DOI: 10.1111/cgf.13279

Pacific Graphics 2017
J. Barbic, W.-C. Lin, and O. Sorkine-Hornung
(Guest Editors)

*Volume 36* (*2017*), *Number 7*

# Albero: A Visual Analytics Approach for Probabilistic Weather Forecasting

A. Diehl[†1] and L. Pelorosso[‡2] and C. Delrieux[§3] and K. Matković[¶4] and J. Ruiz[‖5] and M. E. Gröller[**6] and S. Bruckner[††7]

[1]University of Konstanz, Germany
[2] University of Buenos Aires, Argentina
[3] South National University, Argentina
[4] VRVis Research Center in Vienna, Austria
[5] DCAO (FCEN/UBA) - CIMA (CONICET/UBA), Argentina
[6] Vienna University of Technology, Austria
[7] University of Bergen, Norway

**Abstract**

*Probabilistic weather forecasts are amongst the most popular ways to quantify numerical forecast uncertainties. The analog regression method can quantify uncertainties and express them as probabilities. The method comprises the analysis of errors from a large database of past forecasts generated with a specific numerical model and observational data. Current visualization tools based on this method are essentially automated and provide limited analysis capabilities. In this paper, we propose a novel approach that breaks down the automatic process using the experience and knowledge of the users and creates a new interactive visual workflow. Our approach allows forecasters to study probabilistic forecasts, their inner analogs and observations, their associated spatial errors, and additional statistical information by means of coordinated and linked views. We designed the presented solution following a participatory methodology together with domain experts. Several meteorologists with different backgrounds validated the approach. Two case studies illustrate the capabilities of our solution. It successfully facilitates the analysis of uncertainty and systematic model biases for improved decision-making and process-quality measurements.*

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Viewing algorithms I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques I.3.8 [Computer Graphics]: Applications—Probabilistic Weather Forecasting

## 1. Introduction

Weather predictability influences many aspects of human life and constitutes a grand challenge [Gim13]. The chaotic nature of the atmosphere amplifies weather forecast uncertainty, imposing a limit on predictability [PH06, Kal03]. Numerical Weather Prediction System (NWPS) is the cornerstone of weather forecasts, and consists on the simulation of atmospheric dynamics by the numerical integration of physic-mathematical models of the climate system. In recent years, the amount of data generated by numerical weather prediction systems has been increasing significantly as the temporal and spatial resolution of these models has been continuously refined. The refinement produces a more detailed representation of the atmospheric flow at different scales.

Probabilistic weather forecasts are amongst the most popular ways to quantify forecast uncertainty. In probabilistic forecasts, the NWPS is used to estimate the occurrence probability of a particular weather phenomenon in a certain time window in the future (e.g., the probability of having rain in excess of 20 mm for a particular location in the next 24 hours). Several Model Output Statistics (MOS) [Wil06] methods have been designed to estimate these probabilities as accurately as possible. A combination of ensemble forecasts is employed, calibrated or adjusted using past forecasts and observational data. Hamill and Whitaker [HW06] introduced the “Reforecast Analog Regression” (RAR) method for probabilistic forecasting. It uses a large database of weather forecasts generated by means of a specific NWPS [Roz06] and an associated database of observations [Cli]. One of the main advantages of the RAR method

† e-mail: alexandra.diehl@uni-konstanz.de
‡ e-mail: lpelorosso@dc.uba.ar
§ e-mail: cad@uns.edu.ar
¶ e-mail: Matkovic@vrvis.at
‖ e-mail: jruiz@cima.fcen.uba.ar
** e-mail: groeller@cg.tuwien.ac.at
†† e-mail: stefan.bruckner@uib.no

is that the relationships between systematic model errors and large scale atmospheric circulation patterns can be captured, which leads to a better estimation. Information provided by the analogs and their corresponding observations can be expressed not only as a probability for a particular weather phenomenon, but also as sample mean, standard deviation, and maximum and minimum values. This can help to quantify the forecast uncertainty and to detect the occurrence risk of extreme weather events.

### 1.1. Tasks

The RAR method is successfully used to support forecasters' decision making under uncertainty. Prior to the development of our system, the workflow of the users consisted of setting parameters of the RAR algorithm in a limited way, with the main aim of generating a single probabilistic forecast and to perform a visual comparison by juxtaposing the probabilistic forecasts.

In our approach we propose to put the users in the loop to harness their knowledge and background for improving the task workflow. The goals of our work are to support the expert users in the execution of the following tasks for operational weather forecasting.

- T1: Support a flexible configuration of the RAR method for a particular region, forecast lead time, and purpose.
- T2: Detect the occurrence likelihood of an event.
- T3: Detect possible extreme events.
- T4: Analyze the historical behavior of the numerical forecast model. This task involves the comparison of the current numerical forecast with similar past forecasts and their associated observations.
- T5: Analyze the uncertainty of the forecast expressed through the probabilities of different events.
- T6: Study the systematic model bias and get further insight into model errors.
- T7: Analyze the spatial error distribution for a given accumulation range. This is a measure of the method accuracy for a particular situation.
- T8: Analyze previous similar forecasts in specific sub-regions.
- T9: Inspect and compare the current numerical forecast with historical observations.

### 1.2. Contributions

We propose a novel approach that uses Visual Analytics (VA) to create a new, flexible, and efficient task workflow for probabilistic weather forecasting (see Figure 1). Our main motivation is to use VA to support forecasters in their operational tasks and decision making. The workflow encompasses three loops. The first, Parameterization Loop, presents an overview of the probabilistic forecasts and allows the experts to calibrate parameters of the RAR method. The second, Probabilistic Forecast Loop, goes one level deeper in the analysis of the probabilistic forecasts and the model errors for a specific region of the map. The third, Analogs Loop, provides a detailed view of the construction of a probabilistic forecast, every element used for its calculation, and the associated model biases and uncertainty.

Our tool is designed with a balanced choice of visual components and interactions, leveraging the capabilities of coordinated multiple views (CMV) [Rob07] in an operational weather forecasting environment. It does not comprise new visualization techniques but rather their adaptation to improve decision making based on RAR. Our main contributions can be summarized as follows:

- C1: Enable a flexible configuration of the RAR method for different regions, applications, and purposes.
- C2: Support the visual inspection of the internal parts of the RAR method to improve decision making.
- C3: Support the visual analysis of systematic model errors under different atmospheric circulation regimes.

To illustrate our work, we have chosen the precipitation accumulation variable because it is one of the most influential variables impacting on human activities. It affects different aspects of human life such as planning, decision making, and productivity. Flooding and severe storms can cause disasters and are a big issue in many places, and in particular in South America, where our approach is applied. From 1970 to 2017, there were 75 massive flooding in Argentina, affecting 13 million people and causing the lost of 500 lives. The target beneficiaries of our approach are expert forecasters (EFs) that need to analyze big datasets to produce timely and accurate forecasts for different weather events.

## 2. Related Work

There has been extensive work done in the area of visualization systems and VA applied to weather forecasting and climatology. In the scope of this work, only the most representative examples are summarized as background references to our approach. Recently, Ferstl et al. [FKRW17] have proposed an approach to visualize the uncertainty of ensembles by means of clustering contours as well as juxtaposition and animating spaghetti plots of iso-contours. In our work, the ensemble of analogs can be very heterogeneous. Therefore, the use of spaghetti plots can produce very cluttered and busy visualizations that do not permit to follow the order of the analog forecasts. Instead, we use a list of small multiples ordered by similarity, juxtaposition of small multiples to display probabilistic forecasts, comparative visualization of the probabilistic forecasts, numerical forecasts and method errors, and a detailed visualization of the ensemble of analog forecasts.

Biswas et al. [BLLS17] explore the sensitivity of a model with respect to a specific set of parameters and to a horizontal resolution. In our case, we evaluate –using the RAR method– the probability of occurrence for particular weather events. We assume that there is a user who will make a decision based on probability values. Probability is often used in the field of operational weather forecasting and it is accepted by users of meteorological forecasts for efficient decision making. Other noteworthy previous works [WLSL17, CZC*15] address also the investigation of uncertainties by means of a sensitivity analysis and visual analytics of the parameter space. A sensitivity analysis is a very important step to determine the most influential parameters of a given model. Instead, we use a different approach, the Reforecast Analog Regression, to create probabilistic forecasts for a given variable, a given threshold (event), and a specific lead time. A key point of our method is the usability of probability forecasts compared to deterministic numerical forecasts [Cou06]. We want to provide an interface to allow a large number of users with different requirements to obtain useful information for decision making.

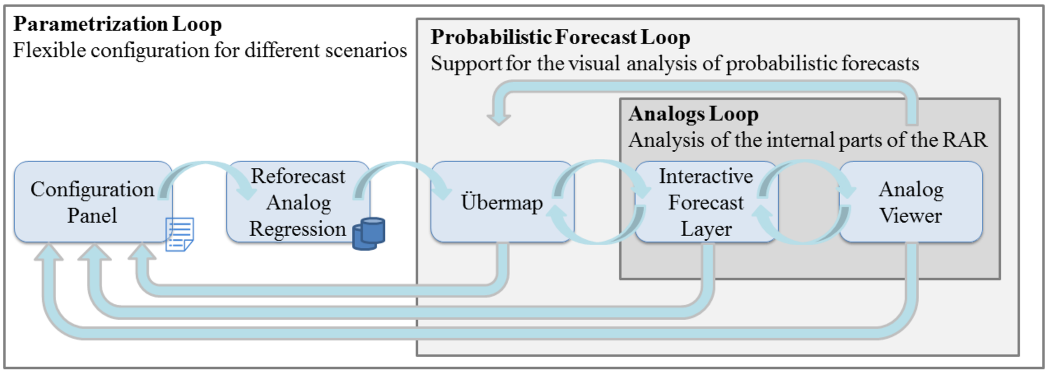



**Figure 1:** *Schema of our VA task workflow. It shows different analysis paths that correspond to different tasks the Expert Forecasters (EFs) want to perform. For example, the parameterization loop covers the complete workflow. The EFs can continue their analysis down to the deepest level of detail and afterwards re-configure the initial settings to improve the analysis and decision making.*

Quinan and Meyer [QM16] described important design aspects of a VA solution to assist experts in the visual comparison of weather features in forecasts. They proposed informed color encoding for different weather variables and the use of contour plots to display multiple isocontour features across an ensemble. In our approach, we also visualize an ensemble of analogs, but the use of contour plots would not be effective given the variability between analogs. The same applies for the ensemble of observations. Also, we only display the mean of the ensemble of reforecasts. It would be an interesting approach for future work to incorporate the members of the ensembles of forecasts. Many tools are available for meteorologists to gain insight and to communicate the uncertainty in ensembles of forecasts [PWB*09], [SZD*10], [HDKS05], [RKSW15], [BGS12].

We focus on the VA of uncertainty using probabilistic weather forecasts, in particular RAR. The NOAA Climate Prediction Center [NOAb] provides visualizations of probabilistic forecasts, but with very little interactivity. The European Centre for Medium-Range Weather Forecasts provides a web-based visualization system named "ecCharts" that enables the user to interactively specify thresholds and display probabilities computed from the ensemble forecasts [SLTO*15]). Our solution allows the visualization of the probabilistic forecasts but also provides a novel VA approach. The user can inspect the intermediate results and better evaluate the performance and quality of the RAR method. Our approach embraces the same idea as expressed through Anscombe's quartet [Ans73]), i.e., to reveal the disparities and outliers in statistical calculations.

## 3. Visual Analytics Workflow

Our solution unfolds the internal steps and data used in the RAR method by means of a balanced combination of linked-views and interactive visualization methods with semi-automatic processing. There are three main points in the task workflow where the user can intervene with the RAR method. The first point is the interactive configuration of the method (Parameterization Loop). The second point is the analysis of the uncertainty and decision making based on the RAR method (Probabilistic Forecast Loop). The third point is the detailed analysis of the analogs (Analogs Loop) to assess the quality of the method itself.

**Parameterization Loop**. This iterative process allows the users to set temporal and spatial parameters of the Hamill and Whitaker algorithm [HW06]. Our solution provides a flexible configuration panel and a comparative view, i.e., the “Übermap” (see Figure 2). It allows users to explore different thresholds, accumulation ranges, and lead times depending on their analysis needs. For example, if the EFs are analyzing a forecast for the next two weeks (lead time), they know that the uncertainty is high, so it does not make sense for them to accumulate a variable per day. In this case the EFs could choose an accumulation range of one-week. If the EFs want to analyze the forecast with a lead time of 24 hours, they will rather choose an accumulation range of 6 hours, for example, to know if there will be rain during the morning, the afternoon, or the night. Furthermore, to perform these analyses the EFs should be able to modify the corresponding thresholds. Setting the thresholds depends on the kind of analysis the EFs want to do and in which region it will be done. A high probability of rainfall more than 1 mm is important for a building construction but may not be for a city, where only a high probability of rainfall more than 100 mm may create an alert.

**Probabilistic Forecast Loop**. EFs analyze the uncertainty of the weather forecast models from different perspectives. They use probabilistic forecasts to explore the uncertainty at different levels of detail and compare the levels of uncertainty both spatially and temporally. In the Probabilistic Forecast Loop, users can visualize a set of probabilistic forecasts for a given meteorological variable, analyze the associated numerical forecast results, and summarize information about observations and uncertainty. They can also visualize and explore a given spatial sub-region used by RAR to construct the probabilistic forecasts (see Figure 4) and continue further analyzing other sub-regions. This iterative process allows the experts a seamless exploration of the aforementioned sub-regions and their associated spatial errors. During this task, forecasters can detect extreme situations such as intense precipitation or flooding in a given sub-region and analyze the spatial distribution of the uncertainty there.

**Analogs Loop**. Meteorologists can improve their decision making process by analyzing how the probabilistic forecast is constructed and how the RAR method performs [TTW15]. Our solution has a break-down view, named the Analog Viewer (see Figure 5), that allows users to access information about forecast analogs and their associated observations, as well as statistical aggregations. For example, they can visualize and analyze the analogs, how the similarity metric works, and the distribution of the systematic model errors under similar atmospheric circulation regimes. The Analog Viewer also shows different statistical summaries of the analogs and the observations.

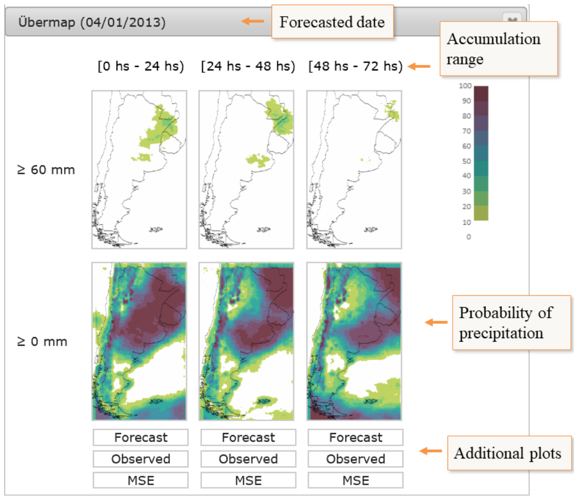



**Figure 2:** *The Übermap provides an overview visualization displaying probabilistic forecasts for various thresholds of precipitation accumulation and accumulation ranges.*

The EFs can interactively adapt the visualization pipeline depending on the task they want to perform. We exemplify the various capabilities of our system in Section 6. The solution is organized as a CMV system with several linked and coordinated views. Among them we may count:

- A map view that allows to zoom in and out and move to different geographical regions.
- The configuration panel that supports the contribution C1 and is used for the Parameterization Loop.
- The Übermap that supports contributions C1 and C2 and is used for the Probabilistic Forecast Loop.
- The Interactive Forecast Layer that supports contributions C2 and C3 and is used for the Probabilistic Forecast Loop and the Analogs Loop.
- The Analog Viewer that supports the contribution C3 and is used for the Analogs Loop.
- The Comparative View that allows the users to analyze at the same time by juxtaposing the probabilistic forecast, the current numerical forecast, and the distribution of the spatial errors.

## 4. Design Process and Choices

We followed a participatory design process [SMM12, SSS*14, Mun14], closely collaborating with seven domain experts who are meteorologists with different backgrounds. Two of them work at the Center of Research of the Sea and the Atmosphere (CIMA), one of them is a senior researcher who participated from the beginning of the project, and the other one has a computer science background. The other five meteorologists are researchers who work at the National Weather Services (NWS) in Argentina. Our work required extensive interdisciplinary efforts between meteorologists and computer scientists.

This section describes different phases of the design process, the results of our meetings, as well as milestones and prototypes built. The work was structured into three phases: inception, design, and evaluation (Figure 6). The goal of the inception phase was to generate ideas for content and design, to gain a better understanding of our target audience, and to get insights into the requirements of our domain experts. The goal of the design phase was to gather detailed requirements from the experts, discuss prototypes with them and jointly elaborate and validate design choices. This phase was performed following an iterative and incremental process. Finally, the goal of the evaluation phase was to present our solution, design case studies, and evaluate them with the domain experts.

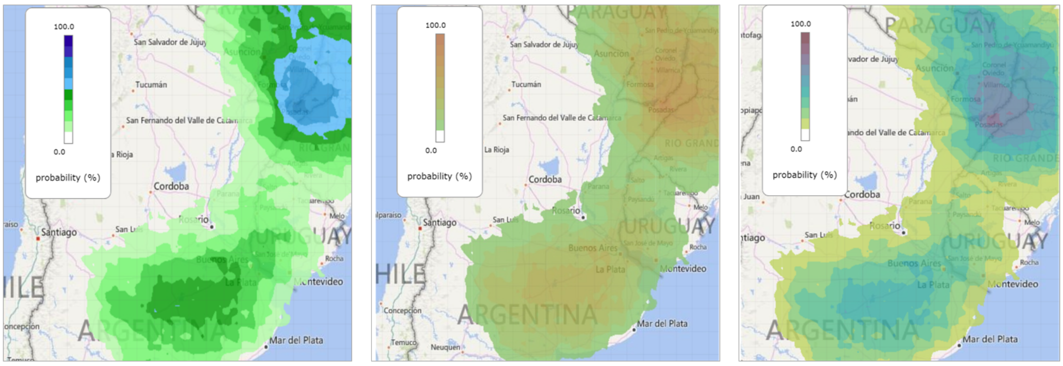
### 4.1. Inception Phase

During the inception phase we discussed a list of different state-of-the-art visualization techniques and practices with our domain experts. Examples include the information-visualization seeking-mantra, focus+context techniques, and linking and brushing. In their previous batch workflow software, they did not have interactivity at all. We provided them with a small and simple set of consistent visual encoding strategies that we devised according to their previous workflow and tasks. We drafted a first prototype based on their tasks and requirements. Our prototype presented a configuration panel that supports the Parameterization Loop. Our domain experts found it very useful because it allows EFs to understand how the method performs under specific atmospheric characteristics in a region, and with a particular application or purpose in mind. We discussed the design of the prototype with our domain experts.

### 4.2. Design Phase

We continued the participatory design process with a series of sessions involving all the experts and iterating over the prototypes. We followed an agile development process. After each iteration, experts evaluated the last prototype, highlighting desirable features, suggesting changes, and proposing new functionalities. All actors benefited from this methodology. Experts discovered new visual features that could improve their work. We learned more about the needs of the experts, which resulted in features that are described in this section. The iterative process allowed us to identify the three main VA loops discussed before.

Our prototypes evolved to incorporate several coordinated and linked views. One gives an overview of probabilistic forecasts for different thresholds, i.e., the Übermap (see Figure 2) . The Übermap



(a) (b) (c)

**Figure 3:** *Different color maps evaluated by our domain experts during the iterative process to refine our design choices. (a) Initial color scheme. (b) Two-hue scheme from green to red. (c) Two-hue scheme from green to violet. For representing the precipitation accumulation probabilities, they preferred the color map (c) with a two-hue HCL color scale ranging from transparent through green to violet.*

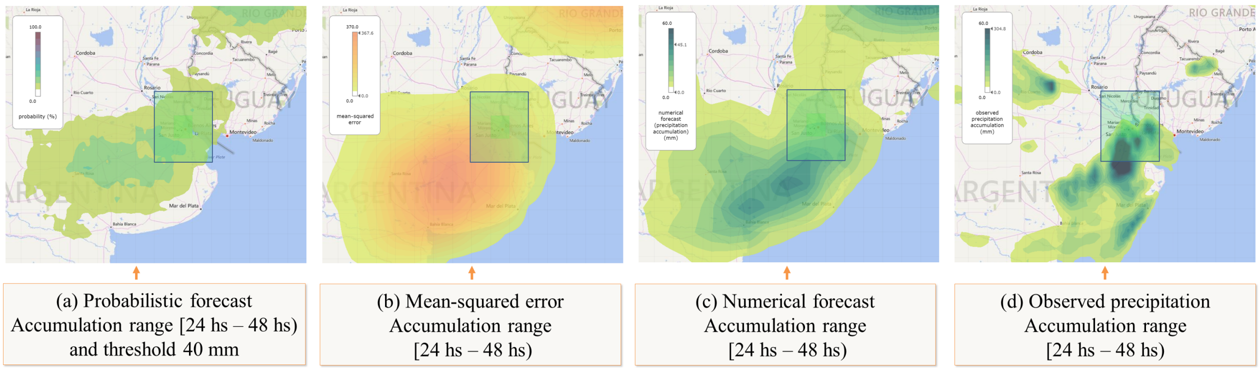



**Figure 4:** *Interactive Forecast Layer close-up. (a) probabilistic forecast divided into the sub-regions used by the RAR, (b) spatial distribution of the Mean Squared Error, (c) numerical forecast, (d) observations. By clicking on one of the sub-regions the user has access to detailed information on the ensemble of analogs and observations used in the computation of the probabilistic forecast for that specific area.*

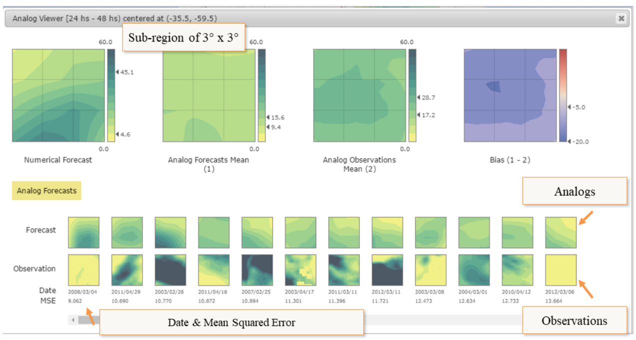



**Figure 5:** *Analog Viewer for the case of a specific selected region on the first accumulation range [24 hours – 48 hours).*

provides a comparative visualization for different thresholds, lead times, and intervals of accumulation precipitation. The experts found this approach useful to quickly perform different tasks. The Übermap allows the EFs to do different types of analyses, dependent on the decision making scenarios. For example, the forecasters may want to detect if there is any sizable probability of rain or not. Goals might be to communicate and broadcast results to different stakeholders, or to analyze specific high thresholds of precipitation accumulation to detect extreme events.

Our domain experts proposed to display the numerical forecast as a 2D scalar field on the map, as well as the spatial distribution of the RAR errors. The analysis of this information provides the EFs with information about the RAR behavior for a particular event. We proposed a comparative visualization by juxtaposition that allows the EFs to analyze all the scalar fields at the same time. For this purpose, we built the Interactive Forecast Layer (see Figure 4) that presents a dashboard with a main view showing the probabilistic forecasts. Three ancillary side views depict the associated numerical weather forecasts, observations if available, and errors.

An important outcome from our meetings was the need to show the EFs the inner workings of the RAR method to convey an under-

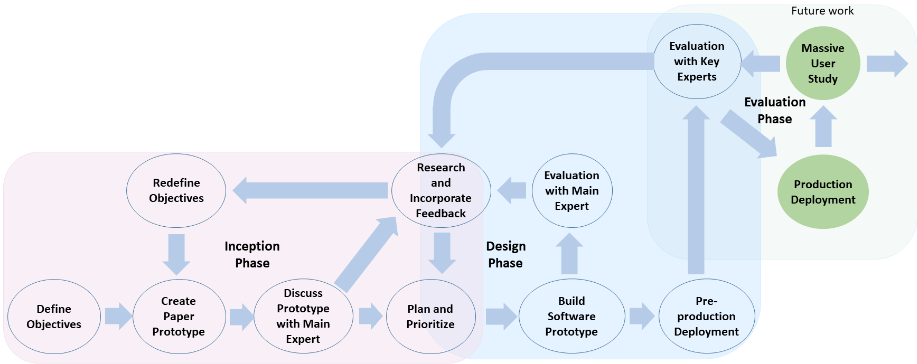



**Figure 6:** *Flowchart describing the three main phases followed during the participatory design process and the evaluation of our solution.*

standing of its behavior and performance. We asked questions to the experts concerning the importance to perform a fast and accurate analysis of the data. We asked also about the relevance of exploring multiple regions or perform a multi-scale analysis. The multi-scale analysis regards the transition from large spatial scales to a more detailed analysis in those areas where large amounts of precipitation are more likely. For this purpose, we propose the Analogs Loop to identify in which sub-regions the method selected similar or disparate analogs. It also enables the experts to analyze the outputs of the NWPS model by comparing analogs and observations. The Analog Viewer displays analogs, their associated errors, and observations involved in the computation of each sub-region with the RAR method. We chose small multiples ordered by increasing similarity to show the ensemble of analogs. In this way, EFs can focus their attention on the forecasts that are more similar to the current situation. We complemented this visualization with summary views to support the analysis of the errors via statistical aggregations.

### 4.3. Lessons Learned

A critical point has been the visualization of detailed information about the analogs and the metrics used to select them. According to the EFs, the Analog Viewer revealed that MSE seems to be a pessimistic metric. It is further heavily affected by spatial shifts in the analogs. Similar phenomena that occurred in the past, but not exactly at the same geographic location, have a strong impact on MSE. This has been a significant discovery facilitated by our tool. Color palettes have been other concern discussed during our interviews. We identified that the initially selected colormap for the probabilistic forecasts was not expressive enough to easily detect probabilities on the map. In general, the variations in our colormaps were considered to be too smooth for forecasters to detect adjacent levels of a variable. Although the rainbow colormap can lead to wrong interpretations, forecasters find it helpful to identify at a glance variations of different thresholds and to quickly spot a severe event. Changing color scales was also not considered to be helpful as this increases the mental load on the users to familiarize themselves with the changes by repeatedly checking the color legends. Figure 3 gives a comparison of different color scales analyzed during the design phase. The color scale in Figure 3c from green to violet was rated to be the best by the domain experts. Our key experts decided to include the tool into their regular workflow, but we still are on pre-production stage. To make our tool operative, we still need a large user study first, where operational forecasters take part to evaluate the reliability and usability of our tool.

## 5. Implementation

Albero (see Figure 7) is an open-source web application for probabilistic weather forecasting, based on a previous client/server model [DPD*15]. The web client offers an interactive visual interface built using HTML, JavaScript and JQuery technologies. BingMaps AJAX Control 7.0 is used to display the Interactive Forecast Layer view. The multi-threaded server is realized in C++. We implemented the Smoothed Rank Analog technique developed by Hamill and Whitaker. The prototype, documentation, and source code is publicly available at: `https://albero.cg.tuwien.ac.at/`.

### 5.1. Data Sources

The system uses retrospective forecasts (shortly reforecasts) generated by the GFS (Global Forecast System) [NOAa]. The reforecast data correspond to the GEFS Reforecast v2 Ensemble Data. The datasets are stored in NetCDF4 format, containing only the mean value of the ensemble. It uses a 30 year historic database of forecasts generated by the system. These forecasts are global and have a maximum lead time of two weeks with a horizontal resolution of approximately 28 kilometers for the first week and approximately 70 kilometers for the second week. They include meteorological variables such as temperature, winds, pressure, humidity, and precipitation among others. In this work, we focus on the precipitation variable. Precipitation forecasts are available as the precipitation accumulated in the last three-hours.

The observations database is constructed using the CMORPH (Climate Prediction Center - NOAA) [Cli] precipitation estimates. The CMORPH observational data is available as an estimated precipitation grid, with a spatial resolution of 0.25 degrees, for every three hours. The resolution of the probabilistic forecast is given by the resolution of the observations of CMORPH, i.e., 0.25 degrees. The use of CMORPH past observations to generate probabilistic forecasts allows us to increase the forecast horizontal resolution. We

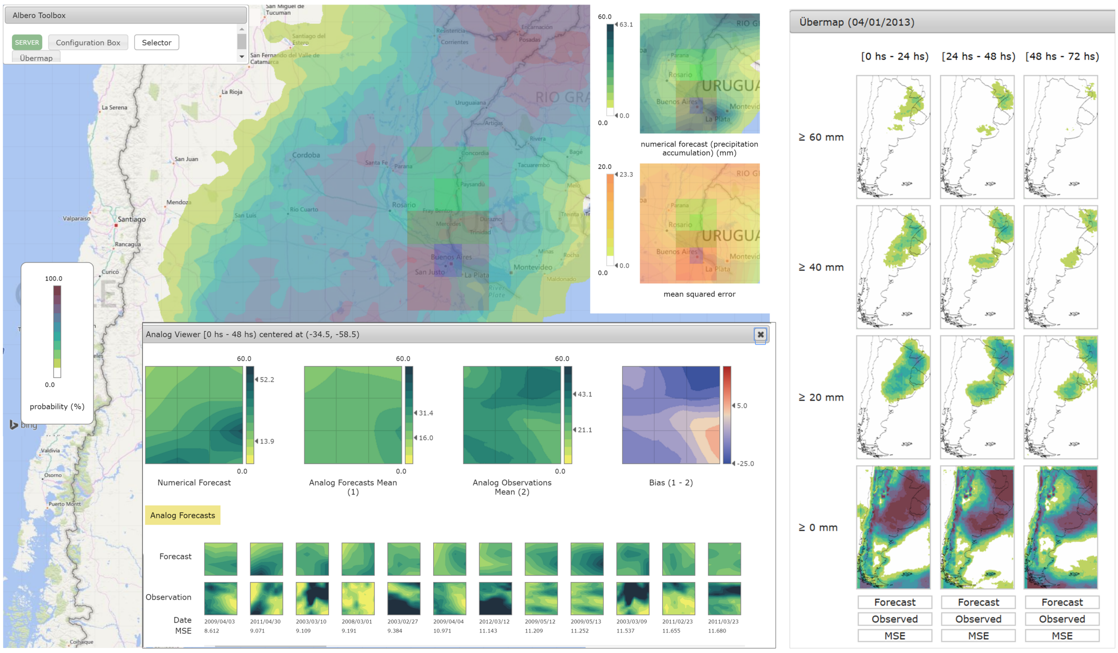



**Figure 7:** *Albero as a visual Analytics solution for the analysis of probabilistic forecasts. Different linked-views are coordinated in a fluent dialog between the users and the application to assist decision making processes and measure the quality of the Reforecast Analog Regression.*

can decrease the resolution from the native 1 degree resolution of the reforecast data to the 0.25 degree resolution of the observations. In this sense, the RAR technique is acting as a statistical down-scaling technique. The precipitation estimates provide data between 40S and 40N latitude with a temporal resolution of as high as 30 minutes and a horizontal resolution of as high as 8 km. In this paper we use a lower-resolution version of the data that has 3 hours and 0.25 degrees resolution. The precipitation estimates are based on passive microwave radiometers on board of several satellites. This data-set is routinely used to verify NWPS precipitation forecasts as well as to study the climatic variability of rainfall at different temporal and spatial scales.

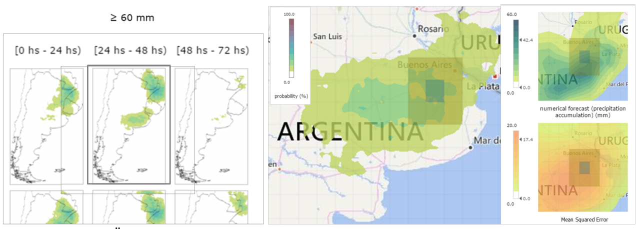
The reforecast database starts in 1984. However, CMORPH estimates began in 2002. Therefore, the analog search in this work is limited to the period 2002–2016 in which both databases are available. Through the combination of NCEP reforecasts and CMORPH databases, the algorithm can be easily implemented for any location between 40S and 40N latitude.

### 5.2. Automatic Processing

The RAR method consists of two main steps. In the first step, the algorithm identifies a given number of past forecasts that are similar to the current deterministic forecast using a specific NWPS and for a particular geographic region. The similarity is evaluated using different metrics such as the Mean Squared Error. At each grid point of the selected region, analogs are identified by computing the similarity between the current forecast and forecasts with the same lead times in the historical database. Calculations are done over a squared domain centered at the grid point and a yearly time window of $M$ days positioned at the forecasted date. Once the $N$ most similar reforecasts are identified, the observed precipitation corresponding to the forecasted period is retrieved from the observations database and a set of $N$ observed precipitation fields is obtained. The probability of the precipitation exceeding the user-selected thresholds as well as some other sample parameters are obtained from the $N$ observed fields. $N$ is the number of retained analogs and is a parameter in the method. If $N$ is too small, the data sample may be too small to produce a robust estimation of the probability or sample parameters. If $N$ is too large, not so good analogs might be included in the sample, which produces a degradation of the estimated probability. By replacing model-produced forecasts by their corresponding observations, the systematic errors of these forecasts are partially removed. This results in a less biased estimation of the probability.

## 6. Case Studies

We exemplify the capabilities of Albero on two scenarios that cover operational weather forecasting and retrospective weather analysis for an event of extreme precipitation in a short period of time. This event took place from the 1st to the 3rd of April in 2013 causing flooding in different areas of the north-eastern part of the region. Both scenarios cover the analysis of extreme precipitation in a short-period of time. The EFs assumed that the selected event could be underestimated by the numerical model, since it was a rare and very extreme situation and, in general, precipitation is one of the most difficult meteorological variables to forecast. We performed two guided interviews with seven domain experts with different backgrounds.




(a) Übermap (b) Interactive Forecast Layer

**Figure 8:** *Albero showing (a) the Übermap for April 1st, 2013 for the next 72 hours and (b) a detailed view of the probabilistic forecast for precipitation above 60 mm.*

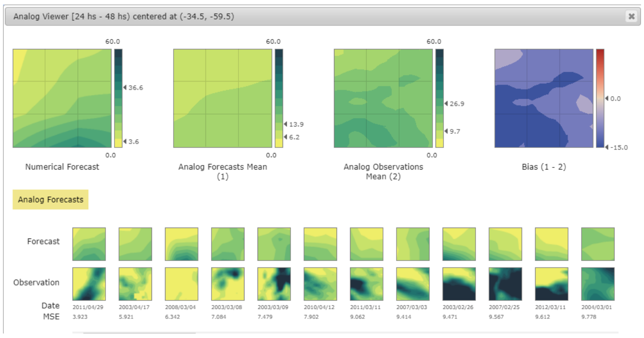



**Figure 9:** *Albero Analog Viewer with RAR results over a region centered at longitude* 59.5°*W and latitude* 34.5°*S, in the province of Buenos Aires.*

### 6.1. Extreme Precipitation Analysis

The goal of the EFs was to evaluate the risk of high precipitations up to 72 hours in advance over southern South America. The EFs set the forecast maximum lead time to 72 hours, dividing this time window into 24 hour intervals. The thresholds from which probabilities will be computed can be adjusted – in this case the selected thresholds were 0 mm, 20 mm, 40 mm, and 60 mm. Setting the precipitation threshold and time intervals allows the EFs to detect different, potentially dangerous events (corresponding to tasks T1, T2 and T3 as identified in Section 1.1). This allowed the EFs to evaluate events of intense precipitations over short periods of time that can be associated with flash floods.

Based on these inputs, Albero presented an overview of the probability associated with the different thresholds grouped by interval-time columns as shown in Figure 8a (task T4). Green and blue shades in the second row and second column of the Übermap indicate high probabilities of precipitation over 20 mm during the second day over a large portion of the selected area. For the first and second forecasted day there was a probability of over 30% for precipitation over 60 mm in 24 hours. Although this may still seem like a low probability of occurrence, it is much higher than the climatological probability of the event. This indicates that the risk of having a high precipitation event for that period is considerable (task T2). The EFs selected a probabilistic forecast from the Übermap where the probabilities were above 60 mm. The EFs focused their attention on a specific sub-region near Buenos Aires city and zoomed into the Interactive Forecast Layer (see Figure 8b). They noticed that for the second time interval, corresponding to the second day, the heaviest rain was forecasted. Figure 10 shows a close-up of the probabilistic forecasts for the second time interval of [24-48 hs), for the thresholds 20 mm, 40 mm, and 60 mm.

Using the Analog Viewer, the EFs analyzed the numerical forecast, the analogs and the observations. Figure 9 depicts the first twelve analogs and their corresponding observations (task T7). The mean difference between the analogs and the observations provides information about the systematic errors in the model associated with situations that are similar to the one anticipated in the current numerical forecast (task T6). By using Albero, the EFs can study systematic model bias and get further insight into their model errors. The bias is also displayed in Figure 9 where large blue and green regions dominate in the selected sub-region. The EFs confirmed that the numerical forecast systematically underestimated the amount of precipitation accumulation since the observations mean is higher than the analogs forecast mean (task T4). The EFs were able to visualize the spatial error distribution for a given accumulation range, as shown in Figure 10d. The highest errors are located in the central part of the country. They are encoded in red shades (task T7).

### 6.2. Visual Quality Analysis of the Analogs

In this scenario, the EFs analyzed previous events (T6). Using the Analog Viewer, they initially compared the numerical forecast, the analogs and observations. The bias, calculated as the difference between the analogs and the observations, provides information about the systematic errors in the model (task T6). This analysis allowed the EF to validate whether the numerical forecast underestimated or overestimated the precipitation accumulation across the three days (task T4). They analyzed the bias between the analogs mean and the observations mean, and as a result they detected large variations between them. Figure 11 shows that the numerical forecast systematically underestimated the amount of precipitation accumulation for a selected sub-region in the given accumulation range, since the observations mean is higher than the analogs mean. This can also be observed in the individual analogs and their associated observations, and can be considered a systematic model bias. For further analysis, the EFs could select the "observation" option at the bottom of the Übermap window to display the averaged observed precipitation corresponding to the analog forecasts. This provides information about the expected value of the precipitation after partially removing the systematic forecast errors (task T8). Moreover, by selecting the MSE option, the EFs display the spatial distribution of the Mean Squared Error (see Figure 10d). This is a measure of the uncertainty associated with the expected precipitation value over the domain. The highest errors are located in the central and north-eastern part of the domain (task T5).

## 7. Conclusions

We presented a VA approach to support probabilistic weather forecasting using the RAR method. We performed a thorough analysis of the tasks workflow of our domain experts, in order to understand how they analyze their data and the associated uncertainty.

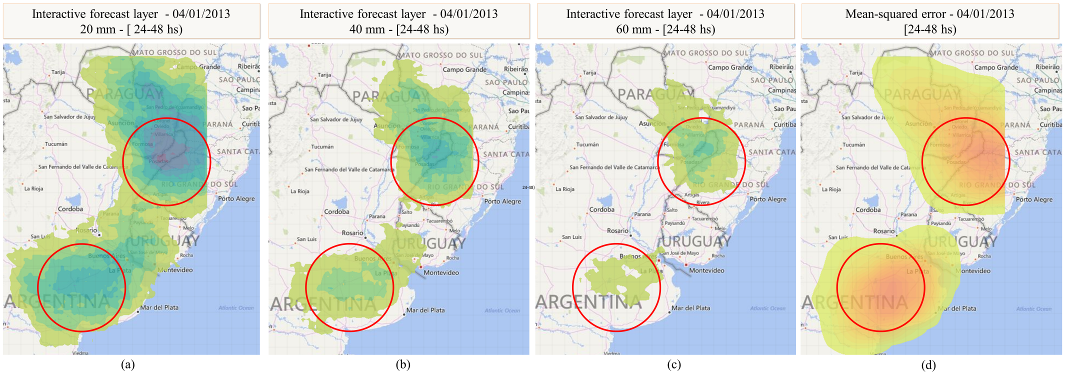
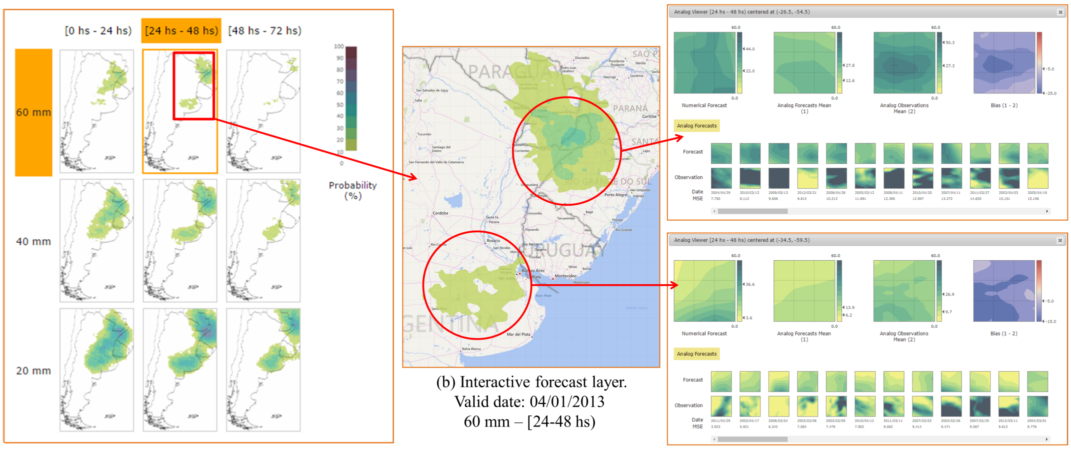



**Figure 10:** *Interactive Forecast Layer showing different forecasts, observations and errors for the second day. The interval of precipitation accumulation is [24-48 hs), valid for April 1st, 2013. Probabilistic forecasts show high probability of precipitation over the north east region.*




(a) Übermap 04/01/2013

(b) Interactive forecast layer. Valid date: 04/01/2013 60 mm – [24-48 hs)

(c) Analog viewer for both salient regions.

**Figure 11:** *Extreme precipitation happened during April 2nd, 2013. The second day is shown because the heaviest rain was forecasted for that date. (a) The user identifies a probability forecast with a large area of precipitation accumulation above 60 mm. (b) Two areas of interest are highlighted. (c) The user finds out that the numerical forecast systematically underestimated the amount of precipitation accumulation since the observations mean is higher than the analogs mean. Also the analog forecasts for both regions presented noticeable differences.*

Our approach is the first interactive VA solution that exposes the internal aspects of the RAR method. It provides meteorologists with capabilities to connect and compare different pieces of information from probabilistic forecasts and internal views of the RAR method. The interactive visual inspection of local analogs and observations, and how they were selected, provides our domain experts with additional information beyond the probabilities themselves. Though it is unpractical to analyze all sub-regions one by one, Albero can be very helpful for the analysis of specific sub-regions of interest. For example, it can be useful to detect high probabilities of strong precipitation in a region that is densely populated. Our solution provides the experts with detailed views of systematic errors and extreme values, among other statistical summaries. We plan to extend our work with additional statistical characteristics for the analogs, the forecast bias with respect to the observations, and the detection of specific weather features. We will also include a statistical evaluation of different similarity metrics. Our domain experts suggest to include an objective validation of the method using statistical scores and machine learning algorithms for parameter optimization. We are now deploying an operative environment of Albero. This will allow us to perform future exhaustive tests and massive user studies of our solution.

## 8. Acknowledgements

We thank Rodrigo Pelorosso and Rohit Choudhary for their technical help. This research has been supported by DFG in Priority Research Programme "Volunteered Geographic Information: Interpretation, Visualisation and Social Computing" (SPP 1894); by the MetaVis project (Nbr. 250133) funded by the Research Council of Norway; by PICT 2014-1000 funded by MinCyT, Argentina; by the VRVis is funded by BMVIT, BMWFW, Styria, SFG and Vienna Business Agency in the scope of COMET - Competence Centers for Excellent Technologies (854174), which is managed by FFG.

## References


[Ans73] ANSCOMBE F. J.: Graphs in statistical analysis. *The American Statistician 27*, 1 (1973), 17–21. URL: http://www.jstor.org/stable/2682899. 3

[BGS12] BOCK L., GÜNTHER M., SCHAAL K.: Gutes Wetter für die Landwirtschaft: Agrarmeteorologische Beratung mit der Visualisierungssoftware Ninjo. In *Informationstechnologie für eine nachhaltige Landbewirtschaftung, Fokus: Forstwirtschaft, Referate der 32. GIL-Jahrestagung, 29. Februar - 1. März 2012, Freising, Germany* (2012), pp. 47–50. 3

[BLLS17] BISWAS A., LIN G., LIU X., SHEN H.-W.: Visualization of time-varying weather ensembles across multiple resolutions. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 841–850. doi:10.1109/TVCG.2016.2598869. 2

[Cli] CLIMATE PREDICTION CENTER, NATIONAL CENTERS FOR ENVIRONMENTAL PREDICTION, NATIONAL WEATHER SERVICE, NOAA, U.S. DEPARTMENT OF COMMERCE: NOAA CPC Morphing Technique (CMORPH) Global Precipitation Analyses, accessed in april 2017. doi:10.5065/D6CZ356W. 1, 6

[Cou06] COUNCIL N. R.: *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts*. The National Academies Press, Washington, DC, 2006. doi:10.17226/11699. 2

[CZC*15] CHEN H., ZHANG S., CHEN W., MEI H., ZHANG J., MERCER A., LIANG R., QU H.: Uncertainty-aware multidimensional ensemble data visualization and exploration. *IEEE Transactions on Visualization and Computer Graphics 21*, 9 (2015), 1072–1086. doi:doi.ieeecomputersociety.org/10.1109/TVCG.2015.2410278. 2

[DPD*15] DIEHL A., PELOROSSO L., DELRIEUX C., SAULO C., RUIZ J., GRÖLLER M. E., BRUCKNER S.: Visual analysis of spatio-temporal data: Applications in weather forecasting. *Computer Graphics Forum 34*, 3 (2015), 381–390. doi:10.1111/cgf.12650. 6

[FKRW17] FERSTL F., KANZLER M., RAUTENHAUS M., WESTERMANN R.: Time-hierarchical clustering and visualization of weather forecast ensembles. *IEEE Transactions on Visualization and Computer Graphics* (2017), 831–840. doi:10.1109/TVCG.2016.2598868. 2

[Gim13] GIMENO L.: Grand challenges in atmospheric science. *Frontiers in Earth Science 1*, 1 (2013). doi:10.3389/feart.2013.00001. 1

[HDKS05] HAO M. C., DAYAL U., KEIM D. A., SCHRECK T.: Importance-driven visualization layouts for large time series data. *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005. 0* (2005), 27. doi:10.1109/INFVIS.2005.1532148. 3

[HW06] HAMILL T. M., WHITAKER J. S.: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. *Monthly Weather Review 134*, 11 (2006), 3209–3229. URL: http://www.esrl.noaa.gov/psd/people/tom.hamill/reforecast_analog_v2.pdf. 1, 3

[Kal03] KALNAY E.: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 2003. URL: https://books.google.com.ar/books?id=Uqc7zC7NULMC. 1

[Mun14] MUNZNER T.: *Visualization Analysis and Design*. A.K. Peters visualization series. A K Peters, 2014. URL: http://www.crcpress.com/product/isbn/9781466508910. 4

[NOAa] NOAA EARTH SYSTEM RESEARCH LAB, PHYSICAL SCIENCES DIVISION, BOULDER, COLORADO, USA: A Description of the 2nd-Generation NOAA Global Ensemble Reforecast Data Set, accessed in Dec. 2016. URL: http://www.esrl.noaa.gov/psd/forecasts/reforecast2/. 6

[NOAb] NOAA EARTH SYSTEM RESEARCH LAB, PHYSICAL SCIENCES DIVISION, BOULDER, COLORADO, USA: Precipitation Forecast Products Based on NCEP GEFS Reforecasts, Version 2, accessed in Dec. 2016. URL: http://www.esrl.noaa.gov/psd/forecasts/reforecast2/analogs/. 3

[PH06] PALMER T., HAGEDORN R.: *Predictability of Weather and Climate*. Cambridge University Press, 2006. 1

[PWB*09] POTTER K., WILSON A., BREMER P.-T., WILLIAMS D., DOUTRIAUX C., PASCUCCI V., JOHHSON C. R.: Ensemble-vis: A framework for the statistical visualization of ensemble data. In *IEEE Workshop on Knowledge Discovery from Climate Data: Prediction, Extremes.* (Oct 2009), pp. 233–240. 3

[QM16] QUINAN P. S., MEYER M.: Visually comparing weather features in forecasts. *IEEE Transactions on Visualization and Computer Graphics (InfoVis '15) 22*, 1 (Jan 2016), 389–398. doi:10.1109/TVCG.2015.2467754. 3

[RKSW15] RAUTENHAUS M., KERN M., SCHÄFLER A., WESTERMANN R.: Three-dimensional visualization of ensemble weather forecasts part 1: The visualization tool met.3d (version 1.0). *Geoscientific Model Development 8*, 7 (2015), 2329–2353. doi:10.5194/gmd-8-2329-2015. 3

[Rob07] ROBERTS J. C.: State of the art: Coordinated & multiple views in exploratory visualization. In *Proceedings of the Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization* (Washington, DC, USA, 2007), CMV '07, IEEE Computer Society, pp. 61–71. doi:10.1109/CMV.2007.20. 2

[Roz06] ROZUMALSKI R.: *WRF Environmental Modeling System - User's Guide*. National Weather Service SOO Science and Training Resource Center, May 2006. 1

[SLTO*15] SAHIN C., LAMY-THÉPAUT S., ORFORD T., VALIENTE C., RAOULT B.: ECMWF forecasts on the web - the ecCharts success story, 2015. 3

[SMM12] SEDLMAIR M., MEYER M., MUNZNER T.: Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012), 2431–2440. doi:10.1109/TVCG.2012.213. 4

[SSS*14] SACHA D., STOFFEL A., STOFFEL F., KWON B. C., ELLIS G., KEIM D.: Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (Dec 2014), 1604–1613. doi:10.1109/TVCG.2014.2346481. 4

[SZD*10] SANYAL J., ZHANG S., DYER J., MERCER A., AMBURN P., MOORHEAD R.: Noodles: A tool for visualization of numerical weather model ensemble uncertainty. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (Nov 2010), 1421–1430. doi:10.1109/TVCG.2010.181. 3

[TTW15] THOMAS TORSNEY-WEIR MICHAEL SEDLMAIR T. M.: Decision making in uncertainty visualization. URL: http://eprints.cs.univie.ac.at/4598/. 4

[Wil06] WILKS D.: *Statistical Methods in the Atmospheric Sciences*. International geophysics series. Academic Press, 2006. 1

[WLSL17] WANG J., LIU X., SHEN H.-W., LIN G.: Multi-Resolution Climate Ensemble Parameter Analysis with Nested Parallel Coordinates Plots. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 81–90. doi:10.1109/TVCG.2016.2598830. 2