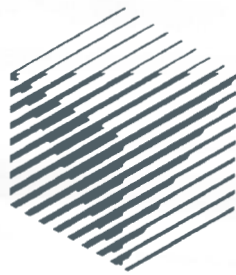


**ERCIM**



**NEWS**

*Special theme:*

# The Climate Action

*Mathematics,  
Informatics  
and Socio-Economics  
Accelerating the Sustainability*



## Editorial Information

ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 6,000 printed copies and is also available online.

ERCIM News is published by ERCIM EEIG  
BP 93, F-06902 Sophia Antipolis Cedex, France  
+33 4 9238 5010, [contact@ercim.eu](mailto:contact@ercim.eu)  
Director: Philipp Hoschka, ISSN 0926-4981

### Contributions

Contributions should be submitted to the local editor of your country

### Copyright notice

All authors, as identified in each article, retain copyright of their work. ERCIM News is licensed under a Creative Commons Attribution 4.0 International License (CC-BY).

### Advertising

For current advertising rates and conditions, see  
<https://ercim-news.ercim.eu/> or contact [peter.kunz@ercim.eu](mailto:peter.kunz@ercim.eu)

ERCIM News online edition: [ercim-news.ercim.eu/](http://ercim-news.ercim.eu/)

### Next issue:

July 2020: Solving Engineering Problems with Machine Learning

### Subscription

Subscribe to ERCIM News by sending an email to [en-subscriptions@ercim.eu](mailto:en-subscriptions@ercim.eu)

### Editorial Board:

Central editor:  
Peter Kunz, ERCIM office ([peter.kunz@ercim.eu](mailto:peter.kunz@ercim.eu))

### Local Editors:

Austria: Erwin Schoitsch ([erwin.schoitsch@ait.ac.at](mailto:erwin.schoitsch@ait.ac.at))  
Cyprus: Georgia Kapitsaki ([gkapi@cs.ucy.ac.cy](mailto:gkapi@cs.ucy.ac.cy))  
France: Christine Azevedo Coste ([christine.azevedo@inria.fr](mailto:christine.azevedo@inria.fr))  
Germany: Alexander Nouak ([alexander.nouak@iuk.fraunhofer.de](mailto:alexander.nouak@iuk.fraunhofer.de))  
Greece: Lida Harami ([lida@ics.forth.gr](mailto:lida@ics.forth.gr)),  
Athanasios Kalogeras ([kalogeras@isi.gr](mailto:kalogeras@isi.gr))  
Hungary: Andras Benczur ([benczur@info.ilab.sztaki.hu](mailto:benczur@info.ilab.sztaki.hu))  
Italy: Maurice ter Beek ([maurice.terbeek@isti.cnr.it](mailto:maurice.terbeek@isti.cnr.it))  
Luxembourg: Thomas Tamsier ([thomas.tamsier@list.lu](mailto:thomas.tamsier@list.lu))  
Norway: Monica Divitini ([divitini@ntnu.no](mailto:divitini@ntnu.no)),  
Are Magnus Bruaset ([arem@simula.no](mailto:arem@simula.no))  
Poland: Hung Son Nguyen ([son@mimuw.edu.pl](mailto:son@mimuw.edu.pl))  
Portugal: José Borbinha ([jlb@ist.utl.pt](mailto:jlb@ist.utl.pt))  
Sweden: Maria Rudenschöld ([maria.rudenschold@ri.se](mailto:maria.rudenschold@ri.se))  
Switzerland: Harry Rudin ([hrudin@smile.ch](mailto:hrudin@smile.ch))  
The Netherlands: Annette Kik ([Annette.Kik@cwil.nl](mailto:Annette.Kik@cwil.nl))  
W3C: Marie-Claire Forgue ([mcf@w3.org](mailto:mcf@w3.org))

Cover illustration: Photo by Annie Spratt on Unsplash.

## JOINT ERCIM ACTIONS

### 4 Foreword from the President

by Björn Levin

### 4 ERCIM Membership

## SPECIAL THEME

The special theme “The Climate Action” has been coordinated by Sobah Abbas Petersen (NTNU) and Phoebe Koundouri (Athens University of Economics and Business and Athena RC)

Introduction to the special theme

### 5 The Climate Action: Mathematics, Informatics and Socio-Economics Accelerating the Sustainability

by Sobah Abbas Petersen (NTNU) and Phoebe Koundouri (Athens University of Economics and Business and Athena RC)

### 8 Urban Vitality: Re-Evaluating the Availability of Urban Parks

by Corine Laan (Amsterdam University of Applied Sciences) and Nanda Piersma (Amsterdam University of Applied Sciences and CWI)

### 9 A Living Lab Approach for Sustainable Forest Management

by Christophe Ponsard and Bérengère Nihoul (CETIC)

### 10 Can Smart Charging Balance the Urban Energy Grid?

by Pieter Bons, Robert van den Hoed (Amsterdam University of Applied Sciences) and Nanda Piersma (Amsterdam University of Applied Sciences and CWI)

### 12 Grid-friendly Sustainable Local Energy Communities

by Michael Kaisers (CWI), Matthias Klein and Alexander Klauer (Fraunhofer ITWM)

### 14 Coastal Ocean Modelling: Uncertainty Representation and Forecast

by Arthur Vidard (Inria)

### 15 Tackling the Multiscale Challenge of Climate Modelling

by Daan Crommelin (CWI; Korteweg-de Vries Institute for Mathematics, University of Amsterdam), Wouter Edeling (CWI) and Fredrik Jansson (CWI)

### 17 A State of the Art Technology in Large Scale Underwater Monitoring

by Gaia Pavoni, Massimiliano Corsini and Paolo Cignoni (ISTI-CNR)

**19 Biogas Residues in the Battle for Terrestrial Carbon Sequestration**

by Bella Tsachidou, Christophe Hissler (Luxembourg Institute of Science and Technology) and Philippe Delfosse (University of Luxembourg)

**20 Circular Intelligence: Digital Services to Facilitate Industrial Symbiosis**

by Theodore Dalamagas, Antonis Kokossis and Thanasis Gentimis (SymbioLabs)

**22 Smart Water Distribution Networks to Reduce Water Loss**

by George Tzagkarakis (FORTH-ICS), Maria N. Anastasiadou and Demetrios G. Eliades (KIOS, University of Cyprus)

**23 Analysis of Extreme Climatic Indexes over Greece for the Period 1979 - 2019**

by Athanasios Karavoulias and Athanassios A. Argiriou

**25 Fostering Smart Cities based on an Enterprise Architecture Approach**

by Bokolo Anthony Jnr. (NTNU), Sobah Abbas Petersen (NTNU) and Markus Helfert (Maynooth University)

**26 Close-The-Loop Model: Social Acceptance of Technology for Sustainability**

by Emanuela Tangari, Carmela Occhipinti and Luigi Briguglio (CyberEthics Lab.)

## RESEARCH AND INNOVATION

**28 Cultural and Heritage Site Defect Detection via 3D Saliency Map Extraction**

by Christos Anagnostopoulos, Aris S. Lalos, George Pavlidis, Christos Koulamas, Athanasios Kalogeras (ISI, Athena R.C), Ana Garcia Lopez (University of Granada) and Giacomo Di Benedetto (Enginlife-Engineering)

**30 Various Approaches to Statistical Calibration of Ensemble Weather Forecasts**

by Sándor Baran and Annette Möller (Faculty of Informatics, University of Debrecen and Institute of Mathematics, Technical University of Clausthal)

**31 Enabling the European Business Knowledge Graph for Innovative Data-Driven Products and Services**

by Dumitru Roman and Ahmet Soylu (SINTEF)

**33 ALerT – Learning about Privacy at the Time of Data Sharing Everywhere**

by Patrick Jost (NTNU), Gisela Böhm (University of Bergen), Monica Divitini (NTNU) and Ingvar Tjostheim (Norwegian Computing Center)

**34 Roadmaps for AI Integration in the Rail Sector - RAILS**

by Francesco Flammini (Mälardalen University and Linnaeus University, Sweden), Valeria Vittorini (CINI and University of Naples Federico II, Italy) and Zhiyuan Lin (University of Leeds, UK)

**35 A Formal Methods Demonstrator for Railways**

by Franco Mazzanti and Davide Basile (ISTI-CNR)

**36 Cyber-Ranges as a Mean of Security Culture Establishment**

by George Hatzivasilis (FORTH), Kostas Fysarakis (Sphynx) and Sotiris Ioannidis (FORTH)

**38 Worried about Data Privacy in Big Data? Don't Be!**

by Sabrina Kirrane, and adapted for ERCIM News by Jessica Michel

**39 MOSAICrOWN - Data Protection in the Data Market Scenario**

by Pierangela Samarati, (Università degli Studi di Milano)

## ANNOUNCEMENTS

**41 ERCIM “Alain Bensoussan” Fellowship Programme**

**42 HORIZON 2020 Project Management**

**42 Dagstuhl Seminars and Perspectives Workshops**

**42 DMS Accelerator Opens its Second Call to Promote Data-Centric Startups and SMEs**

**43 Pan-European Privacy-Preserving Proximity Tracing**

**43 Three Awards for CWI Researchers**

**43 Artificial Intelligence by CWI and Amsterdam UMC Proposes the Best Radiation Treatment Plans in Clinical Practice for the First Time**

# Various Approaches to Statistical Calibration of Ensemble Weather Forecasts

by Sándor Baran and Annette Möller (Faculty of Informatics, University of Debrecen and Institute of Mathematics, Technical University of Clausthal)

**Statistical calibration of ensemble weather forecasts is a rapidly developing research area of statistics as well as atmospheric and water sciences. We are developing and implementing multivariate approaches explicitly accounting for dependencies between weather observation locations and/or between weather variables, including temperature, precipitation and pressure.**

Capturing and modelling uncertainty is essential in any forecasting problem, and, in weather or hydrological predictions, can have enormous economic benefits.

The early 1990s saw an important shift in the practise of weather forecasting, from deterministic forecasts obtained using numerical weather prediction (NWP) models in the direction of probabilistic forecasting. The crucial step was the introduction of ensemble prediction systems (EPSs), which came into operational use in 1992, both at the European Centre for Medium-Range Weather Forecasts (ECMWF) and the U.S. National Meteorological Center.

An NWP model essentially consists of a large and complex set of partial differential equations describing the processes that take place in the planet's atmosphere. An EPS provides not only a single forecast but a range of several forecasts, which are usually generated by running the NWP model multiple times, each time based on another set of assumed initial values obtained by random perturbation of an initial guess derived from the actually available information about the state of the atmosphere.

In recent decades, the ensemble method has become widely used globally, its appeal lying in its ability to easily provide statistical summary measures that explicitly reflect the forecast uncertainty. However, the raw outputs of the EPS often exhibit systematic forecast errors (bias) or cannot properly capture the forecast uncertainty (calibration), thus calling for some form of post-processing. Simple approaches to bias correction or calibration have a long history, and in the first years of the twenty-first century several more sophisticated methods appeared, including statistical models providing full predictive probability distributions of the weather variables at hand. This means that one is not only able to provide a

forecast of tomorrow's temperature in Berlin, but also to forecast the probability that tomorrow's temperature in Berlin will be between 20 and 25 oC with say 80% confidence. Starting with the fundamental works of Tilmann Gneiting and Adrian Raftery that introduced Bayesian model averaging and ensemble model output statistics for ensemble calibration [1], statistical post-processing of ensemble forecasts became a hot topic both in statistics and atmospheric sciences, resulting in a multitude of probabilistic models for different weather quantities, new methods and algorithms for training these models on real weather data and novel approaches to forecast verification [2].

The Hungarian-German research project "Statistical post-processing of ensemble forecasts for various weather quantities", jointly financed by the Hungarian National Research, Development and Innovation Office and the Deutsche Forschungsgemeinschaft, aims to develop and test multivariate post-processing methods that model correlations between different weather variables, and/or incorporate correlations in space, e.g., between observation stations. Further goals are the development of user-friendly software packages for statistical calibration of ensemble weather forecasts,

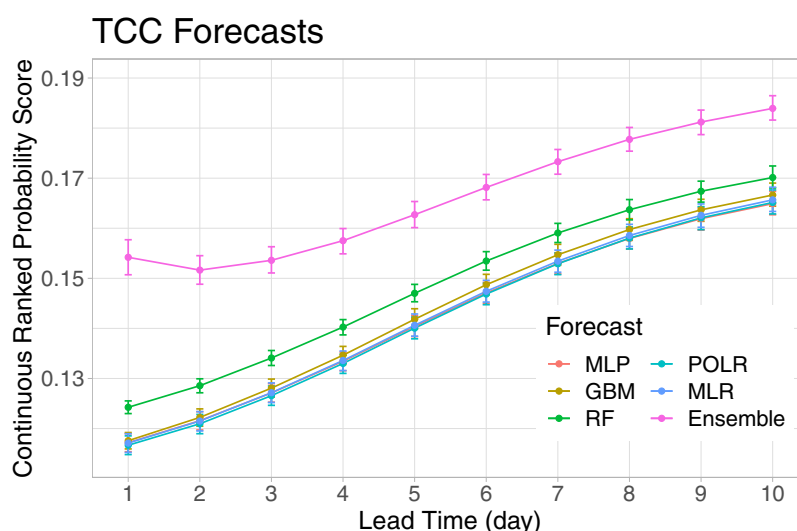


Figure 1: Mean continuous ranked probability score (CRPS) values of the various post-processing approaches and the raw forecasts for different lead times. The CRPS measures the goodness of fit of the predictive distribution to the corresponding observation, the smaller the better.

investigation of approaches that take advantage of local features in the neighbourhood of an observation station, and finally, to create an efficient scientific network.

Most members of the small research group have connections with Tilmann Gneiting's research group at the Heidelberg Institute for Theoretical Studies. The members have a long history of collaboration and include mathematicians and statisticians from Heidelberg University, Karlsruhe Institute of Technology, Technical University of Clausthal, University of Debrecen, University of Hildesheim, and a meteorologist from MeteoSwiss Agency.

Since its inception in April 2018, the project has investigated various angles, including:

- An approach to calibrate hydrological ensemble forecasts, such as water levels of rivers. The challenge when dealing with this type of data is that water level measurements exhibit natural bounds from below and above and are non-Gaussian, so appropriate data transformation schemes are required. The proposed model was applied to predict water levels for the river Rhine at Kaub gauge with great success.
- An approach that accounts for the interdependency between weather variables over time. The temperature today is not independent of the temperature yesterday or other recently observed temperatures. Thus, the forecast errors on different days are not independent from one another. This time dependence was utilised to improve the ensemble forecasts, which worked particularly well for higher lead times when considering temperature forecasts.
- An investigation of the applicability of machine learning approaches to statistical post-processing of ensemble forecasts of total cloud cover (TCC), a variable describing what fraction of the sky is covered by clouds, measured on a nine-point scale: 0, 1/8, 2/8, 3/8, 4/8, 5/8, 6/8, 7/8, 1. Using the ECMWF global TCC ensemble forecasts for the period 2002–2014 and considering different lead times, the predictive performance of multilayer perceptron (MLP) neural networks, gradient boosting machines (GBM) and random forest (RF) algorithms was compared with the forecast skill of the state-of-the-art multiclass and proportional odds logistic regression (MLR and POLR) approaches and the raw ensemble (see Figure 1). To obtain more specific insight, a detailed simulation study was conducted to compare properties and performance of different multivariate post-processing methods (ensemble copula coupling, dual ensemble copula coupling, Schaake shuffle and the Gaussian copula approach, see ([2]) for different situations (e.g., weather variables). The main message gained from the simulation study was that a misspecification in the multivariate dependence structure in the post-processing model leads to a notable deterioration in forecast performance.

Finally, the researchers of our group also complemented existing post-processing software packages [L1] with the necessary algorithms to deal with different types of weather variables.

#### Links:

[L1] <https://kwz.me/hSM>

#### References:

- [1] T. Gneiting, A. E. Raftery: “Weather forecasting with ensemble methods”, *Science* 310 (2005), 248–249.
- [2] S. Vannitsem, D. S. Wilks, J. W. Messner, (eds.): “Statistical Postprocessing of Ensemble Forecasts”, Elsevier, Amsterdam, 2018.

#### Please contact:

Sándor Baran

Faculty of Informatics, University of Debrecen, Hungary

[baran.sandor@inf.unideb.hu](mailto:baran.sandor@inf.unideb.hu)

## Enabling the European Business Knowledge Graph for Innovative Data-Driven Products and Services

by Dumitru Roman and Ahmet Soylu (SINTEF)

*Corporate information, ranging from basic company information such as company name(s) and incorporation date to complex balance sheets and personal data about directors and shareholders, are the foundation that many data value chains depend upon in various sectors. However, collecting and aggregating information about a business entity from relevant public and private sources and especially across borders and languages is a tedious and very expensive task rendering many potential business models non-feasible.*

The euBusinessGraph project [L1], funded by European Union’s Horizon 2020 programme, set the foundations for a business knowledge graph of companies, and delivered a set of innovative data-driven business products and services dealing with company information.

Governments and other public bodies are increasingly publishing open data about firmographics and contextual databases, which reference companies. For example, the UK, Norway, France, and Denmark openly publish records about companies, while other countries have various degrees of openness for their company registries. Examples of contextual databases include the EU TED (Tenders Electronic Daily) public procurement notices and gazette notices.

Unfortunately, firmographics datasets are not yet fully harmonised and interoperable because data differs widely in semantics from one source to another and data formats vary, ranging from UK’s five star Linked Data to poorly accessible and poorly documented datasets. Furthermore, contextual databases are not linked to the national company registries and they still use different company identifier systems or, in some cases, no identifiers at all. Private businesses also produce valuable company-related data, which is seldom linked to the public sources mentioned above. For example, media publishers often reference businesses and legal entities by name (hence ambiguously) even within their digital publications. This occurs because there is no widespread mark-up schema to annotate a digital reference to a company and no standardised way of accessing its information once it’s unambiguously identified. As a result, it is extremely expensive, time consuming, and error prone to find, interpret and reconcile these data from private sector sources.

One of the immediate consequences is that the business information sector is not very cost-efficient in itself, which is reflected in a lack of transparency and efficiency of the markets [1]. Nevertheless, the most relevant consequence in this context is that these inefficiencies severely harm digital