COSMO-EPS results for Poland with ANN-based calibration coupled with space-lag correlation application

A. Mazur[1], G. Duniec[1]

1.Institute of Meteorology and Water Management – National Research Institute, Warsaw, Poland
61 Podlesna str., PL-01-673 Warsaw, Poland

1 Abstract

The results from research on COSMO-EPS, carried out at IMWM, are presented. The operational EPS set-up is based on perturbations of soil surface-area index of the evaporating fraction of grid points over land. Long-term evaluation results of different methods of EPS-post-processing. As a general rule, using Artificial Neural Network (ANN) values of EPS mean are significantly closer to observation of air temperature/dew point temperature or wind speed than those calculated as simple average or Multi-linear Mean. In turn, the application of the Ensemble Prediction System in convection-permitting scale based on time-lagged ICs/BCs allows to improve these forecasts, especially due to the removal of false alarms. The research was carried out using archive data, starting from 2015. The noteworthy correlation between forecasts (ensemble means) and observations was established in this research.

2 Introduction

Extensive tests conducted during the COTEKINO Priority Project proved that small perturbations of selected soil parameter were sufficient to induce significant changes in the forecast of the state of atmosphere and to provide qualitative selection of a valid member of an ensemble *(Duniec and Mazur, 2014)*. Changes of c_soil^) had a significant impact on values of air temperature, dew point temperature and relative humidity at 2m agl., wind speed/direction at 10m agl., and surface specific humidity *(ibidem)*. The usage of an idea of time-lagged initial and boundary conditions allowed obtaining a valid ensemble and using it efficiently in an operational mode. Further work is intended to focus on "tuning" ensemble performance and to provide quantitative quality scores. For this purpose the random number generator combined with perturbations of initial soil surface temperature and the dependence of amplitude of perturbation on soil type will be implemented in the COSMO model. While the set of equally weighted time-lagged forecasts improve short-range forecasts, the further progress may also be sought by adopting a regression approach to compute set of weights for different time-lagged ensemble members. EPS runs operationally at IMWM since January, 2016. It covers 4 runs/day, with 48 hours forecasts, 20 members/4 groups (using Time-lagged Ics/BCs; see Duniec G. et al. (2016); conf. Fig.1 below). Amplitude of perturbation of c_soil depends on type of soil (clay, sand, peat etc). [1]

doi:10.5676/dwd_pub/nwv/cosmo-nl_20_05

[1]*) surface-area index of the evaporating fraction of gridpoints over land

Table 1: Deterministic model(s) – source of ICs/BCs for operational EPS

Model	Grid size NxMxL	Forecast length(h)	Resolution(km)
ICON (DWD)	2949120 triangles	78	13
COSMO v. 5.01	415x460x40	13	7
COSMO v. 5.01*	380x405x50	78	2.8

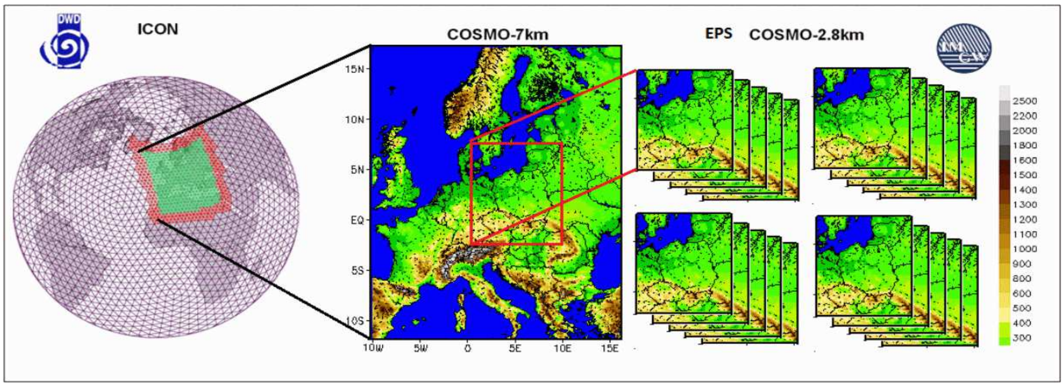


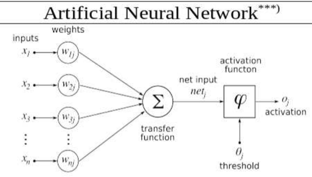
Figure 1: EPS operational configuration

Details of the deterministic models configuration are as follows:

3 Results – comparison of results for three methods of post-processing.

Forecasts of air temperature and dew point temperature at 2m agl., surface pressure and windspeed at 10m agl., as well as other fields are available. As a result, plots/chart of EPS mean, spread, probabilities of threshold exceedance are prepared in the routine manner. Results are subsequently stored for further research (e.g. skill-spread relation) Results of EPS forecasts are subsequently calibrated. Three basic methods of calibration were examined as shown in Table 2 – simple arithmetic mean (SM), multilinear regression mean (MLR) and artificial neural network mean (ANN).

Table 2. Ensemble calibration – Simple Mean (SM) vs. multilinear regression (MLR) mean vs. ANN mean

Simple Mean*)	Multilinear regression **)	Artificial Neural Network***)
$y = \frac{\sum_{i=1}^{m} x_i}{m}$	$\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & \dots & x_{1p} \\ x_{21} & \dots & x_{2p} \\ \dots & \dots & \dots \\ x_{n1} & \dots & x_{np} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_p \end{pmatrix}$	
x – forecast values, *y* – ensemble mean, *m* – # members	y (corrected forecasts) – ensemble mean, x – matrix of raw forecast values, β – weights (from previous fcsts.)	24 input neurons (20 members + λ,φ + ts,tc) 5 neurons in a single hidden layer activation function: hyperbolic tangent

*) Simple avg. – arithmetic mean, m=20 members;

**) # of predictors: n=24=20 members+geo.coords+lead/fcst.time;

***) Trained on data from July 2016 to November 2018, tested on data of 2018

A new element that was introduced as part of the research work was the assessment of the suitability of the space-lag (cross) correlation method in relation to verification against measurements at SYNOP stations or the 'PERUN' Polish lightning detection network.

The basis assumption here was that the spatial distributions of forecast meteorological quantities may be very similar in shape and values to real distributions (from observations), but at the same time shifted in space (Fig. 2).

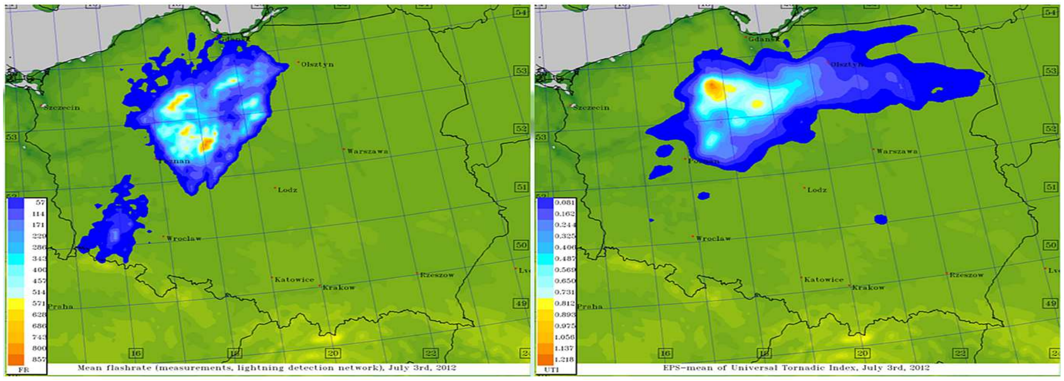


Figure 2: Example spatial distribution of detection (left) and forecast (right) of lightning

A method of determining the Vector of Displacement (VOD) for verification of various types of operational data has been proposed. Because the lightning detection network provides the ability to determine the occurrence of a lightning discharge with high accuracy, in this case the VOD is calculated as a vector between the two "centers of masses" of the forecasted and actual lightning distributions. An example of the effect of such an operation is presented in Figure 3.

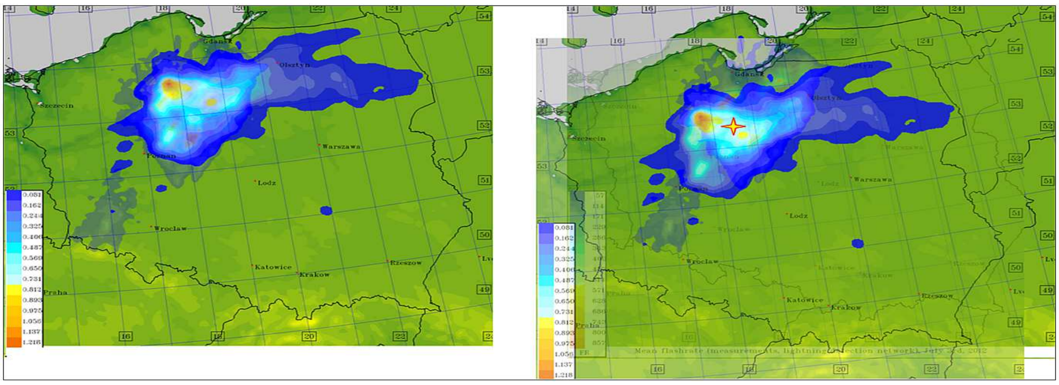


Figure 3: Forecast and detection of lightnings before (left) and after shifting using VOD (right). The centers of masses of both distributions are marked with an asterisk.

For quantities verified against measurements at SYNOP stations (whose number is not large), a different VOD calculation method was used to apply the procedure with mass center calculation.

Table 2: Test ANN for entire year 2018, in comparison with arythmetic mean (AM) and multiline regression method (MLR)

Method	ME	MAE	RMSE	ME	MAE	RMSE	ME	MAE	RMSE
	Dew point			Air temperature			Windspeed		
AM	0.353	1.759	2.431	0.832	2.641	3.623	-0.551	1.837	2.427
MLR	-0.330	1.966	2.321	0.653	2.412	3.112	0.128	1.623	1.901
ANN	-0.212	1.911	2.210	0.046	2.013	2.905	-0.051	1.236	1.557

1. At all stations in a specific environment (red circle, Figure 4) find the grid (with coordinates x, y, horizontal arrow) in which the predicted value of a given size is the to closest measured at the station (xs,ys, vertical arrow)
2. Calculate the VOD for a single station as (x-xs,y-ys); red arrow, left panel.
3. Calculate the average VOD for all stations (red arrow, middle and right panel)
4. "Shift" (displace) the forecast values by the calculated VOD

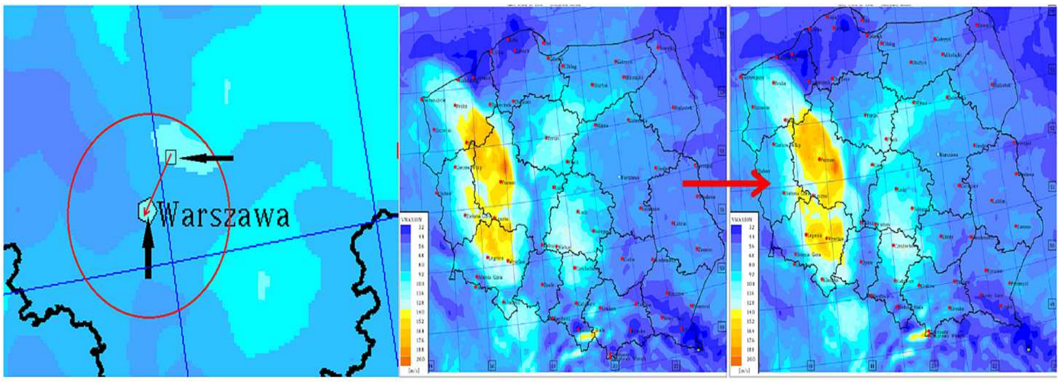


Figure 4: Determining the VOD for a synoptic station (left panel); average shift vector (middle and right)

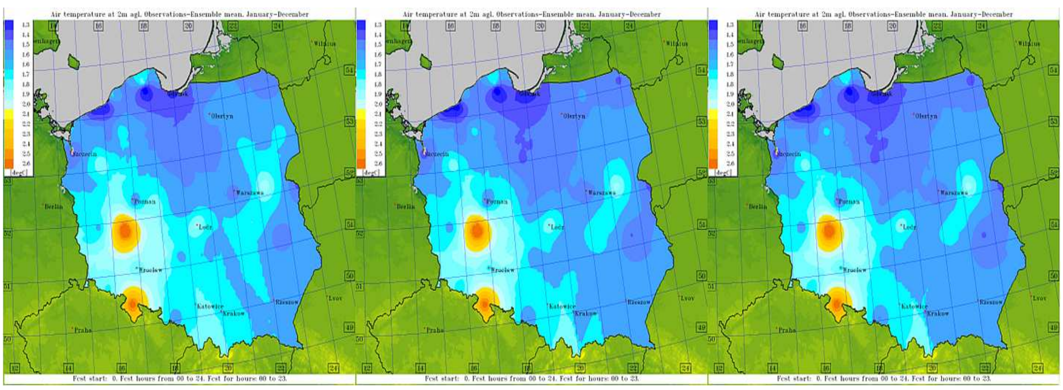


Figure 5: Results of different methods of calculating the average from ensemble forecasts. Left – observations vs. AM. Middle – observations vs. MLR average. Right – observations vs. average ANN. Air temperature forecasts, mean values for 2018.

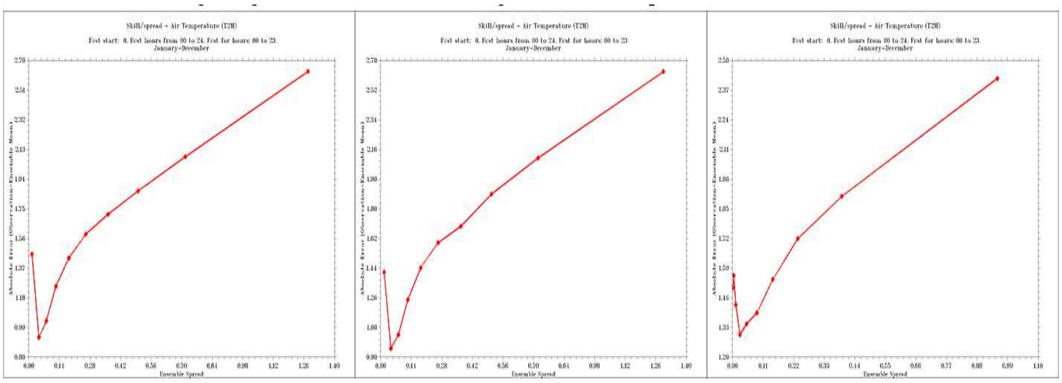


Figure 6: Skill-spread relationship. Left – arithmetic average, middle – MLR average, right – ANN average. Air temperature forecasts, average values for 2018.

Based on the research, it can be concluded that the use of the space-lag correlation procedure improves the forecast values related to measurements. The following figures show examples of the skill value distribution of air temperature, dew point temperature and visibility without using VOD (average for the entire period) and using this method.

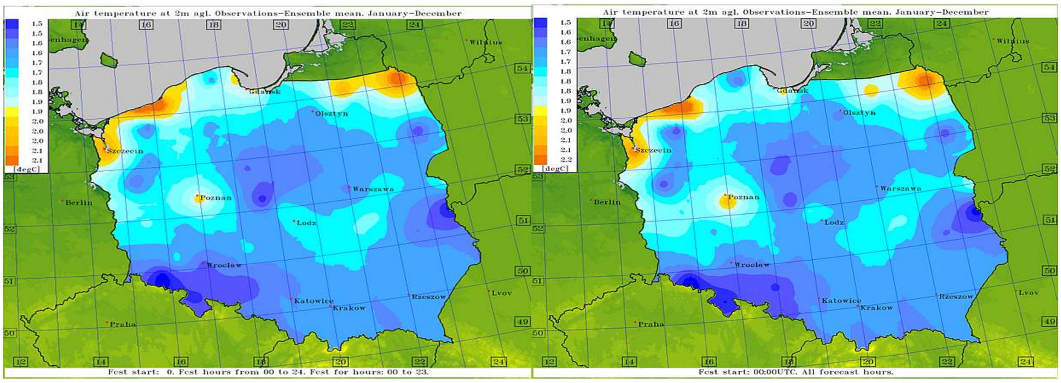


Figure 7: Distribution of skill values of air temperature, left – no VOD used, right – VOD applied, average for the period 2011-2013.

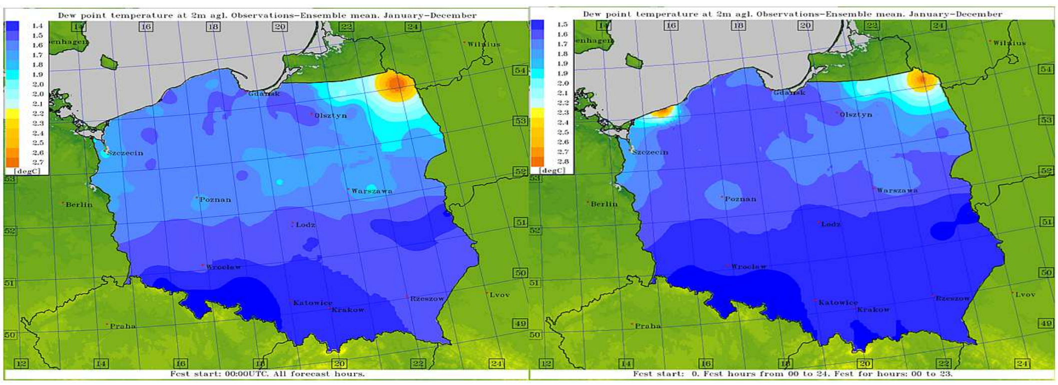


Figure 8: Distribution of skill values of dew point temperature, left – no VOD used, right – VOD applied, average for the period 2011-2013.

5 Conclusions

From an assessment of calibration quality – calculation of the ensemble mean skill and spread, using the neural networks ANN and MLR multilinear regression compared to the usual mean (arithmetic) after the team proved that ANN method was definitely useful both in operational and diagnostic work.

The procedure utilizing the calculated VOD indeed improves forecasts. This pertains to both skill (MAE) and spread (less underdispersivity). With the use of values (prior to the forecasts lead hour) from SYNOP/radar or lightning detection network the computations can be done automatically and relatively easy.

Acknowledgements

The results in the poster/oral form was presented during the 41st EWGLAM/26th SRNWP Workshop in Sofia, Bulgaria, October 2019. The activity and participation in the workshop was reimbursed from COSMO funds.

References

[1] Duniec, G., Interewicz, W., Mazur, A. and Wyszogrodzki, A. (2016): Operational Setup of the COSMO-based, Time-lagged Ensemble Prediction System at the Institute of Meteorology and Water Management – *National Research Institute. Met. Hydrol. Water Manage. (2017) vol. 5; (2)*: **43-51**.

[2] Mazur, A. and Duniec, G. (2017): SPRED PP activities at IMWM-NRI. Presented at COSMO GM, Jerusalem, Israel.

[3] Mazur, A., Duniec, G. and Interewicz, W. (2018): Introductory activities in PP APSU at IMWM-NRI and results of ANN post-processing of EPS forecasts. Presented at COSMO GM, Sankt Petersburg, Russia.

[4] Mazur A. and Duniec G. (2019): Operational- and research- asessment of EPS running at IMWM in the frame of APSU PP. Presented at COSMO General Meeting (WG7 parallel session) in Rome, Italy.