

# Validácia trojrozmernej variačnej asimilácie dát v ansámblovom systéme ALADIN-LAEF.

Martin Imrišek

Slovenský hydrometeorologický ústav, Bratislava

## **Abstract**

---

The 3D variational data assimilation is one of the main approaches how to prepare the most reliable initial conditions of upper air for the integration of numerical weather prediction model. The first task was the implementation of various data types to 3DVar of ALADIN–Limited Area Ensemble Forecasting (LAEF) system. Following task was impact validation of data assimilation to upper air (separately and combined) and implementation of data perturbation. The main goal was verification of ALADIN–LAEF experiment with 3DVar (*phase II*) against ALADIN–LAEF experiment without 3DVar (*phase I*).

**Annotation :** This report discusses about data assimilation and impact validation of various data types in upper air of ensemble weather forecasting system ALADIN–LAEF.

**Keywords:** *data assimilation, ensemble weather forecast*

**Anotácia :** Tento príspevok pojednáva o asimilácii a posúdení dopadu rôznorodých dát na systém ansámblových predpovední počasia ALADIN–LAEF.

**Kľúčové slová :** *asimilácia dát, ansámblové predpovede počasia*

## Pod'akovanie

Chcel by som sa pod'akovať Alenke Trojákovej a Martinovi Bellušovi za poskytnuté rady a podporu pri riešení tejto témy. Taktiež by som chcel pod'akovať programu Regional Cooperation for Limited Area modeling in Central Europe za možnosť pracovať na tejto problematike.

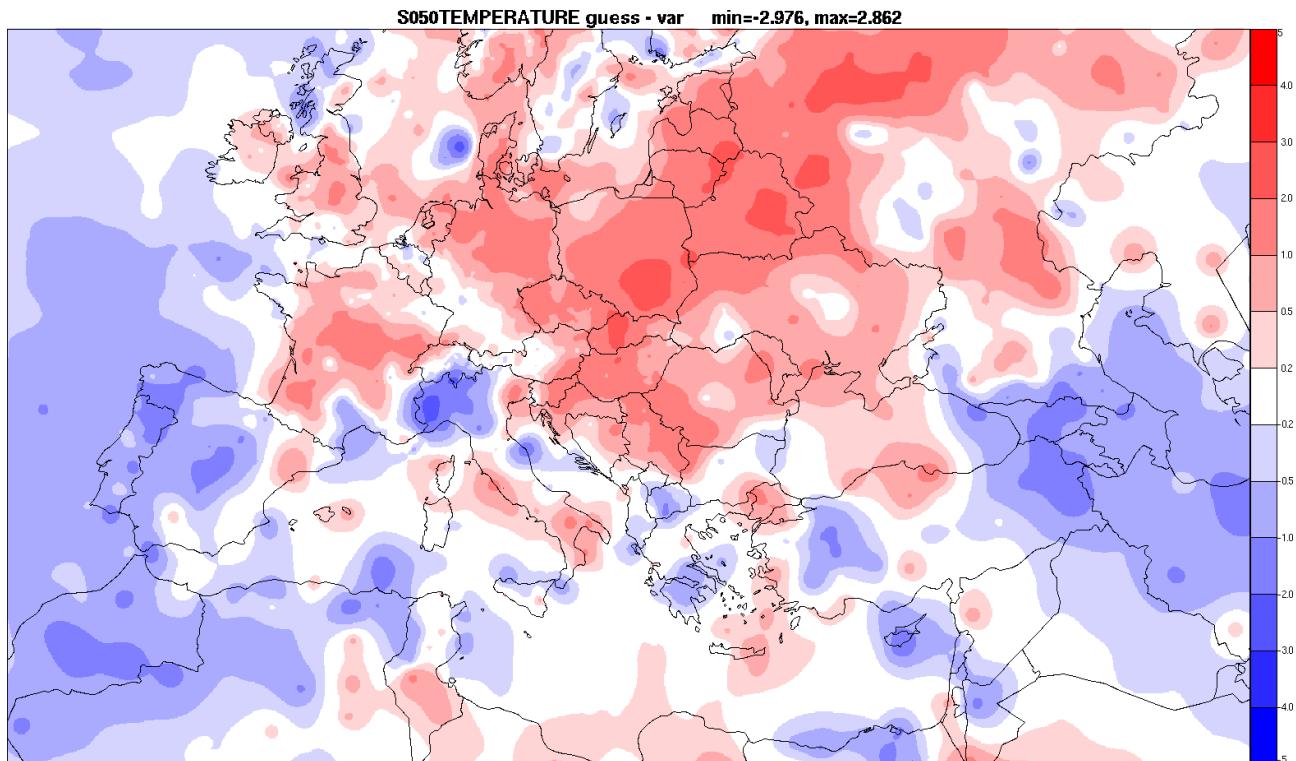
## 1 Úvod

Trojdimentzionalna variačná (3DVar) asimilácia je jednou z hlavných metód spôsobu tvorby počiatočných podmienok (analýzy) atmosféry pre numerické predpovedné modely počasia na ohraničenej oblasti. Prvou úlohou bolo pripravenie dát na asimiláciu do ALADIN–LAEF (Limited Area Ensemble Forecasting – ansámblového predpovedného systému na ohraničenej oblasti). Väčšina použitých dát bola získaná z dátovej služby OPLACE. ALADIN–LAEF systém sa skladá zo šestnástich členov ansámblu, model má  $4.8 \text{ km}$  horizontálne rozlíšenie, 60 vertikálnych hladín a na výpočet bola použitá verzia modelu cy40t1 s ALARO–1 fyzikou. Doména modelu pokrýva takmer celé územie Európy  $\lambda = (351.80^\circ, 67.06^\circ)$ ,  $\phi = 25.76^\circ, 52.85^\circ$ . Ansámblový systém sa skladal z troch krovov ESDA (Ensamble of Surface Data Assimilation) čo je asimilácia pozemných meraní do pôdnich a povrchových parametrov modelu, atmosférické spektrálne zmiešavanie Blending (s ECMWF) s využitím digitálneho filtrovania a integrácia t.j. výpočet predpovedí s 90 sekundovým krokom (12 alebo 24 hodín so šesť hodinovým spájaním s riadiacim modelom, takzvaný Coupling). Do modelu boli asimilované iba vybrané druhy dát. V prvej fáze sme asimilovali dáta jednotlivo do modelu, aby sme overili technickú správnosť asimilácie. Následne sme pristúpili k perturbácii asimilovaných meraní a výpočtu ansámblovej predpovede so šestnásťimi členmi pre obdobie dvoch týždňov od 16.05.2016 do 31.05.2016. Ako posledné bolo urobené porovnanie verifikačných parametrov medzi ansámblovým systémom **bez a s** asimilovanými meraniami pomocou 3DVar.

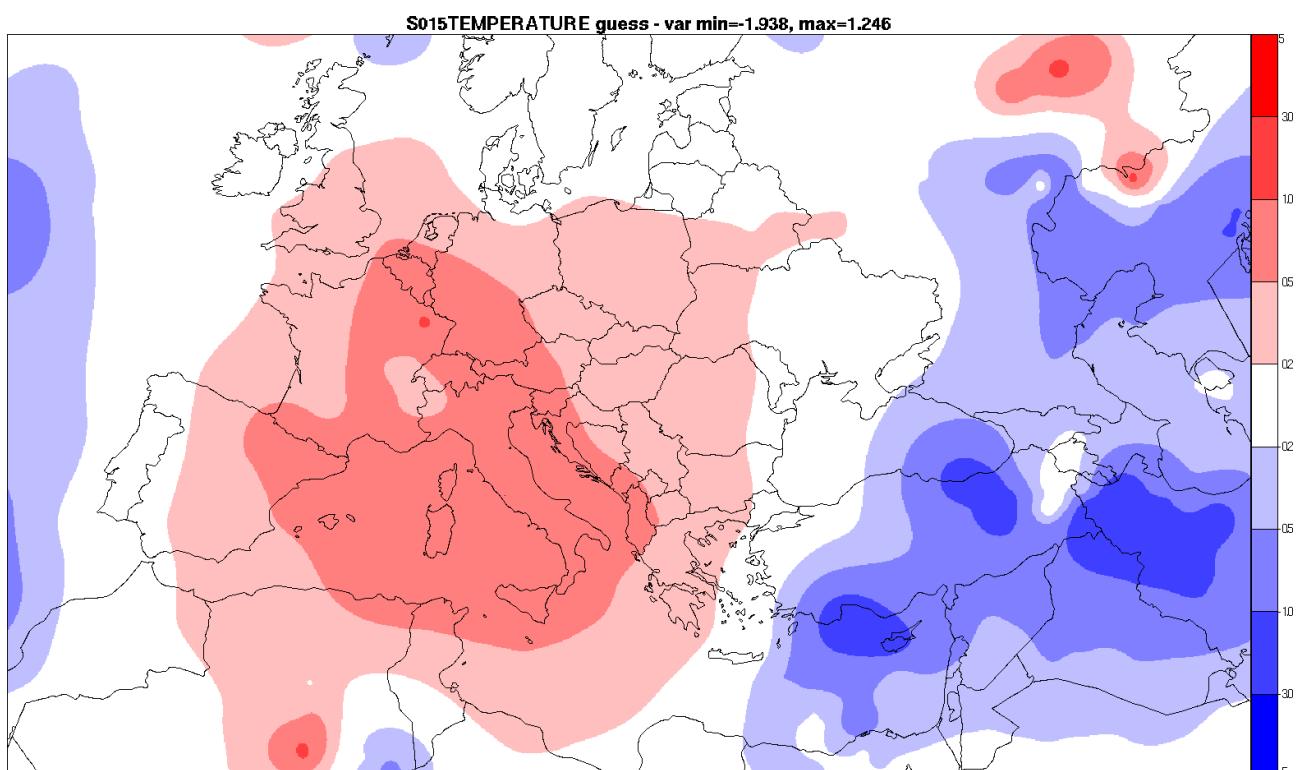
## 2 Technická verifikácia 3DVar asimilácie dát

Prvou úlohou bola verifikácia 3DVar asimilácie vybraných druhov dát ALADIN–LAEF s ALARO–1 fyzikou. Použité boli merania z pozemných automatizovaných meteorologických staníc – SYNOP (OPLACE), dáta z meteorologických snímačov na lietadlach – AMDAR (OPLACE), údaje z rádiosond – TEMP (OPLACE), zložky vetra určené z geostacionárnych družíc – GEOWIND (OPLACE) a totálne zenitové oneskorenia určené pomocou – GNSS (Slovenská technická univerzita). OPLACE je skratka operačného predspracovateľského systému pre projekt RC LACE. Tento systém združuje dáta z Rakúska, Česka, Slovenska, Poľska, Chorvátska, Slovinska, Rumunska a Holandska, jeho hlavnou úlohou je príprava a predspracovanie dát, aby boli priamo použiteľné na asimiláciu do numerických predpovedných modelov počasia. V databáze OPLACE sa nachádzajú SYNOP, SHIP, BUOY, AMDAR, ACARS, TEMP, TEMP MOBIL, EUROPROFILE, GEOWIND, HRWIND, SEVIRI, AMSU-A/B, MHS, HIRS, IASI, ASCAT. V našich experimentoch sme asimilovali iba spomínaných päť druhov dát. Priemerne bolo v bežnom asimilačnom cykle použitých približne: 6200 (37.85%) SYNOP meraní (zahrňujúcich aj ZTD GNSS), 550 (3.35%) AMDAR meraní, 32 (0.20%) GEOWIND párov zložiek vetra a 9600 (58.60%) TEMP záznamov. Na asimiláciu SYNOP, AMDAR a TEMP dát nebolo potrebné vykonať žiadne predspracovanie ani zásahy do zdrojového kódu modelu. Vybrané rozdiely modelových polí medzi analýzou (s asimilovanými dátami) a počiatočným stavom (označovaný ako guess), takzvané inkrementy sú zobrazené na obrázkoch 1, 2 a 3. Počas prvého asimilačného experimentu AMDAR dát z 16.05.2016 o 00UTC boli inkrementy na 55. modelovej hladine sústredené v oblasti letísk v mestách Antalya, Aerhus a Gothenburg.

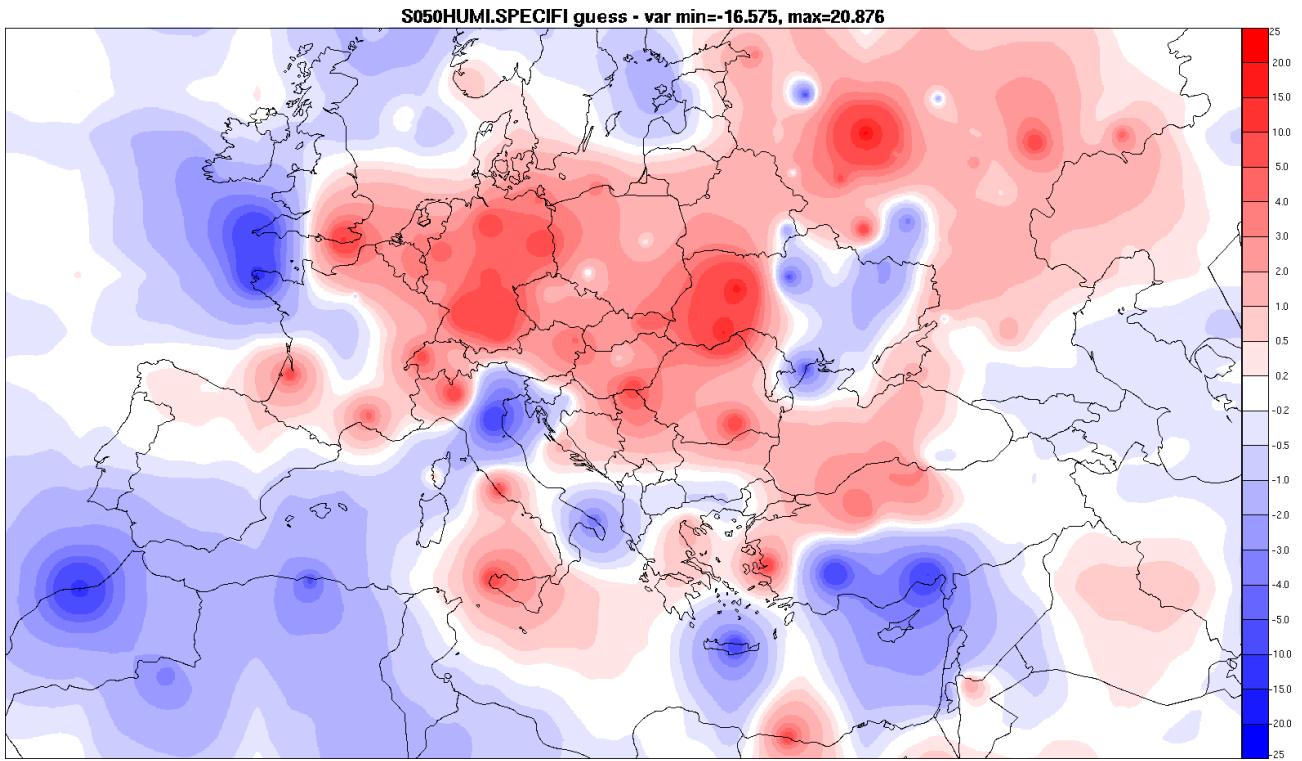
Čo indikuje vzlet lietadiel, pričom najväčšie inkrementy AMDAR dát sú okolo 15. modelovej hladiny, ktorá má štandardnú atmosférickú výšku približne 10300 m na morom.



Obr. 1: Inkrementy teploty [ $^{\circ}\text{C}$ ] na 50. modelovej hladine určené iba na základe asimilácie SYNOP dát.



Obr. 2: Inkrementy teploty [ $^{\circ}\text{C}$ ] na 15. modelovej hladine určené iba na základe asimilácie AMDAR dát.

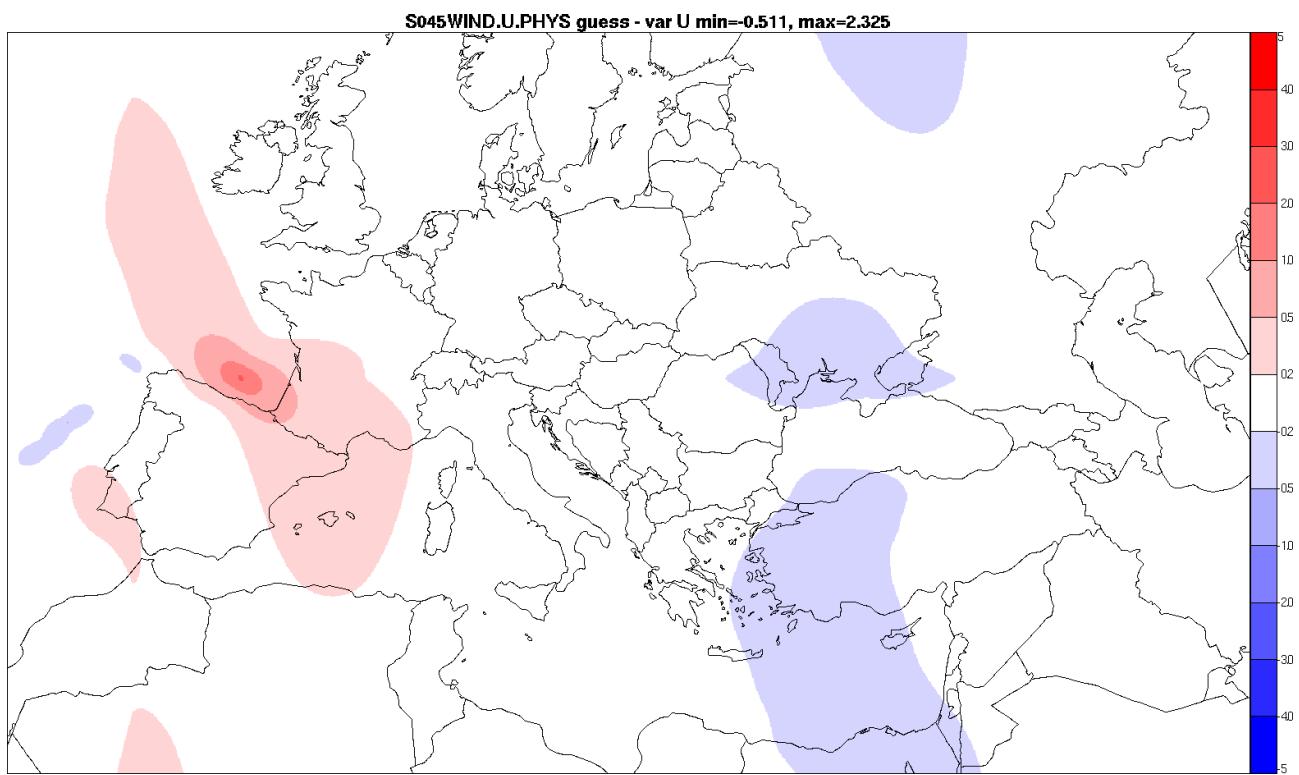


Obr. 3: Inkrementy špecifickej vlhkosti [ $\text{g kg}^{-1}$ ] na 50. modelovej hladine určené iba na základe asimilácie TEMP záznamov.

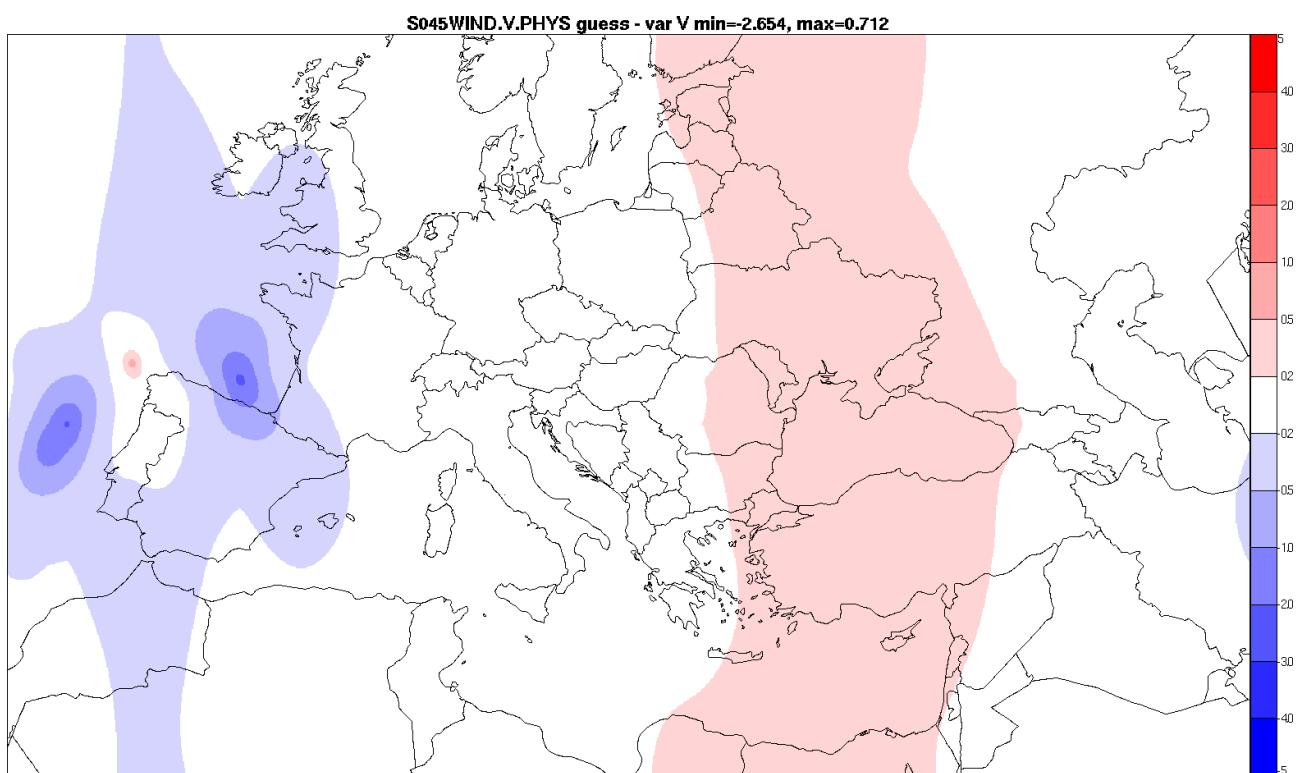
Ako je vidieť, v inkrementoch nie sú zrejmé systematické a ani hrubé chyby a sú v akceptovateľných rozsahoch. Tieto dátu boli následne použité do kombinovanej a perturbovanej asimilácie dát označovanej *phase II* (odsek 3).

## 2.1 Asimilácia GEOWIND dát

Na asimiláciu atmosférických pohybových vektorov bola potrebná úprava zdrojového kódu *mf\_blacklist.b* v časti *Geographical blacklist by satellites*. Išlo o povolenie asimilovania týchto dát nad pevninou, čím sme zabezpečili väčšie množstvo asimilovaných dát. Najväčšie inkrementy, z pohľadu absolútnej hodnoty ( $2.6 \text{ m s}^{-1}$ ) sa nachádzajú na modelovej hladine 45, ako je vidieť na obrázkoch 4 a 5. Táto modelová hladina má štandardnú atmosférickú výšku približne  $860 \text{ m}$  nad morom. Nízky počet asimilovaných dát môže byť zvýšený ladením selekcie dát (thinning) alebo úpravou blacklistovania. Spoľahlivejšiu analýzu je možné dosiahnuť asimilovaním podobných atmosférických pohybových vektorov: HRWIND.



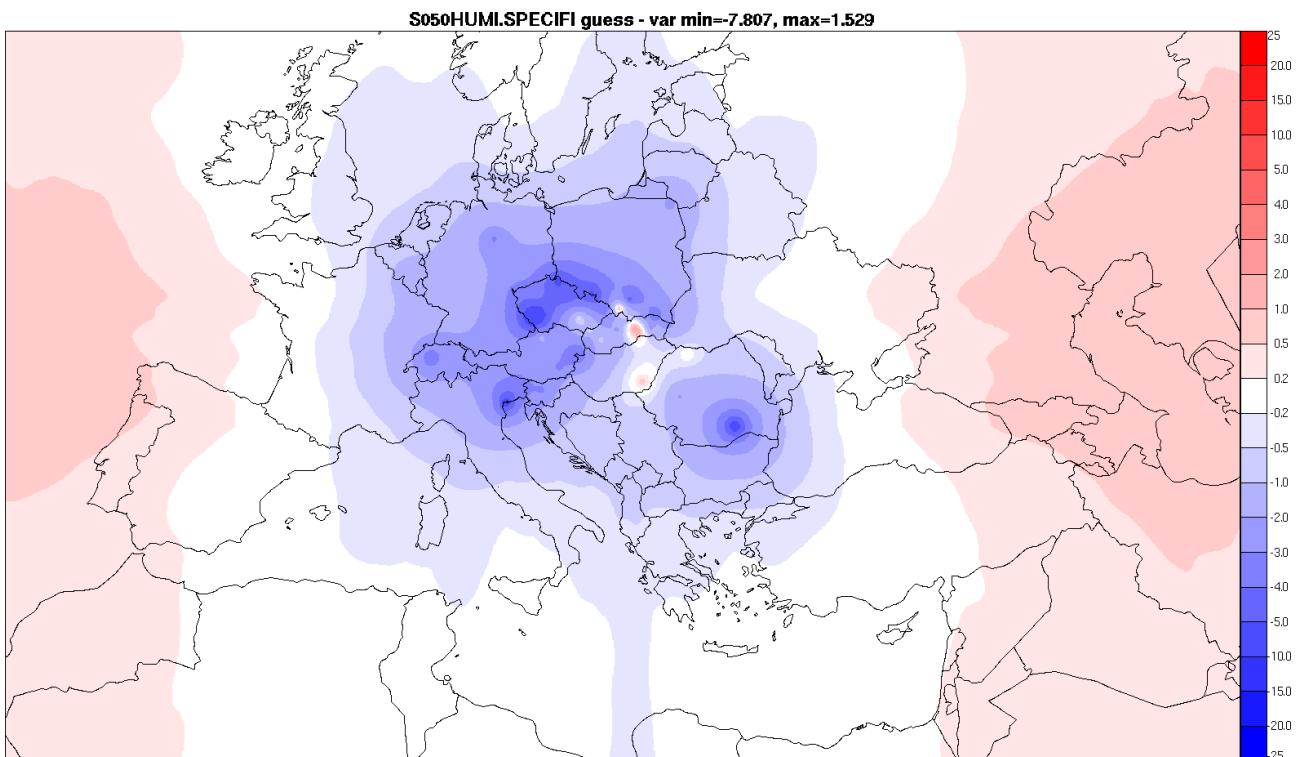
Obr. 4: Inkrementy rýchlosťi vetra [ $\text{m s}^{-1}$ ] v smere zemepisnej šírky na 45. modelovej hladine určené iba na základe asimilácie GEOWIND dát.



Obr. 5: Inkrementy rýchlosťi vetra [ $\text{m s}^{-1}$ ] v smere zemepisnej dĺžky 45. na modelovej hladine určené iba na základe asimilácie GEOWIND dát.

## 2.2 Asimilácia zenitových totálnych oneskorení GNSS

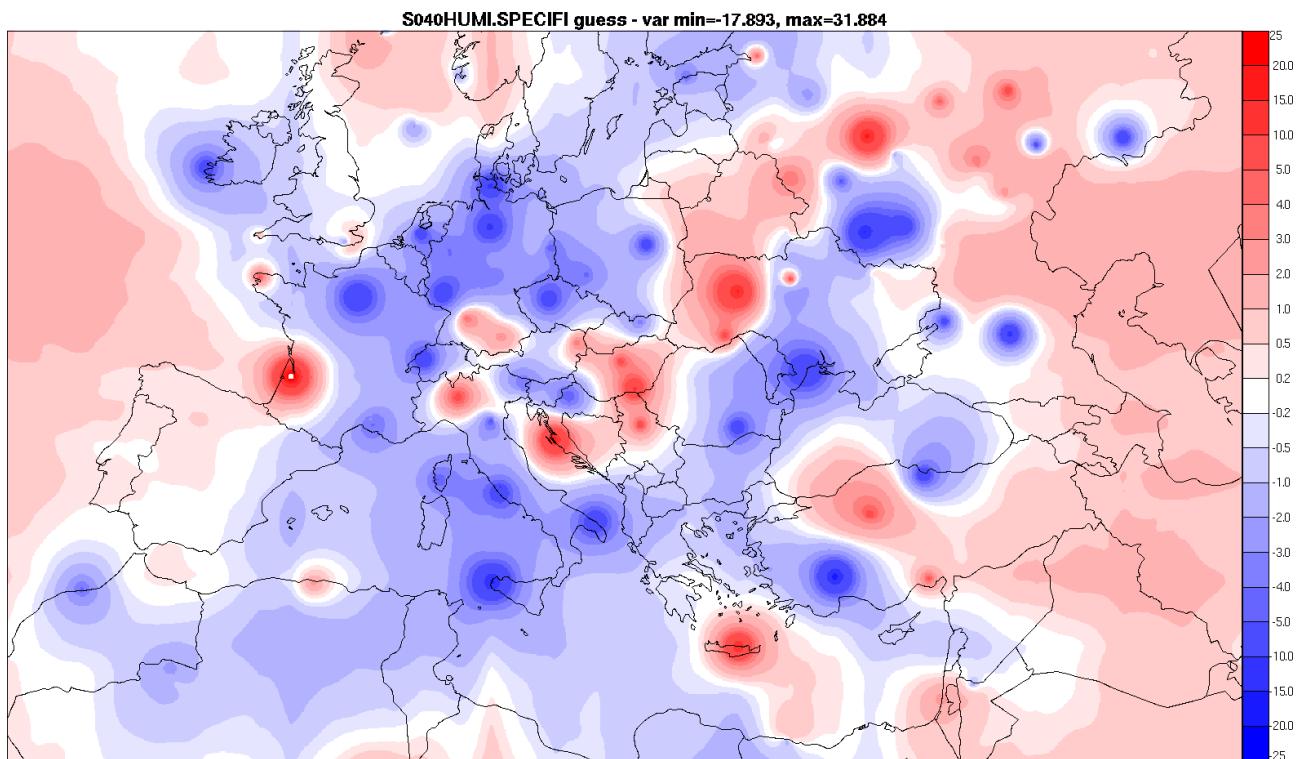
Pred asimiláciou dát GNSS bolo potrebné zostaviť whitelist GNSS permanentných staníc. Whitelist je zoznam staníc, ktoré sú použité do asimilácie a majú určenú korekciu chyby zenitového totálneho oneskorenia. Táto korekcia je spôsobené najmä rozdielom skutočnej nadmorskej a modelovej výšky permanentnej stanice GNSS. Prvým krokom bol výpočet takzvaných first guess (predpovede bez 3DVar asimilácie dát) rozdielov medzi na-mieranými zenitovými totálnymi oneskoreniami a modelovými hodnotami vypočítanými pomocou observačného operátora. Tieto rozdiely sme vypočítali pre všetkých 16 členov ansámblu a pre celé časové obdobie. Prvým zvoleným prístupom bola kombinácia všetkých first guess rozdielov zo všetkých členov a všetkých termínov. Tento prístup sa ukázal ako nevhodný, pretože kombináciou všetkých first guess rozdielov pre jednu stanicu vznikne priveľmi 'špicatý' histogram rozdielov a Pearsonov Chi-kvadrát test ( $H_0$ : Dáta majú normálne roz-delenie) odstráni stanicu pre ďalšie spracovanie. Preto bol navrhnutý ďalší prístup, rozdiely boli testované po jednotlivých členoch ansámblu v rámci celého obdobia. Počet odstránených staníc sa pohyboval od 5 do 12 pre jednotlivé členy. Tretí prístup bol založený na testovaní všetkých členov pre jednotlivé dni pomocou Jar-rque–Bera testu (testovaný je tretí a štvrtý centrálny moment dát,  $H_0$ : Dáta majú normálne rozdelenie). Tento test tak tiež odstránil rôzne množstvo permanentných staníc pre každý deň. Na základe týchto výsledkov bolo rozhodnuté, použiť najlepší deň (so všetkými členmi a najmenším počtom vyradených staníc) a najlepšieho člena (všetky dni s najmenším počtom vyradených staníc) spoločne pri zostavení whitelistu a určení korekcií zenitových totálnych oneskorení. Inkrementy špecifickej vlhkosti po asimilácii ZTD sú na obrázku 6.



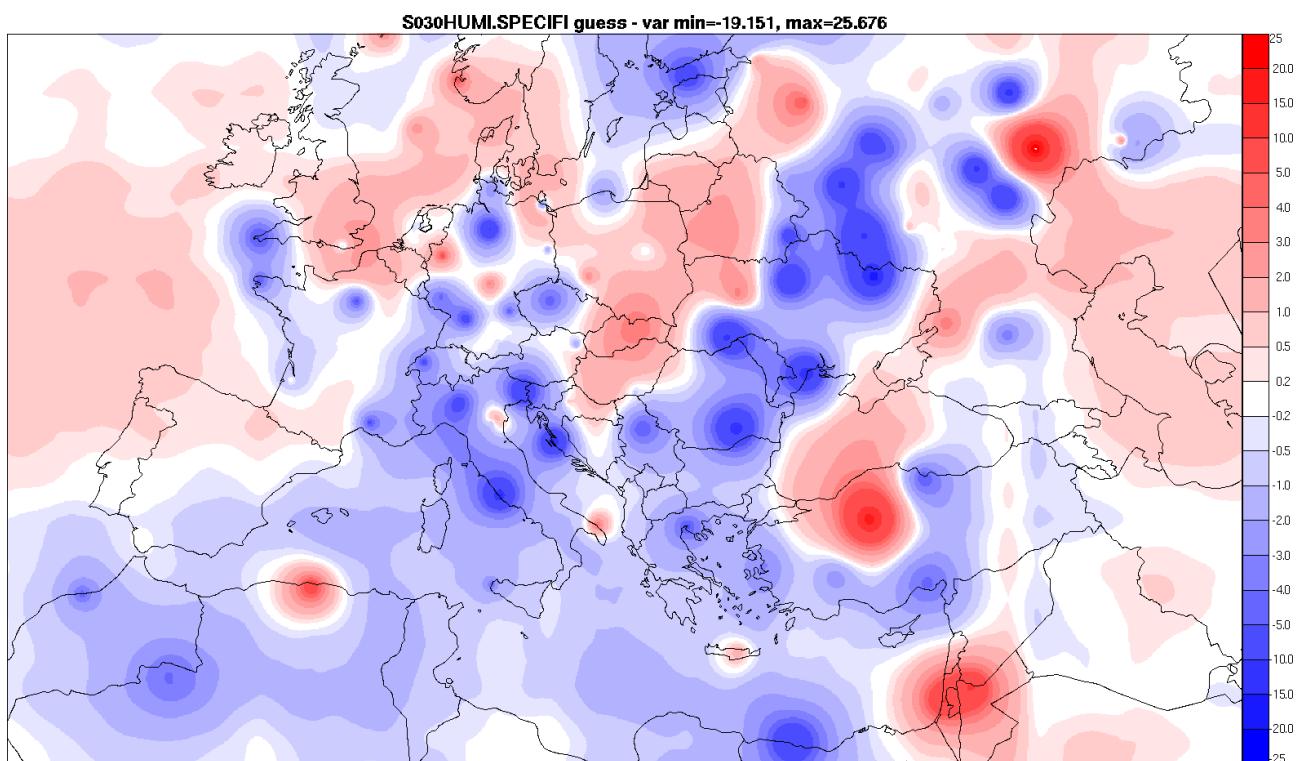
Obr. 6: Inkrementy špecifickej vlhkosti [ $\text{g kg}^{-1}$ ] na 50. modelovej hladine určené iba na základe asimilácie zenitových totálnych oneskorení GNSS.

Na prvý pohľad sú inkrementy špecifickej vlhkosti v centrálnej Európe opačného znamienka ako po asimilácii TEMP záznamov, ale zenitové totálne oneskorenia GNSS reprezentujú celkovú vzdušnú hmotu nad perma-nentnou stanicou, pričom TEMP dáta sú diskrétnie merania rádiosond v priestore. Tieto rádiosondy okrem

vertikálneho stúpania podliehajú aj horizontálnemu posunu, pričom dáta sú asimilované s predpokladom, že sonda sa pohybuje iba vertikálne. Zmeny inkrementov špecifickej vlhkosti na základe asimilácie TEMP dát v rôznych modelových hladinách sú zobrazené na obrázkoch 7 a 8.



Obr. 7: Inkrementy špecifickej vlhkosti [ $\text{g kg}^{-1}$ ] na 40. modelovej hladine určené iba z asimilácie TEMP dát.

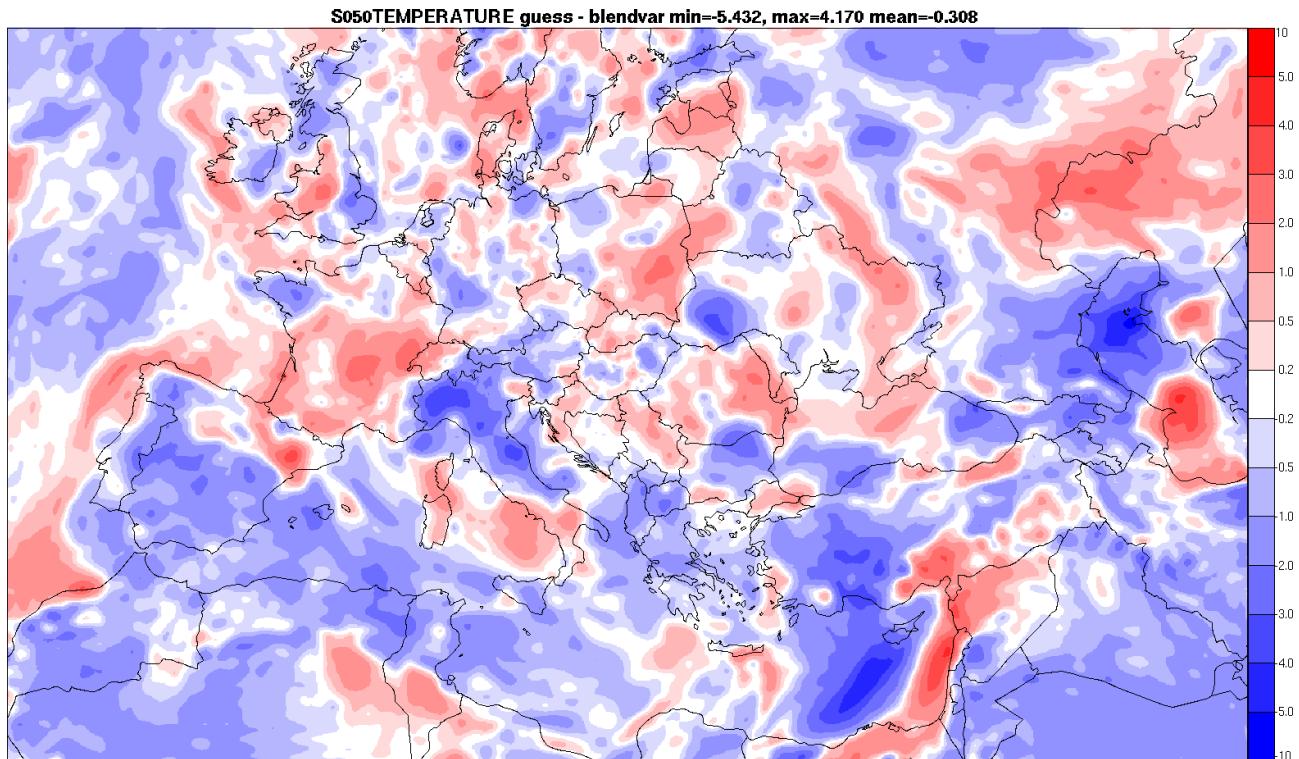


Obr. 8: Inkrementy špecifickej vlhkosti [ $\text{g kg}^{-1}$ ] na 30. modelovej hladine určené iba z asimilácie TEMP dát.

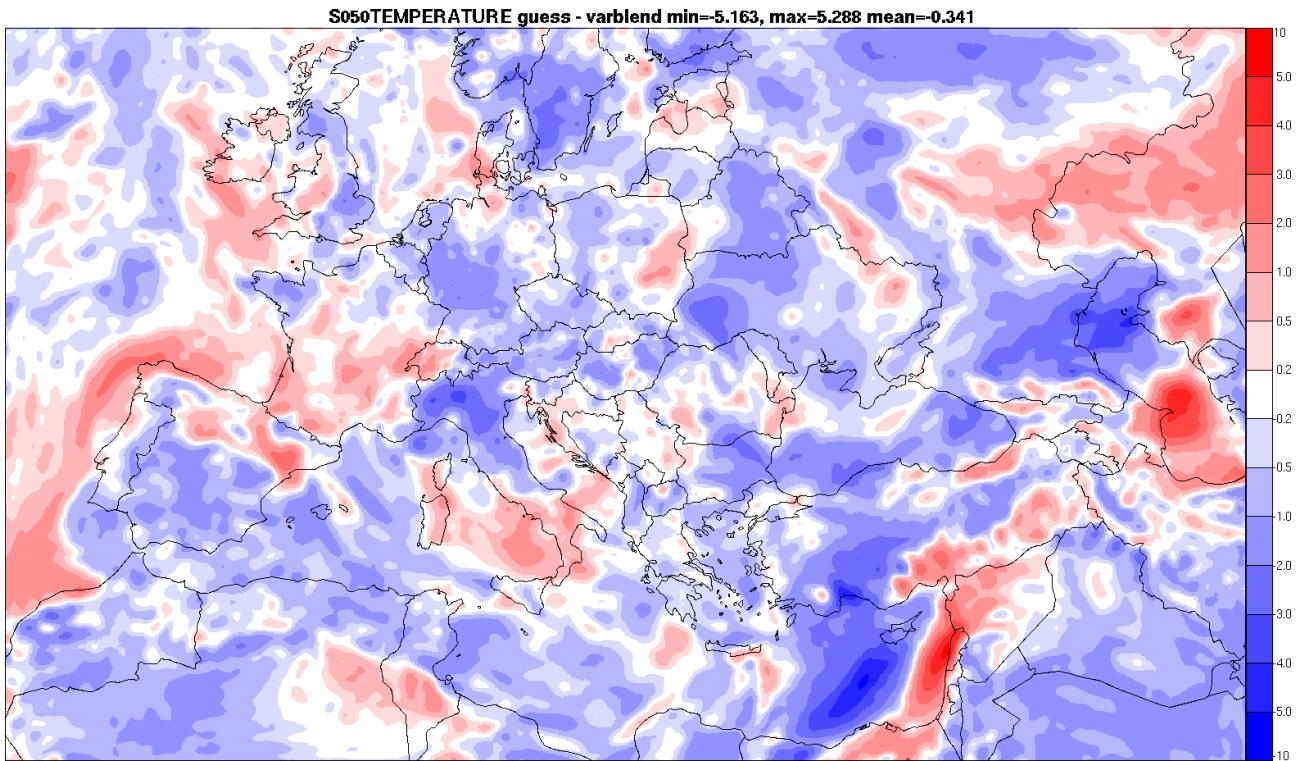
Dve stanice boli vyradené z pri zostavovaní whitelistu pre veľký rozdiel medzi modelovou a nadmorskou výškou stanice. Ďalšie dve stanice boli odstránené na základe Pearsonovho Chi-kvadrát testu, pretože first guess rozdiely nemali normálne rozdelenie pravdepodobnosti.

### 2.3 Implementácia 3DVar asimilácie do ALADIN–LAEF

Po technickej kontrole asimilácie dát sme testovali, či 3DVar asimilácia bude vhodnejšia pred alebo po Blendingu. Boli vykonané dva experimenty pre jeden termín s asimiláciou všetkých dát. Prvý experiment bol *blend-var* a druhý bol označovaný *var-blend*. V *blend-var*, ako z názvu vyplýva je 3DVar asimilácia po zmiešavaní dlhovlnnej zložky riadiaceho modelu s krátkovlnnou zložkou analýzy. V tomto prístupe je dominantnejší efekt asimilácie ako je zobrazené na obrázku 9. Na druhej strane vo *var-blend* (obrázok 10) sú gradienty v inkrementoch teploty na modelovej hladine 50 menšie. Modelové polia sú vyhladenejšie v dôsledku digitálneho filtrovania v Blendingu. Z porovnania stredných hodnôt je zrejmé že strejná hodnota teplotných inkrementov vo *var-blend* je menšia (čo znamená že analýza je horúcejšia) než analýza *blend-var* (rozdiel je  $0.033^{\circ}\text{C}$ ), no tento vplyv môže byť zanedbateľný.



Obr. 9: Inkrementy teploty [ $^{\circ}\text{C}$ ] na 50. modelovej hladine určené z asimilácie všetkých dát s postupnosťou krokov *blend-var*.



Obr. 10: Inkrementy teploty [ $^{\circ}\text{C}$ ] na 50. modelovej hladine určené z asimilácie všetkých dát s postupnosťou krokov *var-blend*.

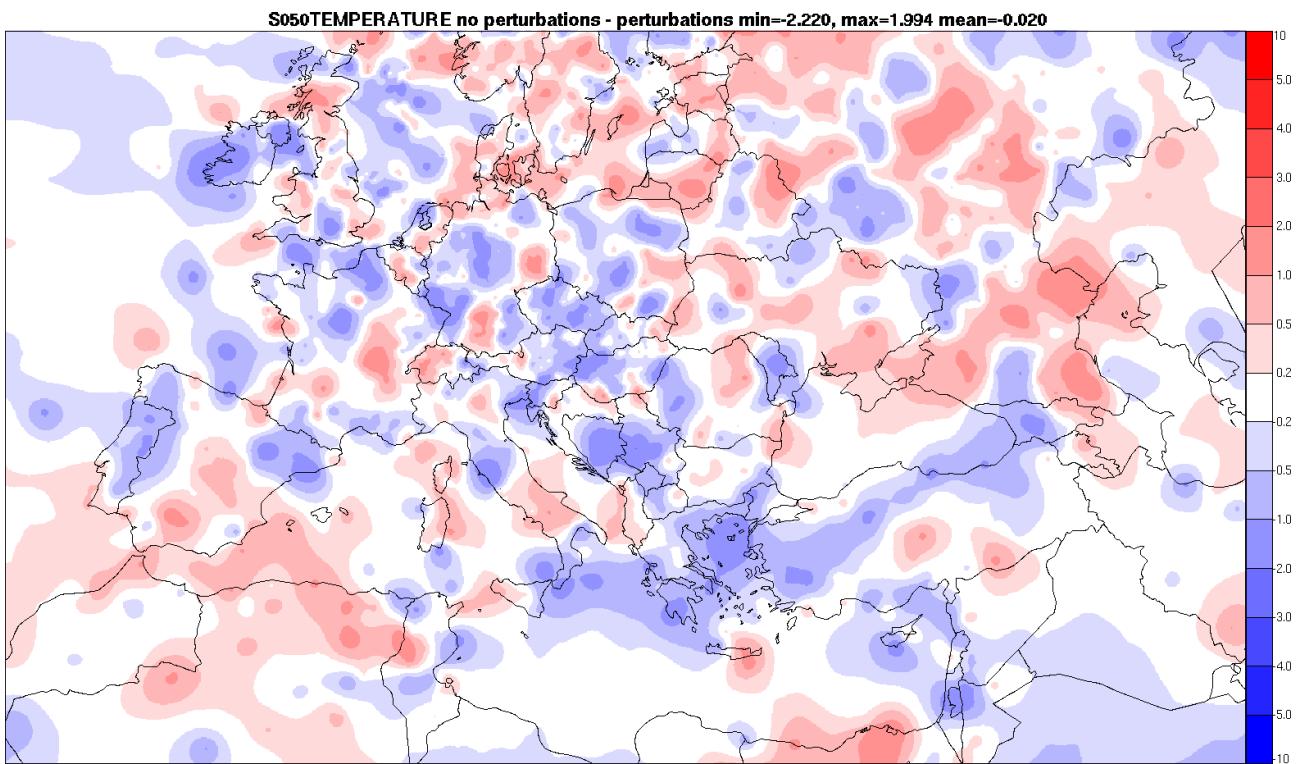
Bolo rozhodnuté, že implementácia 3DVar asimilácie bude pred Blendingom, čo znamená že kroky budú nasledovať v tomto poradí: ESDA, ENS 3DVar, Blending a Integrácia. Táto zvolená postupnosť je experimentálna a môže byť zmenená jednoduchým zásahom do systému spúšťania jednotlivých krokov. Dĺžka integrácie závisí od dennej doby, pre 00UTC bola integrácia predpovedí nastavená na 12 hodín iba pre potreby cyklenia. Pre čas 12UTC bola doba integrácie nastavená na 24 hodín z dôvodu cyklenia a validácie dopadu 3DVar asimilácie na predpovedné vlastnosti modelu.

## 2.4 Perturbácia asimilovaných dát

Ďalšou úlohou bola perturbácia všetkých asimilovaných dát v 3DVar metóde. Perturbácia bola vykonaná v kroku kontroly kvality (Screening) v samotnej 3DVar asimilácií dát pridaním troch položiek do namelistu v sekcií *&NAMSCC*:

- *LPERTURB=.T.*, táto položka umožňuje perturbáciu dát,
- *NAENSEMBLE=1*, táto položka umožňuje zmenu perturbácie v závislosti od čísla členu ansámlu,
- *NAEMEMBER={MEMB}*, táto položka nastavuje číslo členu ansámlu.

Perturbácia je vykonaná pridaním malej neistoty s normálnym rozdelením pravdepodobnosti každému asimilovanému údaju. Vplyv perturbácie na teplotu je uspokojivý (stredná hodnota je  $-0.02^{\circ}\text{C}$ ) ako je zobrazené na obrázku 11.



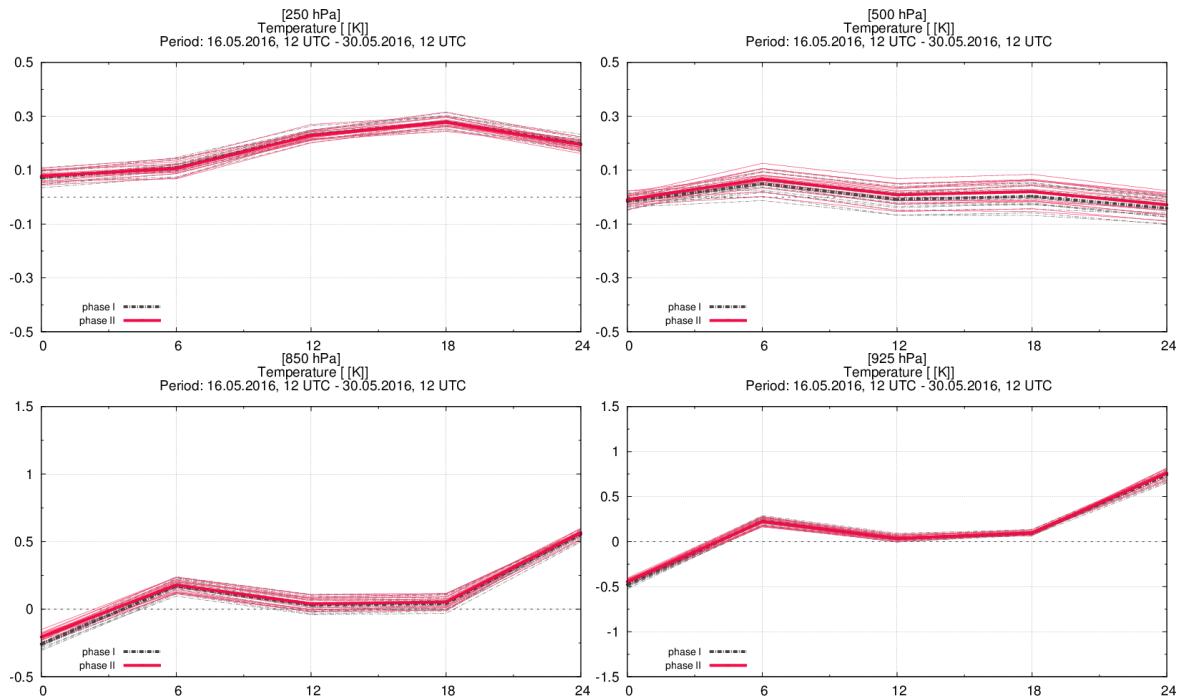
Obr. 11: Teplotné rozdiely [ $^{\circ}\text{C}$ ] v 50. modelovej hladine medzi analýzou všetkých dát s a **bez** perturbácie.

### 3 Dopad 3DVar asimilácie dát

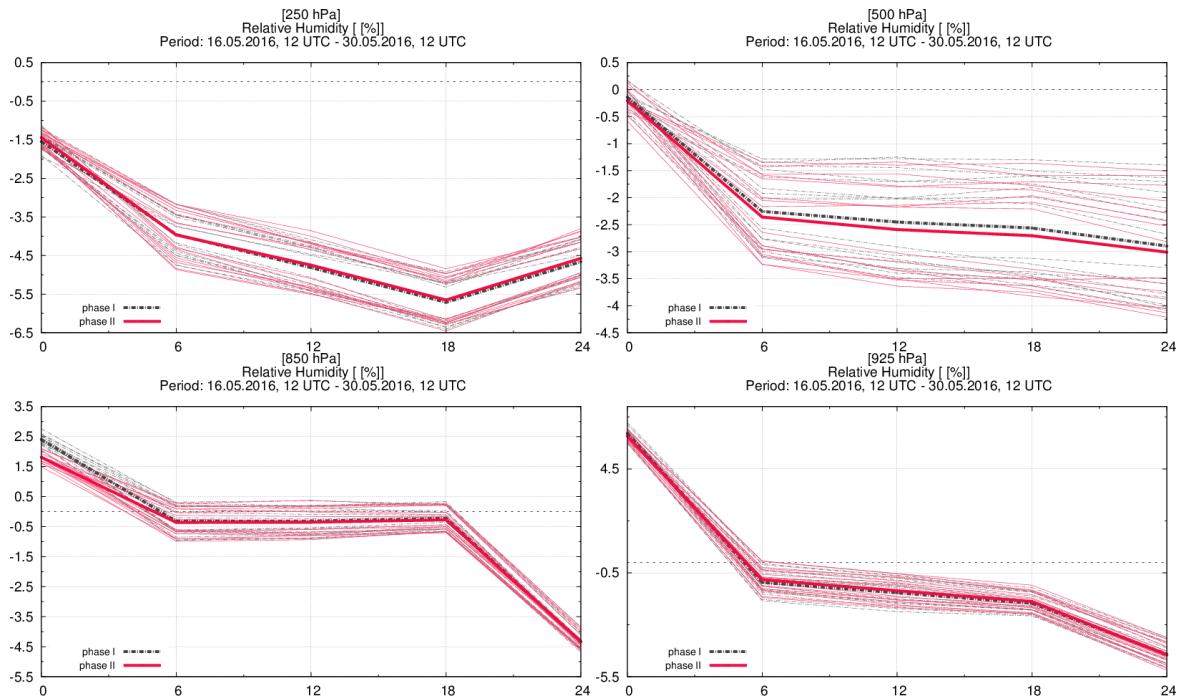
Po validácii perturbácie dát sme pristúpili ku kompletnému LAEF experimentu pre celé časové obdobie. To znamená 16 členov \* 16 dní \* 2 časy integrácie s nasledovnými krokmi:

- ESDA – perturbované SYNOP dát,
- ENS 3DVar – perturbované spomínané asimilované typy dát,
- atmosférický Blending – s ECMWF,
- Integrácia – 12 alebo 24 hodín.

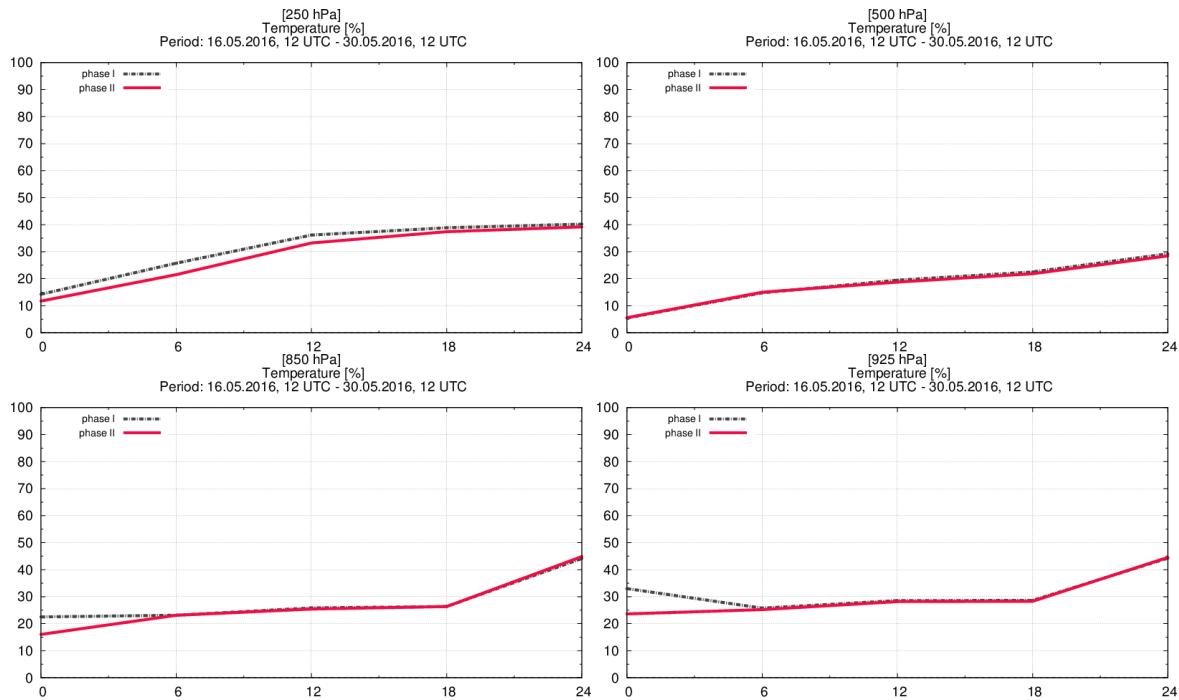
Tento experiment bol verifikovaný voči ECMWF reanalýze pre časové obdobie od 16.05.2016 12UTC do 30.05.2016 12UTC iba pre časy o 12UTC a 24 hodinovou predpoveďou. Tieto verifikačné skóre sú označené *phase II*. Skóre *phase I* boli spočítané z toho istého časového obdobia ale **bez** 3DVar asimilácie dát. Spočítané verifikačné parametre boli: chyba (bias), odľahlé hodnoty (outliers), RMSE a rozptyl (spread) pre tlakové hladiny 925, 850, 500 a 250 hPa pre polia teploty, relatívnej vlhkosti, geopotenciálu a rýchlosťi vetra. Signifikantné a zaujímavé výsledky sú zobrazené na nasledujúcich obrázkoch.



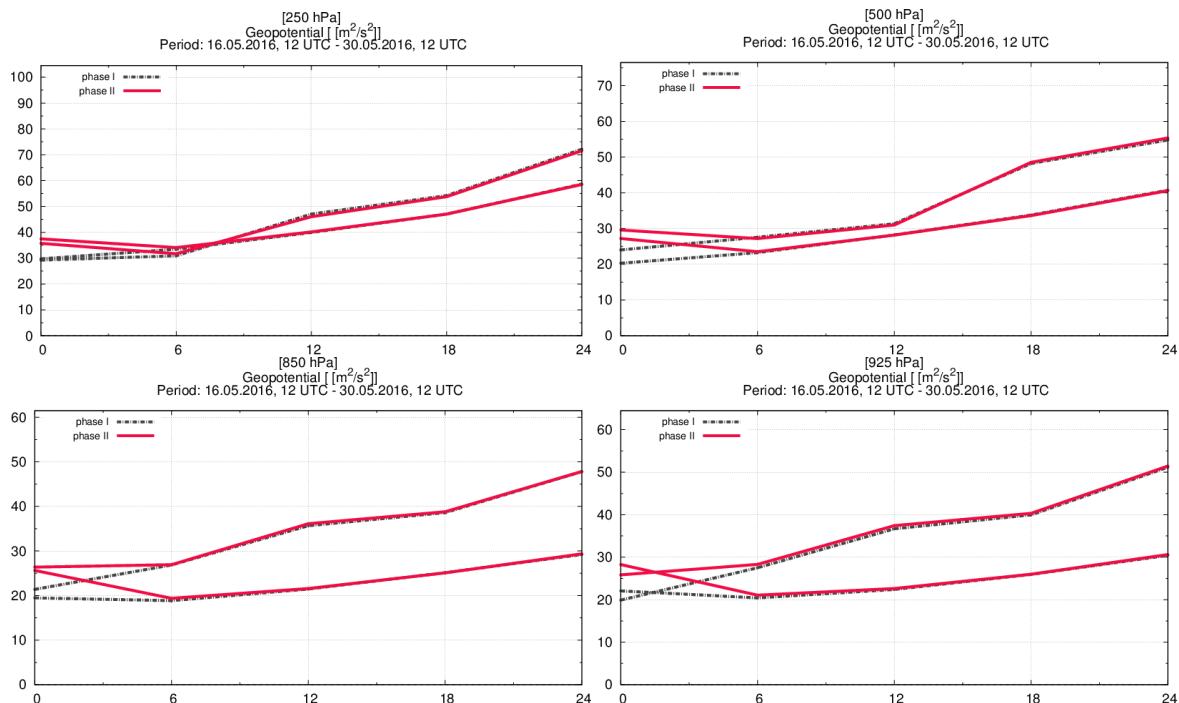
Obr. 12: Chyba teploty. Malé zlepšenie bolo zaznamenané na tlakovej hladine  $850 \text{ hPa}$  v čase analýzy (0. hodina predpovede).



Obr. 13: Chyba relatívnej vlhkosti. Malé zlepšenie bolo zaznamenané na tlakovej hladine  $850 \text{ hPa}$  v čase analýzy, pričom zhoršenie nastalo v 6,12,18 a 24 hodinovej predpovedi na hladine  $500 \text{ hPa}$ .

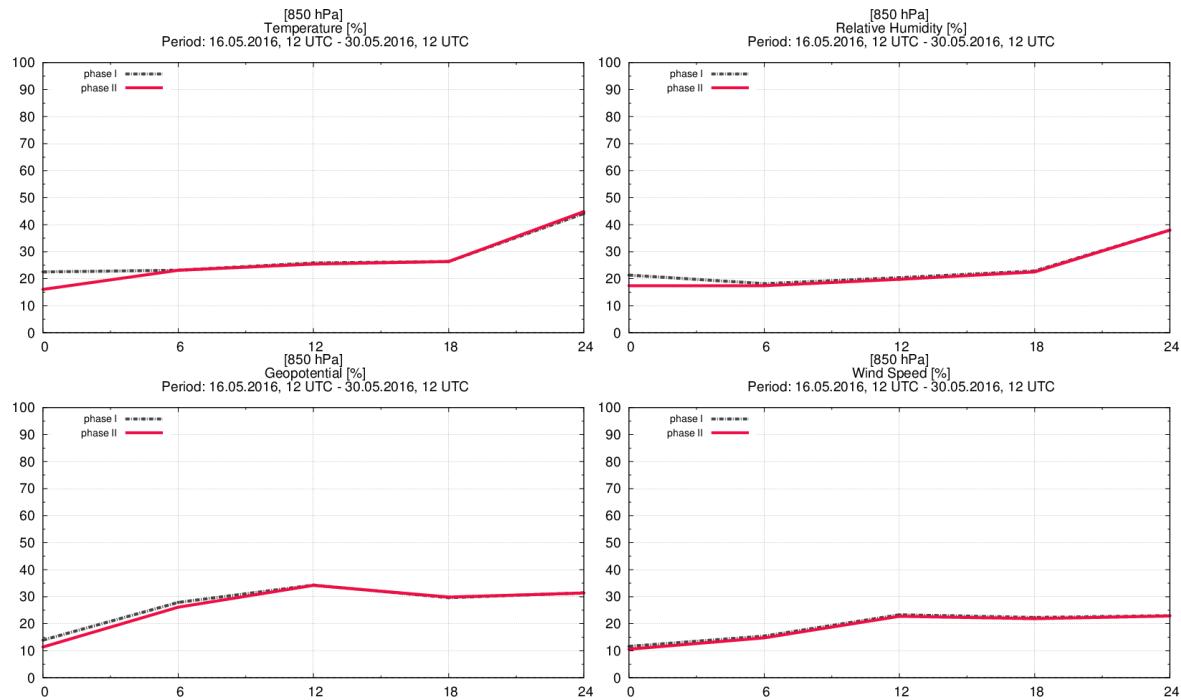


Obr. 14: Odľahlé hodnoty teploty. Všetky skúmané tlakové hladiny vykazujú zlepšenie (pokles) v počte odľahlých hodnôt v čase analýzy, podobné výsledky boli dosiahnuté pre relatívnu vlhkosť, geopotenciál a rýchlosť vetra (nie je zobrazené).

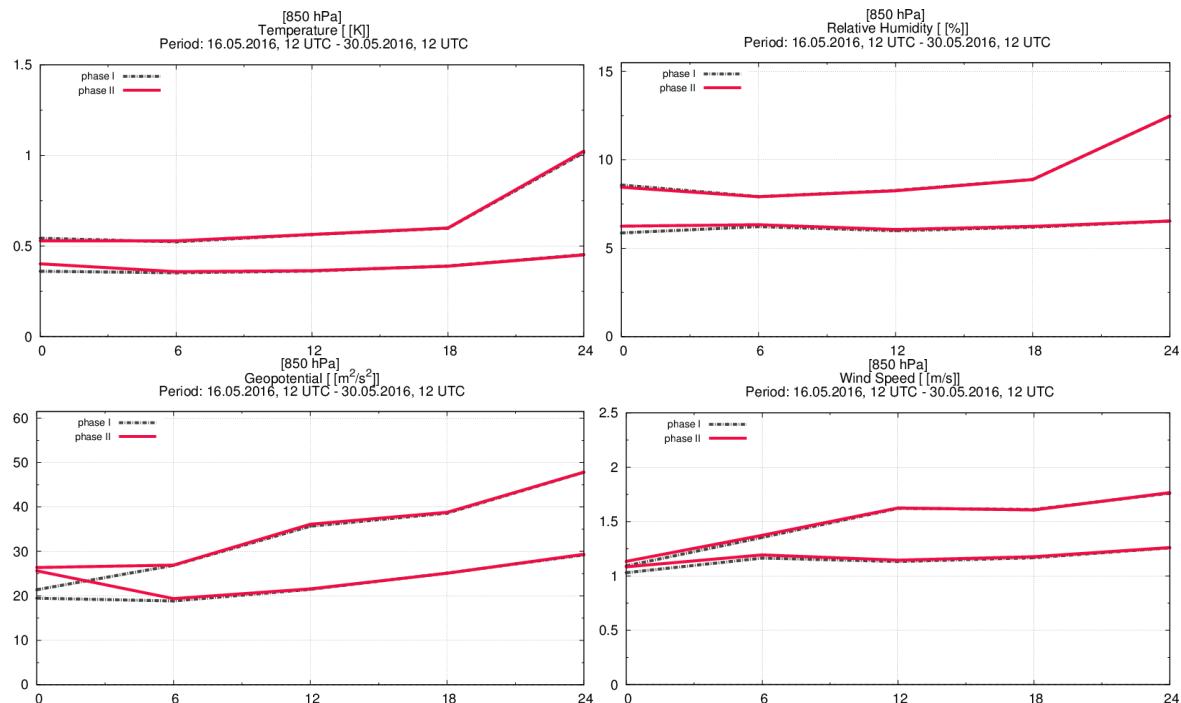


Obr. 15: RMSE a rozptyl geopotenciálu. Zhoršenie RMSE bolo zaznamenané v čase analýzy pre všetky tlakové hladiny (zvýšenie – vrchná čiara), ale zlepšenie rozptylu zároveň (zvýšenie – spodná čiara). Zhoršenie RMSE geopotenciálu je najväčší nedostatok 3DVar asimilácie vyplývajúci z verifikácie.

Na nasledujúcich obrázkoch sú zobrazené verifikačné parametre pre hladinu 850 hPa.



Obr. 16: Zniženie odľahlých hodnôt bolo dosiahnuté v čase analýzy na tlakovej hladine  $850 \text{ hPa}$  pre teplotu, relatívnu vlhkosť, geopotenciál a rýchlosť vetra.



Obr. 17: Vplyv 3DVar asimilácie na teplotu, relatívnu vlhkosť, geopotenciál a rýchlosť vetra pre tlakovú hladinu  $850 \text{ hPa}$ . RMSE (vrchná čiara) a rozptyl (spodná čiara). Vplyv 3DVar asimilácie je neutrálny (okrem spomínaného geopotencálu), rozptyl sa však zlepšíl (nárast).

Vplyv 3DVar asimilácie na ostatné hladiny bol neutrálny, takže nie sú zobrazené ďalšie verifikačné obrázky. Vo všeobecnosti je možné povedať, že 3DVar má malý pozitívny dopad vo všetkých tlakových hladinách na odľahlé hodnoty (pokles) a na rozptyl (nárast). Na druhej strane mierny nárast RMSE geoptenciálu bol zaznamenaný

na všetkých skúmaných tlakových hladinách.

## 4 Záver

Trojdimenzionálna variačná asimilačná metóda bola úspešne verifikovaná a implementovaná do existujúceho 5km LAEF systému. Technická správnosť bola potvrdená preskúmaním inkrementov teploty, špecifickej vlhkosti a rýchlosťi vetra v rôznych modelových hladinách. Asimilované boli vybrané dátové typy: SYNOP, AMDAR, TEMP (časť 2), GEOWIND (časť 2.1) a zenitové totálne oneskorenia GNSS (časť 2.2). ENS 3DVar bol zahrnutý do LAEF systému po ansámblovej asimilácii SYNOP dát do povrchových a podzemných parametrov modelu (ESDA) a pred Blendingom atmosféry, viac v časti 2.3. Taktiež bola úspešne implementovaná perturbácia do kontroly kvality asimilovaných dát, pozri časť 2.4. Po týchto krokoch nasledovala verifikácia dopadu ENS 3DVar asimilácie na 24 hodinové predpovede teploty, relatívnej vlhkosti, geopotenciálu a rýchlosťi vetra (časť 3). Dopad bol hodnotený na štyroch tlakových hladinách: 250, 500, 850 a 925 hPa. Chyba, odľahlé hodnoty, RMSE, a rozptyl boli určené voči reanalýze ECMWF pre ansámblový systém s (*phase II*) a bez (*phase I*) 3DVar asimilácie. Pre väčšinu prípadov bol dopad neutrálny, zaznamenali sme však zlepšenie počtu odľahlých hodnôt (pokles) a zlepšenie rozptylu (nárast). Pričom treba spomenúť aj zhoršenie RMSE geopotenciálu. Je možné, že zmena postupnosti krokov výpočtu analýzy (blend–var schéma) môže priniesť väčší dopad na 6 hodinovú predpoved'. Bolo by vhodné rozšíriť verifikáciu dopadu o prípadové štúdie, kde by sa mohla ukázať pridaná hodnota ENS 3DVar asimilačnej metódy.