

## Experiments with three-dimensional radar reflectivity data assimilation into the COAMPS model

Jakubiak Bogumił

*University of Warsaw, Interdisciplinary Centre for Mathematical and Computational Modelling, Pawińskiego Street 5a, block D, 02-106 Warszawa, Poland, e-mail: B.Jakubiak@icm.edu.pl*

Szturc Jan, Ośródka Katarzyna, Jurczyk Anna

*Institute of Meteorology and Water Management, National Research Institute, Podleśna Street 61, 01-673 Warszawa, Poland, e-mail: jan.szturc@imgw.pl, katarzyna.osrodka@imgw.pl, anna.jurczyk@imgw.pl*

**Abstract.** High temporal and spatial resolution of radar measurements enables to continuously observe dynamically evolving meteorological phenomena. Three-dimensional (3D) weather radar reflectivity data assimilated into the numerical weather prediction model has the potential to improve initial description of the atmospheric model state. The paper is concentrated on the development of radar reflectivity assimilation technique into COAMPS mesoscale model using an Ensemble Kalman Filter (EnKF) type assimilation schemes available in Data Assimilation Research Testbed (DART) programming environment. Before weather radar data enter into the assimilation system, the measurement errors are eliminated through quality control procedures. At first artifacts associated with non-meteorological errors are removed using the algorithms based on analysis of reflectivity field pattern. Then procedures for correction of the reflectivity data are employed, especially due to radar beam blockage and attenuation in rain. Each of the correction algorithms is connected with generation of the data quality characteristic expressed quantitatively by so called quality index (QI). In order to avoid transformation of data uncertainty into assimilation scheme only the radar gates successfully verified by means of the quality algorithms were employed in the assimilation. The proposed methodology has been applied to simulate selected intense precipitation events in Poland in May and August 2010.

**Key words:** weather radar, numerical weather prediction, radar reflectivity, quality control, assimilation

**Submitted** 12 December 2013, **revised** 28 January 2014, **accepted** 15 July 2014

### 1. Introduction

#### 1.1. The problem

Data assimilation as a major component of numerical weather prediction (NWP) models includes the process of creating the best estimate of the initial state of the atmosphere. This involves combining all sources of information from a forecast model itself, together with available observations – thus “analyses” of the atmosphere are provided. Therefore, two types of information are applied: (i) previous short-term model forecasts as the first guess, and (ii) multi-source observations, along with their associated uncertainties. The most important factor in the determination of the initial state are meteorological parameters directly forecast by an NWP model, usually including wind components, temperature, humidity, pressure, and geopotential.

Recent progress in computational possibilities enables the development of models with higher spatial resolution, which require the provision of data describing smaller scale phenomena, e.g. convective processes. High temporal resolution (about 5-10 min) and spatial resolution (mostly 1 x 1 km) of weather radar measurements allows us to observe, nearly continuously, dynamically evolving meteorological phenomena. Such sets of radar reflectivity data assimilated into the NWP model have the potential to improve the initial description of the atmospheric model state.

The main aim of the research was to develop procedures to employ three-dimensional (3D) weather radar reflectivity data for assimilation to NWP models.

#### 1.2. Scheme of weather radar data assimilation

##### *State of the art*

Weather radar reflectivity or radar-based precipitation estimates can be assimilated in NWP models through different approaches. The assimilation of radar information requires complicated observation operators including moist physics. The main techniques are nudging, variational analysis, and ensemble Kalman filtering.

(1) *Nudging* aims to force the model state towards available observations during the model integration. It is achieved by introducing a relaxation term, which measures the “observation minus model” distance. In the nudging scheme the radar-based precipitation estimate is assimilated instead of the reflectivity. One popular method of nudging precipitation information is the Latent Heat Nudging (LHN) technique (Jones, Macpherson 1997).

(2) *Variational assimilation* (3/4-DVAR) solves the analysis problem through the minimization of a given criterion – the so-called cost function which describes the distance between model simulations and available observations. The assumption of the method is that the forecast model is “perfect” within the assimilation win-

dow and the model trajectory which best fits the observations over the window is searched for. Intensive work on four-dimensional variational techniques 4D-VAR was performed by Sun and Crook (1997, 1998) during the late 1990s.

(3) *The ensemble Kalman filter* (EnKF) method transforms the analysis ensemble of model states from one analysis time to the next using the sample mean and covariance of the resulted ensemble as the background mean and covariance for the new analysis (Kalman 1960). Assuming certain idealized conditions, it can be shown that the solution obtained through the EnKF technique is comparable to that obtained using 4D-VAR. The limitation of the technique is the difficulty in estimation of initial conditions before the assimilation of any observations (Jakubiak 2008; Jakubiak, Szpindler 2009).

Modern filtering methods for data assimilation employ information about the structure of preliminary approximation error covariance and observation error covariance. Approaching the ensemble filtering problem as a Monte Carlo approximation to the Bayesian filtering leads to a general solution (Jazwinski 1970). Forecast errors are a result of errors incorporated in the initial data, and of the NWP model itself. The error statistics vary in time and depend on meteorological conditions over a given area.

#### *Description of developed assimilation scheme*

For the aim of the research, a scheme for weather radar data assimilation using square root ensemble filters was developed. The filtering method is based on the Kalman filter and consists of two stages. The first step is a forecast of the model state until the time of the next observation. The ensemble members are evolved independently, according to the nonlinear weather prediction model. The second step is the analysis, i.e. adjustment of the entire ensemble to the information available from new observations. Such an update reflects both the new state estimate and the uncertainty of the observations. The system of data assimilation transforms the ensemble of forecasts into an ensemble of analyses with statistics fitted to observation data. The transformation in this investigation is performed by applying the Monte Carlo technique to square root Kalman filter. In the case of ensemble-based Kalman filter methods applied to complicated NWP models with a large state vector, the need to accurately estimate covariance structures from finite (and small) samples is a challenging task.

Square root ensemble filters are not unambiguous, as different ensembles can generate the same error covariance matrix. This ambiguity led to the development of a number of different algorithms for the modification of an ensemble of atmospheric state analyses. In the research, the following two approaches were tested: (i) the transformed Kalman filter proposed by Bishop et al. (2001), and (ii) the technique proposed by Whitaker and Hamil (2002).

### 1.3. Tools

#### *NWP model COAMPS*

The NWP model COAMPS (Coupled Ocean/Atmosphere Mesoscale Prediction System), developed in the Naval Research Laboratory (Hodur 1997), is a hybrid ocean and atmosphere model employed by the US Navy to provide meteorological protection of its operations over given parts of the world. The model is also used at the Interdisciplinary Centre for Mathematical and Computational Modeling of the University of Warsaw (ICM UW), both operationally and for research work on the development of new algorithms describing physical processes in the atmosphere.

The atmospheric part of the COAMPS models is comprised of non-hydrostatic, fully compressible equations of motion, following the paper of Klemp and Wilhelmson (1978). The model has a total of eleven prognostic variables including three Cartesian velocity components, five mixing ratios (cloud, rain, snow, graupel and water vapor), potential temperature, perturbation Exner function, and turbulent kinetic energy.

The model is constructed in the form of deviations of basic prognostic variables from the time-invariant mean state. The mean state is assumed to be in hydrostatic balance. The equations are solved on a staggered, scheme C grid (Arakawa, Lamb 1977). In general, all derivatives are computed to second-order accuracy. Exceptions are the horizontal diffusion and an option for fourth-order accurate horizontal advection.

The moist physics scheme in COAMPS currently consists of a single-moment bulk prediction of mixing ratio, as developed by Rutledge and Hobbs (1983). The scheme is a bulk cloud microphysical model, based on the Lin et al. (1983) formulation, with single-moment prediction of mixing ratio for five predictive variables in cloud micro-physics (vapour, pristine ice, snow, rain, and cloud water). The primary assumptions used in this scheme include use of the Marshall and Palmer (1948) drop size distribution, the Kessler (1969) autoconversion, and the Fletcher (1962) formulation for the nucleation of pristine ice. Terminal velocities are computed for the rain and snow fields while the remaining fields are treated as scalar tracers.

The bulk scheme is invoked after the model dynamics have been calculated and the scalar prognostic variables have been updated for advection, diffusion, and mixing processes. In addition to the five scalar fields of the original scheme, predictive equations for graupel (heavily rimed ice crystal or frozen drop) (Rutledge, Hobbs 1984), and drizzle (Khairoutdinov, Kogan 2000), have been added.

An important feature of the COAMPS model for this research is that the mixing ratios describing the content of cloud liquid water, ice, rain, snow, and graupel in the atmosphere are prognostic variables (evolving in time), which is a factor that enables us to determine (as simulated by the model) the predicted values of reflectivity and the radial component of the wind speed. Observed and simulated values of reflectivity can be assimilated into the model

directly, without the conversion of reflectivity into precipitation, which allows us to avoid introducing the errors associated with such a transformation. In the data assimilation scheme developed in this research, the observed variable is the radar reflectivity.

#### *System of data assimilation DART*

The Data Assimilation Research Testbed (DART) is an open source community software facility for ensemble data assimilation – it was developed at the National Center for Atmospheric Research (NCAR). It makes it easy to match a variety of data assimilation methods to different numerical models and different kinds of observations. DART produces ensemble-mean analyses providing the initial conditions for ensemble predictions.

In the research for all experiments, the DART was employed to provide real-time, mesoscale analyses that were used as the initial state of the atmosphere for the COAMPS-derived forecasts.

#### *Polish weather radar network POLRAD*

The weather radar data employed in the investigations were generated by the Polish radar network POLRAD, which is operated by the Institute of Meteorology and Water Management – National Research Institute (IMGW-PIB). At present, the POLRAD consists of eight C-Band (5.7 GHz) Doppler radars of Selex SI Gematronik, two of which are polarimetric. The measured quantity is radar reflectivity  $Z$  expressed in dBZ, which can be transformed into precipitation rate  $R$  in  $\text{mm h}^{-1}$  using empirical formulas. The standard range of the measurements is 250 km.

## **2. Corrections of 3D weather radar reflectivity data**

In the research, radar data as 3D reflectivity volumes were employed for assimilation to numerical models. This raw data includes values of measured quantities in spatial points (so called gates) defined by scan strategy (i.e. settings of antenna elevations, azimuths, and bins).

The data are burdened with numerous measurement errors, which are difficult to eliminate due to their very complicated structure. The most important errors are connected with the presence of spurious echoes (biological, ground clutter, interference from external antennas, etc.), the underestimation of measurements due to the partial or total blockage of the radar beam, attenuation in rain, etc. Moreover, the representativeness of measurement depends on beam broadening and the scan strategy employed. Therefore, the correction of radar data is necessary – this is often made using external information – and also the characterization its quality is required. This is a big challenge since there is no benchmark for 3D data such as, for example, rain gauge data for ground precipitation estimates.

The algorithms employed for both the data correction and quality characterization of 3D weather radar reflectivity are listed in Table 1 (Ośródka et al. 2014). They are

presented in order of their implementation in the chain of the data processing. The designed algorithms were partially developed in other research projects, especially the BALTRAD (“An advanced weather radar network for the Baltic Sea Region: BALTRAD”) within the frame of Baltic Sea Region Programme 2007–2013, which was financed by the ERDF (Ośródka et al. 2010; Szturc et al. 2012a).

Each of the correction algorithms is connected with the generation of the data quality characteristic expressed quantitatively by the so called quality index. The quality index (QI) is defined as a unitless metric of data quality based on the analysis of a given quality factor. Its values range from zero to one, for ‘totally uncertain’ and ‘excellent’ data respectively (Einfalt et al. 2010).

All the algorithms, outlined below, are described in detail in the papers by Ośródka et al. (2014) and Szturc et al. (2012).

(1) *Spike echo (interfering signals) recognition and removal.* Signals coming from the Sun and external sources that interfere with the radar signal have frequently become sources of non-meteorological echoes in radar images. These spurious echoes, so called spikes, are characterized by their spatial structure that clearly differs from the precipitation field pattern. The shape of such an echo is very specific: it is similar to a spike along the whole, or a large fraction, of a single or a number of neighbouring radar beams (Zejdlik, Novak 2010). Recognition of this type of echo is not very difficult unless it interferes with a precipitation field.

In the developed algorithm, two categories of spikes are introduced: “wide” and “narrow”. For “wide” spikes, a subalgorithm based on the analysis of the spatial structure of a radar echo is employed. The variability of the echo is examined using reflectivity variances determined within a certain vicinity across and along the radar beam. A given echo is classified as a potential spike if the variance across the radar beam is high while variance along the beam is low. A subsequent subalgorithm is employed to recognize and remove “narrow” spikes, i.e. those not wider than  $7^\circ$  in azimuth. Gates with “narrow” spikes are detected by checking reflectivities at neighbouring beams of azimuth at the same distance to radar.

Reflectivity in gates with spikes is interpolated from values in the nearest gates not burdened with spikes. The quality index for the whole radar beam in which a spike is detected is decreased.

(2) *Non-meteorological echo recognition and removal.* In the algorithm, two stages of non-meteorological echo removal are introduced: one for “low” and one for “high” types of echoes. For “low” non-meteorological echo removal a detection scheme is designed in which the following two parameters are investigated: radar reflectivity and echo height. Values of two specific linear detection functions of the parameters are computed, and their product is compared with the detector threshold. If the product is higher than the threshold

Table 1. Correction algorithms for 3-D weather radar reflectivity data in order of their implementation in chain of the data processing

Algorithm	Problem to solve	Investigated information
SPIKE	Removal of geometrically-shaped non-meteorological echoes spike type	2-D structure of radar reflectivity data
NMET	Removal of non-meteorological echoes	2-D structure of radar reflectivity data
SPECK	Removal of measurement noise (speckle and reverse speckle echoes)	2-D vicinity of radar reflectivity data
BLOCK	Correction of partial and total beam blockage including ground clutter	Radar beam propagation based on digital terrain map
ATT	Correction of attenuation in rain	Precipitation along radar beam

then the echo is considered non-meteorological, and is removed. In the second subalgorithm, all echoes detected at altitudes higher than the preset threshold (where any meteorological echo cannot exist – for a physical reason) are classified as “high” spurious echoes and removed. The quality of measurement gates with echoes identified as non-meteorological is reduced.

(3) *Reverse speckle and speckle echo recognition and removal.* Reverse speckles are isolated gates, without echo, located next to precipitation gates. An algorithm is applied to each elevation separately. A grid of 3 x 3 gates around a given gate is investigated and the number of non-echo gates is calculated. If the number is lower than a preset threshold then the central gate is considered to be reverse speckle and reflectivity from neighbouring gates is taken instead (Jurczyk et al. 2008). Speckles are gates where isolated echo is detected, and then treated as measurement noise. The algorithm for such echo detection is analogous to the one for reverse speckles. The quality of gates with these echoes is reduced.

(4) *Radar beam blockage recognition and correction.* In order to calculate the level of radar beam blockage by terrain obstacles, a geometrical approach is employed. It consists of the calculation applying digital terrain map (DTM) which part of a beam cross-section is blocked – this is expressed by a coefficient called partial beam blockage (PBB) with values between 0 and 1 (Bech et al. 2007). The PBB coefficient is also a basis of the quality index determination.

(5) *Attenuation in rain correction.* Attenuation is defined as a decrease in signal power after passing meteorological objects. The phenomenon results in the underestimation of reflectivity. An iterative algorithm is employed which calculates the two-way (from and to radar) attenuation between two subsequent gates along each radar beam and then a path-integrated attenuation. In the designed algorithm the attenuation at a 1 km distance is estimated for each gate based on empirical formulas that can be found in literature (e.g. Collier 1989). The attenuation magnitude is also a metric for quality index determination.

Apart from quality factors related to the corrections, a quality index related to radar beam properties (beam

broadening with distance from a radar site) is taken into account in data quality evaluation (see Table 1).

After the quality control procedure, the total quality index for each radar gate is estimated by multiplication of all particular indices. In order to avoid the transformation of uncertainty in data into an assimilation scheme in the research, only the radar gates successfully verified by means of the quality algorithms were employed in the assimilation.

### 3. Climatological assessment of precipitation

#### 3.1. Procedures for determination of initial condition

The estimation of the initial conditions for a single forecast of the COAMPS system is implemented with the use of the multi-variable optimal interpolation method (MVOI). In this approach, data from radiosondes, surface synoptic stations, remote satellite and commercial aircraft data can be used. The procedure for radar data assimilation is activated after reaching a balance between basic model variables.

The strategy for achieving such a dynamically balanced stage consists of the performance of two 6-hour traditional data assimilation cycles with the use of all available observations (excluding radar data). To get the next cycle of data assimilation, the 6-hour COAMPS model forecasts are generated. Finally, having an initial state of the model atmosphere prepared in such a manner, the ensemble of initial conditions is produced, and for each of the ensemble members the forecast is continued till the time when the radar reflectivity data are available. Next, in a selected time window, the radar data assimilation is performed with the use of the Ensemble Kalman Filter method. The reflectivity data is the only one parameter assimilated in the DART/COAMPS system. The connection between DART and COAMPS systems is through so-called “restart files”, in which the actual model state is stored. The COAMPS model produces this type of data in selected time steps. The DART system uses restart files to modify the actual model state with new information included in the reflectivity data. In order to do this, the sequential filtering method is implemented. In this method each value of the observation vector (reflectivity) is correlated with each value of the model state vector. Observational data are assimilated every 10 minutes, one pack of data consists

of the radar reflectivity from 6 time-steps for a selected hour. The DART system estimates not only the importance of the relation between the observation vector and model state vector, but also verifies the time of the observation. If the observation time is prior to the actual time of the forecast, the DART starts the COAMPS forecast for all members of the ensemble to ensure that the time of the forecast is equal to the time of the observation.

One of the principles of the DART system is the creation of an ensemble of model states in which the spread of results is sufficiently large. Too small a dispersion of the forecast ensemble gives the wrong results of data assimilation – this is caused by the divergence of the filter. In order to cope with this, the noise typical to the variability of the real observation is added to each single atmospheric state. The choice of the meteorological element to which the noise is added depends on its impact on the final result. Preliminarily, all elements of the model state vector were disturbed – this turned out to be a bad choice, as some instabilities in the solution were observed. Finally, only selected variables were disturbed, in order to get a sufficient spread of solutions.

The achievement of good data assimilation results using the filtering approach, where the dimension of the model state is of the order  $O(10^8)$ , strongly depends on the technical limitations. Preliminary plans to use the ensemble with 50 members were modified (to 13 members) – this still involves about 1.5 TB of disk space for one experiment.

### 3.2. Configuration of the COAMPS model

**Flood** For all experiments with data assimilation, the configuration of the COAMPS model was fixed to the 40 vertical levels, with variable distances between levels. The model atmosphere depth is from ground level to the height of 29 km. The 18 lowest levels are located in the atmospheric boundary layer.

Initially, three nested grids were used (left panel on Fig. 1). The coarse mesh 1a consists of 255 x 255 grid points, with 9km spatial distances between them. The medium mesh 2a, spanned over Poland, has 310 x 310 grid points with spatial distances of 3 km. The finest mesh 3a, located over the Mazovia District, consists of 256 x 256 points with 1km spatial distances. Experiments with data assimilation in this set-up were conducted for the medium and finest grids. The computations were difficult to execute in reasonable time for a few important reasons: the extremely long time of the filter computations caused by the magnitude of the state vector and problems with the development of the nested grid approach in the DART/COAMPS interface. In order to finish some computation, the authors decided to implement a simpler solution – introducing a new configuration of grids, this time with one coarse 6km grid with 128 x 128 grid points, and one nested 2km grid with 128 x 128 grid points radically reduced the dimension of the model state vector. The number of grid points for one field in a new 1b nest was decreased by the rate of 2.4 compared to the number of grid points in the old, 2a, medium nest. Because of the long calculation time for the 2b nest, all described experiments with data assimilation were performed on a coarse 6-km grid in order to conduct simulations for a higher number of events.

## 4. Analysis of developed assimilation algorithms for selected case studies

### 4.1. Input data

Strong convection events from May and August 2010 were selected as case studies in the experiments. The number of test cases was limited to four, due to the long calculation time and the need to repeat them after the alteration of any assimilation parameter or initial conditions. The data from the Legionowo and Brzuchania radars were employed. In each case, three-dimensional radar reflectivi-

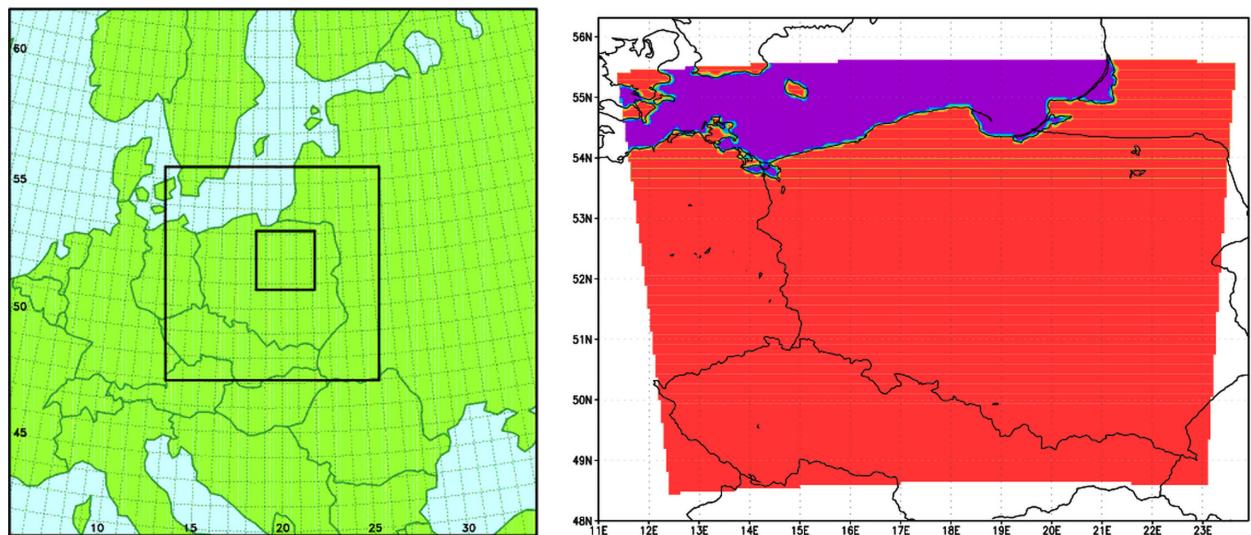


Fig. 1. Domains of COAMPS model in (a-left) the first and (b-right) the second experiments with the assimilation of radar reflectivity

Table 2. The number of data available for experiments

Date	10 May 2010	6 August 2010	15 August 2010		16 August 2010
Time	13 UTC	17 UTC	13 UTC	17 UTC	17 UTC
Radar	Legionowo	Legionowo	Brzuchania	Legionowo	Legionowo
0 min	40 251	99 672	106 123	79 416	36 180
10 min	40 322	104 832	105 992	84 724	37 764
20 min	41 827	108 446	107 377	85 987	40 252
30 min	41 803	112 134	113 059	83 214	42 736
40 min	42 339	115 646	116 737	80 903	45 300
50 min	42 523	113 738	122 994	77 473	48 681

ty data were converted into reflectivity vectors according to the DART format, with the elimination of all the data with reflectivity values below a preset threshold (which was taken as 0 dBZ) and with a low quality index (QI). Data sets were limited to the first five elevations: 0.5°, 1.4°, 2.4°, 3.4°, and 5.3° due to the rapid decrease in the amount of data with the elevation angle. Table 2 summarizes the amount of data supplied to the DART system in the experiments.

#### 4.2. Input data

##### Case study of 10 May 2010

Almost the whole of Europe was under the influence of shallow low pressure areas, one with its centre over Scandinavia, and a second moving from the Bay of Biscay towards France. The weather in Poland was mainly affected by a trough of low pressure between those two centres. The pressure gradient was low, and light winds blew from different directions. During the day, convective precipitation organised in the narrow zone moved over Poland in a north-easterly direction (Fig. 2).

In this case it was difficult for the model to forecast the precipitation zone correctly in space and time. After assimilation, radar data forecasts with lead times above 30 minutes improved. However, the results were not satisfactory.

##### Case study of 6 August 2010

Eastern Europe was influenced by a high with its centre over Russia. A wedge of high pressure developed over Finland, whereas a deep trough of low pressure was observed over northern Italy. The low extended and moved north-east, reaching Poland in the night. That atmospheric circulation triggered an inflow of warm and humid airmasses from Hungary, which resulted in rainfall and thunderstorms (Fig. 3).

Simulations for this case were probably unsuccessful due to the incorrect determination of boundary conditions and also due to limited computational capabilities.

##### Case study of 15 August 2010

Europe was affected by a cold front stretching from Tunis to Austria. When the low pressure trough created

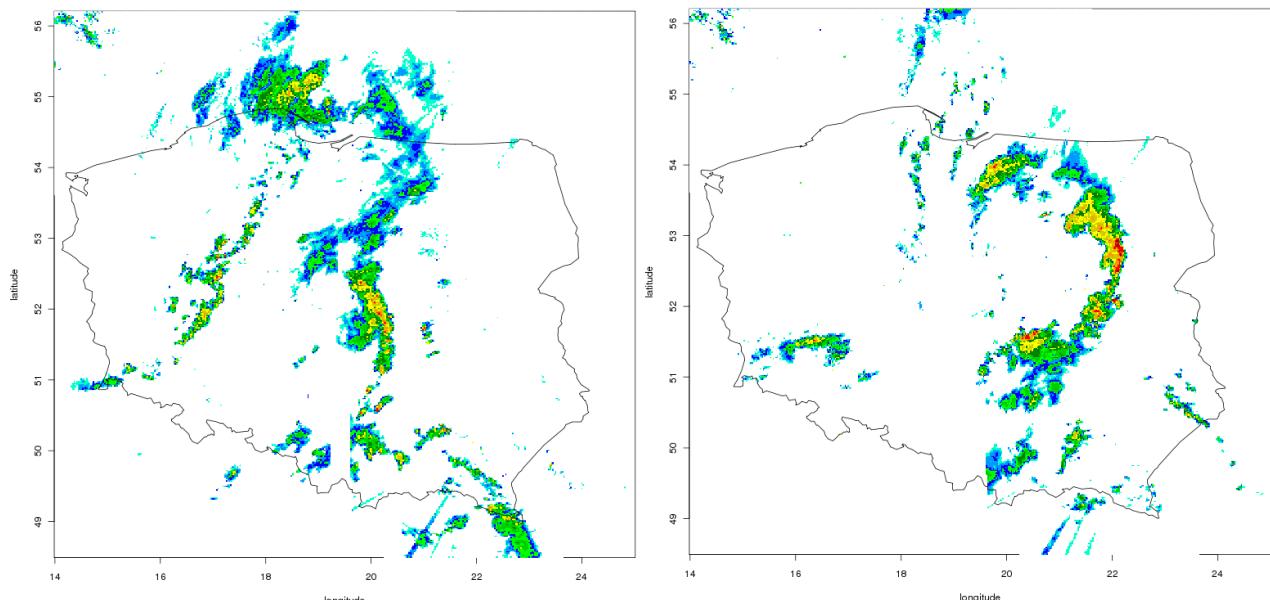


Fig. 2. Radar reflectivity on 500 m a.s.l. observed in Poland, 10 May 2010, 13 UTC (left) and 18 UTC (right)

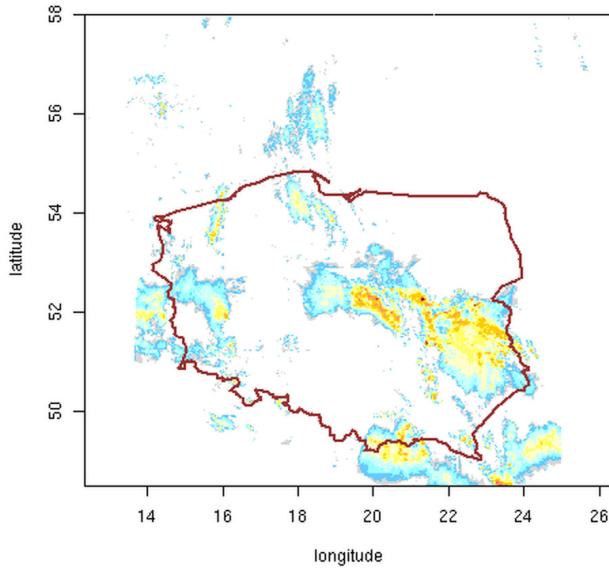


Fig. 3. 1-hour precipitation accumulation observed in Poland, 6 August 2010, 17 UTC

by the front approached Poland, an inflow of Atlantic air began. The observed strong wind convergence resulted in the intensification of convective processes. In Figure 4 an example of the reflectivity field from the Brzuchania radar is presented.

Data from Brzuchania and Legionowo radars were assimilated separately. Similarly to the cases described above, the results of simulations for this event also turned out to be unsatisfactory.

#### *Case study of 16 August 2010*

There was a weak low pressure system over France and Germany, whereas a strong one was developing over northern Scandinavia. The latter was associated

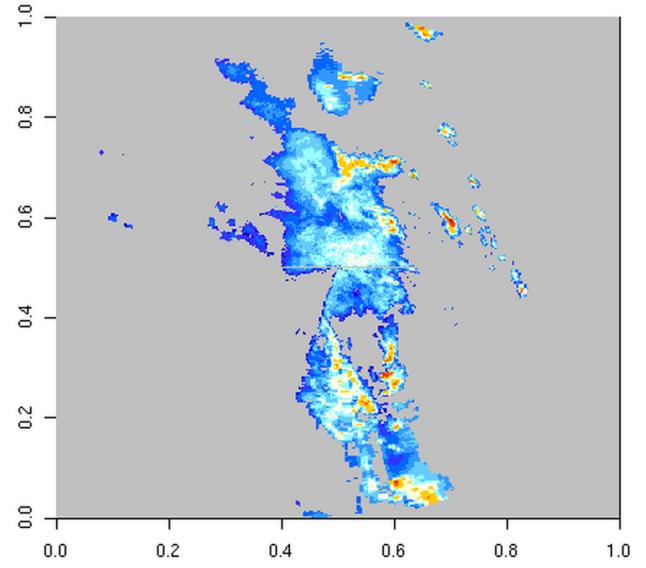


Fig. 4. Radar reflectivity, Brzuchania radar, 15 August 2010, 14:30 UTC, elevation 0.5°

with a cold front passing from the west. In Poland the cold front resulted in an extensive storm zone moving eastwards (Fig. 5).

In Figures 6 and 7, simulated reflectivity as maximum value in the vertical column is presented for two runs: without radar data assimilation (Fig. 6), and with radar data assimilation (Fig. 7), both for the same dates as in Figure 5.

Based on the experience gained during previous experiments for this case, the time period between the start of the model and the first assimilation of radar data was extended from two to five and a half hours. The data was assimilated for two hours with a 10-minute time step, i.e. 13 assimilation cycles were performed. Error covariance matrixes were estimated based on an ensemble of 13 fore-

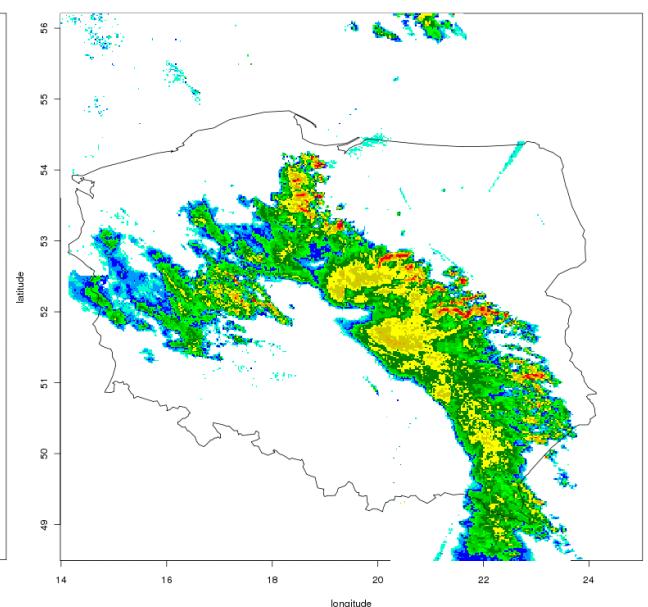
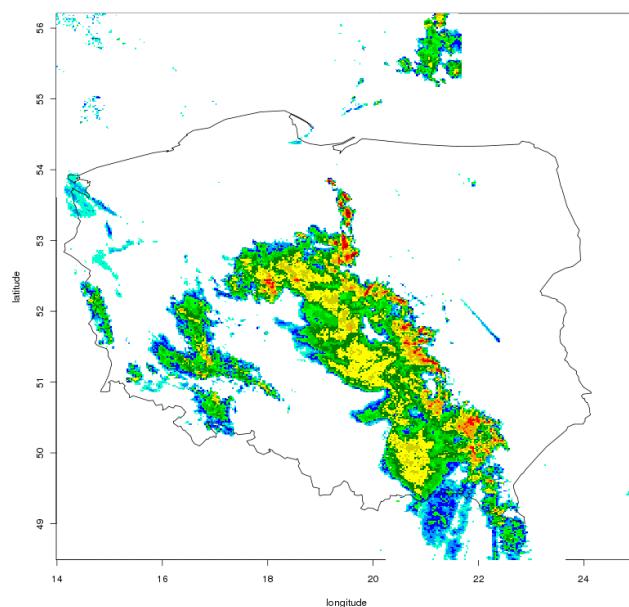


Fig. 5. Composite radar reflectivity on 500 m a.s.l. observed in Poland, 16 August 2010, 18 UTC (left) and 20 UTC (right)

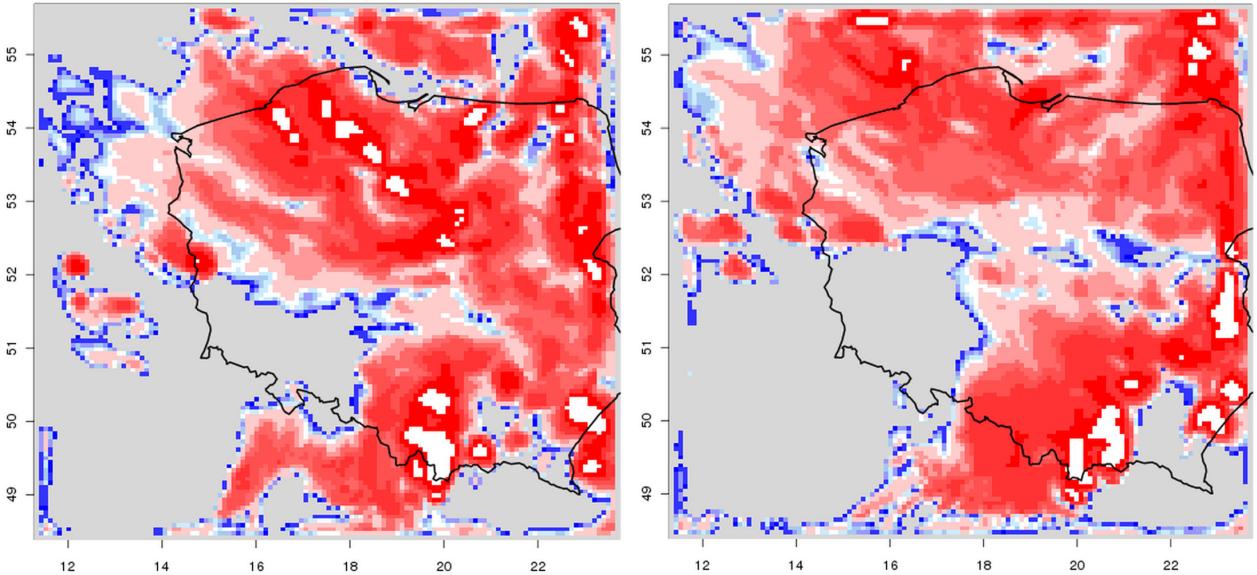


Fig. 6. Maximum reflectivity simulated for 18 UTC (left) and 20 UTC (right), simulation without radar data assimilation (control run), 16 August 2010

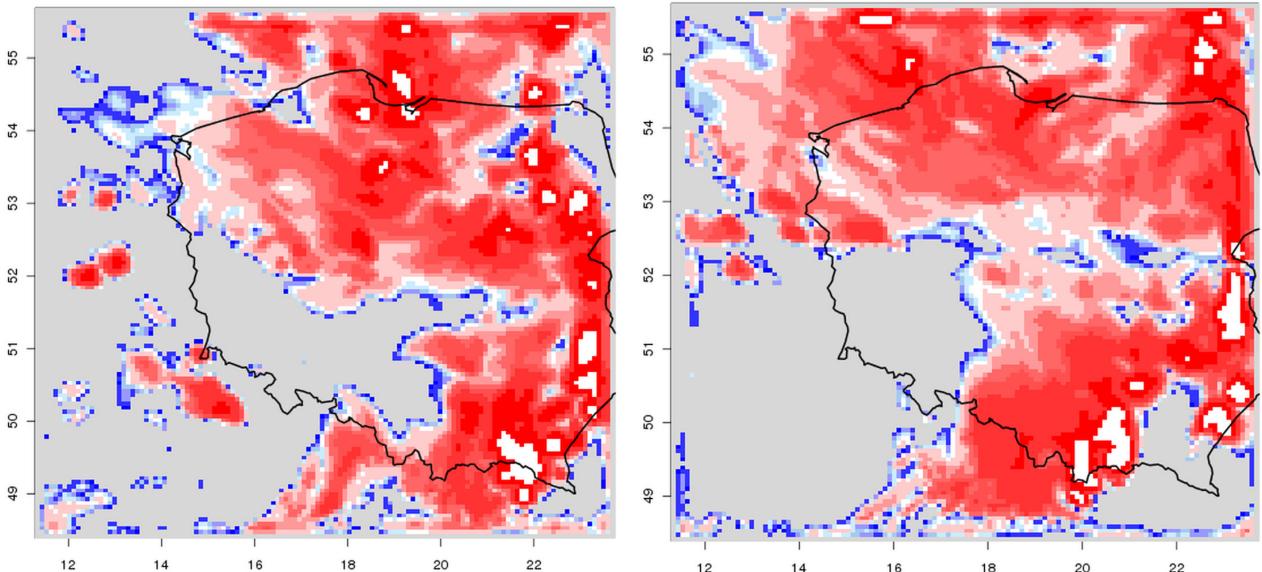


Fig. 7. Maximum reflectivity simulated for 18 UTC (left) and 20 UTC (right), simulation with radar data assimilation, 16 August 2010

casts. After the assimilation cycle, the forecast ensembles were still generated for two hours.

In Figure 7, two examples of simulations are depicted: for 18 UTC, after three cycles of data assimilation, and for 20 UTC, half an hour after the last assimilation cycle. The simulations with the assimilation were generated as an average of the 13 members of ensemble. In Figure 6, for the same time steps, simulations produced by the model without the assimilation as a control run are presented.

The assimilation of radar data influenced the pattern of the simulated radar reflectivity field. The observed differences between simulations with and without assimilation are significant, but local. It is noted that the simulation generated 30 minutes after the end of the assimilation process practically does not differ from the control run.

#### 4.3. Criteria for forecast quality

In order to evaluate the reliability of the presented assimilation algorithms, the observed radar reflectivity data are considered as benchmarks. For the validation of spatial data, a proper technique of spatial validation is needed – one which takes account of both magnitude- and distance-related errors (Venugopal et al. 2005; Ahijevych et al. 2009). Such metrics require that the forecasts are in approximate” agreement with the observations, meaning that they are close in space, time, and intensity (Ebert 2008).

The comparison between a forecast and observed radar image is performed employing different criteria. The first group is based on root mean square error *RMSE*. The applied derivatives are: *M-RMSE* (mean RMSE) – the *RMSE* calculated from mean values in 9 x 9 km grids around each pixel; and *A-RMSE* (area-related RMSE) –

the *RMSE* calculated for each pixel in 9 x 9 km grids around it, after sorting values in the grid in descending order, with the mean for the whole image then calculated (Rezacova et al. 2007).

Other quality criteria applied in the investigation are *BIAS*, correlation coefficient *R*, and standard deviation *s*. The *BIAS* is defined as the difference (in dBZ) between forecast and observed radar reflectivity for the whole computation domain. Moreover, standard deviations of the two fields are calculated: *sObs* and *sFc* for observed and forecast fields, respectively.

Different groups of quality criteria are categorical ones, based on a contingency table (Table 3), stating whether precipitation over a preset threshold was observed/forecast or not.

Table 3. Contingency table

Observation	Forecast	
	yes	no
yes	<i>a</i>	<i>b</i>
no	<i>c</i>	<i>d</i>

Quality criteria obtained from the contingency table employed in this work are listed below (Equations 1-6):

- percentage correct:

$$PC = \frac{a + d}{a + b + c + d} \cdot 100\% \quad (1)$$

- probability of detections:

$$POD = \frac{a}{a + b} \quad (2)$$

- probability of detection of no cases of rain:

$$DPOD = \frac{d}{d + c} \quad (3)$$

- False alarm ratio:

$$FAR = \frac{c}{a + c} \quad (4)$$

- Critical success index:

$$CSI = \frac{a}{a + b + c} \quad (5)$$

- Success ratio:

$$SR = \frac{a}{a + c} \quad (6)$$

#### 4.4. Statistical analysis of results of radar data assimilation

The evaluation of the quality of forecasts generated by the COAMPS model was made by comparing them to the radar data for the relevant time steps. The analyses for two out of four situations (case studies) described in section 4.2, on 10 May and 16 August 2010 are presented in this subsection. Other situations require further work in order to optimize the radar data assimilation scheme.

10 May 2010

Figure 8 shows examples of precipitation forecasts with and without the assimilation of radar data and precipitation observed by the Legionowo radar on 10 May 2010, at 13 UTC.

Figure 9 shows the dependence of the selected quality parameters *A-RMSE* and *BIAS* values for forecasts with and without radar data assimilation on lead-times up to 2 hours.

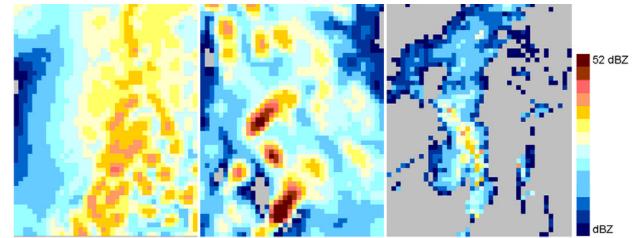


Fig. 8. Example of radar data assimilation to COAMPS model: Legionowo radar, 10 May 2010, run at 12 UTC, forecast for 13 UTC. From the left: forecast without assimilation, forecast with assimilation, radar image

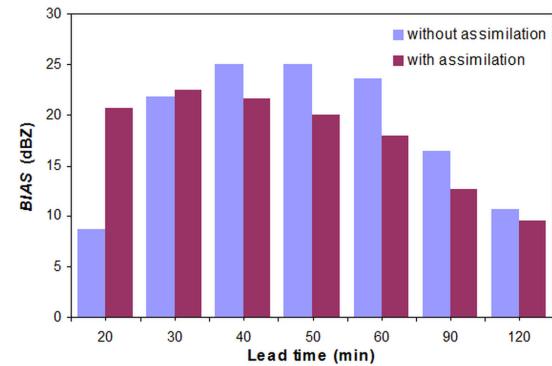
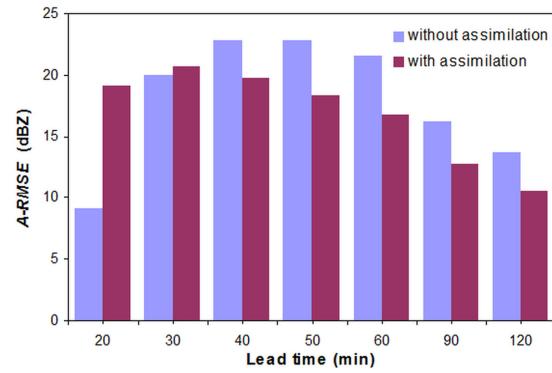


Fig. 9. Results of quality analysis without and with radar data assimilation. On the top A-RMSE (dBZ), on the bottom BIAS (dBZ): range of Legionowo radar, 10 May 2010, run of COAMPS model at 12 UTC

The detailed results of analyses of the forecast quality with and without assimilation in comparison with radar data for relevant time steps are summarized in Table 4.

Generally, *A-RMSE* and *BIAS* values indicate an improvement in forecast quality after radar data assimilation. A more detailed analysis shows that the model without assimilation overestimated the precipitation field, while the assimilation slightly decreased this effect (a smaller

Table 4. Quality statistics for forecasts generated for 10 May 2010, model COAMPS started at 12 UTC with lead time to 120 minutes

Lead time (min)	M-RMSE (dBZ)	A-RMSE (dBZ)	BIAS (dBZ)	R	sObs (dBZ)	sFc (dBZ)	PC	POD	DPOD	FAR	CSI	SR
Forecasts without assimilation												
30	22.65	19.93	21.81	0.05	2.86	3.79	24.18	1.00	0.02	0.77	0.23	0.23
40	25.68	22.76	24.96	0.12	2.93	3.59	23.59	1.00	0.02	0.78	0.22	0.22
50	25.40	22.80	25.02	0.15	3.00	3.65	23.80	1.00	0.02	0.78	0.22	0.22
60	24.54	21.53	23.61	0.14	3.10	3.92	25.49	0.99	0.03	0.76	0.24	0.24
90	26.46	16.22	16.41	0.04	3.30	5.76	28.84	0.91	0.10	0.77	0.23	0.23
120	28.75	13.66	10.65	0.11	3.39	7.27	40.35	0.68	0.32	0.77	0.20	0.23
Forecasts with assimilation												
30	26.94	20.67	22.53	0.26	2.86	5.08	28.53	0.99	0.08	0.76	0.24	0.24
40	27.42	19.72	21.58	0.26	2.93	5.06	26.73	1.00	0.06	0.77	0.23	0.23
50	27.39	18.31	19.95	0.27	3.00	5.45	27.22	1.00	0.07	0.77	0.23	0.23
60	26.74	16.71	17.97	0.26	3.10	6.08	29.76	0.98	0.09	0.75	0.25	0.25
90	27.18	12.78	12.72	0.22	3.30	7.35	46.85	0.92	0.34	0.71	0.28	0.29
120	26.04	10.59	9.56	0.18	3.39	7.59	54.64	0.75	0.49	0.70	0.27	0.30

value of *BIAS*). As a result of assimilation, a more spatially changeable field of precipitation was obtained (higher value of standard deviation *sFc*) (see Fig. 8), in particular, overestimating the reflectivity values in cases of the most intense rainfall. High values of *POD* and *FAR* also indicate too large an area of predicted reflectivity field.

The quality of forecasts measured by correlation coefficient *R* remains at a low level, although the *R* was almost doubled after taking assimilation into account. This shows large errors in predicting the location of precipitation – this is slightly improved after assimilation. The decrease in *A-RMSE* error also suggests an improvement in the location of forecast precipitation with high intensity.

Concluding, the model COAMPS, without assimilation, poorly predicted convective precipitation. The assimilation of radar data resulted in a slight improvement in the quality of the forecasts for lead-times ranging from 0.5 to 2 hours, but this improvement is not satisfactory.

#### 16 August 2010

Figure 10 shows examples of precipitation forecasts (with and without assimilation of radar data) in comparison with rainfall radar observed in Legionowo on 16 August 2010.

Figure 11 presents the dependence of selected quality parameters – *A-RMSE* and *BIAS* – for forecasts on 16 August 2010, with and without radar data assimilation, generated for lead-times up to 9 hours.

The results of statistical analyses of the forecast quality with and without assimilation, in comparison with radar data for relevant time steps, are presented in Table 5, applying the same quality criteria as for the previous case study.

Generally, the results obtained for this case are similar to the previous ones (Table 4), but the influence of radar data assimilation is observed for a longer time – it lasts about 5-6 hours.

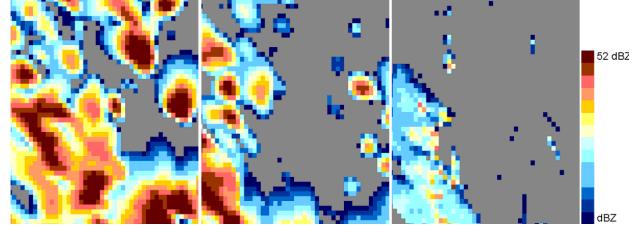


Fig. 10. Example of radar data assimilation to COAMPS model: Legionowo radar, 16 August 2010, run at 12 UTC, forecast for 15 UTC. From the left: forecast without assimilation, forecast with assimilation, radar image

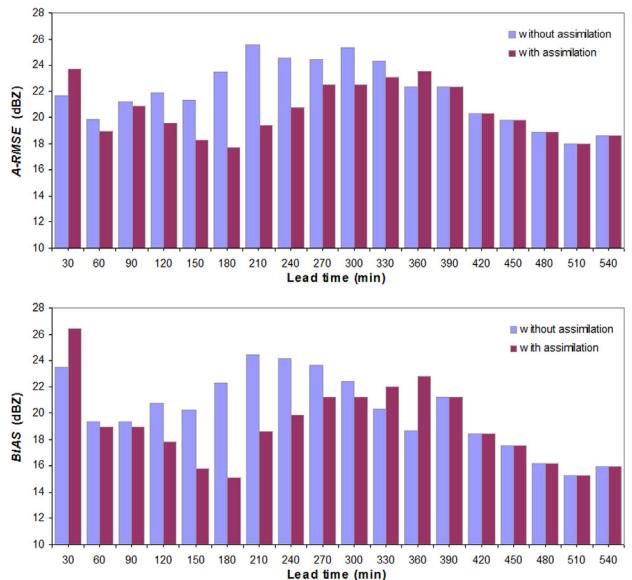


Fig. 11. Results of quality analysis without and with radar data assimilation. On the top A-RMSE (dBZ), on the bottom BIAS (dBZ); range of Legionowo radar. 16 August 2010, run of COAMPS model at 12 UTC

A-RMSE and BIAS show slightly better improvement in the accuracy of forecasts with radar data assimilation than in the previous case, but the error values are still high (Fig. 11). The values of the correlation coefficient R are very low, which indicates that the model COAMPS, in the employed configuration of the model, poorly predicted the location of intense convective precipitation

## 5. Summary

The research work presented in the paper concentrated on the investigation of the effectiveness of weather radar data assimilation into the numerical weather prediction model COAMPS. A filtering approach with the DART tool was employed to modify the initial description of the atmospheric state simulated by the model – this was done by introducing new information derived from the radar observations to make it more reliable.

The proposed methodology was tested on four selected case studies for convective precipitation. The experiments with the assimilation of radar reflectivity to NWP models did not lead to unambiguous conclusions. The accuracy of forecasts with radar data assimilation was slightly improved, however the quality of both forecasts, with and without assimilation, was not satisfactory.

Work on including the observation errors, forecast errors, and model errors in an Ensemble Kalman Filter is a very challenging research field. This method has been applied in a very limited way so far. The evolving of the radar data assimilation scheme for NWP models for practical applications requires a great deal of work and effort, both in a technical sense and also on algorithm improvement.

The major technical limitation of the developed algorithm is its requirement for a high level of computing power. A larger number of processors would enable parallel filtering and the generation of at least a dozen forecasts at a time – which would greatly reduce the time needed for computation.

However these limitations and unsatisfactory results should not discourage further work on radar data assimilation in numerical weather prediction models. It seems that the filtering approach has the potential to ingest radar observations and, in consequence, improve the description of initial atmospheric states for numerical modelling.

## Acknowledgements

The work described in the paper was performed in the frame of the project N N307 467738, granted by the National Science Centre (Poland). Weather radar data from the POLRAD network were provided by IMGW-PIB, and the COAMPS model was made available by ICM UW. Algorithms for the data correction and quality characterization were partially developed within the BALTRAD+ project (“An advanced weather radar network for the Baltic Sea Region: BALTRAD”), in the frame of Baltic Sea Region Programme 2007-2013.

## Bibliography

- Ahijevych D., Gilleland E., Brown B.G., Ebert E.E., 2009, Application of spatial verification methods to idealized and NWP-gridded precipitation forecasts, *Weather Forecasting*, 24, 1485-1497
- Arakawa A., Lamb, V.R., 1977, Computational design of the basic dynamical processes of the UCLA general circulation model, *Methods in Computational Physics*, 17, Academic Press, 173-265
- Bech J., Gjertsen U., Haase G., 2007, Modelling weather radar beam propagation and topographical blockage at northern high latitudes, *Quarterly Journal of Royal Meteorology Society*, 133 (6), 1191-1204, DOI: 10.1002/qj.98
- Bishop C., Etherton B., Majumdar S., 2001, Adaptive sampling with the ensemble transform Kalman filter. part I: Theoretical aspects, *Monthly Weather Review*, 129, 420-435
- Collier C.G. (eds.), 1989, Applications of weather radar system. A guide to uses of radar data in meteorology and hydrology, Ellis Horwood Limited, New York, 294 pp.
- Ebert E.E., 2008, Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework, *Meteorological Applications*, 15, 51-64, DOI: 10.1002/met.25
- Einfalt T., Szturc J., Ośródka K., 2010, The quality index for radar precipitation data: a tower of Babel?, *Atmospheric Science Letters*, 11 (2), 139-144, DOI: 10.1002/asl.271
- Fletcher N.H., 1962, The physics of rainclouds, Cambridge University Press, New York, 386 pp.
- Hodur R.M., 1997, The Naval Research Laboratory Coupled Ocean-Atmosphere Mesoscale Prediction System (COAMPS), *Monthly Weather Review*, 125, 1414-1430
- Jakubiak B., 2008, Data assimilation experiments with ensemble Kalman filter, *Proceedings of the Joint MAP D-PHASE Scientific Meeting COST 731 mid-term seminar: Challenges in hydrometeorological forecasting in complex terrain*, 19-22 May 2008, Conference Centre of CNR, Bologna, Italy, 159-163
- Jakubiak B., Szpindler M., 2009, Real data assimilation experiments using filtering methods, *Geophysical Research Abstracts*, 11, EGU2009-11675
- Jazwinski A.H., 1970, Stochastic processes and filtering theory, Academic Press, New York
- Jones C.D., Macpherson B., 1997, A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model, *Meteorological Applications*, 5, 1-16
- Jurczyk A., Ośródka K., Szturc J., 2008, Research studies on improvement in real-time estimation of radar-based precipitation in Poland, *Meteorology and Atmospheric Physics*, 101 (3-4), 159-173, DOI: 10.1007/s00703-007-0266-3
- Kalman R.E., 1960, A new approach to linear filtering and prediction problems, *Transactions of the ASME – Journal of Basic Engineering*, 82 (series D), 35-45
- Kessler E., 1969, On the distribution and continuity of water substance in atmospheric circulations, *Meteorology Monograph No. 32*, American Meteorological Society, 84 pp.
- Khairoutdinov M., Kogan Y., 2000, A new cloud physics parameterization in a large-eddy simulation model of marine stratuscumulus, *Monthly Weather Review*, 128 (1), 229-243

- Klemp J.-B., Williamson R.-B., 1978, The simulation of three-dimensional convective storm dynamics, *Journal of Atmospheric Science*, 35, 1070-1096
- Lin Y.-L., Farley R.D., Orville H.D., 1983, Bulk parameterization of the snow field in a cloud model, *Journal of Climate and Applied Meteorology*, 22, 1065-1092
- Marshall J.S., Palmer W.M., 1948, The distribution of raindrops with size, *Journal of Meteorology*, 5, 165-166
- Ośródka K., Szturc J., Jurczyk A., Michelson D., Haase G., Peura M., 2010, Data quality in the BALTRAD processing chain, ERAD 2010 – 6th European Conference on Radar in Meteorology and Hydrology, 355-361
- Ośródka K., Szturc J., Jurczyk A., 2014, Chain of data quality algorithms for 3-D single-polarization radar reflectivity (RAD-VOL-QC system), *Meteorology Applications*, 21, 256-270
- Rezacova D., Sokol Z., Pesice P., 2007, A radar-based verification of precipitation forecast for local convective storms, *Atmospheric Research*, 83, 211-224
- Rutledge S.A., Hobbs P.V., 1983, The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. XII: A diagnostic modeling study of precipitation development in narrow cold-front rainbands, *Journal of Atmospheric Science*, 41, 2949-2972
- Sun J., Crook N.A., 1997, Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: Model development and simulated data experiments, *Journal of Atmospheric Sciences*, 54, 1642-1661
- Sun J., Crook N.A., 1998, Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part II: Retrieval experiments of an observed Florida convective storm, *Journal of Atmospheric Sciences*, 55, 835-852
- Szturc J., Michelson D., Koistinen J., Haase G., Peura M., Gill R., Sørensen M., Ośródka K., Jurczyk A., 2012a, Data quality in the BALTRAD+ project, *Proceedings of 7th European Conference on Radar in Meteorology and Hydrology ERAD 2012*, Toulouse, France (CD)
- Szturc J., Ośródka K., Jurczyk A., 2012b, Quality control algorithms applied on weather radar reflectivity data, [in:] *Doppler Radar Observations – Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications*, J. Bech and J.L. Chau (eds.), InTech, 289-306
- Venugopal V., Basu S., Foufoula-Georgiou E., 2005, A new metric for comparing precipitation patterns with an application to ensemble forecasts, *Journal of Geophysical Research*, 110, D08111, DOI: 10.1029/2004JD005395
- Whitaker J., Hamill T., 2002, Ensemble data assimilation without perturbed observations, *Monthly Weather Review*, 130, 1913-1924
- Zejdlik T., Novak P., 2010, Frequency Protection of the Czech Weather Radar Network, *Proceedings of ERAD 2010* (access online)