AtmoSwing (v1.4): Analog Technique Model for Statistical weather forecastING

Pascal Horton[1,2]**, Michel Jaboyedoff**[1]**, and Charles Obled**[3]

[1]University of Lausanne, Lausanne, Switzerland
[2]Terranum, Lausanne, Switzerland
[3]Université de Grenoble-Alpes, LTHE, Grenoble, France

Correspondence to: Pascal Horton (pascal.horton@alumnil.unil.ch)

Abstract.

The Analogue Method allows forecasting local meteorological variables of interest (predictand), such as the daily precipitation, on the basis of synoptic variables (predictors). It relies on the fact that similar meteorological influences are likely to result in similar local effects. This statistical relationship is thus based on archives of observed data, and used in order to operationally forecast the coming days, or to evaluate future conditions under a changing climate.

AtmoSwing is an open source software that implements the Analogue Method in a very flexible way, so that different variants can be handled dynamically, by parameterization through XML files. It is made of 3 tools: the Forecaster to perform operational forecasts, the Viewer to display the results, and the Optimizer to establish the relationship between the predictand and predictors.

The Forecaster handles every required processing internally, such as predictors downloading and reading, grid interpolation, analogy sorting, without external scripts or file conversion. The processing of a forecast is extremely lightweight in terms of IT infrastructure; it can indeed run on almost any computer.

The Viewer displays the forecasts in an interactive GIS environment. It contains several layers of syntheses and details in order to provide a quick overview of the potential critical situations in the coming days, as well as the possibility for the user to go into the details of the forecasted predictand and criteria distributions. Several tips coming from the literature are provided in order to help interpreting the forecasts.

The Optimizer integrates the common semi-automatic sequential approach, as well as Monte-Carlo analyses, and a global optimization technique by means of Genetic algorithms.

1 Introduction

Reasoning by analogy is rather widespread in different domains of sciences or engineering. In hydrometeorology, it consists, given a certain situation, in looking into the past to retrieve situations that can be considered similar and therefore with consequences that may be expected similar to the situation at hand. By situation, one often thinks of the atmospheric situation over a rather wide spatial domain. By consequence, one often focuses on a local variable of interest, such as the occurrence of hail, storm, avalanches, or the amount of wind or precipitation over a given region, cumulated on a given time period.

As a general setting, the required ingredients are a couple of archives, one describing the situation through different variables, called predictors, and another one giving, for those situations, the value of the local variable of interest for engineering purpose, called predictand. Obviously, one understands that the longest and the richest the candidate archive of predictors, the more chance there will be to extract good analogues for a target situation in order to derive a conditional range of predictand values. As for the predictors, one can now hopefully rely on rich reanalysis products, commonly available over 40 years or more, and easily accessible.

Usually, the predictand values could be derived in a rather physical-dynamic approach by modeling the chain of processes linking the predictors to the predictand. The processes involved range from large scale dynamical states of the atmosphere down to very small scale microphysical processes, requiring models that are extremely complex, data-demanding, and time consuming. Conversely, given an appropriate set of predictors archives, enough situations analogous to a target one could be found so that reasonable values will be obtained for the predictand, at a reasonable coding and computing-

time costs. This is particularly true for a predictand much required in hydrometeorological applications, namely the precipitation amount over a given domain and time duration. Incidentally, the forecast will be proposed as a statistical distribution based on the values taken by the predictand in the set of analogues selected, unless one considers only the single best analogue, which may not prove the most efficient.

The purpose of this paper is not to argue in favour or against this approach, which has proved to be used efficiently either in operational forecasting or more recently in climate studies, and which is capable of challenging more classical approaches based on meteorological modeling. In operational forecasting, it has been used mainly by practitioners, notably hydropower companies or flood forecasting services, while in climate studies, it is more and more used to downscale the results of global climate model simulation runs. In operational, it should however not be considered as a substitute for numerical weather prediction models, but as a complement in order to get a fast and partly independent forecast that is known to be accurate several days in advance. It therefore enriches the analysis of a potential critical situation for flood forecasting, for example, and is very interesting in early warning.

However, how efficient and useful it may seem, this approach is not straightforward. Usually the type of predictand is imposed by some practical needs or constraints. Considering a precipitation variable, one has to decide upon the duration of accumulation, (e.g. from 6-hourly to daily or several day totals) according for example to the size or response time of the catchments to be managed. The availability and quality of long enough time series must also be considered. Moreover, the analogy is restricted to a selection of variables at a specific time and over optimized spatial domains, that are compared by means of a chosen criteria that can rank the candidate days from the archive with a distance quantification. The approach appears predictand-specific but also dependent on the archive of predictors.

The Analogue Method (AM) will first be presented along with various implementations (section 2), as the optimal version of the method may vary with the location of interest, and as all these can be processed with AtmoSwing. The software will then be presented (section 3), and its modules detailed, such as the Forecaster (section 4) and the Viewer (section 5) that are tools for operational use, and the Optimizer (section 6) that allows optimizing the method for a given predictand time series.

2   The Analogue Method

AtmoSwing does not rely on a single variant of the AM, but can implement different parameterizations. Thus, several examples of possible implementations are gathered here, as well as the main evolution steps of the AM. This latter summarizes the improvements that were brought, as well as dis-

continued approaches that were found less relevant (such as the use of PCA on the predictors).

2.1   Principles

The AM is based on the principle that two similar synoptic situations may produce similar local effects (Lorenz, 1956). The perfect analogy does not exist, but sufficiently similar situations leading to similar effects can be identified. Thus, two states of the atmosphere that are alike are called analogues (Lorenz, 1969). To be relevant, this analogy must be selected by optimizing the following elements:

– The analogy variable (the predictor) must contain synoptic information having indirect dependency with the local weather and the target predictand (often the precipitation).

– The spatial window is the domain on which the analogy variable is considered. The ideal size of this area is the one that maximizes useful information.

– The temporal window is the hour, or a set of different hours, in the day at which the analogy variables are considered.

– The preselection, which is often a period of the year in which similar situations are being sought. Thus, only situations with a similar distribution of solar energy are compared (Lorenz, 1969).

– The analogy criterion, needed to compare the variables on the chosen spatial and temporal windows, is a score used for ranking past situations according to their degree of similarity with the target situation.

– The ideal number of analogue situations is the best compromise in order to take into account local variability and to maximize useful synoptic information (Bontron, 2004).

Because of the chaotic nature of the atmosphere, two analogue situations quickly diverge over time (Lorenz, 1969). Thus, the AM has strong limitations regarding the temporal extrapolation (Bontron, 2004). Numerical models being more capable of simulating the dynamic evolution of the atmosphere, the temporal extrapolation of the synoptic variables is left to them. The search for analogy aims thus at connecting the forecasted synoptic situation with a local predictand (temperature, precipitation ...), which is more difficult to simulate by numerical models.

The AM is being used for many years at EDF (Électricité de France), to build precipitation and temperature scenarios (Desaint et al., 2008). It is also operational at the Compagnie Nationale du Rhône (CNR, France) and some flood forecasting services (SPC) in France and Switzerland (Marty, 2010; García Hernández et al., 2009; Horton et al., 2012).

2.2 Data

<mark>The AM relies on two types of data: predictors, that are atmospheric variables describing the state of the atmosphere at a synoptic scale, and the predictand, which is the local weather time series one wants to forecast.</mark>

Predictors are generally reanalysis datasets (as archive) and outputs of a global numerical weather prediction models for the target situation in operational forecasting. Reanalysis datasets, such as the NCEP/NCAR reanalysis I (6-hourly, 17 pressure levels at a resolution of 2.5°, see Kalnay et al., 1996), the NCEP/DOE reanalysis II (Kanamitsu et al., 2002), ERA-40 (Uppala et al., 2005) are often used and can be input in AtmoSwing. Other observed predictors can also be used, such as Sea Surface Temperature (SST, Reynolds et al., 2007). The predictors describing the target situation in operational forecasting can, for example, result from GFS (Global Forecast System, Kanamitsu et al., 1991; Kanamitsu, 1989), which is operated by NCEP and NOAA.

Bontron (2004) proposed that the minimum length of the archive should be 30 years to forecast usual situations, and 40 years or more for intense events. He also highlighted the dependency between the optimal number of analogues and the archive size: the shorter the archive, the lower the number of analogues. By varying the resolution of the predictors grids, he concluded that the performance increase is not significant for the analogy of atmospheric circulation below a resolution of 5°. However, it seems more important for the second level of analogy on moisture, which are more local variables. Ben Daoud (2010) assessed the European reanalysis ERA-40 (Uppala et al., 2005) with a resolution of 1.125°. He identified a slight performance improvement, but not substantial enough to really justify an archive change (Ben Daoud et al., 2008).

The predictand is the local meteorological variable of interest. It can be 6-hour data, or more often daily time series, but has to be consistent with the time step of the predictors. The most used predictand is the daily precipitation, usually averaged over subregions in order to smooth local effects (Obled et al., 2002; Marty et al., 2012). These time series are frequently normalized by the precipitation value for a return period of 10 years (Djerboua, 2001). This normalization allows for an easier comparison between subregions subject to different precipitation regimes, and thus to better identify the most important contributions. Most users also apply a root square to this last ratio in order to reduce the weight of the extreme predictand values, but we generally don't, in order to give extremes a certain weighting.

2.3 Present structure

Although multiple variations of the AM exist, they generally share the same basics and the same structure. For example, let's say we want here to forecast the daily precipitation (the predictand) for a target day (at a defined lead time) and have long archives of both the predictand and predictor variables. The generic structure, illustrated for precipitation forecasting, is the following:

1. Preselection: to cope with seasonal effects, candidate dates are extracted from the archive within a period of four months centered around the target date, for every year of the archive. Alternatively, the candidate dates can be selected based on similar air temperature (Ben Daoud, 2010).

2. First level of analogy: N_1 dates are subsampled within all the candidates provided by the preselection, by means of an analogy ranking. The first level of analogy is always based on the atmospheric circulation when it comes to precipitation forecasting. The similarity of the atmospheric circulation of the target date with the candidate situations is thus assessed on the geopotential by means of the S1 criteria (Eq.1, Teweles and Wobus, 1954; Drosdowsky and Zhang, 2003, see also section 2.5), which is a comparison of gradients, over a certain spatial gridded window.

$$S1 = 100 \frac{\sum_i |\Delta \hat{z}_i - \Delta z_i|}{\sum_i max\{|\Delta \hat{z}_i|, |\Delta z_i|\}} \tag{1}$$

where $\Delta \hat{z}_i$ is the forecast geopotential height difference between the ith pair of adjacent points from the grid of the target situation, and Δz_i is the corresponding observed geopotential height difference in the candidate situation. The differences are processed separately in both directions. The smaller the S1 values, the more similar the pressure fields.

The N_1 dates with the lowest values of S1 are considered as analogues, in terms of the general circulation, to the target day. The number of analogues, N_1, is a parameter to calibrate.

3. Subsequent level(s) of analogy: beyond the similarity of airflows, one may look for similarities in other variables of interest . Therefore, the N_1 analogues are subsampled once again on the basis of another variable in order to obtain a lower number of analogue dates, N_2. The S1 criterion is usually not relevant for other predictors that the atmospheric circulation. Other classic criteria represent absolute distances: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), the latest being most often used. These are more relevant to compare moisture variables for example.

This process can be repeated, by subsampling a decreasing number of analogues, N_i, according to a ranking processed on various meteorological variables.

4. Probabilistic forecast: then, the (e.g. daily) observed precipitation amount (or another predictand of interest) of the N_i resulting dates provide the empirical conditional distribution considered as the probabilistic forecast for the target day. The empirical frequencies are processed for every value of the predictand, after classification, based on the Gringorten parameters (for a Gumbel or exponential law) and a probabilistic model can be fitted (e.g. Gamma function, or else).

2.4 Proposed nomenclature

Variants of the AM parameterization are numerous and it is not always easy to reference them in a short and descriptive way. We thus propose a basic nomenclature (Figure 1) in order to express the structure into a simple identifier. This one cannot describe all the parameters of the AM, but quickly illustrate the structure of the implementation. This is particularly useful when working with a global optimization method, where nothing is fixed but the very structure of the AM.
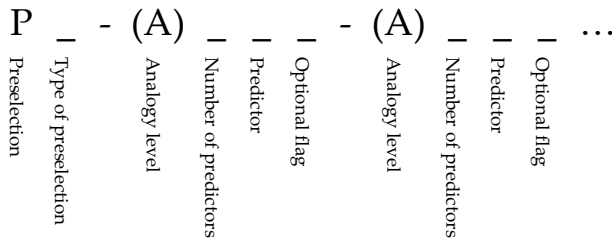

Figure 1. Proposed structure for naming the parameterizations of the AM.

The naming contains different blocs (separated by an hyphen) for the various levels of analogy. It starts with the specification of the preselection (P; can be omitted when comparing AMs with the same preselection approaches), which can be nowadays of 2 types:

- C: calendar period (± 60 days around the target date)
- T: based on air temperature (Ben Daoud, 2010)

Then, the following levels of analogy are listed. They may start with an optional A (for analogy) when other methods are mixed into the parameterization (see eg. Chardon et al., 2014). For every level of analogy, the number of variables used (combination of atmospheric levels and time of observation) is first provided, and then the short name of the variable is given (according eg to ECMWF conventions; in upper case), for example:

- Z : geopotential (circulation)
- MI : moisture index (TPW * RH)
- MF : moisture flux (V * TPW * RH)
- TPW : total precipitable water

- V : wind velocity
- W : vertical velocity

In order to keep the identifier simple, no value of atmospheric level neither time of observation is specified. Moreover, the analogy criterion is not specified and is supposed to be S1 for Z and RMSE for the other variables. If anything changes from these conventions, it can be noted as a flag. The flag (lower case) can also give other information, such as the optimization method:

- sc : sequential calibration (can be omitted as considered as default, see section 6.3)
- go (or just "o") : global optimization (by means of genetic algorithms for example)

This nomenclature can be adapted to specific needs, or simplified for a better readability (eg. by removing the specification of the preselection). Examples are provided in the following section.

2.5 Evolution of the method for precipitation forecasting

The most common usage of the AM is for precipitation forecasting. Thus, the evolution presented hereafter focuses on this predictand. However, the AM, or an equivalent, was also used for short to medium term forecasting of daily temperatures (Radinovic, 1975; Woodcock, 1980; Kruizinga and Murphy, 1983), wind (Gordon, 1987), snow avalanches (Obled and Good, 1980; Bolognesi, 1993), insolation (Bois et al., 1981), and the trajectory of tropical cyclones (Keenan and Woodcock, 1981; Sievers et al., 2000; Fraedrich et al., 2003). Guilbaud (1997) performed a literature review about the use of the AM in long-term forecasting and identified operational applications for monthly forecasts in many countries, including Canada (Shabbar and Knox, 1986), Hungary (Toth, 1989), the Netherlands (Nap et al., 1981), and England (Murray, 1974), as well as seasonal forecasts: Barnett and Preisendorfer (1978), Bergen and Harnack (1982) and Livezey and Barnston (1988).

The evolution of the AM for precipitation forecasting led to several improvements. However, some of these are specific for a certain region and are not relevant for others, and some perform better depending on the lead time. Thereby, one cannot provide a single best parameterization for the AM, neither rank them from worst to best. Some examples of parameterizations of the most recent AM variant will thus be provided, in the chronological order they appeared.

Lorenz (1969) introduced the concept and established the basis of the AM. He considered looking for analogues on the geopotential field (at 200, 500 and 850 hPa) with a corrected Euclidean distance averaged on the 3 levels. In addition, the analogue situations were to belong to the same period of the year to be comparable in terms of distribution of solar energy.

The use of the AM for operational forecasting of daily precipitation originates in the work of Duband (1970, 1974,

1981). It was then designed for operational forecasting at EDF (Électricité de France) in order to better manage water resources and flood risks. The geopotential height at 700 hPa and 00 h was used for its higher stability than the sea level pressure (SLP) field, and its sensitivity to the atmospheric disturbances (Guilbaud, 1997). However, the SLP at 06 h was also introduced for determining the precipitation intensity, as well as the temperature at 700 hPa as indicator of the thermal state of the air mass (Duband, 1974). The data were then based on 37 radiosounding over Europe and condensed by a principal component analysis. The first 25 analogues on the geopotential at 700 hPa were considered, and multiple local regressions between precipitation, SLP, and the temperature at 700 hPa were established on these analogues before being applied to the target situation (Guilbaud, 1997). The archives were split by season to consider comparable situations in terms of distribution of solar energy (Lorenz, 1969). This initially rigid division into seasons was then transformed into a moving selection of more or less two months around the target date. The forecast was then performed only on the basis of radiosonde observations and was temporally extrapolated to the two following days. After a few years, this statistical temporal extrapolation was abandoned in favor of a statistical adaptation using synoptic predictors taken from a numerical model, allowing a forecast over the following 4 days.

Thereafter, the geopotential was considered at 1000 hPa and 700 hPa and the analogues selection was done in two stages: first according to an Euclidean distance in the space of the principal components of the geopotential field at 700 hPa, and then according to a correlation criterion (on the geopotential field at 700 and 1000 hPa, as well as on the thickness between these two levels) in order to remove days close in distance but too dissimilar in pattern (Guilbaud, 1997). The local regression approach to determine precipitation was eventually abandoned in favor of the direct probabilistic forecasting (synthesized by the percentiles 20, 60 and 90 %), as it is performed today by interpolating linearly over the cumulative empirical distribution of daily precipitation measured at the analogue days (Guilbaud, 1997; Guilbaud and Obled, 1998).

Wilson and Yacowar (1980) sought the 20 best analogues on the geopotential field at 500 and 1000 hPa using the Teweles-Wobus (S1) criterion (Eq. 1), which allows for a comparison of the gradients and thus an analogy of the atmospheric circulation. This work has demonstrated the best performance of the AM compared to classical regression based approaches. Woodcock (1980) used the same criterion on the SLP fields to select 50 analogues in order to forecast maximum daily temperatures. After comparing S1 to other criteria from the literature, or combinations of these, Guilbaud and Obled (1998) confirmed that S1 is the most efficient. They also pointed out the benefit of using 2 atmospheric levels, 1000 hPa (Z1000) and 700 hPa (Z700) or 500 hPa (Z500), instead of one (also instead of the thickness), and to also

consider 2 temporal windows, 00 h and 24 h (Obled et al., 2002).

Guilbaud (1997) stopped using PCA in order to work directly with the raw data interpolated on grids, which resulted in an improvement, especially when considering the S1 criterion instead of the Euclidean distance. The number of analogues considered was 50. When using multiple predictors, they combined the criteria values calculated on each atmospheric level and each temporal window using an arithmetic mean.

At the dawn of the new millennium, Bontron (2004) could benefit from a real breakthrough, the newly available NCEP/NCAR reanalysis data. This dataset being more homogeneous, much longer, and containing more variables than what was previously used, he could systematically assess a large selection of variables in the AM combined with different criteria. He thus confirmed the interest of the geopotential height along with the S1 criterion for the first level of analogy. He has also shown that the choice of the temporal window (time of observation) has a higher importance than the atmospheric level for the performance of the AM for daily precipitations. By assessing multiple combinations of atmospheric levels and temporal windows, he concluded that the coupled geopotential heights at 1000 hPa (Z1000) & 12 h and 500 hPa (Z500) & 24 h provided the best performance for the studied region (Table 1). The analogy on the atmospheric circulation proposed by Bontron (2004) is still, at the time of writing, often used operationally, and considered as a reference for benchmarking new parameterizations (Horton et al., 2012; Ben Daoud et al., 2015).

Table 1. Method of type PC-2Z proposed by Bontron (2004). The first column is the level of analogy (0 for preselection), then comes the meteorological variable and its hour of observation (temporal window). The criteria used for the current level of analogy is then provided, as well as the number of analogues.

Level	Variable	Hour	Criteria	Nb
0	±60 days around the target date			
1	Z1000	12 h	S1	50
	Z500	24 h		

Mandon (1985) tried to introduce a second level of analogy and assessed different variables, such as wind, moisture, air temperature, surface temperature of the Mediterranean, the product of wind and the 700 hPa moisture, as well as a second barometric criteria differentiated according to the season. Vallée (1986) continued to work on the second level of analogy and considered the wind at 700 hPa & 12 h in order to improve heavy precipitation events for a region in France susceptible to southerly flows.

Bontron (2004) also assessed several variables for the second level of analogy and identified the moisture variables (precipitable water and relative humidity) and vertical veloc-

ity to be the most relevant. Finally, he found that a moisture index made of the product of relative humidity at 850 hPa (RH850) and total precipitable water (TPW) gave the best performance (Table 2). This index does not represent an actual physical quantity, but expresses the water content of the air column and its proximity to saturation.

Table 2. Method of type PC-2Z-2MI proposed by Bontron (2004). Same conventions as Table 1

Level	Variable	Hour	Criteria	Nb
0	±60 days around the target date			
1	Z1000 Z500	12 h 24 h	S1	70
2	TPW * RH850	12, 24 h	RMSE	30

He noted the relevance of defining different spatial windows for each region, rather than a single window for the entire territory. These windows are more able to represent the specific atmospheric influences for each region. In order to assess the forecast, he introduced the Continuous Ranked Probability Score (CRPS, Eq. 2, Brown, 1974), which discards the issue of determining classes as with the Ranked Probability Score (Epstein, 1969) previously in use.

Gibergans-Báguena and Llasat (2007) implemented the AM for daily precipitation forecasts in Catalonia. They considered the geopotential at 700 and 1000 hPa, and the thickness between these two levels, to search for analogy of circulation, and then applied a second selection on local data from radiosonde: stability indexes, precipitable water, temperature, etc.

Marty (2010) tested other temporal windows for intraday application on the basis of a more comprehensive reanalysis dataset. He first proposed to change the hours of observation for both levels of analogy (both for 06 h and 18 h), and selected the 925 hPa level for the moisture analogy (Table 3). The selected hours and atmospheric levels were not available before in the reanalysis dataset.

Table 3. Method of type PC-2Z-2MI proposed by Marty (2010). Same conventions as Table 1

Level	Variable	Hour	Criteria	Nb
0	±60 days around the target date			
1	Z1000 Z500	06 h 18 h	S1	75
2	TPW * RH925	06, 18 h	RMSE	30

Then, Marty (2010) changed the moisture index into the moisture flux by adding the wind velocity (V700 or V925) component in the multiplication. He considered this flux at 700 hPa or 925 hPa (Table 4).

Table 4. Method of type PC-2Z-2MF proposed by Marty (2010). Same conventions as Table 1

Level	Variable	Hour	Criteria	Nb
0	±60 days around the target date			
1	Z1000 Z500	06 h 18 h	S1	60
2	V700 * TPW * RH700 [†]	06, 18 h	RMSE	25

† or V925 * TPW * RH925 as an alternative

Ben Daoud (2010) applied the AM in the context of large floodplains (Saône, Seine). He evaluated several variables usually used in weather forecasting, amongst them the air temperature and the vertical velocity that presented a potential interest. The temperature (at the nearest grid point, on the 925 hPa & 36 h and 600 hPa & 12 h levels) could replace the rigid calendar preselection of ± 60 days around the target date by a more dynamic screening of similar situations in terms of air masses. The seasonal effect is indeed also present in the temperature data. The number of preselected dates is equivalent to the number of days one would have chosen with the calendar approach, and thus depends on the archive size: $N_0 = 60 \cdot 2 \cdot n_a$ where n_a is the number of years in the calibration period. Ben Daoud (2010) also reconsidered the parameters of the moisture index and ended up with both 925 hPa and 700 hPa levels, at 12 h and 24 h, or at every time step (6 h) between 06 h and 30 h (Table 5).

Table 5. Method of type P2T-2Z-4MI (or P2T-2Z-10MI for the alternative version), by Ben Daoud (2010). Same conventions as Table 1

Level	Variable	Hour	Criteria	Nb
0	T925 T600	36 h 12 h	RMSE	N_0
1	Z1000 Z500	12 h 24 h	S1	70
2	TPW * RH925 TPW * RH700	12, 24 h [†] 12, 24 h [†]	RMSE	25

† or 06-30 h as an alternative

Subsequently, Ben Daoud (2010) added an additional level of analogy, between the circulation and the moisture analogy (Table 6), based on vertical velocity at 850 hPa (W850). This AM was developed for large floodplains in France, but was found to be not relevant for an Alpine environment where topography is mostly responsible for uplift of air masses (Horton et al., 2012).

Bliefernicht (2010) proposed to add a weighting to each grid points of the predictor domain to overcome a compar-

Table 6. Method of type P2T-2Z-4W-4MI, by Ben Daoud (2010). Same conventions as Table 1

Level	Variable	Hour	Criteria	Nb
0	T925 (1 point) T600 (1 point)	36 h 12 h	RMSE	N_0
1	Z1000 Z500	12 h 24 h	S1	170
2	W850	06-24 h	RMSE	70
3	TPW * RH925 TPW * RH700	12, 24 h 12, 24 h	RMSE	25

Table 8. Method of type PC-4Z-2MI, by Horton et al. (2012). Same conventions as Table 1

Level	Variable	Hour	Criteria	Nb
0	±60 days around the target date			
1	Z1000 Z850 Z700 Z500	12 h 24 h 12 h 24 h	S1	N_1
2	V700 * TPW * RH700	12, 24 h	RMSE	N_2

ison on a fixed-size window, and to give more importance to regions with relevant atmospheric circulation information. However, he noticed that there is a risk of overparametrization as the performance was not stable between the calibration and the validation periods, and a simplification of the weighting fields may be necessary. He also compared the performance of the AM with a forecasting method based on a weather types approach, and concluded that the AM was more efficient.

Horton et al. (2012) identified that the analogy on the atmospheric circulation may not be comprehensive, and that a significant gain was obtained by considering 4 atmospheric levels (1000, 850, 700, and 500 hPa at 12 or 24 h) instead of 2 (Table 7). The number of analogues (N_1) was not provided globally, as it was calibrated for every subregion. Considering an AM only based on the atmospheric circulation has an interest for operational forecasting, since moisture variables are less well forecasted for a lead time of a few days. An equivalent structure was also optimized by means of generic algorithms (PC-4Zgo), which is the topic of another paper (Horton et al., 2016b).

Table 7. Method of type PC-4Z proposed by Horton et al. (2012). Same conventions as Table 1

Level	Variable	Hour	Criteria	Nb
0	±60 days around the target date			
1	Z1000 Z850 Z700 Z500	12 h 24 h 12 h 24 h	S1	N_1

Horton et al. (2012) also integrated an analogy on the moisture flux, after the circulation analogy on the 4 atmospheric levels (Table 8). The number of analogues (N_1, N_2) were also calibrated for every subregion. Similarly, an equivalent structure was optimized (PC-4Zgo-2MIgo, Horton et al., 2016b).

The optimal predictors vary from one region to another, along with the leading atmospheric processes. One will thus never get a unique parameterization of the AM valid for any place on earth, but the method needs to be adapted to the local conditions, available data, and to the size of the catchment of interest. Thus, there will always be local adaptations to be made for use in a new region.

2.6 Discussion on the method

A version of the AM was evaluated during the project STARDEX (*STAtistical and Regional dynamical Downscaling of EXtremes for European regions*, see Goodess, 2003; STARDEX, 2005). One of the project goals was to compare various downscaling methods for the determination of weather extremes, and the AM was selected among the most interesting from various techniques (Maheras et al., 2005; Schmidli et al., 2007). Its use as adaptation method of climate models is also the subject of a growing number of studies (Zorita and Storch, 1999; Wetterhall, 2005; Wetterhall et al., 2007; Matulla et al., 2007; Chardon et al., 2014; Dayon et al., 2015).

Hamill and Whitaker (2006) used an analogy based approach on the GFS reforecasts in order to correct systematic errors in the ensemble forecasts of temperature and precipitation. Indeed, the statistical approach pointed at some bias in estimates of the numerical model, which may have been adjusted by taking into account the intrinsic local climatology from the AM. Moreover, the under-dispersion of the ensemble forecast from the numerical model has been corrected using analogues (Hamill and Whitaker, 2006). Correction of ensemble forecast under-dispersion by means of the AM is also used operationally at EDF.

Bliefernicht (2010) observed that the performance of the AM is higher for winter than for summer. The relationship between synoptic predictors and local rainfall is lower in summer, due to convective precipitations that present a higher spatial variability and that depend on other parameters. Variables describing the synoptic circulation are indeed not able to predict the location of thunderstorm cells. This was also observed by Ben Daoud (2010), who set up a specific model for the summer months (June 15 to September 15).

One of the AM limitations is the need for a substantial archive of the predictand variable, for example measured precipitations, that remain to be provided by the user. Without data on several decades, the AM is not applicable. Conversely, long predictor archives are also required, but this problem is now solved with reanalysis data, which may not be perfect in terms of homogeneity, but can be considered of sufficient quality. Moreover, reanalysis data are available all around the world, which represent a great potential for the AM, and some are now extending to the whole 20th century. Another issue, still about data, concerns the operational forecasting: predictors describing forecasted target situations and the ones from the archive are not fully homogeneous, as they usually don't result from the same model and resolution. It is therefore necessary to use robust variables that depend from the numerical model characteristics at a minimum extent.

Another limitation is the fact that extreme events may be under-represented in the considered sample of analogue dates. Indeed, in a limited weather archive, events with high return periods are not so numerous. Their number is certainly lower than the standard number of analogues considered, which can introduce a bias in the forecast. There are however techniques to correct this bias (see Marty, 2010).

It is also legitimate to raise the question of the relevance of an approach based on archives of past situations in a context of climate change. The first potential issue is a possible change in atmospheric circulation. Philipp et al. (2007) observed certain trends in the synoptic atmospheric circulation over Europe, but with moderate frequencies. Moreover, the basic laws governing the atmosphere behavior will not be transformed (Hewitson and Crane, 1996). The assumption is that a large part of local climate change will result from changes in intensity, frequency and persistence of synoptic variables, but with other characteristics substantially similar to the present situation (Hewitson and Crane, 1996). Thus, if the archive of weather situations is long enough, it is reasonable to assume that a large part of future situations is already represented, even those whose frequency will change under different climatic conditions (Wetterhall, 2005). Dayon et al. (2015) has proven these assumptions to be accurate by showing the transferability in the future climate of the AM. They could demonstrate that the AM was able to predict the same trends in precipitations changes than Regional Climate Models. In addition, for operational use, climate change is relatively slow, and through continuous updating of the archive, the AM will incorporate new information progressively with no strong discontinuity. However, this may plea in favor of the approach proposed by Ben Daoud (2010) for defining the potential candidates to analogy: instead of considering a fixed period of two months before and after the target dates, one would prefer to select candidates that have approximately the same range of temperature.

3   Software basics

AtmoSwing is made of 3 main modules (Figure 2) that are standalone, but do share a common code basis: the Forecaster, to process the operational forecasting, the Viewer, to display the forecast in a GIS environment, and the Optimizer, to calibrate/optimize the statistical relationship defining the analogy for a given predictand time series. Separating the Forecaster and the Viewer allows to automate the forecast on a server and to quickly display the results locally. The code is written in object oriented C++ and relies on the wxWidgets (Smart et al., 2006) library to provide a cross-platform native experience to users. The software can thus be compiled on MS Windows, Linux / Unix, Mac (OS X, iOS), both in 32-bit and 64-bit.

Figure 2. The AtmoSwing logo and its 3 modules: the Forecaster, the Viewer and the Optimizer.

The source code in under version control (with Mercurial), and is open source (on Bitbucket, www.atmoswing.org). Developments have been made with a test driven development (TDD) approach. A collection of 500+ tests are frequently evaluated and completed during development, in order to verify that code changes do not introduce regression. Every analogy criterion, prediction score, searching and sorting functions, data manipulation, and so on, is tested. The integration tests specific to the AM rely on the results of another analogue sorting software developed in Grenoble. They result from a comparison work done with the help of R. Marty in order to ensure that the results of AtmoSwing are exactly equivalent to their model, given the same parameters and data. These tests are part of the checks carried out regularly, so that these results are always reproducible.

3.1   Modular approach and implementation

An AtmoSwing great strength is that it is designed to process analogy forecasting in a modular fashion. The structure of the AM (number of analogy levels, number of predictors) is built dynamically (Figure 3), and nothing is fixed a priori. The software then performs successively as many analogy levels as the user specified, with all the predictors he wants. Each level of analogy processed results in an object containing target dates, analogue dates, values of the analogy criteria, and other data. This object can be saved in a NetCDF file and/or can be injected into a new analogy level. The whole structure of the AM is defined in an XML file generated by

the user. Even the time step of the method (6 or 24 hours for example) is a dynamic parameter.
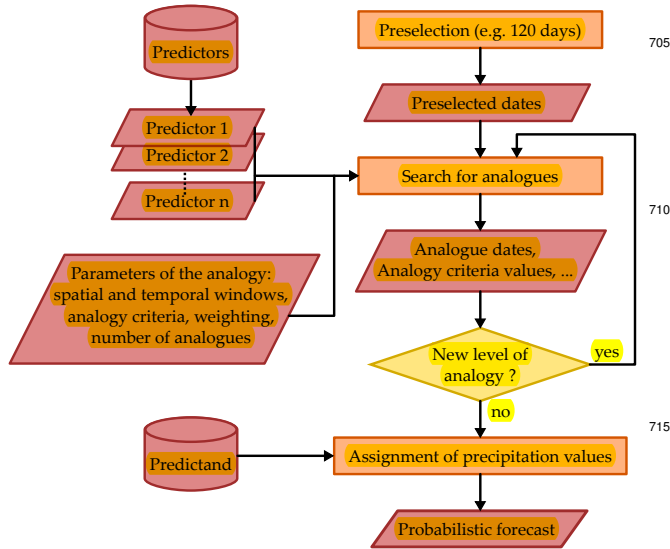

Figure 3. Simplified organizational chart of the AM implementation in AtmoSwing.

Each implementation of the AM (see section 2.5) may enter this scheme, even if it consists of preprocessed variables (e.g. moisture index). Various preprocessing functions are implemented, as the calculation of the moisture index or flux, multiplication operations or calculation of the gradients. The user can specify the preprocessing method and the predictors to use, dynamically, in the XML file.

3.2 Performance

Although processing an analogue forecast for a given target date is fast, when optimizing the AM, forecasts are reprocessed over long periods of several decades, which may become very time consuming. Thus, great effort has been made in order to reduce the processing time to a minimum. Waste of time were identified with profiling tools and cut down at different levels. The software also use the linear algebra library Eigen 3 (Guennebaud et al., 2010) for calculations on vectors and matrices, which resulted in time saving. Obviously, multi-threading is also implemented.

First, every identified redundancy has been removed. Then, when looking for a certain date or data, the search looks first in the proximity where it is likely to be found instead of exploring an entire array. Similar data are not loaded twice, but shared pointers are used. Many other improvements allowed saving time, for example the use of the quicksort method (Hoare, 1962) to sort the date vectors according to the values of analogy criteria. Different implementations were tested (many remain in the code as alternatives) in order to select the most efficient: for example, when storing analogue dates according to their criteria value, it is faster to

insert them in a fixed-size array rather than storing them all and sorting the array subsequently.

The major performance gains were obtained by:

– reduction of the array sizes used in the calculations,
– use of pointers (and therefore the decrease of data copies),
– better search for dates in temporal vectors starting from the previous index (reduction of the range of the array to explore),
– better management of the analogue vectors,
– preprocessed gradients on geopotential fields when using the S1 criteria,
– and finally the implementation of parallel calculations (multi-threading).

4 AtmoSwing Forecaster

The Forecaster module allows to process operational forecasts, defined by an XML file. A list of parameterizations is processed successively. The software can be compiled with a GUI (graphical user interface, Figure 4), or without in order to be used on a headless server through a command line interface. Processing a forecast requires very low computing capabilities and can be done on a low-end computer.
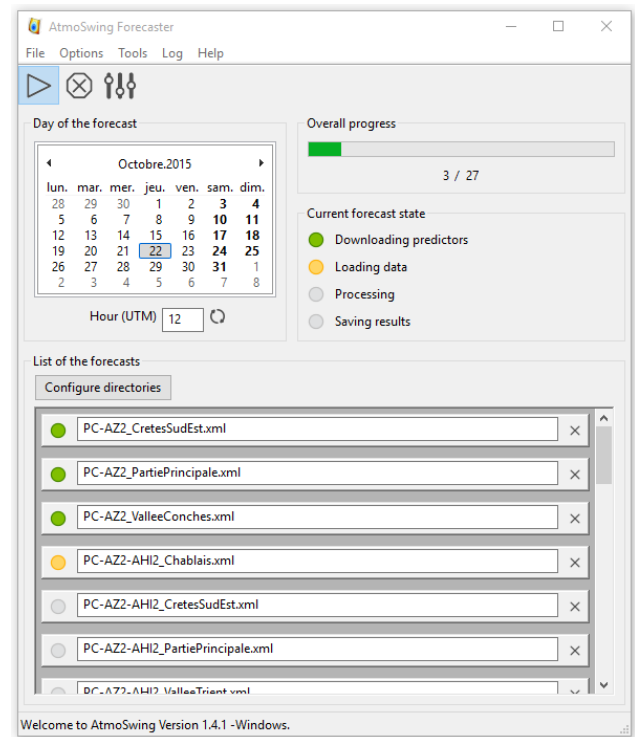

Figure 4. Graphical user interface of the Forecaster module. A list of different parameterizations is being processed.

The software first downloads the predictor describing the target situation, such as 1° GFS outputs (Global Forecast

System, Kanamitsu et al., 1991; Kanamitsu, 1989, see section 2.2). It then interpolates the gridded data to match the resolution of the archive (for example 2.5° in the case of the NCEP/NCAP reanalysis I). The analogues dates are next sought according to the selected parameterization, and the predictand data are associated with the corresponding dates. The results are finally saved in auto-describing NetCDF files. If required, a synthetic XML is generated for an easier integration on a web platform for example. Every step of the forecast, from the predictor downloading to the final results, is done in the software (and controlled through configuration), without use of external scripts (e.g. for data conversion).

When using a given parameterization of the AM, it has previously been optimized in a perfect prognosis context (i.e. by optimizing on the reanalysis dataset; however, this does not take into account the uncertainty of the target situation when taken from operational forecasts of a NWP model, Klein, 1963). In operational forecasting, one should take into account the uncertainty of the predictor increasing with the lead time. A solution, presented in section 6.6, consists in increasing the number of analogue situations to keep with the lead time. This technique is available in AtmoSwing, as the number of analogues can be specified for every lead time, making such approach easily accessible.

The command-line interface allows to start the process for the present date, a selected past date, or the last x days (as long as the predictors are available on the provider server). When there is nothing to process (no new predictor data available), the execution just stops, leading to no loss of computing resources. The recommended use is thus to set up a cron task on a Linux server in order to trigger the forecast every 30 minutes. When using GFS outputs, this would provide 4 forecasts a day with a reduced delay between GFS outputs availability and the analogue forecast.

A user interface allows the creation of the precipitation database. Its generation consists in extracting the time series from text files in order to process them and save a database in the NetCDF format. During the process, Gumbel adjustments are automatically calculated in order to determine the values corresponding to different return periods. The time series are normalized by a selected return period (default 10 years) and their square root is eventually processed. The final database file contains both the raw and the normalized series, as well as characteristics of the gauging stations and some metadata.

5   AtmoSwing Viewer

5.1   The graphical user interface

The Viewer module allows displaying the resulting forecast files in an interactive GIS environment (Figure 5). During the forecast, different parametrization variants of the same type of AM may be applied to a region; the reason being that some parameterizations may be specific for a subregion, when its meteorological influences differ, which is often the case even in relatively small catchments (Horton et al., 2012). The Viewer automatically gathers the similar AMs and provides a composite view of the optimal forecasts per subregion. The user can however choose to display a specific parameterization for the whole region, which is more consistent in terms of homogeneity of the analogue dates.

The software provides several levels of synthesis of the forecasts. When it is opened, it automatically loads the most recent forecasts, eliminating the need to manually open files. It first offers a quick overview of possible alerts by means of color codes on the lead time switcher (upper right in the graphical user interface (GUI), see Figure 5), which synthesizes the worst case scenario, or in the alarms panel (on the left part of the GUI). The alarms panel offers a synthesis of the higher forecasted values over the region, for the different AMs and the different lead times. By default, the colors are expressed relatively to 10-year return period values, for the 90th percentile (can be changed in the preferences). This higher level of synthesis allows to quickly identify potential critical situations in the days ahead.

Then, the user can explore the forecasts more in details, starting from the map provided by the GUI (Figure 5). The map display the forecast of the selected parameterization (upper left) and the selected lead time (upper right). The displayed AM is by default a composite of the parameterization variants for the region, but the user can choose a specific parameterization for the whole region by opening the tree view and selecting a child element. We also allowed a display of all lead times on a single map by means of a symbolic representation on a circular band with a box for every lead time (Figure 6). The number of boxes is dynamically adjusted to the number of lead times. This representation offers a global spatiotemporal display for a chosen parameterization.

Color scales in the map are easily controlled by choosing the predictand ratio (raw value or ratio to different return periods) and the quantile of the distribution (on the left part of the GUI). Representation relatively to a return time better expresses the meaning of a certain precipitation amount for a given region, as climatology can be drastically different between two locations in a mountainous area, even if they are close to each other. All information relative to a rain gauge station (or catchments), such as its location, its name, or the values of different return periods, are stored in the forecast files, in order to be displayable for end users who do not have the predictand database.

By clicking on a station (forecast points) on the map (or by selection in a dropdown list on the left), a new window appears with a plot of the forecasted time series (Figure 7). By default, the plot contains the usual 3 considered percentiles (90th, 60th, and 20th), along with the 10 best analogues with a color code from yellow (10th) to red (1st). The 10 year return period is also displayed in order to put the forecast in perspective. The user can choose to hide any data or to
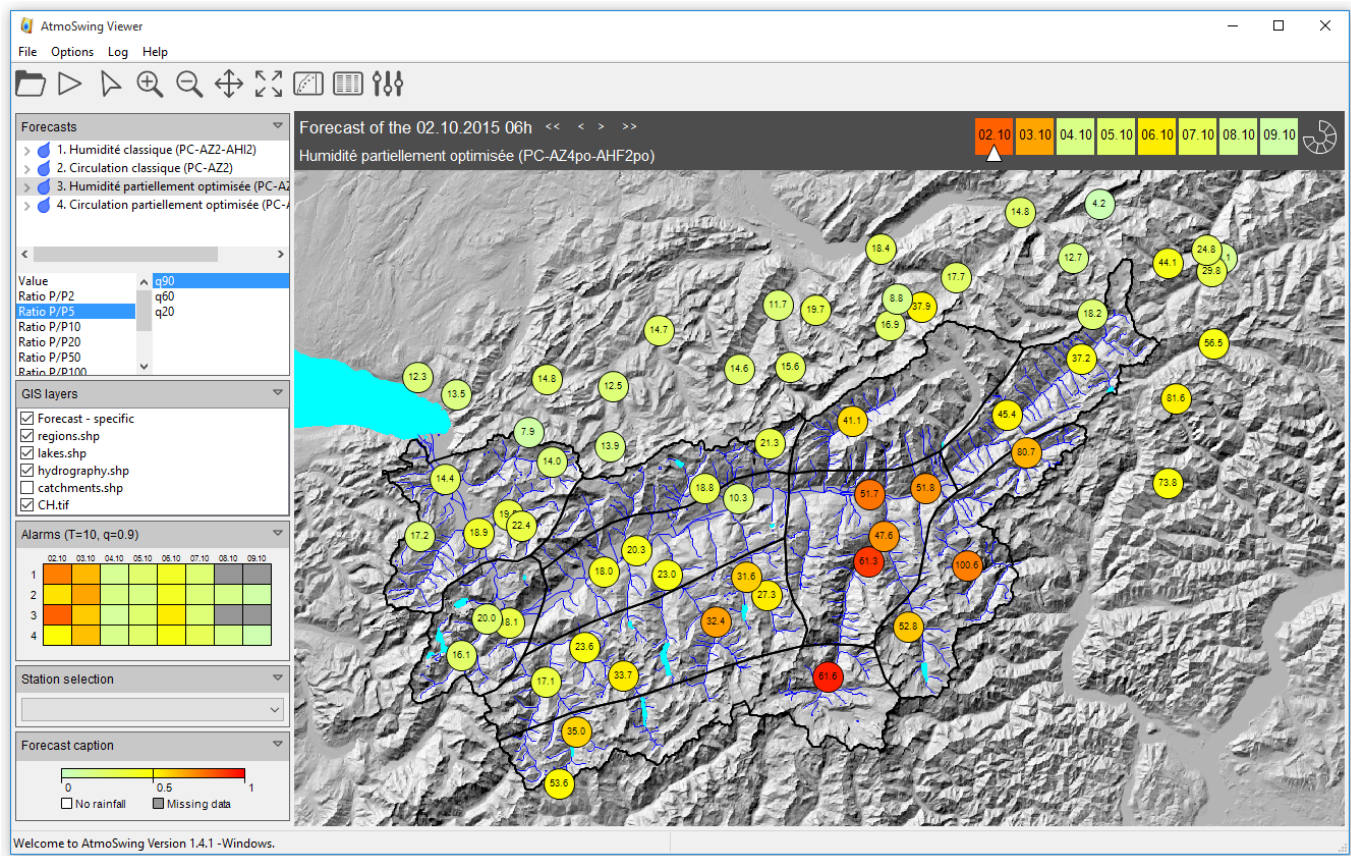
Figure 5. Graphical user interface of the Viewer module (Elevation data from The Shuttle Radar Topography Mission (SRTM), and hydrological network from SwissTopo).
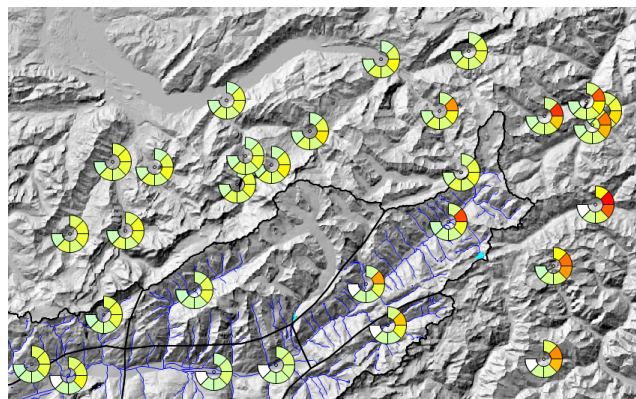
Figure 6. Visualization of multiple lead times on the map (Elevation data from the SRTM, and hydrological network from SwissTopo).
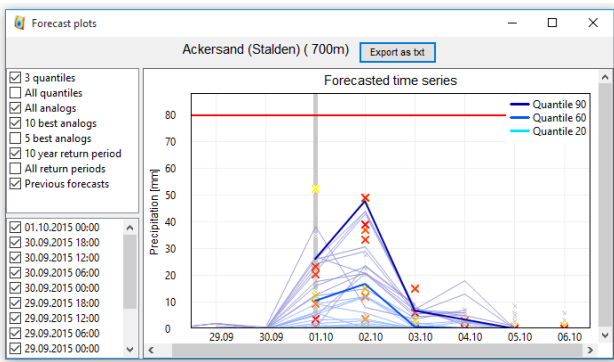
Figure 7. Visualization of the forecasted time series.

display supplementary information (all analogues, all 10th percentiles, or all return periods) in the left panel. Traces of previous days forecasts are also automatically loaded and displayed in order to inform on the stability of the forecasts.

The user can even go into further details and display the predictand cumulative distribution for a given lead time (Fig-ure 8). This is useful when working on extreme precipitation, as comparing the distribution of all analogues versus the 10 best analogues provides information in order to interpret the forecast (trend to under/overestimation). Distribution of the analogy criteria (not shown) can also be displayed in order to identify eventual discontinuities in the criteria values.
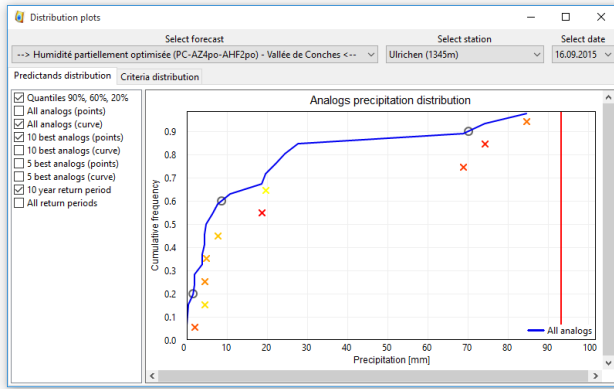
Figure 8. Visualization of the forecasted precipitation distribution for a given lead time.

Finally, one can see the details of the analogue dates, along with the predictand and the criteria values in an interactive spreadsheet (not shown).

AtmoSwing Viewer relies on workspaces to specify the path to the forecast directories and the GIS layers. Many GIS formats are supported thanks to the GDAL and VroomGIS libraries. One can have as many layers as desired, and can control their display properties. It is thus easy to switch from a forecast for a region to another by opening a workspace file (XML format).

5.2 Interpreting a precipitation forecast

Djerboua (2001) applied the AM developed by Guilbaud (1997) in operational forecasting as part of the MAP project (*Mesoscale Alpine Programme*, see Binder and Schär, 1996) during the special observation period (Bougeault et al., 2001). He noted that the 60th percentile was better to forecast precipitation amounts for common situations, but for strong to extreme events, the 90th percentile turned out to be a better indicator. It is therefore necessary to pay particular attention to the 90th percentile which, when reaching high values, may be indicative of a situation related to extreme precipitations, due the presence of several analogue dates with significant observed precipitations amounts in the distribution (Djerboua, 2001). Bontron (2004) made the same observation: he found the 60th percentile to be best for occurrence forecasting, but the 90th percentile to be the most informative in order to forecast medium to high precipitation amounts ($P > P10/10$, one tenth of the 10-year return period value). In this regard, operators should pay particular attention to the distribution of the best analogues. If, for example, it is shifted to higher values, there is a risk of under-estimating the event when considering the middle of the full distribution. Obviously, if the target situation has its best analogues with high precipitations, then this situation is likely to be rare. So, when collecting a fixed number of analogue situations, the worst ones are likely to be more common and will contribute

to lower the resulting precipitation distribution, therefore a trend towards underestimation.

Marty (2010) applied several validation scores to the quantiles of the AM distribution in operational forecasting. He concluded that the detection of the precipitation occurrence was better predicted by median quantiles (50-70 %), with an optimum to 60 %. However, once again, for significant rainfall, the 90th percentile gave the best performance. Comparing the results from the AM to an ensemble forecast, Marty (2010) found the AM to be better than the considered ensembles, particularly for strong precipitation. The gain of the AM comes from a greater accuracy for a poorer sharpness, when the ensemble forecasts are very sharp, but lack a bit of accuracy.

When comparing the performance of Meteo France numerical models along with the AM, Djerboua (2001) concluded that the numerical models were at that time better at predicting the precipitation of the current day, but the AM became more competitive for the following lead times. It is indeed very interesting in early warning approaches, as it may able to identify a potentially critical situation one week ahead.

All AMs rely on atmospheric circulation information, while some add subsequent levels of analogy on moisture or other thermodynamic predictors (see section 2.5). However, these subsequent predictors were found to be relevant only on the first 3 days, and not much further, due to higher uncertainties in their forecast (see section 6.6). This was found to be the case for moisture indexes and vertical velocity. The user has to keep this aspect in mind when interpreting forecasts over a range of lead times.

6 AtmoSwing Optimizer

AtmoSwing Optimizer is a single tool that integrates different optimization methods, presented in section 6.3 to 6.5. Optimization is essentially performed in the calibration framework (section 6.1), where one tries to maximize (by means of performance scores, see section 6.2) the information that can be extracted from the available predictors and predictand archives.

However, in operational use, a further ingredient arises : the model from which the target synoptic situations are taken for the coming days, and the increasing uncertainty related to the lead time. Depending on the properties of these forecasts, the operational scheme has to be tuned (section 6.6).

6.1 Calibration framework

The calibration of the AM is usually carried out in a perfect prognosis (Klein et al., 1959) framework (Bontron, 2004; Ben Daoud, 2010). Perfect prognosis uses observed or reanalyzed data to calibrate the relationship between predictors and predictands. As a result, this framework will identify and

provide us with relationships that are as close as possible to the natural links between predictors and predictands. However, no model is perfect, and even reanalysis data may contain bias that cannot be ignored. The calibration process has to try to cope with such bias, but also with the data spatial resolution, that may be too rough for certain variables like humidity or vertical velocity. For these reasons, the statistical relationships identified in the perfect prognosis framework should hopefully be applied to new target situations (e.g. resulting from an operational NWP model) with characteristics as similar as possible to the those of the reanalyses.

Another reason for working in a perfect prognosis framework is that numerical models evolve continuously, and so does the forecast they provide. Considering re-forecasts would allow us to work on a homogeneous dataset, as they are regularly reprocessed with a defined version of the model. However, one would need to redo the calibration procedure every time a new version is available, in order to reduce the bias (Wilson and Vallée, 2002). Moreover, the length of reanalyses datasets are usually much longer than reforecast datasets, which allows us to identify more robust relationships. The size of the archive is indeed an important criteria for the AM.

The statistical relationship is established on a calibration period that is as long as possible. For every day of this period, a search for analogues is performed, the predictand data are associated with the corresponding dates and a forecast score is calculated. During the search for analogues situations, 120 days around the target date are excluded (thus excluding data in the same year) in order to consider only truly independent candidates days.

A validation period is always considered. It consists of an independent period that is never used as target neither candidate. Validating the parameters of the AM is very important in order to avoid over-parametrization and thus to ensure that the statistical relationship is valid on another period. In order to account for climate change and the evolution of the measuring techniques, it is recommended to use a noncontinuous period for validating, distributed over the whole archive (Ben Daoud, 2010). The user can thus specify a list of the years to keep for the validation and AtmoSwing Optimizer handles this aspect easily by both removing the validation years from possible candidate situations, and by automatically processing the validation score on these years at the end of the optimization.

6.2 Implemented forecasts scores

Most often, the accuracy of the parameters for precipitations forecasting is evaluated by means of the CRPS (Continuous Ranked Probability Score, Brown, 1974; Matheson and Winkler, 1976; Hersbach, 2000). It is the only score that will be detailed hereafter (in section 6.2.4). In addition, other scores are implemented in AtmoSwing Optimizer. The following sections list the scores presently available, but will not detail them.

6.2.1 Discrete deterministic forecasts

Deterministic forecasts consider no uncertainty. A single value is provided for a variable. The first group of this type of forecast is based on discrete values, also called categories. These scores are used, for example, to evaluate a deterministic prediction of threshold exceedance. In our case, the continuous probabilistic nature of an analogy forecast can be transformed into a discrete forecast by considering a threshold exceedance and a fixed percentile from the distribution. These scores rely on a contingency table, which is here a table containing the sums of observed / unobserved and predicted / unpredicted events (Wilks, 2006). On the basis of this table, multiple scores can be processed within AtmoSwing:

- Proportion correct (Finley, 1884)
- Threat Score (Gilbert, 1884)
- Bias
- False Alarm Ratio
- Hit Rate or Probabiliy of Detection
- False Alarm Rate
- Heidke Skill Score (Heidke, 1926)
- Peirce Skill Score (Peirce, 1884)
- Gilbert Skill Score or Equitable Threat Score (Gilbert, 1884)

6.2.2 Continuous deterministic forecasts

The difference of the continuous deterministic forecasts compared to discrete ones is that they provide non-discrete values that must be evaluated with distance measures. In our application, the distribution of the analogues is summarized by a chosen percentile. Available scores are the following:

- Mean Absolute Error
- Root Mean Squared Error

6.2.3 Discrete probabilistic forecasts

In the case of discrete probabilistic forecasts, one works with probabilities of occurrence or probability of belonging to a certain category. The implemented options are:

- Brier Score (Brier, 1950)
- ROC diagramm (Relative Operating Characteristic or Receiver Operating Characteristic, Mason, 1982)
- RPS (Ranked Probability Score, Epstein, 1969)
- SEEPS (Stable Equitable Error in Probability Space, Rodwell et al., 2010, 2011)

6.2.4 Continuous probabilistic forecasts

Continuous probabilistic forecasts of a variable are issued in the form of the expected statistical distribution for this variable. They are expressed either in terms of distribution or by probability density, which needs to be compared to an observed value. This is the situation one meets when forecasting precipitations with the AM.

Most applications mainly rely on the CRPS (Continuous Ranked Probability Score, Brown, 1974; Matheson and Winkler, 1976; Hersbach, 2000). It allows assessing the predicted cumulative distribution functions $F(y)$, for example of the precipitation values y from analogue situations, compared to the observed value y^0. The better the forecast, the smaller the score. The mean CRPS of a forecast series of length n can be written:

$$CRPS = \frac{1}{n} \sum_{i=1}^{n} \left(\int_{-\infty}^{+\infty} \left[F_i(y) - H_i(y - y_i^0) \right]^2 dy \right) \qquad (2)$$

where $H(y - y_i^0)$ is the Heaviside function that is null when $y - y_i^0 < 0$, and has the value 1 otherwise. The mean CRPS is averaged on the calibration, respectively the validation periods, on all days, may they be dry, slightly rainy or with heavy precipitation.

This score is now commonly used for the evaluation of continuous variables prediction systems (Casati et al., 2008; Marty, 2010). To compare the value of the score in regard to a reference, one often considers its skill score value, and use the climatological distribution as the reference. The CRPSS (*Continuous Ranked Probability Skill Score*) is thus defined as following:

$$CRPSS = \frac{CRPS - CRPS_r}{CRPS_p - CRPS_r} = 1 - \frac{CRPS}{CRPS_r} \qquad (3)$$

where $CRPS_r$ is the CRPS value for the reference and $CRPS_p$ would be the one for a perfect forecast (which implies $CRPS_p = 0$).

The CRPS can be decomposed into several indicators, also implemented into AtmoSwing Optimizer, such as:

– Reliability (Hersbach, 2000)
– Resolution / uncertainty (Hersbach, 2000)
– Sharpness (Bontron, 2004)
– Accuracy (Bontron, 2004)

Finally, the rank diagram (Talagrand et al., 1997) and its accuracy as defined by Candille and Talagrand (2005) are also available.

6.3 The sequential calibration

The calibration procedure that we call "sequential" or "classic", but which is also named "optimization" (eg. by Ben

Daoud et al., 2015), was developed by Bontron (2004) at the LTHE laboratory (INPG, Grenoble). It is a semi-automatic procedure that determines the optimal parameters for the different variables of each level of analogy. The analogy levels (eg. the atmospheric circulation or moisture index) are calibrated sequentially. The procedure consists of the following steps (Bontron, 2004):

1. Manual selection of the following parameters:

   (a) meteorological variable (e.g. geopotential, temperature, relative humidity, etc.),
   (b) pressure level (e.g. 700 hPa, ...),
   (c) temporal window (hour of observation - e.g. 6, or 12 h UTC),
   (d) initial analogue numbers (e.g. $N_1 = 50$).

2. For every level of analogy:

   (a) Identification of the most skilled unitary cell (1 point for moisture variables and 4 for the geopotential height when using the S1 criteria) over a large domain. Every point (or cell) of the full domain is assessed jointly on all predictors of the level of analogy (consisting generally of the same variable, but on different pressure levels and at different hours).

   (b) From this most skilled point, the spatial window is expanded by successive iterations in the direction of greater performance gain. The detailed stages are the followings: (i) The unitary spatial window is expanded in every 4 directions successively. The performance score is processed for these 4 windows. (ii) Only the direction providing the best improvement is applied to the spatial window. (iii) From this new spatial window, an increase in every 4 directions is once again assessed, and the best improvement is applied. (iv) The spatial window grows up by repeating the previous steps, until no improvement is reached.

   (c) The number of analogue situations is then reconsidered and optimized for the current level of analogy.

   (d) A new level of analogy can then be added, based on other variables (such as the moisture index) on predefined pressure levels and time frames. The number of analogues for the next level of analogy is initiated at a chosen value. Then, the procedure starts again from step (a) for the new level. The parameters calibrated on the previous analogue levels are fixed and do not change (except the numbers of analogues of all levels of analogy, that are optimized once again at the final stage).

3. Finally, the numbers of analogues are re-assessed for the different levels of analogy. This is done iteratively

by varying the number of analogues of each level in a systematic way.

The calibration is thus done in successive steps, on a limited number of parameters. Previously calibrated parameters are generally not reassessed. The advantage of this method is that it is relatively fast, it provides acceptable results, and it has low computing requirements.

Small improvements were added to this method, then named "classic+", by allowing the spatial windows to do other moves, such as: (1) increase in 2 simultaneous directions, (2) decrease in 1 or 2 simultaneous directions, (3) expansion or contraction (in every direction), (4) shift of the window (without resizing) in 8 directions (including diagonals), and finally (5) all the moves described above, but with a factor of 2, 3, or more. For example, an increase by 2 units in one (or more) direction is assessed. This allows to skip one size that may not be optimal. These supplementary steps often result in spatial windows that are a bit different, but the performance gain is rather marginal.

6.4 Monte-Carlo analysis

A Monte-Carlo analysis is also implemented in AtmoSwing. It performs thousands assessments of random parameters within given ranges. This method is not efficient to find the best parameters set, but helps to better understand the sensitivity of the parameters.

For example, a Monte-Carlo analysis was performed on the PC-2Z (Table 1) AM for the southeastern crests on the upper Rhône catchment in Switzerland, with 10,000 random selections of its parameters according to a uniform law (the analysis was also performed in other subregions of the upper Rhône catchment in Switzerland, but the conclusions depicted here are similar). Unlike conventional calibration, the spatial windows on Z500 and Z1000 were not necessarily overlapping. Thus, 9 parameters were varying together (8 for the spatial windows and 1 for the number of analogues)

Figure 9 presents the resulting scatter of random exploration of the parameters space, along with the results of the sequential calibration (red cross). Results are illustrated here for the 4 parameters defining the spatial window on Z500. Despite the high number of simulations, the upper portion of the distributions seems incomplete, and the score of the sequential calibration is not matched. Although the method is poor to find efficiently the best parameters, it is informative as to the shape of the scatter plots.

The shape of the scatter plots from Figure 9 reveals a low sensitivity as to the exact location of the spatial windows. There is a maximum threshold for minimum longitudes and latitudes and, inversely, a minimum threshold for maximum latitudes and longitudes, but outside of these limits, the distribution slope is more gentle. A spatial window can thus be a bit larger than its optimal size without a significant drop in performance, as long as it includes at least one critical region. This was also observed by Bontron (2004), who noted that "*performance slowly decreases if we consider a slightly too large window, while the use of too small windows results in strong performance loss*". The dilution of a part of the relevant synoptic information has therefore not necessarily a significant negative impact on performance, while ignoring some of this information leads to undesirable loss of performance. It is the same for the analogues number (not shown), for which, past an optimum, the upper slope of the distribution decreases slowly.
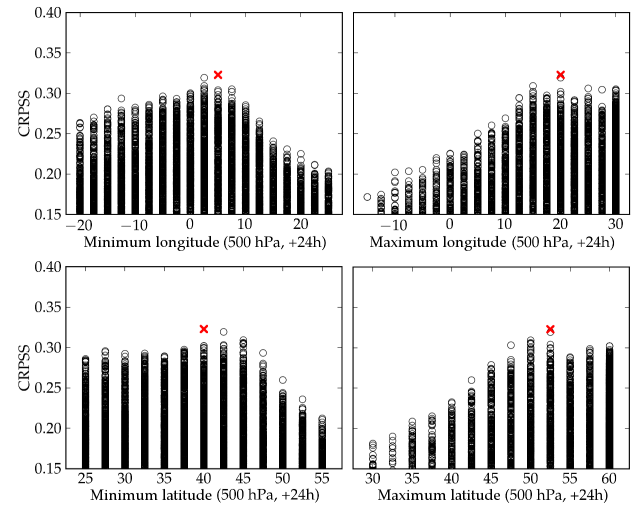
Figure 9. Example of a Monte Carlo analysis on the PC-2Z (Table 1) AM. Results are illustrated for the 4 parameters defining the spatial window on Z500, along with the results of the sequential calibration (red cross).

The same analysis has been conducted on the AM with 2 levels of analogy, PC-2Z-2MI (Table 2), and the results on the number of analogues on the 1st and the 2nd level are illustrated in Figure 10. One can see that the AM is not very sensitive to the number of analogues on the first level of analogy, when considering a second level afterwards, as long as this number is not too small. The trend on the second level is more apparent, but a large tolerance seems possible.

The use of such a random method allows informing on the parameters sensitivity, and so on the importance of the expected precision of the calibration results. It helps understanding how different parameterizations can end up to relatively similar performances. Using Monte Carlo analyses allows to easily display the AM parameters sensitivity.

6.5 Global optimization

Due to limitations in the sequential approach, we decided to explore optimization techniques that could handle all parameters fully automatically. The Nelder and Mead (1965) method was first assessed, but was found unsuitable for this purpose. Because of the complexity of the AM and its non-
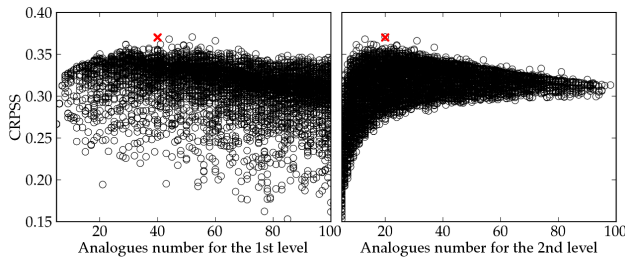
Figure 10. Example of a Monte Carlo analysis on the PC-2Z-2MI (Table 2) AM. Results are illustrated for the number of analogues on the 1st and 2nd level of analogy, along with the results of the sequential calibration (red cross).

linearity, a technique of global optimization is required. Genetic algorithms were thus implemented in AtmoSwing Optimizer in order to perform a global optimization of the AM. This allows for a fully automatic and objective optimization of the AM. However, as this topic is rather consequent, it is the subject of other coming publications (see Horton et al., 2016a, b).

6.6 Optimization for operational use

Up to now, the AM has been optimized essentially by replaying past situations that were taken from a unique meteorological archive. Therefore, the synoptic variables have similar properties in terms of resolution, models that have processed them, etc... However, when it comes to operational forecasting, as already mentioned, the resulting AM will be applied to target situations that are no longer similar to the reanalyzed ones. They will be taken from global model forecasts, which may not be those used to elaborate the archive of reanalyzed situations. Thus, these new targets will display characteristics different from those of the archive, and will in particular contain larger uncertainties, especially at growing lead-times.

Eventually, a meteorological variable that was proven to be a good predictor during the calibration of the AM may be poorly predicted by the forecast NWP model selected beyond a certain lead time. In this regards, it should be dropped after the lead time in question (see section 5.2).

To assess and address these operational aspects, the best option is to collect a dataset of re-forecast from the specific NWP model selected. Then, distinct variants of AM parameterizations can be assessed on different lead times. For example, when using moisture variables for the second level of analogy, Thevenot (2004) showed that beyond a lead time of 3 days, the AM with two levels did not do better that the one with a unique level of analogy. It comes from the fact that large scale geopotential variables are correctly predicted up to 8 or 10 days, while humidity variables are not, at least by the current version of the model considered.

Next, we assume that for the synoptic variables used as predictors, the model forecasts have more uncertainties than the reanalyses, but no systematic bias. To optimize further the AM, it can be considered that the further the lead time, the more uncertain the synoptic forecasts and therefore the more uncertain the precipitation forecasts could be. This can been addressed by reconsidering, for each lead-time, the optimal number of analogues to select, for re-forecast data from a particular NWP model (Thevenot, 2004). The underlying idea is that if the model proves perfect, then the number of analogues should remain close to that obtained in the perfect prognosis framework. While if the forecast becomes irrelevant, for example at distant lead times, the AM should propose a forecast closer and closer to the climatological distribution of the predictand. The optimal number of analogue situations will thus increase with the lead time, and should be optimized for every lead time on the re-forecast dataset (see Thevenot, 2004, for the details). This option is implemented in AtmoSwing Forecaster (see section 4).

It should be stressed that these operational aspects are dependent on the model used to provide the AM with synoptic forecasts, and on the availability of a re-forecast data to perform the tuning of the AM to this particular model.

7 Conclusions and perspectives

Processing operational forecasts by means of AtmoSwing requests in the end very low computing infrastructure and can provide useful information, such as early warning for extreme precipitation, in the case of an application in flood forecasting. The Forecaster is very flexible as it builds the desired method in a dynamic way, and can thus implement multiple variants of the AM. The connection with open access NWP model such as GFS is easy and very efficient (data are collected quickly after being available, and processed in a few minutes). The Viewer offers a user-friendly display of the forecasts, with different levels of synthesis and details. The Optimizer implements different optimization techniques, such as the sequential approach, a Monte-Carlo method, and a global optimization technique. Obviously, it is used only occasionally for a given location: for a first implementation of an AM, in case of a change in the archive, or of a substantial updating, etc.

Meanwhile, several additions would be nice to have in AtmoSwing for operational forecasting. The first one would be using outputs from multiple global NWP models within one forecast, namely making multi-models forecasts, in order to take into account the variability between various providers. In the same direction, Thevenot (2004) showed the benefit of using ensembles of global NWP forecasts as input for the method (in his case the 51 traces from the EFS of ECMWF). This approach takes into account the uncertainties on predictors from the global model. Its implementation is a combination of the selected analogue days associated with each of

the traces. Rainfall distribution is then established on all ana-logues. The forecast on the ensemble was found to be more accurate than the deterministic control for a lead time of 4 days and more (Thevenot, 2004).

Next, a bias correction should be implemented. Marty (2010) has indeed observed that the frequency of the forecasted days without precipitation was biased and underestimated the actual frequency of zero precipitation. On the contrary, the frequency of the forecasted days with precipitation overestimated the observed frequencies. On this basis, Marty (2010) proposed a bias correction technique, that should be implemented as well. This issue occurs in operational forecasting and was not observed in perfect prognosis.

As it was shown in the historical section, the AM evolves, with the availability of new data, computer facilities and progress in NWP models, and still proves to bring complementary information on predictands like precipitation, for example at long lead time. Its interest in climatological downscaling has not been addressed in this paper but should be mentioned. However, there remain still many possibilities to be explored. Amongst them, one can mention:

– an optimization per season in order to get parameterizations specific to the leading processes during the period in question;

– optimizations specific for extreme events, with an objective function tailored to the determination of the most critical situations;

– a better insight in the role of the length of the archive, by using long reanalyses over the 20th century;

– a weighting of the predictor grids, in order to get free of fixed rectangular spatial windows;

– assessment of potential indicators based on the properties of the distribution of all analogues or the 10 best, for example;

– assessment of the relevance of an automatic interpretation of the distribution in order to provide the most likely expected value;

Code availability

AtmoSwing is open source, under the CDDL license. It is developed under the Mercurial distributed revision control, and the source code can be found at:

– http://www.atmoswing.org

– https://bitbucket.org/atmoswing/atmoswing

To get access to the frozen version at the time of writing (v1.4.2), the following address can be used:

– bitbucket.org/atmoswing/atmoswing/commits/tag/1.4.2

References

Barnett, T. and Preisendorfer, R.: Multifield analog prediction of short-term climate fluctuations using a climate state vector, Journal of the Atmospheric Sciences, 35, 1771–1787, http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1978) 035{%}3C1771:MAPOST{%}3E2.0.CO;2, 1978.

Ben Daoud, A.: Améliorations et développements d'une méthode de prévision probabiliste des pluies par analogie., Ph.D. thesis, Université de Grenoble, 2010.

Ben Daoud, A., Sauquet, E., Lang, M., Obled, C., and Bontron, G.: La prévision des précipitations par recherche d'analogues: état de l'art et perspectives, in: Colloque SHF-191e CST - «Prévisions hydrométéorologiques», Lyon, 18-19 novembre 2008, 6, 2008.

Ben Daoud, A., Sauquet, E., Bontron, G., Obled, C., and Lang, M.: Daily quantitative precipitation forecasts based on the analogue method: improvements and application to a French large river basin, Atmospheric Research, 169, 147–159, doi:10.1016/j.atmosres.2015.09.015, http://linkinghub. elsevier.com/retrieve/pii/S0169809515002951, 2015.

Bergen, R. and Harnack, R.: Long-range temperature prediction using a simple analog approach, Monthly Weather Review, 110, 1083–1099, http://journals.ametsoc. org/doi/abs/10.1175/1520-0493(1982)110{%}3C1083: LRTPUA{%}3E2.0.CO{%}3B2, 1982.

Binder, P. and Schär, C.: Mesoscale Alpine Programme (MAP) design proposal, Tech. rep., MAP Programme Office, Zurich, Switzerland, 1996.

Bliefernicht, J.: Probability forecasts of daily areal precipitation for small river basins, Ph.D. thesis, Universität Stuttgart, 2010.

Bois, P., Obled, C., and Thalamy, J.: Etude des liaisons entre champs de pression et températures sur l'Europe avec la durée d'insolation: Application à la possibilité de la prévision d'insolation, in: C.R. du colloque Météorologie de l'énergie solaire, PIRDES-CNRS, Toulouse, pp. 145–177, 1981.

Bolognesi, R.: Premiers développements d'un modèle hybride pour le diagnostic spatial des risques d'avalanches, La Houille Blanche, 8, 551–553, 1993.

Bontron, G.: Prévision quantitative des précipitations: Adaptation probabiliste par recherche d'analogues. Utilisation des Réanalyses NCEP/NCAR et application aux précipitations du Sud-Est de la France., Ph.D. thesis, Institut National Polytechnique de Grenoble, 2004.

Bougeault, P., Binder, P., Buzzi, A., Dirks, R., Houze, R., Kuettner, J., Smith, R., Steinacker, R., and Volkert, H.: The MAP special observing period, Bulletin of the American Meteorological Society, 82, 433–462, http://journals.ametsoc.org/doi/abs/10.1175/1520-0477(2001)082{%}3C0433:TMSOP{%}3E2.3.CO;2, 2001.

Brier, G.: Verification of forecasts expressed in terms of probability, Monthly weather review, 78, 1–3, http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1950)078{%}3C0001:VOFEIT{%}3E2.0.CO;2, 1950.

Brown, T.: Admissible Scoring Systems for Continuous Distributions., http://eric.ed.gov/?id=ED135799, 1974.

Candille, G. and Talagrand, O.: Evaluation of probabilistic prediction systems for a scalar variable, Quarterly Journal of the Royal Meteorological Society, 131, 2131–2150, doi:10.1256/qj.04.71, http://doi.wiley.com/10.1256/qj.04.71, 2005.

Casati, B., Wilson, L. J., Stephenson, D. B., Nurmi, P., Ghelli, a., Pocernich, M., Damrath, U., Ebert, E. E., Brown, B. G., and Mason, S.: Forecast verification: current status and future directions, Meteorological Applications, 15, 3–18, doi:10.1002/met.52, http://doi.wiley.com/10.1002/met.52, 2008.

Chardon, J., Hingray, B., Favre, A.-C., Autin, P., Gailhard, J., Zin, I., and Obled, C.: Spatial Similarity and Transferability of Analog Dates for Precipitation Downscaling over France, Journal of Climate, 27, 5056–5074, doi:10.1175/JCLI-D-13-00464.1, http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-13-00464.1, 2014.

Dayon, G., Boé, J., and Martin, E.: Transferability in the future climate of a statistical downscaling method for precipitation in France, Journal of Geophysical Research: Atmospheres, 120, 1023–1043, doi:10.1002/2014JD022236, http://doi.wiley.com/10.1002/2014JD022236, 2015.

Desaint, B., Nogues, P., Perret, C., and Garçon, R.: La prévision hydrométéorologique opérationnelle: l'expérience d'Electricité de France, in: Colloque SHF-191e CST - «Prévisions hydrométéorologiques», Lyon, 18-19 novembre 2008, Tableau 1, http://www.shf-lhb.org/articles/lhb/abs/2009/05/lhb2009054/lhb2009054.html, 2008.

Djerboua, A.: Prédétermination des pluies et crues extrêmes dans les Alpes franco-italiennes - Prévision quantitative des pluies journalières par la méthode des Analogues, Ph.D. thesis, Institut National Polytechnique de Grenoble, 2001.

Drosdowsky, W. and Zhang, H.: Verification of Spatial Fields, in: Forecast verification: a practitioner's guide in atmospheric science, edited by Jolliffe, I. T. and Stephenson, D. B., chap. 6, pp. 121–136, Wiley, 2003.

Duband, D.: Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations., Ph.D. thesis, Thèse de 3ème cycle de la faculté des sciences de Paris., 1970.

Duband, D.: Reconnaissance dynamique de la forme des situations météorologiques, application à la prévision numérique des précipitations journalières, in: Congrès de la Société Hydrotechnique de France, XIIIèmes journées de l'Hydraulique, Paris, 1974.

Duband, D.: Prévision spatiale des hauteurs de précipitations journalières, La Houille Blanche, 7-8, 497–512, 1981.

Epstein, E.: A scoring system for probability forecasts of ranked categories, Journal of Applied Meteorology, 8, 985–987, http://journals.ametsoc.org/doi/abs/10.1175/1520-0450(1969)008{%}3C0985:ASSFPF{%}3E2.0.CO{%}3B2, 1969.

Finley, J.: Tornado prediction, American Meteorological Journal, 1, 85–88, 1884.

Fraedrich, K., Raible, C. C., and Sielmann, F.: Analog Ensemble Forecasts of Tropical Cyclone Tracks in the Australian Region, Weather and Forecasting, 18, 3–11, doi:10.1175/1520-0434(2003)018<0003:AEFOTC>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0434{%}282003{%}29018{%}3C0003{%}3AAEFOTC{%}3E2.0.CO{%}3B2, 2003.

García Hernández, J., Horton, P., Tobin, C., and Boillat, J.: MINERVE 2010: Prévision hydrométéorologique et gestion des crues sur le Rhône alpin., Wasser Energie Luft – Eau Energie Air, 4, 297–302, http://infoscience.epfl.ch/record/143020/files/Wasser{_}Energie{_}Luft{_}MINERVE{_}JG-PH-CT-JLB.pdf, 2009.

Gibergans-Báguena, J. and Llasat, M.: Improvement of the analog forecasting method by using local thermodynamic data. Application to autumn precipitation in Catalonia, Atmospheric Research, 86, 173–193, doi:10.1016/j.atmosres.2007.04.002, http://linkinghub.elsevier.com/retrieve/pii/S0169809507000695, 2007.

Gilbert, G. K.: Finley's tornado predictions, American Meteorological Journal, 1, 166–172, 1884.

Goodess, C.: Statistical and regional dynamical downscaling of extremes for European regions: STARDEX, European geophysical union information newsletter, 6, 25–29, http://scholar.google.com/scholar?hl=en{&}btnG=Search{&}q=intitle:Statistical+and+regional+dynamical+downscaling+of+extremes+for+European+regions:+STARDEX{#}0, 2003.

Gordon, N. D.: Statistical very short-range forecasting via analogues., European Space Agency Special Publication, pp. 307–312, 1987.

Guennebaud, G., Jacob, B., and Others: Eigen v3, http://eigen.tuxfamily.org, 2010.

Guilbaud, S.: Prévision quantitative des précipitations journalières par une méthode statistico-dynamique des recherche d'analogues - Application à des bassins du pourtour méditerranéen, Ph.D. thesis, Institut National Polytechnique de Grenoble, 1997.

Guilbaud, S. and Obled, C.: Prévision quantitative des précipitations journalières par une technique de recherche de journées antérieures analogues: optimisation du critère d'analogie, Comptes Rendus de l'Académie des Sciences, Series II, A-Earth and Planetary Science, 327, 181–188, http://www.sciencedirect.com/science/article/pii/S1251805098800062, 1998.

Hamill, T. and Whitaker, J.: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application, Monthly Weather Review, 134, 3209–3229, 2006.

Heidke, P.: Berechnung des Erfolges und der Güte der Windstärkevorhersagen im Sturmwarnungsdienst, Geografiska Annaler, 8, 301–349, http://www.jstor.org/stable/519729, 1926.

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, Weather and Forecasting, 15, 559–570, http://journals.ametsoc.org/doi/abs/10.1175/1520-0434(2000)015{%}3C0559:DOTCRP{%}3E2.0.CO;2, 2000.

Hewitson, B. and Crane, R.: Climate downscaling: techniques and application, Climate Research, 7, 85–95, http://scholar.google.com/scholar?hl=en{&}btnG=Search{&}q=intitle:

Climate+downscaling+:+techniques+and+application{#}0, 1996.

Hoare, C. A. R.: Quicksort, The Computer Journal, 5, 10–16, 1962.

Horton, P., Jaboyedoff, M., Metzger, R., Obled, C., and Marty, R.: Spatial relationship between the atmospheric circulation and the precipitation measured in the western Swiss Alps by means of the analogue method, Natural Hazards and Earth System Sciences, 12, 777–784, doi:10.5194/nhess-12-777-2012, 2012.

Horton, P., Jaboyedoff, M., and Obled, C.: Global Optimization of the Analogue Method by Means of Genetic Algorithms, Journal of Climate, submitted, 2016a.

Horton, P., Jaboyedoff, M., and Obled, C.: Using Genetic Algorithms to Optimize the Analogue Method for Precipitation forecasting in the Swiss Alps, Journal of Hydrology, submitted, 2016b.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., and Woollen, J.: The NCEP/NCAR 40-year reanalysis project, Bulletin of the American Meteorological Society, 77, 437–471, http://journals.ametsoc.org/doi/abs/10.1175/1520-0477(1996)077{%}3C0437:TNYRP{%}3E2.0.CO;2, 1996.

Kanamitsu, M.: Description of the NMC global data assimilation and forecast system, Weather and Forecasting, 4, 335–342, http://journals.ametsoc.org/doi/abs/10.1175/1520-0434(1989)004{%}3C0335:DOTNGD{%}3E2.0.CO{%}3B2, 1989.

Kanamitsu, M., Alpert, J., Campana, K., Caplan, P., Deaven, D., Iredell, M., Katz, B., Pan, H., Sela, J., and White, G.: Recent changes implemented into the global forecast system at NMC, Weather and Forecasting, 6, 425–435, http://journals.ametsoc.org/doi/abs/10.1175/1520-0434(1991)006{%}3C0425:RCIITG{%}3E2.0.CO{%}3B2, 1991.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP–DOE AMIP-II Reanalysis (R-2), Bulletin of the American Meteorological Society, 83, 1631–1643, doi:10.1175/BAMS-83-11-1631, http://journals.ametsoc.org/doi/abs/10.1175/BAMS-83-11-1631, 2002.

Keenan, T. D. and Woodcock, F.: Objective Tropical Cyclone Movement Forecasts Using Synoptic and Track Analogue Information, Tech. rep., Bureau of Meteorology, 1981.

Klein, W.: Specification of precipitation from the 700-millibar circulation, Monthly Weather Review, 91, 527–536, http://scholar.google.com/scholar?hl=en{&}btnG=Search{&}q=intitle:Specification+of+precipitation+from+the+700-millibar+circulation{#}0, 1963.

Klein, W., Lewis, B., and Enger, I.: Objective prediction of five-day mean temperatures during winter, Journal of Meteorology, 16, 672–682, http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1959)016{%}3C0672:OPOFDM{%}3E2.0.CO;2, 1959.

Kruizinga, S. and Murphy, A.: Use of an analogue procedure to formulate objective probabilistic temperature forecasts in the Netherlands, Monthly Weather Review, 111, 2244–2254, http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1983)111{%}3C2244{%}3AUOAAPT{%}3E2.0.CO{%}3B2, 1983.

Livezey, R. E. and Barnston, A. G.: An operational multifield analog/antianalog prediction system for United States seasonal temperatures. 1. system design and winter experiments, Journal of Geophysical Research, 93, 10 953–10 974, 1988.

Lorenz, E.: Empirical orthogonal functions and statistical weather prediction, Tech. rep., Massachusetts Institute of Technology, Department of Meteorology, Massachusetts Institute of Technology, Dept. of Meteorology, http://www.o3d.org/abracco/Atlantic/Lorenz1956.pdf, 1956.

Lorenz, E.: Atmospheric predictability as revealed by naturally occurring analogues, Journal of the Atmospheric Sciences, 26, 636–646, http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1969)26{%}3C636:APARBN{%}3E2.0.CO;2, 1969.

Maheras, P., Anagnostopoulou, C., and Tolika, K.: Contribution to D12 (AUTH, APRA-SMR, FIC and UEA) Downscaling of extreme indices in Greece, Tech. rep., STARDEX project, http://www.cru.uea.ac.uk/projects/stardex/deliverables/D12/D12{_}regional{_}Greece.pdf, 2005.

Mandon, S.: Comparaison d'épisodes pluvieux intenses sur le Sud-Est de la France et de situations analogues au sens de la circulation générale. Recherche de variables discriminantes., Tech. rep., Rapport de stage de fin d'études de l'Ecole de la Météorologie Nationale, 1985.

Marty, R.: Prévision hydrologique d'ensemble adaptée aux bassins à crue rapide. Elaboration de prévisions probabilistes de précipitations à 12 et 24 h. Désagrégation horaire conditionnelle pour la modélisation hydrologique. Application à des bassins de la région Cév, Ph.D. thesis, Université de Grenoble, 2010.

Marty, R., Zin, I., Obled, C., Bontron, G., and Djerboua, A.: Toward Real-Time Daily PQPF by an Analog Sorting Approach: Application to Flash-Flood Catchments, Journal of Applied Meteorology and Climatology, 51, 505–520, doi:10.1175/JAMC-D-11-011.1, http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-11-011.1, 2012.

Mason, I.: A model for assessment of weather forecasts, Australian Meteorological Magazine, 30, 291–303, http://www.nssl.noaa.gov/users/brooks/public{_}html/feda/papers/mason82.pdf, 1982.

Matheson, J. and Winkler, R.: Scoring rules for continuous probability distributions, Management science, 22, 1087–1096, http://pubsonline.informs.org/doi/abs/10.1287/mnsc.22.10.1087, 1976.

Matulla, C., Zhang, X., Wang, X. L., Wang, J., Zorita, E., Wagner, S., and von Storch, H.: Influence of similarity measures on the performance of the analog method for downscaling daily precipitation, Climate Dynamics, 30, 133–144, doi:10.1007/s00382-007-0277-2, http://link.springer.com/10.1007/s00382-007-0277-2, 2007.

Murray, R.: Indicators of monthly mean temperature and rainfall for England and Wales based on antecedent monthly pressure anomalies over the Northern hemisphere, The Meteorological magazine, 103, 70, 1974.

Nap, J., van den Dool, H., and Oerlemans, J.: A verification of monthly weather forecasts in the seventies, Monthly Weather Review, 109, 306–312, http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1981)109{%}3C0306{%}3AAVOMWF{%}3E2.0.CO{%}3B2, 1981.

Nelder, J. and Mead, R.: A simplex method for function minimization, The Computer Journal, 7, 308–313, http://comjnl.oxfordjournals.org/content/7/4/308.short, 1965.

Obled, C. and Good, W.: Recent developments of avalanche forecasting by discriminant analysis techniques: a methodological review and some applications to the Parsenn area (Davos, Switzerland), Journal of Glaciology, 25, 315–346, 1980.

Obled, C., Bontron, G., and Garçon, R.: Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach, Atmospheric Research, 63, 303–324, doi:10.1016/S0169-8095(02)00038-8, http://linkinghub.elsevier.com/retrieve/pii/S0169809502000388, 2002.

Peirce, C.: The numerical measure of the success of predictions, Science, 4, 453–454, http://www.nssl.noaa.gov/users/brooks/public{_}html/feda/papers/Peirce1884.pdf, 1884.

Philipp, A., Della-Marta, P. M., Jacobeit, J., Fereday, D. R., Jones, P. D., Moberg, A., and Wanner, H.: Long-Term Variability of Daily North Atlantic–European Pressure Patterns since 1850 Classified by Simulated Annealing Clustering, Journal of Climate, 20, 4065–4095, doi:10.1175/JCLI4175.1, http://journals.ametsoc.org/doi/abs/10.1175/JCLI4175.1, 2007.

Radinovic, D.: An analogue method for weather forecasting using the 500/1000 mb relative topography, Monthly Weather Review, 103, 639–649, http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1975)103{%}3C0639{%}3AAAMFWF{%}3E2.0.CO{%}3B2, 1975.

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G.: Daily High-Resolution-Blended Analyses for Sea Surface Temperature, Journal of Climate, 20, 5473–5496, doi:10.1175/2007JCLI1824.1, http://journals.ametsoc.org/doi/abs/10.1175/2007JCLI1824.1, 2007.

Rodwell, M., Haiden, T., and Richardson, D.: Developments in precipitation verification, ECMWF Newsletter, 128, 12–16, http://scholar.google.com/scholar?hl=en{&}btnG=Search{&}q=intitle:Developments+in+precipitation+verification{#}1, 2011.

Rodwell, M. J., Richardson, D. S., Hewson, T. D., and Haiden, T.: A new equitable score suitable for verifying precipitation in numerical weather prediction, Quarterly Journal of the Royal Meteorological Society, 136, 1344–1363, doi:10.1002/qj.656, http://doi.wiley.com/10.1002/qj.656, 2010.

Schmidli, J., Goodess, C. M., Frei, C., Haylock, M. R., Hundecha, Y., Ribalaygua, J., and Schmith, T.: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps, Journal of Geophysical Research, 112, D04105, doi:10.1029/2005JD007026, http://doi.wiley.com/10.1029/2005JD007026, 2007.

Shabbar, A. and Knox, J.: Monthly prediction by the analogue method, in: Proceedings of the first WMO workshop on the diagnosis and prediction of monthly and seasonal atmospheric variations over the globe, Longe range forecasting Res. Rep. Ser. 6, Vol II, Tech. Doc. WMO/TD87, pp. 672–681, 1986.

Sievers, O., Fraedrich, K., and Raible, C. C.: Self-Adapting Analog Ensemble Predictions of Tropical Cyclone Tracks, Weather and Forecasting, 15, 623–629, doi:10.1175/1520-0434(2000)015<0623:SAAEPO>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0434{%}282000{%}29015{%}3C0623{%}3ASAAEPO{%}3E2.0.CO{%}3B2, 2000.

Smart, J., Hock, K., and Csomor, S.: Cross-Platform GUI Programming with wxWidgets, Prentice Hall, 2006.

STARDEX: STARDEX, Downscaling climate extremes, Tech. rep., STARDEX consortium, 2005.

Talagrand, O., Vautard, R., and Strauss, B.: Evaluation of probabilistic prediction systems, in: Proceedings of the ECMWF Workshop on Predictability, pp. 1–25, 1997.

Teweles, S. and Wobus, H. B.: Verification of prognostic charts, Bulletin of the American Meteorological Society, 35, 455–463, 1954.

Thevenot, N.: Prévision quantitative des précipitations: Adaptation par une méthode d'Analogie de la prévision d'ensemble du CEPMMT. Aspects opérationnels, Ph.D. thesis, INP Grenoble, 2004.

Toth, Z.: Long-range weather forecasting using an analog approach, Journal of Climate, 2, 594–607, http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(1989)002{%}3C0594:LRWFUA{%}3E2.0.CO{%}3B2, 1989.

Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., Mcnally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 reanalysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961–3012, doi:10.1256/qj.04.176, http://doi.wiley.com/10.1256/qj.04.176, 2005.

Vallée, J. L.: Précipitations sur le Sud-Ouest du Massif Central et l'Est des Pyrénées. Optimisation du modèle EDF/DTG de prévision par recherche d'analogues., Tech. rep., Rapport de stage de fin d'études de l'Ecole de la Météorologie Nationale, 1986.

Wetterhall, F.: Statistical Downscaling of Precipitation from Large-scale Atmospheric Circulation: Comparison of Methods and Climate Regions, Ph.D. thesis, Uppsala University, 2005.

Wetterhall, F., Halldin, S., and Xu, C.-Y.: Seasonality properties of four statistical-downscaling methods in central Sweden, Theoretical and Applied Climatology, 87, 123–137, doi:10.1007/s00704-005-0223-3, http://link.springer.com/10.1007/s00704-005-0223-3, 2007.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences: An Introduction. 2nd edition, Elsevier, 2006.

Wilson, L. and Vallée, M.: The Canadian updateable model output statistics (UMOS) system: Design and development tests, Weather and Forecasting, 17, 206–222, 2002.

Wilson, L. J. and Yacowar, N.: Statistical weather element forecasting in the Canadian Weather Service, in: Proceedings of the WMO Symposium on Probabilistic and Statistical Methods in Weather Forecasting, pp. 40–406, Nice, 1980.

Woodcock, F.: On the use of analogues to improve regression forecasts, Monthly Weather Review, 108, 292–297, http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1980)108{%}3C0292{%}3AOTUOAT{%}3E2.0.CO{%}3B2, 1980.

Zorita, E. and Storch, H. V.: The analog method as a simple statistical downscaling technique: comparison with more complicated methods, Journal of climate, 12, 2474–2489, http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(1999)012{%}3C2474{%}3ATAMAAS{%}3E2.0.CO{%}3B2, 1999.