



UNIVERZITA KOMENSKÉHO, BRATISLAVA
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

VIZUALIZÁCIA VERIFIKÁCIE PREDPOVEDNÝCH MODELOV POČASIA

Diplomová práca

Bratislava, 2015

Bc. Marek Kružliak



UNIVERZITA KOMENSKÉHO, BRATISLAVA
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

VIZUALIZÁCIA VERIFIKÁCIE PREDPOVEDNÝCH MODELOV POČASIA

Diplomová práca

Študijný program:	Aplikovaná informatika
Študijný odbor:	2511 Aplikovaná informatika
Školiace pracovisko:	Katedra aplikovanej informatiky
Školiteľ:	RNDr. Andrej Lúčny, PhD.

Bratislava, 2015

Bc. Marek Kružliak

Tu bude zadanie

Čestne prehlasujem, že som túto diplomovú prácu vypracoval samostatne s použitím citovaných zdrojov.

.....

TODO - tu bude podakovanie

Abstrakt

TODO

Kľúčové slová: vizualizácia informácií, verifikácia predpovedí počasia

Abstract

TODO.

Keywords: information visualization, verification of weather forecasts

Obsah

1	Úvod	1
2	Verifikácia predpovedných modelov počasia	2
2.1	Predpovedný model počasia	2
2.1.1	WRF model	4
2.2	Dáta	4
2.2.1	Predpovedané dáta	5
2.2.2	Pozorované dáta	6
2.2.3	Párovanie dát	6
2.3	Meranie chyby predpovede	8
2.3.1	Stredná chyba predpovede	8
2.3.2	Stredná absolútna chyba	9
2.3.3	Stredná kvadratická chyba	9
2.3.4	Všeobecná kumulovaná chyba	9
2.3.5	Medián absolútnych chýb	11
3	Predchádzajúce riešenia	12
3.1	Verifikačný softvér	12
3.1.1	Štatistický softvér	12
3.1.2	Špecializovaný softvér	14
3.2	Techniky vizualizácie vo verifikácii	14
3.2.1	Bodový graf	14
3.2.2	Histogram	15
3.2.3	Krabicový diagram	15

3.2.4	Time series plot	20
4	Návrh riešenia	21
4.1	Návrh systému	21
4.2	Návrh vizualizácie	21
4.2.1	Charakteristika dát	21
4.2.2	Špecifikácia požiadaviek na vizualizáciu	21
4.2.3	Návrh farebnej palety	21
4.2.4	Návrh vizualizácie distribúcie chýb	21
5	Implementácia	23
6	Výsledky	24

Zoznam obrázkov

2.1	Flowchart systému predpovedného model počasia od edukačného programu The COMET [LE11]. Na obrázku je zvýraznená časť, ktorej sa venujeme v tejto práci.	3
2.2	Vizuálne znázornenie dvoch bežne používaných metód na získavanie hodnôt z mriežky	7
3.1	Pôvodný návrh krabicového diagramu, ako bol prezentovaný v práci <i>Exploratory Data Analysis</i> (1977)[Tukey]	17
3.2	a) Klasický krabicový diagram b-f) Vizuálne variácie krabicového diagramu b-c) 2 variácie pre kvartilový graf [Tuf83] c) Skrátený krabicový diagram [RKR06] e) Range-bar chart [Spe] f) Farebná variácia [Carr]	18
3.3	Krubicové diagramy s pridanou informáciou a) Krabicový diagram s variabilnou šírkou b) Vrúbkovaný krabicový diagram c) Krabicový diagram s informáciou o šikmosti dát	19

Zoznam tabuliek

Zoznam skratiek

WRF	Weather Research and Forecasting
NCEP	National Centers for Environmental Prediction
NCAR	National Center for Atmospheric Research

Kapitola 1

Úvod

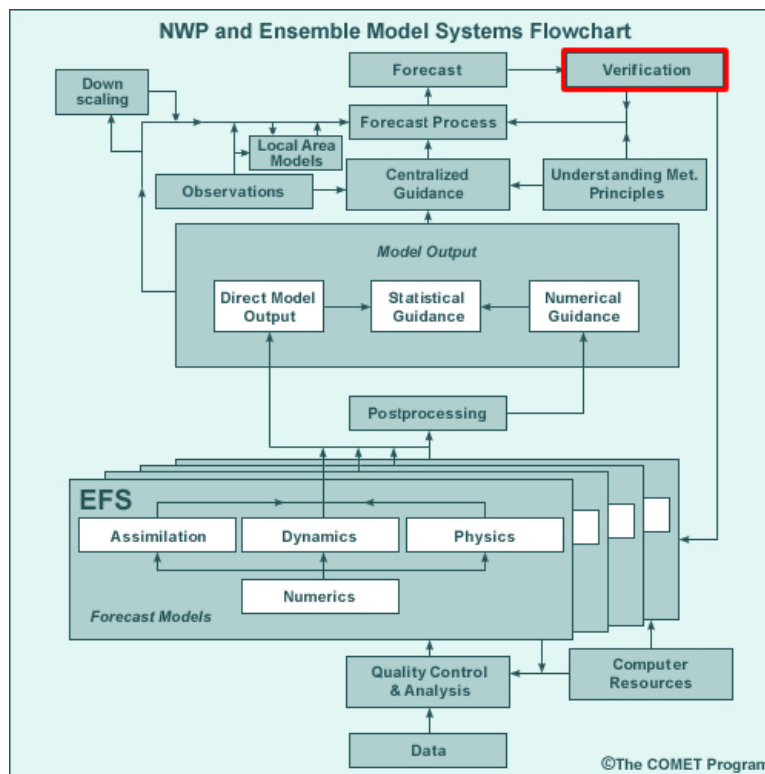
Kapitola 2

Verifikácia predpovedných modelov počasia

Verifikácia je proces, ktorý má overiť správnosť fungovania predpovedného modelu počasia. Z tohto dôvodu je nepostrádateľnou súčasťou meteorologického výskumu a taktiež celkového procesu predpovedania počasia. [CWS⁺08] Ciele verifikácie môžeme rozdeliť do troch skupín: *administratívne*, *vedecké* a *ekonomické*. Medzi *administratívne ciele* patrí monitorovanie úspešnosti predpovedania modelu a nasmerovanie užívateľov na jeho správnu konfiguráciu alebo voľbu iného modelu. *Vedekými cieľmi* sú identifikovanie a oprava slabín modelu a taktiež vylepšovanie predpovedí. *Ekonomickými cieľmi* sú rozhodovanie, kam majú smerovať investície do výskumu a iné závažné ekonomické rozhodnutia. [FJB12]

2.1 Predpovedný model počasia

Už v 19. storočí vývoj termodynamiky na základe Newtonovskej fyziky vyvrcholil v ucelení množiny fundamentálnych princípov, ktoré riadia prúdenie plynov v atmosfére. Začiatkom 20. storočia sa o matematický prístup k predpovedaniu počasia najviac zaslúžili osobnosti ako Vilhelm Bjerknes alebo Lewis F. Richardson. Avšak na ďalší úspech, v tejto oblasti, sa muselo čakať až na vynájdenie prvých počítačov počas 2. svetovej vojny ako bol IAS alebo ENIAC. [Lyn07] Prvá úspešná predpoveď bola vykonaná v 50. rokoch minulého storočia a to hlavne vďaka práci Jula Charneyho. Následný vývoj vo výpočtovej sile počítačov,



Obr. 2.1: Flowchart systému predpovedného model počasia od edukačného programu The COMET [LE11]. Na obrázku je zvýraznená časť, ktorej sa venujeme v tejto práci.

používanie satelitných pozorovaní a vývoj samotnej meteorológie ako vedy zapríčinil, že je numerická predpoveď počasia (NWP) dnes najúspešnejším prístupom ako predpovedať počasie. [Gol]

Odvtedy vzniklo veľké množstvo modelov, ako sú napríklad GFS, NAM, RUC, WRF, SREF, GEFS, ECMWF, ALADIN a mnoho ďalších. Naša práca sa zameriava konkrétne na verifikáciu modelu *WRF*. Taktiež pokračuje neustály vývoj aj vďaka novým modelovacím technikám, novým parametrizáciám, a zvyšovaniu výkonu výpočtových zdrojov.

Ako môžeme vidieť na obrázku 2.1 proces predpovedania počasia má okrem numerického modelu, ktorý je jej jadrom, aj iné časti. Ako príklad môžeme spomenúť získavanie vstupných dát, ich spracovanie, postprocessing, spracovanie výstupu a následne poskladanie samotnej predpovede. Cieľom nášho záujmu, celého procesu predpovedania, je *verifikácia*. Ako môžeme vidieť z obrázka, verifikácia vplýva na vyladenie parametrov modelu, avšak tento proces sa nedeje automaticky, ale vyžaduje prácu meteorológov a ich

chápanie základných meteorologických princípov.

2.1.1 WRF model

Ako sme už spomenuli *The Weather Research and Forecasting* (WRF) model je *numerická predpoveď počasia* (NWP) a systém atmosferickej simulácie.

WRF je podporovaný, ako bežný nástroj pre univerzity, výskum a operačné komunity, pričom sa usiluje o splnenie požiadaviek ich všetkých súčasne. Vývoj WRF modelu bol snahou mnohých spoločností ako napríklad *The National Center for Atmospheric Research's* (NCAR), *Mesoscale and Microscale Meteorology* (MMM), *The National Oceanic and Atmospheric Administration's* (NOAA) *National Centers for Environmental Prediction* (NCEP) a *Earth System Research Laboratory* (ESRL), oddelenie ministerstva obrany *Air Force Weather Agency* (AFWA) a *Naval Research Laboratory* (NRL), *The Center for Analysis and Prediction of Storms* (CAPS) [WCSDG⁺08].

WRF model je vhodný pre širokú škálu aplikácií od *metódy vzdušných vírov* (Large Eddy Simulation - LES) až po globálne simulácie počasia. Takéto aplikácie vyžadujú numerické predpovede v reálnom čase, vývoj a štúdium asimilácie dát, výskum parametrizovanej fyziky, výskum parametrizovanej fyziky, modelovanie kvality ovzdušia, idealizované simulácie, čo všetko WRF model spĺňa.

V roku 2008 evidovala WRF viac ako 6000 užívateľov, no dnes (2014) eviduje viac ako 25000 užívateľov vo viac ako 130 krajinách sveta. Tieto fakty poukazujú na to, že WRF model má nie len veľkú základňu užívateľov, ale aj vývojárov a má v budúcnosti istotne svoje miesto a preto si myslíme, že sa oplatí investovať čas a úsilie do verifikácie tohto modelu.

2.2 Dáta

Na správne zhodnotenie úspešnosti modelu potrebujeme dva druhy dát. V prvom rade sa jedná o dáta, ktoré sú výstupom z daného predpovedného modelu počasia, teda **predpovedané dáta**. Tieto umelo získané dáta chceme konfrontovať s realitou, aby sme si mohli vytvoriť obraz o správnom fungovaní celého modelu. Realitu v našom prípade predsta-

vujú dáta namerané špecializovanými meteorologickými senzormi, ktoré označujeme ako **pozorované dáta** alebo skrátene *pozorovania*.

2.2.1 Predpovedané dáta

Predpovedané dáta z modelu WRF sa ukladajú vo formáte **GRIB**, čo je skratka pre *GRIdded Binary* [WMO94] alebo na iných miestach uvádzané ako *General Regularly-distributed Information in Binary form* [WMO03]. Tento formát je štandardom Svetovej meteorologickej organizácie teda *World Meteorological Organization* (WMO). Jedná sa o pomerne rozšírený formát, používaný pri veľkom množstve meteorologických aplikácií a je taktiež používaný ako výstupný formát pre iné predpovedné modely ako WRF, či už ECMWF, GFS, NAM, SREF alebo mnohé iné [NCE14].

Doteraz boli vyvinuté 3 verzie tohoto formátu od 0 po 2. Verzia 0 bola určená pre malé projekty typu TOGA a to iba s limitovaným použitím a dnes sa táto verzia už vôbec nepoužíva. Verzia grib 1 [WMO94], grib 2 [WMO03] sú dnes bežne používané väčšinou meteorologických centier.

Medzi verziami 1 a 2 nie sú žiadne rozdiely v obsahovej filozofii, preto popis obsahu gribovského formátu, ktorý tu uvádzame je spoločný pre obe tieto verzie. *Gribovský súbor* (ďalej iba *Grib*) pozostáva z viacerých *Gribovských záznamov*, pričom jeden záznam môže existovať ako samostatný Grib. Vďaka tomu je možné ľahko spájať Griby, a to tiež v ľubovoľnom poradí, bez toho, aby sme ich nejako poškodili. Samozrejme musí byť zachovaná homogenita, čo sa týka verzií Gribov, teda verziu 1 nemožno miešať s verziou 2 a naopak. Už samotný názov *Gridded Binary* nám napovedá, že dáta sú usporiadané v pravidelnej mriežke. Každý Gribovský záznam obsahuje dvojrozmernú mriežku (zemepisná šírka x zemepisná dĺžka) hodnôt v určitom čase a vertikálnej hladine. Taktiež v hlavičke záznamu sa nachádzajú metainformácie, ktoré nám hovoria o aké dáta ide, teda o akú premennú sa jedná, čas predpovede, výškovú hladinu a podobne. Grib je zvyčajne z tohto dôvodu 2 až 5 rozmerná dátová štruktúra s veľkým množstvom veličín ako je napríklad teplota, tlak, relatívna vlhkosť, rosný bod, *u* a *v* súradnice vetra a ďalšie, ktoré sú definované v rôznych hladinách. Taktiež je dôležité povedať, že Grib zriedkakedy zachytáva povrch celej planéty, ale iba vymedzenú skúmanú oblasť - *doménu*.

2.2.2 Pozorované dáta

Pozorovania sa získavajú meraním priamo v teréne pomocou špecializovaných meracích zariadení, ktoré sú súčasťou meteo staníc. Každá stanica môže obsahovať iné vybavenie, ku príkladu teplomer, zrážkomer, barometer, vetromer a im podobné [Vas98], ktorými môžeme zachytávať informácie o rôznych skúmaných veličinách.

Majoritná časť meraní sa deje pri povrchu zeme priamo na meteo staniciach a nazývajú sa *surface* merania. Tieto merania najlepšie popisujú dianie v oblasti najväčšieho záujmu (biosfére), avšak neobsahujú informáciu o dianí v iných výškových hladinách. Pozorovania týchto hladín sa dejú pomocou *radiosondy*, ktorá je pripojená k meteo balónu alebo vypustená z lietadla smerom k zemi. Takéto pozorovania sa nazývajú *upper air* merania.

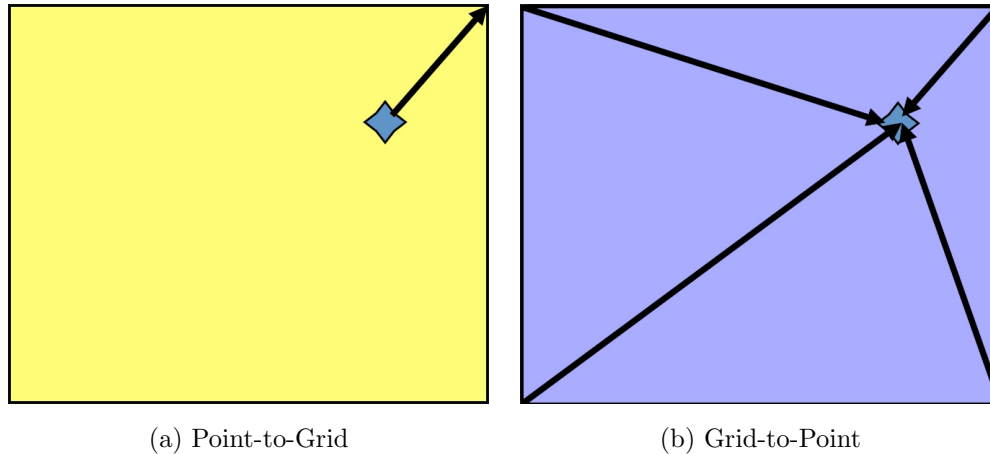
Narozdiel od predpovedaných dát, pozorované dáta nemajú štandardizovaný formát a zvyčajne sa ukladajú do databázy. Aby sme zhrnuli charakteristiku týchto dát, jedná sa o niekoľko meraných veličín, nameraných v konštantných časových krokoch - napríklad každú minútu alebo každú hodinu - v jednom konkrétnom geografickom bode a zvyčajne pri povrchu zeme, teda ak sa nejedná o *upper air* merania, ktoré sa uskutočňujú v štandardných výškových hladinách, ktoré sa merajú v hPa.

2.2.3 Párovanie dát

Z predpovedného modelu a rovnako aj z merania získame veľké množstvo hodnôt. Aby sme mohli korektne porovnať predpovede s pozorovaniami, je nevyhnutné nájsť správne párovanie týchto hodnôt, teda zistiť, ktorú hodnotu porovnať s ktorou, aby sme získali zmysluplný výsledok.

Vždy sa snažíme nájsť správnu predpoveď pre pozorovanie a nie naopak. Dôvodom je, že chceme skúmať vzťah predpovede s realitou a preto v párovaní musí byť zahrnutých čo najviac **meraných** hodnôt, ak nie všetky.

Každá pozorovaná hodnota, ktorú chceme spárovať má štyri kľúče podľa ktorých hľadáme pár: *meraná veličina* (napríklad teplota), *čas merania*, *výšková hladina* a *geografická poloha*. Nájsť všetky hodnoty podľa kľúča meranej veličiny v Grike je ľahké, keďže sa jedná o kategorickú premennú, teda môže nadobúdať iba určitý konečný počet hodnôt.



Obr. 2.2: Vizuálne znázornenie dvoch bežne používaných metód na získavanie hodnôt z mriežky

Toto sa však nedá povedať o čase, hladine a polohe, ktoré sú spojitými premennými.

Pre čas pozorovania, čas predpovede a výškovú hladinu existujú štandardy, ktoré určujú v akých časoch resp hladinách sa robia merania a predpovede, čo nám uľahčuje prácu. Ak sa napriek tomu čas alebo hladina v Grike nevyskytuje, tak pár vyhadzujeme z párovania.

V prípade polohy zo samozrejmych dôvodov neexistuje žiaden štandard a hustota mriežky v Grike nemôže byť nikdy tak veľká, aby poloha našej stanice vždy dopadla na presný bod mriežky. Z tohoto dôvodu získavame hodnoty z mriežky z okolitých bodov a to dvoma metódami *Point-to-Grid* a *Grid-to-Point* [FJB12], ktoré sú znázornené na obrázku 2.2. Jedná sa vlastne o dve interpolačné metódy. Point-to-Grid predstavuje metódu *najbližší sused* (*Nearest Neighbour*) a Grid-to-Point *bilineárnu interpolačnú metódu*.

Výber správnej metódy môže značne ovplyvniť výsledok. Dôvodom je, že môžu byť veľké rozdiely hodnôt v okolitých mrežových bodoch a tak, ak pomocou Point-To-Grid metódy získame nízku hodnotu, tak pomocou Grid-To-Point môžeme získať hodnotu omnoho väčšiu, vplyvom zvyšných troch bodov, ktoré vstúpili do interpolácie. Nemožno však jednoznačne povedať, ktorá z metód je lepšia, keďže obe môžu v istých prípadoch dávať lepšie výsledky.

2.3 Meranie chyby predpovede

Výsledkom procesu párovania je n párov (predpoveď, pozorovanie), ktoré je možné porovnať. Z porovnania týchto dvojíc získame numerickú hodnotu, ktorá nám hovorí o veľkosti chyby predpovede daného modelu pre vybrané predpovedané časy.

Chybu predpovede e_i pre i -tu dvojicu (y_i, \hat{y}_i) definujeme takto:

$$e_i = (y_i - \hat{y}_i)$$

Kde y_i je predpoveď a \hat{y}_i je pozorovanie. Takýmto spôsobom z n párov získame n chýb, ktoré agregujeme pomocou rôznych štatistických metód, ktoré sú bežne používané pri verifikácií predpovedí, ako sa spomína v [Nur03], [FJB12] a [Cas09]. Výsledkom agregácie je numerická hodnota, ktorá sa nazýva *skóre* predpovede.

2.3.1 Stredná chyba predpovede

Budeme ju označovať ako *MFE* z anglického *Mean Forecast Error*, ale v literatúre je možné ju nájsť ako *ME* [Nur03], teda *stredná chyba* alebo ako *Linear Bias* [Cas09], [FJB12]. Vzorec pre výpočet MFE vyzerá nasledovne:

$$MFE = \frac{1}{n} \sum_{i=0}^n e_i$$

MFE je možné vypočítať aj ako rozdiel priemerov predpovedí a pozorovaní.

$$MFE = \bar{y} - \bar{\hat{y}}$$

MFE vyjadruje priemerný smer chyby. To znamená, že pozitívny výsledok indikuje *over-forecast*, teda nadhodnotenú predpoveď a negatívny výsledok *under-forecast*, teda podhodnotenú predpoveď. Avšak MFE **nevyjadruje veľkosť** chyby v tomto smere, keďže kladné a záporné chyby sa navzájom môžu zrušiť. Napríklad máme množinu chýb $E = \{2, -5\}$, tak MFE pre E je -1.5, ale priemerná veľkosť chyby je 3.5.

2.3.2 Stredná absolútna chyba

Budeme ju označovať ako *MAE* z anglického *Mean Absolute Error*. Vzorec pre výpočet MAE vyzerá nasledovne:

$$MAE = \frac{1}{n} \sum_{i=0}^n |e_i|$$

Narozdiel od MFE **neurčuje smer chyby**, ale vyjadruje veľkosť chyby. Z týchto dôvodov je v praxi odporúčané zobrazovať MFE a MAE súčasne [Nur03].

2.3.3 Stredná kvadratická chyba

Budeme ju označovať ako *RMSE* z anglického *Root Mean Square Error*. Vzorec pre výpočet RMSE vyzerá nasledovne:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^n e_i^2}$$

Z povahy vzorca pre RMSE je jasné, že rovnako ako MAE, ani RMSE neurčuje smer chyby, pretože nadobúda vždy iba kladné hodnoty. Ďalšou vlastnosťou RMSE je, že nadobúda hodnoty vždy väčšie alebo rovné ako MAE, pričom výsledok RMSE je citlivý na veľké hodnoty chýb.

V praxi sa zvykne používať aj *MSE* (*Mean Square Error*):

$$MSE = \frac{1}{n} \sum_{i=0}^n e_i^2$$

Má podobné vlastnosti ako RMSE s jediným rozdielom, že RMSE meria veľkosť chyby zachovávajúc jednotky danej veličiny (napr. °C), zatiaľ čo MSE jednotky nezachováva [Nur03]. Preto sme si pre náš účel zvolili RMSE, ktoré je jednoduchšie zobrazíť spolu s MFE a MAE v jednom grafe, keďže sa zachováva konzistentnosť jednotiek veličín.

2.3.4 Všeobecná kumulovaná chyba

V našom systéme sme navrhli všeobecný vzorec na výpočet kumulovaného skóre, ktorým možno vyjadriť ľubovoľnú zo spomenutých štatistických metód. Takéto vyjadrenie umožňuje

nie len všeobecnosť, ale aj jednoduché rozšírenie systému o ďalšie metódy a to nie len programátorom, ale aj samotným užívateľom systému.

Všeobecný vzorec na výpočet *skóre* pre danú predpoveď vyzerá takto:

$$Score = \Phi\left(\sum_{i=0}^n \varepsilon(e_i)\right)$$

Kde Φ je ľubovoľná funkcia z \mathbb{R} do \mathbb{R} , teda $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ a podobne funkcia $\varepsilon : \mathbb{R} \rightarrow \mathbb{R}$. Spomenuté metódy môžeme teda skonštruovať zadaním správneho Φ a ε .

Napríklad pre *MFE*:

$$\begin{aligned}\Phi(x) &= \frac{x}{n} \\ \varepsilon(e) &= e\end{aligned}$$

Pre *MAE*:

$$\begin{aligned}\Phi(x) &= \frac{x}{n} \\ \varepsilon(e) &= |e|\end{aligned}$$

Pre *RMSE*:

$$\begin{aligned}\Phi(x) &= \sqrt{\frac{x}{n}} \\ \varepsilon(e) &= e^2\end{aligned}$$

Pre *MSE*:

$$\begin{aligned}\Phi(x) &= \frac{x}{n} \\ \varepsilon(e) &= e^2\end{aligned}$$

Ako sme spomenuli, je možné rozšírenie o ďalšie metódy a to napríklad o Brownov a Triggov *signál chybných predikcií*, ktorý budeme označovať ako *TS* z anglického *Tracking Signal*. Tieto metódy sme vyššie nespomenuli, keďže sa v meteorologickej praxi nepoužívajú. Uvádzame ich však ako možné rozšírenie, keďže sú tieto metódy bežne používané pri verifikácii iných predpovedných modeloch, ako sú tie meteorologické.

2.3.5 Medián absolútnych chýb

Budeme ju označovať ako MAD z anglického *Median Absolute Deviation*. Vzorec pre výpočet MAD vyzerá nasledovne:

$$MAD = median(|e|) = |\tilde{e}|$$

Nech je daná usporiadaná postupnosť Y_1, \dots, Y_N , tak potom *median* náhodnej premennej x je definovaný rovnako ako v [Wei14]:

$$median(x) = \tilde{x} \equiv \begin{cases} Y_{(N+1)/2} & \text{ak } N \bmod 2 = 0 \\ \frac{1}{2}(Y_{(N+1)/2} + Y_{(N+1)/2+1}) & \text{ak } N \bmod 2 = 1 \end{cases}$$

Z daného vzorca môžeme vidieť podobné vlastnosti ako má MAE, avšak MAD je robustnejší a extrémne chyby nemajú na skóre žiaden efekt.

Kapitola 3

Predchádzajúce riešenia

3.1 Verifikačný softvér

Verifikácia predpovedných modelov počasia je úloha dokonale stvorená pre automatizáciu. Z tohto dôvodu meteorológovia začali využívať dostupný štatistický softvér a neskôr boli taktiež vyvíjané špecializované nástroje určené pre verifikáciu. Môžeme teda rozdeliť verifikačný softvér do dvoch základných kategórií a to *štatistický* a *špecializovaný*, ktorý je zväčša podporovaný rôznymi národnými a medzinárodnými organizáciami.

3.1.1 Štatistický softvér

Spoločnou črtou:

- obmedzená funkcionálna
- obmedzená vizualizácia
- slabé / žiadne GUI
- vyžaduje znalosť špecifického programovacieho jazyka
- ...

Tabuľkový softvér

Napriek tomu, že je tabuľkový softvér na výpočet štatistík zamietnutý komunitou vedcov a štatistikov ako nevhodný a neprofesionálny, tak je využívaný, a to pomerne často, aj vo vedeckých kruhoch. Výhodou je, že novému užívateľovi umožňuje okamžite vidieť všetky

kroky v základných procedúrach verifikácie a teda je výborný pre výučbové účely. [Poc11] Najznámejší kus softvéru z pomedzi komerčných produktov je *Microsoft Excel*[ref] a z voľne dostupných je jeho opensource náprotivok *Open Office Calculate*[ref]. Oba programy zahrňujú základné štatistické funkcie ako napríklad stredná kvadratická chyba (MSE) pre spojité predpovede (pozri odsek 2.3.3) a taktiež umožňujú generovanie jednoduchých grafov na základe tabuľkových dát. Tabuľkový softvér neposkytuje priamo funkcionalitu na výpočet ďalších sofistikovanejších verifikačných štatistík, avšak umožňuje ich implementáciu pomocou makro programovania v špecifickom jazyku. Pre Microsoft Excel je to *Microsoft Visual Basic*[ref] a pre Open Office Calculate zasa *OpenOffice.org Basic*[ref]. Oba jazyky patria do rodiny *Basic* jazykov, takže majú mnoho podobných prvkov.

MATLAB

MATLAB je interaktívne prostredie s vlastným programovacím jazykom, ktorý je využívaný miliónmi inžinierov a vedcov po celom svete [ref] a tým nevynímajúc meteorológov a ďalších odborníkov pracujúcich v atmosférickom výskume. Zvyčajne sa *MATLAB* využíva na výskum a prototypovanie nových metód a procedúr [Poc11], pretože umožňuje rýchlu a jednoduchú implementáciu, keďže jeho súčasťou je mnoho matematických knižníc a je prispôsobený na prácu s maticami dát. Výhodou *MATLABU* je, že umožňuje tvorbu GUI a taktiež poskytuje kreslenie rôznorodých grafov a diagramov. Mali by sme však podotknúť, že podobne ako väčšina štatistického softvéru, aj *MATLAB* je komerčný produkt. Jeho cena za jednu licenciu je \$2,650 (k roku 2015), čo je pomerne vysoká suma, ak vezmeme do úvahy za akým účelom chceme tento softvér využívať a ako dobre je naň prispôsobený.

Minitab

R

Statistical Analysis Software (SAS)

Interactive Data Language (IDL)

3.1.2 Špecializovaný softvér

NCAR Command Language (NCL)

Model Evaluation Tools (MET)

Ensemble Verification System (EVS)

3.2 Techniky vizualizácie vo verifikácii

3.2.1 Bodový graf

Najjednoduchším spôsobom ako analyzovať vzťah dvoch náhodných premenných je *bodový graf*, ktorý je inak nazývaný aj *korelačný diagram* a známy je tiež pod svojim anglickým pomenovaním *scatter plot*. Bodový graf je vhodný na štúdium kolerácie dvoch premenných a taktiež je výborný pri odhaľovaní takzvaných *outlier*-v, teda hodnôt, ktoré sa nejakým spôsobom výrazne odlišujú od tých ostatných. Taktiež nám nepriamo podáva správu o distribúcii hodnôt, čo však pri ich veľkom počte môže byť skreslené, keďže sa body začínú postupne prekrývať, a tak nemožno určiť v akej oblasti je viac, či menej bodov.

Konštrukcia bodového grafu

Bodový graf na svoju konštrukciu využíva karteziánsku sústavu súradníc. Dve náhodné premenné, ktoré chceme porovnať vizualizujeme tak, že spravíme zobrazenie jednej premennej na x -ovú súradnicovú os a zobrazenie druhej premennej na y -ovú os. Následne v danom bode nakreslíme určený symbol, čo zvyčajne býva čierna bodka alebo krúžok. Na obrázku XYZ vidíme príklad výsledného bodového grafu, ktorý vznikne takýmto postupom.

Bodový graf ponúka mnoho spôsobov, ako pridať ďalší rozmer informácie do vizualizácie. Môžeme zvoliť odlišnú súradnicovú sústavu, čím veľmi jednoducho vytvoríme napríklad 3D bodový graf [TODO ref]. Ďalšou možnosťou rozšírenia je zmena vykresľovaného symbolu, ktorému môžeme nastavovať rôzne parametre, do ktorých možno zakódovať nové informácie. Zvyčajne sú týmito parametrami farba[TODO ref], alfa-transparencia[TODO ref], veľkosť[TODO ref], tvar [GSS, ESPUEPPAS] a podobne. Tieto modifikácie si v komunite vizuálnej analýzy vyslúžili aj vlastné názvy a teoretické zázemie, avšak v našej práci sa týmto druhom diagramov nebudeme venovať.

Kantil-kvantil graf

Dôležitou a často využívanou variáciou pre bodový graf je *kvantil-kvantil graf*, skrátene *Q-Q graf* (Q z anglického *quantil*).

Úloha bodového grafu vo verifikácii

3.2.2 Histogram

3.2.3 Krabicový diagram

Krubicový diagram je v anglickej literatúre zvyčajne nazývaný *box plot* alebo na niektorých miestach označovaný tiež ako *box and whisker*¹ *plot*. Odkedy bol prvýkrát publikovaný v roku 1977 [Tukey], uplynulo už takmer 40 rokov a dnes ho považujeme už za štandardnú techniku ako vizualizovať distribúciu hodnôt kompaktným spôsobom. Na svoju reprezentáciu využíva súbor 5 čísel (tzv. *5-number summary*)[Potter], ktoré charakterizujú distribúciu dát robustným spôsobom. Tým, že zredukujeme zvyčajne veľkú dátovú množinu na týchto pár hodnôt ušetríme nielen vzácny vizuálny priestor [Wickham], ale taktiež námahu analytika, ktorý sa snaží preskúmať iba niektoré vybrané charakteristiky.

¹Slovo *whisker* znamená po slovensky fúz, čo naznačuje, že čiary, ktoré spájajú horný a dolný kvartál s hraničnými hodnotami pripomínajú fúzy.

Konštrukcia krabicového diagramu

Na zostavenie krabicového diagramu potrebujeme týchto 5 hodnôt: medián, horný a dolný kvartil, maximum, minimum. (pozri obrázok 3.1) Prvé tri hodnoty sú takzvané kvartily (Q_1 , Q_2 , Q_3), ktoré rozdeľujú súbor dát na 4 rovnako veľké časti a ďalšie dve sú extrémne hodnoty, ktoré ohraničujú celú dátovú množinu.

Zovšeobecnenie kvartilov sú kvantily ktoré rozdeľujú množinu na n rovnakých častí a preto môžeme o kvartile hovoriť ako o 4-quantile. Medián hodnôt je 2-quantil a teda rozdeľuje množinu na 2 rovnaké časti a je definovaný rovnako ako v časti 2.3.5. Ďalej horný (Q_1) a dolný (Q_3) kvartil získame ako medián hodnôt pod a nad hodnotou Q_2 , pričom hodnotu Q_2 nezahrňame do výpočtov.

Na obrázku 3.1 vidíme, že krabica v grafe určuje pozície horného a dolného kvartilu, zatiaľ čo vnútro krabice znázorňuje takzvané *IQR*. Táto skratka označuje *interquartile range*, čo sa dá preložiť ako *medzikvartilový rozsah*. *IQR* definujeme ako rozdiel kvartilov Q_3 a Q_1 :

$$IQR = Q_3 - Q_1$$

IQR nám hovorí o vzdialenosti týchto dvoch kvartilov, preto nám môže byť tento vzorec na pohľad podozrivý, keďže sa javí, že *IQR* by mohlo nadobúdať aj záporné hodnoty. My však vieme z definície Q_3 a Q_1 , že $Q_3 > Q_1$ a ich rozdiel je teda vždy nezáporný (Hovoríme o rozdieli Q_3 od Q_1 , tak ako je definované *IQR*).

Malú obmenu pôvodného návrhu krabicového diagramu od Tukeyho, vidíme na obrázku 3.1 b), kde malé bodky znázorňujú hodnoty nazývané *outlier*, teda hodnoty ležiace ďaleko od hlavného dátového tela, a hviezdička v strede diagramu určuje priemer hodnôt. Môžeme si všimnúť, že konce čiar vychádzajúcich z boxu nemôžu byť extrémny celej množiny dát, ale sú iba extrémami vypočítaných z dát bez *outlier*-ov.

Otázkou zostáva ako určiť, ktorá hodnota je *outlier* a ktorá nie je. Na zodpovedanie tejto otázky sa využíva už spomínaný rozsah *IQR*. Pomocou neho sa definujú hranice *inner fences* (f_1, f_2) a *outer fences* (F_1, F_2), za ktorými hovoríme už o *outlier*-och alebo o ďalekých *outlier*-och. Definované sú nasledovne:



Obr. 3.1: Pôvodný návrh krabicového diagramu, ako bol prezentovaný v práci *Exploratory Data Analysis*(1977)[Tukey]

$$f_1 = Q1 - c \times IQR$$

$$f_2 = Q3 + c \times IQR$$

$$F_1 = Q1 - C \times IQR$$

$$F_2 = Q3 + C \times IQR$$

Konštanty c a C sú v niektorých zdrojoch definované rôzne. Najčastejšie sa však vyskytujú hodnoty $c = 1.5$ a $C = 3$, tak ako ich určil pôvodný autor krabicového diagramu [Tukey].

Ďalšie variácie krabicového diagramu

Popularita krabicového diagramu nevyhnutne viedla k jeho vývoji a modifikáciám. Môžeme hovoriť o dvoch druhoch modifikácií. Jednak *syntaktickej* (vizuálnej), kedy sa zachovávajú všetky vlastnosti a informácie ako v pôvodnom diagrame, len sa menia vizuálne prvky grafu. A modifikácii *sémantickej* pridaním ďalšej popisnej informácie do grafu, čo má na záver vplyv aj na jeho vizuálnu stránku.



Obr. 3.2: a) Klasický krabicový diagram b-f) Vizuálne variácie krabicového diagramu b-c) 2 variácie pre kvartilový graf [Tuf83] c) Skrátенý krabicový diagram [RKR06] e) Range-bar chart [Spe] f) Farebná variácia [Carr]

Na obrázkoch 3.2 b-f) môžeme vidieť niektoré vizuálne variácie krabicového diagramu. Vznik prvých troch motivovala snaha maximalizovať takzvaný *data-ink* [Tuf83], teda množstvo atramentu alebo počet pixlov, ktoré zodpovedajú nejakým dátam. Autori sa teda snažili čo najviac znížiť počet vizuálnych prvkov a ponechať len tie, ktoré skutočne nesú nejakú informáciu.

Na obrázku 3.2 b) a c) vidíme dve z viacerých riešení navrhnutých v knihe *The Visual Display of Quantitative* [Tuf83], ktoré autor nazýva *kvartilový graf*. Preceptuálne štúdie [Stock] však ukázali, že tieto variácie sú výrazne menej presné ako originálny návrh.

Návrh d) ukazuje skrátенý krabicový diagram [RKR06], ktorý sa taktiež snažil zredukovať množstvo okupovaného vizuálneho priestoru. Rozdielom je však to, že jeho účelom nie je existovať samostatne, ale ako súčasť *summary plot*-u, ktorý zahŕňa histogram a ďalšie glify znázorňujúce informácie ako priemer, štandardná odchylka alebo koeficient asymetrie.

Obrázok 3.2 e) znázorňuje predchodcu krabicového diagramu *range* graf alebo tiež nazývaný *range-bar* graf, ktorého autorkou je Mary Eleanor Spear [Spe52].

Ako posledný príklad vizuálnej modifikácie uvádzame pridanie farieb do krabicového



Obr. 3.3: Krabicové diagramy s pridanou informáciou a) Krabicový diagram s variabilnou šírkou b) Vrubkovaný krabicový diagram c) Krabicový diagram s informáciou o šikmosti dát

diagramu. Táto farebná variácia uchováva tvar diagramu, avšak časť nad mediánom je zafarbená inou farbou ako hodnoty pod. Autori článku odporúčajú červenú a modrú farbu, tak ako na obrázku 3.2 f). Cieľom tohto prístupu bolo chápať krabicový diagram ako jednu percepčnú jednotku a nahradiť 5 symbolov jedným, čím sa mala znížiť námaha pozorovateľa pri analýze a taktiež uľahčiť porovnávanie viacerých diagramov navzájom.

Krabicový diagram umožňuje svojim vzhľadom zakódovanie ďalšej informácie do grafu. Na obrázku 3.3 vidíme aspoň niektoré najčastejšie úpravy...

Len rok po oficiálnom publikovaní krabicového diagramu vznikol článok [McGill], ktorý zhŕňa jeho tri najčastejšie používané modifikácie, z ktorých prvé dve môžeme vidieť na obrázkoch 3.3a) a 3.3b) a tretí je ich kombináciou. V prvom prípade sa využíva šírka boxu na zakódovanie veľkosti množiny, ktorú skrýva za sebou diagram. Takýto graf sa nazýva *Krabicový diagram s variabilnou šírkou*. V druhom prípade ide o takzvaný *Vrubkovaný krabicový diagram*. V tomto grafe sú pridané *vrúbky*, ktoré zhruba naznačujú ako výrazné sú rozdiely v miere spoľahlivosti rôznych dátových množín.

V niektorých prípadoch krabicový diagram zakrýva skutočný tvar dát, teda jeho šikmosť alebo modalitu, keďže jeho vzhľad nabáda k tomu, aby si užívateľ myslel, že sú dát zacentrované na stred a unimodálne. V práci s názvom *Can the Box Plot be Improved?* [...]

autor uvádza príklad kedy skutočne rôznorodé dáta generujú rovnaký krabicový diagram. Tento problém rieši elegantným a čistým spôsobom pridaním hrubej čiary na základe koeficientu asymetrickosti γ . Na obrázku 3.3c) môžeme vidieť zľava asymetrické dáta, na stred zarovnané dáta a bimodálne dáta.

Krubicový diagram umožňuje mnoho ďalších rozšírení napríklad pridaním informácie o hustote dát (histogramový krabicový diagram, vázový diagram, huslový diagram [...]), rozšírením pre viacrozmerne dáta (vrecový graf, 2D krabicový diagram [...]) alebo zobrazením hodnôt v inej súradnicovej sústave (napríklad polárnej - vejárový graf [...]). Opis týchto techník je však nad rámec tejto práce, preto ho tu ani nebudeme uvádzať.

Úloha krabicového diagramu vo verifikácii

Ako sme spomenuli v úvode tejto sekcie, vo všeobecnosti je úlohou krabicového diagramu zobrazíť distribúciu dát v kompaktnom tvare a teda slúži na rýchle porovnanie distribúcií viacerých skupín dát. Pri verifikácii spojitej predpovede sa používa na viacero účelov a my tu spomenieme len niekoľko z nich.

V prvom rade ide o porovnanie distribúcie predpovedí s distribúciou pozorovaní za istý časový interval. V takomto prípade máme vedľa seba iba dva krabicové diagramy, ktoré navzájom porovnávame. Použitie krabicového diagramu v takomto prípade, kedy jeden graf pozostáva iba z niekoľkých (2 až 4) krabicových diagramov považujeme za zbytočné. Pri takomto počte nie je potreba na redukcii vizuálnych prvkov a existujú lepšie techniky na vizualizáciu distribúcie, ktoré sprostredkujúviac informácie a teda môže byť analýza efektívnejšia.

Ďalším použitím je porovnanie distribúcie chýb predpovedí, či už pre rôzne merania, konkrétne predpovedané časy, rôzne predpovedné modely a podobne. Tu považujeme použitie krabicového diagramu za opodstatnené, keďže ide zväčša o porovnávanie väčšieho množstva distribúcií, a tak je jeho jednoduchosť, čitateľnosť, kompaktnosť a iné jeho vlastnosti potrebné.

3.2.4 Time series plot

Kapitola 4

Návrh riešenia

4.1 Návrh systému

4.2 Návrh vizualizácie

4.2.1 Charakteristika dát

4.2.2 Špecifikácia požiadaviek na vizualizáciu

4.2.3 Návrh farebnej palety

4.2.4 Návrh vizualizácie distribúcie chýb

Pri verifikácii predpovede spojitej premennej sme použili štatistické metódy spomenuté v sekcii 2.3, ktorých výsledok sme následne vizualizovali. Pôvodné dáta však zostali skryté za použitým matematickým modelom, a tak sme stratili informáciu o distribúcii chyby. Pri verifikácii sa štandardne používajú dve metódy na priamu, či nepriamu vizualizáciu a analýzu distribúcie, ktoré sme opísali v sekcii 3.2. Týmito metódami sú bodový graf (pozri podsekciiu 3.2.1) a krabicový diagram (pozri podsekciiu 3.2.3).

Pri návrh vizualizácie sme vyskúšali niekoľko vizualizačných techník a zvážili ich silné a slabé stránky.

Graf hustoty

Kvantilový diagram

Funkcionálny krabicový diagram

Pre pochopenie dát je dôležité, aby sme sa vedeli pozrieť na hodnoty v ich kontexte. Všetky predošlé techniky uvažovali o chybe ako o samostatnej hodnote pre určitý predpovedný čas, avšak chyby sa nenachádzajú len v kontexte predpovedného času, ale aj v kontexte konkrétnej predpovede. Preto môžeme uvažovať o predpovediach ako o funkciách $x_i(t)$, kde $i \in \{1..n\}$ je poradie predpovede a $t \in I$ je čas predpovede, kde I je časový interval predpovede z \mathbb{R} (v našom prípade sa jednalo o dvojdnú, teda 48 hodinú predpoveď).

Porovnanie metód ...

Tu bude pekný obrázok a obkeci.

Kapitola 5

Implementácia

Kapitola 6

Výsledky

Literatúra

- [Cas09] Barbara Casati. *Verification of continuous predictands*. Joint Working Group on Forecast Verification Research (JWGFVR), Jún 2009.
- [CWS⁺08] B. Casati, L. J. Wilson, D. B. Stephenson, P. Nurmi, A. Ghelli, M. Pocerlich, U. Damrath, E. E. Ebert, B. G. Brown, and S. Mason. Forecast verification: current status and future directions. *Meteorological Applications*, 15(1):3–18, 2008.
- [FJB12] Tressa L. Fowler, Tara L. Jensen, and Barbara G. Brown. *Introduction to Forecast Verification*. 2012.
- [Gol] Professor Brian Golding. Weather forecasting part 1. <http://www.rmets.org/weather-and-climate/weather/weather-forecasting>. [Prístupné online: 6.12.2014].
- [LE11] Dr. Arlene Laing and Dr. Jenni-Louise Evans. *Introduction to Tropical Meteorology 2nd Edition*. UCAR, Október 2011.
- [Lyn07] Peter Lynch. The origins of computer weather prediction and climate modeling. *Journal of Computational Physics* 227 (2008) 3431–3444, Február 2007.
- [NCE14] NCEP. Inventory of Data Products on the NOAA Servers. <http://www.nco.ncep.noaa.gov/pmb/products/>, November 2014. [Prístupné online: 10.11.2014].

- [Nur03] Pertti Nurmi. *Recommendations on the verification of local weather forecasts*. European Centre for Medium Range Weather Forecasts, December 2003.
- [Poc11] Matthew Pocerich. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, chapter Appendix - Verification Software, pages 232–240. John Wiley and Sons, Ltd., 2nd edition, December 2011.
- [Vas98] Tim Vasquez. *Observer Handbook*. International Weather Watchers, 1995, 1998.
- [WCSDG⁺08] Joseph B. Klemp William C. Skamarock, Jimmy Dudhia, David O. Gill, Dale M. Barker, Michael G. Duda, Xiang-Yu Huang, Wei Wang, and Jordan G. Powers. *A Description of the Advanced Research WRF Version 3*. National Center for Atmospheric Research, Jún 2008.
- [Wei14] Eric W. Weisstein. Statistical Median.
<http://mathworld.wolfram.com/StatisticalMedian.html>, 2014.
- [WMO94] WMO. *A GUIDE TO THE CODE FORM FM 92-IX Ext. GRIB Edition 1*. WMO, Máj 1994.
- [WMO03] WMO. *Introduction to GRIB Edition1 and GRIB Edition 2*. WMO, Jún 2003.