Cosmo Application of Rfdbk/MEC on ENS (CARMENs)

Version Date: 09.08.2022

Project Duration: 09.22-09.24

FTEs request: ~3.27

Project Leader: Stefan Ioan Gabrian (NMA)

Main goal

The main goal of the PP is to extend the implementation and usage of the MEC-Rfdbk system to the evaluation of EPS model outputs, which is beyond the scope of the previously completed CARMA PP. Available statistical results for selected time periods of ensemble COSMO and ICON-LAM based systems over national domains will be produced and published on the COSMO Verification web page. Additionally, an extension of Common plot activities to EPS not only over national domains, but selectively even over common areas will be assessed and applied while the requirements for such activity will be defined.

Background Information

In the frame of the CARMA Priority Project (completed 2021), the VERSUS verification software environment was replaced with the MEC-Rfdbk software developed by DWD, as the **Common Verification Software (CVS)** used to perform most of the verification activities in the consortium. The MEC-Rfdbk system is mainly used for the production of the **Common Plot (CP)** verification; spatial verification is performed with other available tools such as VAST.

Following the finalization of the CARMA project, the MEC-Rfdbk system was implemented and is currently running operationally in most countries of the consortium. At present, the following type of products are obtained for **CP activities**:

- Categorical scores for Gust, RR_6h and N;
- Scores for continuous parameters;
- Scores for upper air parameters;
- Comparison between two models showing the trend in various scores;
- Domain average and station based verification;
- Common Area and national domain stratification.

Remaining open issues that concern MEC-Rfdbk adaptation in some services and implementation of new features for CP activities are performed regularly through WG6 SPRT Common Plot activity. Detailed documentation and templates for the use of the MEC-Rfdbk system are also available for use both for CP production and national verification activities (deterministic features).

Motivation

The MEC-Rfdbk system is also suitable and, at some centers, already used for EPS verification. This type of application was not included in the CARMA project, since EPS verification is not currently part of CP activity. As the number of centers producing ensemble forecasts is increasing, so is the demand for a proper ensemble verification framework. Currently, the ensemble systems run operationally by COSMO members are: COSMO-LEPS (COSMO), ICON-D2-EPS (DWD), COSMO-2E and COSMO-1E (MCH), TLE-MVE (IMGW), COSMO-IT-
EPS (COMET), COSMO-2I-EPS (ArpaE), COSMO-IL-ENS (IMS). In order to coordinate the support effort necessary to get the EPS verification running at the individual centers, a follow up project of CARMA, with focus on EPS verification is proposed.

Besides the availability for the different COSMO members of a state-of-the art software package for the verification of the ensembles, the extension of CARMA to EPS will bring two main advantages:

- the usage of the same software for deterministic and probabilistic verification, which guarantees that the data are processed in the same way in the verification process and permits to save resources in the development and implementation of auxiliary routines, like quality control, correction for height, interpolation methods, …
- the usage of the same verification software for assessing ensemble performance in the different COSMO Centres, which ensures that the same processing of the data in the verification process is applied and a unified definition of the indices is used for verification. This makes the scores from the different ensembles more comparable, strengthening the scientific exchange.

Finally, this project is a first step towards extending the Common Plot activity to probabilistic systems both over national and over common areas, whose feasibility will be also explored in the project.

Methodology and System Features

As mentioned, there are a number of operational ensemble forecast systems running from COSMO members that require an approach for their evaluation different from the deterministic ones. In general, the main approaches to evaluate this kind of forecast products is given below:

Verification of an ensemble distribution. The methods used to verify ensemble distributions are: rank histograms (Talagrand diagrams), continuous rank probability score and its related skill score (a measure for the difference between the cumulative distribution function of the forecasts and observations).
Verification measures for the pdf of a generic probability forecast. The methods used are "linear probability" and the "ignorance score". Both the methods are local, evaluating the pdf at the observation value, and both reward sharp distributions which are also accurate.
Verification of the probability of an event. The methods used are: Brier score and its decomposition, Brier skill score, reliability diagrams, ROC curves and area, and the rank probability score.

The above verification techniques and others focused on extreme events will be considered, with emphasis on user needs, in order to be included in the MEC/Rfdbk system. This would ensure the provision of a common "state-of-the-art" tool for EPS verification for the entire COSMO community.

Besides the verification of deterministic forecasts, at DWD the possibility to evaluate ensemble forecasts with the MEC/Rfdbk system has already been implemented. The ENS verification includes all the features already known from the deterministic verification. The basic requirements to a verification system such as stratification, conditional verification or significance tests are possible and in use at DWD for the verification of the operational forecasts. Beyond that, the ensemble verification offers some additional scores and features like:

> Ensemble & probabilistic scores such as the rank-histogram, reliability diagram, Brier score with decomposition, ROC curve, economic value, fair CRPS, outlier statistics, spread/skill and deterministic scores based on the ensemble mean.
> Time series of ensemble scores and station-based ensemble scores.

Evaluation of a subset of ensemble members only.
Conditional verification capabilities.

Features that could be implemented during the project phase, provided positive user feedback and developer time constraints:

- Identification of ENS scores depending on the needs of the users, following an *assessment* ***of required resources both for national and CP domains***.
- Implementing the bootstrapping method to infer on the score uncertainty.
- Implementation of bootstrap confidence intervals to these methods to generalize score estimates over different sample sizes.

A schematic of the additional capabilities of the MEC/Rfdbk system for EPS forecasts that will become available through the suggested project is given in Figure 1.

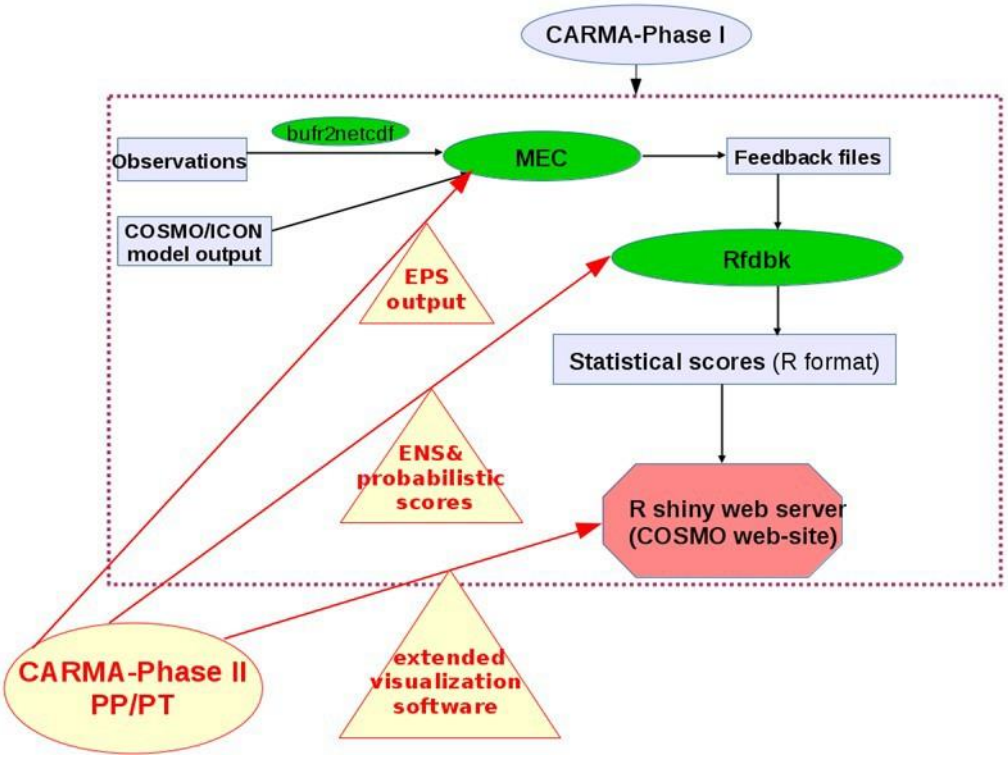

Figure 1. Design of the MEC-Rfdbk system with extension for EPS shown by the red triangles and circles.

Proposed Actions

Within PP-CARMA the technical requirements to run also an ensemble verification are already given at each center of the consortium. This includes the MEC software to produce feedback files from observation and model forecasts, the Rfdbk based scripts to run a verification on top

and the necessary visualization software to interactively browse the verification results. Additional effort from the PP/PT is needed to guide the centers in the adaptation and use of the software for ensemble systems. In summary, the following should be provided:

- Documentation for MEC usage in an ensemble framework;
- Documentation of Rfdbk usage in an ensemble framework;
- Sample seasonal dataset for users with no EPS products and for testing purposes;
- Scripts to produce verification scores for EPS forecasts available to project participants through common WG5 repository;
- Installation of the visualization software for EPS scores at user sites and on the COSMO server;
- Exchange of scores from EPS operational systems over national domains. Explore the possibility of scores calculation over common areas (assessment of required resources).

Links to other projects or activities

- Support Activities - WG5 Common Plot activity for the calculation and representation of verification results of statistical indices for main weather parameters derived using the operational ICON-LAM and/or COSMO model implementations in each service with MEC/Rfdbk system.
- Support Activities – NWP Test Suite: the MEC-Rfdbk system is currently installed and used on the ECMWF platform for the evaluation of new model versions before they are officially released.
- WG7 activities linked to development and evaluation of ENS systems.
- WG5 activities for verification and case studies.

Future actions

- Observation representativeness for different model scales. Not considering the representativeness can lead to misinterpretation of scores like the spread-skill ratio or the CRPS. Quantifying the representativeness could help in model development by offering the possibility to give corrected scores. This is especially important in the ensemble context. Score correction for observation representativeness for the model scales.
- Extend WG6 SPRT Common Plot activity to the collection on a seasonal basis of ENS based FF forecasts.

- Supporting centers to set up their version of an EPS, with or without DA (probably more the job of C2I), where possible.

Description of individual tasks

Task 1. Administrative Tasks and Technical Support

Goal Due to the distributed nature of the project participation team, administrative activities will be included in this task, in order to maintain a good collaboration and information flow between all participants (regular web conferences, workshops, etc.). A mailing list will be used in order to support communication and information exchange between project participants.
Furthermore, continuous technical support will be provided by the PST (Project Support Team) for the update of the MEC-Rfdbk system for EPS output by all the member countries in the consortium and any other technical difficulties that will be revealed.

Sub-Tasks

***Task 1.1* Administrative Tasks**
This task is included to ensure a good collaboration and information flow between all participants (regular web conferences, workshops, etc.)

***Task 1.2* Technical Support**
Update of the MEC-Rfdbk system for EPS output by all the member countries in the consortium will be achieved with support from PST (Project Support Team).

Deliverables:

- Project coordination, meetings, preparation of plans/reports, workshops and regular web conference organization.
- Support in the implementation of the system for all partners.

Contributors:

Task 1.1: Stefan Ioan Gabrian (NMA) – 0. 2 FTE

Start 09.2022 – End 09.2024

Task 1.2: Stefan Ioan Gabrian (NMA) – 0. 25 FTE
Felix Fundel (DWD) – 0. 2 FTE

Start 09.2022 – End 09.2024

Total Resources Task 1: 0.65 FTEs

Task 2. MEC and Rfdbk system adaptations for EPS systems

Goal Preparation of documentation on the installation, adaptation and use of the MEC-Rfdbk system (including CP activities specifications if necessary), assessment of required resources for inclusion of EPS scores into CP activities. Formalize the content of Common Plot/Common Area EPS evaluation guidelines. Templates to run MEC and Rfdbk for ENS and probabilistic products (WG5 repository) over national domains and probabilistic products over Common Area (if such is applicable). A seasonal dataset from COSMO-LEPS will serve as a test bed for users with no EPS products. Seasonal observations necessary for this production are already becoming available through the CP activities.

Sub-Tasks

Task 2.1 Instructions for use and adaptation of MEC software for Feedback Files based on EPS model output (LAM based EPS, IFS ENS) - Problems solving.

Preparation of Instructions on necessary adaptation of MEC software for FFs based on EPS model output. Test of the guidelines/software from the PST on a dataset for a period of a month worth of EPS forecasts, to be defined by the PL for FF production. Special attention for the production of FF based on IFS ENS data. In the case more variables are requested by users, instructions will be provided on how to include them.

Task 2.2 Instructions for adaptation of Rfdbk verification system for production of ENS and probabilistic scores (LAM based EPS, IFS ENS) - Problems solving

Preparation of Instructions on necessary adaptation of Rfdbk libraries for EPS model output evaluation The instructions will be followed by a short description on the definition of scores implemented. Test of the guidelines from the PST on a dataset for a period of a month worth of FF for both LAM and IFS ENS, as produced from **Task 2.1**.

Task 2.3 EPS evaluation guidelines as part of Common Plot activity

Define the content of Common Plot EPS evaluation guidelines. Analyze the possibility of Common Area verification in addition to national EPS domains. Recommendation of a cloud storage possibility for use in the exchange of forecast/observation data for verification purposes.

Task 2.4 Scripts to produce verification scores for EPS forecasts available to project participants through the common WG5 repository.

Adaptation of scripts according to guidelines, to produce the necessary statistical information for the use of the MEC/Rfdbk system for CP ENS requirements over national domains. In the case of more variables requested by users, scripts will be modified accordingly to include them. Additionally, preparation and testing of scripts for semi-automatic use of the Rfdbk for Common Area statistical evaluation in the possibility that a Common Area is applicable.

Deliverables:

- Guidelines for the use/adaptation of MEC or Feedback Files based on EPS model output

- Guidelines for the use/adaptation of Rfdbk for production of ENS and probabilistic scores
- WG5 Common Plot/Area Guidelines for COSMO EPS verification.
- Templates to run MEC and Rfdbk for ENS and probabilistic products (WG5 repository) over national domains, available through common repository.
- Scripts to run Rfdbk for ENS and probabilistic products over the Common Area (if such is applicable).
- Seasonal Sample dataset for users with no EPS products and testing.

Contributors:

Task 2.1: Stefan Ioan Gabrian (NMA) – 0.3 FTE
Start 10.2023 – End 02.2024
Task 2.2: Felix Fundel (DWD) – 0.2 FTE
Stefan Ioan Gabrian (NMA) – 0.05 FTE
Flora Gofa (HNMS) – 0.05 FTE
Start 10.2023 – End 02.2024
Task 2.3: Flora Gofa (HNMS) **–** 0.2 FTE
Start 10.2023 – End 02.2024
Task 2.4: Felix Fundel, DWD – 0.1 FTE
Stefan Ioan Gabrian, NMA – 0.1 FTE
Start 10.2023 – End 02.2024

Total Resources Task 2: 1.0 FTEs

Task 3. Semi-automatic use of the Rfdbk for ENS production and probabilistic scores

Goal Implementation and testing of the production of ENS and probabilistic scores for national domains (and Common Area if applicable according to **Task 2.3**) through practical use.

Sub-Tasks

Task 3.1 Installation and adaptation of MEC-Rfdbk system for EPS over national domains by all participants.
Update of MEC/Rfdbk system by all services so that EPS model outputs will be added in the evaluation process with the use of **Task 2** guidelines and updated software (scripts), with support from the PST (**Task 1.2**).

Task 3.2: Preparation of Seasonal FF and test with Rfdbk
Test of system features on Seasonal EPS forecasts from participating institutes over national domains will be performed following **Task 2.3** guidelines over national domains. COSMO-LEPS

will also be added in the analysis and can serve as dataset for centres with no national ensemble systems.
Testing of system capabilities and features based on EPS data will be performed for at least two seasons. The testing period dataset will be based either on own operational model data (for participants with an ENS implemented) or on **Task 2** deliverables (where operational data not available). This task will include development/adaptation of automatic procedures (scripts adaptation) to local DBs by all involved services. Scripts to run the MEC-Rfdbk system for ENS capabilities are provided from **Tasks 2.1, 2.2 and 2.4**. Rdata statistical indices files will be created and disseminated for visualisation on COSMO Shiny server.

Task 3.3 Adaptation of Shiny visualisation libraries for EPS properties and scores.
The task work will include the installation and update of visualisation software on national servers and COSMO web. R Shiny visualisation libraries for EPS properties and scores will be adapted for usage on the web interface (COSMO server to host CP activity and on individual shiny servers from each centre). The shiny server on the COSMO web-site will be used for visualisation of Task 3.2 outcome and for future use within WG5 CP activity for EPS on a dedicated page (different from the current for deterministic models), while the set-up of individual shiny servers for each interested centre is employed to facilitate the visualisation of individual verification results.

Task 3.4 Test of system features on Seasonal EPS forecasts over Common Area (if applicable).
Test of system features on Seasonal EPS forecasts from participating institutes as part of **Common Area** activity, where possible (including COSMO-LEPS) will be performed, depending on the ***EPS evaluation guidelines as part of Common Plot activity*** delivered from ***Task 2.3.*** If Common Area evaluation is applicable, the FF files produced for Task 3.2 will be disseminated and analysed centrally with Rfdbk scripts following **Task 2.3** guidelines, by the PL team. At least two seasons will be included. COSMO-LEPS will be added in the analysis. Visualisation of statistical results on COSMO shiny server.

Deliverables:

- MEC-Rfdbk system installed and updated in all services (for use with ENS products).
- Updated visualisation software on national servers and COSMO web
- Web interface to host ENS results on COSMO web pages
- Production of FF and Rdata statistical files for seasonal test periods by each service (where applicable).
- Visualization through COSMO Shiny web server or national web servers.

- Production of FF (per country) and Rdata statistical files (centrally) for seasonal test periods for CP (if applicable, depending on results from Task 2.3). Visualisation through COSMO Shiny web server.

Contributors:

Task 3.1: Felix Fundel (DWD) – 0.05 FTE
Stefan Ioan Gabrian (NMA), COSMO-LEPS – 0.05 FTE
J. Linkowska, A. Mazur (IMGW-PIB) – 0.05 FTE
F.Gofa, D. Boucouvala (HNMS), COSMO-LEPS – 0.05 FTE
T. Gastaldo, Arpae-SIMC – 0.05 FTE
A. Pauling, MCH – 0.05 FTE
A. Shtivelman, P. Khain IMS– 0.05 FTE
F. Sudati, CNMCA – 0.05 FTE

Start 07.2023 – End 02.2024

Task 3.2: Felix Fundel (DWD) – 0.14 FTE
Stefan Ioan Gabrian (NMA) – 0.14 FTE
J. Linkowska, A. Mazur (IMGW-PIB) – 0.14 FTE
F.Gofa, D. Boucouvala (HNMS) – 0.14 FTE
T. Gastaldo, Arpae-SIMC – 0.14 FTE
A. Pauling, MCH – 0.14 FTE
A. Shtivelman, P. Khain, IMS – 0.14 FTE
F. Sudati, CNMCA – 0.14 FTE

Start 09.2023 – End 09.2024

Task 3.3: Felix Fundel (DWD) – 0.05 FTE
Stefan Ioan Gabrian (NMA) – 0.01 FTE
Theodore Andreadis (HNMS) – 0.02 FTE
J. Linkowska, A. Mazur (IMGW-PIB) – 0.01 FTE
F.Gofa, D. Boucouvala (HNMS) – 0.01 FTE
T. Gastaldo (Arpae-SIMC) – 0.01 FTE
A. Pauling, MCH – 0.01 FTE
A. Shtivelman, P. Khain, IMS – 0.01 FTE
F. Sudati, CNMCA – 0.01 FTE

Start 10.2023 – End 02.2024

Task 3.4: Stefan Ioan Gabrian (NMA) – 0.14 FTE

Start 05.2024 – End 09.2024

Total Resources Task 3:~1.6 FTEs

Task	Contributing scientist(s)	FTE/ years	Start	Deliverables	Date of delivery	Preceding tasks
1.1	S.I Gabrian (NMA)	0.2	09.2022	- Project coordination, meetings, preparation of plans/reports, workshops and regular web conference organization. - Support in the implementation of the system for all partners.	09.2024	N/A
1.2	S.I Gabrian (NMA) F. Fundel (DWD)	0.45	02.2022		09.2024	N/A
2.1	S.I Gabrian (NMA)	0.3	10.2023	- Guidelines for the use/adaptation of MEC or Feedback Files based on EPS model output. - Guidelines for the use/adaptation of Rfdbk for production of ENS and probabilistic scores. - WG5 Common Plot/Area Guidelines for COSMO EPS verification. - Templates to run MEC and Rfdbk for ENS and probabilistic products (WG5 repository) over national domains, available through common repository. - Scripts to run Rfdbk for ENS and probabilistic products over the Common Area (if such is applicable). - Seasonal Sample dataset for users with no EPS products and testing.	02.2024	N/A
2.2	F. Fundel (DWD) S.I Gabrian (NMA) F. Gofa (HNMS)	0.3	10.2023		02.2024	N/A
2.3	F. Gofa (HNMS)	0.2	10.2023		02.2024	Task 2.1, 2.2
2.4	F. Fundel (DWD) S.I Gabrian (NMA)		10.2023		02.2024	Task 2.1, Task 2.2 Task 2.3

3.1	F. Fundel (DWD) S.I Gabrian (NMA) J. Linkowska, A. Mazur (IMGW-PIB) F.Gofa, D. Boucouvala (HNMS) T. Gastaldo (Arpae-SIMC) A. Pauling (MCH) A. Shtivelman, P. Khain (IMS) F. Sudati (CNMCA)	0.35	07.2023	- MEC-Rfdbk system installed and updated in all services (for use with ENS products). - Updated visualisation software on national servers and COSMO web - Web interface to host ENS results on COSMO web pages - Production of FF and Rdata statistical files for seasonal test periods by each service (where applicable). - Visualisation through COSMO Shiny web server or national web servers. - Production of FF (per country) and Rdata statistical files (centrally) for seasonal test periods for CP (if applicable, depending on results from Task 2.3). - Visualisation through COSMO Shiny web server.	02.2024	Task 1 Task 2
3.2	F. Fundel (DWD) S.I Gabrian (NMA) J. Linkowska, A. Mazur (IMGW-PIB) F.Gofa, D. Boucouvala (HNMS) T. Gastaldo (Arpae-SIMC) A. Pauling (MCH) A. Shtivelman, P. Khain (IMS) F. Sudati (CNMCA)	1.12	09.2023		09.2024	Task 1 Task 2 Task 3.3
3.3	F. Fundel (DWD) S.I Gabrian (NMA) T. Andreadis (HNMS) J. Linkowska, A. Mazur (IMGW-PIB) F.Gofa, D. Boucouvala (HNMS) T. Gastaldo (Arpae-SIMC) A. Pauling (MCH) A. Shtivelman, P. Khain (IMS) F. Sudati (CNMCA)	0.13	10.2023		02.2024	Task 1 Task 2 Task 3.1
3.4	S.I Gabrian (NMA)	0.14	05.2024		09.2024	Task 1 Task 2 Task 3.1, 3.2 Task 3.3
All		**~3.25**	**09.2022**		**09.2024**	

Estimated resources (in FTE-years) needed in total: ~3.27 FTEs

References

W. Chang, J. Cheng, J.J. Allaire, Y. Xie, J. McPherson (2020): shiny: Web Application Framework for R. R package version 1.5.0. Available online at: https://CRAN.R-project.org/package=shiny

F. Gofa (2021): Common Verification Plots Guidelines for 2021-22 (www.cosmo-model.org/content/tasks/verification.priv/common/reports/CP-2021-2022.pdf)

F. Gofa (2016): COSMO Verification Overview, 38th EWGLAM and 23th SRNWP Meeting, Rome, 03-06 October 2016

F. Gofa, U. Pflüger, X. Lapillonne, A. Vocino, D. Boucouvala, J. Linkowska, R. Dumitrache, A. Bundel, M.S. Tesini, E. Oberto, Y. Levi (2015): Recommendations: Strategy on Verification Tools

Feedback-File definition (2012): Supplementary Documentation, DWD, http://cosmo-model.org/content/model/documentation/core/cosmoFeedbackFileDefinition.pdf

F. Fundel (2021): Feedback File Verification Suite at DWD, DWD, available online at http://www.cosmo-model.org/shiny/users/fdbk/RfdbkVeriDoku.html

A. Iriza-Burcă, F. Gofa, D. Boucouvala, T. Andreadis, J. Linkowska, P. Khain, A. Shtivelman, F. Batignani, A. Pauling, A. Kirsanov, T. Gastaldo, B. Maco, M. Bogdan, F. Fundel (2022): The

COSMO Priority Project CARMA: Common Area with Rfdbk/MEC Application Final Report. COSMO Technical Reports No. 46, 2022, DOI: 10.5676/DWD_pub/nwv/cosmo-tr_46

A. Iriza-Burcă, F. Gofa, D. Boucouvala, T. Andreadis, J. Linkowska, P. Khain, A. Shtivelman, F. Batignani, A. Pauling, A. Kirsanov, T. Gastaldo, B. Maco, M. Bogdan, F. Fundel (2022): Summary of Results from the COSMO Priority Project CARMA: Common Area with Rfdbk/MEC Application. COSMO Newsletter, No.21 (p.21-27), DOI: doi:10.5676/dwd pub/nwv/cosmo-nl 21 04

A. Iriza-Burcă, J. Linkowska, F. Fundel (2020): Common Area Verification Activity with Rfdbk/MEC Application: MAM 2020 First Results. COSMO Newsletter, No.20 (p.18-30), DOI: doi:10.5676/dwd_pub/nwv/cosmo-nl_20_04

P. Patruno, D. Cesari (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. Available on-line at: http://cosmo-model.org/content/consortium/generalMeetings/general2011/wg1.htm.

R. Potthast (2019): Model Equivalent Calculator (MEC) overview, 21st COSMO General Meeting, 9-13 September 2019, Rome (Italy). Available on-line at: http://cosmo-model.org/content/consortium/generalMeetings/general2019/default.htm