, J. M. and Troupin, C. (2008). Introduction to Optimal Interpolation and Variational Analysis, 68 pp.
- Bégin-Drolet, A., Ruel, J. and Lemay, J. (2011). Off-axis characterization of ice-free anemometers. *Journal of Wind Engineering and Industrial Aerodynamics*, 99(8), 825-832.
- Begert, M., Zenklusen, E., Häberli, C., Appenzeller, C. and Klok, L. (2008). An automated procedure to detect discontinuities; performance assessment and application to a large European climate data set. *Meteorologische Zeitschrift*, 17(5), 663-672.

- Beirlant, J., Goegebeur, J. S. and Teugels, J. (2004): Statistics of extremes: theory and applications, *Wiley*, 514 pp.
- Brázdil, R., Chromá, K., Dobrovolný, P. and Tolasz, R. (2009). Climate fluctuations in the Czech Republic during the period 1961–2005. *International Journal of Climatology*, 29(2), 223-242.
- Brázdil, R., Dobrovolný, P., Štekl, J., Kotyza, O., Valášek, H., Jež, J. (2004). History of Weather and Climate in the Czech Lands VI: Strong Winds, *Masarykova Univerzita*, 378 pp.
- Brázdil, R., Kirchner, K. a kol. (2007). Vybrané přírodní extrémny a jejich dopady na Moravě a ve Slezsku, Masarykova univerzita, *Český Hydrometeorologický ústav, Ústav geoniky Akademie věd ČR, v. v. i.* , 435 pp.
- Bush N. E. and Kristensen L. (1976): Cup Anemometer Overspeeding. *Journal of Applied Meteorology*, 15, 1328-1332.
- Camuffo, D. (2002). Errors in early temperature series arising from changes in style of measuring time, sampling schedule and number of observations. In *Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources* (pp. 331-352). *Springer Netherlands*.
- Chaudhuri, S. and Middey, A. (2012). A composite stability index for dichotomous forecast of thunderstorms. *Theoretical and Applied Climatology*, 110(3), 457-469.
- Choi, E. C. and Hidayat, F. A. (2002). Gust factors for thunderstorm and non-thunderstorm winds. *Journal of wind engineering and industrial aerodynamics*, 90(12), 1683-1696.
- Chung K. L. (2000). A Course in Probability Theory, Third Edition. *Academic Press*, 419 pp.
- Coles, S., Bawa, J., Trenner, L. and Dorazio, P. (2001). An introduction to statistical modeling of extreme values (Vol. 208). London: *Springer*, 224 pp.
- Coles, S. G. and Walshaw, D. (1994). Directional modelling of extreme wind speeds. *Applied Statistics*, 139-157.
- Cook, N. J. (1982). Towards better estimation of extreme winds. *Journal of Wind Engineering and Industrial Aerodynamics*, 9(3), 295-323.
- Cook, N. J. (1983). Note on directional and seasonal assessment of extreme winds for design. *Journal of wind engineering and industrial aerodynamics*, 12(3), 365-372.
- Cook, N. (2011). Comments on "Plotting positions in extreme value analysis". *Journal of Applied Meteorology and Climatology*, 50(1), 255-266.

- Costa, A. C. and Soares, A. (2009). Homogenization of climate data: review and new perspectives using geostatistics. *Mathematical Geosciences*, 41(3), 291-305.
- Costa P., Miranda P. and Estanqueiro A. (2006). Development and validation of the Portuguese Wind Atlas, *EWEC Proceedings*, Athens, Greece.
- ČHMÚ (2007). Metodický návod NMK č.j.: P07211000062, Zpráva SYNOP.
- ČMES (2015). Meteorologický slovník výkladový a terminologický (eMS), ČMeS, dostupný na: <http://slovník.cmes.cz>.
- Darling, D. A. (1957). The kolmogorov-smirnov, cramer-von mises tests. *The Annals of Mathematical Statistics*, 823-838.
- David H. A. and Nagaraja H. N. (2003). Order Statistics – 3rd editon, *Wiley-Interscience*, 488 pp.
- Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, 393-442.
- Deaves, D. M. (1993). Analysis of gust factors for use in assessing wind hazard. *Journal of Wind Engineering and Industrial Aerodynamics*, 45(2), 175-188.
- Della-Marta, P. M. and Wanner, H. (2006). A method of homogenizing the extremes and mean of daily temperature measurements. *Journal of Climate*, 19(17), 4179-4197.
- Dellwik, E., Mann, J. and Larsen, K. S. (2010). Flow tilt angles near forest edges–Part 1: Sonic anemometry. *Biogeosciences*, 7(5), 1745-1757.
- Diffenbaugh, N. S., Scherer, M. and Trapp, R. J. (2013). Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proceedings of the National Academy of Sciences*, 110(41), 16361-16366.
- Domonkos, P. and Štěpánek, P. (2009). Statistical characteristics of detectable inhomogeneities in observed meteorological time series. *Studia Geophysica et Geodaetica*, 53(2), 239-260.
- Durst, C. S. (1960). Wind speeds over short periods of time. *Meteorological Magazine*, 89(1960), 181-186.
- European Environment Agency (2006). Land accounts for Europe 1990-2000: Towards intergrated land and ekosystém accounting. *EEA Report 11/2006*, 107 pp.
- Francis, J. A. and Vavrus, S. J. (2012). Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophysical Research Letters*, 39(6).

- Fueyo, N., Sanz, Y., Rodrigues, M., Montañés, C. and Dopazo, C. (2010). High resolution modelling of the on-shore technical wind energy potential in Spain. *Wind energy*, 13(8), 717-726.
- Galambos, J. and Seneta, E. (1973). Regularly varying sequences. *Proceedings of the American Mathematical Society*, 41(1), 110-116.
- Ginger, J. D., Cechet, B., Sanabria, A., Kepert, J., Holmes, J. and Henderson, D. (2011). Extreme wind speed baseline climate investigation project. *A report for the Commonwealth Department of Climate Change and Energy Efficiency, Australia*.
- Greenwood, J. A., Landwehr, J. M., Matalas, N. C. and Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15(5), 1049-1054.
- Grell, G. A., Dudhia, J. and Stauffer, D. R. (1994). A description of the fifth-generation Penn State/NCAR mesoscale model (MM5), 128 pp.
- Gruber, C. and Haimberger, L. (2008). On the homogeneity of radiosonde wind time series. *Meteorologische Zeitschrift*, 17(5), 631-643.
- Guentchev, G., Barsugli, J. J. and Eischeid, J. (2010). Homogeneity of gridded precipitation datasets for the Colorado River basin. *Journal of Applied Meteorology and Climatology*, 49(12), 2404-2415.
- Gujarati, D., N. (2004). Basic Econometrics, 4th ed., New York, *McGraw-Hill*, 1032 pp.
- Gumbel E. J. (1958). Statistics of Extremes. *Columbia University Press*, 375 pp.
- Hanel, M., Buishand, T. A. and Ferro, C. A. (2009). A nonstationary index flood model for precipitation extremes in transient regional climate model simulations. *Journal of Geophysical Research: Atmospheres* (1984–2012), 114(D15).
- Hanslian D. (2014). Analýza výsledků měření větru. Disertační práce, *Univerzita Karlova v Praze, Matematicko-fyzikální fakulta*, 219 pp.
- Hanslian, D., Chládková, Z., Pop, L. and Hošek, J. (2012). Modely pro konstrukci větrných map v ČR, *Meteorologické zprávy*, ročník 65, číslo 2, 36-44.
- Hanslian, D. and Hošek, J. (2015). Combining the VAS 3D interpolation method and Wind Atlas methodology to produce a high-resolution wind resource map for the Czech Republic. *Renewable Energy*, 77, 291-299.
- Hanslian D., Hošek, J., Chládková Z. and Pop, L. (2013a). Větrné podmínky v České republice ve výšce 10 m nad povrchem I. *TZB-info*, 15.4.2013. <http://oze.tzb-info.cz/vetrna-energie/9770-vetrne-podminky-v-ceske-republice-ve-vysce-10-m-nad-povrchem-i>.

- Hanslian D., Hošek, J., Chládová Z. and Pop, L. (2013b). Větrné podmínky v České republice ve výšce 10 m nad povrchem II. *TZB-info*, 22.4.2013. <http://oze.tzb-info.cz/vetrna-energie/9770-vetrne-podminky-v-ceske-republice-ve-vysce-10-m-nad-povrchem-ii>.
- Harper B. A., Kepert J. D., Ginger J. D. (2008). Guidelines for converting between various wind averaging periods in tropical cyclone conditions, *World Meteorological Organization*, 60 pp.
- Harris, R. I. (1996). Gumbel re-visited-a new look at extreme value statistics applied to wind speeds. *Journal of Wind Engineering and Industrial Aerodynamics*, 59(1), 1-22.
- Harris, R. I. (1999). Improvements to the Method of Independent Storms. *Journal of Wind Engineering and Industrial Aerodynamics*, 80(1), 1-30.
- Hoff P. D. (2009). A First Course in Bayesian Statistical Methods, *Springer Texts in Statistics*, 272 pp.
- Holmes, J. D., Allsop, A. C. and Ginger, J. D. (2014). Gust durations, gust factors and gust response factors in wind codes and standards. *Wind and Structures*, 19(3), 339-352.
- Holton J. R. (2004). An Introduction do Dynamic Meteorology, Fourth Edition, *Academic Press*, 535 pp.
- Hosking, J. R. (1984). Testing whether the shape parameter is zero in the generalized extreme-value distribution. *Biometrika*, 71(2), 367-374.
- Hosking J. R. M. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B. Methodological*, 52, 105-124
- IEC 61400-1 (2005). Wind Turbines Standarts – Design requirements – 3rd ed., *International Electrotechnical Commision*, 99 pp.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., ... and Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society*, 77(3), 437-471.
- Kalvová, J. and Sobíšek, B. (1981). Periodicity in time series of wind direction data. *Studia Geophysica et Geodaetica*, 25(3), 275-283.
- Kotz S. and Nadarajah N. (2000). Extreme Value Distributions – Theory and Applications, *World Scientific Publishing Company*, 185 pp.
- Kozumplík S. (1950). Univerzální anemograf a anemografické dvojče. *Meteorologické zprávy*, ročník 4, číslo 3-4, 93-98.

- Král, J. (2007). Ten-minute wind speeds and gusts in the Czech Republic. *Journal of Wind Engineering and Industrial Aerodynamics*, 95(9), 1216-1228.
- Kristensen L. (1993). The Cup Anemometer and Other Exciting Instruments. *Risø National Laboratory*, 85 pp.
- Kristensen, L. (1998). Cup anemometer behavior in turbulent environments. *Journal of Atmospheric and Oceanic Technology*, 15(1), 5-17.
- Kwon, D. K. and Kareem, A. (2014). Revisiting gust averaging time and gust effect factor in ASCE 7. *Journal of Structural Engineering*, 140(11).
- Kysely, J., Gaál, L. and Picek, J. (2011). Comparison of regional and at-site approaches to modelling probabilities of heavy precipitation. *International Journal of Climatology*, 31(10), 1457-1472.
- Landberg, L. (2000). The mast on the house. *Wind Energy*, 3(3), 113-119.
- Landberg, L., Myllerup, L., Rathmann, O., Petersen, E. L., Jørgensen, B. H., Badger, J. and Mortensen, N. G. (2003). Wind resource estimation—an overview. *Wind Energy*, 6(3), 261-271.
- Larsén, X. G. and Mann, J. (2006). The effects of disjunct sampling and averaging time on maximum mean wind speeds. *Journal of wind engineering and industrial aerodynamics*, 94(8), 581-602.
- Lewis, P. A. (1961). Distribution of the Anderson-Darling statistic. *The Annals of Mathematical Statistics*, 1118-1124.
- Lombardo, F. T., Main, J. A. and Simiu, E. (2009). Automated extraction and classification of thunderstorm and non-thunderstorm wind data for extreme-value analysis. *Journal of Wind Engineering and Industrial Aerodynamics*, 97(3), 120-131.
- Mantel, N. and Myers, M. (1971). Problems of convergence of maximum likelihood iterative procedures in multiparameter situations. *Journal of the American Statistical Association*, 66(335), 484-491.
- Miller, C., Holmes, J., Henderson, D., Ginger, J. and Morrison, M. (2013). The response of the dines anemometer to gusts and comparisons with cup anemometers. *Journal of Atmospheric and Oceanic Technology*, 30(7), 1320-1336.
- Moriarty, W. W. and Templeton, J. I. (1983). On the estimation of extreme wind gusts by direction sector. *Journal of Wind Engineering and Industrial Aerodynamics*, 13(1), 127-138.

- Mortensen, N. G., Landberg L., Troen, I. and Petersen, E., L. (1993). Wind atlas analysis and application program (WASP): Vol. 2.: User's guide, *RisØ National Laboratory*, 133 pp.
- NÁVORKA, M. (2011). Aplikace statistických metod na problematiku predikce nárazů větru – disertační práce, *ZČU Plzeň*, 2011, 164 s.
- Pahlow M. and Parlange M. B. (2001). 3-D sonic anemometer Standard Operating Procedure (SOP) for the Baltimore PM Supersite, *Johns Hopkins Univerzity*, 6 pp.
- Palutikof, J. P., Brabson, B. B., Lister, D. H. and Adcock, S. T. (1999). A review of methods to calculate extreme wind speeds. *Meteorological applications*, 6(02), 119-132.
- Papadopoulos, K. H., Stefanos, N. C., Paulsen, U. S. and Morfiadakis, E. (2001). Effects of turbulence and flow inclination on the performance of cup anemometers in the field. *Boundary-Layer Meteorology*, 101(1), 77-107.
- Perera, M. D. A. E. S. (1981). Shelter behind two-dimensional solid and porous fences. *Journal of Wind Engineering and Industrial Aerodynamics*, 8(1), 93-104.
- Perrin, O., Rootzén, H. and Taesler, R. (2006). A discussion of statistical methods used to estimate extreme wind speeds. *Theoretical and applied climatology*, 85(3-4), 203-215.
- Pešice, P., Sulan, J. and Řezáčová, D. (2003). Convection precursors in the Czech territory. *Atmospheric research*, 67, 523-532.
- Pindado, S., Barrero-Gil, A. and Sanz, A. (2012). Cup anemometers' Loss of performance due to ageing processes, and its effect on Annual Energy Production (AEP) estimates. *Energies*, 5(5), 1664-1685.
- Pop L. (2007). Vertikální profil nárazů větru podle stožárového měření, *Větrná energie*, 23, 16-17.
- Reiss, R. D. and Thomas, M. (2001). Statistical analysis of extreme values – 2nd edition. *Birkhauser Boston*, 472 pp.
- Rhoades, D. A. and Salinger, M. J. (1993). Adjustment of temperature and rainfall records for site changes. *International Journal of climatology*, 13(8), 899-913.
- Robert, C. and Casella, G. (2004). Monte Carlo Statistical Methods 2nd ed., *Springer texts in Statistics*, 649 pp.
- Rogers, A. L., Rogers, J. W. and Manwell, J. F. (2005). Comparison of the performance of four measure–correlate–predict algorithms. *Journal of Wind Engineering and Industrial Aerodynamics*, 93(3), 243-264.

- Řepka M. (2011). Přehled měření větru v České republice. *Meteorologické zprávy*, ročník 64, číslo 4, 97-106 .
- Sathyajith, M. (2006). *Wind Energy: Fundamentals, Resource Analysis and Economics*, Springer, 246 pp.
- Siedlecki, M. (2009). Selected instability indices in Europe. *Theoretical and applied climatology*, 96(1-2), 85-94.
- Sokol, Z. and Štekl, J. (1995). Estimation of annual mean ground wind speed over the territory of the Czech Republic. *Meteorologische Zeitschrift*, 4(5), 218-222.
- Sokol, Z., Štekl, J., Hanslian, D. and Hošek, J. (2007). Časová a prostorová variabilita rychlosti větru se zřetelem na činnost větrných elektráren. Výzkumná zpráva pro ČEPS, a. s., ÚFA AV ČR, v.v.i., 60 pp.
- Sterl, A. (2004). On the (in) homogeneity of reanalysis products. *Journal of Climate*, 17(19), 3866-3873.
- Sugahara, S., Da Rocha, R. P., Ynoue, R. Y. and da Silveira, R. B. (2012). Homogeneity assessment of a station climate series (1933–2005) in the Metropolitan Area of São Paulo: instruments change and urbanization effects. *Theoretical and Applied Climatology*, 107(3-4), 361-374.
- Svoboda, J., Chladova, Z., Pop, L. and Hosek, J. (2013). Statistical-dynamical downscaling of wind roses over the Czech Republic. *Theoretical and applied climatology*, 112(3-4), 713-722.
- Štěpánek, P., Řezníčková, L., Brázdil, R. (2009). Homogenization of daily air pressure and temperature series for Brno (Czech Republic) in the period 1848–2005. In: *Proceedings of the Fifth seminar for homogenization and quality control in climatological databases* (Budapest, Hungary, 29 May – 2 June 2006), WCDMP-No. 71. WMO, Genova, 107-122.
- Taylor, P. A. and Salmon, J. R. (1993). A model for the correction of surface wind data for sheltering by upwind obstacles. *Journal of Applied Meteorology*, 32(11), 1683-1694.
- Tolasz, R., Brázdil, R., Bulíř, O., Dobrovolný, P., Dubrovský, M., ... and Žalud, Z (2007). Atlas podnebí Česka. 1. vydání. Praha, Olomouc: Český hydrometeorologický ústav, *Universita Palackého*, 256 pp.
- Troen, I. E. L. P. and Petersen, E. L. (1989). European wind atlas, *Risø National Laboratory*, 660 pp.
- Twisdale, L. A. and Vickery, P. J. (1992). Research on thunderstorm wind design parameters. *Journal of Wind Engineering and Industrial Aerodynamics*, 41(1), 545-556.

- Uppala, S. M., Kállberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V., Fiorino, M., ... and Woollen, J. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612), 2961-3012.
- Walmsley, J. L., Troen, I. B., Lalas, D. P. and Mason, P. J. (1990). Surface-layer flow in complex terrain: comparison of models and full-scale observations. *Boundary-Layer Meteorology*, 52(3), 259-281.
- Wang, X. L., Chen, H., Wu, Y., Feng, Y. and Pu, Q. (2010). New techniques for the detection and adjustment of shifts in daily precipitation data series. *Journal of Applied Meteorology and Climatology*, 49(12), 2416-2436.
- Wieringa, J. (1980). Representativeness of wind observations at airports. *Bulletin of the American Meteorological Society*, 61(9), 962-971.
- Wieringa, J. (1983). Description requirements for assessment of non-ideal wind stations—for example Aachen. *Journal of Wind Engineering and Industrial Aerodynamics*, 11(1), 121-131.
- Wieringa, J. (1993). Representative roughness parameters for homogeneous terrain. *Boundary-Layer Meteorology*, 63(4), 323-363.
- Wieringa, J. (1996). Does representative wind information exist?. *Journal of Wind Engineering and Industrial Aerodynamics*, 65(1), 1-12.
- Wieser, A., Fiedler, F. and Corsmeier, U. (2001). The influence of the sensor design on wind measurements with sonic anemometer systems. *Journal of atmospheric and oceanic technology*, 18(10), 1585-1608.
- Wilks D. S. (2011). Statistical methods in the atmospheric sciences – 3rd ed., *Academic Press*, 704 pp.
- WMO (2010). Guide to meteorological instruments and methods of observation, 2008 edition updated in 2010, *World Meteorological organization*, 716 pp.
- Żurański, J. A. and Jaśpińska, B. (1996). Directional analysis of extreme wind speeds in Poland. *Journal of wind engineering and industrial aerodynamics*, 65(1), 13-20.

Přílohy

Příloha A: Problém domén atrakce

Tato příloha popisuje řešení problému domén atrakce rozdělení extrémních hodnot představeného v podkapitole 3.3. Toto řešení se v literatuře objevuje v různých tvarech, zde bude uvedena verze převzatá z knihy Beirlant et al. (2004), a to zvláště pro Fréchetovo, Gumbelovo a Weibullovo rozdělení, protože obecná formulace pro GEV rozdělení sice existuje, ale je komplikovanější. Konvergencí blokových maxim ke GEV rozdělení budeme v tomto dodatku rozumět konvergenci ve smyslu vztahu (3.2.1).

A.1 Nutná a postačující kritéria konvergence

(i) Případ $\xi > 0$ (Fréchetovo rozdělení): blokovaná maxima nezávislých náhodných hodnot s distribuční funkcí F konvergují ke GEV rozdělení s kladným parametrem tvaru ξ právě tehdy, pokud pro nějakou kladnou konstantu w platí:

$$\frac{1-F(xw)}{1-F(x)} \rightarrow w^{-1/\xi}, \quad (\text{A.1})$$

pro $x \rightarrow \infty$.

(ii) Případ $\xi < 0$ (Weibullovo rozdělení): pro tento případ budeme potřebovat definici pomalu se měnící funkce, která byla zavedena v Galambos et Seneta (1973).

Funkci L , která je konečná, měřitelná a kladná na intervalu $[A, \infty)$ nazveme pomalu se měnící funkcí, pokud pro všechna $\lambda > 0$ platí vztah

$$\lim_{x \rightarrow \infty} \frac{L(\lambda x)}{L(x)} = 1. \quad (\text{A.2})$$

Pomalu měnící funkce jsou zřejmě všechny funkce, které mají v nekonečnu konečnou kladnou limitu, ale není to nutná podmínka, např. funkce $\ln(x)$ je pomalu se měnící, i když nemá v nekonečnu limitu v oboru reálných čísel, a funkce $1/\ln(x)$ je také pomalu se měnící, i když má v nekonečnu limitu 0.

Kritérium konvergence blokových maxim k Weibullovu rozdělení zní takto: Blokovaná maxima nezávislých náhodných hodnot s distribuční funkcí F konvergují ke GEV rozdělení se záporným parametrem tvaru ξ a pravou hranicí rozdělení x_* právě tehdy, pokud platí:

$$1 - F\left(x_* - \frac{1}{x}\right) \rightarrow x^{\frac{1}{\xi}} L_f(x), \quad (\text{A.3})$$

pro $x \rightarrow \infty$, kde L_f je pomalu se měnící funkce. Pravou hranici rozdělení pak definujeme vztahem

$$x_* := \inf\{x: F(x) = 1\}, \quad (\text{A.4})$$

v případě, že $F(x) < 1$ pro všechna reálná x , pak $x_* := \infty$.

(iii) $\xi < 0$ (Gumbelovo rozdělení): Nejabstraktnější kritérium konvergence je v případě Gumbelova rozdělení: *Bloková maxima nezávislých náhodných hodnot s distribuční funkcí F konvergují ke Gumbelovu rozdělení právě tehdy, pokud pro nějakou pomocnou funkci b a pro každé $\lambda > 0$ platí vztah*

$$\frac{1-F(y+b(y)\lambda)}{1-F(y)} \rightarrow e^{-\lambda}, \quad (\text{A.5})$$

pro $x \rightarrow \infty$.

A.2 Postačující kritéria konvergence (Von Misesův teorém)

Uvedená nutná a postačující kritéria konvergence blokových maxim jsou v praxi velmi těžkopádná. Pro důkaz konvergence se proto často používá tzv. Von Misesův teorém, což je postačující podmínka konvergence blokových maxim, která bývá většinou snadněji použitelná. Nejprve si definujme tzv. funkci hazardu takto: předpokládejme, že distribuční funkce F je diferencovatelná, a její derivaci (hustotu pravděpodobnosti) označme ρ . Pak funkci hazardu definujeme vztahem

$$r(x) = \frac{\rho(x)}{1-F(x)}. \quad (\text{A.6})$$

Von Misesův teorém pro tři různé druhy GEV rozdělení (Fréchetovo, Weibullovo a Gumbelovo) pak zní takto (symbolem x_* opět značíme pravou hranici rozdělení):

Pokud $x_ = \infty$ a pro funkci hazardu $r(x)$ platí*

$$\lim_{x \rightarrow \infty} xr(x) = \alpha > 0, \quad (\text{A.7})$$

pak blokovaná maxima konvergují ke GEV rozdělení s parametrem tvaru $\xi = 1/\alpha$.

Pokud $x_ < \infty$ a platí*

$$\lim_{x \rightarrow x_*-} (x_* - x)r(x) = \alpha > 0, \quad (\text{A.8})$$

pak blokovaná maxima konvergují ke GEV rozdělení s parametrem tvaru $\xi = -1/\alpha$.

Pokud $r(x)$ je kladné a diferencovatelné v nějakém okolí bodu plus nekonečno, a platí

$$\lim_{x \rightarrow \infty} \frac{dr(x)}{dx} = 0, \quad (\text{A.9})$$

pak blokovaná maxima konvergují ke Gumbelovu rozdělení.

Příloha B: Důkaz věty o průměrné době návratu určené proložením GEV rozdělení náhodnými hodnotami (sekce 3.3.5)

Věta: *Nechť náhodná veličina x má hustotu pravděpodobnosti $\rho(x)$ takovou, že*

$$\rho(x) > 0 \quad \forall x \in (x_{inf}, x_{sup}), \quad (\text{B.1})$$

kde x_{inf}, x_{sup} jsou z R^ , tedy buď reálná čísla, nebo $\pm\infty$, a $x_{inf} < x_{sup}$ a současně je $\rho(x)$ na intervalu (x_{inf}, x_{sup}) spojitá. Dále nechť n je přirozené číslo a G_n je metoda přiřazující distribuční funkci GEV rozdělení naměřeným datům sestávající se z funkcí $R^n \rightarrow R$ μ_n, σ_n a ξ_n . Současně předpokládejme, že existuje taková n -tice reálných čísel $X_1, X_2, \dots, X_n = \mathbf{X}$, která nejsou všechna stejná taková, že $\xi_n(\mathbf{X}) < 0$ a μ_n, σ_n a ξ_n jsou v bodě \mathbf{X} spojitě, G_n nezávisí na pořadí proměnných a škáluje a posouvá se stejně jako data, tedy platí:*

$$G_n(x, \mathbf{X}) = G_n(ax + b, a\mathbf{X} + b) \quad \forall a, b \in R, a \neq 0, \quad (\text{B.2})$$

kde x je hodnota, pro kterou se distribuční funkce počítá, a X je n -tice hodnot, kterými se GEV rozdělení prokládá. Dobu opakování určíme podle vztahu (3.3.11), ve kterém zvolíme $T_0 = 1$, přičemž definujeme podle obvyklé praxe dobu opakování jako $+\infty$, pokud je hodnota distribuční funkce rovna 1. Potom průměrná doba opakování určená z n náhodných hodnot s hustotou pravděpodobnosti $\rho(x)$ metodou proložením GEV rozdělení G_n je nekonečná pro všechna $y > x_{inf}$.

Důkaz: Protože G_n nezávisí na pořadí argumentů, seřadíme si hodnoty \mathbf{X} vzestupně. Dále nechť x_{sup} je bez újmy na neurčitosti konečné. Protože pokud rovnice (B.1) platí s $x_{sup} = \infty$, tak platí i pokud je x_{sup} jakékoliv reálné číslo.

Dobu opakování určenou pro hodnotu y metodou G_n z dat \mathbf{X} si označíme jako $T(y, G_n, \mathbf{X})$. Definujme si $L(\mathbf{X})$ jako horní mez Weibullova rozdělení proloženého hodnotami \mathbf{X} , tedy

$$L(\mathbf{X}) = \mu_n(\mathbf{X}) - \frac{\sigma_n(\mathbf{X})}{\xi_n(\mathbf{X})}. \quad (\text{B.3})$$

$G_n(x, \mathbf{X}) = 1$ pro všechna $x \geq L(\mathbf{X})$ a pro tyto hodnoty x je tedy odhad doby opakování nekonečný. Definujme si \mathbf{X}_{max} jako

$$H = \max(X_n, L(\mathbf{X})), \quad (\text{B.4})$$

a K jako

$$K = \min(y, x_{sup}), \quad (\text{B.5})$$

kde y je konkrétní hodnota, pro kterou dokazujeme, že průměrná doba opakování výše popsanou metodou je nekonečná. Doba opakování je ze své definice neklesající funkce, a pokud je nekonečná pro K , musí být nekonečná i pro y .

Nyní si hodnoty X lineárně transformujeme na hodnoty Y . Položíme $Y = aX + b$ tak, aby platilo že

$$1. a > 0, \tag{B.6}$$

$$2. Y_1 > x_{inf}, \tag{B.7}$$

$$3. aH + b < K. \tag{B.8}$$

První podmínka zaručuje, že proložením $G_n(Y)$ bude stále Weibullovo rozdělení pro maxima. Druhá podmínka zaručuje, že nejmenší z hodnot Y bude stále v intervalu (x_{inf}, x_{sup}) a konečně třetí podmínka zajišťuje, že největší z hodnot Y a také horní hranice Weibullova rozdělení proloženého hodnotami Y bude také v intervalu (x_{inf}, x_{sup}) .

Pokud $x_{inf} = \infty$, splní tyto podmínky např. hodnoty

$$a = 1 \tag{B.9}$$

$$b = K - H - 1, \tag{B.10}$$

A pokud je x_{inf} konečné, splní je např. hodnoty

$$a = \frac{K - x_{inf}}{2 * (H - X_1)} \tag{B.11}$$

$$b = \frac{x_{inf} + 3K - 4aX_1}{4}. \tag{B.12}$$

Za podmínek (B.6-8) tedy platí $T(x, G_n, Y) = \infty$ pro všechna $x > K - \varepsilon_K$. Protože funkce μ_n , σ_n a ξ_n jsou v bodě Y spojitě a horní mez L Weibullova rozdělení daná vztahem (B.3) je spojitá pro hodnoty $\mu_n(Y)$, $\sigma_n(Y)$ a $\xi_n(Y)$ (protože se jedná o Weibullovo rozdělení, platí $\xi_n(Y) < 0$). Platí, že $T(K, G_n, Z) = \infty$ pro všechny Z náležející do nějakého ε_1 okolí bodu Y . Pro jednoduchost uvažujme definici okolí s normou

$$\|X\| \equiv \max(|x_1|, \dots, |x_n|). \tag{B.13}$$

Nyní zvolme ε jako

$$\varepsilon \equiv \min(\varepsilon_1, \frac{Y_1 - x_{inf}}{2}, \frac{x_{sup} - Y_n}{2}). \tag{B.14}$$

Definujme okolí O jako uzavřené ε okolí bodu Y a m_i jako

$$m_i \equiv \min(\rho(x), x \in \langle Y_i - \varepsilon, Y_i + \varepsilon \rangle) \tag{B.15}$$

a m jako minimum z m_i , $i = 1, \dots, n$. Protože platí předpoklad ze vzorce (B.1) a při volbě ε podle vzorce (B.14) náleží všechny uzavřené intervaly $\langle Y_i - \varepsilon, Y_i + \varepsilon \rangle$ do intervalu (x_{inf}, x_{sup}) , bude m existovat a bude větší než 0.

Nyní pro průměrnou dobu opakování $E(T(Z, G_n))$, kde Z je n -tice hodnot s rozdělením $\rho(x) (Z_1, \dots, Z_n)$ platí rovnice

$$\begin{aligned}
 E(T(Z, G_n)) &= n! \int_{R^n} T(Z, G_n) \rho(Z_1) \dots \rho(Z_n) dZ_1 \dots dZ_n = \\
 &n! \int_{Z \in O} T(Z, G_n) \rho(Z_1) \dots \rho(Z_n) dZ_1 \dots dZ_n + \\
 &+ n! \int_{Z \notin O} T(Z, G_n) \rho(Z_1) \dots \rho(Z_n) dZ_1 \dots dZ_n \geq \\
 &n! \infty 2^n \varepsilon^n m = \infty,
 \end{aligned} \tag{B.16}$$

kde jsme využili faktu, že druhý integrál v prostředním řádku musí být nezáporný, protože doba opakování je z definice vždy nezáporná. Člen $n!$ ve vzorci je proto, aby byly zahrnuty všechny možné permutace n hodnot.

Příloha C: Vliv hodnot parametrů Gumbelova a GEV rozdělení na odhad hodnoty se zadanou dobou opakování

Tato příloha vychází z analýzy aproximace extrémních dat GEV a Gumbelovým rozdělením metodou L-momentů a ML na uměle vygenerovaných datech (podkapitola 3.8).

C.1 Gumbelovo rozdělení

Podle tabulky 3.8.6 je pro odhad rychlosti větru s dobou opakování 50 let u_{50} pomocí aproximace rozdělení dat Gumbelovým rozdělením zásadní odhad parametru σ , protože mezi odhady u_{50} a σ je vyšší korelace než mezi odhady u_{50} a μ . V této kapitole je tento poznatek zobecněn i pro jiné doby opakování (dobu opakování zde chápeme jako dobu opakování ročních maxim).

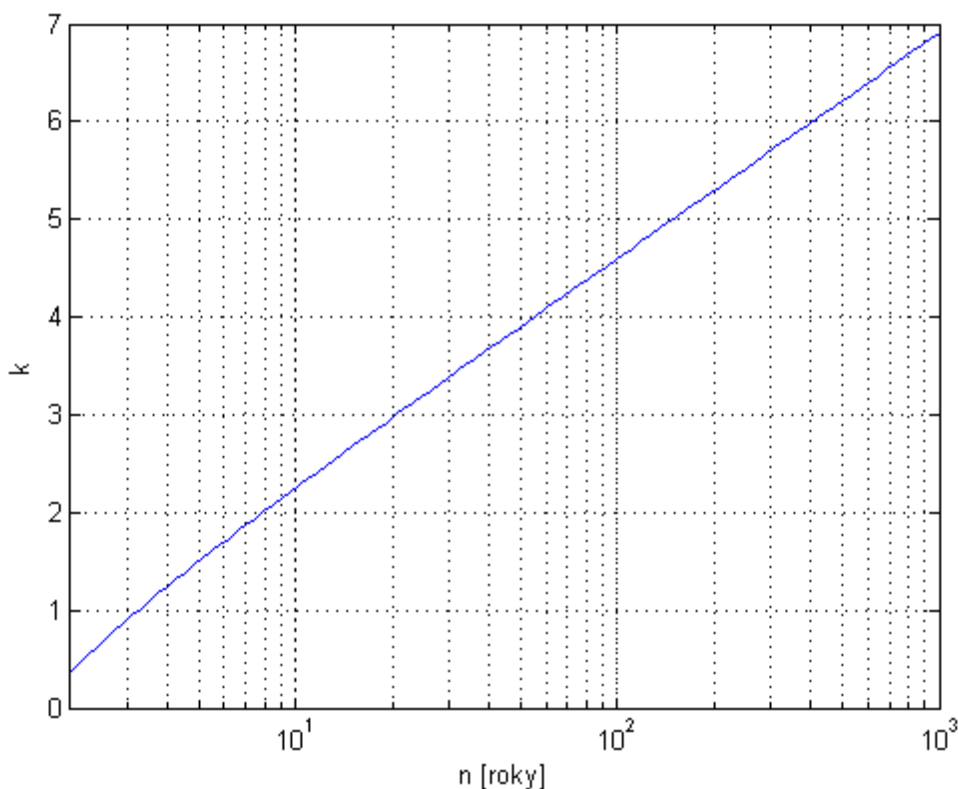
V následujícím textu použijeme hodnoty, které byly získány aproximací dat generovaných Gumbelovým rozdělením pomocí Gumbelova rozdělení metodou L-momentů. Velmi podobné výsledky byly získány i pro data generovaná ostatními způsoby a aplikací ML (viz podkapitola 3.8).

Podle tabulek 3.8.2 a 3.8.3 je ve vybraném případě směrodatná odchylka odhadu parametru μ přibližně 1,28 krát větší než směrodatná odchylka odhadu parametru σ .

Rychlost s dobou návratu n let, u_n , se dá v případě Gumbelova rozdělení vypočítat přímo z jeho parametrů. Dosazením distribuční funkce Gumbelova rozdělení (3.2.3) do vzorce pro výpočet doby návratu 3.1.1 dostaneme

$$u_n = \mu - \ln\left(-\ln\left(1 - \frac{1}{n}\right)\right) \sigma = \mu + k\sigma, \quad (\text{C.1})$$

což znamená, že chyba v určení σ má k krát větší vliv než chyba určení μ . Pro $n = 50$ je k přibližně rovno 3,902, což je podstatně více než podíl směrodatných odchylek parametrů μ a σ . Graf závislosti koeficientu k na době návratu n je na obr. C.1. Vliv odhadu σ na určení u_n roste s rostoucím n .



Obr. C.1: Graf koeficientu k ve vzorci (C.1) v závislosti na době návratu n . Stupnice k je logaritmická.

C.2 GEV rozdělení

V případě GEV rozdělení neexistuje vzorec analogický (C.1). Podle tabulky 3.8.7 je odhad u_{50} pro data generovaná Gumbelovým rozdělením pomocí GEV rozdělení nejvíce korelován s odhadem ξ .

Empirická zkušenost ukazuje, že se vliv odhadu parametru ξ na odhad u_n zvyšuje s dobou opakování. To potvrzuje tabulka C.1, ve které jsou korelace odhadů u_n s odhady parametrů GEV rozdělení pro metody L-momentů i ML pro $n = 10, 50$ a 250 . Zatímco pro $n = 10$ je korelace odhadu u_n nejvyšší s odhadem σ , pro delší doby návratu je nejvyšší s parametrem ξ . Mezi oběma metodami aproximace nejsou znatelné rozdíly.

	u_{10} lm	u_{10} ml	u_{50} lm	u_{50} ml	u_{250} lm	u_{250} ml
μ	0,473	0,439	0,058	0,008	-0,117	-0,156
σ	0,679	0,647	0,306	0,259	0,078	0,042
ξ	0,319	0,384	0,744	0,785	0,865	0,875

Tabulka C.1: Korelace odhadů rychlosti větru s dobou opakování 10, 50 a 250 let s odhady parametrů GEV rozdělení metodami L-momentů (lm) a maximální věrohodnosti (ml).

Příloha D: Odhad parametrů tvaru GEV rozdělení Bayesovským přístupem

Podle kapitoly 3.8 a přílohy C je odhad hodnot s vysokou dobou opakování v případě aproximace extrémních hodnot GEV rozdělením silně závislý na určení parametru tvaru GEV rozdělení ξ . Pro určení hodnoty tohoto parametru se obvykle používají dvě metody:

1. Použije se Gumbelovo rozdělení, tj. položí napevno $\xi \equiv 0$.
2. ξ se určí proložením obecného GEV rozdělení naměřenými daty.

V podkapitole 3.8 byla metoda 1 zobecněna tak, že pro parametr ξ je možné použít libovolnou pevně zadanou hodnotu, dále tedy bude za metodu 1 považováno toto zobecnění. V případě metody 1 se parametr ξ určí výhradně pomocí apriorního předpokladu, v případě metody 2 výhradně pomocí naměřených hodnot.

Výsledky simulací v podkapitole 3.8 jsou motivací pro použití Bayesovského přístupu (Hoff, 2009). Vycházíme ze dvou předpokladů:

1. je k dispozici apriorní znalost parametru ξ GEV rozdělení daná normálním rozdělením s parametry ξ_h a σ_h ,

2. chyba určení hodnoty parametru ξ GEV rozdělení má normální rozdělení s parametry θ a σ_s .

Potom z Bayesovské analýzy pro hustotu pravděpodobnosti ξ plyne

$$\begin{aligned} q(\xi) &\approx e^{-\frac{(\xi-\xi_h)^2}{2\sigma_h^2}} * e^{-\frac{(\xi-\xi_s)^2}{2\sigma_s^2}} = e^{-\frac{(\xi-\xi_h)^2}{2\sigma_h^2} + \frac{(\xi-\xi_s)^2}{2\sigma_s^2}} = \\ &= e^{-\frac{(\sigma_h^2 + \sigma_s^2) * (\xi^2 - \frac{2\xi(\sigma_h^2 \xi_s + \sigma_s^2 \xi_h)}{\sigma_h^2 + \sigma_s^2}) + k(\xi_h, \sigma_h, \xi_s, \sigma_s)}{2 * \sigma_h^2 * \sigma_s^2}}, \end{aligned} \quad (D.1)$$

kde \approx značí přímou úměrnost a μ_s je odhad parametru ξ z naměřených dat a k značí konstantu nezávislejší na ξ . Hodnoty σ_h a σ_s budeme nazývat apriorní a výběrovou směrodatnou odchylkou. Porovnáním výrazu (D.1) se vzorcem pro hustotu pravděpodobnosti normálního rozdělení dostaneme, že hustota pravděpodobnosti parametru ξ je dána normálním rozdělením s parametry

$$\sigma^2 = \frac{\sigma_h^2 \sigma_s^2}{\sigma_h^2 + \sigma_s^2}, \quad (D.2)$$

a

$$\mu = \frac{\sigma_s^2 \xi_h + \sigma_h^2 \xi_s}{\sigma_h^2 + \sigma_s^2}. \quad (D.3)$$

Odhad parametru ξ je pak roven průměru tohoto rozdělení, tj. μ ze vzorce (D.3). Jedná se tedy o lineární kombinaci odhadů z metod 1 a 2. Samotné metody 1 a 2 jsou speciálními případy. Metodu 1 dostaneme volbou hodnot $\sigma_h = 0$ a metodu 2 limitou $\sigma_h \rightarrow \infty$.

Navržená metoda výpočtu ξ vyžaduje znalost apriorní a výběrové směrodatné odchylky (STD). Apriorní STD je nutné určit na základě specifických úvah k dané konkrétní aplikaci, např. na základě regionálního přístupu. Výběrovou STD můžeme odhadnout na základě numerických simulací provedených v podkapitole 3.8 s GEV rozdělením s parametry $\xi = 0, -0,1, 0,1$. Odhadnuté výběrové STD jsou v tabulce D.1.

Předpoklad o normálním rozdělení odhadů parametru ξ GEV rozdělení nelze exaktně potvrdit, avšak z histogramů na obr. 3.8.14 (histogramy jsou pro hodnoty $\xi = 0$, tvary histogramů pro hodnoty $\xi = -0,1$ a $0,1$ jsou velmi podobné) vyplývá, že tento předpoklad je akceptovatelný.

Přibližná chyba odhadu parametru ξ GEV rozdělení z výběru 50 hodnot je podle tabulky D.1 rovna 0,115. Protože chyby výběru jsou obecně nepřímo úměrné odmocnině

počtu hodnot (Wilks, 2011), lze výběrovou chybu $\sigma_s(N)$ z N hodnot odhadu parametru ξ vyjádřit přibližným vztahem

$$\sigma_s(N) \cong \frac{0.115}{\sqrt{N/50}} \quad (\text{D.4})$$

Z tohoto vztahu plyne, že v pro roční maxima se σ_s snižuje s rostoucí délkou měřené řady. Proto podle vztahu (D.3) se bude při konstantní hodnotě σ_h s rostoucí délkou měření zvětšovat vliv odhadu ξ z daného výběru na úkor jeho apriorního odhadu. Vztahy (D.3) a (D.4) určují konkrétní tvar této závislosti.

ξ	-0,1	0	0,1
Metoda			
L-momenty	0,106	0,110	0,117
ML	0,113	0,118	0,125

Tabulka D.1: Odhady výběrových směrodatných odchylek parametru ξ GEV rozdělení.

Postup odhadu ξ se skládá z následujících kroků:

1. Zvolíme parametry ξ_h a σ_h normálního rozdělení parametru ξ GEV rozdělení.
2. Naměřené hodnoty aproximujeme obecným GEV rozdělením a tím dostaneme střední hodnotu odhadu μ_s výběrového parametru ξ . Směrodatnou odchylku odhadu ξ zjistíme podle vztahu (D.4).
3. Získané hodnoty dosadíme do vzorce (D.3) a tím získáme odhad parametru ξ .
4. Daty proložíme GEV rozdělení s parametrem ξ získaným v kroku 3.