

Oponentský posudek disertační práce Mgr. Davida Hansliana „Analýza výsledků měření větru“, zpracované na Katedře meteorologie a ochrany prostředí MFF UK, červen 2014.

Autor si v disertační práci klade tři cíle:

- analyzovat nežádoucí jevy ovlivňující výsledky měření větru (kap. 2),
- analyzovat metody efektivního doplnění či prodloužení časových řad větru (kap. 3),
- analyzovat faktory vytvářející větrné klima ČR (kap. 4).

K jednotlivým kapitolám práce:

Úvod: Dovoluji si autora upozornit, že nejen vítr je vektor, ale např. i intenzita přímého slunečního záření je vektor.

Kap. 2: Nejde o analýzu všech nežádoucích jevů, ale o rešerši specifické oblasti měření větru, a to o výčet a analýzu vlivu umístění měrných čidel na naměřená data s ohledem na překážky v jejich bezprostředním i vzdálenějším okolí. Práce se nezabývá se např. problémem citlivosti a setrvačnosti měření apod. Tato kapitola je zajímavá a nemám k ní zásadní připomínky. Z formálního hlediska postrádám označení abscisy v levé části obrázku č. 2.20 na str. 35, místo obratu „umístění měření větru“ by snad byl případnější obrat „umístění měrného čidla“. Speciálně oceňuji nápad simulovat různé modely tvaru profesionální meteorologické stanice na str. 30 metodou WASP.

Kap. 3 se zabývá metodami prodloužení či doplnění časových řad měření větru. V rámci rešerše dostupné literatury jsou nejprve diskutovány okolnosti ovlivňující přesnost simulace a uveden přehled existujících metod simulace a doplňování větroměrných řad. Ve vzorci (3) str. 49 mne poněkud překvapilo, že průměrné rychlosti v trénovacím období nejsou vztaženy k termínovým hodnotám, ale ke všem hodnotám. Jde o překlep?

Str. 46: V práci je zmiňován požadavek homogenity dat. Zajímalo by mne, jakým způsobem autor tento problém konkrétně řešil.

Str. 59 až 63. Kromě literárního přehledu metod simulace v předchozí části práce autor zde navrhuje dvě vlastní metody simulace. U metody 1 bych uvítal bližší vysvětlení následujících větných sekvencí (str. 58): „*Přiřazení termínů T_t k termínům T_d je prováděno tak, aby všechny termíny T_t spadající do kategorie k byly při doplňování zastoupeny shodně*“ a „*termínům T_d přiřazovat termíny T_t z odpovídajícího kvantilu rozdělení rychlostí větru*“. Možnost náhodného přiřazování termínu považuji za dosti dubiózní. Na str. 59 mi není zřejmé, co je považováno za skupinu nízkých rychlostí větru? Ke str. 60: „termín“ nemá délku, minimálně jde o nešťastnou formulaci. Prosím o bližší vysvětlení.

Z hlediska původnosti metod jde o jisté modifikace publikovaných metod, přičemž uvnitř těchto metod existuje značná variabilita přístupů v oblasti členění dat do kategorií.

Str. 65: Při popisu vstupních měrných dat postrádám informaci o tom, zda na všech použitých stanicích byla stejná čidla a s jakými vlastnostmi.

Str. 67: Osobně bych si netroufal dělat ze šestihodinových dat hodinová data (viz reanalýza NCEP/NCAR), pokud bych chtěl tuto reanalýzu využít, tak bych separátně řešil úlohu pro šestihodinové rozlišení. Jinak využití dat z reanalýz pro referenční řady považuji za rozumné.

Str. 68 až 69: Nesouhlasím s autorovým názorem (bod 1ii), že *„aplikace většího počtu částečně se překrývajících kombinací trénovacích dat přináší novou informaci a má lepší vypovídací schopnost pro vzájemné srovnání různých metod či referenčních dat než by mělo použití pouze 4 navzájem nezávislých testů“*. V každém případě jde o nedoložené tvrzení. Podobně bod 2) obsahuje tvrzení, která jsou vzájemně poněkud sporné. Na jedné straně tvrzení, že rozkouskování trénovacího období zvyšuje zachycenou variabilitu podmínek, na druhé straně zdůraznění vhodnosti použít souvislý blok (nebo několik souvislých bloků) dat.

Autor na konci této subkapitoly shrnuje: *„Popsaným způsobem je možno ze 4 let dat vytvořit celkem $3 \times 4 \times 4 \times 4$, tedy 768 různých trénovacích souborů. Ty se však v některých případech liší pouze jediným měsícem a kalkulace všech variant by proto přinášela značné duplicity. Mimo jiné také z důvodu výpočetní náročnosti bylo proto testování prováděno pouze pro omezený počet variant (běhů). Jednotlivé běhy byly voleny náhodně skrze náhodnou volbu prvního měsíce trénovacího období a skrze náhodnou volbu roků pro první měsíce zbývajících tří bloků. V rámci provedeného testování bylo počítáno vždy 50 náhodně vybraných běhů. Výběr běhů byl v rámci jednoho porovnání vždy shodný, různá porovnání však byla prováděna s různými sestavami“*. K tomu mám několik výhrad. Jsem toho názoru, že náhodný výběr 50 ze 768 vzájemně více či méně se překrývajících souborů je v daném případě velmi problematická metoda, která v daném případě spíše vnáší další zmatky a nežádoucí až nekontrolovatelný statistický šum do dalších analýz. Totéž ovšem platí i v případě, kdyby bylo zpracováno všech 768 variant (jen to zašumění by bylo alespoň ve všech případech stejné). Problém je navíc umocněn i tím, že různá porovnání byla prováděna s různými trénovacími soubory. V práci naprosto postrádám jakékoli bližší údaje o frekvenčním rozložení jednotlivých výchozích souborů, tím méně porovnání frekvenčních rozložení jednotlivých variant trénovacích souborů. Jinými slovy, naprosto není jasné jak tento přístup a tento „náhodný výběr“ deformuje celkové výsledky a celkově ho považuji za scestný. Očekával bych i zpracování situace, kdy testovací řada zahrnuje celé čtyřleté období.

Str. 70, nulová metoda, citace z disertace: „Bylo předpokládáno, že větrné podmínky v cílovém období jsou shodné s podmínkami v období trénovací. Konkrétně byla časová řada konstruována tak, že byly každému termínu doplňovaného období přiřazeny směr a rychlost větru na cílové řadě v náhodně vybraném termínu trénovacího období. Vlivem náhody se výsledky Nulové metody mohou navzájem nepatrně lišit i v případě dvou totožných výpočtů.“ Pokud autorovi správně rozumím, tak touto metodou byla vytvořena naprosto nerealistická časová řada, jejíž frekvenční rozložení se podle autora blíží frekvenčnímu rozložení trénovacího období. Osobně bych za nulovou metodu považoval buď prosté převzetí frekvenčního rozdělení trénovacího období cílové řady nebo doplnění cílové řady např. metodou1, ale s využitím trénovacího období cílové, nikoli referenční řady. Je žádoucí, aby autor konkrétní konstrukci jasně vysvětlil, nejlépe na příkladě.

Ke straně 71: autor definuje Standardní variantu, vzhledem k níž dále porovnává ostatní varianty. Pominu-li, že i tato Standardní varianta v sobě zahrnuje spoustu problémů praktické realizace, pak jediným rozumným a nejjednodušším srovnávacím etalonem je skutečně napozorovaná řada.

Str.72až74, hodnocení rozdělení rychlostí větru: K použití statistiky Chí-kvadrát bych prosil autora o přesnější údaje, jak ji použil. Statistika zdaleka nemusí být tak nevhodná, jak autor uvádí, zejména při použití „testu dobré shody při známých parametrech“. Kritéria pro hodnocení přesnosti metod MCP celkem vhodná, i když si lze představit i jiné charakteristiky, nezávislé na tvaru rozdělení rychlostí větru. Pokud jde o normování, pak při porovnávání metod stejné dvojice referenční a doplňované stanice resp. stejné cílové řady a různých referenčních řad normovat není třeba (situace tab. 3.3). V opačném případě vzniká otázka vhodného postupu normování s přihlédnutím k normalitě rozložení rychlosti větru apod. (situace tab. 3.4). K prostému posunutí o rozdíl průměrné rychlosti jsem dosti skeptický, pomineme-li i možnost vzniku záporných rychlostí.

Str. 75 až 77 – porovnání výsledků pro různé referenční a cílové řady. Souhlasím s autorem, že pro interpretaci výsledků jsou souhrnné tabulky 3.3 a 3.4 jen velmi orientační. Nicméně i zde bylo možno a žádoucí použít deskriptivní charakteristiky, umožňující výrazně lepší souhrnný pohled na výsledky dosažené pro různá trénovací období. Z tabulek je také vidět, že použití různých hodnotících statistik u stejné řady může vést k rozdílným výsledkům (viz např. sloupec bias versus ostatní sloupce). Alespoň pokus autora o nějaké souhrnné kritérium využívající různých metod bych očekával. Nicméně kladně kvituji, že autor si je problematických možností interpretace vědom, i když výpovědní možnosti prezentovaných tabulek patrně hodnotím ještě níže než on. Samostatná publikace tabulek 3. 3., 3.4, 3.7 by byla vysloveně zavádějící. Naštěstí autor se při podrobnějších interpretacích výsledků drží spíše

jednotlivých stanic, jejichž výsledky uvádí v přílohách. I zde ale postrádám alespoň pokus o analýzu závislosti výsledků na volbě testovacího období.

Str. 82 až 88 – porovnání základních metod MCP. Zde autor zpracoval jiné sestavy testovacích řad než v předchozím případě, což samo o sobě znemožňuje jakékoli srovnání s předchozími výsledky. Zpracování provedl pouze pro čtyři cílové řady v různých geografických poměrech, což je samo o sobě rozumné. V rámci těchto prací porovnává různé metody simulace, včetně různého členění směrů větru. Uvedení souhrnné tabulky 3.7 pro všechny řady v tomto případě vůbec není na místě, daleko vhodnější by bylo rovnou použít tabulky pro jednotlivé cílové řady s rozlišením referenčních řad.

Pokud jde o výsledky porovnání „vlastních metod“ s ostatními, autor spíše potvrzuje to, co píše v rešerši o jiných metodách, tj. v každé práci vychází autorova metoda lepší než ostatní.

Kap. 4, větrné poměry ČR. Zde autor uvádí již publikované výsledky zpracování mapy rychlosti větru na území ČR a využitím kombinace několika standardních modelů (VAS, WASP a PIAP), přičemž autor uvádí, že je autorem koncepce myšlenky použití této kombinace a velmi mnoho práce udělal při přípravě výchozích dat. Chtěl bych zdůraznit, že tato práce bývá často podceňována, nicméně často tvoří až 90% náročnosti celé práce, je zcela podstatná pro výsledek práce a s výjimkou odstranění zcela zjevných elementárních chyb vyžaduje i značnou odbornou erudici a zkušenosti s prací s daty. Autor krom toho uvádí řadu publikovaných větroměrných studií pro území ČR. Pochopitelně neuvádí (ani dost dobře nemůže), nepublikované studie zpracované např. v rámci ČHMÚ jako zakázky (včetně mapy průměrných rychlostí větru ČR začátkem devadesátých let). Samu myšlenku použití kombinace různých modelů považuji za nosnou a výsledná mapa průměrné rychlosti větru je zatím nejkvalitnější svého druhu.

Souhrnné hodnocení

Oponovaná disertační práce se zabývá velmi aktuálním a stále živým tématem, který vystupuje do popředí v souvislosti s praktickými technickými a zemědělskými aplikacemi, jako jsou větrné elektrárny, extrémní rychlosti větru, výsušné vlastnosti větru, modely šíření emisí apod. Problematikou extrémních rychlostí větru se tato práce nezabývá, což je v pořádku, neboť jde o specifickou problematiku a i tak je zpracovávané téma velmi široké.

Z formálního hlediska práci posuzuji takto: Práce je dobře členěna. Ojedinele se v ní vyskytují překlepy a někdy nešťastné nebo nejasné formulace, které zde až na výjimky neuvádím, protože je považuji za marginální. Na závadu čitelnosti práce je i to, že autor podstatnou část obrazové a tabelární dokumentace uvádí správně v závěru práce, ale řadu tabulek a obrázků má v textu, aniž by při odkazu na tyto tabulky a obrázky uváděl, na které straně jsou. Nejde ale o

podstatnou námitku. Text disertace je dosti rozvláčný na úkor přehlednosti a srozumitelnosti. Jistě by bylo možno text práce zpracovat úsporněji a přehledněji, některé vývody se opakují nebo jsou vzájemně v rozporu. Literatura je bohatá a dobře citovaná.

Práce obsahuje velmi dobrou rešerši a rozbor některých stránek problematiky měření větru, týkajících se překážek v bezprostředním i vzdálenějším okolí měrného čidla. K použitým metodám a statistickému zpracování empirického materiálu mám však řadu podstatných výhrad, které jsem zmínil v posudku. Tyto nedostatky, jsou v interpretacích výsledků zeslabeny, neboť autor při nich vychází především z výsledků pro jednotlivé stanice a ze zkušeností z práce s konkrétními daty. Z hlediska požadavků matematicko-fyzikální fakulty a publikovatelnosti výsledků jsou však uváděné výhrady závažné.

Práce má rozhodně význam pro další rozvoj vědního oboru, i když zejména z hlediska inspirativního. Jednoznačný význam pro praktickou aplikaci má větrná mapa ČR, v jejímž rámci autor nepochybně využil i metod, prezentovaných v kapitole 3 disertační práce.

Aplikace navržených metod doplňování řad je možná, nicméně bych v daném stavu nepovažoval za účelné je nějak formálně preferovat resp. zobecnit ve formě metodického pokynu, čehož je si autor patrně vědom. Krom toho v běžné klimatologické praxi se často používá řada nepublikovaných metod na stejné úrovni, které ovšem z hlediska vědecké obce vlastně neexistují. Další tvůrčí vývoj metod uvedených v práci by byl žádoucí. Možnosti v tomto směru jsou obrovské.

Disertační práce prokazuje předpoklady autora k rešeršní činnosti a rovněž v interpretační oblasti individuálních výsledků, pokud nevyžadují složitější statistiku. Kladně hodnotím i snahu autora uvažovat o kvalitě výchozích dat a určitě autorovi nelze upřít jisté tvůrčí schopnosti. Stěžejní část disertace v kap. 3 mne však neuspokojila a neprokázala dostatečné odborné znalosti či schopnosti jak při formulaci dílčích úkolů a volbě vhodných datasetů, tak zejména při jejich vyhodnocování. Při doktorandském studiu na MFF UK se nelze omezit na „okometrické“ zhodnocení aritmetických průměrů. Je nutné, aby do oponentury autor byl schopen na výše uvedené výhrady konstruktivně reagovat formou doplňujících výpočtů a vyhodnocení, alespoň na části empirického materiálu. Pouze v takovém případě, budu moci konstatovat, že autor má předpoklady k samostatné tvořivé práci, jinak nikoli.

V Praze 12. července 2014

RNDr. Vít Květoň, CSc.