Stochastic Ensemble Climate Forecast with an Analogue Model

Pascal Yiou[1] and Céline Déandréis[2]

[1]Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, IPSL and Université Paris-Saclay, CE l'Orme des Merisiers, 91191 Gif-sur-Yvette, France

[2]ARIA Technologies, 8-10 Rue de la Ferme, 92100 Boulogne-Billancourt, France

Correspondence: P. Yiou (pascal.yiou@lsce.ipsl.fr)

Abstract. This paper presents a system to perform large ensembles climate stochastic forecasts. The system is based on random analogue sampling of sea-level pressure data from the NCEP reanalysis. It is tested to forecast an NAO index and the daily average temperature in five European stations. We simulated 100 member ensembles of averages over lead times from 5 days to 80 days in a hindcast mode, i.e. from a meteorological to a seasonal forecast. We tested the hindcast simulations with usual forecast skill scores (CRPSS or correlation), against persistence and climatology. We find significantly positive skill scores for all time scales. Although this model cannot outperform numerical weather prediction, it presents an interesting benchmark that could complement climatology or persistence forecast.

1 Introduction

Stochastic weather generators (SWG) have been devised to simulate many and long sequences of climate variables that yield realistic statistical properties Semenov and Barrow (1997). Their main practical use has been to investigate the probability distribution of local variables such as precipitation, temperature or wind speed, and their impacts on agriculture (Carter, 1996; Semenov, 2006), energy (Parey et al., 2014) or ecosystems (Maraun et al., 2010). Such systems can simulate hundreds or thousands of trajectories on desktop computers and propose cheap alternatives to climate model simulations.

There are many categories of SWGs (Ailliot et al., 2015). Some SWGs are explicit random processes, whose parameters are obtained from observations of the variable to be simulated (Parey et al., 2014). Some SWGs are based on a random resampling of the observations (Iizumi et al., 2012). Some SWGs simulate local variables from their dependence to large-scale variables such as the atmospheric circulation (Kreienkamp et al., 2013). This allows to simulate spatially coherent multivariate fields (Yiou, 2014; Sparks et al., 2018) and can be used for downscaling (Wilks, 1999).

SWGs that use observations as input could in principle be used to forecast variables. This is the case for analogue weather generators (Yiou, 2014). Methods of analogues of atmospheric circulation were first devised for weather forecast (Lorenz, 1969; van den Dool, 1989). They were abandoned when numerical weather prediction was developed and implemented, because their performance was deemed inadequate (van den Dool, 2007). However, recent studies on *nowcasting* have shown

that analogue based methods could outperform numerical weather prediction for precipitation (Atencia and Zawadzki, 2015).
Yiou (2014) showed some skill for temperature simulations in Europe of an analogue SWG.

Due to uncertainties in observations and the high sensitivity to initial conditions (van den Dool, 2007) weather forecasts estimate probability density functions rather than deterministic meteorological values. Therefore, weather forecasts examine the

5   properties of all possible trajectories of an atmospheric system from an ensemble of initial conditions. Such properties include the range and the median, for example. Then one can compare how the ensemble of trajectories compares to observations, and other reference forecasts. Numerical weather forecasts rely on large ensembles of model simulations and require a massive use of supercomputers in order to provide estimates of the probability density function (pdf) of variables of interest, for various lead times. Being able to increase the ensemble size of weather forecast systems in order to lower the bias of the forecast skill

10   has been a challenge of major centers of weather prediction (Weisheimer and Palmer, 2014).

The most trivial prediction systems are based on either climatology (i.e. predicting from the seasonal average) or persistence (i.e. predicting from the past observed values) (Wilks, 1995). Probabilistic and statistical models can provide more sophisticated benchmarks for weather forecast systems, still without simulating the underlying primitive hydrodynamic equations and using supercomputers. For example, statistical models of forecast for precipitation based on analogues (of precipitation) were tested

15   for North America (Atencia and Zawadzki, 2015). Such systems tend to outperform numerical weather forecast systems, although their computing cost is steeper than most SWGs. Therefore the potential of analogue based methods can be useful to assess probability distributions, rather than a purely deterministic forecast.

Machine learning algorithms were recently devised to simulate complex systems (Pathak et al., 2018a) with surprising performances. Such algorithms are sophisticated ways of computing analogues of observed trajectories in a learning step, and

20   simulating potentially new trajectories from this learning. The main drawback is that such algorithms generally require a tricky tuning of parameters that might not be based on a physical intuition. From the inspiration of machine learning algorithms, we propose to devise a weather forecast system based on a stochastic weather generator that uses analogues of circulation to generate large ensembles of trajectories. The rationale for using analogues, rather than more sophisticated machine learning, is that they correspond to a physical interpretation of relations between large scale and regional scales. Moreover, mathematical

25   results in dynamical system theory (Freitas et al., 2016; Lucarini et al., 2016) suggest that properties of recurring patterns is asymptotically independent on the distance that is used to compute analogues.

This paper presents tests of such a system to forecast temperatures in Europe and an index of the North Atlantic Oscillation (NAO). The NAO controls the strength and direction of westerly winds and location of storm tracks across the North Atlantic in the winter (Hurrell et al., 2003). It is therefore strongly tied to temperature and precipitation variations in Europe (Slonosky

30   and Yiou, 2001).

Since the set up of such a system is fairly light, it is possible to test it for time leads from a meteorological forecast (5 days ahead) to a seasonal forecast (80 days ahead). We test this system in hindcast experiments to forecast climate variables between 1970 and 2010. The tests are performed with usual skill scores (continuous rank probability score and correlation).

The paper is organized as follows. Section 2 presents the datasets that are used as input of the system. Section 3 presents the forecast system based on analogues, the skill scores and the experimental protocol. Section 4 presents the results on simulations of the NAO index and European temperatures.

2 Data

5   We used data from different sources for sea-level pressure (SLP), NAO index and temperatures. SLP data are used for analogue computations as predictor. The NAO index and temperatures are the predictands. It is important that they share a common chronology, in order to allow their simulation because the NAO index and temperatures are simulated from from SLP analogues.

2.1 Sea-level pressure

We use the reanalysis data of the National Centers for Environmental Prediction (NCEP) (Kistler et al., 2001). We consider the
10   sea-level pressure (SLP) over the North Atlantic region. We used SLP daily averages between January 1st 1948 and April 30th 2018. The horizontal resolution is 2.5° in longitude and latitude. The rationale of using this reanalysis is that it covers more than 60 years and is regularly updated, which makes it a good candidate for a continuous time forecast exercise.

    One of the caveats of this reanalysis dataset is the lack of homogeneity of assimilated data, in particular before the satellite era. This can lead to breaks in pressure related variables, although such breaks are mostly detected in the southern hemisphere
15   and the Arctic regions (Sturaro, 2003). We are not interested in the evaluation of SLP trends, therefore breaks should only marginally impact our results.

2.2 NAO index

The North Atlantic Oscillation (NAO) is a major mode of atmospheric variability in the North Atlantic (Hurrell et al., 2003). Its intensity is determined by an index that can be computed as the normalized sea-level pressure difference between the Azores
20   and Iceland (Hurrell, 1995). The NAO index is related to the strength and direction of the westerlies, so that high values correspond to zonal flows across the North Atlantic region, stormy conditions and rather high temperatures in Western Europe (Slonosky and Yiou, 2001; Hurrell et al., 2003).

    We retrieved the daily NAO index from the NOAA web site:

http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml.

25   The procedure to calculate the daily NAO teleconnection indice is detailed on the NOAA web site. In short, a Rotated Principal Component Analysis (RPCA) is applied to monthly averages of geopotential height at 500 hPa (Z500) anomalies (Barnston and Livezey, 1987) in the 20N–90N region, between January 1950 and December 2000, from the NCEP reanalysis. The empirical orthogonal functions (EOFs) provide climatological monthly teleconnection patterns (Wilks, 1995). Those monthly teleconnection patterns are interpolated for every day in the year. Then daily Z500 anomaly fields are projected onto
30   the interpolated climatological teleconnection patterns in order to obtain a daily NAO index.

The geographical domain on which this NAO index is computed is larger than the one for SLP data. Scaife et al. (2014) used an NAO index to test the UKMO seasonal forecast system. The index they used is based on monthly SLP differences between the Azores and Iceland, and is therefore different from ours.

2.3 European temperatures

5 We took daily averages of temperatures from the ECAD project (Klein-Tank et al., 2002). We extracted data from Berlin, De Bilt, Toulouse, Orly and Madrid (Fig. 1). Those stations cover a large longitudinal and latitudinal range in western Europe. These datasets were also chosen because

– they start before 1948 and end after 2010. This allows the computation of analogue temperatures with the SLP from the NCEP reanalysis, which includes that period,

10 – they contain less than 10% of missing data.

These two criteria allow keeping 528 out of the 11422 ECAD stations that are available in 2018.

3 Methods

3.1 Analogues of circulation

Analogues of circulation are computed on SLP data from NCEP (Sec. 2.1). For each day between Jan. 1st 1948 and Dec. 31st
15 2017, the best 20 analogues (with respect to a Euclidean distance) in a different year are searched. This follows the procedure of (Yiou et al., 2013). The analogues are computed over two regions (large region: North Atlantic region (80W–30E; 30–70N); small region: Western Europe (30W–20E; 40–60N)). The large region is used to simulate/forecast the NAO index. This choice is justified by the fact that the North Atlantic atmospheric circulation patterns are well defined over that region (Michelangeli et al., 1995). The small region is used to simulate/forecast continental temperatures, following the domain recommendations
20 of the analysis of Jézéquel et al. (2018).

3.2 Forecast with analogue stochastic weather generator

Ensembles of simulations of temperature or the NAO index can be performed with the rules illustrated by Yiou (2014), with an analogue-based stochastic weather generator. This stochastic weather generator can be run in so called *dynamical* mode. For each day t, we have N best SLP analogues. We randomly select one of those N analogues and time $\tilde{t}$, with a probability
25 weight that is

1. inversely proportional to the calendar distance of the analogues dates $\tilde{t}$ to t. This constrains the time of analogues to move forward.

2. inversely proportional to the correlation of the analogue with the SLP pattern at time t. This constraint favors analogues with the best patterns, among those with the closest distance.
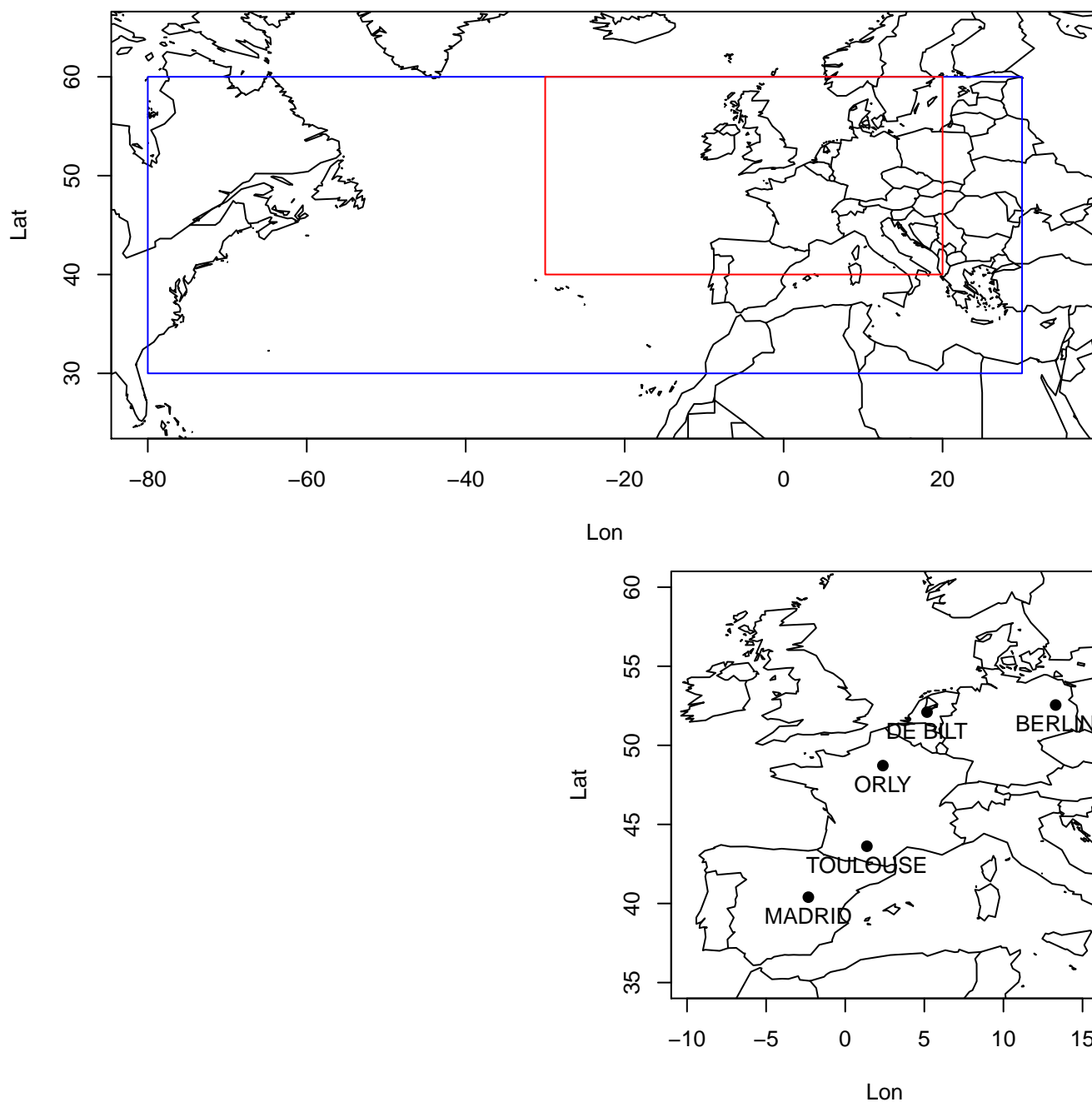
Figure 1. Upper panel: North Atlantic region (blue rectangle) and Western European region (red rectangle) on which analogues are computed.
Lower panel: European stations used for daily mean temperature.

3. a zero weight if $\tilde{t}$ is larger than t. This ensures that no information coming from times beyond t is used in the simulation process.

The simulated SLP at $t+1$ is then the next day of the selected analogue. We repeat this operation until a lead time T. This generates one random trajectory between t and $t+T$. The random sampling procedure is repeated S times to generate an

5 ensemble of trajectories. Here, $S = 100$.

If we want to simulate a sequence starting at time t and until $t+T$, we have excluded all analogues whose date falls in $[t, t+T]$ in the random analogue selection. This provides a simple way of performing hindcast forecast for temperature or NAO index.

In this paper, the lead time T is 5, 10, 20, 40 and 80 days ahead. For each daily trajectory starting at t, we compute the

10 average between t and $t+T$. Therefore, we go from a meteorological forecast (5 days) to a seasonal forecast (80 days) of averaged trajectories.

The $S = 100$ simulations at each time steps allow computing medians and quantiles of the averaged trajectories.

For comparison purposes, climatological and persistence forecasts are also computed. The climatological forecast for a lead time T is determined from the seasonal cycle of T averages of the variable we simulate. For each time t, the *climatological*

15 forecast for $t+T$ is the mean seasonal cycle of T averages at the calendar day of t. The *persistence* forecast at time t for a lead time of T is the observed average between $t-T$ and t. Those two types of forecasts are illustrated in Fig. 2. These two definitions ensure a coherence between the predictand (averages over T values ahead) and predictors (mean of averages over T values for climatology, or average over T preceding values for persistence).

3.3 Forecast skill

20 The simplest score we use is the temporal correlation between the median of the ensemble forecast and the observations. Due to the autocorrelation and seasonality of the variables we try to simulate (temperature and NAO index), we consider the correlations for the forecast in the months of January and July.

The continuous rank probability score (CRPS) compares the cumulated density functions of a forecast ensemble and observations y_t, for all times t (Ferro, 2014).

25 $$\mathrm{CRPS}(t) = \int_{-\infty}^{\infty} \left(F_t(x) - \mathbf{1}(x \geq y_t) \right)^2 dx. \tag{1}$$

F_t is the cumulated density function of the ensemble forecast at time t. It is obtained empirically from the ensemble of simulations of the model. $\mathbf{1}(x \geq y_t)$ is the empirical cumulated density function of the observation y_t.

The score is the average over all times:

$$\mathrm{CRPS} = \frac{1}{N} \sum_{t=1}^{N} \mathrm{CRPS}(t). \tag{2}$$

30 The CRPS is a fair score (Ferro, 2014; Zamo and Naveau, 2018) in that it compares the probability distributions of forecasts and observations and it is optimal when they are the same. Discrete estimates of CRPS can yield a bias for small ensemble
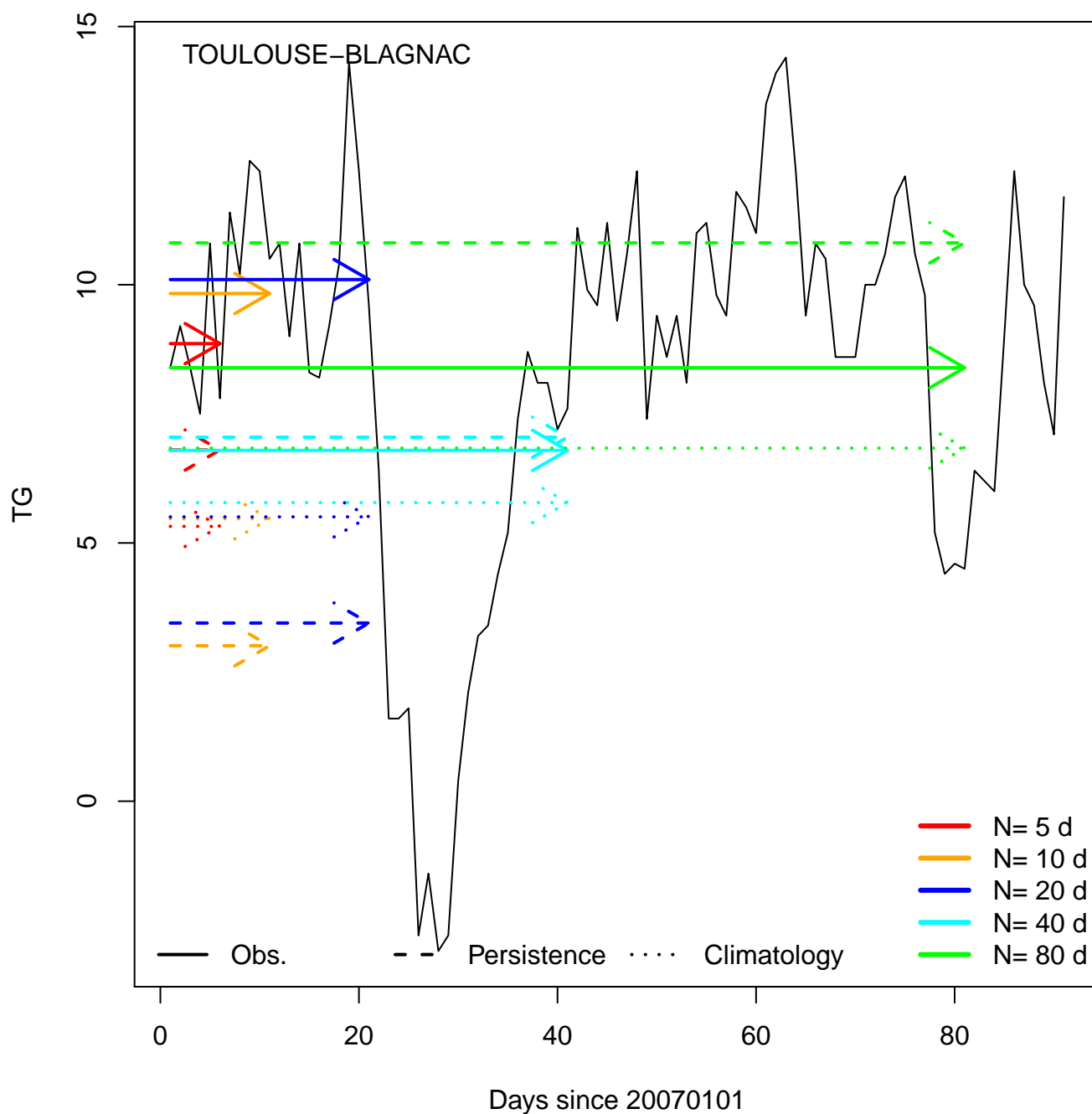
Figure 2. Illustration of average forecast for daily mean temperature (TG) in Toulouse, for Jan. 1st 2007. The continuous black line indicates the observations of TG for the first 90 days of 2007. The colors indicate lead times T. The continuous arrows are for observed averages of TG on Jan. 1st 2007. The dashed lines are for the persistence forecast of TG and the dotted lines are for the climatology forecast of TG on Jan. 1st 2007.

sizes M. We simulate $M = 100$ trajectories for each forecast. This is more than most ensemble weather forecasts (typically, $M = 51$ for the European Center for Medium Range Forecast (ECMWF) ensemble forecast) and guaranties that the bias due to the number of samples is negligible.

A perfect forecast gives a CRPS value of 0. The units of CRPS are those of the variable to be forecasted therefore its

5   interpretation is not universal, and comparing the CRPS values for NAO index and temperatures is not directly possible. Therefore, it is useful to compare the CRPS of the forecast with the one of a reference. A normalization of CRPS provides a skill score with respect to that reference:

$$\text{CRPSS}_{\text{ref}} = 1 - \frac{\text{CRPS}}{\text{CRPS}_{\text{ref}}}. \tag{3}$$

The CRPSS indicates an improvement over the reference forecast. A perfect forecast has a CRPSS of 1. A positive im-

10   provement over the reference yields positive a CRPSS value. A value of 0 or less indicates that the forecast is worse than the reference.

We compute CRPSS for the climatological and persistence references. We used the package "SpecsVerification" in R to compute CRPS and CRPSS scores. Hence we compare our stochastic forecasts with forecasts made from climatology and persistence. By construction, the persistence forecast shows an offset with the actual value ahead, because the persistence is

15   the value of the average of observations between $t - T$ and t. The variability of the climatological forecast is low because it is an average of T long sequences.

3.4   Protocole

We tested the ensemble forecast system on the period between 1970 and 2010. We simulate $N = 100$ trajectories of length $T \in \{5, 10, 20, 40, 80\}$ days for a given date t, and average each trajectory over T. The dates t are shifted every $\delta t \in \{2, 5, 10, 10, 20\}$

20   days, respectively for each different value of lead times T.

We recall that the tests we perform are on the *average* of the forecast between t and $t + T$, not on the *value at* time $t + T$.

The CRPSS is computed for each value of lead times T, with references of climatology and persistence. We plot quantile-quantile plots for observed and forecast values of the averages. This allows assessing biases in simulating averages. Variables such as temperature yield a strong seasonality, which is larger than daily variations. It is hence natural to have very high

25   correlations or skill scores if one considers those scores over the whole year. Therefore we compare the skill scores for January and July, in order to avoid obtaining artificially high scores.

4   Results

We performed our stochastic forecasts on the NAO index and European temperatures. The two datasets are treated separately because the simulations are done with two different analogue computations (Sec. 3.1).

4.1 NAO index

For illustration purposes, we comment on the skill on simulations of 2007. Fig. 3 shows the simulated and observed values of averages of the NAO index, for five values of T (5, 10, 20, 40 and 80 days). This example suggests a good skill to forecast the NAO index from SLP analogues, especially at lead times of $T = 5$ to 10 days.

5   The q-q plots of the median of simulations versus observations show a bias that reduces the range of variations (Fig. 3, right column). There are two reasons for this reduction of variance, which is proportional to the lead time T:

    1. individual simulated trajectories tend to "collapse" toward a climatological value after ≈ 10 days,

    2. taking the median of all simulations also naturally reduces the variance.

The q-q plots are almost linear. This means that the bias could in principle be corrected by a linear regression. We will not
10   perform such a correction in the sequel.

The correlation and CRPSS values for NAO index forecast are shown in Fig. 4. The values of $\mathrm{CRPSS_{pers}}$ (for a persistence reference) are rather stable (with a slight increase) near 0.45, and the climatology score slightly decreases with T although positive.

One the one hand, the $\mathrm{CRPSS_{clim}}$ values do not depend on the season (identical triangles in Fig. 4). On the other hand,
15    $\mathrm{CRPSS_{pers}}$ values are higher in July than January for lead times $T \leq 10$ days, and lower for lead times $T \geq 40$ days (squares in Fig. 4). This means that the climatology forecast tends to be better than the persistence forecast for $T > 5$ days (squares higher than triangles in Fig. 4), which can be anticipated because of the inherent lag of the persistence forecast.

The correlation scores decrease with lead time T. The correlation skill is higher in January than in July. It is no longer significantly positive for T larger than 40 days (confidence intervals contain the 0 value). The correlation score values range
20   between 0.65 and 0.82 for $T = 5$ day forecasts, and 0.45 and 0.77 for $T = 10$ day forecasts, depending on the season. This is consistent with the NAO forecast of the Climate Prediction Center ($r = 0.69$ for a 10 day forecast). The correlation score is still significantly positive for $T = 20$ days. The higher correlation scores over the whole year (black circles) reflect a (small) seasonal cycle of the NAO index. This artificially enhances the score for those lead times because SLP analogue predictands tend to reproduce the seasonality of the SLP field (by construction of the simulation procedure), and the NAO index and SLP
25   variations are closely linked on monthly time scales (by construction of the NAO index).

4.2 European temperatures

The correlation and CRPSS values for temperature forecast are shown in Fig. 5. The values of $\mathrm{CRPSS_{pers}}$ (for a persistence reference) increases with lead time T. This is not surprising because the forecast for the next T days is based on the average of the past T days. Therefore, the persistence forecast is always "late" due to the strong seasonality of temperature variations.
30   The $\mathrm{CRPSS_{clim}}$ values decrease with T and plateau near ≈ 0.2. This skill score is still positive (albeit small) for a seasonal forecast. This positive average skill ($\mathrm{CRPSS} > 0.2$) illustrates that the stochastic weather generator follows the seasonality of
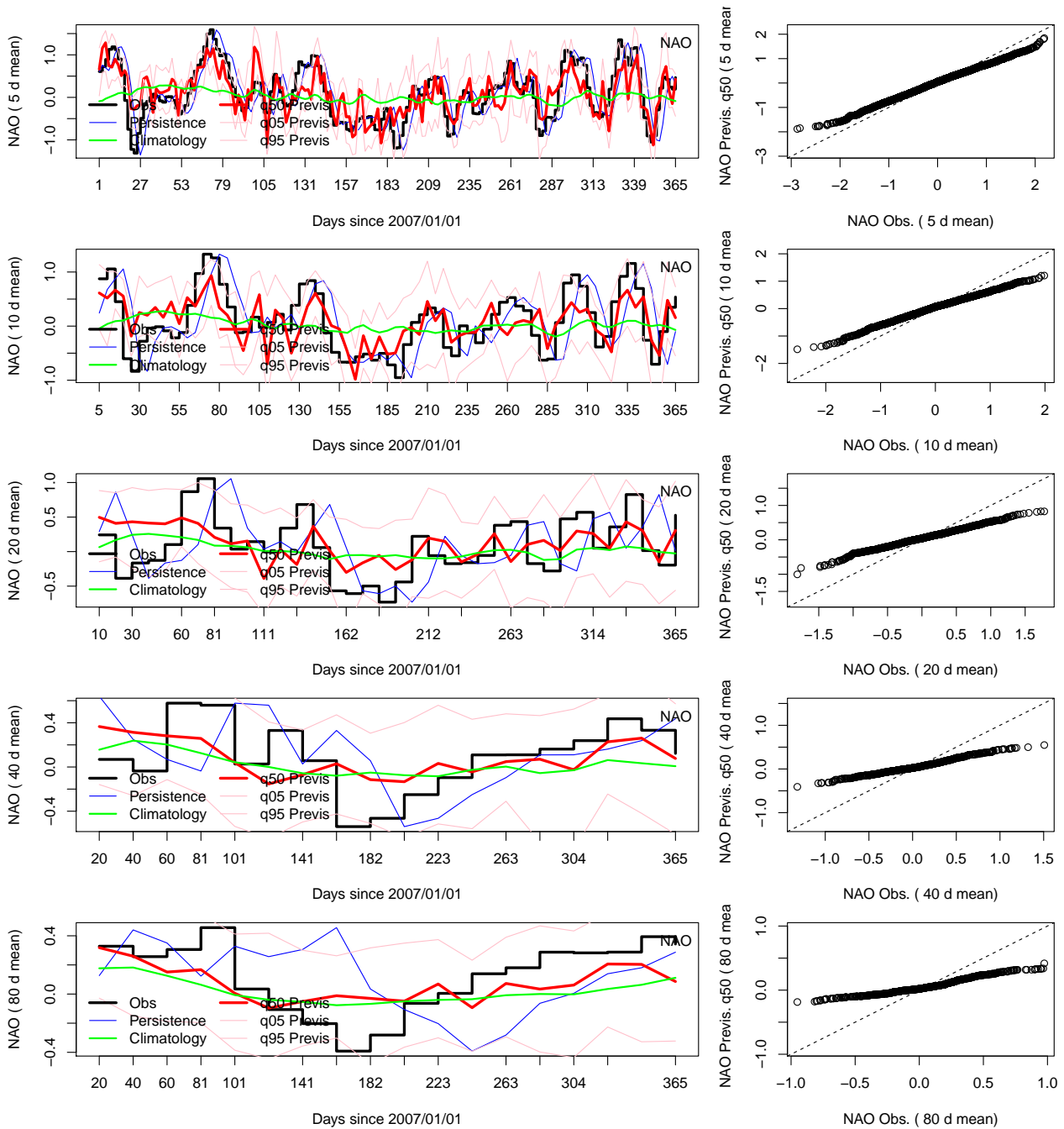
Figure 3. Left column: time series of ensemble forecasts for 2007, for lead times $T \in \{5, 10, 20, 40, 80\}$ days. Right column: q-q plots of NAO forecasts vs. observed values for all years in 1970–2010 for all lead times. The dotted line is the first diagonal.
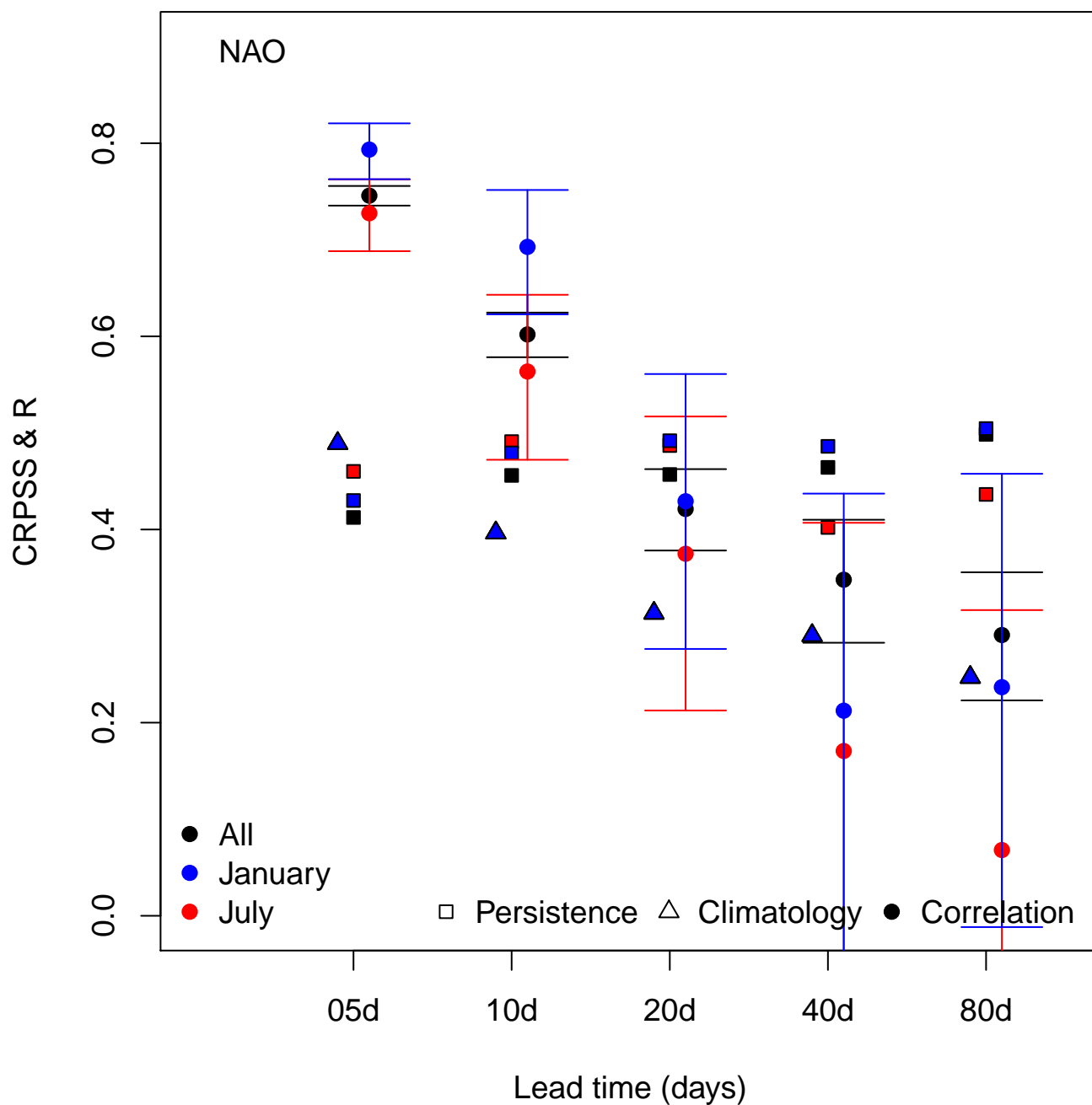
Figure 4. Skill scores for NAO index for lead times T of 5, 10, 20, 40 and 80 days for all days (black), January (blue) and July (red). Square indicate $CRPSS_{pers}$, triangles $CRPSS_{clim}$ and circles are fore correlation. Triangles are identical for all days, January and July. The error bars for the correlation indicate a 95% confidence interval (von Storch and Zwiers, 2001).

temperature variations. We note that the CRPSS values for temperature are higher than for the NAO index. This is explained by the seasonality of temperatures, which is more pronounced than in the daily NAO index.

The CRPSS values are rather consistent for the four of the stations (Toulouse, De Bilt, Berlin and Orly). The stochastic model CRPSS fares slightly worse at Madrid station.

5 The correlation scores for January and July decrease with lead time T. The correlation score values for all days are above 0.97 due to the seasonality of temperatures and forecasts. Since this is not informative, this is not shown in Fig. 5. The correlations are always significantly positive for Toulouse, De Bilt, Berlin and Orly. The summer correlation intervals contain the 0 value at Madrid. This is probably due to the fact that temperature is not linked to the atmospheric circulation in the summer, but rather to local processes of evapo-transpiration (Schaer et al., 1999; Seneviratne et al., 2006).

5   Conclusions

We have presented a system to generate ensembles of stochastic simulations of the atmospheric circulation, based on pre-computed analogues of circulation. This system is fairly light in terms of computing resources as it can be run on a (reasonably powerful) personal computer. The most fundamental assumption of the system is that the variable to be predicted is linked to the atmospheric circulation. The geographical window for the computation of analogues needs to be adjusted to the variable to be predicted, so that a prior expertise is necessary for this analogue forecast system. This implies that this approach would not be adequate for variables that are not connected in any way to the atmospheric circulation (here approximated by SLP). The use of other atmospheric fields (e.g. geopotential heights) might increase the skill of the system. The computation of analogues with other parameters (geographical zone, atmospheric predictand, reanalysis, climate model output, etc.) can be easily performed with a web processing service (Hempelmann et al., 2018).

20 We have tested the performance of the system to simulate an NAO index and temperature variations in five European stations. The performance of such a system cannot beat a meteorological or seasonal forecast with a full-scale atmospheric model (Scaife et al., 2014), but its skill is positive, even at a seasonal time scale, with a rather modest computational cost.

The reason for the positive skill (especially against climatology) remains to be elucidated, especially for long lead times. We conjecture that the information contained in the initial condition (as done with regular weather forecasts) actually controls the 25 mean behavior of the trajectories from that initial condition. But such a skill is actually "concentrated" in the first few days, because the trajectories tend to converge to the climatology after 15 days.

Recent experimental results in chaotic systems have shown that a well tuned neural network algorithm could simulate efficiently the trajectories of a chaotic dynamical system (Pathak et al., 2018b). Our system is an extreme simplification of an artificial intelligence algorithm, but it does demonstrate the forecast skill of such approaches. The advantage here is the 30 physical constraint between the atmospheric circulation and the variables to be simulated.

This system was tested on temperature for five European datasets. This could be extended to precipitation or wind speed. If a real-time forecast is to be performed, we emphasize that only the predictand (here, SLP) needs to be regularly updated for the computation of analogues.
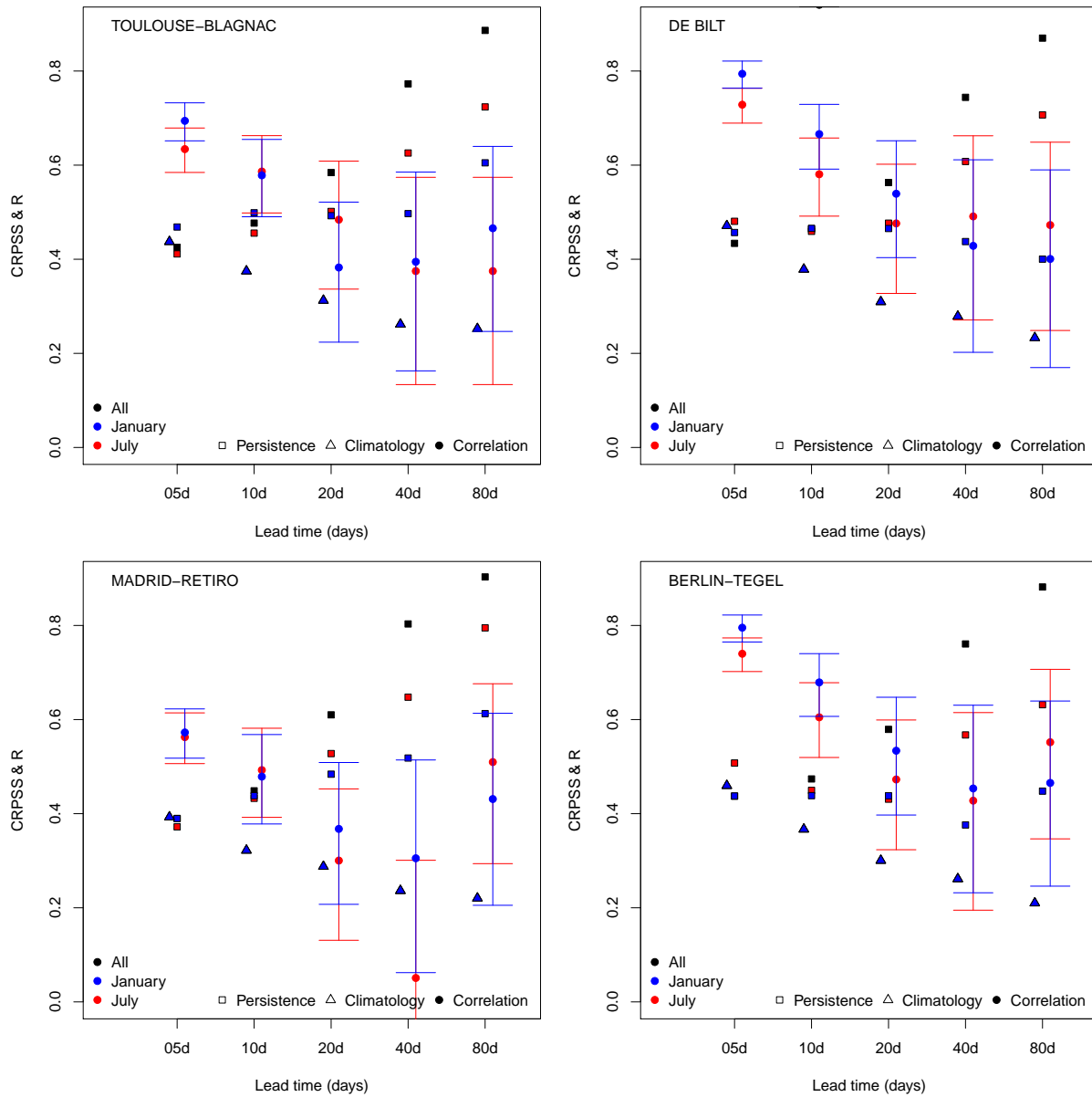
Figure 5. Skill scores for mean daily temperature in Toulouse, De Bilt, Madrid and Berlin, lead times T of 5, 10, 20, 40 and 80 days. Square indicate CRPSS$_{pers}$, triangles CRPSS$_{clim}$ and circles are fore correlation. Black symbols are for all days, blue symbols are for January and red symbols are for July. Triangles are identical for all days, January and July. The correlations for "all days" are very close to 1, due to the seasonal cycle and do not appear on the figure. The error bars for the correlation indicate a 95% confidence interval (von Storch and Zwiers, 2001).
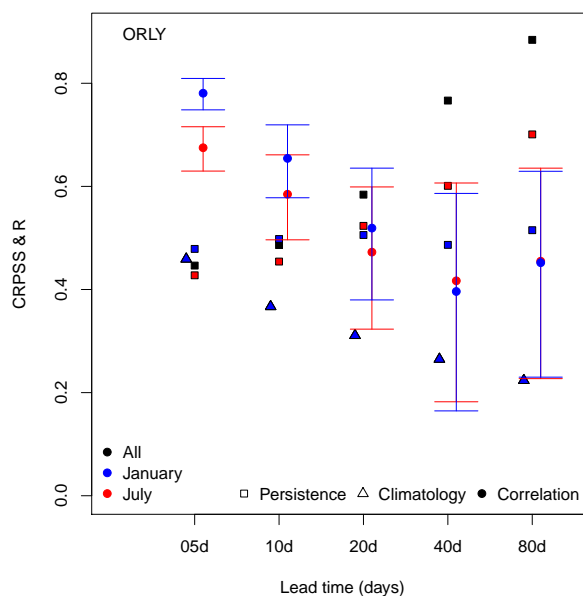
Figure 5. Skill scores for temperature in Orly (continued).

The goal of such a system is not to replace ensemble numerical weather/seasonal forecast. Rather, it can refine the usual references (climatology and persistence) for the evaluation of skill scores. This would create a third "machine learning" reference for CRPSS that might be harder to beat than the classical references.

Code and data availability. The code for the computation of analogues is available at (free CeCILL license):

5   https://a2c2.lsce.ipsl.fr/index.php/deliverables/101-analogue-software

The temperature data are available at: https://www.ecad.eu

The NAO index data are available at: http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml

The NCEP reanalysis SLP data is available at: https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html

Author contributions. PY wrote the codes, designed the experiments. CD participated to the writing of the manuscript.

10 *Competing interests.* The authors declare no competing interest.

Disclaimer.

Acknowledgements. This work was supported by a grant from the Labex-IPSL and ERC grant No. 338965-A2C2. We thank Mariette Lamige and Zhongya Liu who performed preliminary analyses during their training periods.

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.

Barnston, A. G. and Livezey, R. E.: Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns, Monthly Weather Review, 115, 1083–1126, <GotoISI>://A1987H976400002, 1987.

Carter, T. R.: Developing scenarios of atmosphere, weather and climate for northern regions, Agricultural and Food Science in Finland, 5, 235–249, 1996.

Ferro, C. A. T.: Fair scores for ensemble forecasts, Quarterly Journal of the Royal Meteorological Society, 140, 1917–1923, 2014.

Freitas, A. C. M., Freitas, J. M., and Vaienti, S.: Extreme Value Laws for sequences of intermittent maps, arXiv preprint arXiv:1605.06287, 2016.

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, http://www.sciencedirect.com/science/article/pii/S0098300416302801, 2018.

Hurrell, J.: Decadal trends in the North-Atlantic Oscillation - regional temperatures and precipitation, Science, 269, 676–679, <GotoISI>://A1995RM70200029, 1995.

Hurrell, J., Kushnir, Y., Ottersen, G., and Visbeck, M., eds.: The North Atlantic Oscillation : Climatic Significance and Environmental Impact, vol. 134 of *Geophysical monograph*, American Geophysical Union, Washington, DC, 2003.

Iizumi, T., Takayabu, I., Dairaku, K., Kusaka, H., Nishimori, M., Sakurai, G., Ishizaki, N. N., Adachi, S. A., and Semenov, M. A.: Future change of daily precipitation indices in Japan: A stochastic weather generator-based bootstrap approach to provide probabilistic climate information, J. Geophys. Res.-Atmospheres, 117, Doi 10.1029/2011jd017 197, 2012.

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

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.

Klein-Tank, A., Wijngaard, J., Konnen, G., Bohm, R., Demaree, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Muller-Westermeier, G., Tzanakou, M., Szalai, S., Palsdottir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., Van Engelen, A., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Lopez, J., Dahlstrom, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L., and Petrovic, P.: Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment, Int. J. Climatol., 22, 1441–1453, 2002.

Kreienkamp, F., Spekat, A., and Enke, W.: The Weather Generator Used in the Empirical Statistical Downscaling Method, WETTREG, Atmosphere, 4, 169–197, doi:10.3390/atmos4020 169, 2013.

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

Lucarini, V., Faranda, D., Freitas, A. C. M., Freitas, J. M., Holland, M., Kuna, T., Nicol, M., Todd, M., and Vaienti, S.: Extremes and recurrence in dynamical systems, John Wiley & Sons, 2016.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themessl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation Downscaling under

5   Climate Change: Recent Developments to Bridge the Gap between Dynamical Models and the End User, Reviews of Geophysics, 48, https://doi.org/Artn Rg3003 Doi 10.1029/2009rg000314, <GotoISI>://000282328000002, 2010.

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

Parey, S., Hoang, T. T. H., and Dacunha-Castelle, D.: Validation of a stochastic temperature generator focusing on extremes, and an example of use for climate change, Climate Research, 59, 61–75, 2014.

10   Pathak, J., Hunt, B., Girvan, M., Lu, Z., and Ott, E.: Model-Free Prediction of Large Spatiotemporally Chaotic Systems from Data: A Reservoir Computing Approach, Physical Review Letters, 120, 024 102, 2018a.

Pathak, J., Wikner, A., Fussell, R., Chandra, S., Hunt, B. R., Girvan, M., and Ott, E.: Hybrid forecasting of chaotic processes: using machine learning in conjunction with a knowledge-based model, Chaos: An Interdisciplinary Journal of Nonlinear Science, 28, 041 101, 2018b.

Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., Eade, R., Fereday, D., Folland, C. K., and Gordon, M.:

15   Skillful long-range prediction of European and North American winters, Geophysical Research Letters, 41, 2514–2519, 2014.

Schaer, C., Luthi, D., Beyerle, U., and Heise, E.: The soil-precipitation feedback: A process study with a regional climate model, J. Clim., 12, 722–741, <GotoISI>://000079181700004, 1999.

Semenov, M. A.: Using weather generators in crop modelling, Proceedings of the VIIth International Symposium on Modelling in Fruit Research and Orchard Management, pp. 93–100, <GotoISI>://000239370200011, 2006.

20   Semenov, M. A. and Barrow, E. M.: Use of a stochastic weather generator in the development of climate change scenarios, Climatic Change, 35, 397–414, 1997.

Seneviratne, S. I., Luthi, D., Litschi, M., and Schaer, C.: Land-atmosphere coupling and climate change in Europe, Nature, 443, 205–209, <GotoISI>://000240467000045, 2006.

Slonosky, V. and Yiou, P.: The North Atlantic Oscillation and its relationship with near surface temperature, Geophys. Res. Lett., 28, 807–810,

25   <GotoISI>://000167229700016, 2001.

Sparks, N. J., Hardwick, S. R., Schmid, M., and Toumi, R.: IMAGE: a multivariate multi-site stochastic weather generator for European weather and climate, Stochastic Environmental Research and Risk Assessment, 32, 771–784, 2018.

Sturaro, G.: A closer look at the climatological discontinuities present in the NCEP/NCAR reanalysis temperature due to the introduction of satellite data, Climate dynamics, 21, 309–316, https://doi.org/10.1007/s00382-003-0334-4, 2003.

30   van den Dool, H. M.: A new look at weather forecasting through analogs, Monthly Weather Review, 117, 2230–2247, https://doi.org/10.1175/1520-0493(1989)117<2230:ANLAWF>2.0.CO;2, 1989.

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

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

Weisheimer, A. and Palmer, T. N.: On the reliability of seasonal climate forecasts, Journal of the Royal Society Interface, 11, 20131 162,

35   2014.

Wilks, D.: Statistical Methods in the Atmospheric Sciences: An Introduction, Academic Press, San Diego, 1995.

Wilks, D. S.: Multisite downscaling of daily precipitation with a stochastic weather generator, Climate Res., 11, 125–136, 1999.

Yiou, P.: AnaWEGE: a weather generator based on analogues of atmospheric circulation, Geoscientific Model Development, 7, 531–543, https://doi.org/10.5194/gmd-7-531-2014, http://www.geosci-model-dev.net/7/531/2014/, 2014.

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.

5 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.