ABSTRACTS

EUMETNET 2019-2023

SRNWP-EPS workshop 2021

Workshop on "*LAM-EPS verification on extremes and forecasting value chain for High Impact Weather*"

26-28 October 2021, BlueJeans video-conference meeting

Session 2: EFI and SOT

https://bluejeans.com/736316721/2982

The Extreme Forecast Index and Shift Of Tails at ECMWF (invited)

Ivan Tsonevsky, ECMWF

The Extreme Forecast Index (EFI) and its complementary Shift Of Tails index (SOT) have been designed to provide early warning signals of extreme weather by measuring the difference between the real-time forecast and the model climate used as a reference. The model climate has a crucial role for the quality of the EFI. ECMWF uses reforecasts, from the same model version that provides the forecasts, to ensure a high level of compatibility between the model climate and the real-time forecast which is essential for a good quality EFI. Some examples of incompatibility and other issues will be shown. The procedure of creating the model climate for the EFI at ECMWF will be provided. If one wants to generate an EFI from LAM ensemble output, and producing re-forecasts from LAMs is not viable, then there are potentially other approaches. For any such approach to be viable it needs to retain some similarity of spatial scale between the LAM forecast resolution and the resolution of the reference climate. To this end one could try to use "ecPoint" post-processed (downscaled) output from ECMWF, from either ECMWF re-forecasts or from ERA5, at least for some variables, notably rainfall although this approach needs to be tested. Some important considerations for the EFI computation will be highlighted including the numerical procedure, ensemble size and the meteorological parameter of interest. ECMWF performs verification of the EFI for some parameters such as temperature, total precipitation and mean wind using the area under the ROC curve as a skill measure. There is ongoing work on upgrading the operational EFI verification at ECMWF. Some results and upcoming changes of the EFI verification will be presented. Recently the EFI and SOT have been used to highlight the risk of severe convection. Some practical considerations and verification results for these convective EFIs will be given.

Defining an Extreme Forecast Index (EFI) and Shift Of Tails Index (SOT) for EUMETNET LAM-EPS (Short poster presentation)

Joan Montolio Llenas, AEMET - EUMETNET SRNWP-EPS, *Spain*

The aim of the Application Task “EPS_8” (hereon referred to as “AT”) is the design of a software that performs the calculation of the EFI and SOT indexes within the domain of a Limited Area Model Ensemble Prediction System (LAM-EPS). Given that these indexes have already been designed and implemented in relevant meteorological offices such as the European Center for Medium-Range Weather Forecasts (ECMWF) and Méteo-France, the starting point for the AT consisted in gathering information and scientific articles about the formal definition of both indexes and how they were applied to different EPS. At that point, the article written by Laure Raynaud et al. “Detection of Severe Weather Events in a High-Resolution Ensemble Prediction System Using the Extreme Forecast Index (EFI) and Shift of Tails (SOT)” proved to be a nice reference for the development of the AT as it presents the methodology followed to construct the EPS climatology -which is the key point in the implementation of the EFI and SOT indexes- from Méteo-France LAM “AROME-EPS” historical archives. Having acquired the main theoretical aspects about EFI and SOT indexes, it was time to translate this knowledge into computational language: Python is, at the moment, the programming language used to perform the implementation of EFI and SOT. As a first experiment, a set of functions have been designed to compute the calculation of EFI and SOT indexes for a single grid point from the Spanish Meteorological Agency (AEMET) LAM “γSREPS” domain. The objective of this exercise is to set the basis of the functions that perform the different operations to obtain EFI and SOT values. Special attention is being given in the robustness and synthesis of the programming code in order to optimize the computational execution time of the programs and make the Python code appear user-friendly.

Session 3: SRNWP-EPS LAM-EPS database and Research Plan

https://bluejeans.com/553546542/0555

An Update on Multi-Model Ensemble Evaluation over the British Isles

David Flack and Aurore Porson, *Met Office, UK*

There is an ongoing debate about the best design for convective-scale ensembles to improve a multi-centre problem of lack of spread. Using the SRNWP-EPS convective-scale ensemble database a multi-model ensemble is created over the British Isles. Work is underway with two different approaches to test the "usefulness" of the multi-model ensemble. One aspect takes a more "traditional" approach to the evaluation of ensembles at the convective-scale using neighbourhood techniques to look at the spatial spread. The second considers a more process- and object- based examination of the properties of the ensembles, rather than a comparison with observations. The latter is the focus of the work presented here. It is shown that the different models have statistically different underlying distributions and are not equally contributing to the ensemble, although this is strongly case and threshold dependent. Therefore, it is plausible that, traditional ensemble scores may require a different interpretation for a multi-model ensemble due to the "equally-likely" assumption no longer holding. It is suggested that greater "benefit" of multi-model ensembles may be obtained by considering process- and object-based techniques. These techniques indicate that the individual models can be clearly identified within the multi-model ensemble based on factors such as the number, structure and intensity of convective events. Thus, showing an advantage of using a multi-model ensemble for operational meteorologists. However, further consideration of the impact that non-equally likely members have on the interpretation of our current scores is required to fully assess the "usefulness" of a multi-model ensemble. Adapting our current scores to take into account the non-equally likely members may also be possible but will need greater consideration.

Session 4: EPS applications

https://bluejeans.com/576254957/0679

New convective guidance at the Australian Bureau of Meteorology (invited)

Harald Richter, Rob Warren (+Monash University), Dean Sgarbossa, Ivor Blockley, and Maree Carroll, Australian Bureau of Meteorology, *Australia*

Over the past two years the operational guidance for lightning and severe convective hazards at the Bureau has either undergone a major uplift or has been extended through the introduction of new diagnostics. The new suite of the guidance is now fully based on either the global or the convection-allowing ensembles of the Australian Community Climate and Earth System Simulator (ACCESS).

The lightning guidance (Calibrated Thunder) runs off a 36-member ~33 km lag ensemble of the global ACCESS model and has been extended to predict total lightning instead of cloud-to-ground lightning delivered through the previous system. A more sophisticated calibration methodology is under development to further increase the skill of the system. Based on the same global ensemble is a large suite of convective parameters useful for the prediction of severe convection and its associated storm modes or convective hazards.

While the previously described guidance estimates the probability of lightning or hazardous convective weather based on conductive convective environments, storm attribute guidance based on the 12-member ~2.2 km convection-allowing ACCESS ensemble exploits specific properties of the storms produced by the model.

This overview will present the broad features of these new total lightning and severe convective guidance systems and provide examples of their application, their forecast quality and their strengths and weaknesses.

Session 5: Value Chain

https://bluejeans.com/576254957/0679

Using Value Chain Approaches to Evaluate End-to-End Warning Systems (**invited**)

Beth Ebert, Science and Innovation Group, Bureau of Meteorology, *Australia*

David Hoffman, World Weather Research Programme

Carla Mooney, Community Services Group, Bureau of Meteorology, *Australia*

The weather information value chain provides a framework for characterising the production, communication, and use of information by all stakeholders in an end-to-end warning system. Since the generation of weather warning and climate services has become more complex, both technically and organizationally, the value chain concept has become a popular tool for describing and assessing the production, use and benefits of such services.
The end-to-end warning system for high impact weather brings together hazard monitoring, modelling and forecasting, risk assessment, communication and preparedness activities and systems and processes which enable timely action to reduce risks. Weather and associated warning services are typically developed and provided through a multitude of complex and malleable value chains (networks), often established through co-design, co-creation and co-provision.

A new international project under the WMO World Weather Research Programme is using value chain approaches to describe and evaluate the end-to-end warning system for high impact weather. Its aims are

(1) To review value chain approaches used to describe weather, warning and climate services to assess and provide guidance on how they can be best applied in a high impact weather warning context that involves multiple users and partnerships;

(2) To create a searchable end-to-end Warning Chain Database for scientists and practitioners involved in researching, designing and evaluating weather-related warning systems to review previous experience of high impact weather events and assess their efficacy using value chain approaches.

We encourage the research and operational community, including the SRNWP-EPS community, to participate in this project by contributing case studies of high impact events and collaborating in their analysis. Integration of the physical and social sciences in this project will lead to new insights that we hope will ultimately improve the effectiveness of warnings for high impact weather.

Value Chain in Warm Convection over Barcelona (value chain case)

Sergi González Herrero, DT Catalonia, AEMET, Barcelona, *Spain*

During the early morning of July 16, 2020, heavy rains in Barcelona of more than 70 mm in 3 hours occurred. However, no warnings were issued during this event. Although environmental conditions were favorable to heavy precipitation, with high precipitable water values and a low-level jet, deterministic NWP predicted modest precipitation values of less than 5 mm. Only one member of the AEMET γSREPS model simulated a precipitation amount greater than 40 mm in 3 hours. Therefore, the observations fell into the tail of the pdf of the forecasts. The analysis of the value chain of the event shows that the local forecaster did not issue a warning due to the low precipitation outputs in most NWPs. However, he suggested there was a low probability of heavy rains to civil protection and the forecaster in charge of the watching. Despite advice, the watching forecaster in Madrid during the event did not realize that heavy rains were occurring in Barcelona due to the low reflectivity of the radar output during the event caused by the microphysics of warm convection. This case shows that a combination of unfortunate circumstances, such as rainfall in the tail of the numerical forecast distribution and problems in the observing systems, can lead to a failure in the value chain of heavy rainfall events.

Session 6: LAM-EPS, Post-processing and applications

https://bluejeans.com/275654685/7937

Applying LSTM to High Resolution Numerical Weather Products for Convection Prediction (invited)

Aniel Jardines, University Carlos III in Madrid, *Spain*

The aim of this research is to explore how machine learning techniques can improve the use of numerical weather products to forecast convective events. In this research, we use forecast data from a high resolution ensemble prediction system (EPS), AEMET-γSREPS, produced by AEMET and which covers roughly the Iberian peninsula. This EPS forecast is coupled with storm observation data from satellite and lighting detection to train a Long Short Term Memory (LSTM) machine learning model to predict convection characteristics relating to cloud top altitude and severity, and convective elements, lightning and overshooting tops, up to 36 hours in advance.

Design of an object-based probabilistic convection product for air traffic management using AROME-EPS

Matteo Ponzano, Laure Raynaud and Lucie Rottner, CNRM, Météo-France/CNRS, *Toulouse, France*

The accurate prediction of the atmospheric conditions plays a key role in the air traffic management. Among the variables predicted by numerical prediction models, convection forecasts are particularly useful for the pre-tactical (one day before) planning support. A probabilistic convection product has been designed in the framework of SESAR 2020 ISOBAR project. It is based on a post-processing of forecasts issued by the convection permitting LAM-EPS AROME. The maximum reflectivity forecast along grid-points vertical section and the ICA (Index of Convection AROME) are used as proxy model variables of convection. These variables are processed following a fuzzy object-oriented approach in order to cope with small-scale unpredictable details of mesoscale structures. This procedure allows to extract coherent features, which are associated with a higher predictability than the direct model output. Object-based method is applied through an adaptation of the similarity-based method described in Raynaud et al. (2019) and Rottner et al. (2019) to take into account reflectivity field. In the final stage, convection objects computed for each member of AROME-EPS are merged to draw two probability fields for total and moderate-intensity of convection over the Western Europe. The product has been calibrated using a one-month period (July 2019). Forecasts are compared to satellite observations collected into the RDT (Rapidly Developing Thunderstorm) product. The validation test was performed for a one-month period during the summer 2021. It has highlighted beneficial information provided by post-processed forecasts, with area under the ROC curve values ranging between 0.75 and 0.80. A visual inspection of some case studies showed some promising results in the terms of similarity between probability forecasts and RDT observations.

Analysis of multiple ensemble prediction systems for Mediterranean extreme weather forecasting

Alejandro Hermoso, Víctor Homar, *University of Balearic Islands, Spain*

The Mediterranean basin is frequently affected by heavy precipitation episodes, causing substantial human and material losses. Despite recent progresses in the modelling and prediction of such high-impact events, precisely predicting socially relevant aspects of convective phenomena, such as intensity, location, and timing remains challenging and unsolved. The deficiencies of current numerical weather prediction systems arise from inaccuracies in the estimation of the initial atmospheric state across all the scales of interest and from the lack of full understanding of the small-scale highly nonlinear processes involved in the genesis and evolution of deep moist convection. In this context, ensemble prediction systems (EPSs) represent a practical approach to sample the inherent uncertainties of numerical weather forecasts.

In this work, we investigate the properties of perturbations to initial conditions and/or model parameterizations of subgrid processes in a convection-permitting setting for a set of three heavy precipitation episodes that have impacted various areas of western Mediterranean in recent years. In particular, the analysed strategies comprise dynamical downscaling from a coarser resolution global model and tailored bred vectors, to sample initial condition uncertainties. Furthermore, the performance of stochastic methods to perturb some key subgrid parameterization is assessed. The characteristics of individual perturbations (i.e., initial conditions or model parameterizations) and combined perturbations, considering both error sources, are analysed in terms of ensemble diversity and skill. This study contributes to the identification of the most relevant error sources for the prediction of Mediterranean heavy precipitation events, providing useful guidance to design efficient and skilful convection-permitting EPSs.

Ensemble forecasting of the SOFOG 3D experiment

Anne McCabe, Jenna Thornton, Jeremy Price, Paul Field, Adrian Lock, Steve Derbyshire, Adrian Hill, *Met Office, UK*

Over a 6 month period during the winter of 2019 – 2020, Météo-France led an extensive observation campaign in the South West of France known as the South west FOGs 3D (SOFOG 3D) experiment. The aim of the campaign was to measure the boundary layer during fog events in order to develop further our understanding of fog processes and to improve the forecast of fog by Numerical Weather

Prediction (NWP) models. Using data from the UK Met Office (UKMO) observational site in the SOFOG 3D experiment, we explore the potential benefit and challenges of using ensemble to forecast fog. Results are shown comparing the ensemble output with observations at three different horizontal length scales: 2.2km, 300m and 100m. The first matches the UKMO's operational convective scale ensemble, MOGREPS-UK, while the higher resolutions are of interest for producing detailed output over specific sites, such as airports. We explore the sensitivity of the ensemble forecast to changes in the model physics through use of a perturbed parameter scheme and different cloud schemes. We consider the difficulties operational meteorologists face in picking out storylines from ensemble forecasts when there is a lot of spread, and discuss the potential benefits of a scenario-based approach.

Session 7: Calibration on Extremes

https://bluejeans.com/729366412/2658

AI methods for post-processing ensemble predictions (invited)

Sebastian Lerch, Karlsruhe Institute of Technology, *Germany*

The weatherPost-processing ensemble weather predictions to correct systematic errors has become an integral part of weather forecasting and standard practice in research and operations. Most commonly used post-processing methods are distributional regression approaches from statistics, however, the use of modern machine learning methods for post-processing has been the focus of much recent research interest. By enabling the incorporation of additional predictor variables beyond ensemble forecasts of the target variable, these methods have the potential to overcome inherent limitations of traditional approaches. In this talk, I will give an overview of our research on developing neural network-based methods that enable nonlinear relationships between arbitrary predictor variables and forecast distribution parameters to be learned in a data-driven way. I will specifically focus on recent work on wind gust forecasting, where we conducted a systematic comparison of state of the art methods from statistics and machine learning based on six years of data from a high-resolution, convection-permitting ensemble prediction system that was run operationally at the German weather service, and hourly observations at 175 surface weather stations in Germany. The case study demonstrates that incorporating information from additional meteorological predictor variables beyond wind gusts leads to significant improvements in forecast skill and enables the AI methods to learn physically consistent relations. With a *focus on high-impact extreme events*, I will further discuss directions for future work, in particular on incorporating physical information and expertise into post-processing models.

Extreme events and ensemble post-processing: from forecast to evaluation (invited)

Maxime Taillardat, Météo-France, *France*

Forecast and verification of ensemble prediction systems for extreme events remain a difficult question in the numerical weather prediction (NWP) community.

The general public as well as the media pay particular attention on extreme events and conclude about the global predictive performance of NWP models, which can be unskillful when they are needed.

Statistical post-processing of ensemble forecasts, from simple linear regressions to more sophisticated techniques, is now a well-known procedure in order to correct biased and poorly dispersed ensemble weather predictions.

However, the ability of these machine learning algorithms to extrapolate beyond the range of the trained data is challenging.

In this talk, some post-processing methods for the ensemble forecast of extremes used by Météo-France are presented, and some ideas for the future are discussed.

Thus, the paradigm of maximizing the sharpness subject to calibration should be associated with the paradigm of maximizing the value for extreme events subject to a good overall performance.

A comparison on ensemble post-processing methods using the SRNWP-EPS database

Maria Cortès Simó and Alfons Callado-Pallarès, *AEMet, Spain*

The main aim of the SRNWP-EPS 2019-2023 module/project is to contribute to the cooperation building reliable and valuable high-resolution limited-area ensemble systems in Europe which resolves the convection-permitting scale phenomena to improve the high impact weather (HIW) predictions. With this aim, a database on common convection-permitting LAM-EPS has been created. The database collects the most relevant surface parameters in GRIB format from all the convection-permitting ensembles in a unique repository. The purpose is to facilitate the comparison of the ensembles and their behavior to sustain and foster coordinated research.

To benefit from the full potential of an ensemble prediction system, statistical post-processing is a necessary step. In the past decades, several calibration methods have been developed for a wide range of weather quantities, however few studies have focused on ensemble post-processing of extreme events (e.g. wind gusts, precipitation, maximum and minimum temperature).

Here, we provide a comparison of different statistical and machine learning methods applied to calibrate daily precipitation and maximum wind gust forecasts. Common stations from two EPS of the SRNWP-EPS database (γSREPS and MFArome-EPS) are used to perform the comparison between the post-processing approaches. The performance of ensemble model output statistics (EMOS) and Bayesian Model Averaging (BMA) considering different distributions (including GEV distributions) are compared with Quantile Regression Forests. Classical scoring rules together with scores centered on high values are used to verify predictive forecasts.