Consortium

for

Small-Scale Modelling

Newsletter

December 2020 *No. 20*

Deutscher Wetterdienst

Ufficio Generale Spazio Aereo e Meteorologia

Instytucie Meteorogii i Gospodarki Wodnej

Agenzia Regionale per la Protezione Ambientale dell'Emilia Romagna Servizio Idro Meteo Clima

Centro Italiano Ricerche Aerospaziali

Israel Meteorological Service

MeteoSwiss

ΕΘΝΙΚΗ ΜΕΤΕΩΡΟΛΟΓΙΚΗ ΥΠΗΡΕΣΙΑ

Administratia Nationala de Meteorologie

Federal Service for Hydrometeorology and Environmental Monitoring

Amt fur GeoInformationswesen der Bundeswehr

Agenzia Regionale per la Protezione Ambientale del Piemonte

www.cosmo-model.org

Editor: Mihaela Bogdan(NMA)

Publisher

COSMO
Consortium for Small Scale Modelling
cosmo-model.org

Editor

Mihaela Bogdan, NMA
mihaela.bogdan@meteoromania.ro

Contributions to COSMO Newsletter No. 20 have DOIs.
These are indicated on the title page of each contribution.
The DOI format is 10.5676/dwd_pub/nwv/cosmo-nl_20_NN, where NN is the contribution number.

Table of Contents

Consortium for Small-Scale Modelling mounts further efforts towards replacing the limited-area model COSMO with the Limited Area Mode of the ICON modelling framework ICON (ICON-LAM). Considerable progress is made.First-hand information is available at the web page of the priority project C2I (http://www.cosmo-model.org/content/tasks/priorityProjects/c2i), see also the PP C2I presentation at the 22nd COSMO General Meeting (http://www.cosmo-model.org/content/consortium/generalMeetings/general2020/plenary/C2I_GM20_Overview.pdf).

As regards the COSMO Model, recall that the version 6.0 will be the last official COSMO-Model release (expected in early 2021). Beyond the version 6.0, only the maintenance of the COSMO code, including bug fixes, will be performed over some years to come. The COSMO documentation (with DOIs) will be updated and "frozen" together with the model code. Meanwhile the documentation of the COSMO version 5.05 with DOIs has become available. It can be downloaded from the DWD landing page, https://www.dwd.de/EN/ourservices/cosmo_documentation/cosmo_documentation.html, or from the COSMO web page, http://cosmo-model.org/content/model/documentation/core.

An important step of legal order was made. The development partnership agreement between DWD and COSMO was signed in March 2020. The contractual partners of the cooperation agreement for the ICON Modelling Framework, viz., DWD, MPI-M, KIT, and DKRZ, granted "the participating national meteorological services as well as other major members of COSMO (see http://cosmo-model.org/) the right, which is non-exclusive, free of charge and unlimited in time, to use the ICON Modelling Framework for non-commercial Research and Development and for their Official Duty purposes". The liaisons between the COSMO consortium and the ICON community go through DWD that has the responsibility to ensure that the development partnership between DWD and COSMO is in accordance with the cooperation agreement between the ICON contractual partners.

Three new priority projects and one new priority task are launched (start in September 2020). These are PP KENDAScope ("KENDA from Surface to Cloud Observations Progressive Extension", implemented within the framework of Working Group 1, PP leader is Christoph Schraff, DWD), PP MILEPOST ("MachIne LEarning-based POST-processing", WG4, Andrzej Mazur, IMGW), PP PROPHECY ("PRObabilistic Prediction at High-resolution with EnhanCed perturbation strategY", WG7, Chiara Marsigli, DWD), and PT VAINT ("Vegetation Atmosphere INTeractions", WG3b, Merja Tölle, University of Kassel, Germany). Detailed information about the above PPs and PT is available at the COSMO web (see http://www.cosmo-model.org/content/tasks/priorityProjects and http://www.cosmo-model.org/content/tasks/priorityTasks).

Because of the COVID-19 pandemic, the 22nd COSMO General Meeting went online. It was held over two weeks, from 31 August through 11 September 2020. The WG, PP and PT meetings were held 1-7 September, followed by the plenary sessions 8-11 September. In this way, the overlap between the WG, PP and PT meetings was largely avoided, and numerous COSMO scientists were able to attend several meetings (which ran mostly in parallel on Monday and Tuesday afternoon during the previous in-person COSMO General Meetings). The online WG, PP and PT meetings proved to be rather efficient and were generally a success. Admittedly, this was not quite the case for the plenary sessions, where the interaction between the plenary speakers and the audience was not sufficiently intensive (very few or no questions, etc.). This problem associated with the online plenary-session format seems difficult to overcome. On the other hand, the final discussion was remarkably vivid and fruitful.

doi:10.5676/dwd_pub/nwv/cosmo-nl_20_01

Also worthy of mention is the STC meeting whose second part was held on Thursday (September 10) after the plenary sessions. This is at variance with the previous STC meetings (during the COSMO General Meetings) that were held on Monday and Tuesday, i.e. prior to the plenary sessions. Holding the second part of the STC meeting after the plenary sessions is very advantageous. Since the STC members attend the plenary sessions, the SPM should not reiterate, in his report to the STC, what is conveyed by the WG co-ordinators and the PP and PT leaders during their overview plenary talks. This makes the STC meeting shorter and more targeted and generally more efficient. Note that the SMC also took advantage of the online format and organized their meeting after the COSMO GM (the SMC meeting took place 21 September and 5 October). All in all, the online COSMO General Meeting 2020 appeared to be rather successful. The organization of the online GM took a great deal of effort, however; arguably, it took somewhat too much effort given the outcome. Improvements are clearly required should the next COSMO General Meeting go online. During the teleconference held on October 12, the STC reviewed the organization process of the virtual COSMO GM 2020 and summarized the lesson learnt.

Finally, I would like to thank the colleagues who contributed to the current issue of the COSMO Newsletter.

Dmitrii Mironov
COSMO Scientific Project Manager

Calibration of high resolution COSMO model over Switzerland: CALMO-MAX results

Voudouri A.[1], Carmona I.[2], Kaufmann P.[3], Avgoustoglou E.[1], Levi Y.[2], E. Bucchignani [4] and J.M.Bettems [3]

1 Hellenic National Meteorological Service
2 Israel Meteorological Service
3 MeteoSwiss
4 CIRA
e-mail: antigoni.voudouri@hnms.gr

1 Introduction

One of the main goals of the CALMO-MAX priority project is a robust implementation of the objective calibration method originally performed by Bellprat et al., 2012a and 2012b on COSMO-CLM and later by Voudouri et al., 2017 and 2018 on COSMO-NWP. More specifically, in this work, the calibration procedure has been applied to a fine horizontal resolution of 0.01o (approximately 1km) over a mainly continental domain covering the Alpine Arc. The CALMO methodology aims at substituting expert tuning, by which free or poorly confined model parameters are tuned using mainly expert knowledge (Duan et al., 2006; Skamarock, 2004;Bayler et al., 2000), with a more replicable and automatic approach. This methodology optimizes an overall model performance score by adjusting the values of a set of unconfined model parameters. The core of the calibration process is the determination of the metamodel (model emulator), which represents (using a simple mathematical function) the dependency of some representative model fields on the selected model parameters. The mathematical function at the core of the metamodel is constrained by a set of full model simulations over a time period long enough to represent the variability of the atmospheric conditions. Once fully specified, the metamodel supports a fast sampling of the parameter space to find an optimal combination of the model parameters. Detailed description of the procedure is available in Khain et al. (2015, 2017) and Voudouri et al. (2017a).

In the present work, the calibration is performed using a set of five unconfined model parameters. The selection of parameters is constrained by the fields used in the overall performance score, which should be sensitive to the chosen parameters. Because the overarching goal of the calibration in this project is to improve the quality of daily operational forecasts, the fields considered in the performance score are meteorological quantities used by bench forecasters, such as minimum and maximum 2m temperature, precipitation and wind speed. Although the number of parameters is limited, the main parameterization schemes affecting turbulence, soil-surface exchange and radiation are represented by these parameters. It is worth mentioning at this point that a strong dependency of the parameters optimum on the time of the year has been observed, as described in Voudouri et al. (2019); this reflects the dependency of the optimum on the atmospheric flow (or weather pattern). For this reason, if the primary goal of the calibration is to improve the daily operational forecast with a unique set of parameters, a climatologically representative set of weather patterns should be used in the calibration.

The steps followed, such as model setup and selection of parameters, are briefly described in Section 2, while in Section 3 results of CALMO-MAX applied over Switzerland are discussed. Conclusions are given in Section 4 and perspectives on further developments are summarized in Section 5.

doi:10.5676/dwd_pub/nwv/cosmo-nl_20_02

2 Data and Methodology

2.1. Model setup

COSMO model was computed for the years 2013 and 2017, with a horizontal resolution of 0.01o (approximately 1km) over a domain including the Alpine Arc (in particular the wider area of Switzerland and Northern Italy), in hindcast mode (in particular no data assimilation active). The grid extends vertically up to 23.5 km (30hPa) with 80 model levels. Initial and boundary fields for all tests are provided by the MeteoSwiss operational forecasting archive system at 0.01ohorizontal resolution (1km). Note also that the soil history is considered for all the CALMO-MAX simulations, and a prior 3 years soil spin up has been computed using terra standalone (TSA). The code used is the refactored version of the COSMO model (Lapillonne and Fuhrer, 2014) based on the version 5.03 of the model, capable of running on GPU-based hardware architectures, operationally used by MeteoSwiss.

2.2. Data and selected parameters

NWP models, including COSMO, describe physical processes through parameterization schemes in which many unconfined, ‘free’ parameters exist. These parameters are related to sub-grid scale turbulence, surface layer parameterization, grid-scale cloud formation, moist and shallow convection, precipitation, radiation and soil schemes (Doms et al., 2011, Gebhardt et al., 2011).

In the framework of CALMO-MAX, an extended preliminary set of twelve parameters covering turbulence (tur_len, tkhmin, tkmmin), surface layer parameterization (rat_sea, rlam_heat, crsmin), grid-scale precipitation (v0_snow), moist and shallow convection (entr_sc), radiation (rad_fac, uc1) and the soil scheme (c_soil) have been tested. Several sensitivity experiments have been performed and the most sensitive parameters have been selected for calibration (Avgoustoglou et al., 2020). Specifically, the five model parameters chosen for CALMO-MAX are:

1. Minimal diffusion coefficients for heat, tkhmin[m2/s].
2. Scalar resistance for the latent and sensible heat fluxes in the laminar surface layer, rlam_heat [no units].
3. A factor in the terminal velocity for snow, v0_snow[no units].
4. Parameter controlling the vertical variation of critical relative humidity for sub-grid cloud formation, uc1 [no units].
5. The fraction of cloud water and ice considered by the radiation scheme rad_fac [no units].

These parameters were calibrated with respect to daily minimum and maximum 2m temperature (T_max and T_min respectively), hourly, 6h and 24h accumulated precipitation(Prec), and vertical profiles (TEMP). For temperature, available measurements of daily mean surface air temperature selected at the station network of MeteoSwiss were used.

For precipitation, observations over Switzerland were available through the gridded MeteoSwiss radar composites corrected by rain gauges and interpolated to the model grid. Over Northern Italy, observations interpolated to the model grid were used. In addition, vertical model profiles at grid points near soundings locations were considered.

2.3. Methodology

The calibration methodology used in CALMO-MAX was discussed in details in Voudouri et al. 2017 and 2018. It relies on the metamodel proposed by Neelin et al., (2010 and 2010a) and modified by Bellprat et al.,(2012a) that approximates the parameter space using a multivariate quadratic regression in an n-dimensional model. The gain/loss in model quality is assessed using the “COSMO Index” score (COSI) developed by Ulrich Damrath (2009). The score is a combination of both RMSE-type for continuous variables and ETS (Equitable Threshold Score) for categorical fields, and has been used to estimate the overall model performance.

3 Results

The aim of this work was to calibrate COSMO-1, using a full year of statistics. The year 2013 has been chosen as climatologically representative for the target area. Once the simulations for the 5 parameters (tkhmin, rlam_heat, v0_snow, rad_fac and uc1) have been completed, the optimum set of parameters was calculated using the metamodel. It should be noted that although calibration is performed over the entire year, optimum parameter values are extracted over sets of 10-days periods.

An average for these 36 periods is then produced to extract the best optimum parameter set over the entire year. The optimum parameter values are as follows: tkh_min = 0.279 (m2/s), rlam_heat= 0.929, v0_snow = 18.95, rad_fac = 0.6775 and uc1= 0.7686.

The default parameter values (proposed by model developers) were replaced by these “optimal” values, and model simulations for 2013 have been performed again to investigate the improvement in model performance. Additionally, simulations for 2017 have been performed to examine whether the optimum parameter set, calculated for the year of the calibration, is also beneficial for a different independent year.

The verification of simulations using default parameter values (tkh_min = 0.4 (m2/s), rlam_heat = 1, v0_snow = 20, rad_fac = 0.6 and uc1=0.8) (DEF) against the one using optimum parameter set (BEST) for 2m temperature, and hourly accumulated precipitation are presented in Table 1 and Table 2 for both years.

More specifically, statistical measures such as mean error (ME), root mean square error (RMSE), minimum (MINMOD) and maximum (MAXMOD) model values, minimum (MINOBS) and maximum (MAXOBS) observed values are shown. Categorical scores such as Equitable threshold (ETS), False Alarm Ratio (FAR) and Probability of detection (POD) for a small threshold (0.1mm) are also calculated for hourly precipitation in Table 2.

A decrease in the mean error of the 2m temperature is observed when using the optimized configuration that is 0.12°C instead of 0.43°C for 2013, and 0.14°C instead of 0.24°C for 2017. An improvement of the yearly maximum 2m temperatureis also observed for the 2013 experiment (forecasted 38.4°C and 37.4°C for DEF and BEST simulations respectively, against the observed 37.1°C).
An improvement of approximately 0.3°C in RMSE during daytime is also well visible in the daily cycle of the 2m temperature RMSE for the year 2013 (figure 1,RMSE averaged over a full year, blue line is with DEF and red line with BEST parameters). Statistics of hourly accumulated precipitation when using the set of optimum parameter values, for both years, against the values recommended in the default model setup are presented in Table 2.

Table 1: Statistics of 2m temperature for years 2013 and 2017

Year	2013		2017	
Measure/Simulation	DEF	BEST	DEF	BEST
ME	0.43	0.12	0.24	0.14
RMSE	2.2	2.16	2.35	2.33
MINMOD	-28.67	-28.67	-29.64	-28.77
MINOBS	-28.7		-29.5	
MAXMOD	38.43	37.41	40.0	40.0
MAXOBS	37.1		36.9	

Table 2: Statistics of the hourly accumulated precipitation for years 2013 and 2017

Hourly precipitation	2013		2017	
Measure/Simulation	DEF	BEST	DEF	BEST
ME	0.032	0.029	0.027	0.025
RMSE	0.771	0.771	0.8	0.8
MAXMOD	56.07	47.17	48.59	58.24
MAXOBS	48.5		60.8	
ETS(0.1)	0.35	0.33	0.35	0.35
FAR(0.1)	0.44	0.45	0.45	0.45
POD(0.1)	0.64	0.62	0.63	0.63

Unlike the 2m temperature, no clear benefit from the calibration is visible. Although ME for both years is slightly reduced and maximum modeled values are closer to the observed ones when using the optimum values, categorical scores such as ETS and POD are degraded. More specifically, when choosing a small threshold such as 0.1mm, ETS slightly decreases from 0.35 to 0.33, for 2013. This is also the case for thresholds of 1mm and 10mm (not shown in Table 2). For year 2017, no effects on these categorical scores have been observed.This model response could be attributed to deficits of the precipitation scheme in representing the prevailing weather patterns during 2013 and 2017, however further investigation is needed.

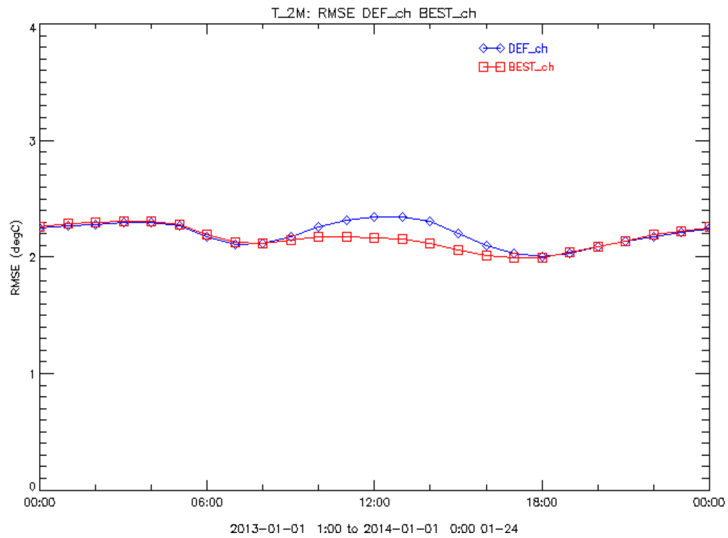


Figure 1: Daily cycle (averaged over the entire year) of 2m temperature RMSE when using default (blue line) and optimum (red line) parameter values for 2013

4 Conclusions

CALMO-MAX is the COSMO priority project for the implementation and consolidation of a robust objective calibration method. In the present work the MeteoSwiss COSMO-1 configuration has been calibrated, selecting five model parameters, using a full year statistic, with the history of the soil included (hindcast),to demonstrate the benefits of the methodology. A different year has been used to have an independent assessment of the impact of the optimization process. Although the chosen model configuration, based on the operational model of MeteoSwiss and close to the DWD configuration, was already a well-tuned configuration, results showed that a slight model performance gain is obtained by using the CALMO methodology.

A remaining issue for a broader use of this methodology is its computational cost, due to the necessity to run multi-years simulation of a high-resolution model to constrain the meta-model. A first consideration which may alleviate this issue is to consider calibrating a lower resolution configuration than the target model; a factor of two in the horizontal resolution does not significantly change the characteristics of the forecasts (e.g. as observed at MeteoSwiss), but reduces the cost of a single simulation by a factor eight. Another consideration, if the history of the soil is not a dominant factor, is to restrict the weather sampling to a set of representative periods, instead of using a full year; typically, choosing 60 days reduces the cost of the calibration by a factor six. A last consideration is to partition the set of considered parameters into (relatively) independent subsets, and to calibrate each subset in turns; this approach reduces the number of simulations required, given that the number of full model simulations is O(N2) where N is the number of parameters to calibrate.

Besides the costs associated with the meteorological model, the specification and the use of the meta-model may also become expensive, in particular when a large number of observations are considered. A lot of meta-model optimizations have already been done and further ideas about the optimization process are under consideration.

To demonstrate the feasibility of these ideas, a new calibration is currently applied over a large Central- and Eastern-Mediterranean domain, covering mainly marine instead of continental area. This application will prove whether the CALMO methodology can be used as an affordable and useful tool to define the optimal calibration over a different target area of interest (or a significantly different model configuration).

5 Future Work

A new Swiss National Fund project in the group of Prof. C. Schaer / ETHZ has been accepted (trCLIM), and, in particular, a 3 years PhD focusing on calibration will start in autumn 2020. Furthermore, as already stated, a lot of developments have been done in the meta-model by the CALMO team, and further ideas and considerations about the optimization process have been proposed at a BTU / Cottbus meeting beginning of 2020. This shows the necessity to synchronize the COSMO and the ETHZ developments anew, and to provide a unified, consolidated, portable (Octave or Python) and well documented (user guide) meta-model code, with the possibility to define any meaningful model performance score in an easy way. This will be a very useful tool for both the Climate and the NWP communities and this could be implemented in a future COSMO PT or PP.

An important aspect shown in previous work is the strong dependency of the parameters optimum on the time of the year, which reflects the dependency of the optimum on the atmospheric flow (or weather pattern), and the implicit dependency of some tuning parameters to the model state variable. This was expected, but the intra-annual fluctuations of the optimum are large. Practical consequences of this fact on the use of the CALMO methodology have to be considered in the future.

Finally, it should be noted that this methodology is essentially “model independent” and can be applied to any NWP or climate model. The only pre-requisite is an up-to-date and well-documented list of tunable model parameters, which supports a first screening of relevant parameters for the planned calibration. In fact, plans to calibrate ICON with the CALMO methodology are already considered by some COSMO partners.

Acknowledgments The present work is part of CALMO-MAX priority project of COSMO. CSCS and MeteoSwiss are acknowledged for providing computer resources.

References

[1] Avgoustoglou E., A. Voudouri, I Carmona, E. Bucchignani, Y. Levy, J. -M. Bettems, 2020: A methodology towards the hierarchy of COSMO parameter calibration tests via the domain sensitivity over the Mediterranean area. COSMO Technical Report 42. $https://doi.org/10.5676/dwd\_pub/nwv/cosmo\_tr\_42$

[2] Bayler, C., Gail, M., Aune R. M., Raymond, W. H., 2000. NWP Cloud Initialization Using GOES Sounder Data and Improved Modeling of Non precipitating Clouds. Mon. Wea. Rev., **128**, 3911-3921.

[3] Bellprat, O., Kotlarski, S., Lüthi, D., Schär, C., 2012a. Objective calibration of regional climate models. Journal of Geophysical Research, **117**, D23115.

[4] Bellprat, O., Kotlarski, S., Lüthi, D., Schär, C., 2012b. Exploring perturbed physics ensembles in a regional climate model. Journal of Climate, **25**, 4582-4599.

[5] Damrath U., 2009, Presentation at 11$^{\text{th}}$ COSMO-GM Offenbach, Germany. $www.cosmo-model.org/content/consortium/generalMeetings/general2009/wg5\_versus2/Ulrich_Damrath\_Long\_term\_time\_series\_of\_COSMO\_EU.ppt$

[6] Doms, G., Foerstner, J., Heise, E., Herzog, H-J., Mironov, D., Raschendorfer, M., Reinhardt, T., Ritter, B., Schrodin, R., Schultz, J.-P., Vogel, G., 2011. A description of the nonhydrostatic regional COSMO model. Part II: Physical Parameterization. $http://www.cosmo-model.org$

[7] Duan, Q., Schaake, J., Andreássian, V., Franks, S., Goteti, G., Gupta, H.V., Gusev, Y.M., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., Wood, E.F., 2006. Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. Journal of Hydrology, **320**, 3–17.

[8] Khain P., I. Carmona, A. Voudouri, E. Avgoustoglou, J.-M. Bettems, F. Grazzini, 2015: The Proof of the Parameters Calibration Method: CALMO Progress Report. COSMO Technical Report, 25. $https://doi.org/10.5676/dwd\_pub/nwv/cosmo\_tr\_25$

[9] Khain P., I. Carmona, A. Voudouri, E. Avgoustoglou, J.-M. Bettems, F. Grazzini, P. Kaufman, 2017:CALMO Progress Report. COSMO Technical Report 31.$https://doi.org/10.5676/dwd\_pub/nwv/cosmo\_tr\_31$

[10] Lapillonne, X., Fuhrer, O., 2014: Using compiler directives to port large scientific applications to GPUs: An example from atmospheric science. Parallel Processing Letters, **Vol. 24**, No. 1 , 1450003

[11] Neelin, J. D., Bracco, A., Luo, H., McWilliams, J.C., Meyerson, J.E., 2010. Considerations for parameter optimization and sensitivity in climate models. Proc. of the National Academy of Sciences of the United States of America, **107**, 21349-21354.

[12] Neelin, J. D., Bracco, A., Luo, H., McWilliams, J.C., Meyerson J. E., 2010a. Supporting information for:Considerations for parameter optimization and sensitivity in climate models. $http://www.pnas.org/content/suppl/2010/11/24/1015473107.DCSupplemental/Appendix.pdf$

[13] Skamarock, W.C., 2004. Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra. Mon. Wea. Rev., **132**, 3019-3032.

[14] Voudouri, A., Khain, P., Carmona, I., Bellprat, O., Grazzini, F., Avgoustoglou, E., Bettems, J. M. and Kaufmann, P., 2017. Objective calibration of numerical weather prediction models, Atm. Res., 190, pp 128-140.

[15] Voudouri A., P. Khain, I. Carmona, E. Avgoustoglou, J.M. Bettems, F. Grazzini, O. Bellprat, P. Kaufmann and E. Bucchignani, 2017a:Calibration of COSMO Model, Priority Project CALMO, Final report. COSMO Technical Report,**32**. $https://doi.org/10.5676/dwd\_pub/nwv/cosmo\_tr\_32$

[16] Voudouri A., Khain P., Carmona I., Avgoustoglou E., Kaufmann P., Grazzini F.,andBettems J.M, 2018: Optimization of high resolution COSMO model performance over Switzerland and Northern Italy. Atm. Res. 213, 70-85, $https://doi.org/10.1016/j.atmosres.2018.05.026$

[17] Voudouri A., I. Carmona, E. Avgoustoglou, Y. Levi, J.M. Bettems and E. Bucchignani, 2019: Impacts on model performance score from CALMO and CALMO-MAX. COSMO Newsletter 19, pp 32-36 $https://doi.org/10.5676/dwd\_pub/nwv/cosmo\_nl\_19$

Common Verification Activity during JJA 2018 - MAM 2019: Main Results

Kirsanov A., Gofa F., Boucouvala D., Batignani F., Bogdan M., Linkowska J., Pflüger U., Tesini M.S., and A. Bundel

1 Introduction

As decided during COSMO General Meeting in Lugano 2012, the performance of COSMO models is demonstrated through the Common Plot activity. Verification results of statistical indices for main weather parameters are derived using the operational COSMO model implementations in each service. The domain, resolution, statistical scores/methods, frequency and graphical representation, are decided on an annual basis from WG5 and listed in the guidelines document (http://www.cosmo-model.org/view/repository/wg5/commonPlots/ reports).

The main findings of this organized analysis are presented during the GM plenary session together with the long-term trend of the scores, providing a basis to track the performance of COSMO model as well its systematic errors. As a Common Verification Software (CVS) is used by all services, this allows for a homogeneous, standardized and objective way to apply, calculate and present the verification scores.

The common geographical areas for the coarser and the higher resolution models that are used in this analysis are shown in Figure 1. ECMWF (IFS), ICON-EU and ICON results are also included and compared to COSMO models. The common area includes part of Northern Italy, Austria, Slovenia, Croatia, Germany, Bosnia and Herzegovina, Hungary, Slovakia, and Check Republic.

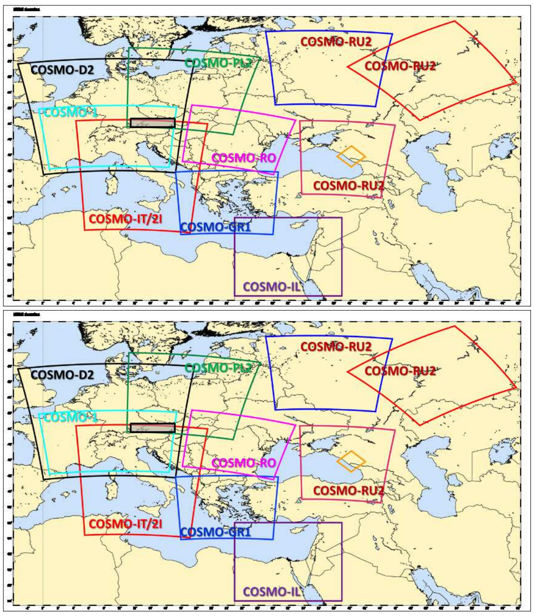


Figure 1: Common Area 1 (upper) and 2 (lower) domains as located within each country simulation area.

doi:10.5676/dwd_pub/nwv/cosmo-nl_20_03

While the transition from COSMO to ICON model is ongoing, the focus of Common Verification activity for the period Summer 2018 – Spring 2019 was also on the comparison of COSMO and ICON models as this would provide a perspective on the existing and the resolved biases that models exhibit.

2 Methodology and Data

The models participating in verification of Common Area 1 (coarse resolution) are COSMO-GR4, COSMO-5M, ICON-EU, COSMO-ME, COSMO-PL, COSMO-RU7 and the global models IFS (ECMWF), ICON(DWD). The fine resolution models participating in verification of Common Area 2 are COSMO-D2, COSMO-IT, COSMO-P. Comparison with the larger scale models IFS, ICON in this domain is also performed. The forecast parameters were interpolated to the observation point by using the 3D method height optimized, except for TCC (15km radius method) and Precipitation (8km radius method). The forecast quality was estimated by applying continuous and dichotomic scores on a 3-hourly time step, depending on the parameter type as follows:

ME and RMSE scores were calculated for Continuous parameters: T2m (2m Temperature), MSLP (Mean Sea Level Pressure), Td (dew point Temperature), WS (Wind Speed), TCC (Total Cloud Cover).

FBI, ETS scores were calculated for Dichotomic parameters: 6h cumulated Precipitation, main thresholds: 0.2, 1, 5, 10, 15, 20 mm/6h.

Dichotomic scores FBI, ETS, CSI were also applied to continuous parameters TCC and Wind Gust by setting intervals. Specifically, for TCC intervals are set to [0,25], (25, 75), [75,100], and for Wind Gust $>$ 12.5, $>$15, $>$20 m/sec.

For more information about scores analysis, see the interactive plots and final report for 2018-2019 and previous years on [1].

3 Results

Detailed analysis on the models performance and comparison with previous years can be found in [1]. Only the main results for Common Area 1 are presented in this short report.

2m Temperature: Despite the differences among models, the common feature is that T diurnal cycle amplitude is underestimated, similarly to the previous years. In JJA (Figure 2a) and SON, the bias is positive in the daytime (overestimation), and slightly negative (underestimation) or close to zero during night hours. In DJF (Figure 3b) and MAM, the temperature is underestimated in the daytime, while the values are close to zero at night in DJF and slightly overestimated at night in MAM. The RMSE also exhibits a diurnal cycle with maxima depending on the season and following the absolute maxima of the ME cycle. ICON and ICON-EU perform relatively better compared to COSMO models: The bias cycle is weaker with values closer to zero. RMSE values are also lower (Figure 3).

It is worth noting, that there are some differences in RMSE of ICON model diurnal cycle between JJA and DJF. During the summer, ICON and ICON-EU performance follows COSMO models one exhibiting similar RMSE diurnal cycle, with higher values in the daytime. However, in winter period, ICON and ICON-EU RMSE values exhibit a slight peak during night hours. Regarding the tendency of the scores compared to the previous year, RMSE slightly decreased for JJA, SON and MAM. A small increase was noticed for DJF period though[1].

Dew point temperature at 2 m. Similarly, to the previous years Td exhibits a diurnal cycle in errors, with ME and RMSE peaking in early afternoon.

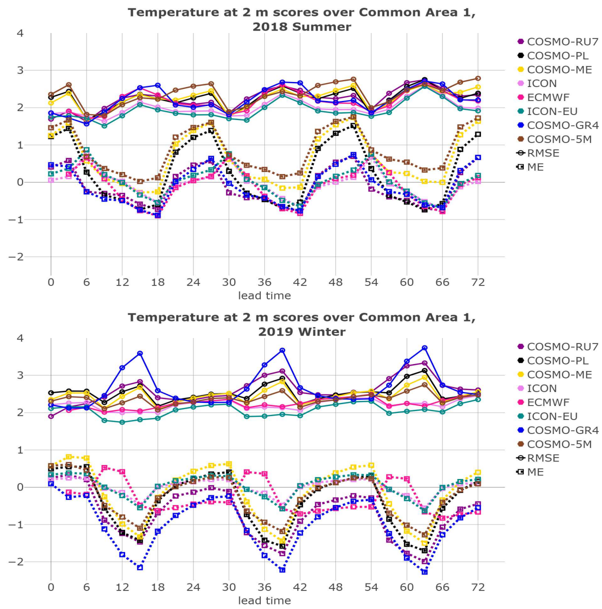


Figure 2: Temperature at 2 m Continuous scores during summer(a) and winter(b)

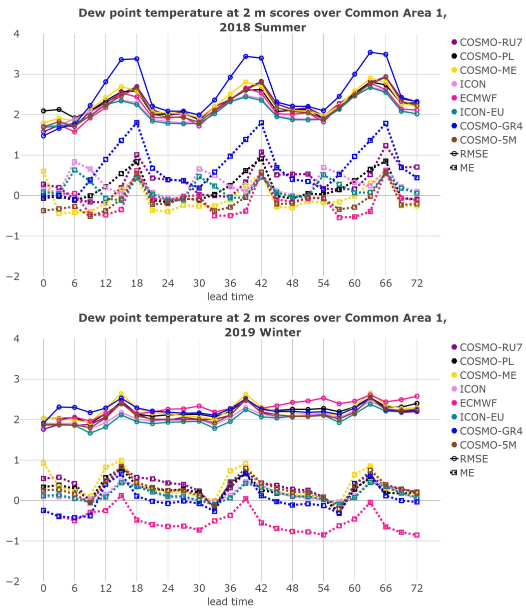


Figure 3: Dew point temperature at 2 m Continuous scores during summer(a) and winter(b)

All models have the tendency to overestimate dew point temperature at 2 m. However, COSMO-PL and COSMO-RU7 underestimate observed values in fall season [1] while IFS underestimates it in winter. ME has negative values increasing with lead time (Figure 3). ICON and ICON-EU scores are better than COSMO for all seasons. Regarding the score tendency compared to the previous year, RMSE slightly decreased for JJA season while no significant changes are shown for MAM. However, a small increase was noticed for DJF and SON seasons [1].
It is worth noting that ICON and ICON-EU elevated ME values found in the year 2016-2017, are no longer present in the two more recent periods.

Mean Sea Level Pressure. The Bias evolution with lead time differs among models and seasons and does not have a very clear diurnal cycle. The only season where a noticeable trend is apparent is SON, that ICON, ICON-EU and IFS underestimate MSLP while all COSMO models tend to overestimate it.

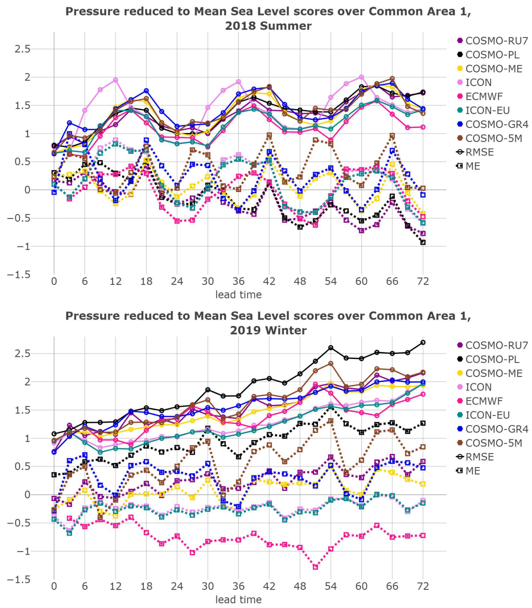


Figure 4: MSLP at 2 m Continuous scores during summer(a) and winter(b)

The RMSE variability is more obvious for this parameter. Similarly to the previous years, RMSE increases with forecast time especially in SON, DJF, MAM [1] and exhibits an afternoon maximum in JJA and in MAM seasons (ICON-EU, ICON and IFS peak only during JJA) (Figure 4). ICON, ICON-EU and IFS RMSE values are lower than COSMO models, especially at higher lead times.
Regarding the RMSE trend compared to the previous years, there was a decrease of RMSE in JJA, SON and DJF for all models. However, a slight RMSE increase is noticed for all models in MAM.

Total Cloud Cover (Note that nighttime observations are limited). ICON and IFS performance is generally better than the one of COSMO models for all seasons especially in JJA, with ME values closer to zero. (Figure 5).

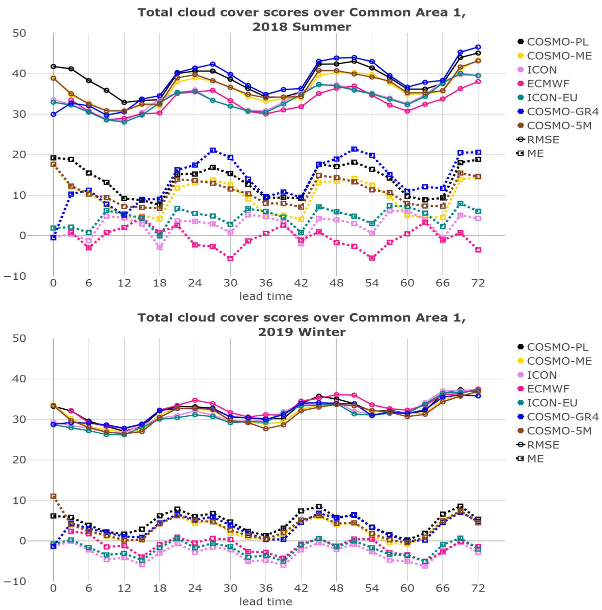


Figure 5: Total Cloud Cover continuous scores during summer and winter

RMSE values for ICON model are slightly lower than COSMO in warm seasons and quite comparable in cold seasons. RMSE for all models is slightly higher in JJA especially during nighttime. Regarding the tendency compared to the previous year, RMSE decreased during MAM for all models, COSMO RMSE increased during JJA and SON (ICON did not change significantly), and there were no significant changes for DJF season.

The dichotomic scores TCC analysis showed that:
TCC<25% (including clear sky) events are generally underforecasted especially at night in JJA when the FBI values are lower.
TCC >75% (including overcast) events are generally overforecasted especially in the summer around noon (FBI >1).
For 25-75% TCC events, the FBI values (mainly<1) indicate underforecasting of events especially for COSMO models. This category is the hardest to forecast especially in the summer, exhibiting the lowest TS (Threat score) [1]. Overall, ICON models perform better than COSMO. In Figure 6, the FBI values for ICON are shown.

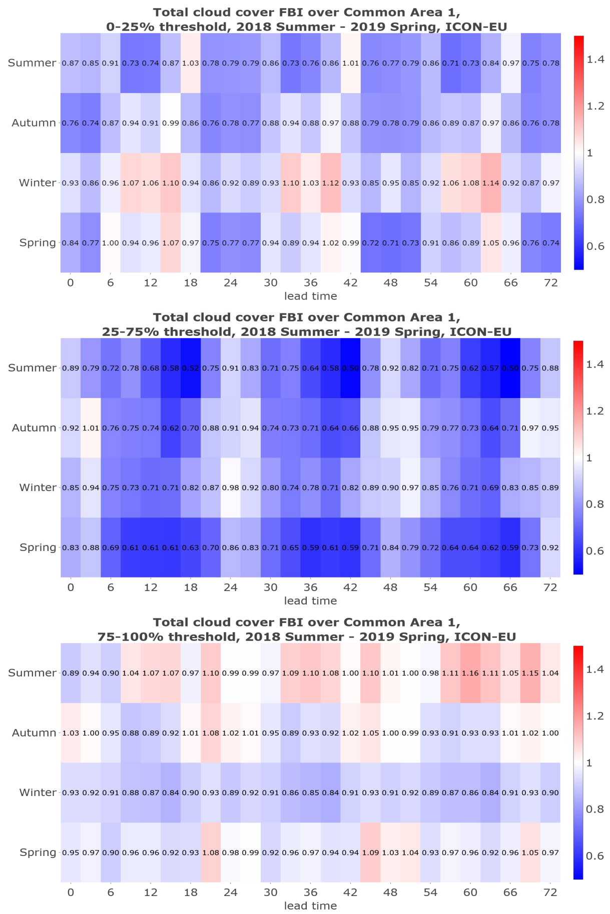


Figure 6: ICON-EU Total Cloud Cover FBI (Frequency Bias Index) for different thresholds

Wind speed at 10 m: COSMO and IFS models exhibit a diurnal bias cycle, with a slight wind speed overestimation at night and early morning, and a slight underestimation around noon hours. ICON models generally underestimate wind speed especially at night, with a minimal diurnal cycle (Figure 7). All models have similar RMSE values and variation, with a slight diurnal cycle in summer period. There is a notable RMSE difference between ICON and ICON-EU during winter, with ICON exhibiting higher values than ICON-EU limited area model. The RMSE trend compared to the previous year was a slight decrease in JJA and SON seasons, and a slight increase in spring. No significant changes for winter months.

10m Wind Gust: The dichotomic scores (FBI, TS) of the 10m wind gust simulations [1] exhibit a general tendency to underestimate 10m wind gust FBI especially for higher thresholds. TS values decrease with increasing threshold.

6h precipitation: Overestimation of low thresholds and underestimation of higher thresholds of precipitation events, is a common feature that is shown similarly to the previous years, derived from the fact that FBI decreases for higher precipitation thresholds (Figures 8, 9). FBI values exhibit a pronounced diurnal cycle in JJA season with higher values around noon which means that mainly noon time precipitation events are overestimated.

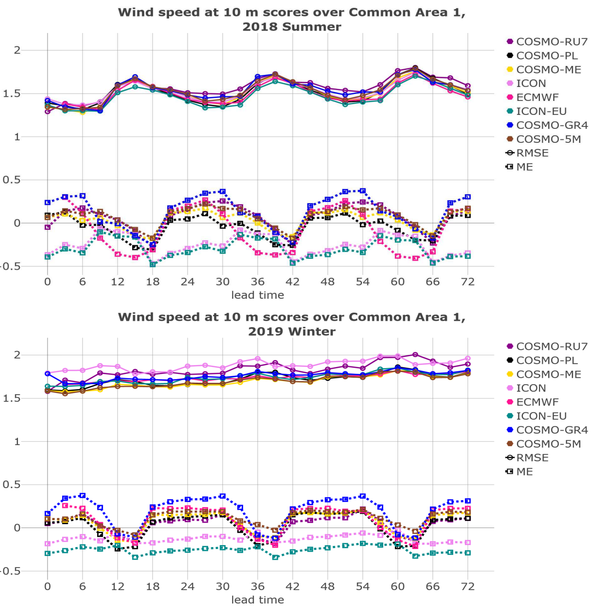


Figure 7: Wind speed at 10 m Continuous scores during summer and winter

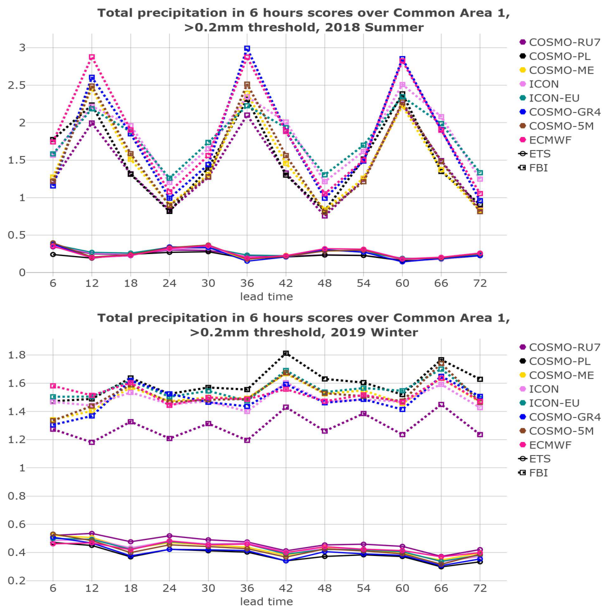


Figure 8: Total precipitation in 6 hours >0.2 mm threshold scores (FBI-ETS) during summer and winter

COSMO models underestimate the frequency of higher threshold precipitation events during summer and autumn, less than ICON model does (better FBI values). ETS (Equitable Threat score) for all models is lower for higher thresholds which means that the forecast quality is worse with increasing threshold. COSMO ETS values are worse than for ICON models.

ETS precipitation scores are improved during winter and are worsen during summer period. The worst ETS values generally occur at forecast times 12-18 UTC, while the best values during the period 00-06 UTC.

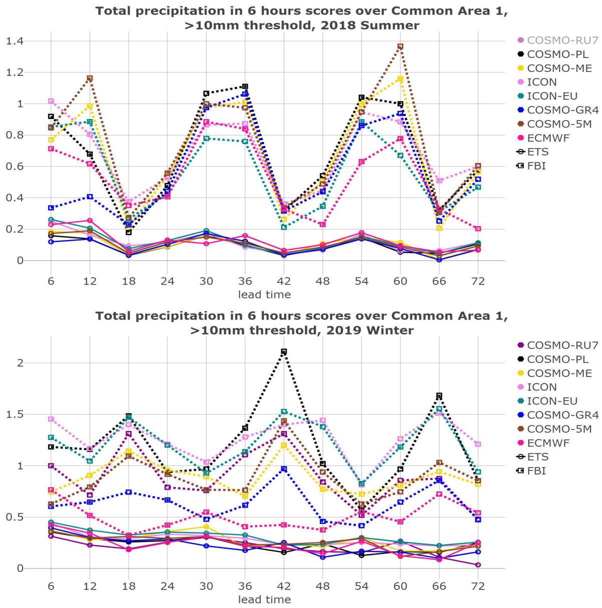


Figure 9: Total precipitation in 6 hours >10 mm threshold scores during summer and winter FBI

4 Conclusions

Common Plots Verification activity which initiated after the 14th COSMO General Meeting in 2012, aims to track systematic errors and long term trends of COSMO model, During Summer 2018 – Spring 2019 while transition from COSMO to ICON model is ongoing, one of the main points of Common Verification activity was the comparison of COSMO and ICON models over the same geographic region and time periods.

COSMO statistical scores behavior (diurnal cycle, evolution with forecast time, seasonal variability) for almost all parameters are persistent over the years. However, the scores absolute values have the tendency to decrease over the years, which denotes the model improvement in subsequent model versions.

ICON model(s) performance is overall improved compared to that of COSMO and IFS models especially for continuous parameters with reduced diurnal cycle of errors and lower RMSE values. Precipitation scores are however comparable. There was some improvement of ICON performance over the last 2 years due to tuning that has been applied as shown in [1] (e.g. dew point ME reduction).

References

[1] Consortium for Small-scale Modeling Verification tasks page.http://cosmo-model.org/content/tasks/verification.priv/default.htm

Common Area Verification Activity with Rfdbk/MEC Application: MAM 2020 First Results

A. Iriza-Burcă[1], J. Linkowska[2] and F. Fundel[3]

1. National Meteorological Administration, Bucharest, Romania
2. Institute of Meteorology and Water Management National Research Institute, Warsaw, Poland
3. Deutscher Wetterdienst, Offenbach am Main, Germany

1 Introduction

The Common Plots (CP) Verification Activity aims to evaluate the performance of the various operational model implementations of COSMO over the same geographical areas, based on the same set of observations. For these activities, a common verification tool called VERSUS is used, ensuring the use of the same verification practices. This allows for a fair comparison of the results and an easy assessment of model performance.

The CARMA priority project (PP CARMA) aims to replace VERSUS with the MEC-Rfdbk software developed by DWD, as a Common Verification Software. This software uses the Model Equivalent Calculator (MEC, [1]) to produce Feedback Files [2], which are then used by R verification scripts based on the Rfdbk package [3] to produce verification scores. The MEC-Rfdbk system is used operationally at DWD for the verification of both the COSMO [4] and ICON [5] model chains.

The advantages of this verification system proposed to be adopted for CP preparation activity include the data pre-processing where all the observations are quality controlled by data assimilation, co-located observational and model data using the same observation operator to produce one combined file, allowing a fast and simple calculation of standard verification scores and online visualization of results.

The first results of a cross model verification (with the MEC-Rfdbk software) for the Common Plot activities for the 2020 spring season (MAM 2020) are presented below. For these results, a set of three COSMO model runs are considered: COSMO-D2 (DWD), COSMO-PL (IMGW) and COSMO-RO (NMA).

2 Brief Description of Data and Methodology

COSMO 00 UTC model runs are evaluated, with forecast step every 3 hours. The horizontal resolution is 7km for the COSMO-PL and COSMO-RO models and 2.2km for COSMO-D2. The integration time for COSMO-PL and COSMO-RO is 72 hours, while for COSMO-D2 27 hours of forecast are available. As a consequence, scores are computed every 3 hours for 27 hours forecast when comparing all three models. However, for a comparison only between COSMO-PL and COSMO-RO, scores could be computed for 72 hours of forecast.

The integration domains for COSMO-D2 and COSMO-PL (figure 1, top row) are the operational ones (also evaluated for the official Common Plots activities), while the COSMO-RO integration domain for the current evaluation (figure 1, bottom row) differs from the operational set-up of the model employed in NMA. This was done in order to cover the Common Areas, which are otherwise outside of the COSMO-RO domain. No other modifications are done compared to the operational set-up of the model.

doi:10.5676/dwd_pub/nwv/cosmo-nl_20_04

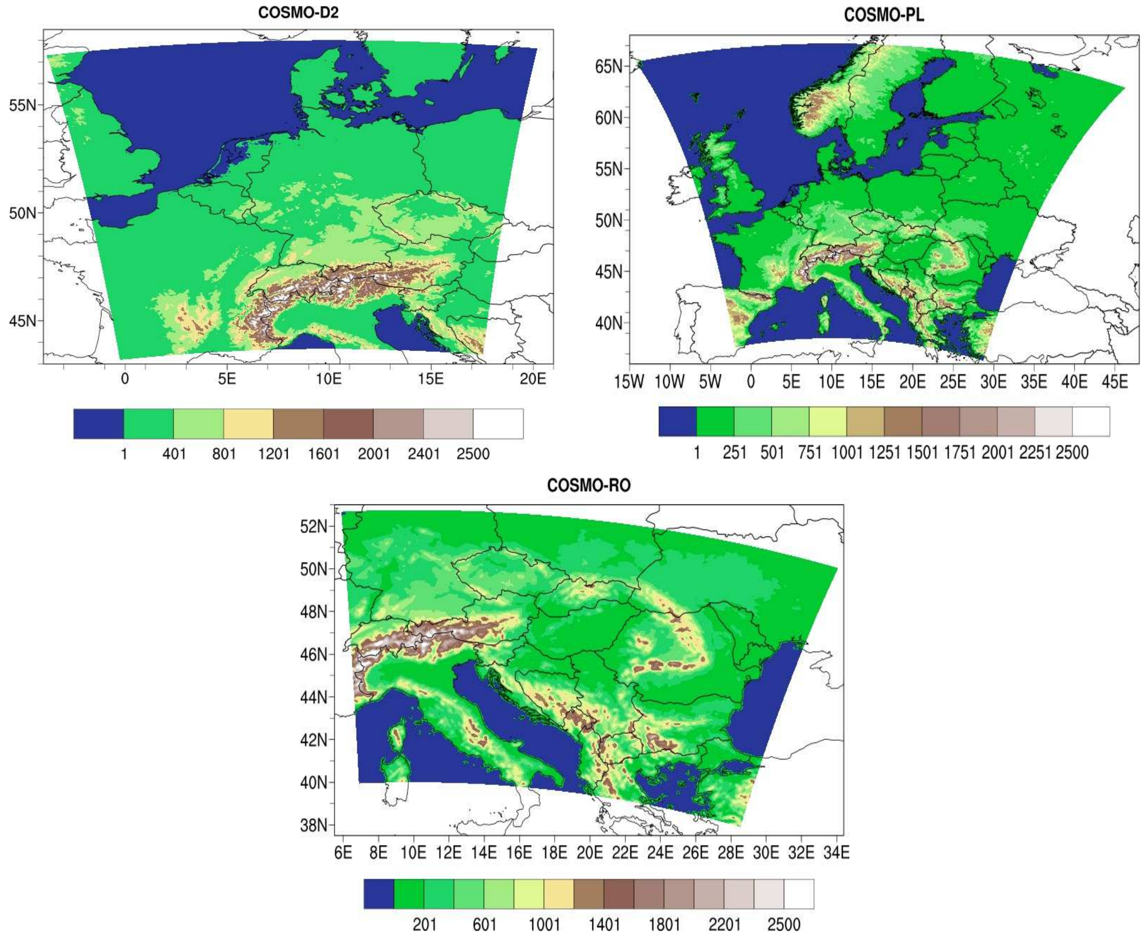


Figure 1: Integration domains for COSMO-D2 (top, left), COSMO-PL (top, right), COSMO-RO-7km (bottom) used for verification.

An overview of the model configurations can be seen in Table 1, while a more detailed description can be found on the COSMO web-site [6].

	res.	ie_tot	je_tot	ke_tot	ICLBC	DA	ant.
COSMO-D2	0.02	651	716	65	ICON-EU	KENDA,LHN	27h
COSMO-PL	0.0625	415	460	40	DAC/ICON	nudging	72h
COSMO-RO	0.0625	201	177	40	ICON	nudging	72h

Table 1: Overview of the model configurations.

As described in [7], COSMO models are generally verified over the same areas, with the same set of stations. Observations are retrieved in BUFR format from the MARS database and converted to netcdf format using the bufr2netcdf software [8]. The Rfdbk system allows to align the data, so that only those available for all evaluated experiments are used.

The two common areas for which the scores are computed are presented in figure 2 and include parts of Northern Italy, Austria, Slovenia, Croatia, Germany, Bosnia and Herzegovina, Hungary, Slovakia, and the Czech Republic.

For Common Area 1 (CA1, figure 2, left), 96 stations are selected, with altitudes between 0m and 753 meters. Although slightly more restricted geographically, Common Area 2 (CA2, figure 2, right) includes a larger number of stations than CA1. The stations in CA2 also cover a greater range of altitudes compared to CA1 (maximum altitude 3114m in CA2 compared to 753m in CA1). Table 2 offers an overview of station distribution for the two areas.

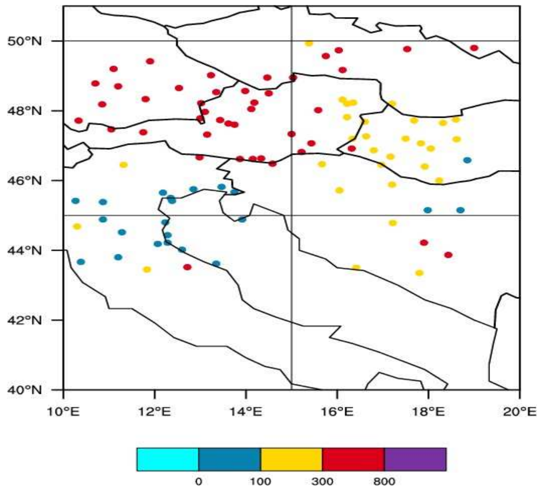


Figure 2: Location and altitude of synoptic stations in CA1 (left) and CA2 (right) used for the verification.

Common Area	Number of stations					Alt.	
	Total	alt≤100m	100m<alt≤300m	300m<alt≤800m	alt<800m	min	max
CA1	96	23	30	43	0	0m	753m
CA2	205	2	41	107	55	0m	3114m

Table 2: Overview of stations distribution for the 2 Common Areas, depending on altitude.

The results presented in this paper are computed taking into account all stations of interest. However, results are also available stratified by station altitude.

Standard Verification (seasonal) was performed for the following continuous parameters: 2 meter temperature – T2M (deg K), 2 meter dew point temperature - TD2M (deg K), surface pressure - PS (Pa), total cloud cover - N (octa), 10 meter wind speed - FF (m/s), wind direction - DD (deg) and wind gust - Gust (1 hour, m/s). Scores for these parameters include the mean error (ME), root mean squared error (RMSE) and mean absolute error (MAE).

Categorical scores are also computed, for the following parameters: 6 hour accumulated precipitation - RR_6h (thresholds: 0.2, 1, 5, 10, 20 mm/6h), total cloud cover (thresholds >=1, >=4 and >=7) and 10 meter wind gust (thresholds: >=12.5, >=15, >=20 m/s). Dichotomic scores computed for these variables include the probability of detection (POD), false alarm rate (FAR), equitable threat score (ETS), frequency bias (FBI). Other scores based on the number of hits, misses, false alarms and correct negatives respectively are available but not shown.

Both for the standard verification and the dichotomic scores, the number of observations used in computations (LEN) can be shown. General information of scores and a detailed description of the common plot verification procedures can be found in [7]. The verification results presented below (figures 3-12) are a sample of the derived statistics that were obtained. A complete set of statistical scores obtained with the MEC-Rfdbk verification system for the three models considered here is available on the COSMO web-site [6].

3 Considerations Regarding the MEC-Rfdbk System

As mentioned before, currently, CP verification activities are carried out using the VERSUS verification software environment. Because of the technical limitations of VERSUS and lack of further development, it has been decided within the consortium that it should be replaced with the MEC-Rfdbk software. The MEC-Rfdbk system uses small files for fast calculation of verification scores, while the results can be browsed interactively online. The verification is based on the use of feedback files, that hold information on observations (including meta-data) and their usage in data assimilation, as well as the corresponding model analysis, first-guess and past forecasts, in NetCDF format. These files are produced for each valid time and observation type and can be used for various verification tasks.

MEC [1] is a Fortran 2003/2008 and C - based binary that produces feedback files using observations in netcdf format and model data in grib format. These files can be produced for any model (COSMO, ICON, IFS, etc.). Some mandatory parameters from the model of interest must be available on all model levels: PS, T, U, V, P, Q, while others are optional, depending on the available observations and user needs (T2M, TD2M, PS, N, FF, DD, Gust, RR, etc). MEC applies the observation operators from the data assimilation scheme to model fields and stores the results in feedback-files. The software can use as input observations, model runs (deterministic or ensemble), analysis runs or even another MEC run, depending on the user needs and can be applied to interpolate between two or more time periods [9].

The advantages of using the MEC software as part of the verification system for CP activities are related to data pre-processing (all data in one place) and ensuring observation and forecasts are correctly assigned to each other, with quality control done by data assimilation.

The Rfdbk package [3] is an R interface that aims to exploit the information contained in the feedback files and can be used to perform feedback file based verification. Rfdbk is the basis to a set of verification scripts that use as input the feedback files obtained from MEC (one file for each validity date and observation type) and outputs score files (again, for each validity date and observation type). Based on Rfdbk, verification scripts are available for various types of observations (SYNOP, radiosondes, radio occultation, aircraft, wind profiler and so on). Continuous scores are computed for various types of observations, while categorical scores are also available for SYNOP observations. Verification can be performed for deterministic runs (forecast or hindcast) or ensemble, for any model (COSMO, ICON, IFS, etc.), while cross model verification (e.g. COSMO vs. ICON vs. IFS) is also possible. Rfdbk -based scripts can be used for:

- domain average verification (function of forecast lead-time for a user defined verification period), including domain stratification,
- time series (function of valid time in the verification period or as a function of the forecast lead-time)
- station based verification (function of observation station).
- aggregation on sub-domains, height bins, levels or periods
- significance test
- conditional verification

The advantage of using Rfdbk based verification scripts is their flexibility, which means they can be modified and adjusted according to the needs of each user and (UNIX) system. Finally, the centralized, online, interactive visualization of the results on the COSMO web-site using the R Shiny server [10] allows for an easier evaluation of the results.

4 First Results

Results for continuous parameters - Common Area 1 (figures 3 and 4)

For 2 meter temperature, similar behaviour can be observed for all three models, that is overestimation of forecasted values during night hours and underestimation during the day. The amplitude of errors is of about 2 K, with slightly higher error amplitude for COSMO-NMA during the day (+15, +18 hours).

With regards to 2 meter dew point temperature, there is a similar behaviour for COSMO-D2 and COSMO-RO, mainly overall overestimation of forecasted values, with higher errors during the afternoon. Smaller errors are registered for COSMO-PL compared to the other 2 models, with overestimation of temperatures for the first forecast interval (up to +9hours) and for +18 hours anticipation, while for the remaining intervals the tendency of the model is that of underestimating the values for this parameter. The amplitude of errors slightly larger than that for 2 meter temperature, again with slightly higher error amplitude for COSMO-NMA during the day (+15, +18 hours).

An underestimation of surface pressure values can be observed during the day for all three models, especially COSM-PL The latter also exhibits a larger amplitude of errors in the afternoon.

On the other hand, total cloud cover values are overestimation by all three models, especially COSM-DE, while the amplitude of errors is similar for all models considered.

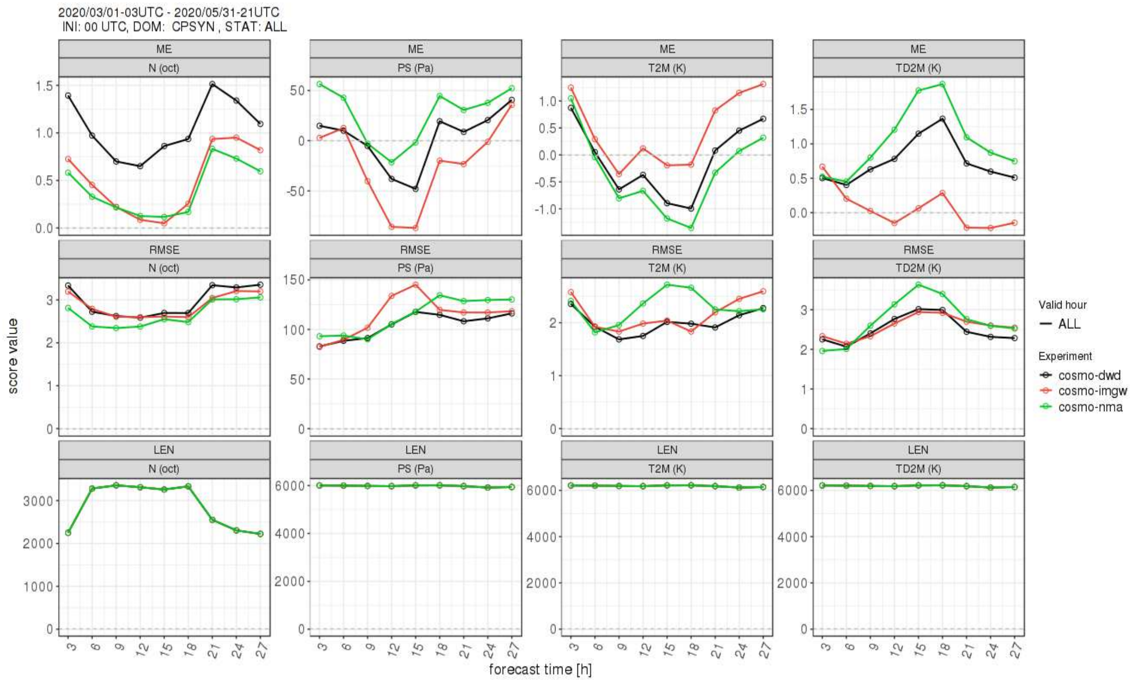


Figure 3: ME (top row), RMSE (middle row) and LEN (bottom row) values for CA1; left to right: N, PS, T2M, TD2M. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

With regards to wind related parameters (FF, DD, Gust, figure 4), all three models exhibit a general tendency to overestimate forecasted values compared to observations, with slight underestimation of wind speed values during the day. In this case, the amplitude of errors is low and comparable for all three models. The varying number of observations (LEN) for DD is due to the limit of $FF > 3m/s$ which is imposed to eliminate unreliable wind direction observations in case of too little wind speed.

Results for continuous parameters - Common Area 2 (figures 5 and 6)

Again, for 2 meter temperature similar behaviour for all three models is observed also for CA2, mainly overestimation of forecasted values during night hours and underestimation during the day; with slightly higher error amplitude for COSMO-NMA during the day (+15, +18 hours).

Also for 2 meter dew point temperature there is a similar behaviour between all three models, with and overall overestimation. Generally, higher errors are registered in the afternoon. As for CA1, there are smaller errors from COSMO-PL compared to the other 2 models, with slight underestimation starting with +21 hours anticipation.

Both COSM-PL and COSMO-D2 mostly underestimate surface pressure values during the day, while COSMO-RO exhibits a behaviour of overestimation for this parameter for the entire forecast period, also with slightly

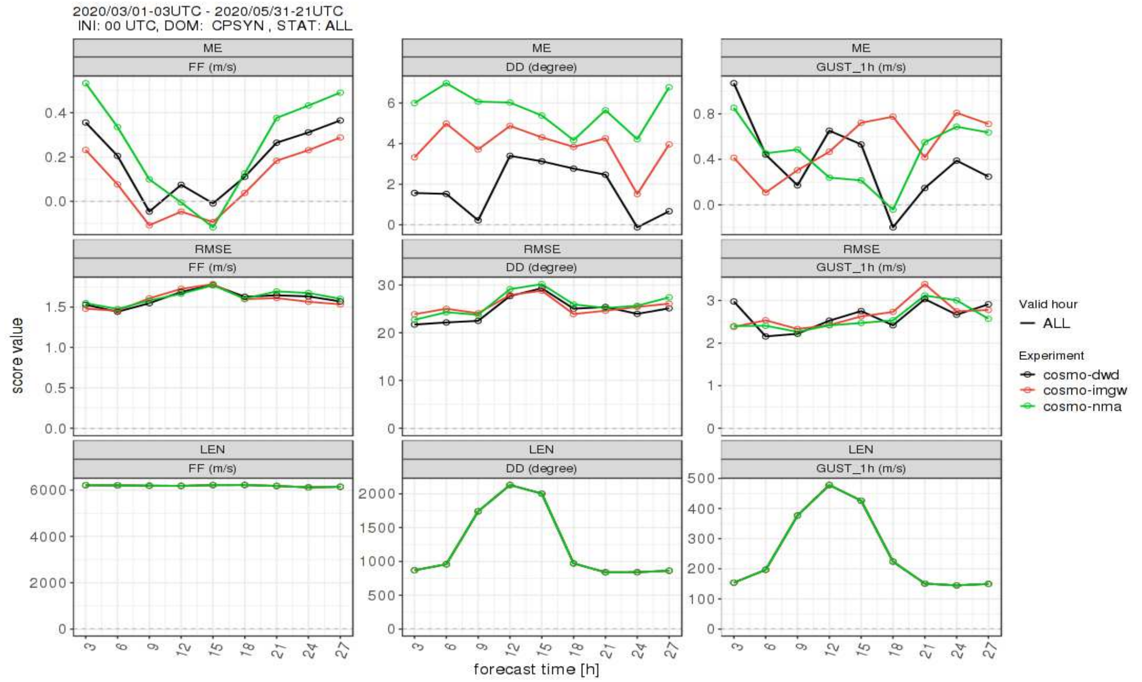


Figure 4: ME (top row), RMSE (middle row) and LEN (bottom row) values for CA1; left to right: FF, DD, Gust. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

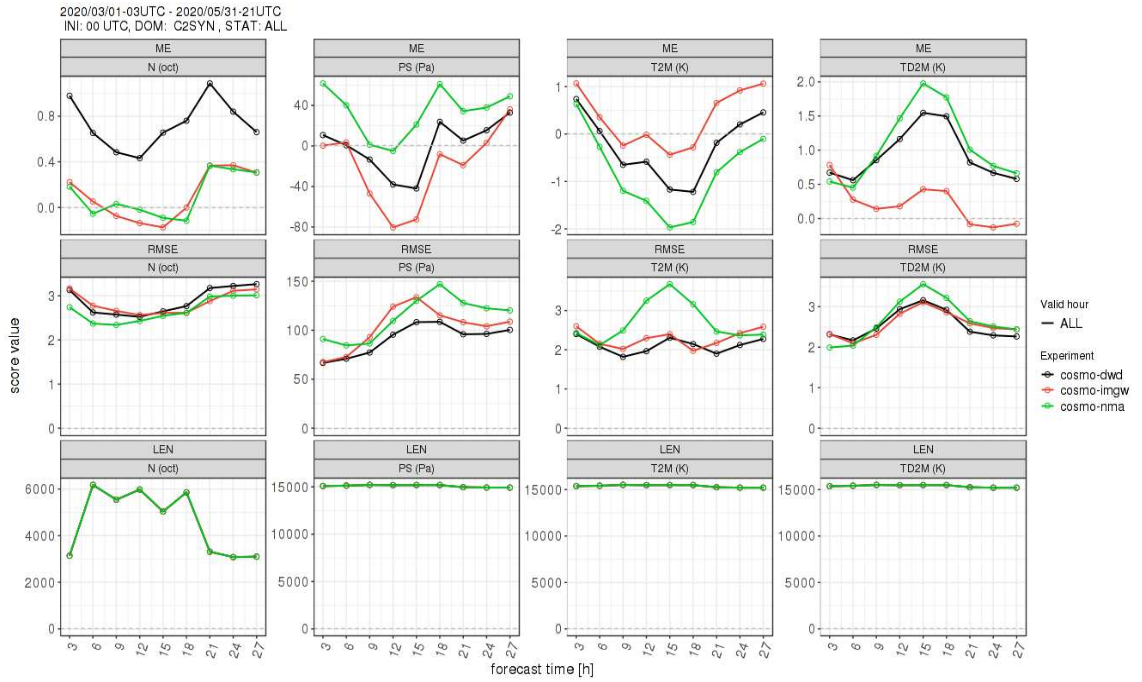


Figure 5: ME (top row), RMSE (middle row) and LEN (bottom row) values for CA2; left to right: N, PS, T2M, TD2M. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

larger amplitude of errors.

Overestimation of total cloud cover values from COSM-DE is shown for the entire interval; for the other two models, there is a slight underestimation of total cloud cover values during the day, while amplitude of errors is comparable between the three.

For 10 meter wind direction (for CA2), an overestimation from COSMO-PL and COSMO-RO for the entire forecast interval is exhibited, while for COSMO-D2 this can be seen only during the day. 10 meter wind speed forecasts have a very low error. The general tendency is that of overestimation of the values forecasted for this parameter from COSMO-D2. COSMO-RO overestimated the values forecasted for this parameter only during the day, while from COSMO-PL we notice underestimations. Overestimation of wind gust values is seen during the night and early morning for all three models; during the day, the general tendency is that of overestimation of values from COSMO-PL and underestimation for COSMO-RO. For these last three parameters, the amplitudes of errors are comparable in all analyzed models.

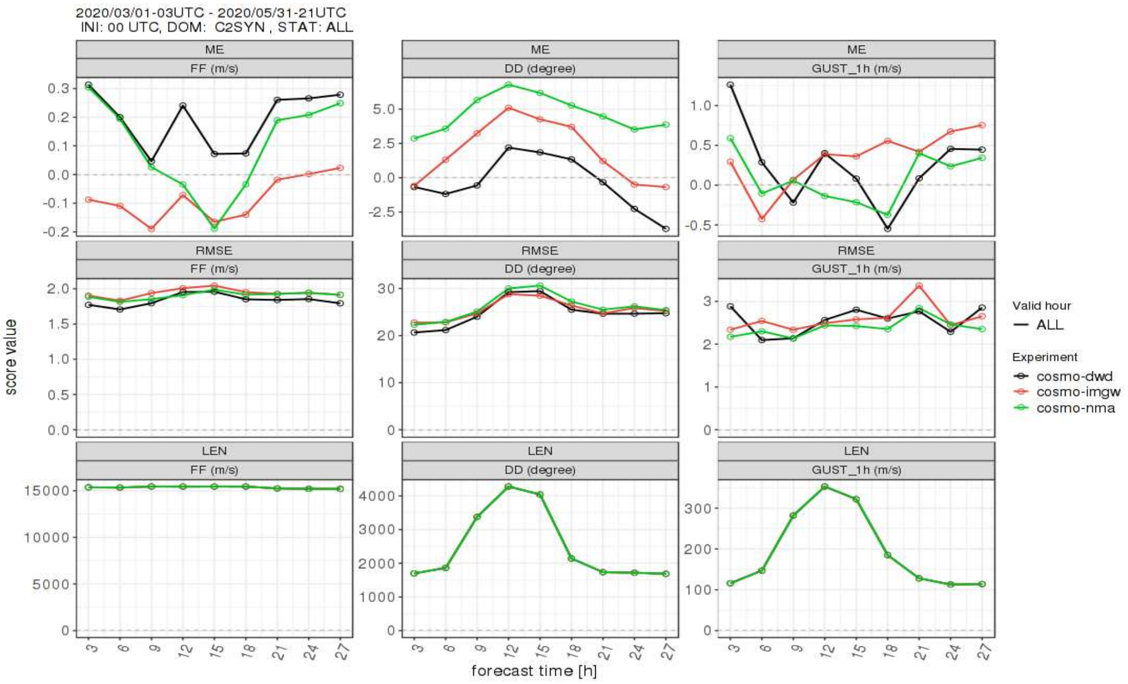


Figure 6: ME (top row), RMSE (middle row) and LEN (bottom row) values for CA2; left to right: FF, DD, Gust. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

As mentioned before, the observations used to compute the scores are the same for all three models, with lower number of observations for N, DD and especially Gust.

Categorical scores for cloud cover (figures 7 and 8) suggest a a high probability of detection from all three models, for both common areas (all thresholds). A similar behaviour can be observed from all three models, with slightly better performance from COSMO-D2, especially for the highest threshold. POD values are consistent with results for FAR, that show a low false alarm rate, again for all three models and both common areas. Slightly higher false alarm rates were obtained when forecasting high cloud cover values. FBI values show a good performance from all three models for this parameter, with a slight tendency to overcasting especially for the higher cloud cover values.

Categorical scores for 10 meter wind gust (1 hour) computed for CA1 (figure 9) show a generally higher probability of detection for the 12.5 m/s threshold than for the 15m/s one. For the first threshold, highest POD values are obtained during the day, with lower values starting with +21 hours anticipation, especially for COSMO-D2. For CA2 (figure 9), a lower probability of detection is observed especially for the 15m/s threshold. A high false alarm rate from all three models can be seen for the 15m/s threshold, especially from COSMO-RO for CA1 and all three models for CA2 in the second part of the considered forecast interval. For this threshold, a slightly better behaviour is exhibited by COSMO-RO for the first hours of forecast, while for the 12.5m/s threshold, the behaviour of the three models is similar for both areas. Values for the FBI score suggest a general tendency of overcasting from the three models for both areas, especially for the 12.5m/s threshold, with some undercasting of wind gust frequencies from COSMO-RO for the 15m/s threshold for the first anticipations, especially for CA2. For both areas, the behaviour of the model (mainly for the 12.5 m/s threshold) is more similar between COSMO-D2 and COSMO-RO.

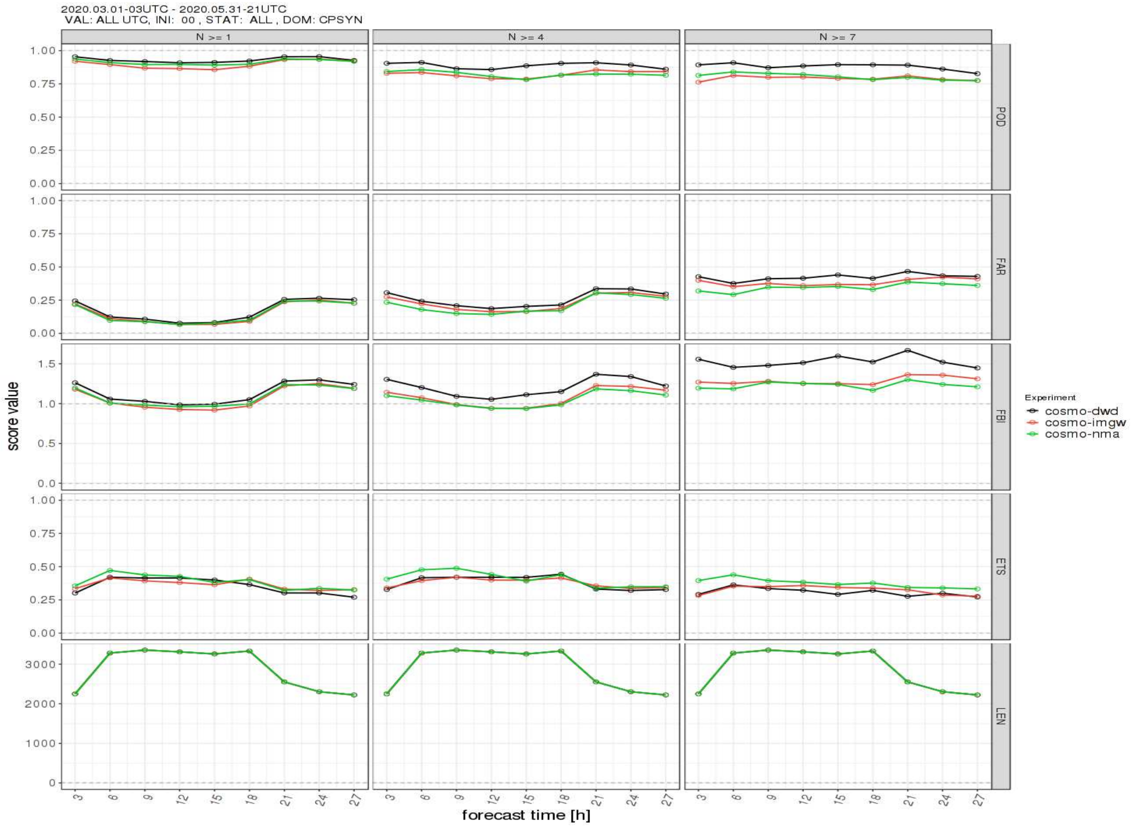


Figure 7: Categorical scores for N (CA1); top to bottom: POD, FAR, FBI, ETS and LEN for CA1; left to right: cloud cover >=1, >=4, >=7. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

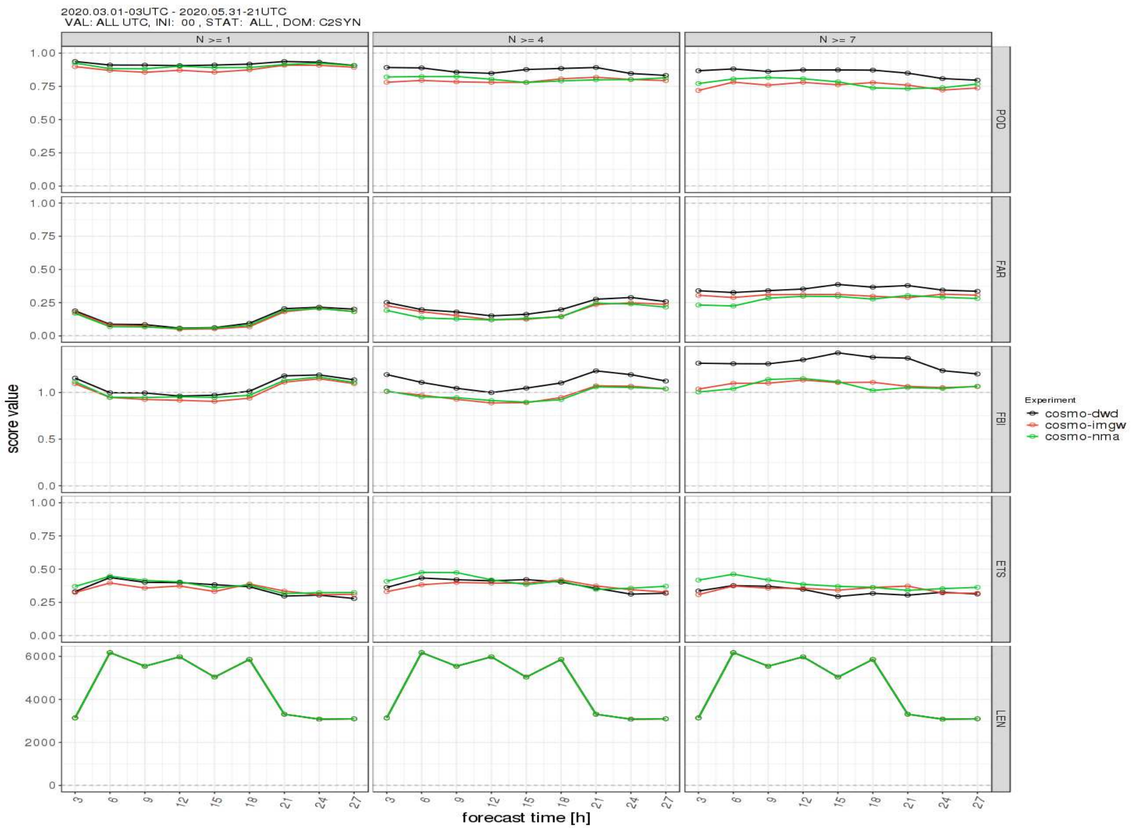


Figure 8: Categorical scores for N (CA2); top to bottom: POD, FAR, FBI, ETS and LEN for CA2; left to right: cloud cover >=1, >=4, >=7. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

Figure 9: Categorical scores for Gust (CA1); top to bottom: POD, FAR, FBI, ETS and LEN for CA1; left to right: cloud cover >=1, >=4, >=7. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

Figure 10: Categorical scores for Gust (CA2); top to bottom: POD, FAR, FBI, ETS and LEN for CA2; left to right: cloud cover >=1, >=4, >=7. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

Categorical scores for 6-hour accumulated precipitation are similar for both common areas (figures 11 and 12). Probability of detection values are better for the lower thresholds (0.2 mm/6h, 1 mm/6h) and worsen with the increase of the thresholds. POD values are consistent with results for FAR, indicating a low false alarm rate for the lower thresholds, with slightly better scores for COSMO-PL. FAR scores also worsen with the increase of threshold (both areas), with slightly smaller false alarm rates from COSMO-RO. FBI values suggest a general tendency to overcasting from all three models, with some undercasting from COSMO-PL and COSMO-RO for the higher thresholds and a slightly better performance from COSMO-D2. ETS values indicate a low skill for the upper thresholds (10mm/6h, 20mm/6h) and forecast quality drops significantly.

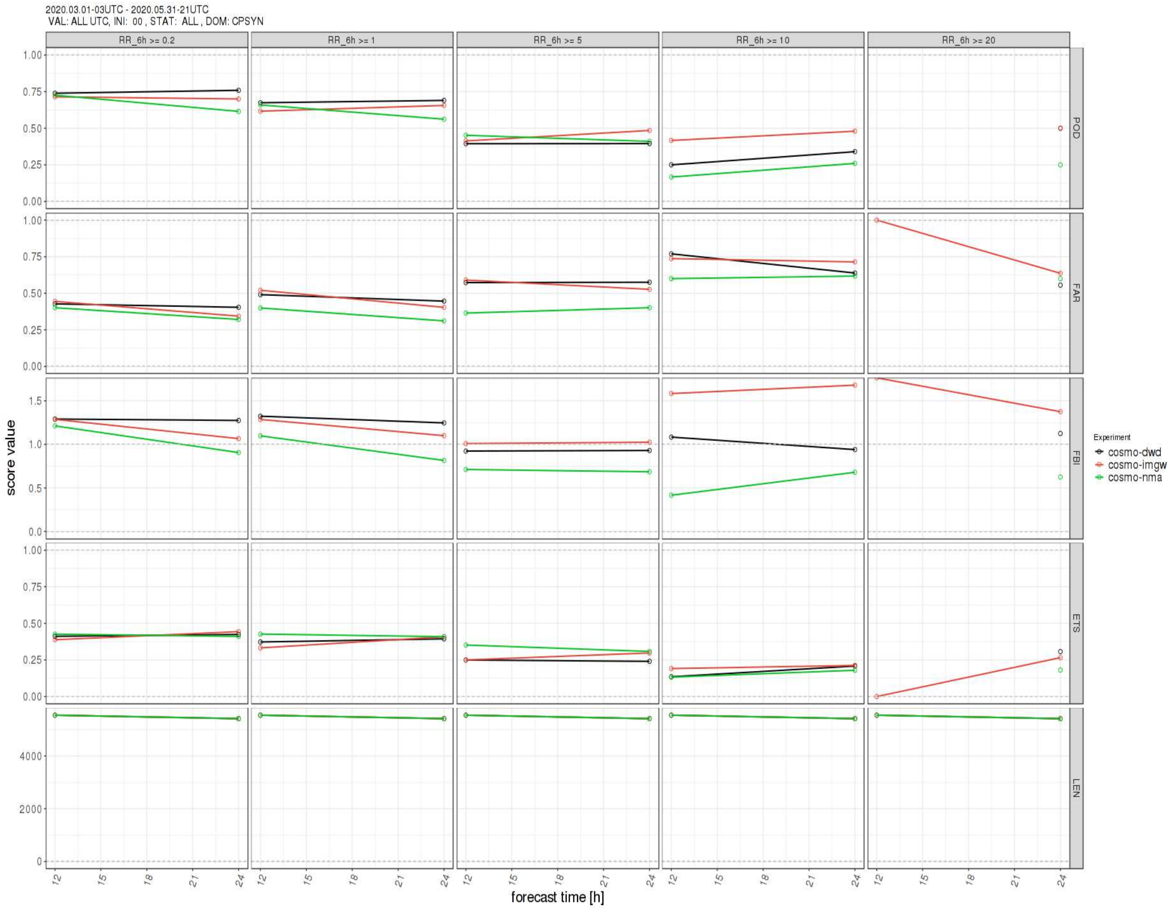


Figure 11: Categorical scores for 6-hour accumulated precipitation (CA1); top to bottom: POD, FAR, FBI, ETS and LEN for CA1; left to right: cloud cover >=1, >=4, >=7. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

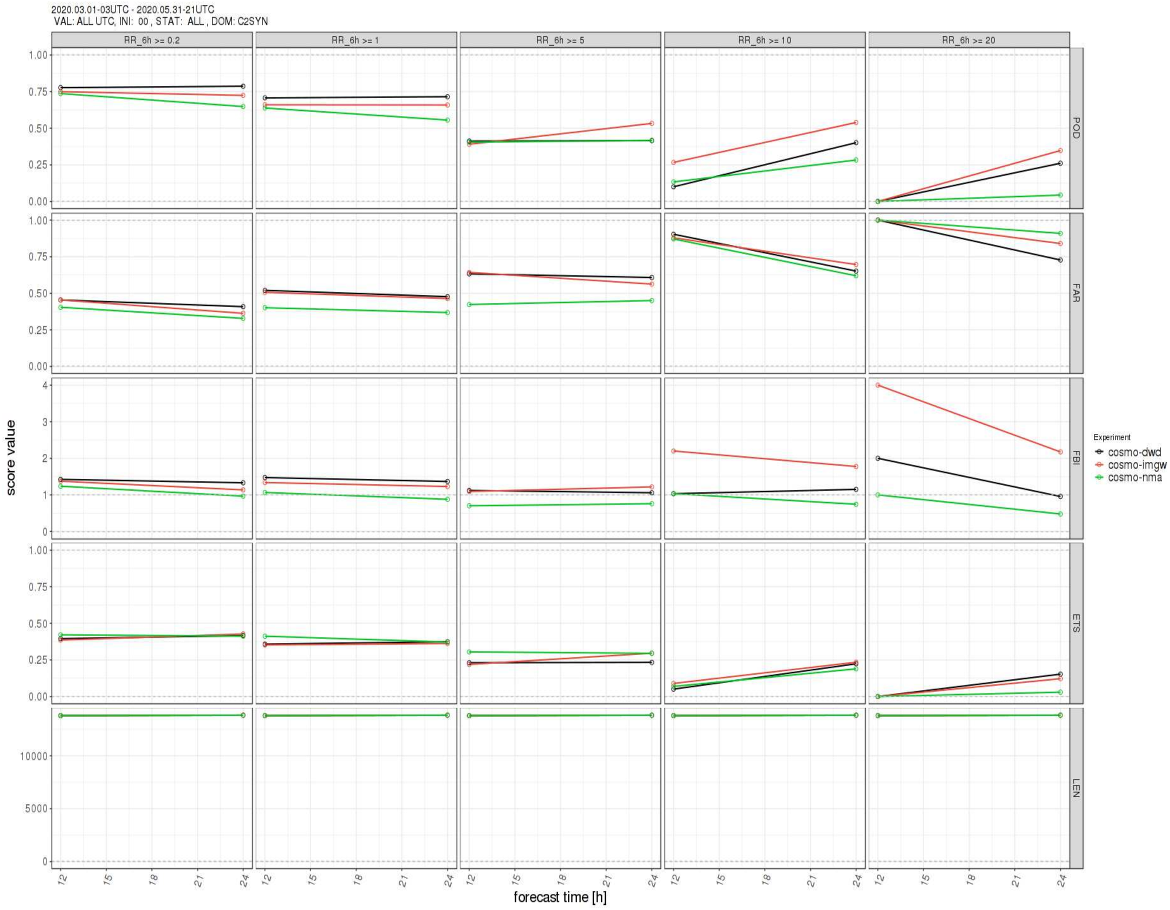


Figure 12: Categorical scores for 6-hour accumulated precipitation (CA2); top to bottom: POD, FAR, FBI, ETS and LEN for CA2; left to right: cloud cover >=1, >=4, >=7. COSMO-D2 (black), COSMO-PL (red) and COSMO-RO (green).

5 Conclusions and Outlook

Evaluation of operational COSMO (or ICON-LAM) model implementations in each service is part of the Common Plots Verification Activity, currently performed with the VERSUS software. These activities offer the opportunity to assess the performance of various operational COSMO (and ICON) model implementations over the same geographical areas, with the same set of observations. The existing VERSUS verification software environment is being replaced with the MEC-Rfdbk system as a Common Verification Software, with the latter being currently under implementation in all the member countries of the consortium.

For this purpose, cross model verification for the Common Plot activities for the 2020 spring season (MAM 2020) were performed with the MEC-Rfdbk software, with comparative results from three models already available. The verification activities with MEC-Rfdbk are carried out following the criteria and evaluation scores from the Common Plot tasks, in order to test the implementation of the new verification system.

The first results presented in this study show the comparative evaluation of three (COSMO) operational suites, while more operational COSMO (and ICON) model implementations will be further included in order to obtain a performance overview similar to that currently available from the Common Plot activities.

References

[1] Rhodin, A., 2015: MEC Manual, *DWD*

[2] Feedback-file definition, 2012: Supplementary Documentation, *DWD*

[3] Fundel, F., 2020: Feedback File Verification Suite at DWD, *DWD*

[4] Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., Reinhardt, T., 2011: Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Wea. Rev.*, 139, **3887–3905**, https://doi.org/10.1175/MWR-D-10-05013

[5] Zängl, G., Reinert, D., Ripodas, P., Baldauf, M., 2015: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quart. J. of the Royal Met. Soc.*, 141(687), **563–579**, https://doi.org/10.1002/qj.2378

[6] Consortium for Small-scale Modeling web-site: www.cosmo-model.org

[7] Consortium for Small-scale Modeling Verification tasks page: http://cosmo-model.org/content/tasks/verification.priv/default.htm

[8] Patruno, P., Cesari, D., 2011: Wreport-bufr2netcdf: a free library and tools for decoding BUFR reports and creating input files for COSMO-Model assimilation, 13th COSMO General Meeting, 5-9 September 2011, Rome (Italy), Parallel session: WG1 and PP KENDA

[9] Potthast, R., 2018: Introduction to Using MEC, datool, BACY and DACE, *DWD*

[10] Shiny Server page https://rstudio.com/products/shiny/shiny-server/

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

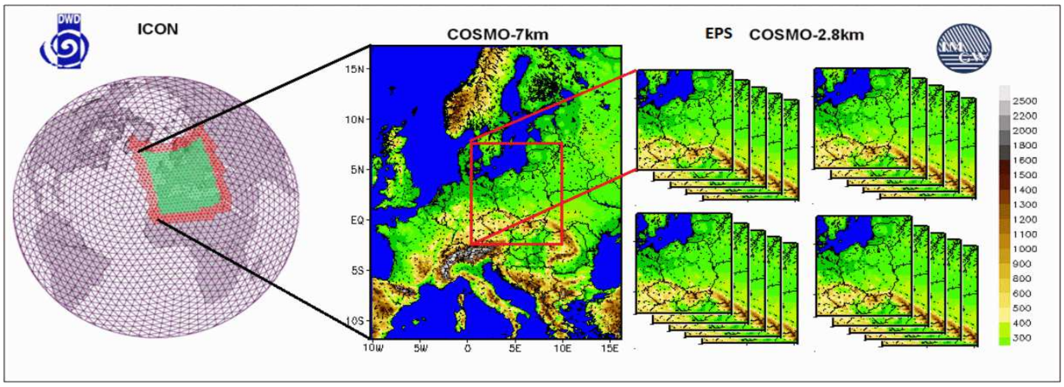


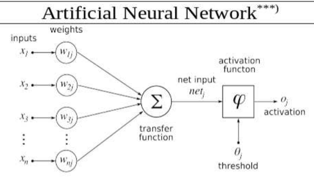
Figure 1: EPS operational configuration

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***)
$y = \frac{\sum_{i=1}^{m} x_i}{m}$	$\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ x_{21} & \cdots & x_{2p} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdots \\ \beta_p \end{pmatrix}$	
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).

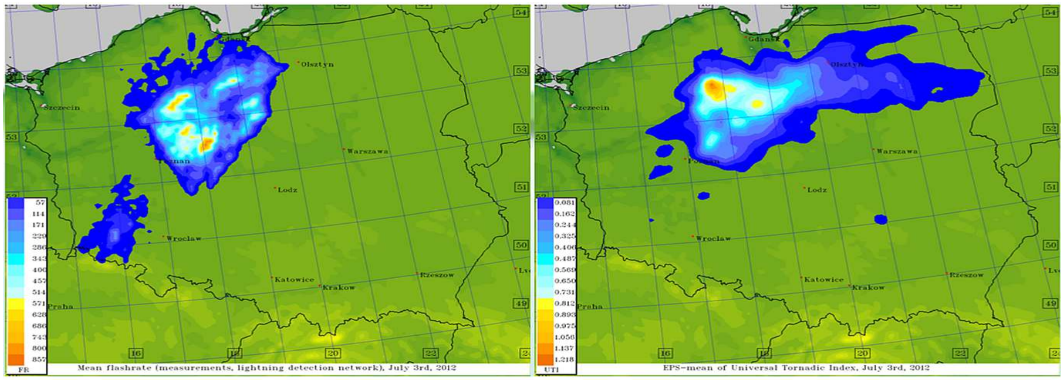


Figure 2: Example spatial distribution of detection (left) and forecast (right) of lightning

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.

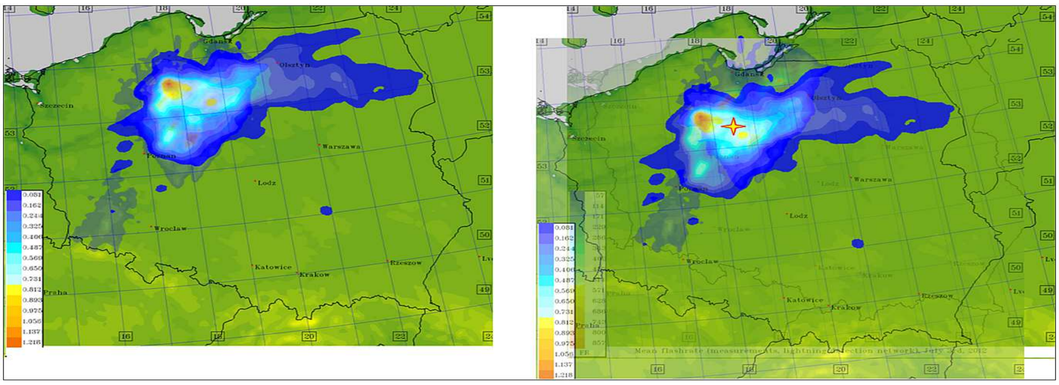


Figure 3: Forecast and detection of lightnings before (left) and after shifting using VOD (right). The centers of masses of both distributions are marked with an asterisk.

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

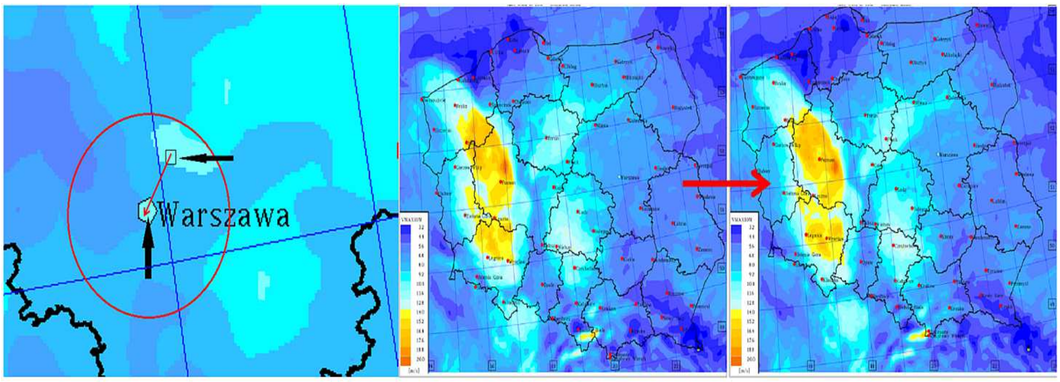


Figure 4: Determining the VOD for a synoptic station (left panel); average shift vector (middle and right)

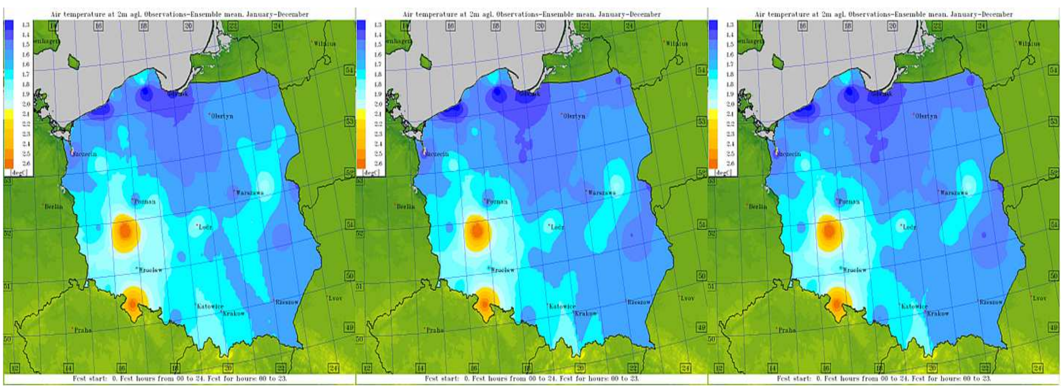


Figure 5: Results of different methods of calculating the average from ensemble forecasts. Left – observations vs. AM. Middle – observations vs. MLR average. Right – observations vs. average ANN. Air temperature forecasts, mean values for 2018.

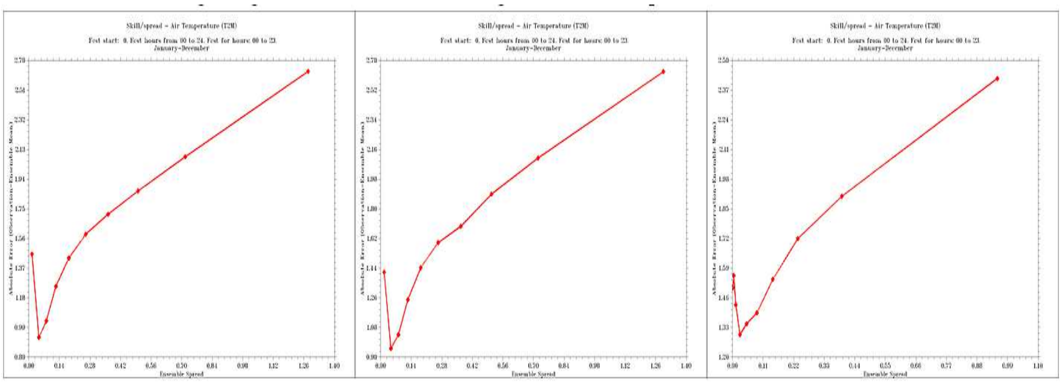


Figure 6: Skill-spread relationship. Left – arithmetic average, middle – MLR average, right – ANN average. Air temperature forecasts, average values for 2018.

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.

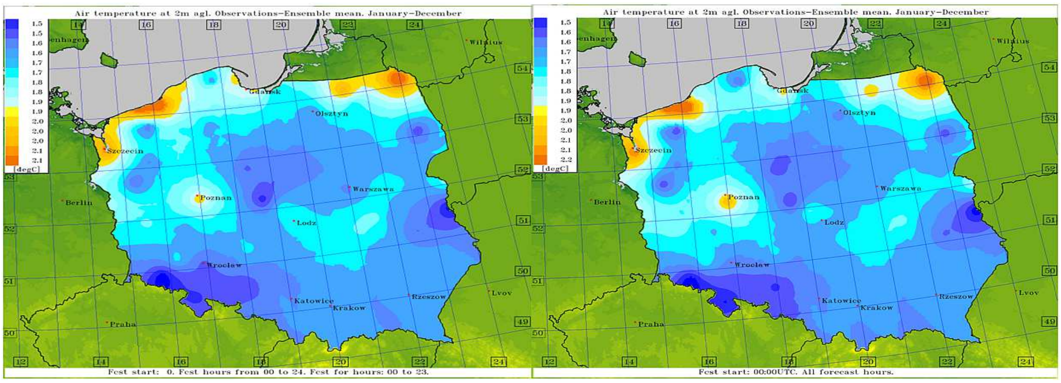


Figure 7: Distribution of skill values of air temperature, left – no VOD used, right – VOD applied, average for the period 2011-2013.

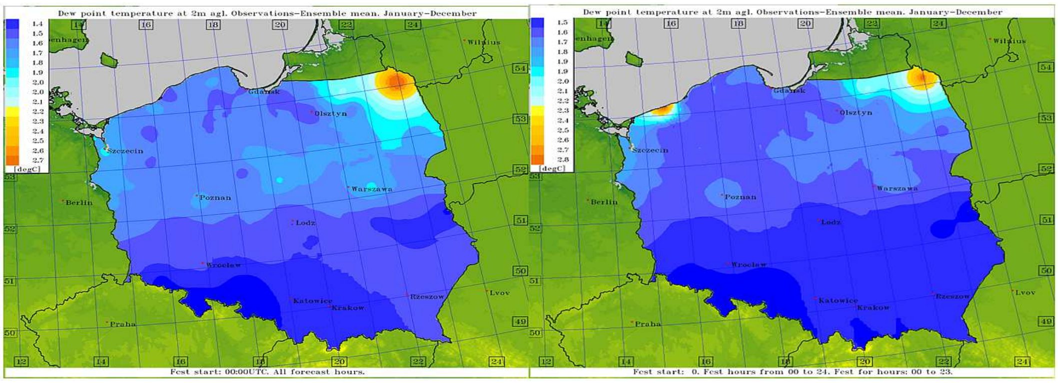


Figure 8: Distribution of skill values of dew point temperature, left – no VOD used, right – VOD applied, average for the period 2011-2013.

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.

EPS-Case study of serious HIW event in Poland, August 11th, 2017. Increasing resolution approach ˘ from 7 km to 0.7km

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

Abstract

The results from research on COSMO-EPS, carried out at IMWM, are presented. The operational EPS (Ensemble Prediction System) set-up is based on perturbations of soil surface-area index of the evaporating fraction of grid points over land. In this paper High Impact Weather event that occured on August 11th, 2017 in northwestern Poland is analyzed. This event resulted in two deaths, many people were injured, the final consequences are to be presented later on the basis of the prosecutor's investigation. The case was examined in both deterministic and EPS approaches, with a resolution increasing from 7 to 0.7 km.

1 Introduction

On August 27th, 2017, around 20:30 UTC, a very strong storm went over the Kujawsko-Pomorskie and Pomorskie voivodeships. The effects of this storm are being analyzed to this day. Two people were killed at the scout camp in Suszek, and many were wounded as a result of being crushed by breaking trees. Since then, the ongoing prosecutor's investigation led at its current stage to accusations against weather forecasters, claiming that they had not issued the highest level warning (red alarm), despite the fact that the wind speed exceeded 25 m/s, i.e. the threshold obliging such a warning. Measurements (not forecasts) at the nearest synoptic stations (Chojnice or Gdańsk) found wind gusts of up to 50 m/s, however, these measurements were made already during the period when the storm was passing through this area. This work attempts to answer the question whether numerical weather forecasts, in particular ensemble forecasts, predicted such wind speed, that is, whether there was a basis for prior announcement of warning of the highest degree. In this work, the ensemble prediction system (EPS) was used, which has been running in IMGW-PIB since 2013. A detailed description of the system can be found in the work (Duniec et al, 2016). In this approach, a cascading sequence of nested domains from a resolution of 7 km through 2.8 km to 0.7 km was used. The cascade of domains is shown in Figure 1, with the location of the town of Suszek marked.

doi:10.5676/dwd_pub/nwv/cosmo-nl_20_06

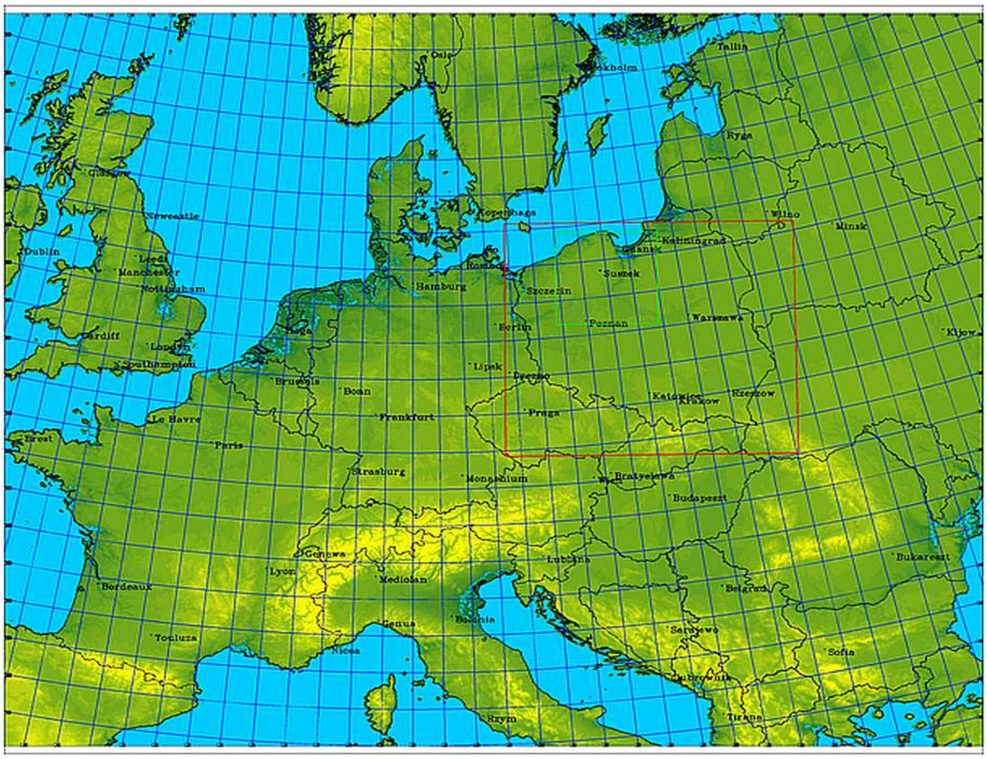


Figure 1: The basic computational domain of the COSMO model at a resolution of 7x7km, mesh size: 385x321 points. The red rectangle marked the nested domain 2.8x2.8km, (285x255 points), the green – nested domain 0.7x0.7km (400x400 points).

The basic input data to the cascade (EPS input data with a resolution of 7km) were obtained from the results of the ICON model (Zängl et al, 2015), calculations at a resolution of 2.8km were performed using the results of the model at 7km resolution as initial and boundary conditions. Finally, the 2.8km model results served as boundary and initial conditions for EPS at a resolution of 0.7km. The result of the model's work were the values (spatial distributions) of parameters related to the description of convective phenomena, including Total Precipitation, Windspeed at 10m agl., Maximum windspeed at 10m agl., CAPE_3KM, CAPE_ML, CAPE_MU, Derecho Composite Potential, Supercell Detection Index, Showalter Index, Lifted Index, Universal Tornadic Index (conf. Taszarek and Kolendowicz, 2013), Radar Reflectivity and Wind Shear up to 6 km. In this paper, however, only the results of the two basic parameters are presented: forecasts of maximum wind speed (VMAX) and radar reflectance (REFL).

2 Results

Figures 2-7 provide forecasts for the distribution of mentioned parameters (VMAX, REFL) for all resolutions, from 19:00 to 22:00 UTC. Forecasts are presented both as EPS results and as results of a deterministic approach. In the case of EPS, the distributions of the ensemble mean and the maximum predicted values are presented. Additionally, Figure 8 presents comparisons of radar reflection forecasts with real images obtained with the Polish radar network, hourly accumulated from 15:00 to 23:00 UTC (see also Taszarek et al, 2019). The exemplary areas and convective structures (with locations similar for ensemble mean/max/deterministic spatial distributions), that can be identified from reflections forecasts vs. the observations, are marked with arrows.

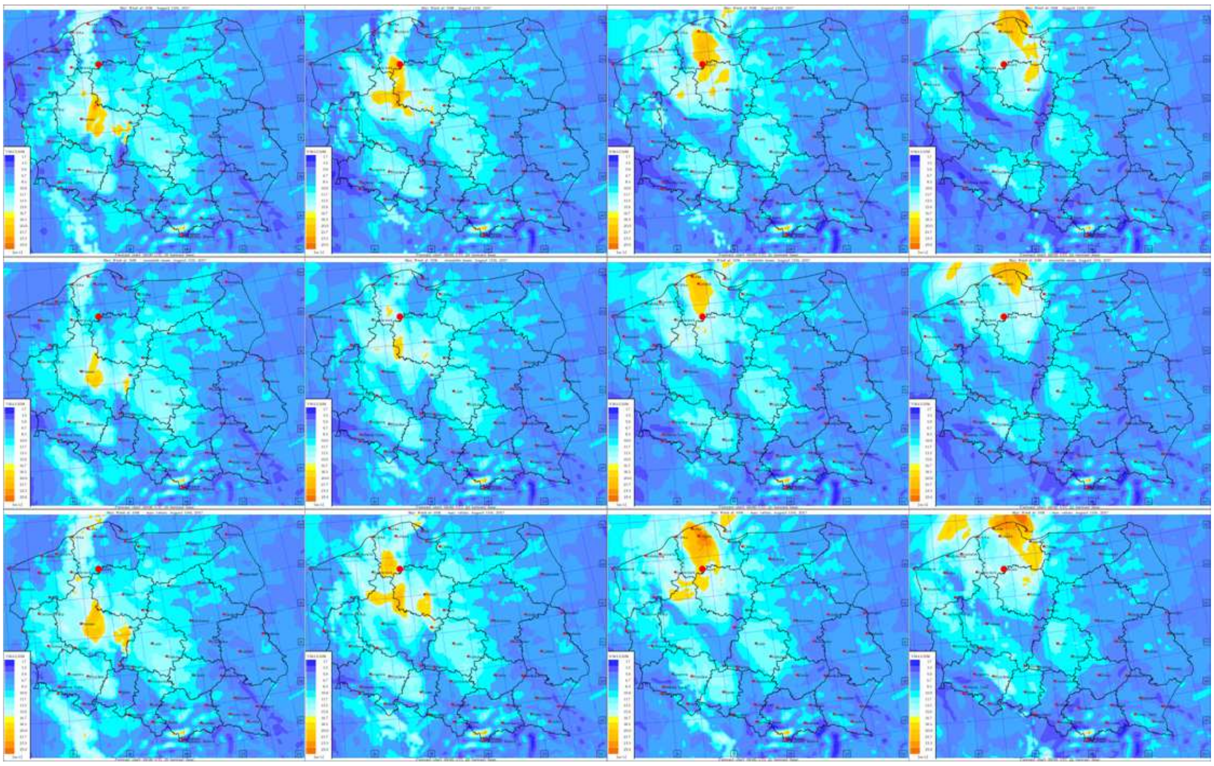


Figure 2: VMAX forecasts from 19:00 (leftmost) to 22:00 (rightmost) UTC. Top to bottom: deterministic forecast, ensemble mean, ensemble maximum values. Resolution 7km.

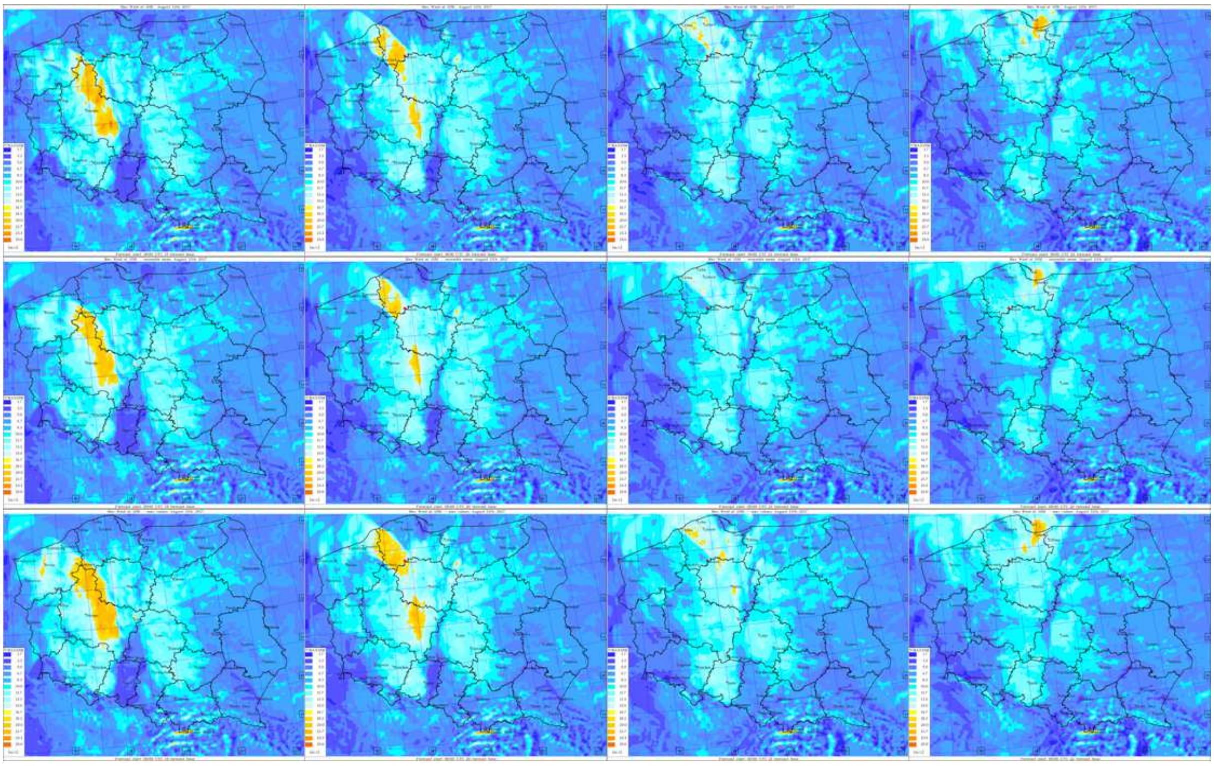


Figure 3: VMAX forecasts from 19:00 (leftmost) to 22:00 (rightmost) UTC. Top to bottom: deterministic forecast, ensemble mean, ensemble maximum values. Resolution 2.8km.

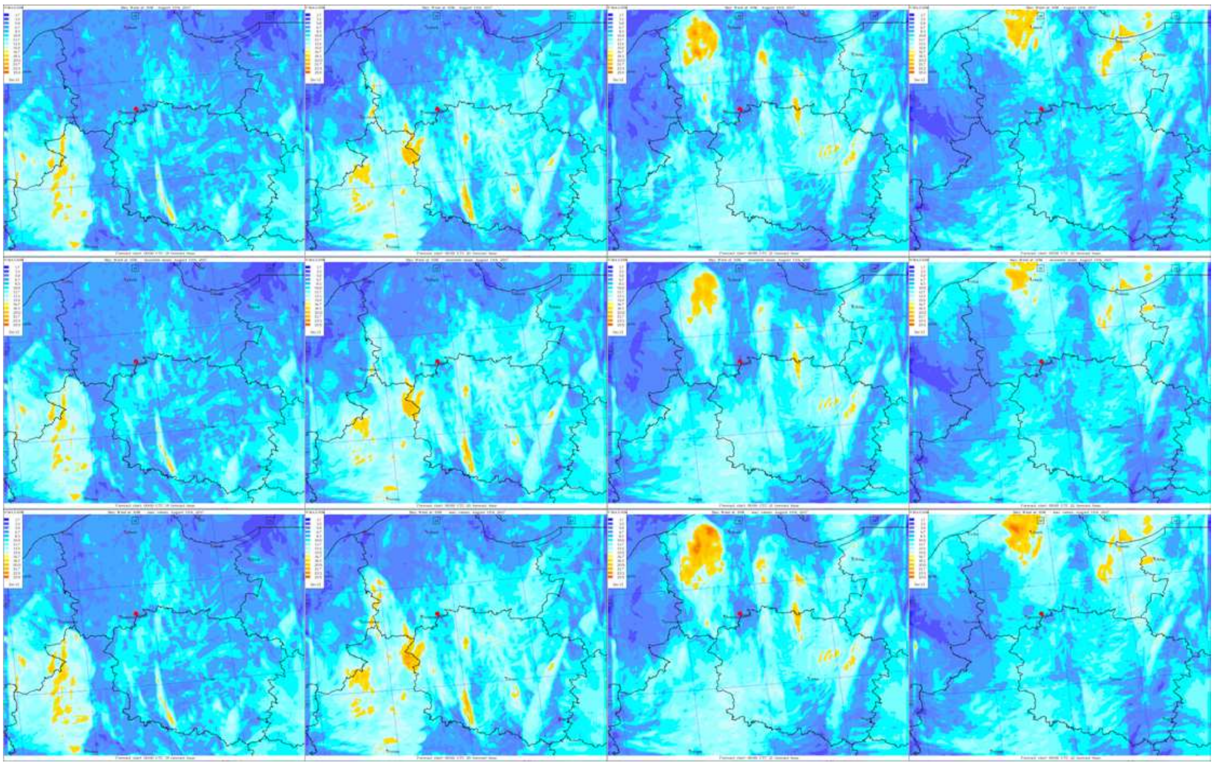


Figure 4: VMAX forecasts from 19:00 (leftmost) to 22:00 (rightmost) UTC. Top to bottom: deterministic forecast, ensemble mean, ensemble maximum values. Resolution 0.7km.

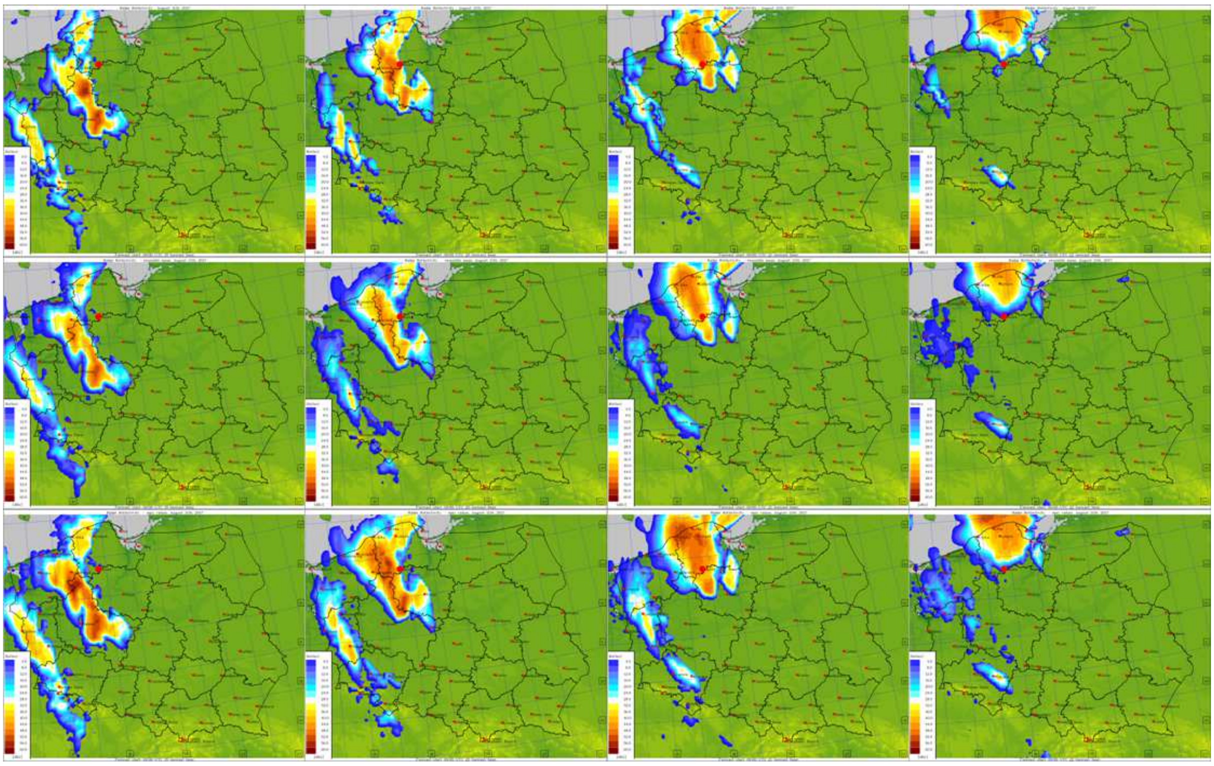


Figure 5: REFL forecasts from 19:00 (leftmost) to 22:00 (rightmost) UTC. Top to bottom: deterministic forecast, ensemble mean, ensemble maximum values. Resolution 7km.

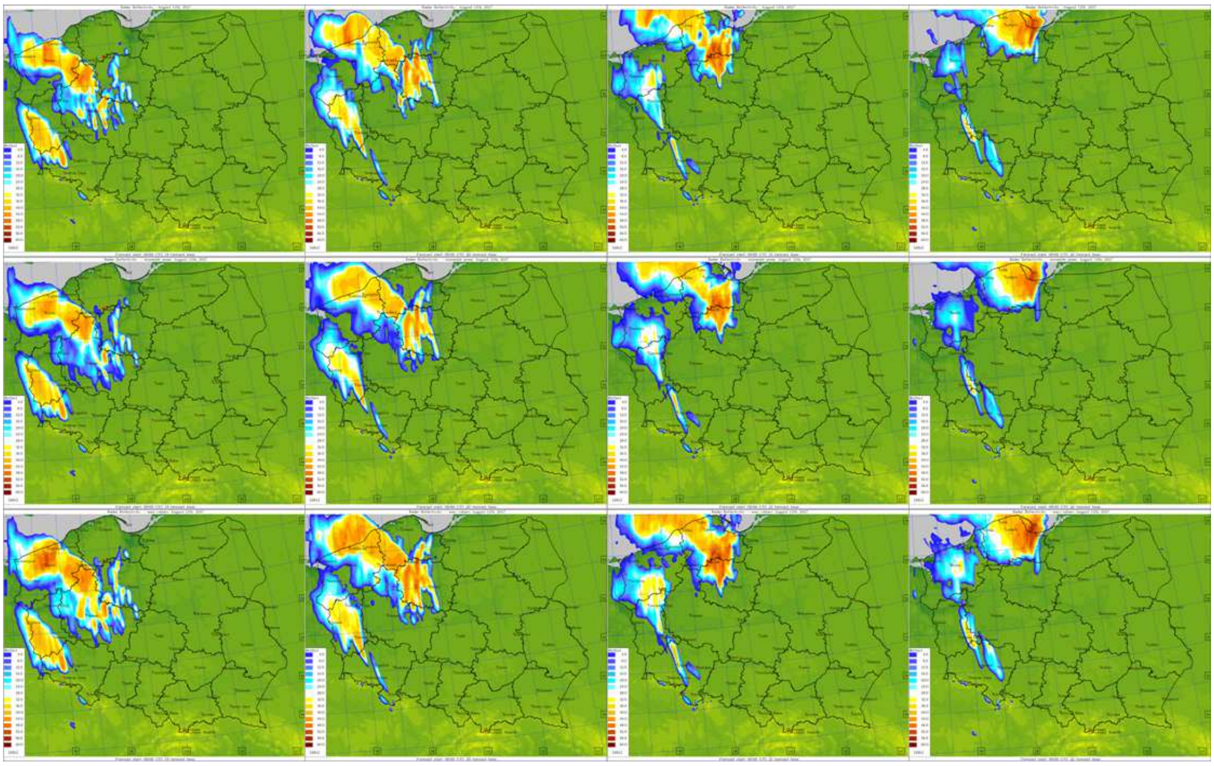


Figure 6: REFL forecasts from 19:00 (leftmost) to 22:00 (rightmost) UTC. Top to bottom: deterministic forecast, ensemble mean, ensemble maximum values. Resolution 2.8km.

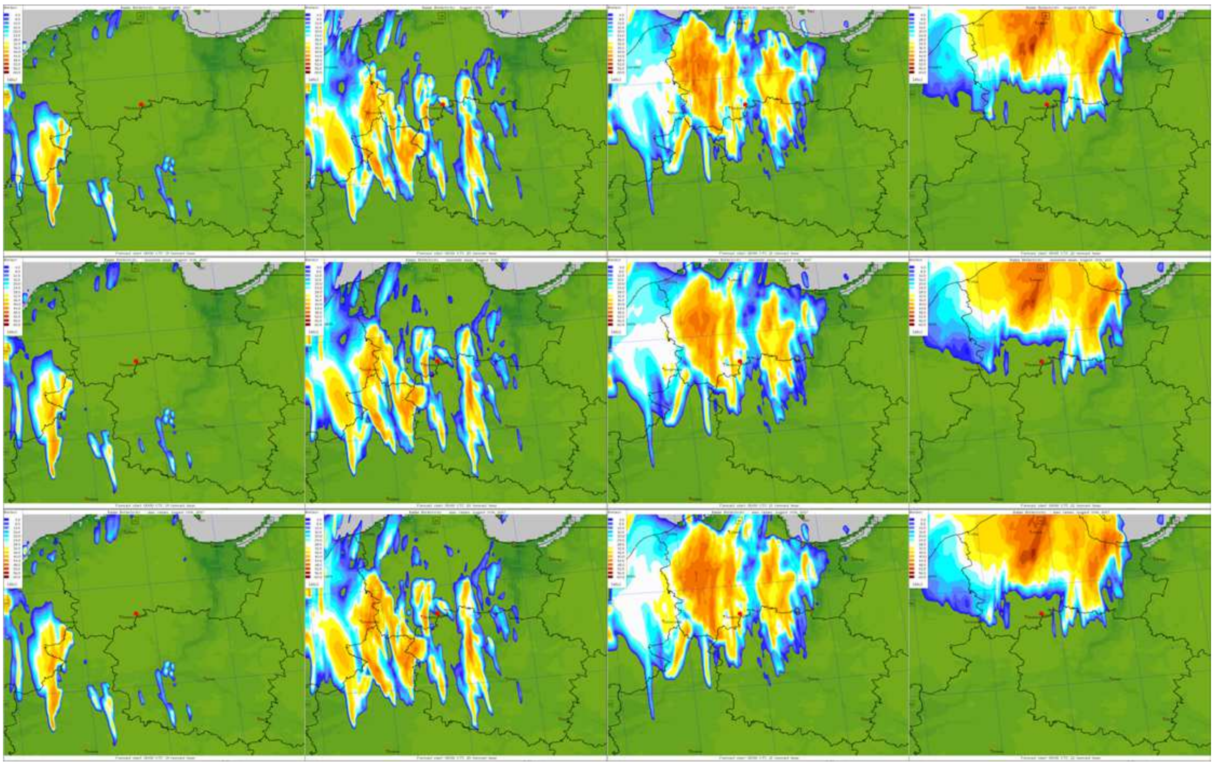


Figure 7: REFL forecasts from 19:00 (leftmost) to 22:00 (rightmost) UTC. Top to bottom: deterministic forecast, ensemble mean, ensemble maximum values. Resolution 0.7km.

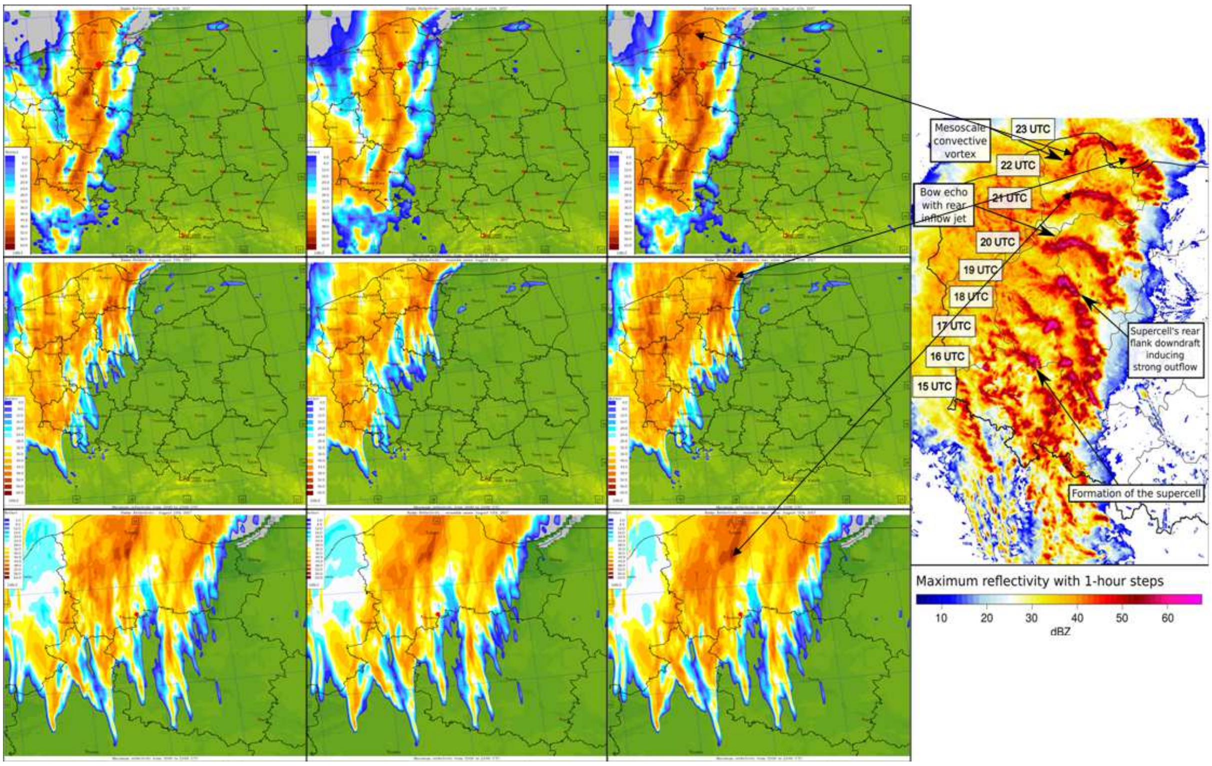


Figure 8: Radar reflectivity forecasts from 15:00 to 23:00 UTC. Top to bottom: resolution 7.0km, 2.8km, 0.7km. Left to right: deterministic forecast, ensemble mean and maximum values. Right-most: Maximum reflectivity as observed with Polish radar network.

3 Discussion

It should be stressed that most likely this event was cause by a supercell that rapidly moved north-northeast. Important information about the event was provided by VMAX forecasts(as DMO) and reflectivity forecasts, which could be verified in relation to measurements at SYNOP stations and the values in the Polish radar network (conf. Taszarek et al, ibid). The information on reflectivity forecasts may serve as the method to identify structures relevant to high impact weather events, especially an intensive convective phenomena. Results of verification are similar for both EPS-and deterministic forecasts for the case. Low wind speeds (below 5.0 m/s) are overestimated by all instances of EPS deterministic forecasts whilst larger wind speeds (above 5 m/s) are underestimated. The skill decreases with wind speed and smaller spatial scales. Wind gusts above 6 m/s are underestimated and the degree of underestimation is increasing with gust speed. The best skill was achieved for low wind gust speeds. A small variation in skill with spatial scale, with greater skill at larger scales, can be seen.

4 Summary

However, forecasts of the maximum wind gusts, both in statistical and deterministic approach, in all considered resolutions did not exceed 25 m/s. Hence, the maximum forecast wind was not strong enough to give substantial reason for the warning of the highest level. In the light of this study, the accusations against forecasters should be considered unfounded.

Thus, the answer the basic question asked in this article (whether there was a basis for issuing of warning of the highest degree) should read as follows: from all of the above it can be stated that no forecast, neither deterministic nor EPS, predicted winds faster than 25m/s (regardless, mean or gusts) that would be a basis for giving out the top-level – red – alert.

As it was already said, the transition from 7.0 km via 2.8 km to 0.7 km was not just the increasing the resolution. It was also the transition from parameterized to explicit convection, that changes the results with every change of resolution.

Finally, it should be stressed that at IMWM-NRI this work was the very first attempt of forecasting model with resolution higher than 1km. Further work on this topic is expected.

Acknowledgements

The results in a oral form was presented during the Workshop on "LAM-EPS prediction of high impact weather and extremes" in th frame of EUMETNET SRNWP-EPS 2019-2023 programme in Madrid, Spain, October 2019. The activity and participation in the workshop was reimbursed from COSMO funds.

References

[1] Duniec G, Interewicz W, Mazur A, Wyszogrodzki A (2017) 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. vol. 5; (2): 43-51

[2] Taszarek M, Kolendowicz L (2013) Sounding-derived parameters associated with tornado occurrence in Poland and Universal Tornadic Index. Atmos. Res., **134**, 186-197.

[3] Taszarek M, Pilguj N, Orlikowski J, et al (2019) Derecho Evolving from a Mesocyclone - A Study of 11 August 2017 Severe Weather Outbreak in Poland: Event Analysis and High-Resolution Simulation. Mon.Wea.Rev., $https://doi.org/10.1175/MWR-D-18-0330.1, 2019$.

List of COSMO Newsletters and Technical Reports

(available for download from the COSMO Website: `www.cosmo-model.org`)

COSMO Newsletters

No. 1: February 2001.

No. 2: February 2002.

No. 3: February 2003.

No. 4: February 2004.

No. 5: April 2005.

No. 6: July 2006; Proceedings from the COSMO General Meeting 2005.

No. 7: May 2008; Proceedings from the COSMO General Meeting 2006.

No. 8: August 2008; Proceedings from the COSMO General Meeting 2007.

No. 9: December 2008; Proceedings from the COSMO General Meeting 2008.

No.10: January 2010; Proceedings from the COSMO General Meeting 2009.

No.11: February 2011; Proceedings from the COSMO General Meeting 2010.

No.12: March 2012; Proceedings from the COSMO General Meeting 2011.

No.13: April 2013; Proceedings from the COSMO General Meeting 2012.

No.14: April 2014; Proceedings from the COSMO General Meeting 2013.

No.15: July 2015; Proceedings from the COSMO General Meeting 2014.

No.16: June 2016; Proceedings from the COSMO General Meeting 2015.

No.17: July 2017; Proceedings from the COSMO General Meeting 2016.

No.18: November 2018; Proceedings from the COSMO General Meeting 2017.

No.19: October 2019; Proceedings from the COSMO General Meeting 2018.

No.20: December 2020; Proceedings from the COSMO General Meeting 2019.

COSMO Technical Reports

No. 1: Dmitrii Mironov and Matthias Raschendorfer (2001):
Evaluation of Empirical Parameters of the New LM Surface-Layer Parameterization Scheme. Results from Numerical Experiments Including the Soil Moisture Analysis.

No. 2: Reinhold Schrodin and Erdmann Heise (2001):
The Multi-Layer Version of the DWD Soil Model TERRA_LM.

No. 3: Günther Doms (2001):
A Scheme for Monotonic Numerical Diffusion in the LM.

No. 4: Hans-Joachim Herzog, Ursula Schubert, Gerd Vogel, Adelheid Fiedler and Roswitha Kirchner (2002):
LLM — the High-Resolving Nonhydrostatic Simulation Model in the DWD-Project LITFASS.
Part I: Modelling Technique and Simulation Method.

No. 5: Jean-Marie Bettems (2002):
EUCOS Impact Study Using the Limited-Area Non-Hydrostatic NWP Model in Operational Use at MeteoSwiss.

No. 6: Heinz-Werner Bitzer and Jürgen Steppeler (2004):
Description of the Z-Coordinate Dynamical Core of LM.

No. 7: Hans-Joachim Herzog, Almut Gassmann (2005):
Lorenz- and Charney-Phillips vertical grid experimentation using a compressible nonhydrostatic toy-model relevant to the fast-mode part of the 'Lokal-Modell'

No. 8: Chiara Marsigli, Andrea Montani, Tiziana Paccagnella, Davide Sacchetti, André Walser, Marco Arpagaus, Thomas Schumann (2005):
Evaluation of the Performance of the COSMO-LEPS System

No. 9: Erdmann Heise, Bodo Ritter, Reinhold Schrodin (2006):
Operational Implementation of the Multilayer Soil Model

No. 10: M.D. Tsyrulnikov (2007):
Is the particle filtering approach appropriate for meso-scale data assimilation?

No. 11: Dmitrii V. Mironov (2008):
Parameterization of Lakes in Numerical Weather Prediction. Description of a Lake Model.

No. 12: Adriano Raspanti (2009):
Final report on priority project VERSUS (VERification System Unified Survey).

No. 13: Chiara Mirsigli (2009):
Final report on priority project SREPS (Short Range Ensemble Prediction System).

No. 14: Michael Baldauf (2009):
COSMO Priority Project "Further Developments of the Runge-Kutta Time Integration Scheme" (RK); Final Report.

No. 15: Silke Dierer (2009):
COSMO Priority Project "Further Developments of the Runge-Kutta Time Integration Scheme" (RK); Final Report.

No. 16: Pierre Eckert (2009):
COSMO Priority Project "INTERP"; Final Report.

No. 17: D. Leuenberger, M. Stoll, A. Roches (2010):
Description of some convective indices, implemented in the COSMO model.

No. 18: Daniel Leuenberger (2010):
Statistical Analysis of high-resolution COSMO Ensemble forecasts, in view of Data Assimilation.

No. 19: A. Montani, D. Cesari, C. Marsigli, T. Paccagnella (2010):
Seven years of activity in the field of mesoscale ensemble forecasting by the COSMO-LEPS system: main achievements and open challenges.

No. 20: A. Roches, O. Fuhrer (2012):
Tracer module in the COSMO model.

No. 21: M. Baldauf (2013):
A new fast-waves solver for the Runge-Kutta dynamical core.

No. 22: C. Marsigli, T. Diomede, A. Montani, T. Paccagnella, P. Louka, F. Gofa, A. Corigliano (2013):
The CONSENS Priority Project.

No. 23: M. Baldauf, O. Fuhrer, M. J. Kurowski, G. de Morsier, M. Muellner, Z. P. Piotrowski, B. Rosa, P. L. Vitagliano, D. Wojcik, M. Ziemianski (2013):
The COSMO Priority Project 'Conservative Dynamical Core' Final Report.

No. 24: A. K. Miltenberger, A. Roches, S. Pfahl, H. Wernli (2014):
Online Trajectory Module in COSMO: A short user guide.

No. 25: P. Khain, I. Carmona, A. Voudouri, E. Avgoustoglou, J.-M. Bettems, F. Grazzini (2015):
The Proof of the Parameters Calibration Method: CALMO Progress Report.

No. 26: D. Mironov, E. Machulskaya, B. Szintai, M. Raschendorfer, V. Perov, M. Chumakov, E. Avgoustoglou (2015):
The COSMO Priority Project 'UTCS' Final Report.

No. 27: Jean-Marie Bettems (2015):
The COSMO Priority Project 'COLOBOC' Final Report.

No. 28: Ulrich Blahak (2016):
RADAR_MIE_LM and RADAR_MIELIB - Calculation of Radar Reflectivity from Model Output.

No. 29: M. Tsyrulnikov, D. Gayfulin (2016):
A Stochastic Pattern Generator for ensemble applications.

No. 30: Dmitrii Mironov, Ekaterina Machulskaya (2017):
A Turbulence Kinetic Energy - Scalar Variance Turbulence Parameterization Scheme.

No. 31: P. Khain, I. Carmona, A. Voudouri, E. Avgoustoglou, J.-M. Bettems, F. Grazzini, P. Kaufmann (2017):
CALMO - Progress Report.

No. 32: A. Voudouri, P. Khain, I. Carmona, E. Avgoustoglou, J.M. Bettems, F. Grazzini, O. Bellprat, P. Kaufmann and E. Bucchignani (2017):
Calibration of COSMO Model, Priority Project CALMO Final report.

No. 33: Naim Vela (2017):
V.A.S.T. (Versus Additional Statistical Techniques) User Manual (v2.0).

No. 34: C. Marsigli, D. Alferov, M. Arpagaus, E. Astakhova, R. Bonanno, G. Duniec, C. Gebhardt, W. Interewicz, N. Loglisci, A. Mazur, V. Maurer, A. Montani, A. Walser (2018):
COsmo Towards Ensembles at the Km-scale IN Our countries" (COTEKINO), Priority Project final report.

No. 35: G. Rivin, I. Rozinkina, E. Astakhova, A. Montani, D. Alferov, M. Arpagaus, D. Blinov, A. Bundel, M. Chumakov, P. Eckert, A. Euripides, J. Foerstner, J. Helmert, E. Kazakova, A. Kirsanov, V. Kopeikin, E. Kukanova, D. Majewski, C. Marsigli, G. de Morsier, A. Muravev, T. Paccagnella, U. Schaettler, C. Schraff, M. Shatunova, A. Shcherbakov, P. Steiner, M. Zaichenko (2017):
The COSMO Priority Project CORSO Final Report.

No. 36: A. Raspanti, A. Celozzi, A. Troisi, A. Vocino, R. Bove, F. Batignani(2018):
The COSMO Priority Project VERSUS2 Final Report

No. 37: INSPECT Final Report A. Bundel, F. Gofa, D. Alferov, E. Astakhova, P. Baumann, D. Boucouvala, U. Damrath, P. Eckert, A. Kirsanov, X. Lapillonne, J. Linkowska, C. Marsigli, A. Montani, A. Muraviev, E. Oberto, M.S. Tesini, N. Vela, A. Wyszogrodzki, M. Zaichenko, A. Walser(2019):
The COSMO Priority Project INSPECT Final Report

No. 38: G. Rivin, I. Rozinkina, E. Astakhova, A. Montani, J-M. Bettems, D. Alferov, D. Blinov, P. Eckert, A. Euripides, J. Helmert, M.Shatunova(2019):
The COSMO Priority Project CORSO-A Final Report

No. 39: C. Marsigli, D. Alferov, E. Astakhova, G. Duniec, D. Gayfulin, C. Gebhardt, W. Interewicz, N. Loglisci, F. Marcucci, A. Mazur, A. Montani, M. Tsyrulnikov, A. Walser (2019):
Studying perturbations for the representation of modeling uncertainties in Ensemble development (SPRED Priority Project): Final Report

No. 40: E. Bucchignani, P. Mercogliano, V. Garbero, M. Milelli, M. Varentsov, I. Rozinkina, G. Rivin, D. Blinov, A. Kirsanov, H. Wouters, J.-P. Schulz, U. Schaettler(2019):
Analysis and Evaluation of TERRA_URB Scheme: PT AEVUS Final Report

No. 41: X. Lapillonne, O. Fuhrer(2020):
Performance On Massively Parallel Architectures (POMPA): Final report

No. 42: E. Avgoustoglou, A. Voudouri, I Carmona, E. Bucchignani, Y. Levy, J. -M. Bettems (2020):
A methodology towards the hierarchy of COSMO parameter calibration tests via the domain sensitivity over the Mediterranean area