https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development Discussions

Assessment of stochastic weather forecast of precipitation near European cities, based on analogs of circulation

Meriem Krouma [1,2], Pascal Yiou [2], Céline Déandreis [1], and Soulivanh Thao [2]

[1]ARIA Technologies, 8 Rue de la Ferme, 92100 Boulogne-Billancourt, France
[2]Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, IPSL '&' Université Paris-Saclay, 91191 Gif-sur-Yvette, France

Correspondence: M.Krouma (meriem.krouma@lsce.ipsl.fr)

Abstract.

In this study, we aim to assess the skill of a stochastic weather generator (SWG) to forecast precipitation in several cities of Western Europe. The SWG is based on random sampling of analogs of the geopotential height at 500 hPa. The SWG is evaluated for two reanalyses (NCEP and ERA5). We simulate 100-member ensemble forecasts on a daily time increment. We evaluate the performance of SWG with forecast skill scores and we compare it to ECMWF forecasts. Results show significant positive skill score (continuous rank probability skill score and correlation) for lead times of 5 and 10 days for different areas in Europe.

We found that the low predictability of our model is related to specific weather regimes, depending on the European region. Comparing SWG forecasts to ECMWF forecasts, we found that the SWG shows a good performance for 5 days. This performance varies from one region to another. This paper is a proof of concept for a stochastic regional ensemble precipitation forecast. Its parameters (e.g. region for analogs) must be tuned for each region in order to optimize its performance.

1 Introduction

Ensemble weather forecasts were designed to overcome the issues of meteorological chaos, from which small uncertainties in initial conditions can lead to a wide range of possible trajectories (Sivillo et al., 1997; Palmer, 2000). Hence, from a sufficiently large ensemble of initial conditions, it is in principle possible to sample the probability distribution of future states of the system.

Forecasts issued by meteorological centers are obtained by computing several simulations with perturbed initial conditions, in order to sample uncertainties. Those experiments are rather costly in terms of computing resources and are generally limited to a few tens of members (Hersbach et al., 2020; Toth and Kalnay, 1997), which can hinder a proper estimate of probability distributions of trajectories. Moreover, obtaining information at local spatial scales can be difficult because the horizontal resolution of the atmospheric models is around 18 km, e.g. for the European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble forecast system.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions

From a mathematical point of view, computing the probability distribution of the trajectories of a (deterministic) system makes the underlying assumption that the system behaves like a stochastic process, for which statistical properties are defined naturally (Ruelle, 1979; Eckmann and Ruelle, 1985). This has justified the development of stochastic weather generators (SWG), which are stochastic processes that emulate the behavior of key climate variables (Ailliot et al., 2015). The advantages of stochastic models are a relative simplicity of implementation and a low computing cost. The challenge of their development is to verify that the behavior of the simulations are realistic, according to well-defined criteria (van den Dool, 2007; Jolliffe and Stephenson, 2011).

The first stochastic weather generators were devised to simulate rainfall occurrence by Gabriel and Neumann (1962) and to simulate rainfall amounts by Todorovic and Woolhiser (1975). SWGs were developed and used to estimate the probability distributions of climate variables such as temperature, solar radiation, and precipitation through extensive simulations (Richardson, 1981).

Stochastic weather generators can be useful complements to atmospheric circulation models, in order to simulate large ensembles of local variables, as they can be calibrated for small spatial scales comparing to numerical models (Ailliot et al., 2015). This explains their wide applications in impact studies.

A successful simulation with SWG relies on the choice of inputs. One of them consists on the use of the atmospheric circulation as a predictor for other local variables. The (loose) rationale for this choice is that the circulation is modeled by prognostic equations (Peixoto and Oort, 1992), that drive the other physical variables. Therefore the primitive equations of the atmosphere (Peixoto and Oort, 1992, Chap. 3) suggest that reproducing temporal variability on daily time scales requires considering circulation variables. The influence of large-scale circulation on local climate variables has been proven in previous studies such as the influence of atmospheric circulation on eastern Mediterranean Basin and Greece precipitation (Xoplaki et al., 2000; Türkes et al., 2002). Similar influences have been found on precipitation and temperature over the North Atlantic region (Jézéquel et al., 2018b).

Analogs of circulation were initially designed to provide "model-free" forecasts, by assuming that similar situations in atmospheric circulation may lead to similar local weather conditions (Lorenz, 1969). The potential to simulate large ensembles of forecasts temperature with circulation analogs was explored by Yiou and Déandréis (2019), by considering random resamplings of K best analogs (rather than only considering the best analogue). This has lead to the development of a SWG in "predictive" mode, which uses updates of reanalysis datasets (Kistler et al., 2001) as input.

Alternative systems of analogs to forecast precipitation have been proposed by Atencia and Zawadzki (2015). Those systems are based on analogs of precipitation itself. Such systems are very efficient for nowcasting, i.e. forecasting precipitation within the next few hours. Considering the atmospheric circulation analogs allows to focus on longer time scales.

Yiou and Déandréis (2019) evaluated ensemble forecasts of the analog SWG with classical probability scores against climatology and persistence. Reasonable scores were obtained up to 20 days. Through this study, we aim to assess the skill of this SWG to forecast precipitation in different areas of Europe and for different lead times. The previous study on this forecast tool was a proof of concept for temperature. In this study we will adapt the parameters of the analog SWG to optimize the

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions

simulation of European precipitations. We then analyse the performance of this SWG for lead times of 5 to 20 days, with the forecast skill scores used by Yiou and Déandréis (2019).

We will evaluate the seasonal dependence of the forecast skills of precipitation and the conditional dependence to weather regimes. Finally, comparisons with medium range precipitation forecasts from the ECMWF will be performed.

The paper is divided as follows: Section 2 is dedicated to describe the data used for the experiments. Section 3 explains the methodology (analogs and stochastic weather generator) and experimental set up. Section 4 details results of simulations and the evaluation of the ensemble forecast. Section 5 contains the main conclusions of the analyses.

2 Data

Daily precipitation data was obtained from the European Climate Assessment and Data (ECAD) project (Klein-Tank et al., 2002) for four locations in western Europe (Berlin, Madrid, Orly, Toulouse), which are subject to various meteorological influences (Figure 1). The data were available at a daily time step from 1948 to 2019. The choice of those stations was based on the availability of large and common period of observations with a low rate of missing data (less than 10%).

We recovered the geopotential height at 500 hPa (Z500) and sea level pressure (SLP) fields from the reanalysis of the National Centers for Environmental Prediction (NCEP: Kistler et al. (2001)) with a spatial resolution of 2.5° $\times$ 2.5° from 1 January 1948 to 31 December 2019.

We also used the atmospheric reanalysis (version 5) of the European Centre for Medium-Range Weather Forecasts (ECMWF) (ERA5; Hersbach et al. (2020)). ERA5 data are available from 1979 to present with a horizontal resolution of 0.25° $\times$ 0.25°. There are fundamental differences between the two reanalyses, in the atmospheric models, assimilated data, and assimilation schemes.

Most of the computations were performed with the NCEP reanalysis because it covers a longer period. The ERA5 data were used for verification purposes. The higher spatial resolution of ERA5 has little impact on the results since the Z500 fields yield a smooth spatial variability.

We considered the daily averages of Z500 from NCEP and ERA5, over the region covering 30°W – 20°E and 40°– 60°N to compute circulation analogs. Daily averages of SLP were used over the region covering 80°W – 20°E and 30° – 70°N to define weather regimes.

In order to assess the predictive skill of our precipitation forecast model, a comparison with another forecast was made. There are many available datasets that can be used for deriving this information. We considered the ECMWF ensemble forecast dataset system 5 (Hersbach et al., 2019). It is a daily gridded dataset interpolated over Europe to provide information covering the all the domain. Data are available through the Copernicus Climate Data Store including forecasts created in real-time (since 2017) and hindcast forecasts from 1993 to 2019 (Hersbach et al., 2019). The data are provided at an hourly time step with a horizontal resolution of 0.25° $\times$ 0.25°. We considered the grid points that include Berlin, Orly, Toulouse and Madrid, which were identified in the ECAD database.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions

3 Methodology

3.1 Analogs

The first step is to build a database of analogs of the atmospheric circulation. We outline the procedure of Yiou and Déandréis (2019), summarized in Figure 1a. For a given day t, we determine the similarity of Z500 for all days t' that are in a year different from t, but within 30 calendar days of t. This similarity is quantified by a Euclidean distance (or root mean square) between the daily Z500 maps. Other types of distances are possible (Blanchet et al., 2018), but the expected impact on the results is often marginal (Toth, 1991). We believe that the simplicity of the a Euclidean distance makes it more robust to changes in horizontal resolution (from NCEP to ERA5), compared to more sophisticated distances that include local spatial gradients, which would require adjustments and additional tuning. This choice can be left open for future fine tuning, depending on the region.

For each day t, the $K = 20$ best analogs are the days t' for which the distance is the smallest. We compute the spatial rank correlation between the Z500 best analogs and the Z500 at time t for a posteriori verification purposes.

As a refinement over the study of Yiou and Déandréis (2019), a time embedding of 4 days was used for the search of analogs dates. This means that the field $X(t)$ for which we compute analogs is $X(t) = (Z500(t), Z500(t+1), \ldots, Z500(t+3))$. This ensures that temporal derivatives of the atmospheric field are preserved (Yiou et al., 2013). Hence the distance that is optimized to find analogs of the $Z500(x,t)$ field is:

$$D(t,t') = \left[\sum_{x} \left(\sum_{i=0}^{3} |Z500(x,t+i) - Z500(x,t'+i)| \right)^2 \right]^{\frac{1}{2}}, \tag{1}$$

where x is a spatial index.

We consider different geographic domains as showed in Figure 1 for the computation of analogs and weather regimes. The computation of circulation analogs was performed with the "blackswan" Web Processing Service (WPS, Hempelmann et al. (2018)). The "blackswan" WPS is an online tool that helps computing circulation analogs on various datasets (reanalyses, climate model simulations) with a user friendly interface.

3.2 Configuration of stochastic weather generator

As detailed by Yiou and Déandréis (2019), the stochastic weather generator (SWG) we use is based on a random reshuffling of circulation analogs. For a given day t_0, we perform an ensemble of simulations until a lead time $t_0 + T$, with $T \in \{5, 10, 20\}$ days. In order to go from $t \in [t_0, t_0 + T]$ to $t + 1$, we sample one analog (out of $K = 20$) at day t,with a probability that is proportional to / depends on the correlation between the Z500 fields at time t and at the analog date t'. Then we exclude samples that are in $[t_0, t_0 + T]$, so that this procedure reflects a hindcast forecast from t_0.

This procedure is iterated from $t = t_0$ to $t = t_0 + T$, to generate one trajectory. It is then repeated $N = 100$ times to generate an ensemble of daily trajectories starting at t_0. Each daily trajectory is time averaged between t_0 and $t_0 + T$. In this paper, we hence analyse the properties of an ensemble forecast of the mean precipitation between t_0 and $t_0 + T$.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions

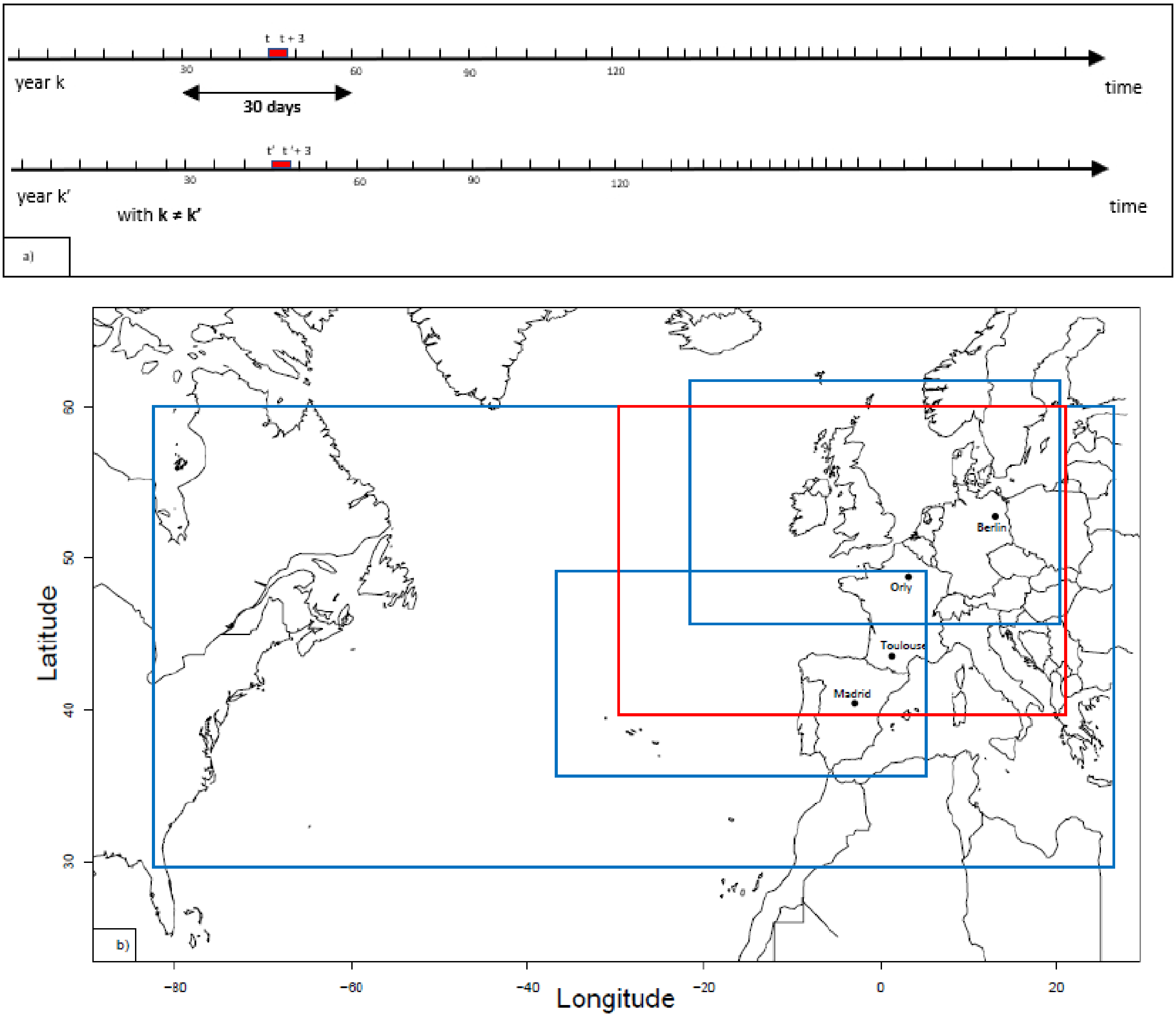

Figure 1. (a) For each day t at year k, we chose an analog day t' of a similar atmospheric circulation condition at year k' selected among database of n years; the variation of Z500 at 4 days after t' are considered; t' is selected within 30 calendar days of t'. (b) Domains of computation of analogs, we computed analogs over different domains, each one includes a part of the Atlantic and focus in a part of Western Europe, in order to test the sensitivity of our model to different geographic areas, the optimising area was [30°W-20°E; 40°-60°N], indicated by the red rectangle.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Then t_0 is varied from Jan. 1st 1948 to Dec. 31st 2019, with increments of $\approx T/2$. This produces a stochastic ensemble hindcast forecast of precipitation and atmospheric circulation (Z500). We will focus on the properties of the properties of precipitation forecast.

For each day t_0, we also compute persistence and climatological forecasts for the average between t_0 and $t_0 + T$. The persistence forecast consists in using the average value between $t_0 - T$ and t_0. The climatological forecast take the climatological average between t_0 and $t_0 + T$. Both "control" forecasts are randomized by adding a small Gaussian noise, whose standard deviation is estimated by bootstrap over T long intervals. Hence we generate ensembles of persistence and climatological forecasts that are consistent with observations (Yiou and Déandréis, 2019).

3.3 Domains of computation of analogs

We computed sample trajectories of the SWG for the four domains outlined in Figure 1b. We used different domains in order to find an optimal region which allows verifying the relationship between precipitation and Z500. Each domain includes a part of the Atlantic and a part of western Europe. The widest domain of the coordinates 80°W – 20°E and 30°– 70°N did not give good results for precipitation forecasting for the four studied areas in western Europe, while the other two smaller domains (in blue) gave good forecasts for specific locations. However, in order to make a forecast for the whole of Europe, we found that the domain for Z500 analogs that optimizes precipitation correlations is 30°W – 20°E and 40° – 60°N. We, therefore, kept this domain for the subsequent analyses.

To sum up, the protocol is as follows:

- Analogs computed from Z500 over region 30°W – 20°E; 40° – 60°N (red rectangle in Figure 1) from 1948 to 2019.
- 100 trajectories generated from random analogs of Z500 and averaged for a lead times of $T = \{5, 10, 20\}$ days are used to simulate precipitation. The simulation of the precipitation was done from 1948 to 2019.
- comparisons with reference forecasts ("persistence" and "climatology").

The simulations of this stochastic model will be called "SGW forecasts", as opposed to ECMWF forecasts.

3.4 Forecast Verification

Forecast verification is the process of determining the statistical quality of forecasts. A wide variety of ensemble forecast verification procedures exists. They involve measures of the relationship between a set of forecasts and corresponding observations. To assess the quality of precipitation forecasts, we compute indicators such as the Correlation and Continuous Rank Probability Skill Score (CRPSS) for each lead time T, for different seasons and months.

The temporal rank correlation is calculated between the precipitation observations and the median of 100 simulations. This simple diagnostic is often used to assess forecast skills of indices (Scaife et al., 2014).

The continuous ranked probability score (CRPS) represents the most used score for probability forecast verifications Ferro (2007). It is sensitive to the distance between forecast and observation probability distributions.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

If the ensemble forecast yields a probability distribution $P(x)$, the CRPS measures how the observations the probability distribution of x (Hersbach, 2000).

The CRPS is computed as:

$$CRPS(P,x_a) = \int_{-\infty}^{+\infty} (P(x) - P_a(x))^2 dx, \tag{2}$$

where P_a represents the Heaviside function of the occurrence of x. The decomposition and properties of the CRPS have been investigated by (Ferro, 2007; Hersbach, 2000; Zamo and Naveau, 2018). A perfect forecast would have a CRPS equal to 0, but the CRPS value obviously depends on the units of the variable to forecast. It is hence difficult to compare CRPS values for temperature and precipitation, within the same ensemble forecast. This issue is also acute for non Gaussian variables with heavy tails (Zamo and Naveau, 2018), so that the interpretation of a given CRPS value might not always be informative.

One way of circumventing this difficulty is to compare CRPS values to reference forecasts, such as persistence or seasonality. The continuous rank probability skill score (CRPSS) is a normalization of Eq. (2) with respect to such a reference.

The CRPSS is hence computed by:

$$CRPSS = 1 - \frac{CRPS}{CRPS_{ref}} \tag{3}$$

where the $CRPS_{ref}$ is the CRPS of reference forecast (climatology or persistence). The CRPSS is interpreted as a percentage of improvement over a reference forecast.

The values of the CRPSS varies between $-\infty$ and 1. The forecast is considered to be an improvement over the reference when the CRPSS value is close to 1 (i.e. when the CRPS is 0). Values of CRPSS equal to 0 indicates no improvement over the reference. Values inferior to 0 mean that the forecast is worse than the reference.

We use the CRPSS values to determine the maximum lead time T for which the SWG forecast is better than references. Then the SWG assessments will use the CRPS and directly compare the probability distributions of precipitation ensemble forecasts.

4 Results

4.1 Sample of forecast

As an example, we illustrate the behavior of the trajectories in Orly for the summer and winter of 2002. Figure 2 shows the observed and simulated values of precipitation for lead times of 5 and 10 days for summer (June–July–August: JJA) and winter (December–January–February: DJF), for Orly precipitation data. We observe significantly positive correlations between observed values and the median of the forecasts, for the four data sets as represented in Table 1 . The correlation is generally smaller in the summer than in the winter.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

The correlation skill is low for some extremes values of precipitation. For a lead time of 10 days, SWG simulation still show capacity to predict precipitation especially for winter with a correlation equal to 0.23 (Orly), 0.30 (Berlin), 0.43 (Madrid), 0.31 (Toulouse).

Table 1. Correlation between observations and the median of 100 simulations for both seasons winter (DJF) and summer (JJA) for a lead time of 5 days

Location	**Correlation DJF**	p-values	**Correlation JJA**	p-values
Berlin	0.42	2.2^{-16}	0.22	2.2^{-16}
Madrid	0.58	2.2^{-16}	0.29	2.2^{-16}
Orly	0.58	2.2^{-16}	0.23	2.2^{-16}
Toulouse	0.43	2.2^{-16}	0.18	2.2^{-16}

We observe that the 5th and 95th quantiles of simulations include the different values of observations. This heuristically confirms the good skill of SWG to forecast precipitation from Z500 for several seasons (winter and summer) in several locations for $T = 5$ and $T = 10$ day lead times.

The difference of the forecast correlation skills between the six studied locations may be related to the variation of the local climate from one region to an other. The studied areas are in different climate types according to Köppen-Geiger's climate classification map (Peel et al., 2007). From the south western side of Europe, Madrid is in the arid zone (Peel et al., 2007), which indicates that convective rains are less significant, so that the origin of precipitation might be the result of humidity coming from the Atlantic. Conversely, Berlin is located in a cold zone characterised by warm summer and the absence of a dry season (Peel et al., 2007), so that the precipitation could be the result of both convective rains and Atlantic humidity.

In this paper, we decided (for simplicity) to use the same analogs to forecast precipitation for those four stations. A refinement of the analog regions would be necessary when focusing on Madrid vs. Berlin.

4.2 Forecast probability skill

We first computed the CRPSS for precipitation in Orly for lead times from 5 to 20 days (Figure 3). The skill score was also computed for Berlin, Madrid and Toulouse, as shown in illustrations are represented in Figure A1. We represent skill scores for January and July in order to evaluate the skill of the SWG to predict precipitation in both seasons (winter and summer). For comparison purposes, SWG simulations are obtained using analogues computed from reanalyses on the NCEP and ERA5 reanalyses.

Comparing their skill scores, we found that CRPSS and correlation between observations and simulations are positive in both cases, and showing positive improvement comparing to persistence and climatology forecasts. The CRPSS and correlation for simulations with analogs of NCEP are slightly higher than with ERA5, due to the longer length of the NCEP reanalysis, which has a better potential to find good analogues.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

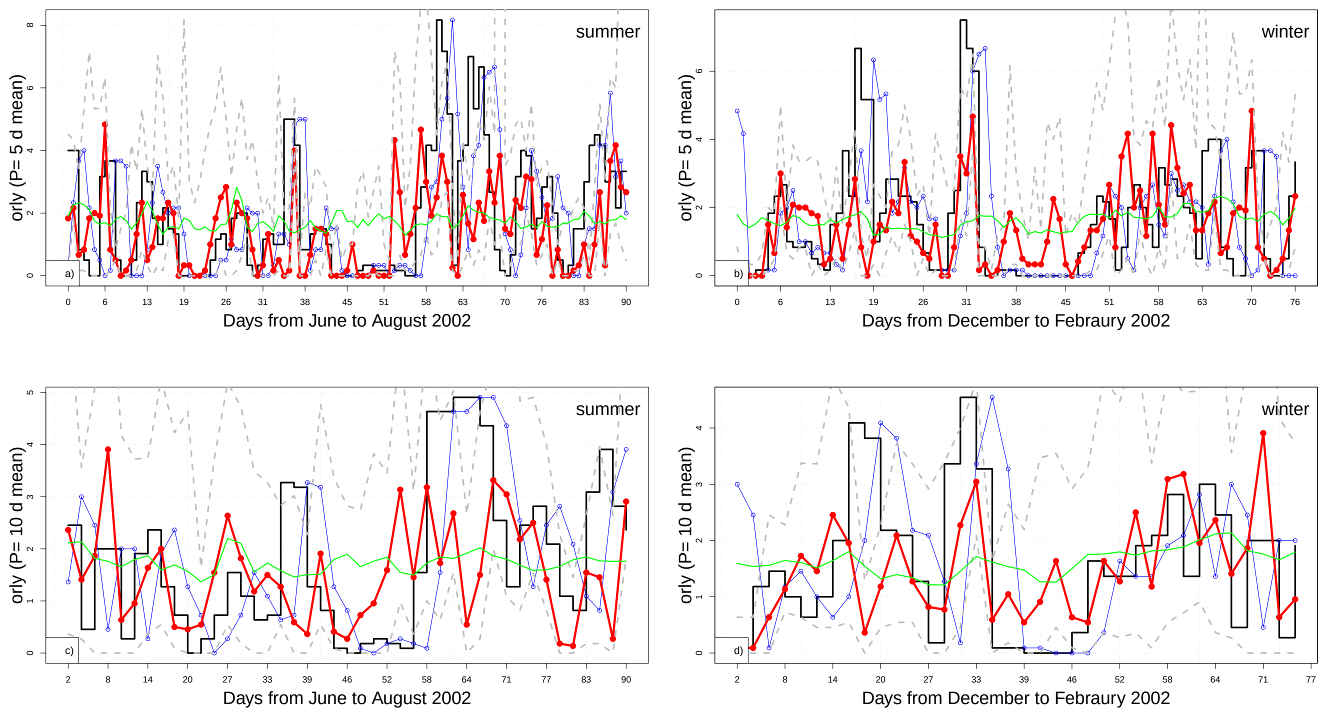

Figure 2. Time series of analogue ensemble forecasts for 2002, for lead times of 5 days (top) and 10 days (bottom) for summer (June to August) a) and c) and winter (December to February) b) and d) for Orly. The median of 100 simulations is represented by red line. Black line represent observations values. Gray lines represent the 5th and 95th quantiles. Blue lines represent persistence forecasts and green lines represent the climatology forecasts.

We determined that the SWG simulations showed better skills for the geographic domain outlined in red, in Figure 1b) as it allows to make forecasts for all the studied areas and we find that the skill scores over this geographic domain remained the highest ones as represented in the following Table 2. Therefore, we focus on SWG simulations with analogs from the NCEP reanalysis in the sequel.

The CRPSS for persistence and climatology references show positive values for lead times of up to 20 days (Figure 3). The values of CRPSS with persistence reference (represented by squares) decrease with lead times, showing high values over 5 days. The CRPSS for climatology (triangles) show lower values, although positive. The correlation skill is positive for both seasons but higher in winter (January) than in summer (July). For a lead time of 5 days, the correlation is equal to 0.59 for Madrid, 0.50 for Berlin and to 0.40 for Toulouse. For a lead time of 10 days, it is equal to 0.42 for Madrid, 0.30 for Berlin and to 0.41 for Toulouse.

The SWG was tested in previous work Yiou and Déandréis (2019) to forecast North Atlantic oscillation (NAO) and temperature in western Europe. Comparing the performance of the SWG to forecast those different meteorologic variables, we

Table 2. Correlation between observations and the median of 100 simulations for the winter (DJF) for the different studied domains represented in figure 1b, with the coordinates 80°W – 20°E ; 30° – 70°N for the largest one (blue) and 30°W – 20°E ; 40° – 60°N for the red rectangle for a lead time of 5 days.

Location	The domain 80°W – 20°E ; 30° – 70°N		The domain 30°W – 20°E ; 40° – 60°N	
	Correlation	95% confidence interval	Correlation	95% confidence interval
Berlin	0.32	0.30 – 0. 35	0.58	0.55 – 0.60
Madrid	0.35	0.33 – 0. 39	0.66	0.64 – 0.68
Orly	0.39	0.37 – 0. 41	0.56	0.54 – 0.59
Toulouse	0.34	0.31 – 0.36	0.59	0.56 – 0.61

notice that the model shows good performance to forecast the temperature and NAO in the winter, also the best performance of the model is at a lead time of 5 days. We find that the skill scores (CRPSS and correlation) decrease with lead of times. The forecast skill of the SWG shows variability from one locations to another. However, the model was able to forecast temperature until 40 days in Berlin, Orly, Toulouse and De Blit with positive skill scores.

From a visual inspection of the CRPSS and correlations, we chose to focus on lead times of $T = 5$ days, for which the correlation exceeds 0.5 in the winter. It is rather low in the summer, due to convective events leading to a high precipitation variability (from no rain to very high values). Correlation scores become barely significant for lead times of 20 days, so that, like temperature, the SWG should not be used beyond that horizon.

4.3 Relation between weather regimes and CRPS

In this subsection, we investigate the role of North Atlantic weather patterns on the CRPS of the SWG precipitation simulations.

We use weather regimes, which are defined as large-scale quasi stationary atmospheric states. They are characterised by their recurrence, persistence and stationarity (Michelangeli et al., 1995). They help describing the features of the atmospheric circulation. Surface variables like temperature and precipitation are largely correlated with weather regimes (van der Wiel et al., 2019) .

The North Atlantic weather regimes were computed with the procedure of (Yiou et al., 2008), with the NCEP reanalysis. The first 10 principal components of SLP (large region in Figure 1b) are classified with a k-means algorithm onto four classes, over a reference period between 1970 and 2010. The procedure is repeated 100 times with random k-means initialization, so as to determine the most probable classification. Figure 4 shows four weather regimes for each season (winter and summer) that are coherent with the literature (Cassou; Ghil et al., 2008; Kimoto, 2001; Michelangeli et al., 1995)

The winter weather regimes are the Scandinavian blocking (BLO), Atlantic ridge (AR), negative phase of the North Atlantic oscillation (NAO-) and Zonal flow (ZO). The summer weather regimes are the negative phase of the NAO (NAO-), Atlantic ridge (AR), Scandinavian blocking (BLO) and Atlantic low (AL). The regimes are not the same in both seasons, due to the seasonality of the large scale atmospheric circulation.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions

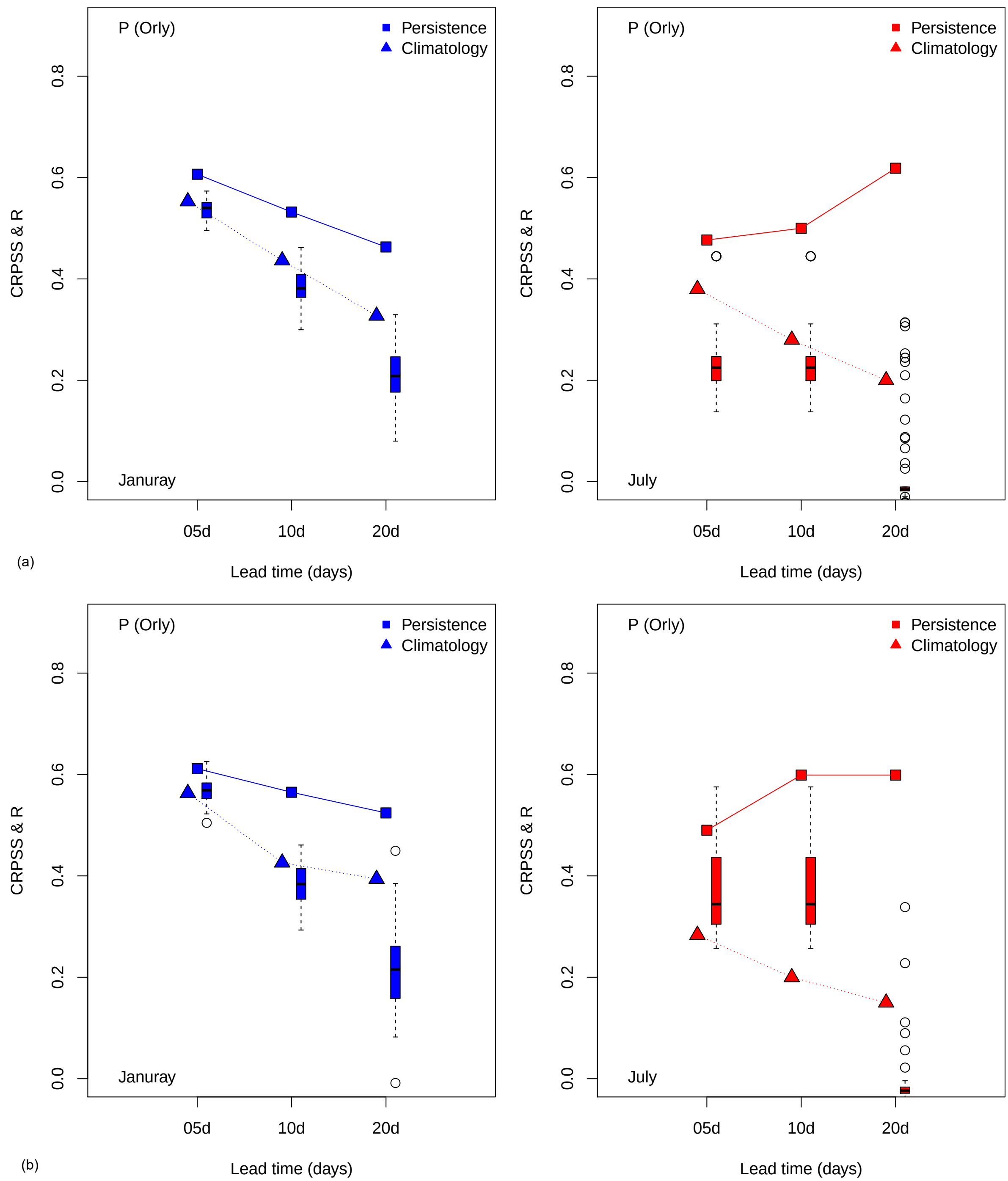

Figure 3. Skill scores for the precipitation of Orly for lead times of 5, 10, 20 days for January (blue) and July (red) for analogs computed from reanalyses of (a) NCEP and (b) ERA5. Squares indicate CRPSS where the Persistence is the baseline, triangles indicates CRPSS where the climatology is the reference, and box-plots indicates the correlation between observation and median of 100 simulations.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development Discussions

For each day (in winter and summer) between 1948 and 2019, we classify the SLP by minimizing the root mean square to four reference (1970–2010) weather regimes.

We evaluate the influence of weather regimes on the SWG forecast quality by plotting the probability distribution of CRPS values *conditional* to each weather regime. This is done separately for "good" forecasts (low CRPS values) and "bad" forecasts (high CRPS values). Weather regimes were considered at the time of the forecast at $t = t_0 + T$.

Hence identify those two classes of predictability from CRPS values:

- Low predictability is related to high values of CRPS that exceed the 75th quantile,
- High predictability is linked to low values of CRPS, below the 75th quantile.

This procedure helps identifying atmospheric patterns that could lead to low/high predictability with the SWG model.

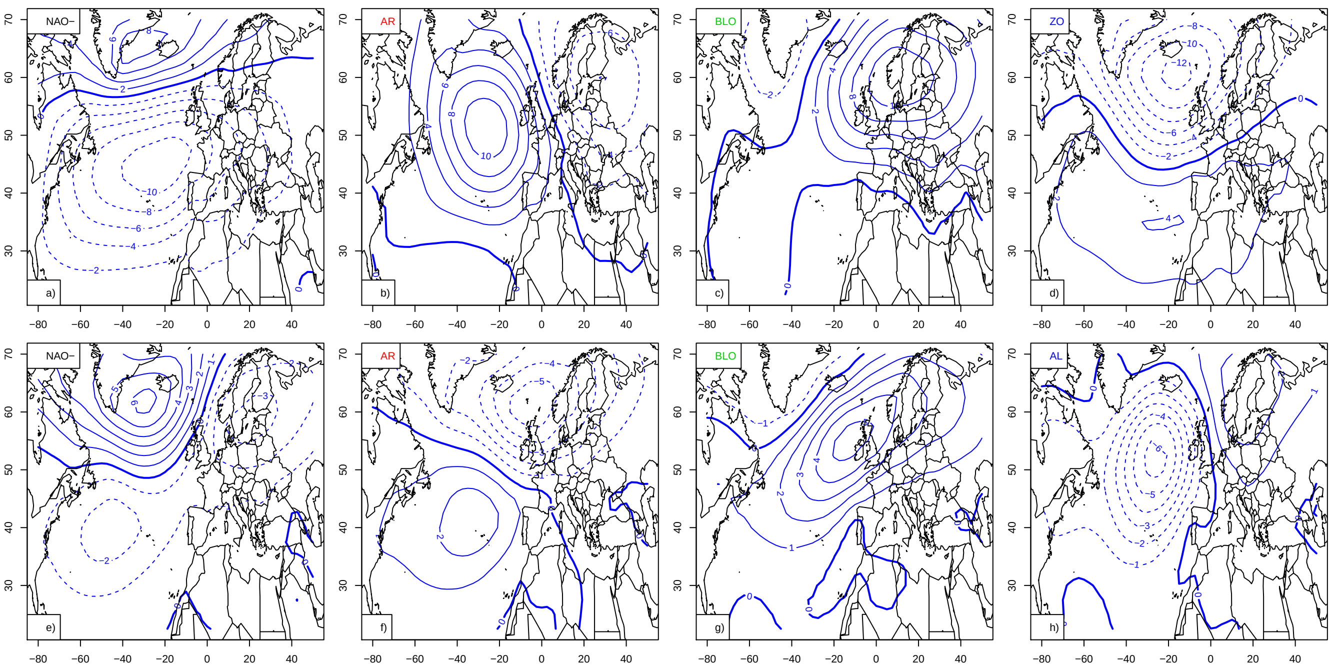

Figure 4. Weather regimes over Europe from SLP fields: North Atlantic oscillation (NAO-),the Atlantic ridge (AR), the Scandinavian blocking (BLO),and Atlantic zonal (NAO+). The figure summarises the different states of the atmosphere during summer (a to d) and winter (e to h). It indicates the low and the high pressure over Europe and the direction of flow from the west (Atlantic) to the east. The isolines show seasonal anomalies with respect to a June-July-August and December-January-February means, in hPa with 2 hPa increments.

We find a small, albeit significant, influence of specific weather regimes on the CRPS distribution. This relation is represented in Figure 5a for Orly and for the rest of the studied stations in Figure B1. Good forecasts (low quantiles of CRPS) are mainly related to the Scandinavian blocking for Berlin and Orly in winter and summer, while they are related to the Atlantic ridge

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

weather regime in the winter and to the Atlantic Low in the summer for Madrid and to the Atlantic ridge in the winter and to NAO- in the summer for Toulouse.

The low quality forecasts (high quantiles of CRPS) are related to the Atlantic ridge in both seasons for Berlin, to Atlantic Zonal (NAO+) in the summer for Orly (Figure 5b), to the Atlantic Zonal in the winter for Madrid and to NAO- for Toulouse.

This relation between predictability (or the CRPS distribution) and weather regimes is consistent with previous work (Faranda et al., 2017). We find that the sensitivity of the forecast to weather regime is larger for low values of CRPS and in the winter. The sensitivity of forecast skill to weather regimes is rather small on average, even for low lead times.

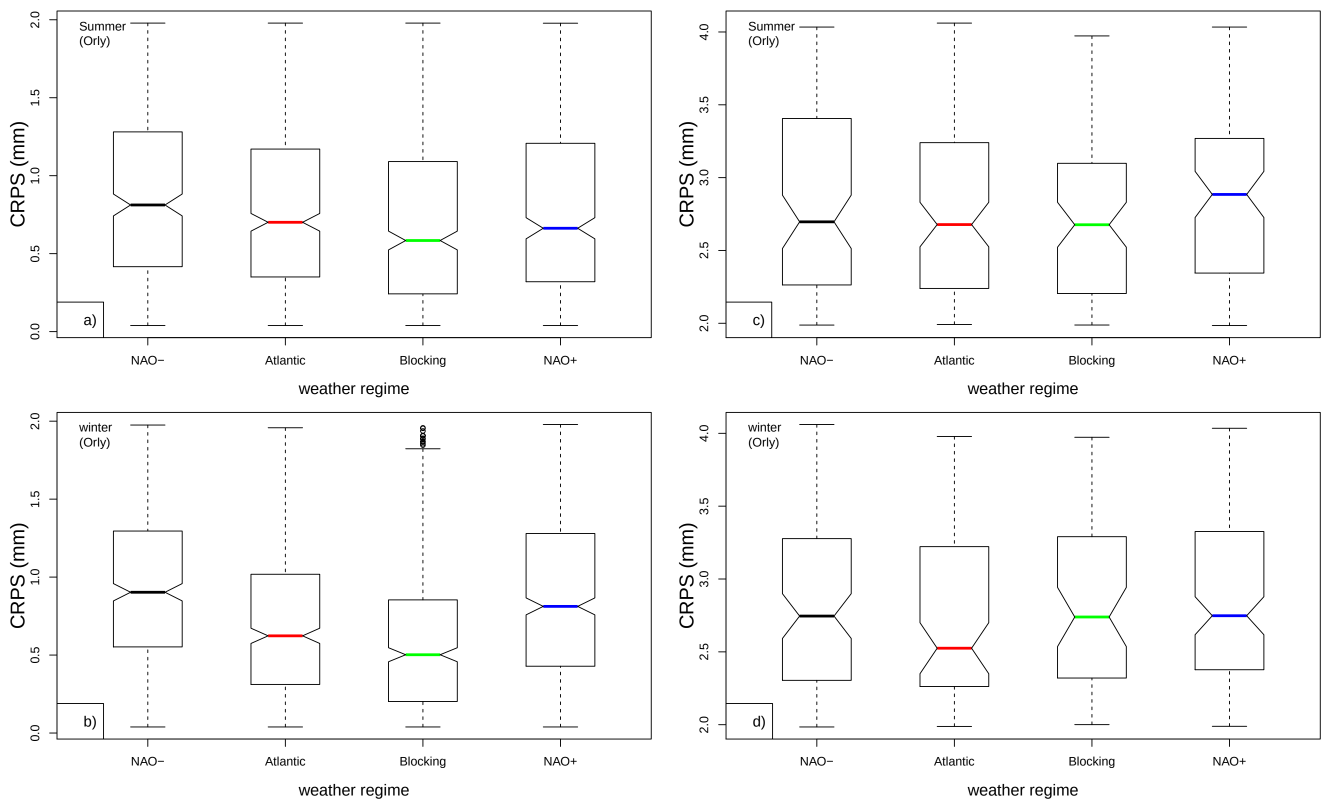

Figure 5. Evaluation of the forecast quality by weather regimes. (a and b) low values of CRPS and (c and d) high values of CRPS for summer and winter.

4.4 Comparison with ECMWF forecast

We first compared the CRPSS of SWG forecasts for winter and summer with the CRPSS of ECMWF forecasts.

The CRPSS of ECMWF forecast is operationally computed for different lead times going from 1 day to 10 day for the precipitation. It uses the climatology as a reference (Haiden et al., 2018). The values of CRPSS for Europe for 2020 decrease with lead times. They are about 0.16 in the summer (JJA) and 0.25 in the winter (DJF) for a lead time of $T = 5$ days. The values of CRPSS for ECMWF for both seasons are computed over whole Europe (Haiden et al., 2019).

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions
Open Access
EGU

Table 3. The CRPSS of SWG for a lead time of $T = 5$ days for winter (DJF) and summer (JJA)

Location	CRPSS DJF	CRPSS JJA
Berlin	0.42	0.21
Madrid	0.57	0.25
Orly	0.53	0.23
Toulouse	0.41	0.24

The CRPSS of SWG for a lead time of $T = 5$ days showed in Table 3, and this suggests that the predictive skill of SWG is qualitatively promising for short lead times, compared with ECMWF forecasts.

A quantitative comparison was made by comparing the empirical cumulative distribution function (ECDF; Hersbach,2000) of the CRPS of ECMWF and SWG forecasts for 5 days (Figure 6). We found that the values of CRPS of ECMWF forecast and SWG forecast are 80%, 39% 50% and 40 % equal or near to zero for respectively Orly, Berlin, Madrid and Toulouse, which indicates the small variations of the CRPS.

We used the Kolmogorov-Smirnov test (von Storch and Zwiers, 2001) to compare the probability distributions of the CRPS of SWG and ECMWF forecasts.The null hypothesis was defined as the two series of CRPS have the same distribution. It was verified with $p.values = 2.2^{-16}$. We found that the maximum distance between both ECDFs is ≈ 0.2. This confirm the overall good skill of the SWG to forecast precipitation, compared to ECMWF. One notable difference between SWG and ECMWF forecasts is that although the proportion of CRPS values close to zero is higher in ECMWF, the CRPS for the worse forecasts are much higher than those of SWG.

Finally, we computed the CRPSS for ECMWF forecasts taking as a baseline the CRPS of SWG (Figure 7).

This evaluates the added value of the deterministic ECMWF forecast over the SWG forecast. We found that the SWG is still showing a positive improvement especially for a lead time of 5 days for the different studied areas.

5 Conclusions

In this work, we showed the performance of a stochastic weather generator (SWG) to simulate precipitation over different locations in western Europe and for various times scales from 5 to 20 days. The input of our model was analogs of geopotential heights at 500 hPa (Z500). The choice of such input was made in order to evaluate the impact of large scale circulation on local weather variables. SWG showed a good skill to predict precipitation for a lead time of 5 and 10 days from analogues of Z500.

This study complements the work of Yiou and Déandréis (2019), for precipitation. We explored the sensitivity of the SWG model on analogs computed from different geographical areas and from different reanalyses (ERA5 and NCEP). We found that the NCEP reanalyses provide better performances for simulations, due to its larger length (≈ 70 years in NCEP vs. ≈ 40 years in ERA5). Therefore the length of the analog database does make a difference, as already suggested by Jézéquel et al. (2018a).

We evaluated the relation between the quality of the forecast and weather regimes over Europe, we found that low and high predictability was slightly related to specific weather regimes, although this dependence is weak.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions

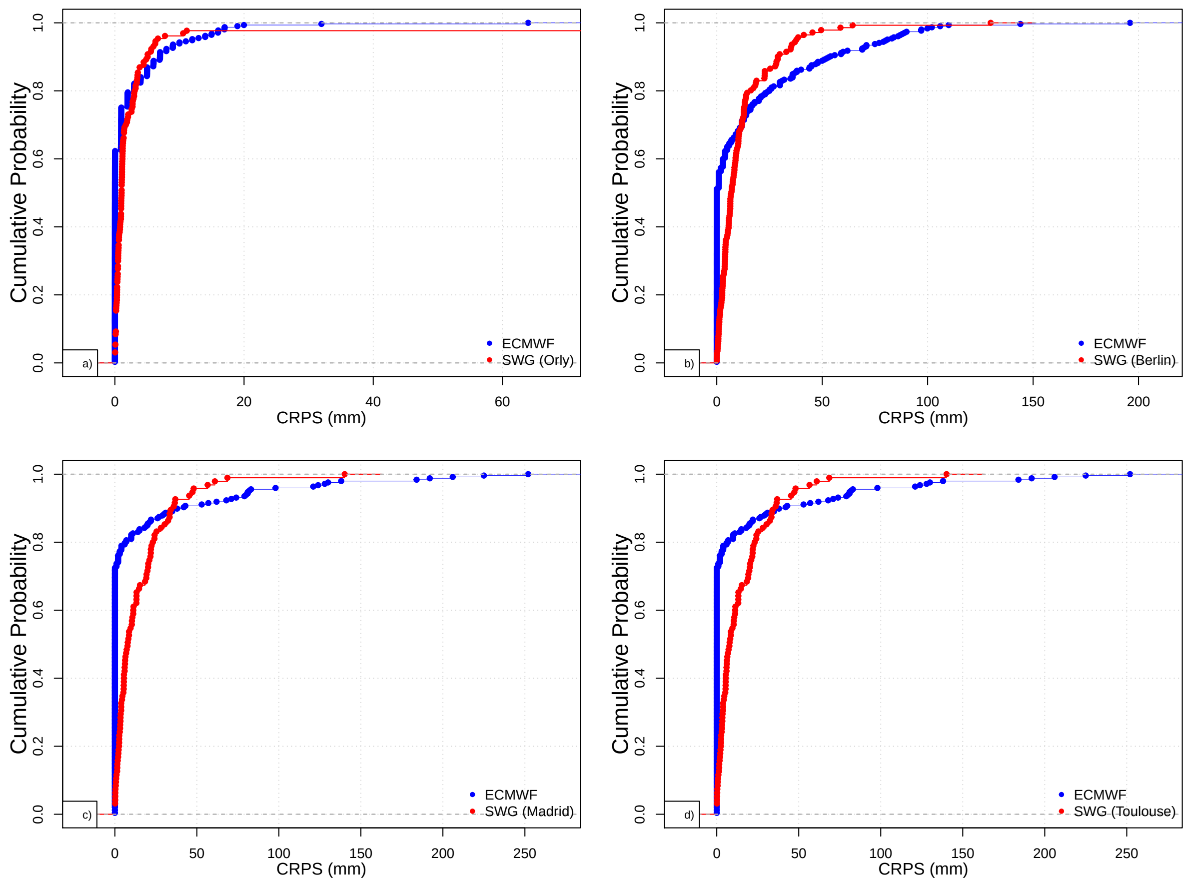

Figure 6. Empirical cumulative distribution function of the CRPS of ECMWF and SWG forecasts for 5 days for Orly (a), Berlin (b), Madrid (c) and Toulouse (d)

A comparison with the ECMWF forecast system over Western Europe confirmed the good performance of the SWG quantitatively and qualitatively, for lead times $T \leq 10$ days. Of course, the SWG model cannot replace a numerical weather prediction, as the SWG parameters (e.g. region of analogues) are tuned to local variables, and rely on the existence of a fairly large database to compute analogues. Here we used the same domain of circulation analogues for stations from Madrid to Berlin. Obviously, this region should be optimized for each individual station. Therefore, the main utility of the SWG forecast system is to make local ensemble simulations, where its performances can challenge a numerical weather prediction, if the parameters are well tuned.

This paper hence confirms the proof of concept to generate ensembles of (local) precipitation forecasts from analogs of circulation. Its performance relies on the relation between precipitation and the synoptic atmospheric circulation, which is verified for western Europe. Transposing this SWG to other regions of the globe requires observations covering several decades. Numerical weather models obviously do not yield this constraint.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions

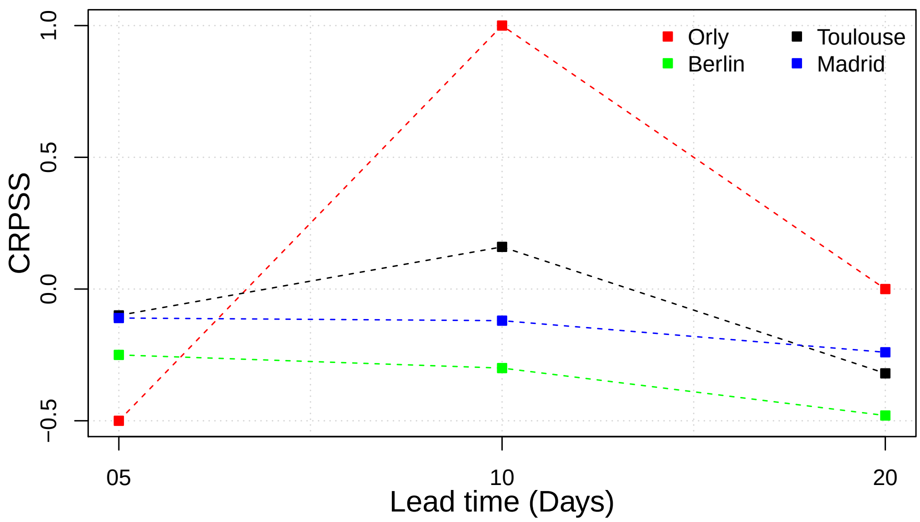

Figure 7. CRPSS of ECMWF forecasts using as a baseline the CRPS of SWG, for lead times $T = 5, 10$ and 20 days.

Code availability. The code and data files are available at http://doi.org/10.5281/zenodo.4524562

Author contributions. MK performed the analyses. PY co-desgined the analyses. CD and ST participated to the manuscript preparation.

Competing interests. The authors declare no competing interest.

Acknowledgements. This work is part of the EU International Training Network (ITN) Climate Advanced Forecasting of subseasonal Extremes (CAFE). The project receives funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 813844. We thank L. Magnusson for helpful discussions on the ECMWF data.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development Discussions
Open Access

Appendix A: Skill scores for other stations

To avoid a tedious redundancy we deferred the figures of the individual CRPSS and correlation scores to this appendix section.

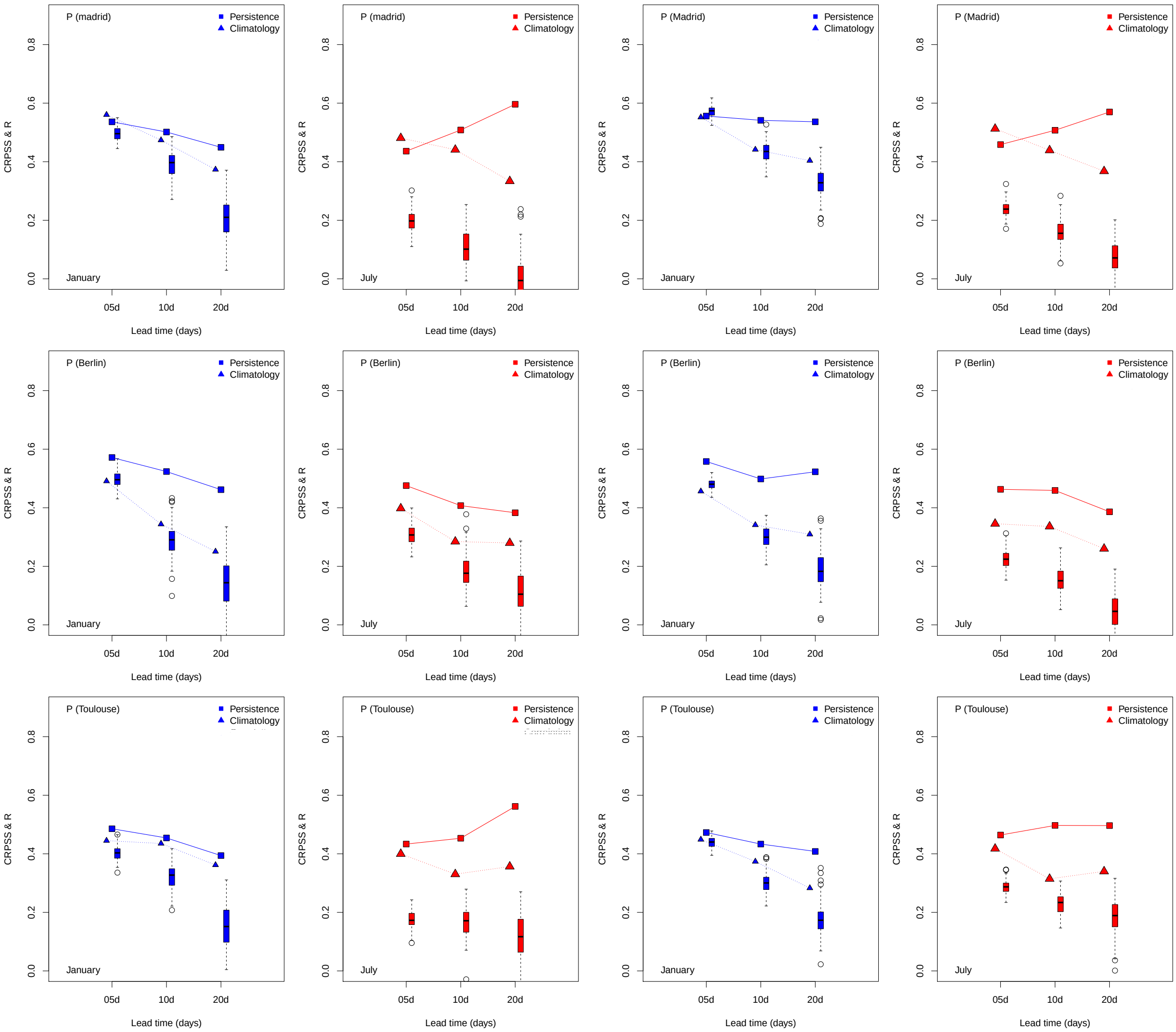

Figure A1. Skill scores for the precipitation of Madrid, Berlin and Toulouse for lead times of 5, 10, 20 days for January (blue) and July (red) for analogues computed from reanalyses of ERA5 (left) and NCEP (right). Squares indicate CRPSS where the Persistence is the baseline, triangles indicates CRPSS where the climatology is the reference, and box-plots indicates the correlation between observation and median of 100 simulations.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions
Open Access
EGU

Appendix B: CRPS and weather regimes

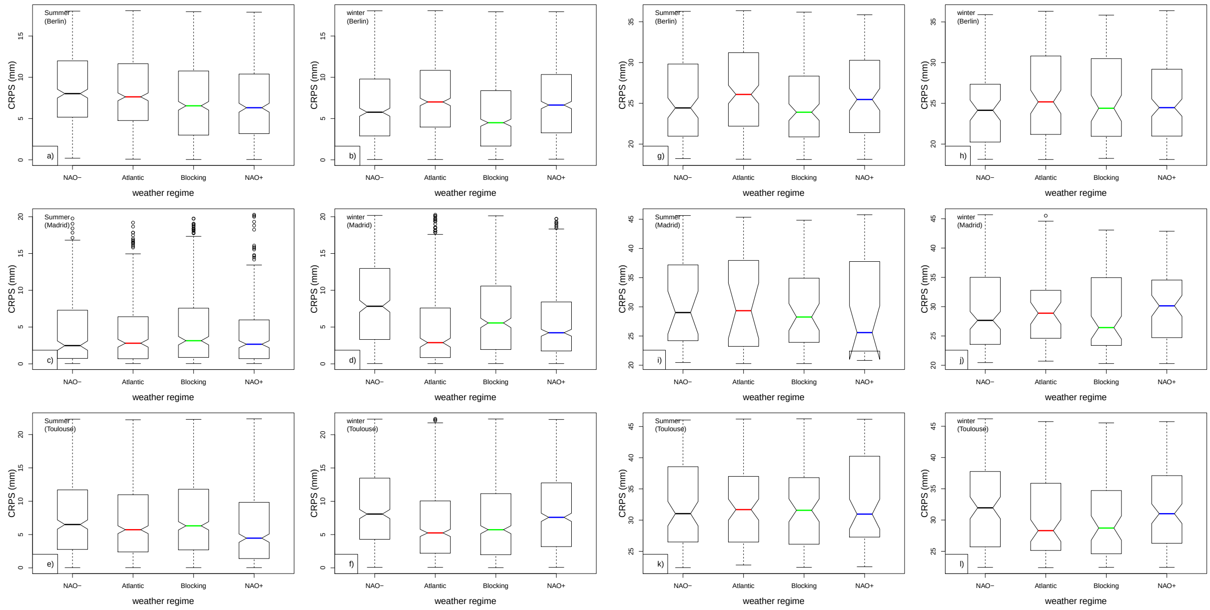

Figure B1. Evaluation of the forecast quality by weather regimes, for Berlin, Madrid and Toulouse low values of CRPS (a to f) and high values of CRPS (g to l)

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Geoscientific Model Development
Discussions

References

Ailliot, P., Allard, D., Monbet, V., and Naveau, P.: Stochastic weather generators: an overview of weather type models, Journal de la Société Française de Statistique, 156, 101–113, 2015.

Atencia, A. and Zawadzki, I.: A comparison of two techniques for generating nowcasting ensembles. Part II: Analogs selection and comparison of techniques, Monthly Weather Review, 143, 2890–2908, 2015.

Blanchet, J., Stalla, S., and Creutin, J.-D.: Analogy of multiday sequences of atmospheric circulation favoring large rainfall accumulation over the French Alps, Atmospheric Science Letters, 19, e809, 2018.

Cassou, C.: The weather regime paradigm: description and properties, Nature.

Eckmann, J.-P. and Ruelle, D.: Ergodic theory of chaos and strange attractors, in: The Theory of Chaotic Attractors, pp. 273–312, Springer, 1985.

Faranda, D., Messori, G., and Yiou, P.: Dynamical proxies of North Atlantic predictability and extremes, Scientific reports, 7, 41 278, https://www.nature.com/articles/srep41278.pdf, 2017.

Ferro, C. A. T.: A probability model for verifying deterministic forecasts of extreme events, Weather and Forecasting, 22, 1089–1100, <GotoISI>://000250413300011, 2007.

Gabriel, K. R. and Neumann, J.: A Markov chain model for daily rainfall occurrence at Tel Aviv, Quarterly Journal of the Royal Meteorological Society, 88, 90–95, iSBN: 0035-9009 Publisher: Wiley Online Library, 1962.

Ghil, M., Chekroun, M., and Simonnet, E.: Climate dynamics and fluid mechanics: Natural variability and related uncertainties, Physica D, 237, 2111–2126, https://doi.org/10.1016/j.physd.2008.03.036, 2008.

Haiden, T., Janousek, M., Bidlot, J., Buizza, R., Ferranti, L., Prates, F., and Vitart, F.: Evaluation of ECMWF forecasts, including the 2018 upgrade, European Centre for Medium Range Weather Forecasts, 2018.

Hempelmann, N., Ehbrecht, C., Alvarez-Castro, C., Brockmann, P., Falk, W., Hoffmann, J., Kindermann, S., Koziol, B., Nangini, C., Radanovics, S., Vautard, R., and Yiou, P.: Web processing service for climate impact and extreme weather event analyses. Flyingpigeon (Version 1.0), Computers & Geosciences, 110, 65–72, https://doi.org/10.1016/j.cageo.2017.10.004, https://github.com/bird-house/blackswan, 2018.

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, Weather and Forecasting, 15, 559–570, 2000.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., and Schepers, D.: The ERA5 global reanalysis, Quat. J. Roy. Met. Soc., 146, 1999–2049, iSBN: 0035-9009 Publisher: Wiley Online Library, 2020.

Jolliffe, I. T. and Stephenson, D. B.: Forecast verification: a practitioner's guide in atmospheric science, John Wiley & Sons, 2011.

Jézéquel, A., Yiou, P., and Radanovics, S.: Role of circulation in European heatwaves using flow analogues, Climate Dynamics, 50, 1145–1159, 2018a.

Jézéquel, A., Yiou, P., Radanovics, S., and Vautard, R.: Analysis of the exceptionally warm December 2015 in France using flow analogues, Bulletin of the American Meteorological Society, 99, S76–S79, 2018b.

Kimoto, M.: Studies of Climate Variability Using General Circulation Models, in: Earth Planets and Space, 2001.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, Bulletin of the American Meteorological Society, 82, 247–267, <GotoISI>://000166742900003, 2001.

https://doi.org/10.5194/gmd-2021-36
Preprint. Discussion started: 24 March 2021

Lorenz, E. N.: Atmospheric Predictability as Revealed by Naturally Occurring Analogues, J. Atmos. Sci., 26, 636–646, 1969.

Michelangeli, P., Vautard, R., and Legras, B.: Weather regimes: Recurrence and quasi-stationarity, J. Atmos. Sci., 52, 1237–1256, 1995.

Palmer, T. N.: Predicting uncertainty in forecasts of weather and climate, Reports on Progress in Physics, 63, 71–116, <GotoISI>://000085479700001, 2000.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, 2007.

Peixoto, J. P. and Oort, A. H.: Physics of climate, American Institute of Physics, New York, 1992.

Richardson, C. W.: Stochastic simulation of daily precipitation, temperature, and solar radiation, Water resources research, 17, 182–190, iSBN: 0043-1397 Publisher: Wiley Online Library, 1981.

Ruelle, D.: Ergodic theory of differentiable dynamical systems, Publications Mathématiques de l'Institut des Hautes Études Scientifiques, 50, 27–58, 1979.

Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., Eade, R., Fereday, D., Folland, C. K., and Gordon, M.: Skillful long-range prediction of European and North American winters, Geophysical Research Letters, 41, 2514–2519, 2014.

Sivillo, J. K., Ahlquist, J. E., and Toth, Z.: An ensemble forecasting primer, Weather and Forecasting, 12, 809–818, <GotoISI>://000071577700008, 1997.

Todorovic, P. and Woolhiser, D. A.: A stochastic model of n-day precipitation, Journal of Applied Meteorology, 14, 17–24, iSBN: 0021-8952, 1975.

Toth, Z.: Intercomparison of Circulation Similarity Measures, Mon. Wea. Rev., 119, 55–64, <GotoISI>://A1991FA68200003, 1991.

Toth, Z. and Kalnay, E.: Ensemble forecasting at NCEP and the breeding method, MONTHLY WEATHER REVIEW, 125, 3297–3319, <GotoISI>://A1997YJ55500014, 1997.

Türkes, M., Sümer, U. M., and Kiliç, G.: Persistence and periodicity in the precipitation series of Turkey and associations with 500 hPa geopotential heights, Climate Research, 21, 59–81, iSBN: 0936-577X, 2002.

van den Dool, H. M.: Empirical Methods in Short-Term Climate Prediction, Oxford University Press, Oxford, 2007.

van der Wiel, K., Bloomfield, H. C., Lee, R. W., Stoop, L. P., Blackport, R., Screen, J. A., and Selten, F. M.: The influence of weather regimes on European renewable energy production and demand, Environmental Research Letters, 14, 094 010, iSBN: 1748-9326 Publisher: IOP Publishing, 2019.

von Storch, H. and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge University Press, Cambridge, 2001.

Xoplaki, E., Luterbacher, J., Burkard, R., Patrikas, I., and Maheras, P.: Connection between the large-scale 500 hPa geopotential height fields and precipitation over Greece during wintertime, CLIMATE RESEARCH, 14, 129–146, <GotoISI>://000086740200005, 2000.

Yiou, P. and Déandréis, C.: Stochastic ensemble climate forecast with an analogue model, Geoscientific Model Development, 12, 723–734, 2019.

Yiou, P., Goubanova, K., Li, Z. X., and Nogaj, M.: Weather regime dependence of extreme value statistics for summer temperature and precipitation, Nonlinear Processes in Geophysics, 15, 365–378, 2008.

Yiou, P., Salameh, T., Drobinski, P., Menut, L., Vautard, R., and Vrac, M.: Ensemble reconstruction of the atmospheric column from surface pressure using analogues, Clim. Dyn., 41, 1333–1344, https://doi.org/10.1007/s00382-012-1626-3, 2013.

Zamo, M. and Naveau, P.: Estimation of the Continuous Ranked Probability Score with Limited Information and Applications to Ensemble Weather Forecasts, Mathematical Geosciences, 50, 209–234, 2018.