IDŐJÁRÁS
Quarterly Journal of the Hungarian Meteorological Service
Vol. 119, No. 2, April – June, 2015, pp. 215–239

Overview of mesoscale data assimilation developments at the Hungarian Meteorological Service

Máté Mile[1*], **Gergely Bölöni**[1], **Roger Randriamampianina**[2], **Roland Steib**[1], and **Ersin Kucukkaraca**[3]

[1]*Hungarian Meteorological Service*
P.O. Box 38, H-1525 Budapest, Hungary

[2]*Norwegian Meteorological Institute*
P.O. Box 43, Blindern, N-0313 Oslo, Norway

[3]*Turkish State Meteorological Service*
P.O. Box 401, Ankara, Turkey

[*]*Corresponding authors E-mail: mile.m@met.hu*

(Manuscript received in final form June 03, 2014)

Abstract—The operational AROME (Applications of Research for Operations at MEsoscales) mesoscale numerical weather prediction (NWP) model has been run using interpolated analyses of the ALADIN (Aire Limitée Adaptation Dynamique Développement International) NWP model for its initialization since the end of 2010 at the Hungarian Meteorological Service (HMS). In order to improve the initial conditions, a local three-dimensional variational (3DVAR) data assimilation system was developed for the Hungarian version of AROME (AROME-Hungary). Regarding the data assimilation cycling strategy, it was shown that 3 hourly rapid update cycling (RUC), which was implemented operationally in March 2013 using conventional observations, outperforms 6 hourly cycling method. This paper describes at length the main characteristics of this local data assimilation system and its impact on the model short-range forecasts. Although the forecasts of AROME-Hungary based on a local data assimilation were already improved compared to the previous implementation (initialization via interpolated analyses of the ALADIN model), there is still a way to go to exploit the full benefit of the local 3DVAR assimilation cycle. Current development works aim at improving the system through exploitation of remote sensing observations (radar, GPS, satellite AMVs), with a special emphasis on humidity information. All tested observations showed promising performance on both the analyses and forecasts of the AROME-Hungary model, which should lead to their respective operational implementation in the near future.

Key-words: operational numerical weather prediction, mesoscale data assimilation, rapid update cycle, remote sensing observations

1. *Introduction*

State of the art mesoscale numerical weather prediction (NWP) models, such as ALADIN (*Horányi et al.,* 1996) and AROME (*Seity et al.,* 2010), describe the time evolution of small scale processes in the atmosphere (e.g., convection, sea breeze, fog), through the applied prognostic microphysical parametrization schemes and the non-hydrostatic dynamics. Advanced model dynamics and physics are, however, in vain, if the initial state does not contain appropriate information regarding the small-scale weather systems we aim to describe. The simplistic approach for the initialization of limited area models (LAMs) is to interpolate the analysis or the forecast of the driving model (i.e., a global model or another LAM) to the mesoscale grid. This approach, often referred to as spin-up initialization, is computationally cheap (i.e., there is no need to run expensive data assimilation schemes), nevertheless it implies several drawbacks, which will be demonstrated in this section. The sophisticated and scientifically sound alternative of the spin-up initialization is to run a local data assimilation system in the LAM, combining the high-resolution first guess of the mesoscale model with the available high-density observations.

Data assimilation is achieved by solving the BLUE (Best Linear Unbiased Estimation) analysis equation, which shows that the two main information used for estimating the initial state (x_a) are the background (x_b) and the actual observation set (y) (see e.g., *Kalnay*, 2003; *Lorenc*, 1986; *Evensen,* 2009):

$$x_a = x_b + K\big(y - H(x_b)\big) \tag{1}$$

In Eq.(1), K stands for the Kalman gain determining the weight of the background and the observations in each gridpoint in an optimal way, i.e., based on their reliability in a statistical sense. H denotes the observation operator, which enables the comparison of the observations with the background through a projection from the model space to the observation space.

In practice local data assimilation is substantially more expensive to implement than spin-up initialization (both computationally and regarding manpower), but in turn, it enables a more accurate representation of small-scale phenomena in the initial state. In order to demonstrate this, the kinetic energy spectra for a spin-up analysis (ALADIN analysis interpolated to the AROME-Hungary grid) and the local 3DVAR analysis of AROME-Hungary are plotted in *Fig. 1*. It can be seen that the energy spectra for the AROME analysis follows rather close the theoretical slope of energy cascade, i.e., k^{-3} for small k wavenumbers (large scales) and $k^{-5/3}$ for high k wavenumbers (small scales). In contrast, the energy spectra of the spin-up analysis is far from the theoretical slopes, especially at mesoscales (above wavenumber 50, which corresponds to spatial scales smaller than 25 km), where the energy curve is rather noisy. The mesoscale noise in the spin-up analysis is introduced by the interpolation, and it reflects that the ALADIN analysis with its 8 km grid-length (30–45 km effective resolution) do not hold physical information on the mesoscales resolved by the AROME model with its 2.5 km grid-length. Among others, this diagnostic comparison gave a great motivation for implementing a local 3DVAR data assimilation system for the AROME-Hungary, which finally became operational in March 2013 due to approved ability to improve the forecast performance compared to the spin-up initialization approach. Similar positive impact of mesoscale data assimilation implementations have been reported by *Benjamin et al.,* (2004), *Fischer et al.,* (2005), *Bölöni* (2006), *Randriamampianina* (2006a), *Brousseau et al.* (2011).

The structure of the article is as follows: In the next section, the operational DA system will be described, with special emphasis on rapid update cycling, and also, the added value and the impact of the AROME-Hungary data assimilation system will be detailed. In Section 3, we briefly review experimental data assimilation studies with non-conventional observations added to the existing operational system. Finally, the last section gives a summary of the presented work and provides corresponding conclusions.

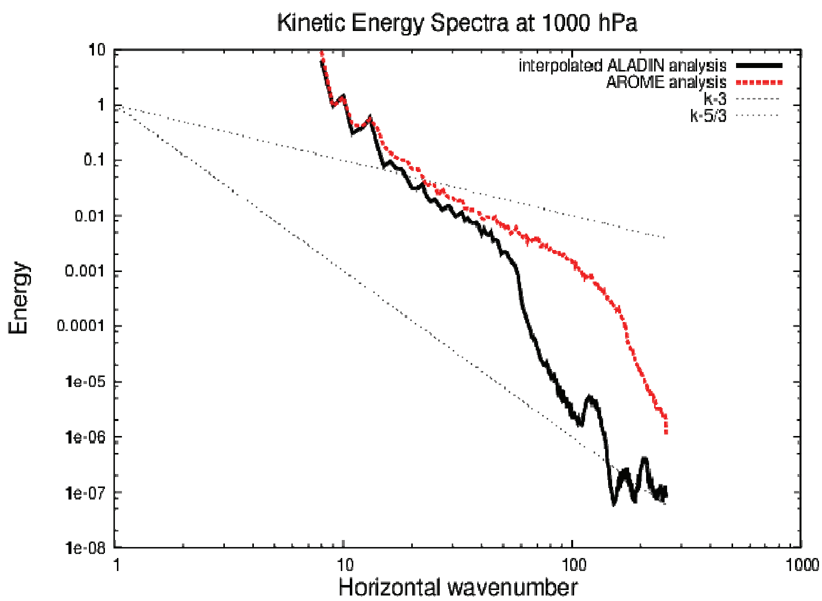

Fig. 1. Wind energy spectra (kg m^2 s^{-2}) at 1000 hPa for the AROME (dx = 2.5 km) analysis (red) and for an interpolated ALADIN (dx = 8 km) analysis (black). Dotted straight lines correspond to the theoretical slopes of kinetic energy at large- (k^{-3}) and mesoscales ($k^{-5/3}$) if using a log-log scale.

2. ***Operational data assimilation system of the AROME mesoscale model***

2.1. The assimilation system

The local data assimilation system of AROME-Hungary is based on a 3 hourly rapid update cycle (RUC). The organization of one particular assimilation step at 00 UTC is shown schematically in *Fig. 2*, where the lateral boundary conditions in the assimilation cycle are provided by the global ECMWF/IFS (European Centre for Medium-Range Weather Forecasts / Integrated Forecast System) model, and surface parameters are initialized either using the surface analysis of the operational ALADIN model where available (at synoptic times, i.e., 00, 06, 12, and 18 UTC) or using the previous AROME/SURFEX (SURFace Externalized) forecast (at sub-synoptic times i.e., 03, 09, 15, and 21 UTC).

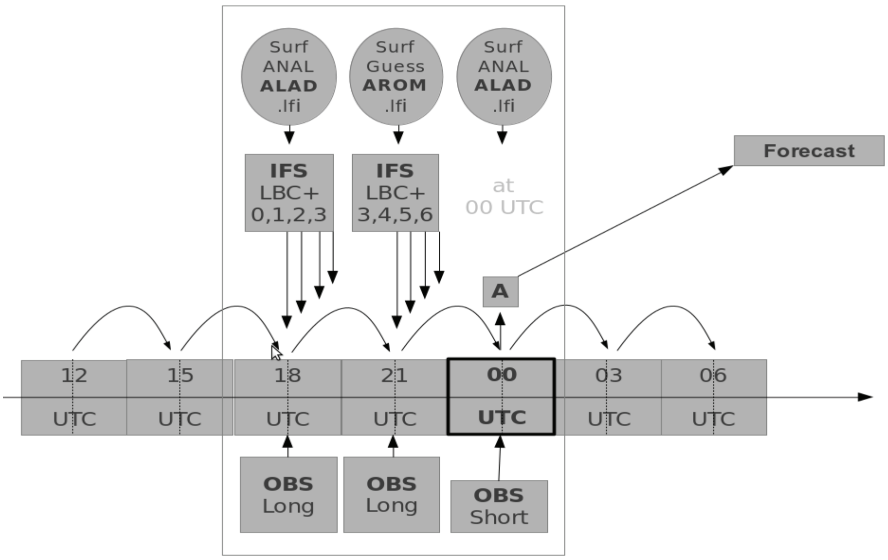

Fig. 2. Schematic figure of the data assimilation cycle applied for AROME showing the elements of a 00 UTC assimilation run.

The core of AROME-Hungary data assimilation system is an incremental 3DVAR method, where the basic mathematical formulation and its corresponding implementation is very similar to the one used in the IFS, ARPEGE, and ALADIN models (*Courtier et al.,* 1998; *Fischer et al.,* 2005; *Vasiliu* and *Horányi*, 2005; *Bölöni* 2006). An important component of the 3DVAR is the representation of background error statistics which plays a key role in filtering the information coming from observations and spreading it out to the model grid (see e.g., *Berre* 2000; *Brousseau et al.*, 2011). In the current operational system, the background error covariance matrix was sampled from the downscaling of an ALADIN Ensemble Data Assimilation (EDA) experiment, run for a summer period using 3-hour forecast ranges and 5 EDA members to get sufficient statistical sample (*Bölöni et al.,* 2014). Input observations for the AROME-Hungary 3DVAR suite are provided by the OPLACE (Observation Preprocessing for LACE (Limited Area modeling for Central Europe)) system, which includes both conventional and non-conventional observations. Although the currently operational AROME-Hungary data assimilation system uses only conventional observations, the system is able to assimilate non-conventional observations, as described in Section 3.

Delivering the mesoscale NWP forecasts as early as possible is of high priority for every operational forecasting centre. In order to find the optimal observation cut-off time (waiting time after the nominal analysis time) of the operational AROME-Hungary data assimilation system, the timeliness of the incoming observations was diagnosed for the area of interest. In *Fig. 3*, the availability of conventional observations at 00 UTC is shown, based on the amount of data received by the OPLACE system. After gaining this result, the observation (short) cut-off time has been set to one hour, which brings the fastest possible production of AROME forecast with an almost complete input observation set. This rather early cut-off enables a faster AROME forecast production compared to the former operational version of AROME-Hungary, which was based on a spin-up initialization from the 3DVAR analysis of the ALADIN model, because the ALADIN data assimilation system uses longer cut-off.

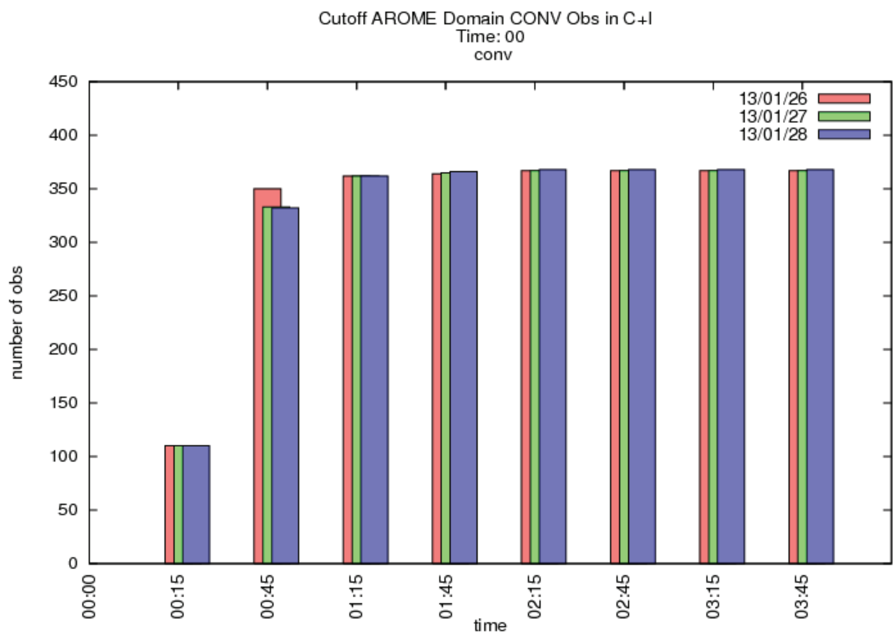

Fig. 3. Estimation of the optimal short cut-off time (considering AROME integration domain at 00 UTC) on January 26 (red columns), 27 (green columns) and 28 (blue columns) 2013.

2.2. The Rapid Update Cycle

In mesoscale NWP models, the accuracy of initial conditions is getting more and more crucial with an increasing resolution, hence small scale processes of the atmosphere have less and less predictability (*Fabry* and *Sun*, 2009). The rapid update cycling approach with increased analysis frequency in the assimilation cycle aims to involve more observations with reduced representativity error in time. It is assumed that the more observations we consider for the update of the model state, the better initial conditions we will get when starting a forecast from an analysis of the assimilation cycle. Considering conventional observations for instance, the different amount of used observations between a 6 hourly data assimilation cycle and a 3 hourly RUC is plotted in *Fig. 4* for a short period. At sub-synoptic times, mainly aircraft and SYNOP reports provide almost a double amount of data per day, due to the 4 extra analyses. To further emphasize the benefits of a RUC, it should be mentioned that many remote sensing observations are available with high temporal frequency, which are potentially beneficial in a data assimilation system with high frequency cycling. Also, many of the remote sensing observations are available in a very timely manner, almost immediately after analysis time, which allows to keep the operational observation cut-off time rather short, and thus to provide the forecasts with an early delivery.

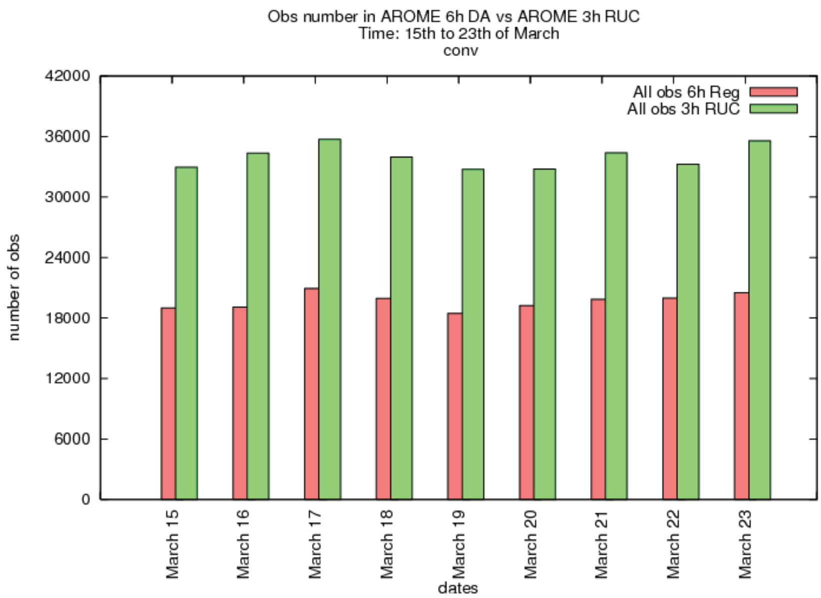

Fig. 4. The number of conventional observations used in a day in the RUC implementation (green) and in a 6 hourly cycling (red). The amount of data was counted in the period of March 15-23, 2014.

Beside these attractive features of the RUC system, there are some issues which have to be treated carefully in case of using a frequent assimilation cycle, and thus, shorter background forecast lengths. The forecast model integration may imply spin-up effects (noise due to spurious gravity waves or imbalances between dynamics, atmospheric and soil physics) at the very short ranges, which can be accumulated in the assimilation cycle through the background forecasts, and thus, lead to degradation in the analyses and the forecasts. A usual practice in NWP, and in the Hungarian version of the ALADIN model as well, is the use of initialization techniques e.g., digital filter initialization (DFI) (*Daley*, 1991; *Lynch et al*, 1997), which is a low-pass spectral filter removing high-frequency components of the initial conditions. In case of AROME-Hungary, no DFI is applied because such filtering is assumed to be too strong in case of a mesoscale non-hydrostatic model, where gravity waves are described by the dynamics. To diagnose spin-up effects in AROME-Hungary, surface pressure tendencies have been examined for very short-range forecasts (+2 hours) for an arbitrarily chosen case. In *Fig. 5*, time evolution of the surface pressure tendency provided by three different forecasts is shown for a particular gridpoint over orography. The red curve corresponds to a forecast, which was started from an AROME 3DVAR analysis, using a time-consistent coupling scheme, i.e., when the lateral boundary condition (LBC) at initial time is the interpolated global IFS forecast. The blue curve stands for a similar run with the only difference of using a space-consistent coupling scheme, i.e., when the lateral boundary condition (LBC) at initial time is the AROME 3DVAR analysis itself. As an additional reference, the tendency from an AROME forecast using a spin-up initialization is added (green curve), i.e., where both the initial condition and the LBC at initial time is the interpolated global IFS forecasts. As the high amplitude oscillation in the time evolution of surface pressure tendency is an indicator of noise, it has been concluded based on *Fig. 5*, that AROME forecasts using a RUC assimilation with a space-consistent coupling scheme imply less noise than a RUC with a time-consistent coupling approach or the spin-up initialization. Supposedly, the higher amplitude oscillation in case of the spin-up initialization is caused by the interpolation noise which is more emphasized over orography. It should be also mentioned that plotting the evolution of surface pressure tendency on a horizontal map (not shown) supports the choice for the space-consistent coupling scheme. Namely, in the time-consistent case, noise patterns (indicated by large tendencies) penetrate from the domain borders towards the middle of the domain in a rectangular shape. In *Fig. 5*, this is captured by the red curve at integration steps 40–45 as an outstanding wave. An explanation for this structured noise might be that imbalances between the local 3DVAR analysis and the LBC at initial time arise, due to the inconsistent model states of the AROME and IFS models near the boundaries. The final decision on implementing a 3 hourly RUC instead of a traditional 6 hourly cycling was based on a comparison study over a 1 month period during summer 2012, where the skill of these two cycling options was measured. Verification results – which will be studied in the next section – reflected a better performance of the 3 hourly RUC, leading to its operational implementation.

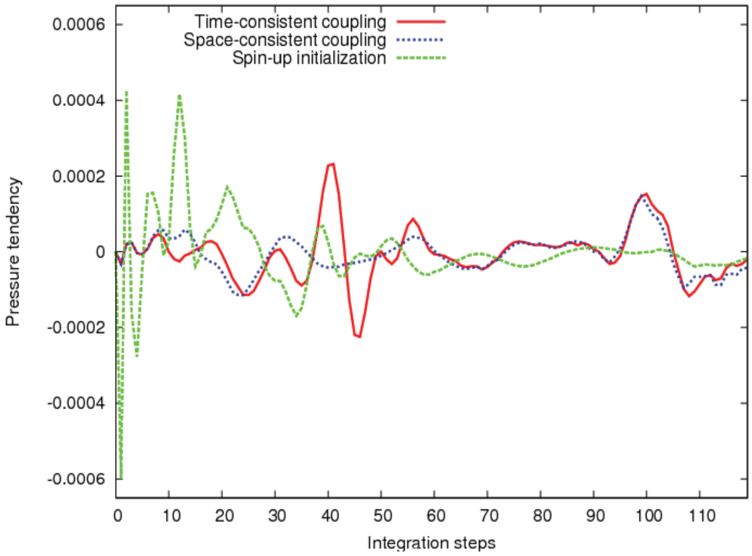

Fig. 5. Temporal evolution of surface pressure tendency (Pa/min) over orography during the first 2 hours of a forecast concerning spin-up initialization (green dashed line), time-consistent coupling approach (red line), and space-consistent coupling approach (blue dashed line).

2.3. The impact of local data assimilation scheme on the analysis and forecast

For measuring the impact of the local data assimilation scheme, three experiments with AROME-Hungary have been run and compared for several periods based on objective verification scores primarily. The three experiments are an AROME suite based on spin-up initialization (called DYNA), and two AROME suites based on local 3DVAR data assimilation, one of them using a 6 hourly cycling (called CONV6H) and another one using a 3 hourly rapid update cycling (called CONV). Forecasts have been run up to +36 hours starting at 00 UTC network times. Concerning the verification, both point-based and object-based scores have been computed. In the point-based verification, surface and radiosonde observations have been used as reference. The applied object-based method calculates average precipitation of forecasted weather objects (or alternatively the full domain average), and compares with calibrated radar precipitation measurements as reference. In *Figs. 6a* and *6b*, 10 m wind and mean sea level pressure scores are shown for the period between June 25 and July 25, 2012. It is rather clear from these figures that both the 6 hourly and the 3 hourly RUC 3DVAR provide an added value compared to the spin-up initialization with respect to both RMSE and BIAS. Moreover, the 3 hourly RUC provides slightly better scores than the 6 hourly cycling during daytime, while during the night, scores show a rather low sensitivity to the cycling frequency. To demonstrate the impact on precipitation, object-based verification score (domain average precipitation) is shown in *Fig. 7*. It can be seen, that in comparison with the spin-up initialization, both the 6 hourly 3DVAR cycling and the 3 hourly RUC could reduce the overestimation of the precipitation maximum linked to convective activity in the afternoon. It is also clear from *Fig. 7*, that the 3 hourly RUC provides slightly better precipitation forecasts than the 6 hourly cycling, by reducing the overestimation further.

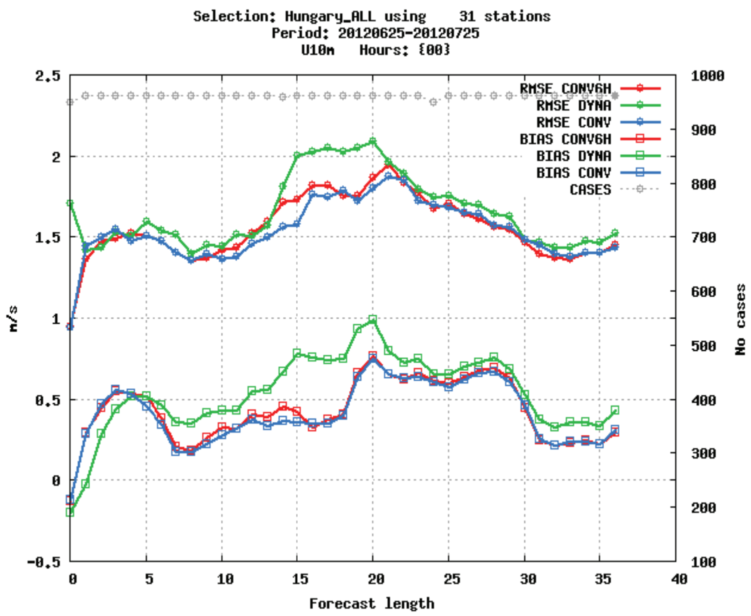

Fig. 6a. RMSE and BIAS scores corresponding to the spin-up initialization scheme (green), 3 hourly RUC (blue), and 6 hourly cycling (red) for 10 m wind speed (m/s).

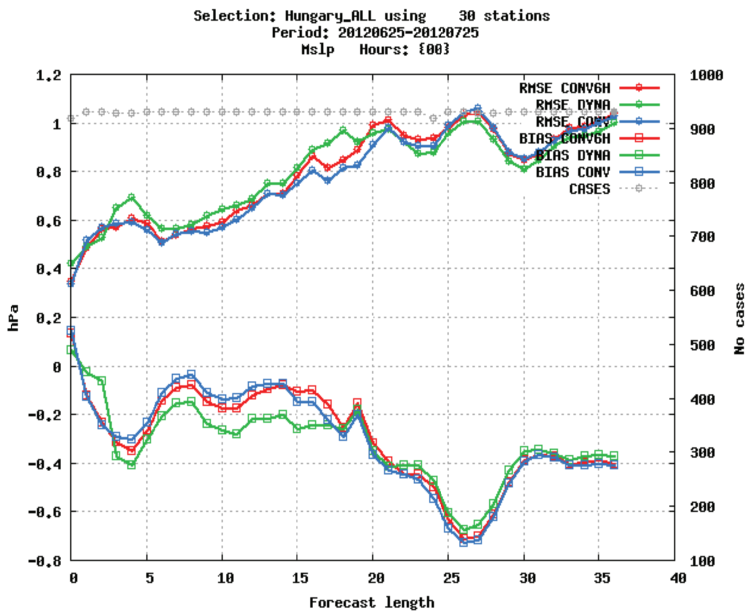

Fig. 6b. RMSE and BIAS scores corresponding to the spin-up initialization scheme (green), 3 hourly RUC (blue), and 6 hourly cycling (red) for mean sea level pressure (hPa).

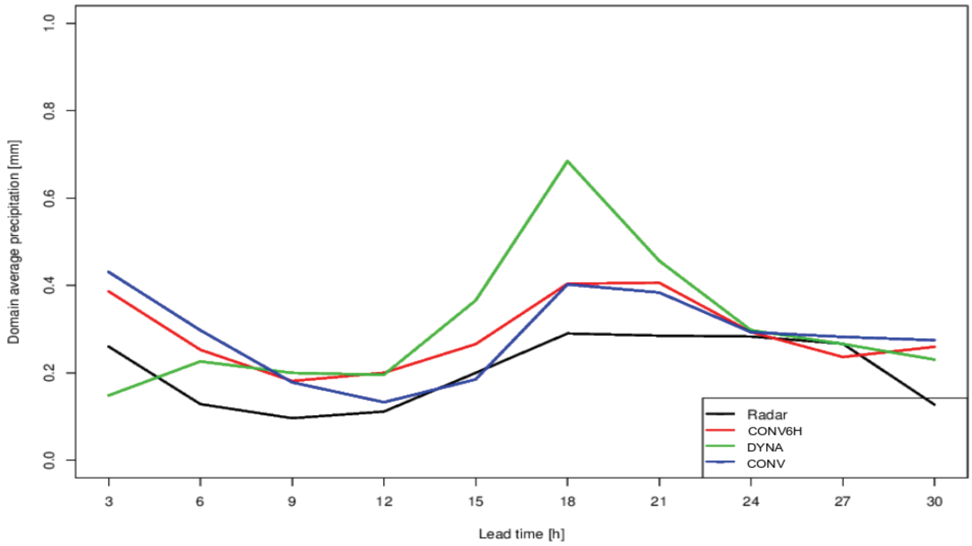

Fig. 7. Comparison of 3h RUC (blue), 6 hourly cycling (red) and the spin-up initialization scheme (green) regarding domain average precipitation (mm/h). The reference is the calibrated radar measurement (black).

Based on the success of the RUC approach (*Figs. 6a, 6b,* and *7*), a parallel suite with a 3 hourly RUC data assimilation system was compared with the former operational AROME-Hungary system (based on spin-up initialization) over the period from February 20 to March 12, 2013. This parallel suite was set up in a fully operational environment, providing real time outputs for the forecasters of the Hungarian Meteorological Service (this time both for 00 and 12 UTC base times), with the main aim to make a final decision about the operational implementation of the RUC system, in case of preferable scores and positive feedbacks from the forecasters. Based on the verification results, the RUC system clearly outperformed the forecasts of the former operational AROME-Hungary suite with spin-up initialization, which is demonstrated in *Figs. 8a, 8b* for 2 m temperature and in *Fig. 9* for precipitation. Feedbacks of the forecasters also confirmed the slight but consistent improvements implied by the RUC system, and this led to its operational implementation on March 17, 2013.

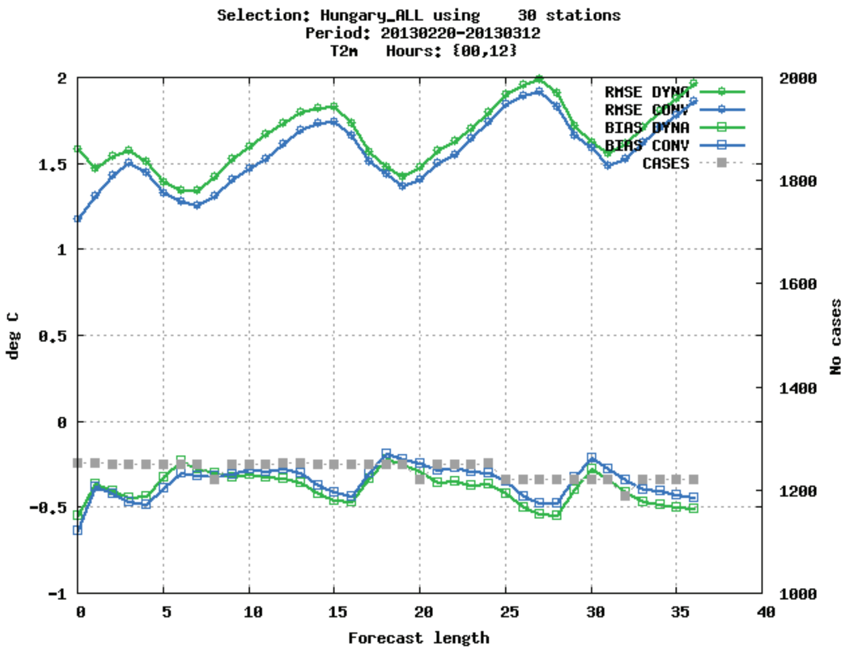

Fig. 8a. RMSE and BIAS scores of AROME forecasts according to the spin-up initialization (DYNA – green), and the 3 hourly rapid update cycle (CONV – blue) for 2 m temperature (^{o}C).

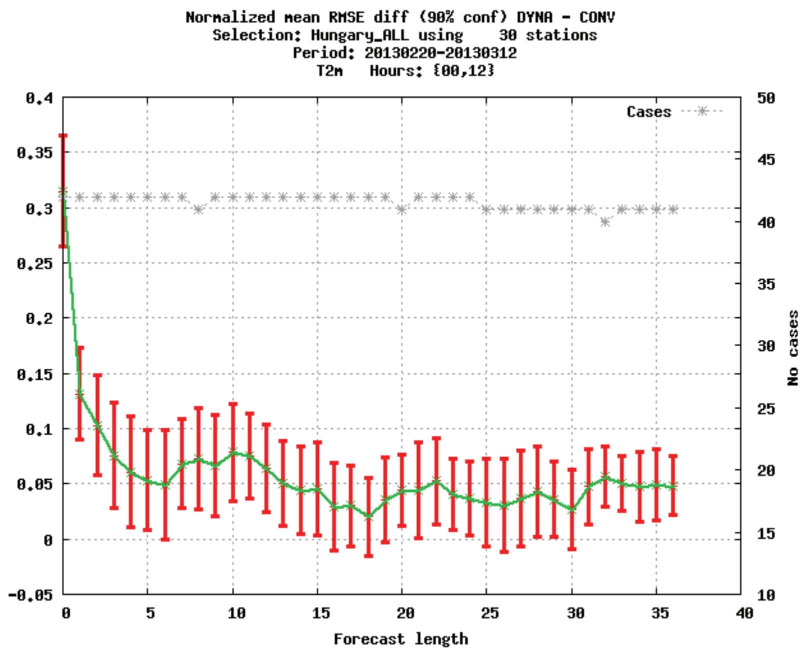

Fig. 8b. Normalized RMSE differences according to the spin-up initialization (DYNA – green), and the 3 hourly rapid update cycle (CONV – blue) for 2 m temperature (^{o}C).

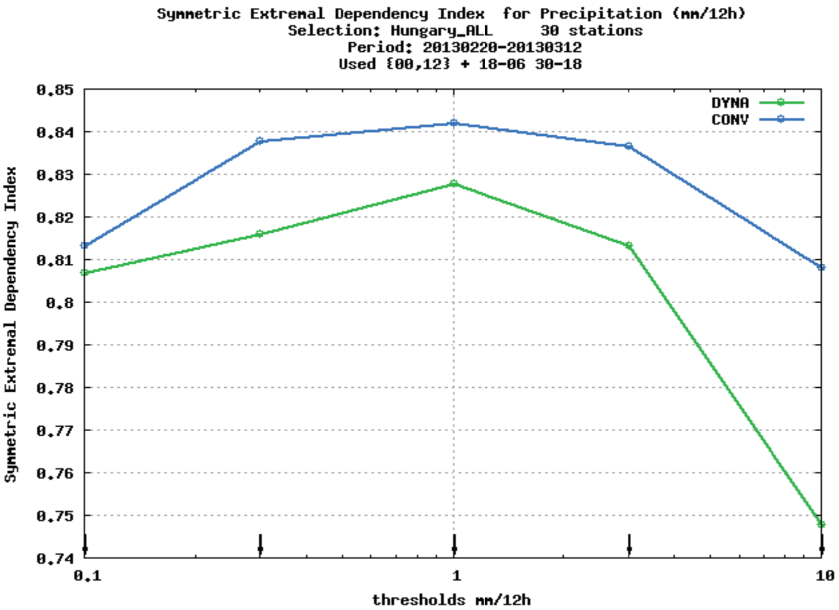

Fig. 9. Symmetric Extremal Dependency Index (SEDI) for precipitation according to the spin-up initialization (DYNA – green), and the 3 hourly rapid update cycle (CONV – blue) for precipitation (mm/12h).

2.4. *Diagnosing analysis sensitivity to observations*

An obvious way for the further development of the operational RUC is to bring non-conventional observations to its analysis system. In order to figure out, which observations would contribute the most to the analysis, the DFS (Degrees of Freedom for Signal) diagnostic tool (*Chapnik et al.,* 2006; *Cardinali et al.*, 2004) has been adapted and applied at the Hungarian Meteorological Service. The DFS tool diagnoses the observation influence on the analysis, thus, if applied for the available observation types, it provides an indication on their relative contribution. The DFS diagnostic is computed as the trace of the Kalman gain matrix projected to observation space:

$$DFS = Tr(HK) \tag{2}$$

where K and H denote respectively the Kalman gain matrix and the observation operator introduced in Eq. (1). In practice, this trace cannot be computed, because the gain matrix K is not known explicitly. *Girard* (1987) suggested a solution which enables to evaluate the above mentioned trace with the following approximation:

$$Tr(HK) \approx (y' - y)^T R^{-1}\left(H(x_a') - H(x_a)\right) \tag{3}$$

where $H(x_a)$ gives the analysis state at observation locations using a background and the observations (y), and $H(x'_a)$ stands similarly for the analysis at observation space but using the perturbed observations (y') (and the same background). In Eq. (3), R stands for the observation error covariance matrix.

Therefore, DFS can be calculated through a random perturbation of the observations, and the method is flexible in a sense that DFS values can be computed for any subset of the available observational data. For a given date, the DFS was computed with conventional and some experimentally used non-conventional observations. In order to verify the influence of the available observations, both the absolute and the relative DFS diagnostics were computed (*Fig. 10*). Absolute DFS stands for what have been explained above in Eqs. (2) and (3), while relative DFS is the absolute DFS normalized by the number of observations within the given observation subset. The first conclusion based on the absolute DFS is that the largest contribution to the analysis is provided by wind observations, i.e., the largest amount of information is extracted from these observations in the current data assimilation system. On the other hand, relative DFS reflects the importance of humidity (from surface SYNOP stations and TEMP radiosondes), RADAR reflectivity (RADAR-Z), and GNSS (global navigation satellite system) ZTD (zenith total delay) observations. In conclusion, DFS diagnostics show that radar reflectivity and GNSS ZTD observations are promising candidates for assimilation in the future version of the RUC system. It should be mentioned that DFS provides a theoretical measure of the information content projected from the observations to the analysis and it does not provide any indications on the impact attributed to the forecasts. Another point to be added here is that no radiance observations have been considered and diagnosed in the recent RUC system.

3. *Use of non-conventional observations in the AROME 3DVAR system*

In the following section, we present the latest developments of the RUC system since its operational implementation. The need of using more observations, especially non-conventional ones, was already mentioned. Specifically, the observations measuring humidity are potentially good candidates based on the results of the DFS analysis sensitivity study shown in the previous section. At the same time, to gain advantage of the RUC system, early accessible observations with high frequency and high density are also required. Taken into account these objectives, the atmospheric motion vectors (AMV) from Meteosat Second Generation (MSG) geostationary satellite, RADAR measurements as reflectivity and radial wind, and GNSS ZTD observations have been investigated in the operational data assimilation system of AROME-Hungary.

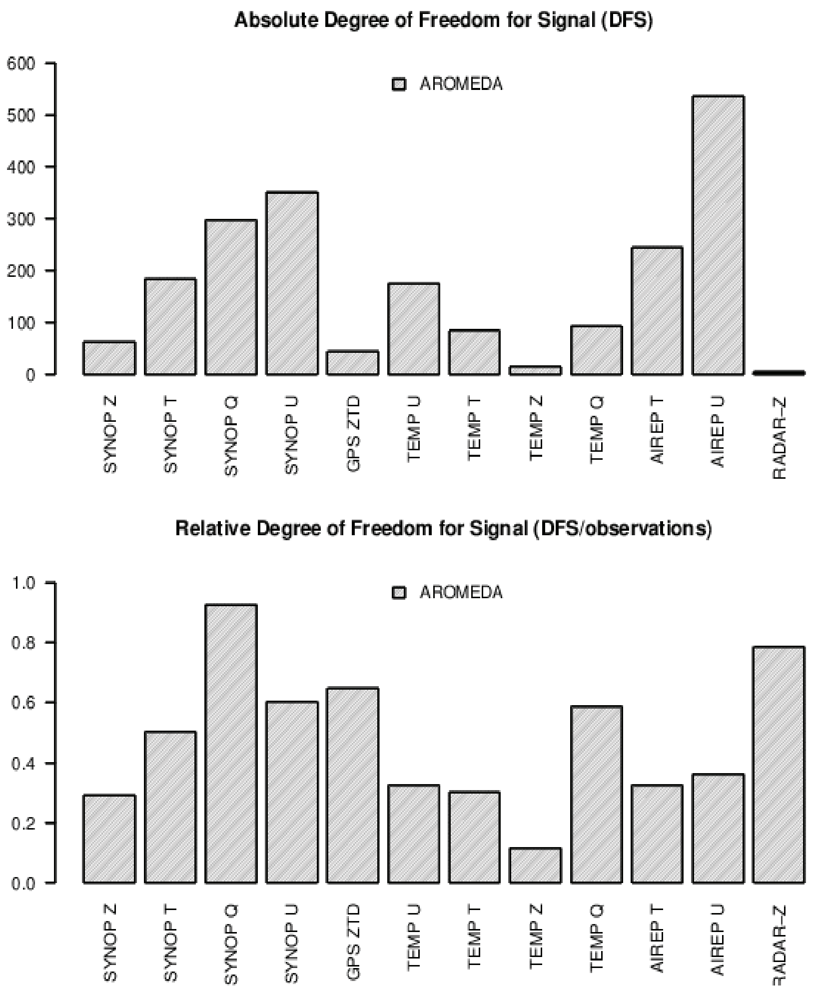

Fig. 10. Absolute and relative degrees of freedom for signal (DFS) for experimental data assimilation of AROME-Hungary at 12 UTC, January 3, 2014.

3.1. The impact of the Atmospheric Motion Vectors

EUMETSAT (European Organization for the Utilization of Meteorological Satellites) MSG provides sets of satellite winds (AMVs) extracted from sequences of well-navigated and calibrated images produced by the SEVIRI (Spinning Enhanced Visible and Infrared Imager) instrument. Accordingly, AMVs are derived from SEVIRI infrared, water vapor, and visible channels. At the HMS MSG AMV date is received through the EUMETCast broadcasting service of EUMETSAT with hourly frequency and processed in OPLACE for data assimilation purposes.

MSG AMV data is proved to be beneficial in nowcasting applications and in data assimilation systems (*Randriamampianina*, 2006b). Furthermore, numerous examples exist (*Forsythe et al.,* 2014), where AMVs are operationally assimilated in a similar way like the adopted technique used in the ALADIN model in Hungary. As MSG AMV observations possess many advantages needed for a RUC data assimilation, a summertime impact study has been run with AROME-Hungary, where MSG AMV observations were added to the conventional observations. The implementation of the AMV data in AROME-Hungary was done according to *Randriamampianina* (2006b). To assess the impact of AMVs two experiments were conducted during the period of June 25 – July 25 in 2012. The AROME-Hungary forecasts initialized at 00 and 12 UTC were verified against SYNOP and radiosonde observations. Regarding the impact of MSG AMV observations, in *Figs. 11a* and *11c* RMSE and BIAS scores are plotted with the corresponding (*Figs. 11b* and *11d*) normalized RMSE differences for 10 m wind speed and 2 m dew point temperature forecasts. In these figures, CONV stands for the operational AROME-Hungary and GEOW denotes the experimental run, where MSG AMV was assimilated as well. In case of 10 m wind speed forecasts, the AMV experiment provides better skill for the shorter range forecasts up to 15 hours. Concerning the 2 m dew point temperature verification scores, the overall decrease of the error is perceptible and it is even statistically significant for some longer ranges. To conclude, AMV data in AROME-Hungary provides small contribution with respect to the amount of assimilated data, but with positive signal on the short-range forecasts.

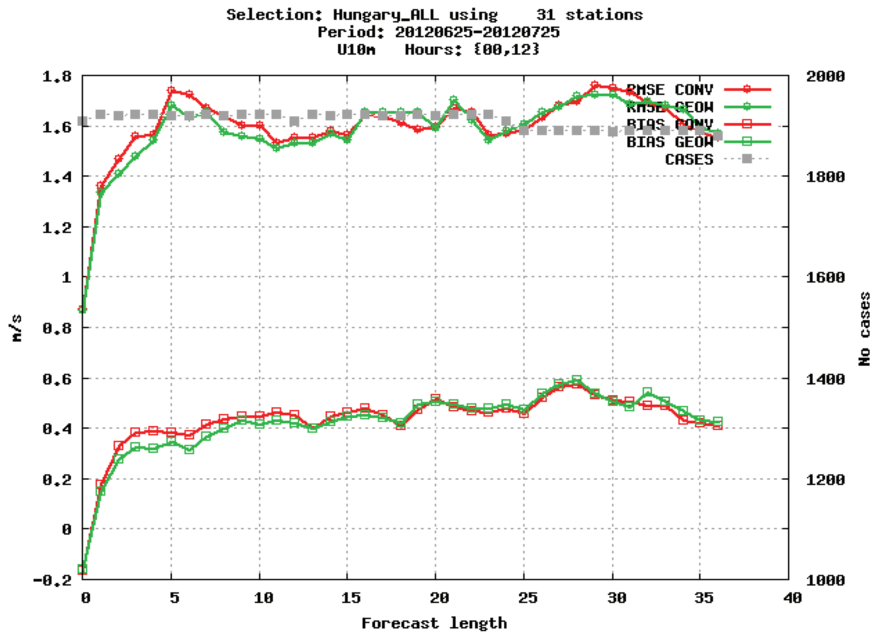

Fig. 11a. Experimental assimilation study of AROME-Hungary for the period between June 25 and July 25, 2013. RMSE and BIAS scores of AROME forecasts corresponding operational AROME with conventional observations (CONV – red) and AROME with conventional plus AMV observations (GEOW – green) are plotted for 10 m wind speed (m/s).

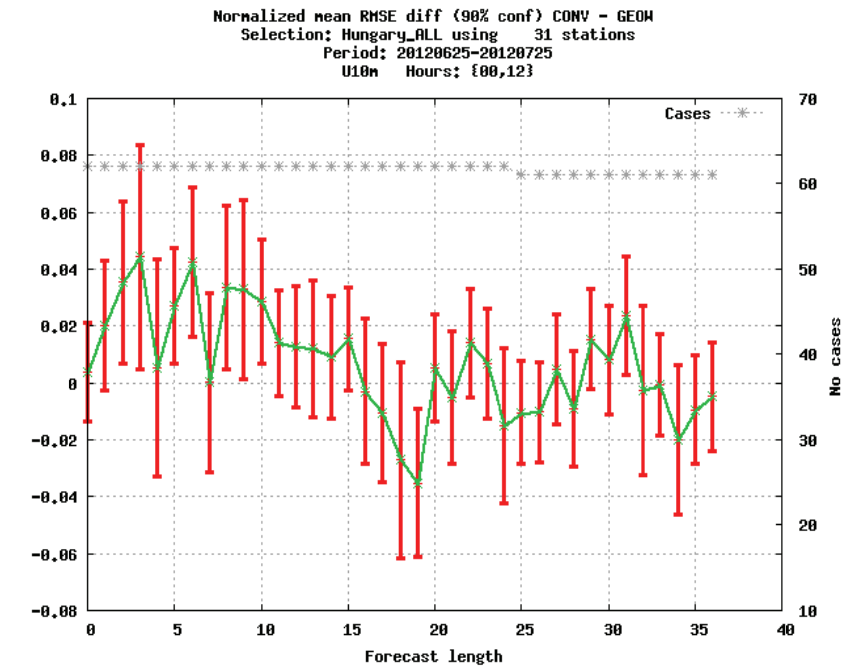

Fig. 11b. Experimental assimilation study of AROME-Hungary for the period between June 25 and July 25, 2013. Normalized RMSE differences between operational AROME with conventional observations (CONV – red) and AROME with conventional plus AMV observations (GEOW – green) are plotted for 10 m wind speed (m/s).

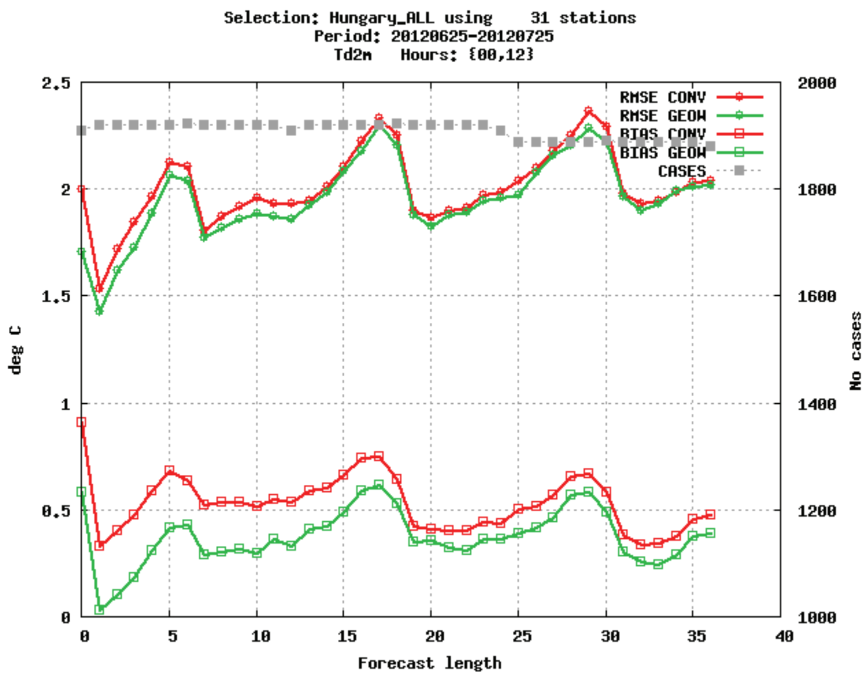

Fig. 11c. Experimental assimilation study of AROME-Hungary for the period between June 25 and July 25, 2013. RMSE and BIAS scores of AROME forecasts corresponding operational AROME with conventional observations (CONV – red) and AROME with conventional plus AMV observations (GEOW – green) are plotted for 2 m dew point temperature (^{o}C).

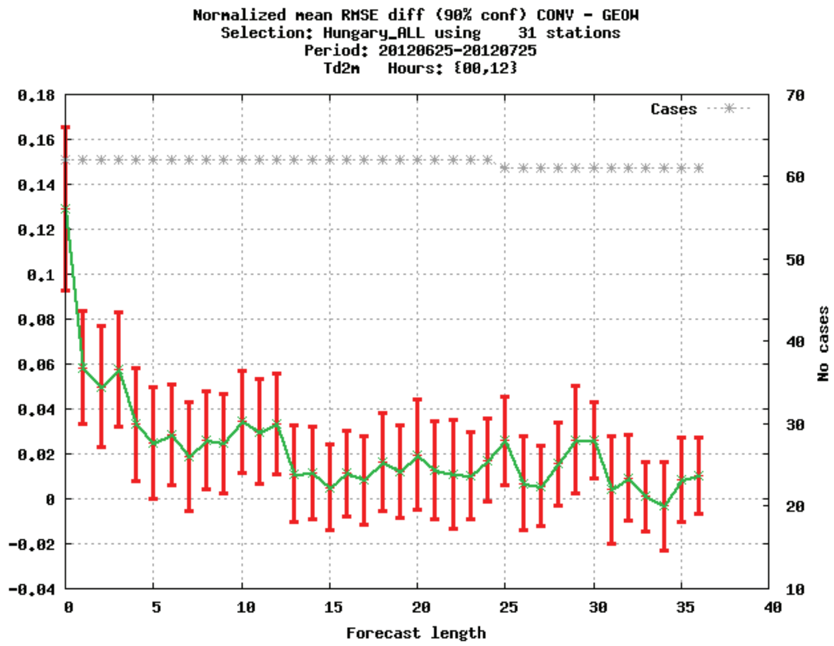

Fig. 11d. Experimental assimilation study of AROME-Hungary for the period between June 25 and July 25, 2013. Normalized RMSE differences between operational AROME with conventional observations (CONV – red) and AROME with conventional plus AMV observations (GEOW – green) are plotted for 2 m dew point temperature (^{o}C).

3.2. The impact of the radar reflectivity and radial wind observations

Radar measurements play an important role in nowcasting, and nowadays they also contribute to the initial conditions of mesoscale NWP models. The weather RADAR instrument receives emitted electromagnetic signal to measure the reflectivity of the atmosphere's elements along the emitted ray's path. From the backscattered radiation one can estimate the reflectivity, i.e., the precipitation intensity, and from the phase shift of the backscattered signals the radial wind can be measured using Doppler's law.

Focusing on data assimilation, the utility of radar observations has been already demonstrated by different studies (see, e.g., *Lindskog et al,* 2004; *Snyder* and *Zhang,* 2003; *Montmerle* and *Faccani,* 2008). However, assimilating the observed quantities of the radar is not straightforward since the relationship between the measured quantities and the control variables of the data assimilation scheme is complex and non-linear. This relation in case of radial wind observation is less complex than that with reflectivity, where the radar equation gives the direct relationship between the observed hydrometeors and the 3DVAR control variables. Instead of extending the control variables to account also for hydrometeors, an alternative solution was worked out by *Caumont et al.* (2010) and *Wattrelot et al.* (2014) using 1D+3DVAR method, which enables to retrieve columns of relative humidity and temperature from reflectivity profiles as pseudo-observations. This approach is based on a 1D Bayesian estimate, which uses the assumption that a well-chosen linear combination of model simulated reflectivities in the neighborhood of the observation provides comparable quantities to what is observed (see *Wattrelot et al.* (2014) for more details).

In the observing system of the Hungarian Meteorological Service dual-polarized Doppler radars are used which provide reflectivity and radial wind observations with 240 km and 120 km range, respectively. Raw radar data requires specific pre-processing in consideration of data assimilation which consists of the elimination of non-meteorological and noisy signals. Due to this quality control, for instance, reflectivity data under 7 dBz is filtered to avoid clear-sky echo and also unwanted RLAN (Radio Location Area Network) signals are rejected. After a thorough pre-processing with quality control, RADAR data is presented in Cartesian coordinates and in BUFR, which is one of the accepted format of the current 3DVAR system.

An observing system experiment with AROME-Hungary was made for an early, but convective summer period of May 15 – June 18, 2012. Point- and object-based verifications were computed to evaluate the performance of the operational AROME-Hungary and experimental runs including a combination of radar reflectivity and radial wind. In the first experiment with assimilation of both the radar Doppler wind and reflectivity, skill scores showed positive impact on forecasts of precipitation, but we observed also a cold and wet bias for surface parameters (not shown). A possible explanation of the observed bias might be that the assimilation of reflectivity data over-saturates the planetary boundary layers (PBL), which degrades the forecast of surface parameters through physical process along the model integration. To verify this assumption, another experiment was run avoiding the use of reflectivity observations below 1000 m from all 3 used radar stations. As a result, no degradation on surface parameters was observed, but on the other hand, the impact on precipitation forecasts was also reduced. The average intensity of the precipitation objects was verified against objects measured by the radar (*Fig. 12*). The four curves are respectively the operational AROME-Hungary (AromeCONV), the experimental AROME with complete set of radar data (AromeFULL), AROME runs with blacklisted reflectivity (AromeBLACK), and radar observations (RADAR). In *Fig. 12*, one can see that AromeBLACK provides the closest estimation to radar, however, the diurnal cycle of the maximum precipitation is still slightly shifted with delay in time. Additionally, a case study is shown in *Fig. 13*, where 3 hourly accumulated precipitation forecasts are plotted for all the three tested runs. One can see that the AromeFULL run predicts more realistic precipitation over north-eastern Hungary than AromeCONV, but it overestimates slightly in the mid-western part of the country. AromeBLACK is able to correct this overestimation, but the positive signal is also suppressed by filtering reflectivity. To conclude, the assimilation of radar data has major impact on forecasts of precipitation, but the quality control has to be further investigated and improved for better accounting of all potential measurements.

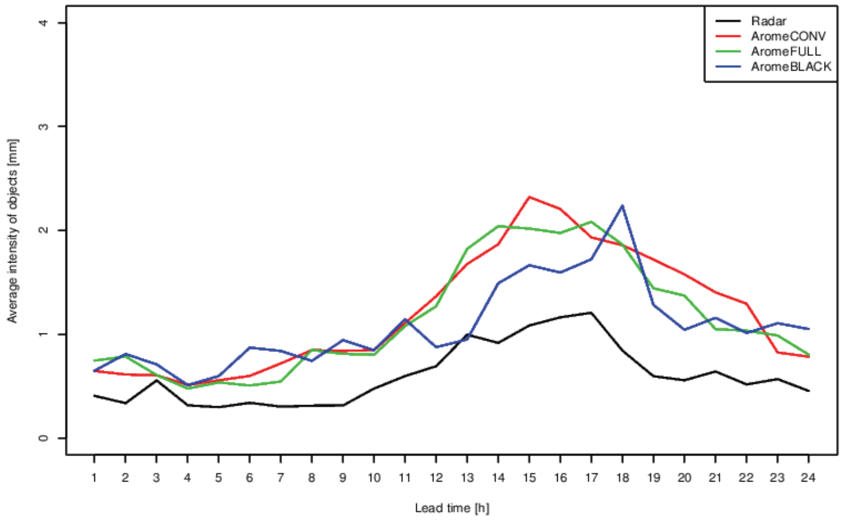

Fig. 12. Object-based verification of radar data assimilation experiments where average intensity of precipitation (mm/h) objects is verified against radar measurements for the period June 7 – June 18, 2012. AromeCONV: Operational AROME-Hungary (red line), AromeFULL: experimental AROME with RADAR reflectivity and radial wind observations added to conventional ones (green line), AromeBLACK: experimental AROME with same set of observations except reflectivity which was blacklisted below 1000 m elevation (blue line).

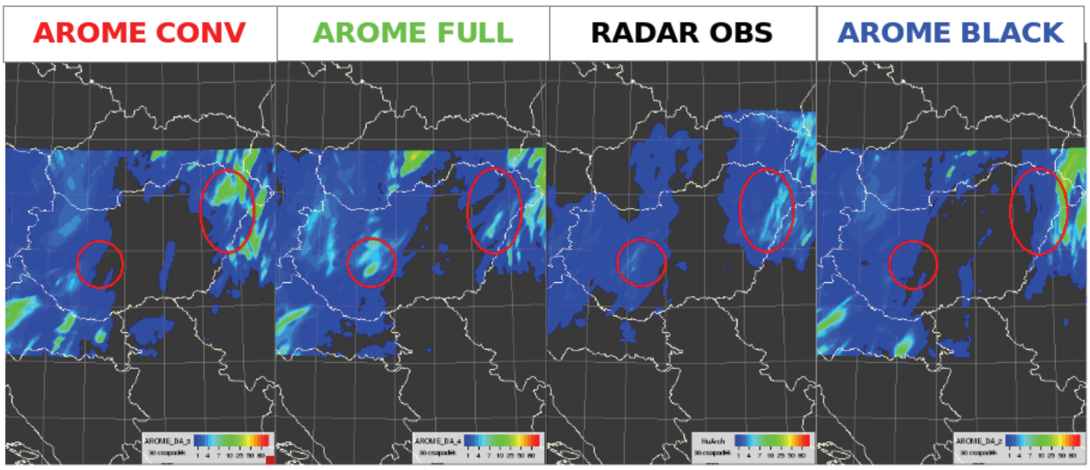

Fig. 13. A case study at 03 UTC, June 5, 2012 for 3 hourly accumulated precipitation forecast according to AROME model with operational configuration (AROME CONV), experimental AROME with radar reflectivity and radial wind (AROME FULL), experimental AROME with blacklisted low level reflectivity (AROME BLACK), and radar composite image (RADAR OBS).

3.3. The impact of the GNSS ZTD observations

Signal delay originating from different constituents of the troposphere and stratosphere can be extracted from satellite constellations of GNSS. The zenith tropospheric delay (ZTD), which is the converted-to-distance time delay, provides valuable information on atmospheric water vapor content expressed in length units along the zenithal direction above the ground-based GPS receiver station. *Bevis et al*., (1992) describes at length the principle of such measurement. The number of ground-based GPS stations over Europe has been increasing during the last years, and their use for meteorological purposes is coordinated by EUMETNET GNSS Water Vapour Programme (E-GVAP). E-GVAP also provides a data hub allowing the assimilation of GNSS ZTD observations with high spatial and temporal resolution. The Hungarian GNSS network (so called SGOB) operated by the Satellite Geodetic Observatory of Hungary was added to E-GVAP officially at the end of 2013, which provides access to a dense station network of ground-based GPS over the Carpathian Basin. This was a good motivation for us to assimilate the GNSS ZTD data.

The impact of GPS ZTD observations in data assimilation systems has been already investigated in the ARPEGE/ALADIN/AROME model family (see e.g., *Yan et al,* 2008; *Poli et al,* 2007; *Storto* and *Randriamampianina,* 2010). For the assimilation of E-GVAP ZTD data, a whitelist approach is used containing only stations with good-quality measurements. The whitelist is created according to the following criteria evaluated during a passive assimilation for a period of 15 days: i) the availablility of data is more than 40%, ii) observation minus background departures have Gaussian distribution, the absolute bias and also the standard deviation are both less than 40 mm, iii) the difference between station altitude and corresponding model orography height is less than 250 meter. We were able to choose 67 active stations inside our area of interest. The computed bias at each selected station is used as static bias correction in the assimilation scheme.

The impact of ground-based GNSS ZTD was investigated with AROME-Hungary over a winter period of 2014, namely January 5 to 27. The operational AROME-Hungary forecasts and the experimental AROME run with GNSS ZTD were compared with verification against SYNOP and radiosonde observations. In *Fig. 14a*, RMSE and BIAS scores are plotted for 2 m dew point temperature forecasts and the corresponding (*Fig. 14b*) normalized RMSE differences with significance test check. It can be seen that the experimental run (marked PGPS) has better skill scores on forecast of surface dew point temperature than the operational one (marked CONV), however, it is not statistically significant. In addition, one case study is presented in *Fig. 15*, showing the accumulated precipitation during the first 3 hours of the forecasts. In this case, the operational AROME-Hungary (CONV) provided a strong overestimation of precipitation as compared to the measured SYNOP observations plotted with numbers, probably due to spin-up effects. By assimilating ZTD observations (PGPS), the AROME forecast became more realistic with a reduction in the amount of predicted precipitation. This example showed, that the assimilation of GNSS ZTD observations is advantageous for improving short-range model forecasts, particularly regarding humidity, which is very promising to further improve the current operational AROME system.

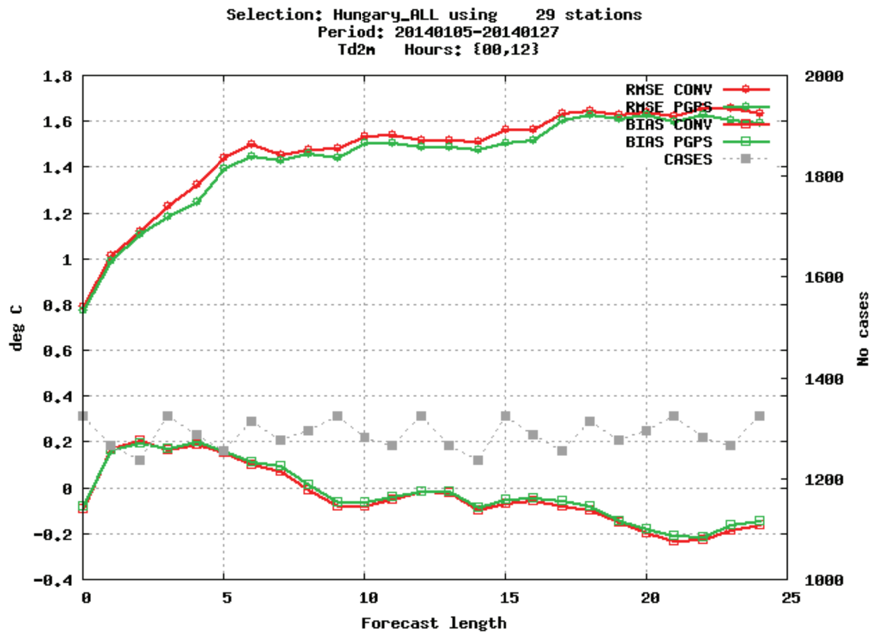

Fig. 14a. RMSE and BIAS scores of AROME forecasts corresponding to the operational AROME with conventional observations (CONV – red) and AROME with conventional plus GNSS ZTD observations (PGPS – green) for 2 m dew point temperature (°C).

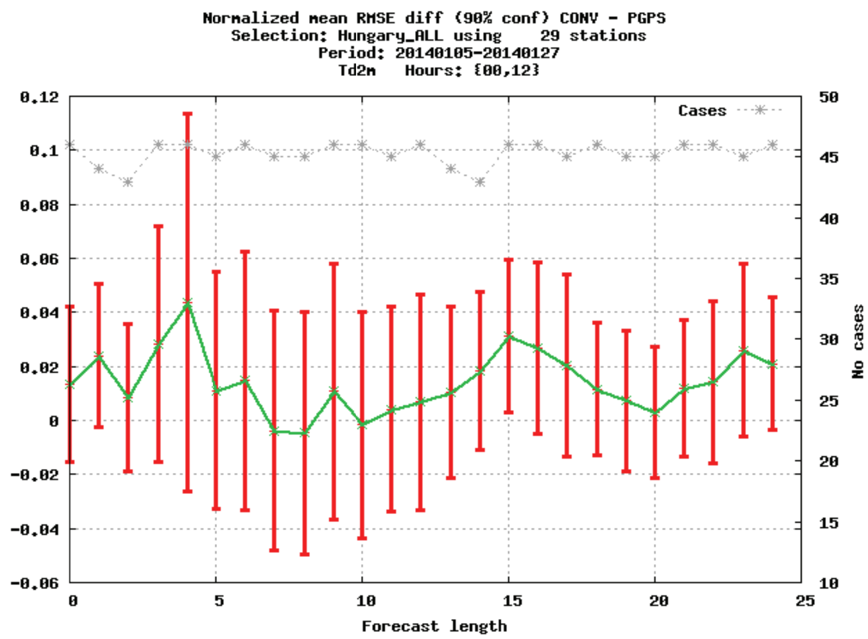

Fig. 14b. Normalized RMSE differences between operational AROME with conventional observations (CONV – red) and AROME with conventional plus GNSS ZTD observations (PGPS – green) for 2 m dew point temperature (°C).

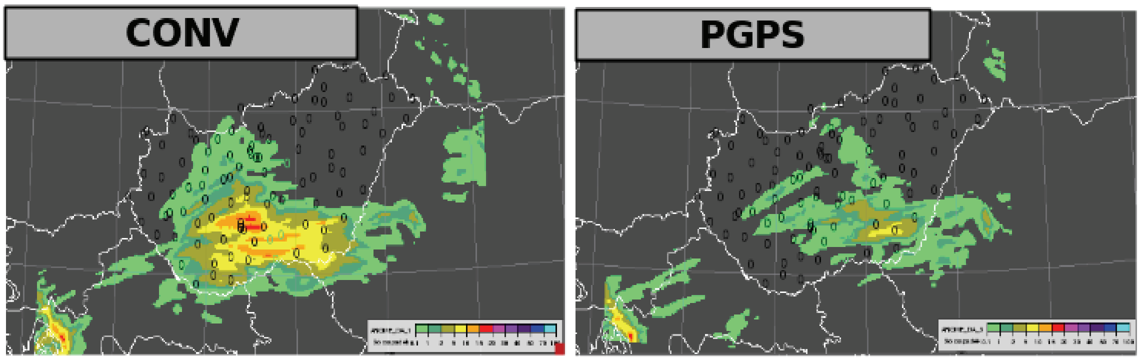

Fig. 15. A case study at 15 UTC, January 3, 2014 for 3 hourly accumulated precipitation forecast, AROME model with operational configuration (CONV), experimental AROME with GNSS ZTD (PGPS) and SYNOP precipitation (in numbers) are plotted.

4. *Summary and conclusions*

The current operational RUC data assimilation system of the AROME-Hungary mesoscale model has been described with a special emphasis on the design of the assimilation cycle and the use of observations. It has been demonstrated, that the RUC system using conventional observations (surface, radiosonde, and aircraft measurements) improves the reliability of short-range forecasts compared to the spin-up initialization technique (former operational configuration) and also compared to the use of a 6 hourly data assimilation cycle.

The most important attempts for improving the current operational RUC system so far consisted of impact studies using remote sensing observations, such as MSG AMV, radar reflectivity, radial wind, and GNSS ZTD. The impact of AMV data assimilation was found to be significantly positive on the forecast of surface parameters, up to a forecast range of 15 hours. These results imply an operational use of MSG AMV data in the near future. The assimilation of radar data has been found to be useful in ameliorating precipitation forecasts, however, as a side effect of radar data assimilation, a bias have been found in surface parameters. The cause of these systematic errors has to be understood in order to achieve an operational implementation of radar data assimilation. The impact of GNSS ZTD data assimilation has been found to be slightly positive regarding the forecasts of surface parameters. Given that GNSS ZTD data provide information on atmospheric humidity also in clear-sky conditions, their importance is high in mesoscale data assimilation. This is reflected in some of our case studies through the preferable feature that ZTD data assimilation allows to reduce possible humidity and precipitation overestimations originating from the model first guess. Based on the overall impact of ZTD data, they are anticipated for an operational implementation in the near future.

Apart from the observation impact studies, an overview has been given about the relative importance of observing networks and observed variables based on the DFS method. The main message to be extracted out of these analysis sensitivity studies is that humidity observations are really important in mesoscale data assimilation, as they have relatively large influence on the analysis as compared to other observed variables. This indicates that the density of humidity observations have to be increased in the coming years either by using cloudy information from satellites or by implementing humidity sensors on board European aircrafts, similarly to the practice applied at the USA.

The paper gives an indication, that by increasing the resolution of mesoscale models, it becomes highly important to implement local data assimilation at the full resolution of the model, using high-resolution observations. It is shown that doing so, the spin-up initialization scheme can be outperformed both in terms of verification scores and case studies. This experience justifies that data assimilation will remain one of the major directions for improving mesoscale forecasts at the Hungarian Meteorological Service, with a special emphasis on remote sensing data. Besides the implementation of new observations to the RUC system, attention will have to be paid to the development of the background error covariance representation, which is responsible for the efficient filtering and spreading of observed information to the model space. It is foreseen that the background error covariance matrix for AROME-Hungary will be recalculated based on AROME ensembles of data assimilations similarly to the work of *Brousseau et al.* (2011).

Acknowledgements: The authors would like to thank *András Horányi, Edit Adamcsek, Helga Tóth* for their earlier work on data assimilation developments at the Hungarian Meteorological Service. Also we are grateful to *Balázs Szintai* for providing the object-based verification for RADAR assimilation experiments, *Ulf Andrae* for installing the Harmonie verification package at HMS, *Szabolcs Rózsa* to preparing test ZTD data for the assimilation study, and to *Xin Yan* who provided useful advices for ZTD pre-processing. We would like to thank also *László Kullmann* for the great support of AROME configurations and *István Sebők* for the important development of RADAR BUFR converter. Last but not least we are grateful to the organizers of the 39th Meteorological Scientific Days to found the possibility of this IDŐJÁRÁS special issue.

References

Benjamin, S.G., Dévényi D., Weygandt S.S., Brundage K.J., Brown J.M., Grell G.A., Kim D., Schwartz B.E., Smirnova T.G. and *Smith T.L.,* 2004: An hourly assimilation-forecast cycle: The RUC. *Mon. Weather Rev. 132,* 495–518.

Berre, L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Mon. Wea. Rev. 128*, 644-667.

Bevis, M., Businger, S., Herring, T., Rocken, C., Anthes, R., and *Ware, R.H.,* 1992: GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system*. J. Geophys. Res. 97,* 15787–15801.

Bölöni, G., 2006: Development of a variational data assimilation system for a limited area model at the Hungarian Meteorological Service. *Időjárás 110*, 309–327.
Bölöni, G., Berre, L., and *Adamcsek, E.,* 2014: Comparison of static mesoscale background error covariances estimated by three different ensemble data assimilation techniques. *Q. J. Roy. Meteor. Soc., 141*, 413-425.
Brousseau, P., Berre, L., Bouttier, F., and *Desroziers, G.,* 2011: Background-error covariances for a convective-scale data-assimilation system: AROME–France 3D-Var, *Q. J. Roy. Meteor. Soc., 137,* 409–422.
Cardinali, C., Pezzuli, S. and *Andersson, E.* 2004: Influence matrix diagnostic of data assimilation system. *Q. J. Roy. Meteor. Soc. 130,* 2767–2786.
Chapnik, B., Desroziers, G., Rabier, F., and *Talagrand, O.,* 2006: Diagnosis and tuning of observational error in a quasi-operational data ssimilation setting, *Q. J. Roy. Meteor. Soc. 132,* 543–565.
Caumont, O., Ducroq, V., Watterlot, E., Jaubert, G., and *Pradier-Varbe, S.,* 2010: 1D+3DVar assimilation of radar reflectivity data: a proof of concept, *Tellus A 62*, 173–187.
Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic. D., Hamrud, M., Hollingsworth, A., Rabier, F., and *Fisher, M.,* 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Q. J. Roy. Meteor. Soc., 124*, 1783–1808.
Daley, R., 1991: Atmospheric Data Analysis. Atmospheric and Space Science Series, Cambridge.
Evensen, 2009: Data Assimilation, The Ensemble Kalman Filter, Springer.
Fabry, F. and *Sun, J.,* 2009*:* For How Long Should What Data Be Assimilated for the Mesoscale Forecasting of Convection and Why? Part I. *Mon. Weather Rev. 138,* 244–255.
Fischer, C., Montmerle, T., Berre, L., Auger, L., and *Stefanescu, S.* 2005: An overview of the variational assimilation in the Aladin/France numerical weather prediction system, *Q. J. Roy. Meteor. Soc. 131*, 3477–3492.
Forsythe M, C Peubey, C Lupu and *J Cotton,* 2014: Assimilation of wind information from radiances: AMVs and 4D-Var tracing, *ECMWF Annual Seminar*, September 2014, available in-line at: http://www.ecmwf.int/sites/default/files/ecmwf_sep14.pdf.
Girard, D., 1987: A fast Monte-Carlo cross-validation procedure for large least-squares problems with noisy data. *Technical Report 687-M, IMAG, Grenoble, France*
Horányi, A., Ihász, I. and *Radnóti, G.*, 1996: ARPEGE/ALADIN: a numerical weather prediction model for Central-Europe with the participation of the Hungarian Meteorological Service. *Időjárás 100*, 277–301.
Kalnay, E., 2003: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press.
Lindskog, M., Salonen, K., Järvinen, H.,Michelson, D., B., 2004: Doppler Radar Wind Assimilation with HIRLAM 3DVAR. *Mon Wea. Rev., 132,* 1081-1092.
Lorenc, A., 1986: Analysis methods for numerical weather prediction*. Q. J. Roy. Meteor. Soc. 112,* 1177–1194.
Lynch, P., D. Giard, and *V. Ivanovici,* 1997: Improving an efficiency of Digital Filtering Scheme for Diabatic Initialization. *Mon. Wea. Rev. 125*, 1976–1982.
Montmerle T. and *Faccani, C.,* 2008: Assimilation of Doppler wind into the French mesoscale model AROME. Proceedings of the ERAD 2008 conference *(*http://erad2008.fmi.fi/proceedings/index/session_wednesday.html*)*
Poli, P., Moll, P., Rabier, F., Desroziers, G., Chapnik, B., Berre, L., Healy, S. B., Andersson, E., and *EI Guelai, F.-Z.* 2007: Forecast impact studies of zenith total delay data from European near realtime GPS stations in Meteo France 4DVAR, *J. Geophys. Res. 112*, D06114.
Randriamampianina, R., 2006a: Impact of high resolution observations in the ALADIN/HU model. *Időjárás 110*, 329–349.
Randriamampianina, R., 2006b*:* Investigation of the AMV data derived from Meteosat-8 in the ALADIN/HU data assimilation system. Eighth International Winds Workshop. 24-28 April 2006 Bejing, China, Available from: http://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_CONF_P47_S2_06_RANDRIAMA_V&RevisionSelectionMethod=LatestReleased&Rendition=Web

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F. Lac, C., and *Masson, V.,* 2010: The AROME-France Convective-Scale Operational Model, *Mon. Weather Rev. 139*, 976–991.
Snyder, C. and *Zhang, F.,* 2003: Assimilation of Simulated Doppler Radar Observations with an Ensemble Kalman Filter, *Mon. Weather Rev. 131,* 1663–1677.
Storto, A. and *Randriamampianina, R.,* 2010: A new bias correction scheme for assimilating GPS zenith tropospheric delay estimates, *Időjárás 114,* 237–250.
Vasiliu S. and *Horányi A.,* 2005: An evaluation of the performance of the three-dimensional variational data assimilation scheme for the ALADIN/HU spectral limited area model, *Időjárás 109,* 233–257.
Wattrelot, E, O Caumont, and *J-F Mahfouf,* 2014: Operational Implementation of the 1D+3D-Var Assimilation Method of Radar Reflectivity Data in the AROME Model. *Mon. Weather Rev. 142,* 1852–1873.
Yan, X., Ducrocq, V., Poli, P., Jaubert, G., and *Walpersdor,f A.,* 2008: Mesoscale GPS Zenith Delay assimilation during a Mediterranean heavy precipitation event, *Adv. Geosci.* 17, 71–77.