Clim. Past, 12, 635–662, 2016
www.clim-past.net/12/635/2016/
doi:10.5194/cp-12-635-2016

Climate of the Past
Open Access
EGU

Probabilistic precipitation and temperature downscaling of the Twentieth Century Reanalysis over France

Laurie Caillouet[1], Jean-Philippe Vidal[1], Eric Sauquet[1], and Benjamin Graff[2]

[1]Irstea, UR HHLY, Hydrology and Hydraulics Research Unit, 5 rue de la Doua – BP32108, 69916 Villeurbanne CEDEX, France
[2]Compagnie Nationale du Rhône (CNR), 2 rue André Bonin, 69004 Lyon, France

Correspondence to: Laurie Caillouet (laurie.caillouet@irstea.fr)

Received: 7 August 2015 – Published in Clim. Past Discuss.: 18 September 2015
Revised: 15 February 2016 – Accepted: 29 February 2016 – Published: 16 March 2016

Abstract. This work proposes a daily high-resolution probabilistic reconstruction of precipitation and temperature fields in France over the 1871–2012 period built on the NOAA Twentieth Century global extended atmospheric reanalysis (20CR). The objective is to fill in the spatial and temporal data gaps in surface observations in order to improve our knowledge on the local-scale climate variability from the late nineteenth century onwards.

The SANDHY (Stepwise ANalogue Downscaling method for HYdrology) statistical downscaling method, initially developed for quantitative precipitation forecast, is used here to bridge the scale gap between large-scale 20CR predictors and local-scale predictands from the Safran high-resolution near-surface reanalysis, available from 1958 onwards only. SANDHY provides a daily ensemble of 125 analogue dates over the 1871–2012 period for 608 climatically homogeneous zones paving France. Large precipitation biases in intermediary seasons are shown to occur in regions with high seasonal asymmetry like the Mediterranean. Moreover, winter and summer temperatures are respectively over- and under-estimated over the whole of France.

Two analogue subselection methods are therefore developed with the aim of keeping the structure of the SANDHY method unchanged while reducing those seasonal biases. The calendar selection keeps the analogues closest to the target calendar day. The stepwise selection applies two new analogy steps based on similarity of the sea surface temperature (SST) and the large-scale 2 m temperature (T). Comparisons to the Safran reanalysis over 1959–2007 and to homogenized series over the whole twentieth century show that biases in the interannual cycle of precipitation and temperature are reduced with both methods. The stepwise subselection moreover leads to a large improvement of interannual correlation and reduction of errors in seasonal temperature time series. When the calendar subselection is an easily applicable method suitable in a quantitative precipitation forecast context, the stepwise subselection method allows for potential season shifts and SST trends and is therefore better suited for climate reconstructions and climate change studies.

The probabilistic downscaling of 20CR over the period 1871–2012 with the SANDHY probabilistic downscaling method combined with the stepwise subselection thus constitutes a perfect framework for assessing the recent observed meteorological events but also future events projected by climate change impact studies and putting them in a historical perspective.

1 Introduction

Studying the influence of different modes of climate variability on hydrometeorological events like droughts and low flows requires numerous, long and reliable series to derive robust relationships with predictive capacity (see, e.g. Giuntoli et al., 2013; Boé and Habets, 2014). However, even in a data-rich country like France, few local observations covering the entire twentieth century are currently available in databases, as shown by Fig. 1. Less than one precipitation station per 600 km^2 was available on average between 1871 and 1930. The number of temperature observations shows an increase after 1950 only. The data sparseness before 1950 does not only come from the low number of meteorological

observations but also from their poor spatial coverage in spite of recent data rescue efforts (Moisselin and Jourdain, 2007). The recent release of global reanalyses extending backwards to the late nineteenth century provides some exciting opportunities to overcome the limitations of such an observation sparseness and to increase our understanding of local-scale climate variability. Global reanalyses, providing atmospheric data at coarse spatial and fine temporal resolutions, can be used to reconstruct local climate. A downscaling step is however required to bridge the gap from the reanalysis scale to the scale relevant for catchment hydrology for example. This can be done using statistical methods that establish a relationship between large-scale predictors from the global reanalyses and observed local-scale predictands, and the present work indeed follows this path.

Statistical downscaling methods have been widely used in France in various contexts, like quantitative precipitation forecast (see, e.g., Ben Daoud et al., 2011; Marty et al., 2012) or climate change impact studies (see, e.g., Timbal et al., 2003; Boé and Terray, 2008; Quintana-Seguí et al., 2011). Applications in reconstruction contexts were until recently limited by the availability of large-scale predictors from global reanalyses that covered only the second part of the twentieth century. As a consequence, existing studies for earlier periods generally focused on a specific event using rescued predictor observations. Yiou et al. (2014) for example reconstructed meteorological variables over the North Atlantic region during the Laki volcanic eruption in 1783 based on a specific gridded data set (Kington, 1988). Auffray et al. (2011) reconstructed hydrometeorological conditions having led to the 1859 flood of the Isere river (French Alps) based on purposely rescued pressure data. The release of the two extended global reanalyses – the Twentieth Century Reanalysis (20CR, Compo et al., 2011) and the European Reanalysis of the Twentieth Century (ERA-20C Poli et al., 2013) – spanning the entire twentieth century (respectively since 1871 and 1900) opens new ways to derive continuous time series of local weather thanks to downscaling methods. The 20CR reanalysis has already been downscaled for specific regions, as the south-eastern US by DiNapoli and Misra (2012) and Misra et al. (2013) using dynamical downscaling, the North East of Spain by Turco et al. (2014) using statistical downscaling, or the Durance river basin by Kuentz et al. (2013, 2015) using probabilistic statistical downscaling. It has also been recently used for downscaling wave climate over locations in the North-Eastern Atlantic (Camus et al., 2014) or station wind speed in Spain (Kirchner-Bossi et al., 2013). Few studies have been performed at a country scale, with the exception of Minvielle et al. (2015) who used a deterministic statistical downscaling approach optimized over France as a whole.

This paper proposes a daily high-resolution downscaling of the 20CR reanalysis to reconstruct precipitation and temperature gridded series from 1871 over France. Compared to previous downscaling studies, this experiment has three main combined specificities: (1) the 20CR reanalysis is downscaled for the whole of mainland France and Corsica; (2) it is performed with SANDHY (Stepwise ANalogue Downscaling method for HYdrology, Ben Daoud et al., 2011, 2016; Radanovics et al., 2013), a statistical downscaling method locally optimized over 608 zones paving France to take into account regional climate features; (3) the downscaling method provides 25-member ensemble reconstructed fields to reflect the uncertainty in the downscaling step.

The first objective of this paper is to present refinement steps through post-processing methods of the standard SANDHY method that were required in this context of historical reconstruction. Indeed, SANDHY has been mainly used in contexts of quantitative precipitation forecasting (Ben Daoud et al., 2011, 2016) and reanalysis of recent past precipitation (Radanovics et al., 2013). This is moreover the first time that this method is tested for downscaling temperature. The second objective is to test the quality of reconstructed ensemble fields with respect to different data sets including long homogenized series. The last objective is to illustrate some characteristics of the probabilistic reconstructed data sets through examples in both the temporal and spatial aspects.

The paper is structured as follows. Section 2 introduces the different reanalysis and observation data sets used during all the steps of the study. Section 3 describes the downscaling method SANDHY, the two alternative post-processing methods developed here as well as the performance scores used to assess the quality of the reconstruction. Section 4 compares reconstructed precipitation and temperature fields to different data sets, over a recent period as well as across the twentieth century. Examples of these reconstructions are then provided in Sect. 5 and results are finally discussed in Sect. 6.

2 Data

This section presents the large-scale data sets providing the large-scale atmospheric and oceanic data as well as the local-scale data sets providing the local-scale meteorological data used in this study.

2.1 NOAA Twentieth Century Reanalysis

Version 2 of 20CR (Compo et al., 2011) from the National Oceanic and Atmospheric Administration (NOAA) is used here as the source for atmospheric predictors in the downscaling step. Outputs from this global atmospheric reanalysis are available at 2.0° spatial resolution and 6 hourly temporal resolution from the 1 January 1871 to present. This reanalysis is the first global reanalysis spanning the entire twentieth century and only using the 6-hourly sea level pressure (SLP) from the International Surface Pressure Databank (ISPD v2.2.4, Compo et al., 2010) for the assimilation step and the monthly sea surface temperature (SST) and sea-ice concentration fields from the Hadley Centre global sea ice

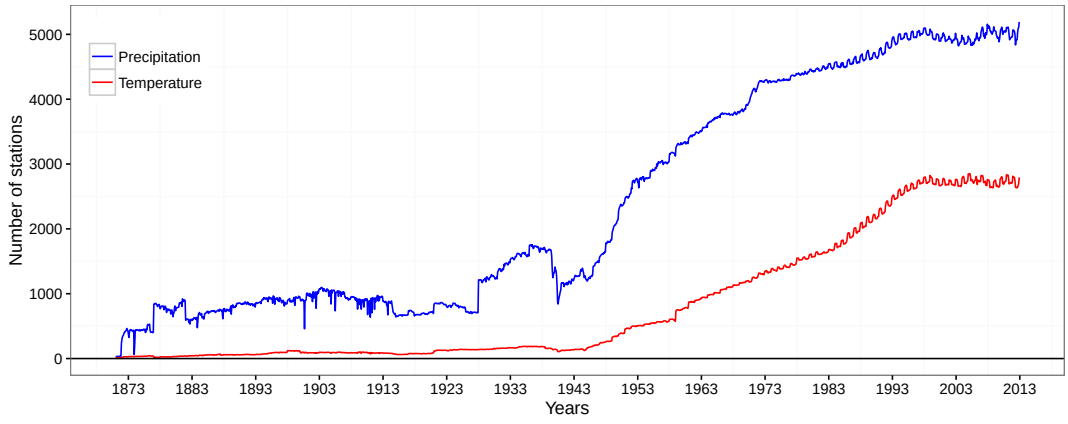

Figure 1. Evolution of the monthly averaged number of available precipitation and temperature stations in the Météo-France database (as of March 2015) since 1871.

and sea surface temperature (HadISST, Rayner et al., 2003) as boundary conditions. The feasibility of such a reanalysis has been demonstrated by Compo et al. (2006) and this choice has been done to avoid non-climatic bias due to the assimilation of observations from different systems. Outputs are available as either the original 56-member ensemble derived from the Ensemble Kalman Filter assimilation process or as the single ensemble mean. In this study, only the latter was used, as in all other downscaling applications to date. This will be further discussed in Sect. 6.

Six-hourly predictor variables from 20CR used by the SANDHY statistical downscaling method are the temperature at 925 and 600 hPa, the geopotential height at 1000 and 500 hPa, the vertical velocity at 850 hPa, the precipitable water content and the relative humidity at 850 hPa. The large-scale 2 m temperature (T) is also used as an additional predictor in the stepwise post-processing method described in Sect. 3.3.

2.2 NOAA Extended Reanalysis sea surface temperature

The second large-scale data set considered in this study is the NOAA Extended Reanalysis sea surface temperature version 3b (ERSST, Smith and Reynolds, 2003; Smith et al., 2008). This reanalysis is a global monthly sea surface temperature reanalysis, available at a 2.0° spatial resolution since the 1 January 1854. This reanalysis is derived from the Comprehensive Ocean–Atmosphere Data set release 2 (COADS, Woodruff et al., 1998) with missing data filled in by statistical methods. Like 20CR, this version does not use satellite data to avoid homogeneity biases.

ERSST provides the SST predictor used in the stepwise selection (see Sect. 3.3). A spline interpolation of the monthly means is considered to derive time series at the daily time step required by the downscaling method.

2.3 Safran

Safran French near-surface reanalysis data are used as predictands in the downscaling step. Safran is a meteorological reanalysis available at an 8 km spatial resolution and at the hourly temporal resolution from the 1 August 1958 to present (Vidal et al., 2010a). This data set has been computed using the Safran analysis system (Quintana-Seguí et al., 2008) which performs an optimal interpolation between all available surface observations in the Météo-France database and first-guess from the ERA-40 reanalysis (Uppala et al., 2005) for each of the 608 climatically homogeneous zones mapped in Fig. 2. A temporal interpolation step followed by a spatial interpolation step using an 8 km resolution orography are applied to obtain hourly data on a regular grid over France.

Vidal et al. (2010a) performed a detailed validation of the gridded Safran data set with both dependent and independent data. They showed that the errors on precipitation are low and constant over the 1958–2008 period. Errors on temperature are decreasing with the increasing number of available surface observations.

Daily precipitation is the predictand used in the SANDHY downscaling method. In this application and for the first time, temperature values from the analogue days are also considered for reconstructing downscaled temperature.

2.4 Surface observations

Raw meteorological surface observations are used here only for qualitative and limited comparison to reconstructed fields in Sect. 5.2. Long homogenized series are here favoured for quantitatively assessing the long-term quality and temporal homogeneity of meteorological reconstructions. Indeed, surface observations are generally affected by inhomogeneities due to their environment modifications over time. The French homogenized series (Moisselin et al., 2002; Moisselin and Schneider, 2002) have been derived using a statistical procedure detecting breaks and outliers (Caussinus and Mestre,

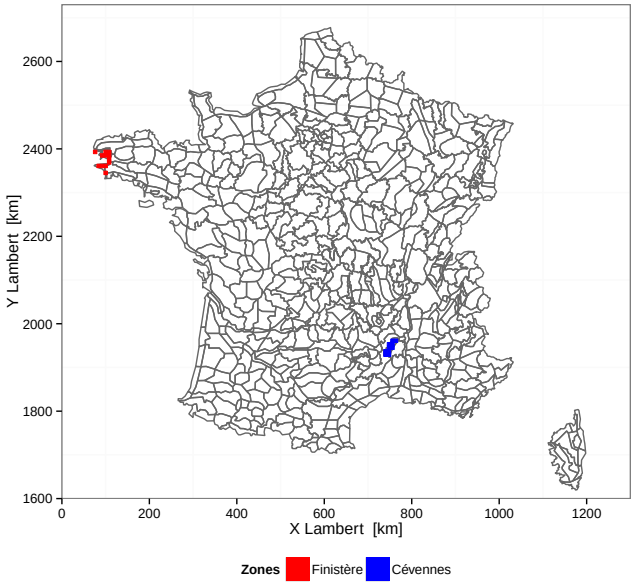

Figure 2. Map of the 608 climatologically homogeneous zones defined in the Safran data set. Coloured zones highlight case study zones used in Sect. 4: Finistère in Brittany and Cévennes in the Massif Central mountain range.

2004). Such series have been widely used for example for regional climate change detection (see e.g. Spagnoli et al., 2002; Ribes et al., 2010). The 323 monthly precipitation series and 65 monthly time series of minimum and maximum temperature spanning the whole twentieth century have been retained here. Monthly mean temperature time series have been obtained by averaging minimum and maximum temperature as done by Moisselin et al. (2002). Their locations are shown in Fig. 3. Note that no homogenized series are available in Corsica.

3 Methods

This section first presents the SANDHY downscaling method and the two post-processing methods developed. Synthetic diagrams in Appendix A summarize the different methodological steps described in this section as well as their sequence. The performance scores used in the following study are then presented.

3.1 SANDHY

The SANDHY method (Ben Daoud et al., 2011) follows an analogue approach based on the idea introduced by Lorenz (1969) that similar atmospheric situations lead to similar local effects. The analogue approach uses two concurrent data sets over an archive period, from a large-scale reanalysis and a local-scale meteorological data set. The large-scale reanalysis should also be available over the period to reconstruct, called the target period. Large-scale predictors (like atmospheric circulation patterns) from any date in the target period are compared to those of the archive period and dates with the most similar predictors are chosen as analogues. Local-scale variables (or predictands) like precipitation or temperature from the analogue dates are taken as plausible values for the target date. Numerous applications of the analogue method with different parametrizations have been developed over the last decades, notably in France (see, e.g., Dayon et al., 2015; Chardon et al., 2014; Yiou et al., 2014). Most of them use a unique analogy distance computed from one or several predictors to find the most relevant analogues, on the contrary to SANDHY that uses a stepwise approach where analogue subsets are sequentially refined with respect to different predictors.

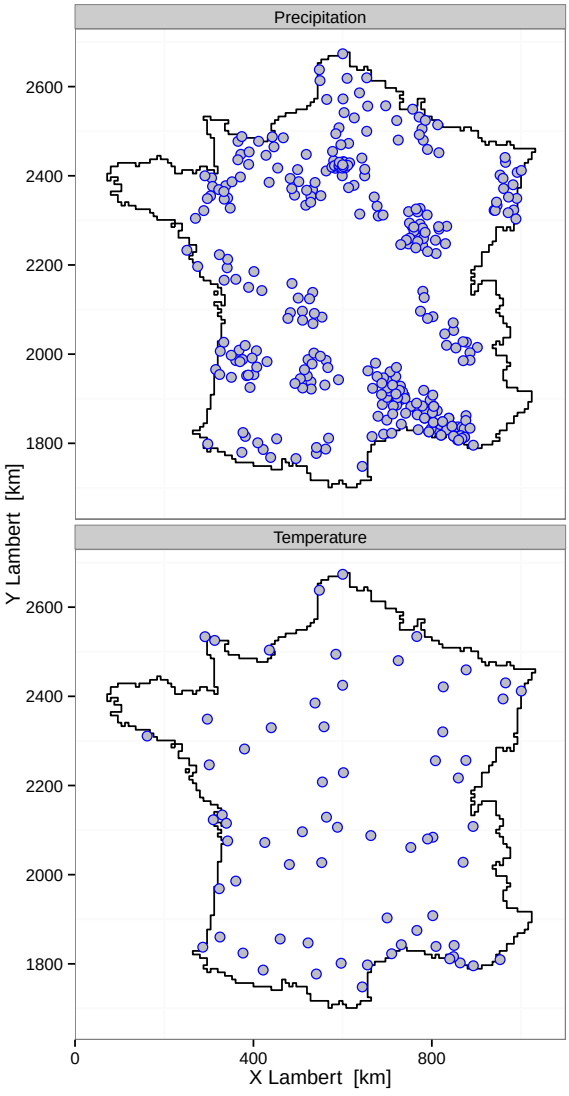

Figure 3. Spatial coverage of the available precipitation and temperature homogenized series produced by Météo-France during the full 1900–2000 period in France.

The SANDHY method is described by four analogy levels optimized by Ben Daoud et al. (2011, 2016) on the Seine and Saône basins and summarized in Table 1. The predictand considered is daily precipitation, as the initial aim of this method was quantitative precipitation forecast. The first step is a selection of N analogue days on temperature at 925 and 600 hPa with the exclusion of a 4-day window around the target date. N is taken as 100 × the number of years in the archive period. The second step is a selection on geopotential height at 500 and 1000 hPa with 170 analogue days retained. The third step selects 70 analogues thanks to an analogy on vertical velocity at 850 hPa and the final step selects 25 analogues on humidity, considered as the product of the precipitable water content and the relative humidity at 850 hPa. Predictors were extracted from ERA-40 and the precipitation predictand from Safran. The similarity criterion used for the analogy levels on temperature, vertical velocity and humidity is the Euclidean distance, with equal weights when different pressure levels are used. The analogy on geopotential height is measured through the similarity between fields shape with the Teweles and Wobus (1954) criterion. The synthetic diagram in Fig. A1 summarizes the different steps described above.

The spatial domain for a predictor where the analogy is looked for (spatial extent and position) should be defined and potentially optimized for each of the 608 climatically homogeneous zones shown in Fig. 2, considered as individual target locations, following Radanovics et al. (2013). The predictor domain for the first, third and fourth analogy levels is chosen as the closest large-scale grid point to each zone. Radanovics et al. (2013) moreover optimized the geopotential predictor domains with an algorithm of growing rectangular domains, leading to the selection of five near-optimum domains for each zone in France. The performance criterion for the optimization was the continuous ranked probability score (CRPS Brown, 1974; Matheson and Winkler, 1976), widely used for probabilistic verification forecast. Domains found from neighbouring zones were also considered if they provide a better performance. The different downscaling steps are therefore applied 5 times using a different predictor domain at each run, thus providing 125 analogue days for each target date and each climatically homogeneous zone in France.

3.2 Application of SANDHY for reconstruction

In this work, 20CR variables are used as predictors whereas Safran precipitation is kept as predictand. The predictor domains per climatically homogeneous zones have been here re-optimized over the same 20-year period as in Radanovics et al. (2013) – 1 August 1982 to 31 July 2002 – but with 20CR predictors in order to use the same reanalysis in all steps of the downscaling method. The target period considered is the whole period spanned by 20CR data, i.e. 1 January 1871 to 31 December 2012. Following Minvielle et al. (2015), the archive period has been chosen as the 50-year period running from 1 August 1958 to 31 July 2008 in order to maximize the pool of analogue situations. This specific period has previously been the target of several climatological assessments (see e.g. Vidal et al., 2010a, b). Reconstructions are expressed as an ensemble of 125 equally plausible gridded precipitation and temperature series over the whole 1871–2012 period, combining analogue days independently from one zone to another and from one day to the next. This set-up will be further discussed in Sect. 6.4.

Previous applications, including optimization, used ERA-40 variables as predictors. Tests preliminary to this work showed that the overall performance of SANDHY outputs in terms of CRPS are lower (around 0.05 mm day^{-1} on average) when using 20CR variables as predictors. These tests also showed that the precipitation for intermediary seasons in areas with a high seasonal asymmetry – e.g., Mediterranean areas including the Cévennes zone highlighted in blue in Fig. 2 – is not well simulated, with an over-estimation of precipitation in spring and an under-estimation of precipitation in autumn. It is important to note that these biases also occur using predictors from ERA-40. Moreover, winter and summer temperatures are respectively over- and under-estimated. This last result was not unexpected since SANDHY predictors were chosen for their strong relation to precipitation, and not temperature.

Post-processing approaches were therefore explored with the aim of keeping unchanged the structure of the SANDHY method while reducing those seasonal biases. Two analogue subselection methods were thus developed: for each target date, a specific number of analogue dates are retained among the 125 available ones.

3.3 Subselection methods

Subselection methods described below attempt to retain a number of analogues called $N2$ out of the initial 125 ones. When two successive subselection levels are used, the number of analogues are called $N1$ ($N1 \leq 125$) then $N2$, with $N2 \leq N1$. The synthetic diagram in Fig. A2 summarizes the different steps described below.

3.3.1 First predictor domain selection

The first method considered is the benchmark method against which the actual subselection methods described in the subsections below are tested. Radanovics et al. (2013) demonstrated that the five predictor domains optimized locally lead to very similar performances and can be considered as near-optimal. This benchmark subselection thus consists in retaining only the 25 analogue dates derived from the optimal predictor domain per zone. In the following, this method is called first domain selection.

Table 1. Characteristics of SANDHY predictors. N depends on the length of the archive period.

Predictor	Pressure level (hPa) and time (UTC)	Similarity criterion	Number of analogues
Temperature	925 at +36 h, 600 at +12 h	Euclidean distance	N
Geopotential	1000 at +12 h, 500 at +24 h	Teweles and Wobus Score	170
Vertical velocity	850 at +6 h +12 h +18 h +24 h	Euclidean distance	70
Humidity (PWC × RH)	850 at +12 h +24 h	Euclidean distance	25

3.3.2 Calendar selection

The second method considered is a calendar selection that keeps the $N2$ analogues closest to the target calendar day. $N2$ has been optimized to 25, based on the France-averaged skill score in root mean square error (RMSE) over precipitation and temperature monthly time series, with respect to the benchmark first domain selection (see Appendix C). This calendar approach had been initially included in the precursors of the SANDHY method where the search for analogue dates were restricted within a four-month window around the target date (Obled et al., 2002; Bontron and Obled, 2005). Ben Daoud et al. (2011) subsequently replaced this preliminary calendar step by a selection on large-scale temperature in order not to rule out potentially relevant analogue situations from one intermediary season to another and actually achieve better overall performances in terms of CRPS. In the following, this method is called calendar selection.

3.3.3 Stepwise selection

The third method follows the stepwise approach at the heart of SANDHY by adding two analogy levels, a first one on SST and a second one on the T. The SST and the T have been chosen because of their influence on both local precipitation or temperature.

The influence of the SST on continental precipitation has been studied for years, and according to Gimeno et al. (2010), 10 % of water evaporated from the oceans is transported to continents where it precipitates. To give some examples, Wilby (2001) used the North Atlantic ocean temperature as predictor in a downscaling method to reconstruct precipitation in the UK. Colman and Davey (1999) used the preceding winter North Atlantic ocean temperature to predict summer temperature, rainfall and pressure in Europe. van der Ent and Savenije (2013) identified the sources of continental precipitation and showed that the North Atlantic ocean and the Mediterranean sea are the regions contributing the most to European continental precipitation. Gimeno et al. (2010) identified what oceanic regions affect the most continental regions in terms of oceanic moisture. The North Atlantic ocean affects Europe in winter whereas the Mediterranean sea rather affects Europe in summer, and they both have an equal influence on Europe in spring and autumn.

Precipitation is not the only variable affected by the SST. Hoerling and Kumar (2003) and Cassou et al. (2005) demonstrated the contribution of the SST on the 2003 summer heat wave. Large-scale T, which contains direct information about the seasonal cycle of local temperature, is also selected to correct the local-scale temperature with a second subselection level. The other possible combination, with the T as first level and the SST as second level has also been studied but resulted in lower performances (not shown).

In practice, for each target date, the $N1$ analogue days giving the closest simulated SST to the observed one in terms of Euclidean distance are selected among the 125 available. Then, the $N2$ analogue days giving the closest simulated T to the observed one – also in terms of Euclidean distance – are selected among the $N1$ available. $N1$ and $N2$ have been optimized to respectively 80 and 25 (see Appendix C). For the sake of consistency over France and parsimony of the stepwise subselection method parameters, a single common grid point is considered for deriving the ERSST SST time series used for computing the SST analogy for all climatically homogeneous zones in France. This point has been chosen based on its highest correlation with precipitation over France (see Appendix B). It is located in the Atlantic ocean south of Brittany (4° W, 46° N). The grid point for deriving the time series for computing the T analogy is chosen as the closest land grid point to each climatically homogeneous zone, following the approach used for levels 1, 3 and 4 in the standard SANDHY method.

In the following, this method is called stepwise selection.

3.4 Performance indicators

The quality of precipitation and temperature reconstructions from the different methods described above was first assessed in terms of interannual regime by looking at annual and seasonal bias, and at the temporal evolution of root mean square error (RMSE). Rank correlation on annual and seasonal series were also computed to look at the natural variability derived from the hydrometeorological chain 20CR-SANDHY-subselection. All above indicators aimed at comparing reconstructions (1) to Safran pseudo-observations over the second part of the twentieth century, and (2) to homogenized series over the whole century.

An evaluation at multiple temporal scales of the three selection methods against Safran data has also been devised

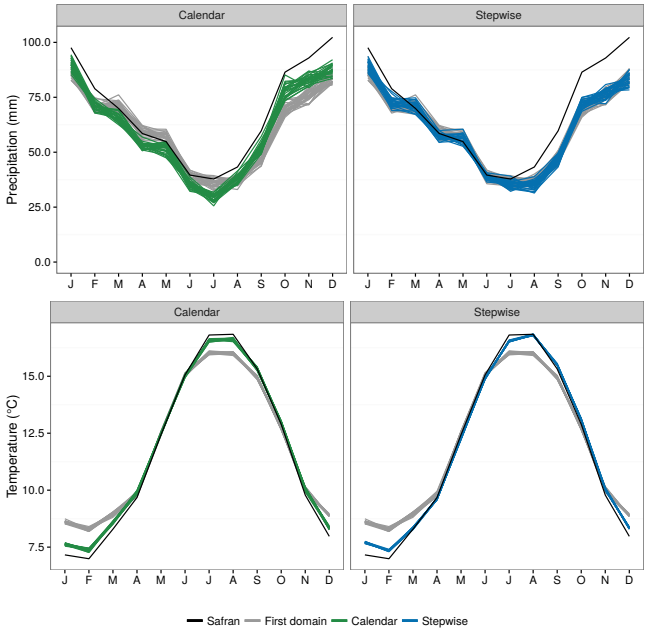

Figure 4. Precipitation and temperature interannual regimes between 1959 and 2007 for the Finistère case study zone from Safran and the three subselection methods, through their 25 individual realizations.

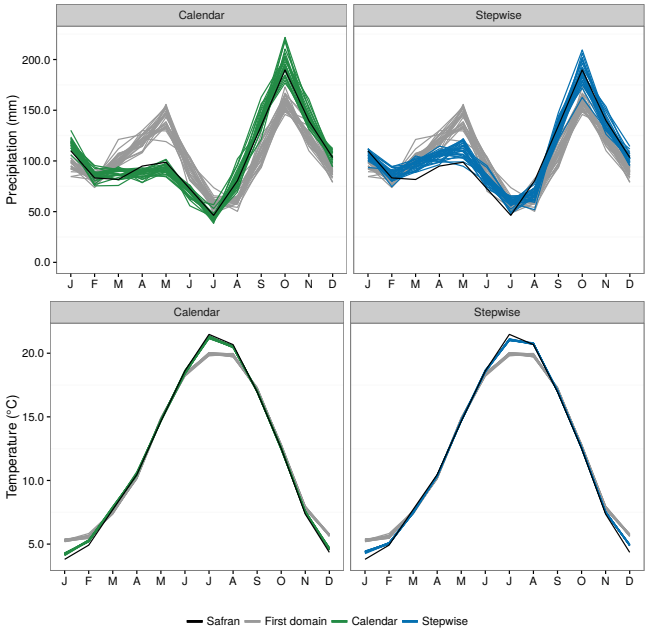

Figure 5. As for Fig. 4 but for the Cévennes case study zone.

based on the continuous ranked probability score (CRPS) as:

$$\mathrm{CRPS} = \int_{-\infty}^{\infty} [F(x) - H^0_{x_{\mathrm{obs}}}(x)]^2 \mathrm{d}x, \quad (1)$$

where $F(x)$ is the simulated cumulative distribution function of the variable x, x^0_{obs} the observed value and $H^0_{x_{\mathrm{obs}}}(x)$ the Heaviside function of $x - x^0_{\mathrm{obs}}$. This score is equivalent to the mean absolute error in a deterministic context. The CRPS was computed for each selection method at the daily monthly and annual scales and for each climatically homogeneous zone, and over the 1 August 1958–31 July 2008 archive period with reference to Safran data. The continuous ranked probability skill score (CRPSS) is used here for comparing zones with different climatological characteristics. It is computed as:

$$\mathrm{CRPSS} = 1 - \frac{\overline{\mathrm{CRPS}}}{\overline{\mathrm{CRPS}_{\mathrm{clim}}}}, \quad (2)$$

where $\mathrm{CRPS}_{\mathrm{clim}}$ is the reference climatological distribution over the archive period. For the daily timescale, the $\mathrm{CRPS}_{\mathrm{clim}}$ is calculated over the archive period using data from ± 60 days around the target date to take seasonality into account, as in Radanovics et al. (2013). This score is only used to compare the different subselection methods to each other. Therefore, the number of days chosen to compute the $\mathrm{CRPS}_{\mathrm{clim}}$ does not change the relative evaluation of the different methods and only serves to normalize results among zones with different magnitudes of precipitation and/or temperature values, and thus with different magnitudes of CRPS. For the monthly timescale, it is calculated from the interannual distribution of values from the specific month considered over the archive period. For the annual timescale, it is calculated from the interannual distribution of annual values over the archive period. The CRPSS thus expresses the skill of each selection method with respect to the information given by the climatology.

4 Performance assessment of the subselection methods

This section presents the assessment of reconstructed precipitation and temperature fields using SANDHY and the three subselection methods described in Sect. 3.3. As mentioned above, it consists of two parts: (1) a detailed spatial assessment with respect to Safran data over the 1959–2007 period, and (2) a more focused temporal assessment over the 1900–2000 period with respect to homogenized series.

4.1 Comparison to Safran

Figure 4 shows the precipitation and temperature interannual regimes for the Finistère case study zone (see Fig. 2) from Safran data as well as from the reconstructions based on the three subselection methods. Raw outputs from SANDHY (first domain selection) show a slight underestimation of precipitation in autumn and winter and a too flat seasonal cycle of temperature, as already mentioned in Sect. 3.2. Both the calendar and stepwise methods improve the temperature cy-

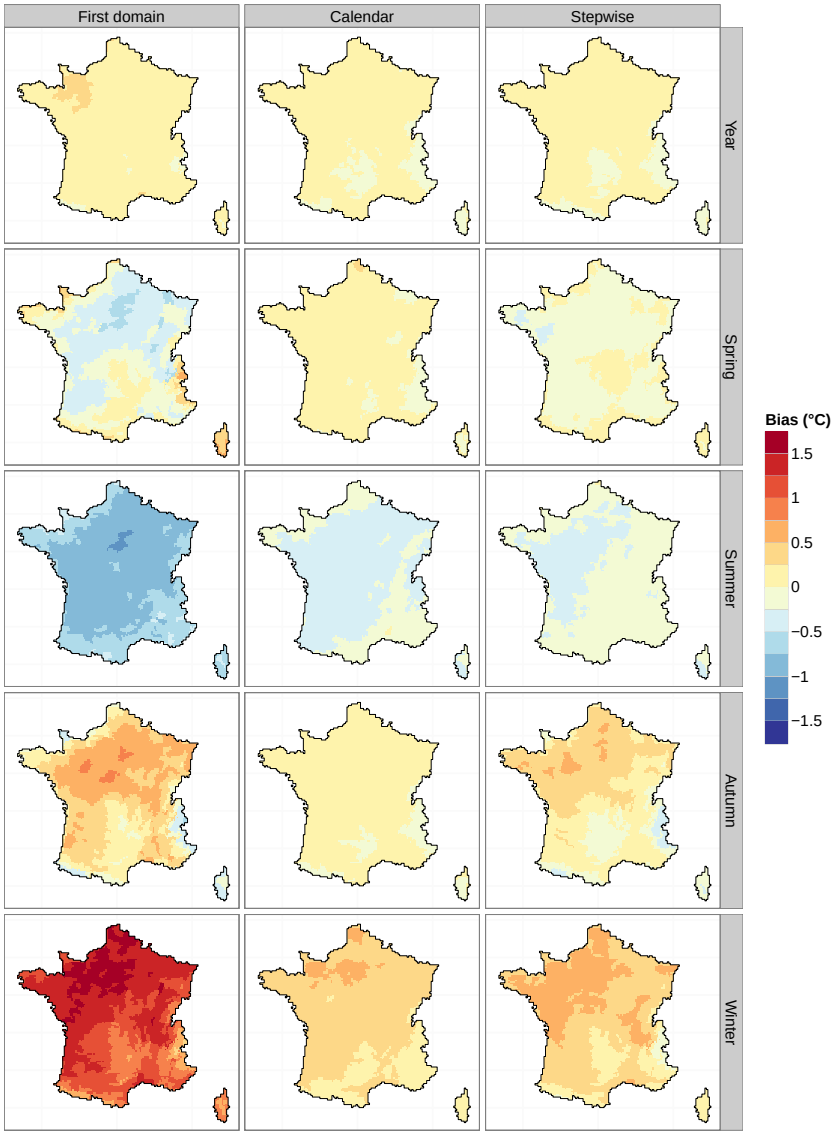

Figure 6. Median of annual and seasonal temperature bias between Safran and the three subselection methods for the 1959–2007 period. Red corresponds to an overestimation of the reconstructed temperature. Minimum bias of −1.0 °C for the first domain selection in summer and maximum bias of 1.7 °C for the first domain selection in winter.

cle, while only the calendar selection slightly improves the precipitation signal.

Figure 5 shows corresponding results for the Cévennes case study zone. The first domain selection shows a bias in the temperature signal similar to the other case study zone, and strong biases in precipitation in the two intermediary seasons. The latter is mainly due to the strong seasonal cycle in this region (see Vidal et al., 2010a) and the fact that SANDHY resamples analogue dates based on the temperature without considering the actual target season. As a consequence, a target date in spring may have analogue dates from autumn where precipitation is higher on average, leading to the positive bias in reconstructed spring precipitation. Both the calendar and the stepwise methods manage to drastically reduce such precipitation biases and to correct the temperature cycle.

Figure 6 extends the above results to the whole of France by showing the median annual and seasonal temperature bias from the three subselection methods. When all methods have a very low bias at the annual timescale, raw SANDHY outputs (first domain selection) show a strong positive (resp. negative) bias in winter (resp. summer), more pronounced in the northern part of the country. Both the calendar and stepwise methods bring sizeable improvements in all seasons and generally limits absolute biases under 0.5 °C except in winter in the northern half of France.

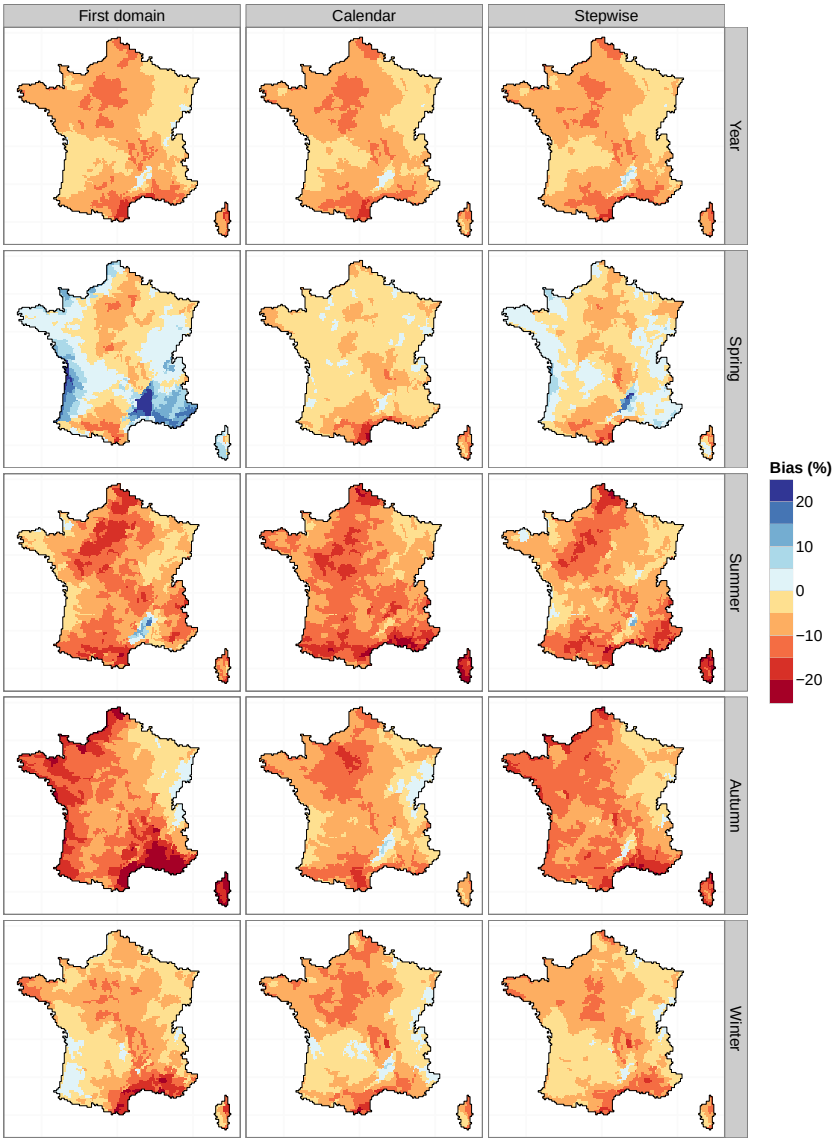

Figure 7. As for Fig. 6, but for precipitation. Red corresponds to an underestimation for precipitation. Minimum bias of −34 % for the first domain selection in autumn and maximum bias of 41 % for the first domain selection in spring.

Figure 7 shows in a similar way the precipitation bias from the three subselection methods. At the annual scale, biases are generally around −10 % (with a minimum around −20 % and a maximum around +5 %), with a slight overestimation in the Cévennes area and an underestimation elsewhere, more pronounced around the Mediterranean. Raw SANDHY outputs (first domain selection) show substantial biases in all seasons, with an overestimation in spring around the Cévennes area, the Atlantic and Mediterranean coasts and an underestimation in summer (except in the Cévennes area), autumn (except in the most eastern part of the country) and winter (except along the south-western coast). These spatial patterns are generally preserved by the stepwise selection but with reduced intensity. The calendar selection shows a similar spatial pattern in summer and winter but also more homogeneous biases over France in spring and autumn. This is most noticeable in spring when the calendar selection mainly shows absolute values under 10 %. The three methods show the same spatial pattern for the annual biases. In autumn, the calendar selection performs better than the stepwise selection while the opposite is observed in summer. Finally, relatively little improvement is proposed by the calendar and stepwise methods in summer (except in the Cévennes area) and winter (except along the Mediterranean coast).

The assessment of subselection methods in terms of interannual variability of temperature is shown in Fig. 8. The lowest rank correlation values are generally found in summer (especially in the eastern part of the country), followed

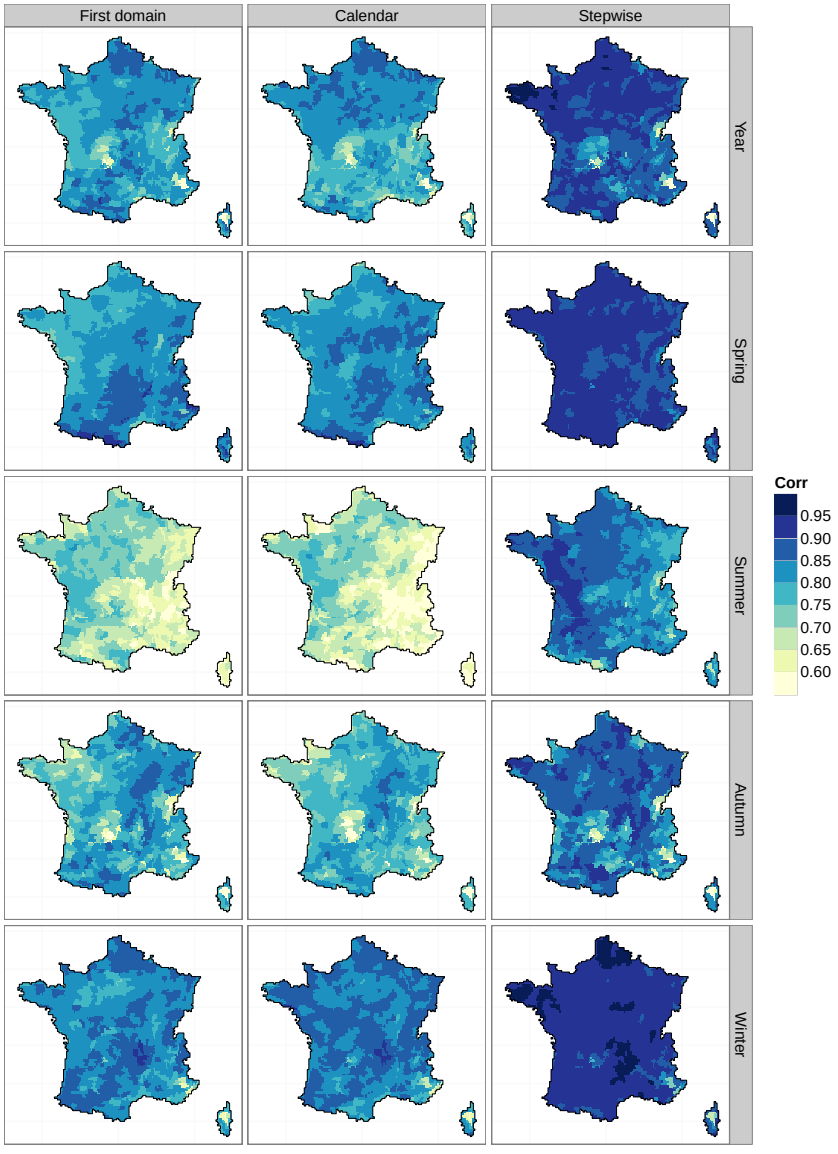

Figure 8. Median of the annual and seasonal temperature rank correlation between Safran all subselection methods for the 1959–2007 period. Minimum value of 0.35 for the first domain selection (year) and maximum value of 0.97 for the stepwise selection (year).

by autumn. For the first domain selection, values may reach 0.9 only in some regions in both spring and winter. Using the calendar selection does not provide any improvement and even reduces correlations in summer. On the other hand, the stepwise selection generates much higher correlations in all seasons all over the country, with values in spring and winter generally above 0.9, and a 0.1 to 0.2 improvement in summer with respect to raw SANDHY outputs.

Figure 9 shows the rank correlations in precipitation between Safran and the three subselection methods. The highest correlations are found in winter with values above 0.9 along the Atlantic coast, the north-east and the Alps. This spatial pattern – along with lower values around the Mediterranean and in the Loire and Allier valleys on the lee side of Massif Central mountains – may be found in all seasons. Summer correlations are generally under 0.6, presumably due to the prominence of convective events hardly explainable by large-scale predictors during this specific season. Correlations are very similar between the three methods, with only slight local differences hardly detectable on Fig. 9.

The above comparisons were only based on median values from the 25 realizations of each subselection method. Figure 10 now provides a probabilistic assessment of the quality of the reconstructions at different timescales by comparing the CRPSS (see Eq. 2) of raw SANDHY outputs (first domain selection) to the other subselection methods, for both temperature and precipitation. Considering precipitation at the daily timescale, CRPSS values are lower using the two

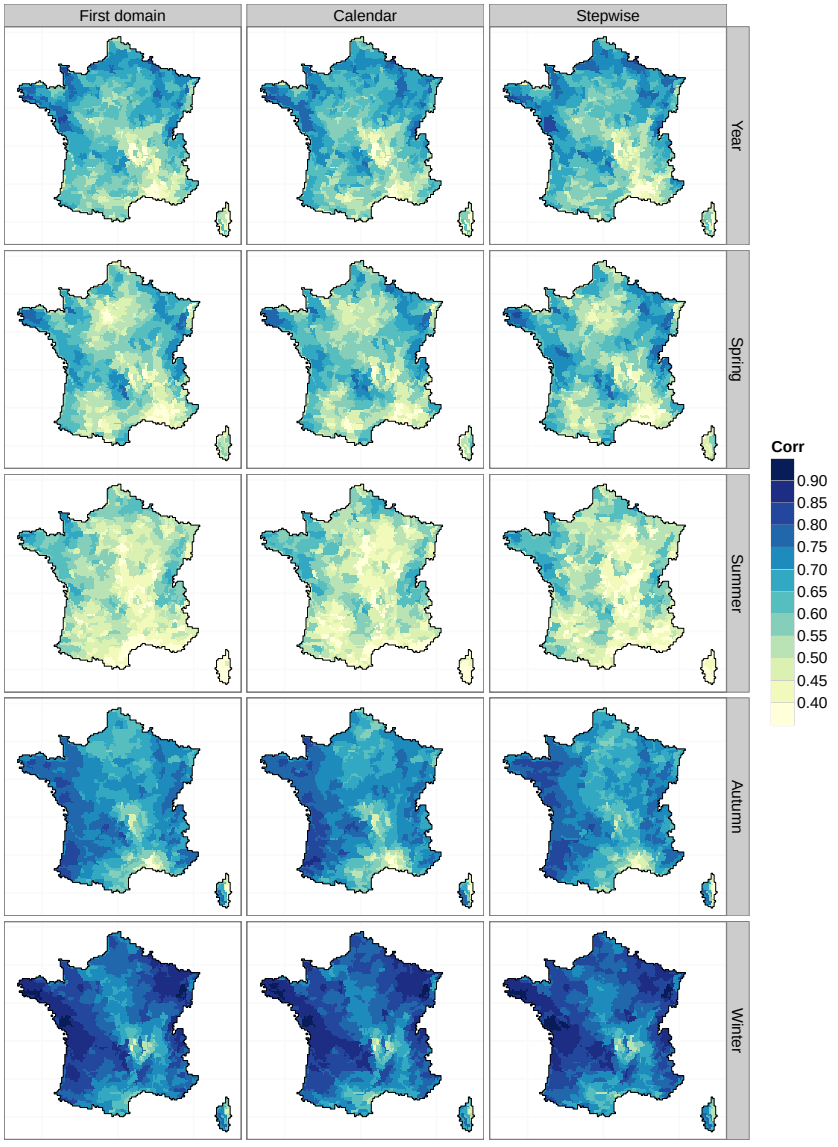

Figure 9. As for Fig. 8, but for precipitation. Minimum value of 0.19 for the first domain selection in summer and maximum value of 0.92 for the first domain selection in winter.

subselection methods rather than the first domain selection. Further investigations show that higher reliability values are obtained when considering the CRPS decomposition (Hersbach, 2000). Indeed, the two subselection methods lead to a more drastic selection of precipitation values than simply taking the first domain selection, which is representative of the 125 available values. As a consequence and because of the slight annual precipitation bias (see Fig. 7), observed values are more often outside the simulations range when the calendar and stepwise selection are considered. At larger timescales, the skill in reconstructing precipitation is generally slightly increased when using either the calendar or stepwise selections. Note that the average skill for all three methods is higher at the monthly timescale than at the daily or annual scale.

The bottom row of Fig. 10 shows CRPSS results for temperature. At the daily and monthly timescales, CRPSS values are strongly increased – and interestingly spatially homogenized over France – when using either the calendar and stepwise selections with respect to the first domain selection. The picture is rather different at the annual timescale, with only the stepwise selection generating higher skill scores without however homogenizing values spatially. The CRPSS increase is much stronger for the stepwise selection at all timescales, with an improvement of around 0.25 with respect to the first domain selection at the daily timescale, and up to 0.5 at the monthly timescale. It has to be noted that nega-

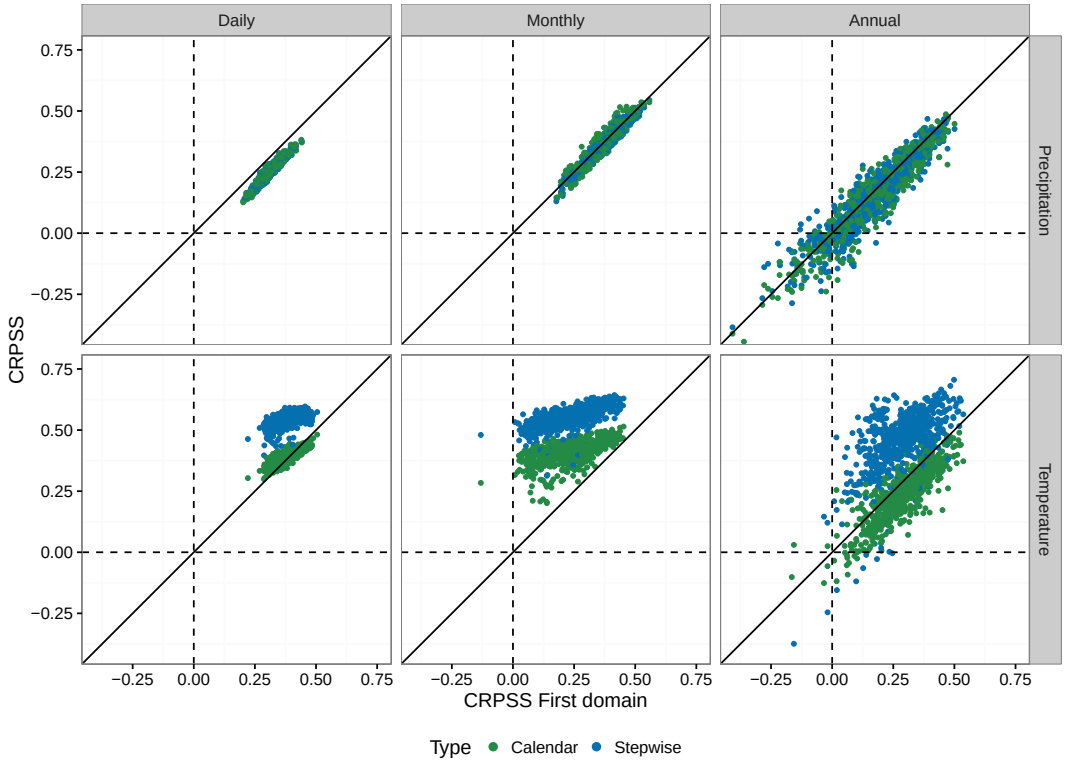

Figure 10. Daily, monthly and annual precipitation and temperature CRPSS for the calendar and stepwise methods compared to the first domain selection over the 1959–2007 period. One point for each climatically homogeneous zone mapped in Fig. 2.

tive CRPSS values are generated by the stepwise selection for a few specific and spatially scattered zones at the annual scale for temperature. A thorough investigation on these zones showed that the recent warming trend observed everywhere else in France is actually simulated by the stepwise selection while not being present on these Safran time series (not shown). This presumably originates from the inhomogeneities in Safran data that have already been documented by Vidal et al. (2010a) and that result from the evolution of the density of surface observations analysed by the system.

Except for the CRPSS, this section mainly compared the different subselection methods (First domain, Calendar and Stepwise) to Safran at the monthly/seasonal timescales where differences are the most prominent and improvements first needed. A comparison of daily cumulative distribution functions points to a similar overestimation of dry days whatever the method considered. Concerning daily temperatures, conclusions are very similar to those drawn from monthly temperatures (not shown).

4.2 Comparison to homogenized series

Figure 11 shows the annual and seasonal rank correlation of temperature between homogenized series described in Sect. 2.4 and all three subselection methods. Seasonal and spatial patterns follow the same conclusions highlighted by Fig. 8 over the 1959–2007 period. Correlations are indeed improved in a similar way by using the stepwise selection compared to the other two methods. Correlation values for the stepwise selection are around 0.05 lower than the ones with Safran data (1959–2007) plotted in Fig. 8 for all seasons except winter.

The annual and seasonal rank correlation of precipitation between homogenized series and the three subselection methods are shown in Fig. 12. Correlations are very similar amongst the three methods, like they were with Safran as a reference over a shorter and recent period (see Fig. 8). Values from all seasons are again slightly lower (around 0.05) than the ones obtained with reference to Safran over the 1959–2007 period.

The evolution of the RMSE for both Safran time series and the reconstructions with all three subselection methods, using the homogenized series as a reference, are presented in Fig. 13. Monthly RMSE are first calculated and then averaged yearly for each realization. To have an idea of the average daily error for precipitation, the RMSE in mm year^{-1} is divided by the number of days in each corresponding year to get an RMSE in mm day^{-1}. Ribbons then show the range between minimum and maximum annual values of the 25 realizations from each method. Looking first at precipitation results, reconstruction errors are relatively constant over time, with an average value around 1.3 mm day^{-1}. Errors are how-

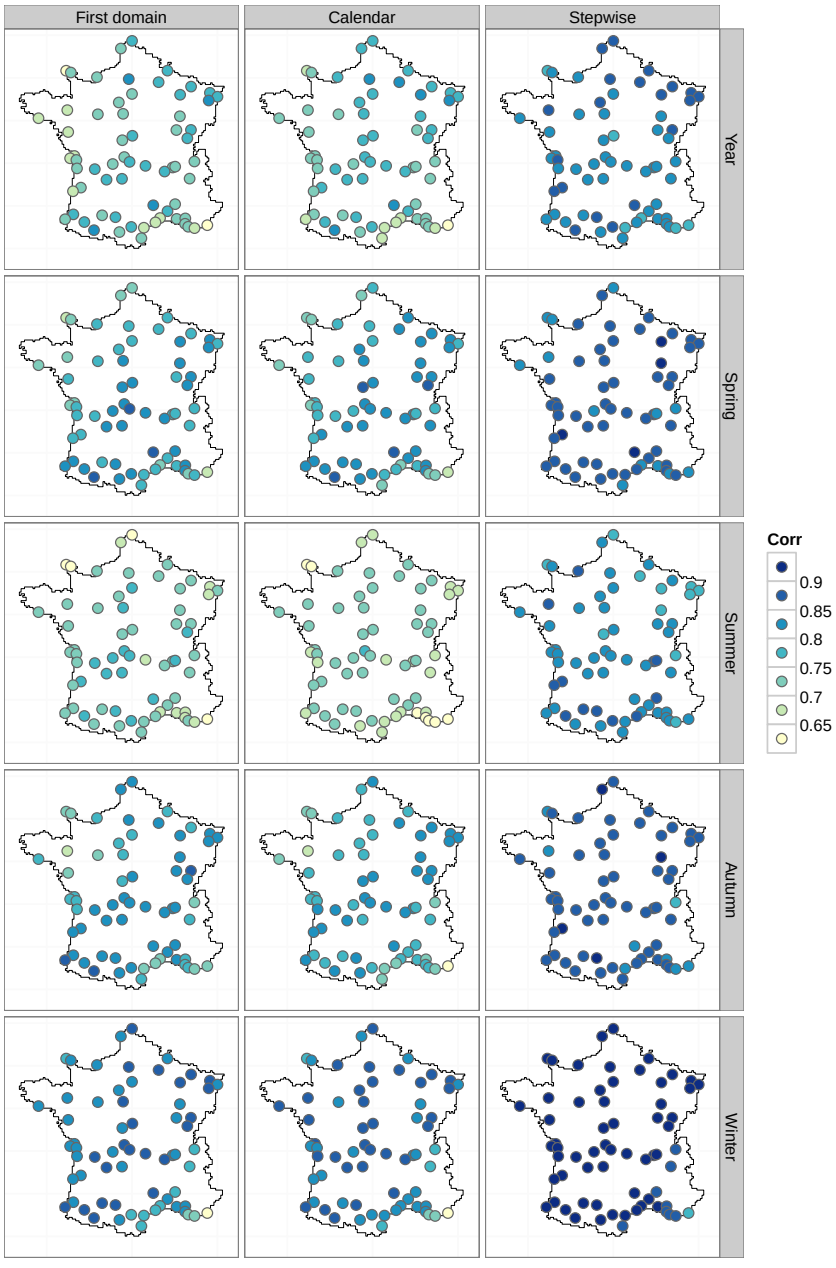

Figure 11. Median of the annual and seasonal temperature rank correlation between homogenized series and all three subselection methods for the 1900–2000 period. Minimum correlation of 0.58 for the first domain selection in winter and maximum correlation of 0.95 for the stepwise selection in winter.

ever slightly higher before the 1940s. It is important to note that Safran errors account for more than one-third of reconstruction errors. These errors have been documented by Vidal et al. (2010a) and stems mainly from the spatial interpolation step of the analysis system, but also in part here from the temporal inhomogeneities in Safran input data. Note that RMSE values provided by Vidal et al. (2010a) and Quintana-Seguí et al. (2008) were respectively computed at the daily and hourly timescale and are thus not directly comparable to the ones in Fig. 13 that are computed at the monthly timescale. The stepwise selection generates slightly lower errors than the two other methods.

The picture is quite different for temperature. Figure 13 indeed shows large differences between the three selection methods, with an average RMSE of 1.3 °C for the first domain selection, 1.1 °C for the calendar selection, and 1.0 °C for the stepwise selection. The latter even reaches 0.9 °C over 1959–2000 where Safran errors are around 0.8 °C. The step-

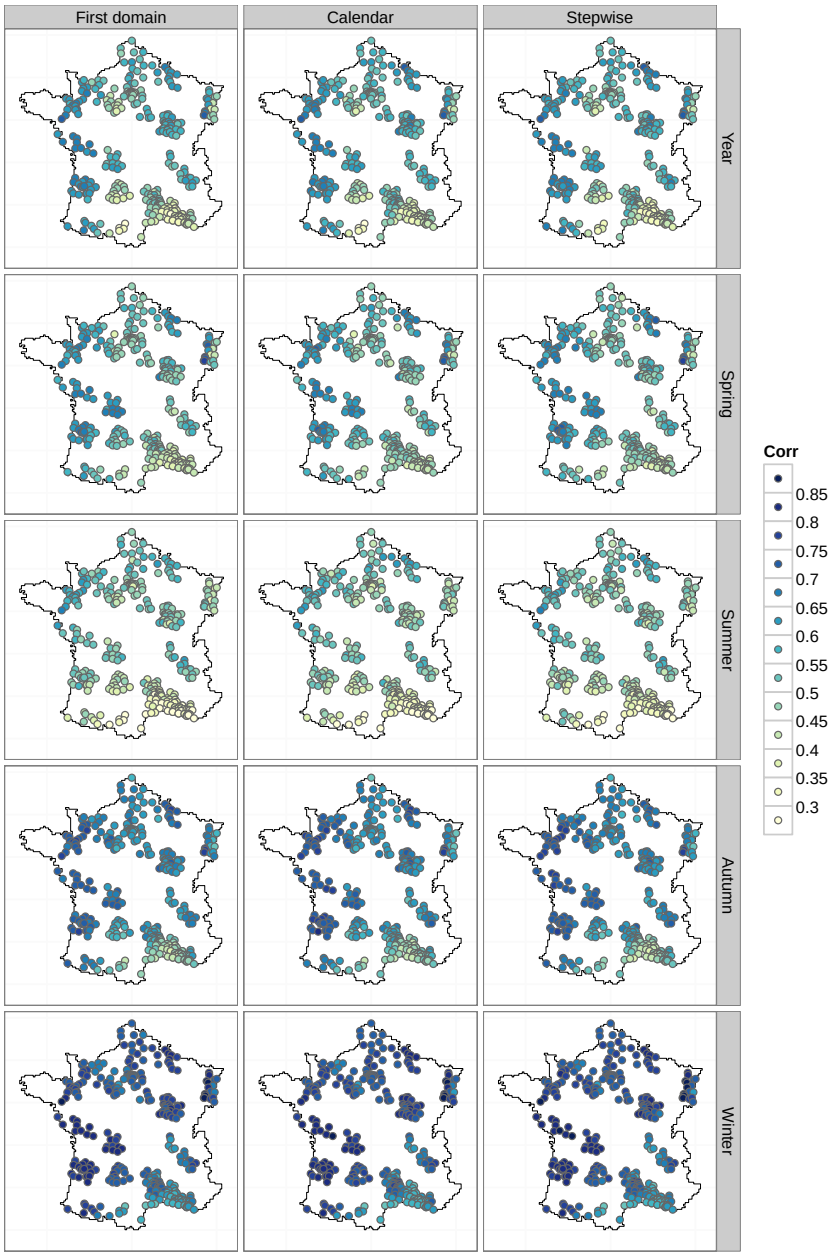

Figure 12. As for Fig. 11 but for precipitation. Minimum correlation of 0.10 for the calendar selection in summer and maximum correlation of 0.91 for the first domain selection in winter.

wise selection thus clearly outperforms the other two methods for temperature, with errors of the same order of magnitude as Safran itself.

5 Reconstruction examples

This section presents two examples of reconstructed precipitation and temperature through the stepwise selection. This specific method has been selected for its higher performance compared to the other subselection methods (see Sect. 4). This section aims at providing qualitative insights on the reconstructed outputs and at looking in more detail into features along (1) the local-scale temporal dimension and (2) the spatial dimension for a temporal snapshot. The focus is here on the monthly and/or seasonal timescale at which reconstructions are expected to deliver relevant additional information.

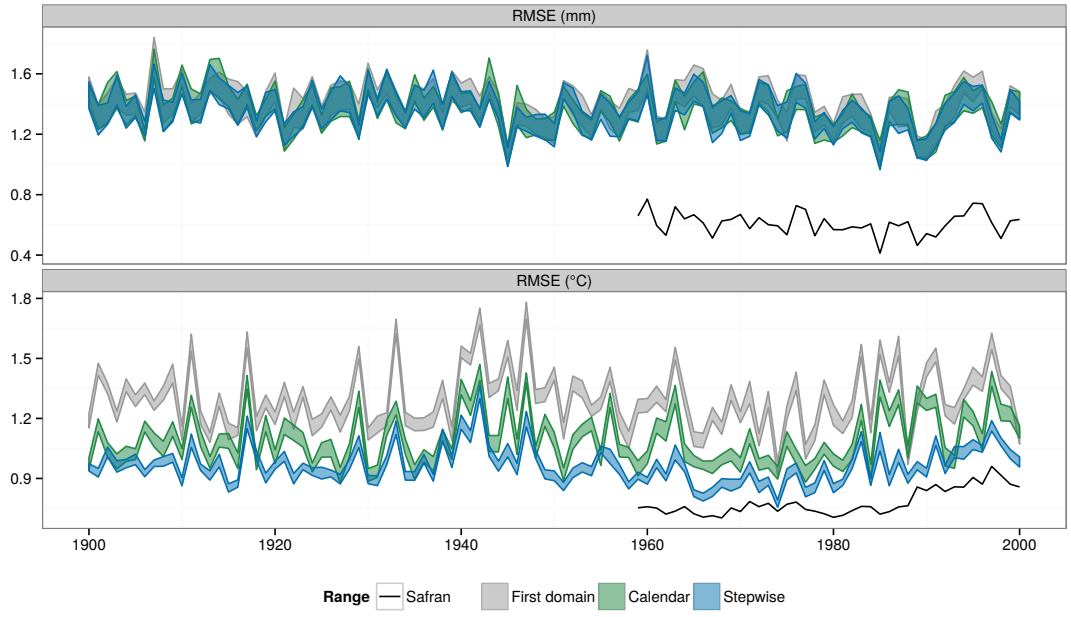

Figure 13. Temporal evolution of the precipitation and temperature RMSE for both Safran and the reconstructions with all three subselection methods, with the homogenized series as a reference. Values are initially computed at the monthly timescale. See text for details.

5.1 Temporal reconstruction: time series at Paris Montsouris

Paris Montsouris is a station with available data since 1873 with good quality data due to the low number of detected breaks over the years (Moisselin et al., 2002). This station has for example been used by Slonosky (2002) to reconstruct 3 centuries of precipitation data in Paris blended from different sources. Homogenized temperature and precipitation data from this station have also been used to perform climate variability assessments (Dieppois et al., 2013; Lüdecke et al., 2013). It is included in the list of homogenized stations described in Sect. 2.4 and used in Sect. 4.2.

Figure 14 compares annual and seasonal evolution of temperature from the Paris Montsouris homogenized series, Safran data and the range of reconstructed series from the stepwise selection. The variability is very satisfactorily simulated at the annual timescale, in coherence with Figs. 8 and 11. The uncertainty in reconstructions seems however underestimated, with observations being too often out of the range of the 25 realizations. Winter temperature reconstructions moreover suffer from a hot bias, leading to a systematic overestimation of temperature in cold years. However, the recent trend in spring temperature is well captured by the reconstruction method.

Some more information may be drawn from examining specific features of the seasonal time series. Some extremely cold seasons are reasonably well simulated, like winter and spring 1963 (see e.g. Greatbatch et al., 2015). The temperature for other events like the extremely cold winter 1879–1880, and especially December 1879 (Angot, 1881) are however largely overestimated with, for example, an overestimation of approximately 4 °C in December 1879. In this particular case, this may be due to the peculiar convergence of meteorological features (snow cover feedback, temperature inversion, etc., see U., 1884; Angot, 1881) that may hardly be captured by 20CR and/or our downscaling method. Extremely hot seasons are generally well simulated, especially recent summers like 1976 (Brochet, 1977) or 2003 (Trigo et al., 2005), but also much older events like spring 1893 (Plumandon, 1893).

Figure 15 shows the corresponding evolution of precipitation for Paris Montsouris. As for temperature, the interannual variability is satisfactorily simulated at the annual timescale, in coherence with Figs. 9 and 12. The main difference lies in the uncertainty range which is much higher than for temperature, leading to a quite reliable reconstructed ensemble. The reconstruction method has unsurprisingly difficulties in simulating the extreme wet autumn of 1896 (Mascart, 1898). Nevertheless, other extreme events like the dry summer 1949 (Sanson, 1950) or the exceptionally dry year of 1921 (Duband et al., 2004) are well simulated, even if observed values are consistently overestimated considering the median of the simulations.

For both precipitation and temperature, springs and summers around year 2000 are underestimated in comparison to homogenized time series. This specific discrepancy may be related to soil-atmosphere retroactions which are not taken into account in the one-way downscaling approach considered.

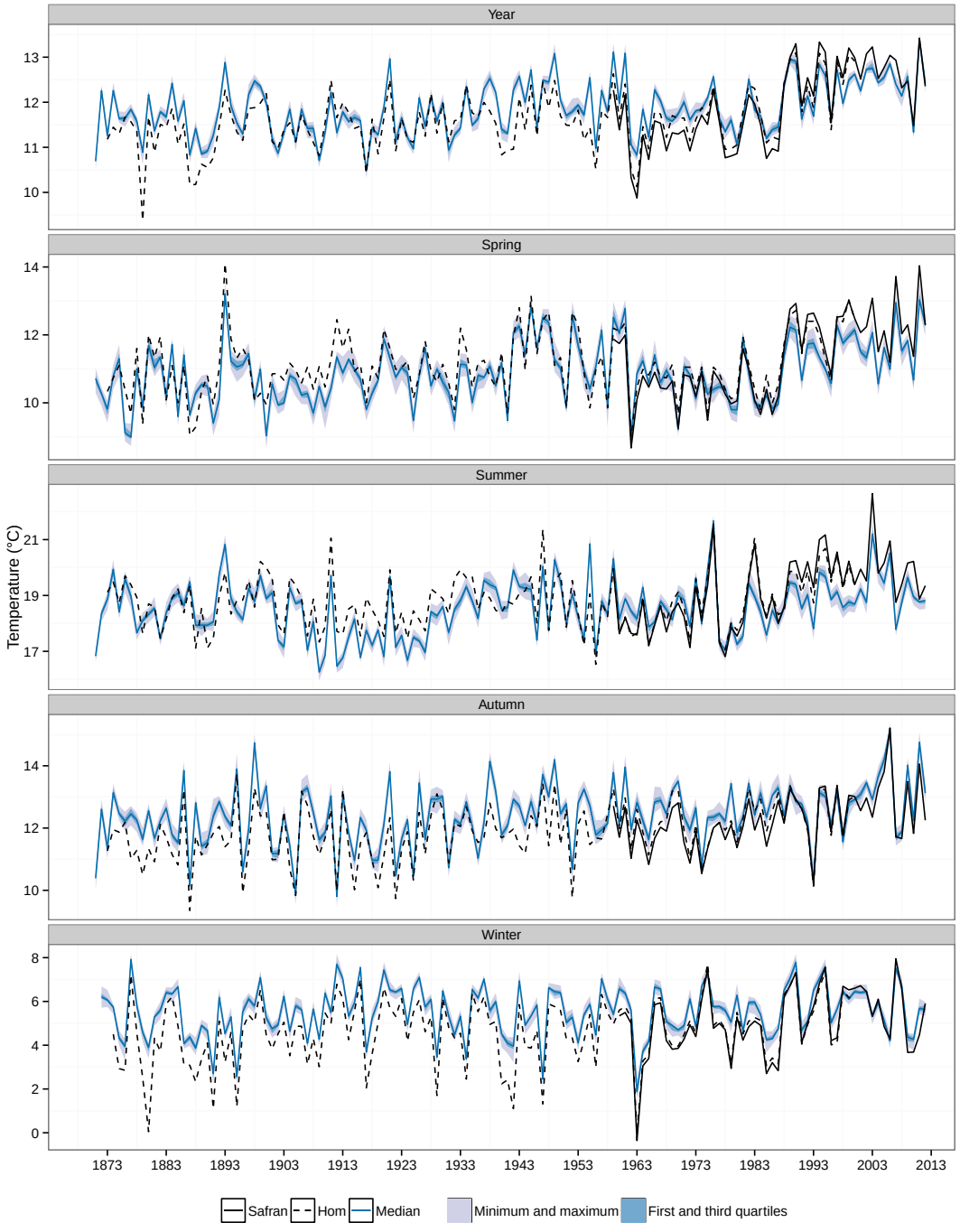

Figure 14. Paris Montsouris temperature homogenized time series, corresponding Safran data and reconstructed series from the stepwise selection at the annual and seasonal timescales over the 1871–2012 time period. Grey and blue ribbons define the range and the interquartile range, respectively, of values from all 25 realizations from the stepwise selection. Note the different scales for the y axes.

5.2 Spatial reconstruction: December 1909 precipitation over France

December 1909 precipitation over France has here been chosen as an example reconstruction based on its hydrological and socio-economic consequences. Indeed, high (but not extreme) precipitation amounts fell in this particular month over the north-eastern part of France (Angot, 1911). At the end the month, soils were saturated and heavy rain in January 1910, combined with frozen grounds and melting snow, led to widespread floods in the region (Lang et al., 2013). December 1909 thus appears as a key factor for understanding the early 1910 floods (Schneider, 1997), and especially

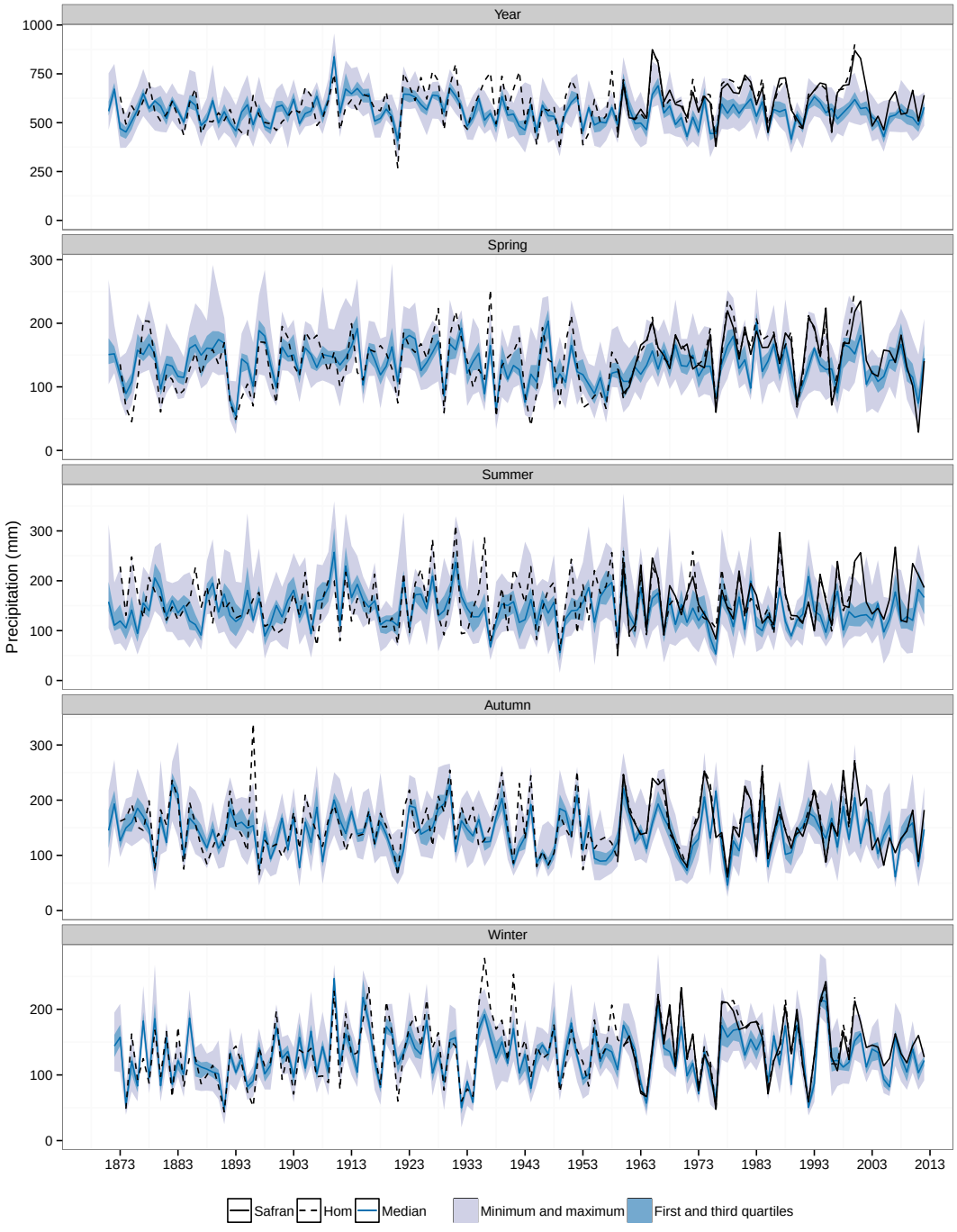

Figure 15. As for Fig. 14 but for precipitation.

the most studied 100-year flood of the Seine in Paris (see e.g. Marti and Lepelletier, 1997) but also other important floods on the Rhine (Martin et al., 2011) or Rhône (Pardé, 1925) tributaries.

Figure 16 shows the map of December 1909 precipitation adapted from Angot (1911, Pl. III) and compiled with observations available at that time. Heavy and unusual amounts have been recorded over several mountain ranges, (Morvan, Southern Vosges, Jura and Northern French Alps). The map of observations currently available in the Météo-France database is shown in the top row of Fig. 17 with the same colour scale. The Seine basin is well covered with observations, following the Météo-France data rescue efforts for understanding the flood in Paris (see the "extreme rainfall"

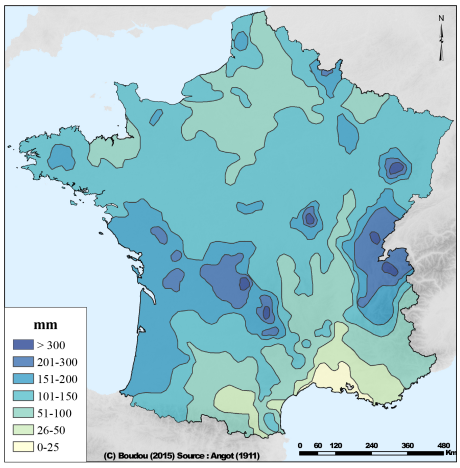

Figure 16. Map of December 1909 precipitation, adapted from Angot (1911).

database and website[1]). However, several regions appear as void of observations: Picardie (north), the main part of the Loire basin (centre), Jura mountain range (east), most of the Alps, and Provence (south-east).

Rows 2 to 4 in Fig. 17 show maps compiling local minimum, median and maximum of the 25-member ensemble monthly precipitation amounts for December 1909 from the stepwise selection. Spatial patterns are in very good agreement with the observations from both Fig. 16 and the first row of Fig. 17. Observed values are very close or slightly lower than median simulated values, which demonstrates the small bias of the method. The last three maps in Fig. 17 also confirm the occurrence of large rainfall amounts in the areas void of currently available observations, especially the Jura mountain range where an exceptional flooding of the town of Besançon was recorded in January 1910 (Allard, 1910; Boudou et al., 2016).

6 Discussion

6.1 Homogeneity and uncertainty of 20CR predictors

This section aims at summarizing literature findings about potential inhomogeneities of 20CR predictors and associated uncertainties that may directly affect downscaled reconstructed products assessed in Sect. 4.

Compo et al. (2011) compared 20CR to other reanalyses like ERA-40 (Uppala et al., 2005), ERA-Interim (Dee et al., 2011) or NCEP/NCAR (Kalnay et al., 1996) and demonstrated its good quality on common recent periods. However, this quality should actually be lower in the first part of the simulated period due to the lower overall density of assimilated SLP data (Compo et al., 2010). Ferguson and Villarini

[1]http://pluiesextremes.meteo.fr/1910-01-18/crue-historique-de-la-seine.html

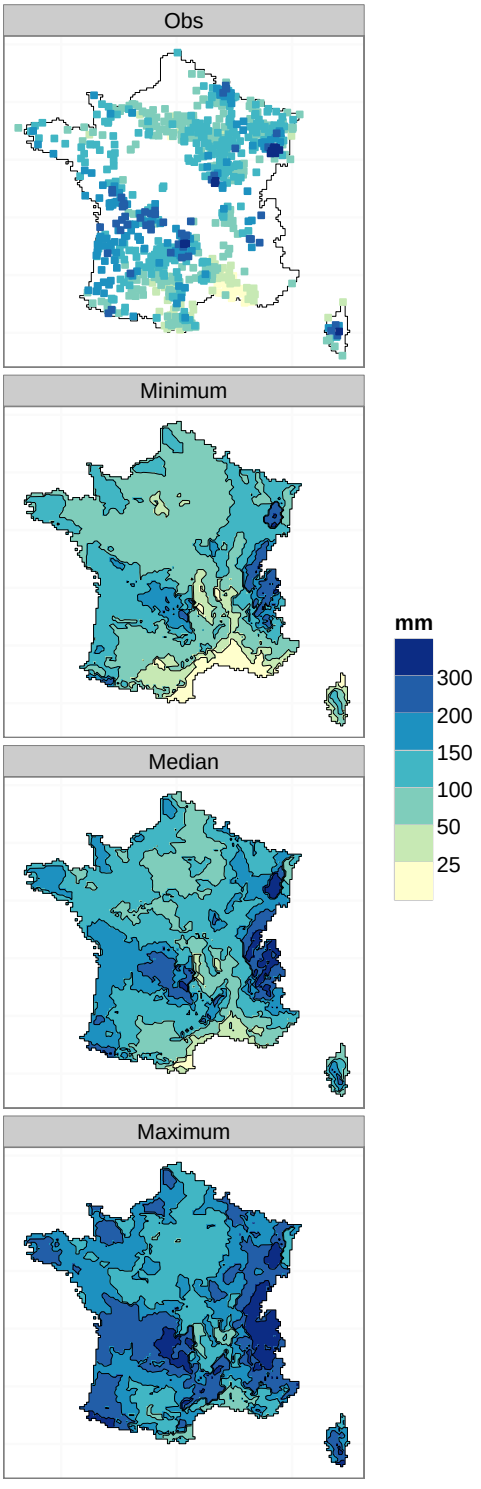

Figure 17. December 1909 precipitation. Top: observations currently available from the Météo-France database. Below: minimum, median and maximum monthly precipitation from the 25 realizations of the stepwise selection. Maximum values from observations, minimal, median and maximum simulations of 520, 409, 533 and 625 mm respectively.

(2014) and Lee and Biasutti (2014) showed that such impacts are rather limited on surface variables over Western Europe thanks to the good early data coverage in ISPD (see Yin et al., 2008, Appendix C). This notably increases our confidence on the T used as an additional predictor in the stepwise subselection method. However, no comprehensive study focused on upper-air variables such as the ones used as predictors by the original SANDHY downscaling method like geopotential heights. Several authors also looked in detail into the homogeneity of different indices of storminess in Northern Europe, with features not very distant from the shape of the 1000 hPa geopotential height used as predictor from SANDHY. Krueger et al. (2013) found that extreme geostrophic wind speed over the North Atlantic suffers from inhomogeneities that are likely due to the increasing number of observations assimilated into 20CR over time. Wang et al. (2013) also found some inhomogeneities in cyclone activity in the 20CR ensemble mean before the 1930s in Northern Europe (see also the discussions in Krueger et al., 2013; Wang et al., 2014). All studies mentioned above are in accordance with our findings in Figs. 13 to 15 which show systematic biases before the 1930s/1940s.

The 20CR inhomogeneities discussed in the references above partly relate to using the ensemble mean as in the present study. Wang et al. (2013) for example showed that the evolution of cyclone activity over the Northern Hemisphere is dramatically different when considering the ensemble mean or individual members, with the latter giving much more homogeneous results over the twentieth century. Krueger et al. (2013) also found that the spread between ensemble member extreme geostrophic wind speed increases before the 1940s in the North Atlantic. These findings strongly suggest that downscaling studies should benefit from using predictors from individual 20CR members instead of the ensemble mean in order to propagate the reanalysis uncertainty and hopefully reduce systematic biases in the late nineteenth–early twentieth century. However, doing so would first require all specific predictor variables to be available from individual members. In the case of SANDHY, one should necessarily use a degraded version as the vertical velocity at 850 hPa is not available. Second, the computation time would therefore dramatically increase compared to downscaling the ensemble mean only.

6.2 Other potential sources of reconstruction uncertainty

The assessment of reconstructions in the beginning of the target period presented in Sects. 4.2 and 5.1 may be influenced by other factors. For example, homogenized series taken in these sections as long-term references may suffer from issues related to the homogenization method used. Indeed, Moisselin et al. (2002) pointed out inconsistencies in the evolution of linear precipitation trends due to the homogenization performed independently inside different regional administrative divisions. New data sets of homogenized series are being developed by Météo-France in order to remove such inconsistencies but are currently limited to the second part of the twentieth century (Gibelin et al., 2014).

Besides such issues, the discrepancies in reconstructed and observed precipitation in the late nineteenth century (e.g. in Figs. 13 to 15) might also be due to peculiar teleconnections with the Tropics (Bichet et al., 2014) that may weaken the temporal transferability of the SANDHY downscaling method.

6.3 Choosing the appropriate reconstruction method

Section 4 compared three different post-processing methods of SANDHY outputs. The paragraphs below attempt to provide some advice in using one or another method in different downscaling contexts.

Section 4 showed that the use of either the calendar and stepwise methods strongly reduces precipitation and temperature biases in comparison to the first domain selection (raw SANDHY outputs). Even if the stepwise selection shows better performances for temperature reconstructions – and especially in terms of rank correlations or RMSE – the calendar selection may be retained in specific cases. First, if one is interested in simulating precipitation only, the calendar selection offers a much easier implementation, without the need to resort to additional predictors (T and SST). This may for example be the case in the context of quantitative precipitation forecast for which SANDHY has been originally developed.

Nevertheless, the stepwise selection appears much more suitable and adapted to contexts where seasons are exposed to potential shifts, like in long-term historical reconstructions as here or in climate change impact studies. Indeed, the selection of analogue days may in this case closely follow the decadal to multi-decadal variability as well as the anthropogenic warming signal carried out by both the T and the SST.

6.4 Using the reconstruction data sets

As mentioned in Sect. 3.2, analogue days are randomly combined from one climatically homogeneous zone to the next for all three subselection methods. This specific feature may lead to a lack of spatial continuity in the associated meteorological time series. For local studies (e.g. a specific climatically homogeneous zone), precipitation and temperature data can be used at a daily timescale. For regional to national studies, using a monthly scale like in Fig. 17 balances the random sampling effect observed at the daily scale. On-going works on different and very promising methods will allow resampling local ensemble members to derive an ensemble of spatially coherent daily precipitation and temperature fields.

7 Conclusions

This paper describes a daily high-resolution reconstruction of precipitation and temperature fields since 1871 over France through a statistical downscaling of the Twentieth Century Reanalysis. The motivation for this study comes from the lack of available surface observations before the 1950s and the poor understanding of local-scale climate variability from the late nineteenth century onwards.

The downscaling of the 20CR with SANDHY provides a daily set of 125 analogue dates for each day in the 1871–2012 period, over 608 climatically homogeneous zones paving France. The ensemble of analogue dates is then converted into probabilistic meteorological daily fields. A warm bias in winter and a cold bias in summer are identified over France in SANDHY reconstructed series. Additionally, wet and dry biases in intermediary seasons are identified for regions with a high seasonal variability. Post-processing approaches have been developed with the aim of keeping the structure of the SANDHY method unchanged while reducing those seasonal biases. Two analogue subselection methods are therefore developed. The stepwise selection adds two new analogy levels with a first selection of 80 analogues based on the ERSST SST similarity and a second selection of 25 analogues based on the 20CR T similarity. The calendar selection selects the 25 analogues closest to the target calendar day. These two approaches are then compared to the 25 analogues derived from the first predictor domain considered as representative of the 125 analogues from SANDHY outputs.

A performance assessment using both dependent and independent data showed an improvement in terms of precipitation and temperature biases, mean errors and rank correlations using both subselection approaches with respect to the first domain selection. The lower quality of reconstructions identified before the 1930s might be explained by potential inhomogeneities of 20CR predictors. Few differences are observed between the two subselection approaches considering precipitation data. However, the stepwise selection leads to higher rank correlations and lower mean errors between reconstructed and observed temperatures. Even without assimilating T observations, 20CR outputs obviously provide quite useful information on this variable. Examples of a temporal and spatial reconstruction outputs from the stepwise selection showed a persistent warm bias in winter and some issues to reconstruct seasonal extreme values, but also the overall good quality of the reconstructions.

The two approaches developed in this study can be applied in different contexts: the calendar selection is an easily applicable method ideal in a quantitative precipitation forecast context whereas the stepwise selection is more suitable for historical reconstructions or climate change studies with potential season shifts. The probabilistic downscaling of 20CR will constitute a perfect framework for assessing the recent observed events but also future events projected by climate change impact studies and putting them in a historical perspective. More specifically, the coupling between the SANDHY probabilistic downscaling method and the stepwise subselection – called SANDHY-SUB from now on – may be used in various contexts, such as past hydrological studies or specific reconstructions of meteorological events and will improve the knowledge of the French local climate over the last 140 years.

The reconstructed meteorological fields derived from the modelling chain 20CR-SANDHY-SUB will now serve as forcings for hydrological models to obtain daily probabilistic river flow reconstructed series since 1871 for a large number of near-natural catchments in France. Such a hydrometeorological reconstruction will benefit from the inter-variable consistency inherently provided by the analogue approach and from the improvement brought by the compromises across temperature and precipitation reconstruction sought after in the present work. It will first be used to study historical drought and low-flow events that strongly depend on both temperature – through evapotranspiration processes – and precipitation.

Appendix A: Outline of the different downscaling methods

The four analogy levels describing the SANDHY method in Sect. 3.1 are illustrated in Fig. A1. To reconstruct precipitation over a climatically homogeneous zone for a specific target date, a selection of 25 analogues is provided through analogue subselections. For all these steps, an atmospheric reanalysis provides the corresponding target date predictor and the meteorological archive predictor where the analogue situations are searched for. The only step requiring an analogy domain definition – where the analogy is looked for – is the one involving geopotential height predictors. The local optimization process providing this definition is described in detail in Radanovics et al. (2013).

The overall process described in Sect. 3 is illustrated in Fig. A2. The "Simulation with SANDHY" process described in Fig. A1 is applied five times using five different geopotential height domain definitions and 20CR as inputs. The First domain selection, Calendar selection and Stepwise selection are then derived using the output of the five simulations.

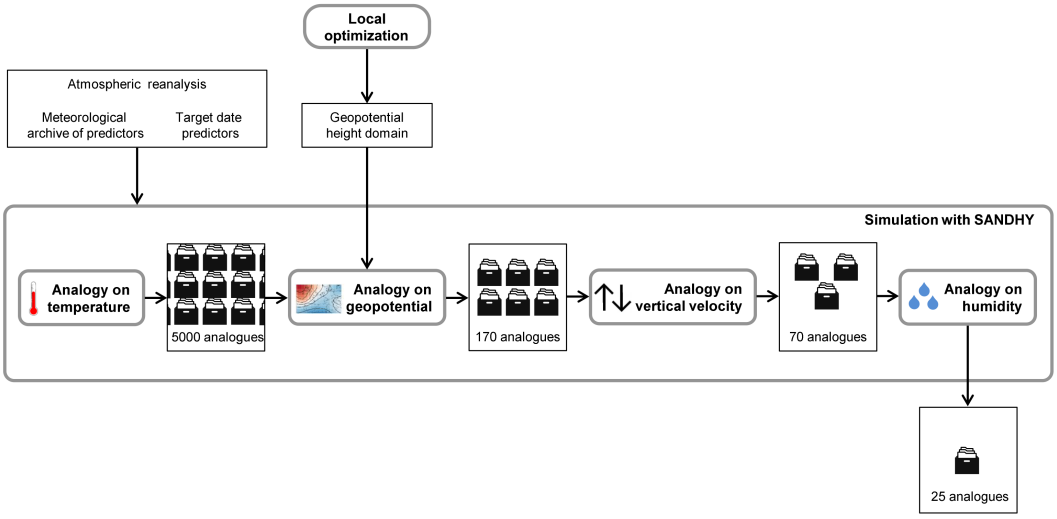

Figure A1. Synthetic diagram showing the sequence of analogy steps in the SANDHY method to reconstruct precipitation and temperature over a climatically homogeneous zone for a specific target date.

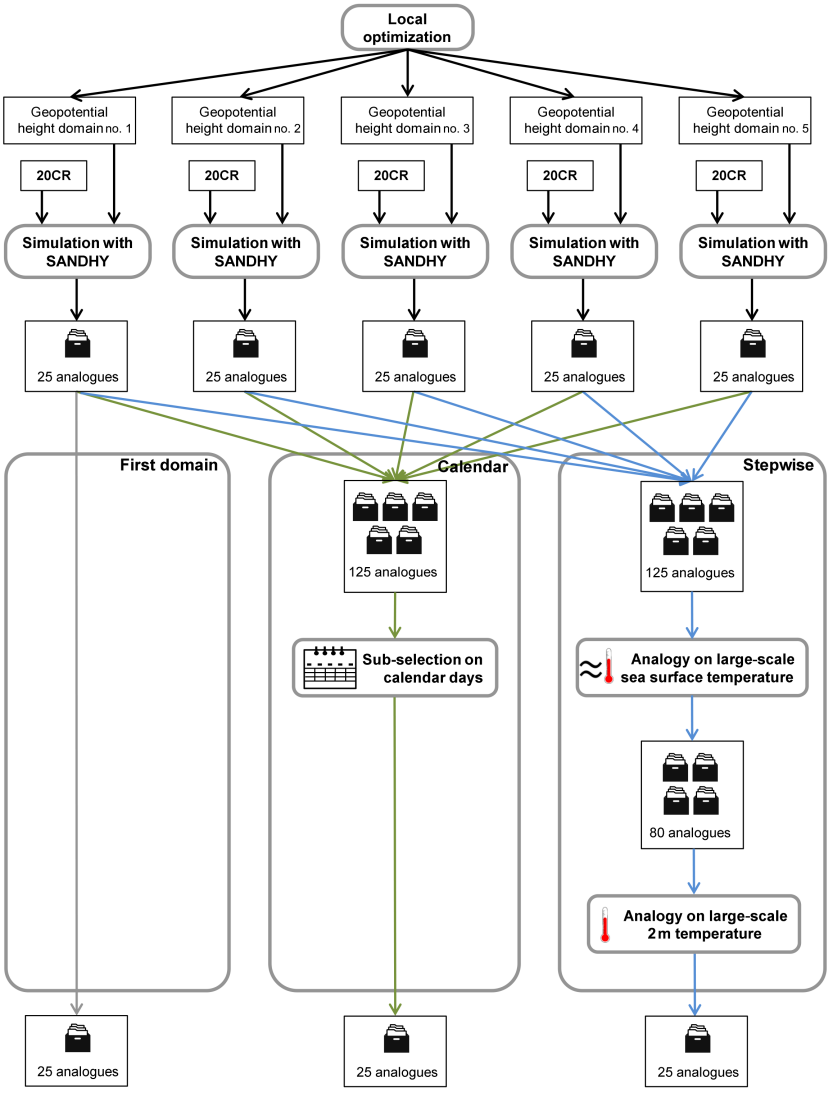

Figure A2. Synthetic diagram showing the sequence of steps for reconstructing precipitation and temperature over a climatically homogeneous zone for a specific target date using the three subselection methods described in the main text.

Appendix B: Choice of the SST predictor spatial domain for the stepwise selection

The stepwise subselection method as described in Sect. 3.3.3 includes an analogy step based on SST similarity. As mentioned above, for sake of parsimony, a single SST grid point is considered for computing the similarity of analogue dates. The most informative point for simulating precipitation has therefore to be selected. Absolute values of seasonal rank correlations between different SST grid points from the NOAA ERSST (Smith et al., 2008) and precipitation from all Safran climatically homogeneous zones (see Fig. 2) have been computed over the 1959–2007 period. As shown in Fig. B1, the median correlation is highest in autumn. The spatial pattern of correlations is moreover quite different from one season to another, making the choice of the most informative grid point difficult.

Two grid points close to France – in the Atlantic and the Mediterranean, respectively, see Fig. B1 – and with reasonably high correlations for all seasons have been selected for further investigation. A subselection of 25 analogue dates from the 125 analogue dates provided by SANDHY and based on SST similarity on each of the two points have been performed over the 1959–2007 period. Figure B2 displays the daily RMSE skill score in both precipitation and temperature for this subselection with respect to the first domain selection. It compares the subselection using the two points for each climatically homogeneous zone. It shows that SST from both grid points are informative additional predictors by displaying skill scores that are mainly positive over France. Moreover, the Atlantic and the Mediterranean grid points show similar skill scores. The final choice towards the Atlantic grid point (4° W, 46° N) was finally made following conclusions from van der Ent and Savenije (2013) who demonstrated that the Atlantic Ocean has a larger influence area over France.

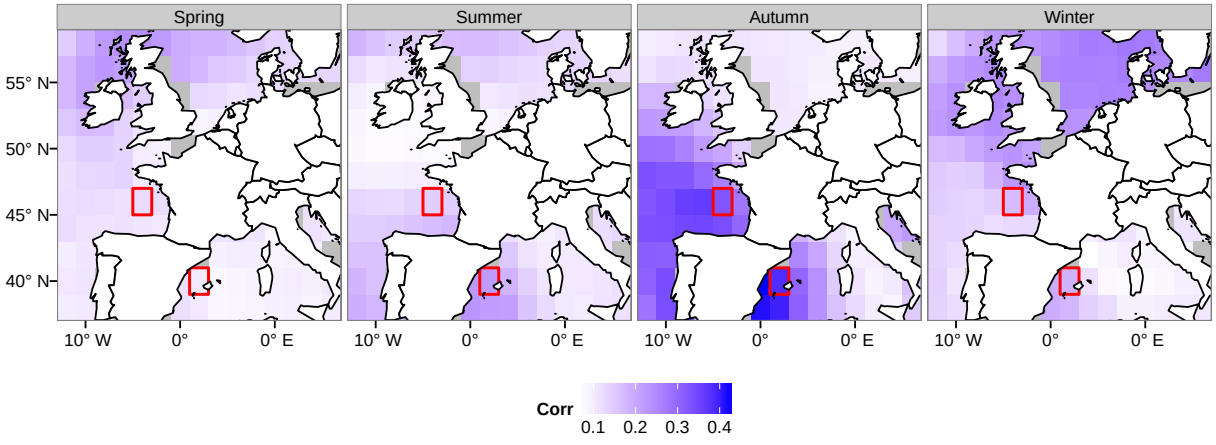

Figure B1. Median of the absolute values of the rank correlations between SST and precipitation from the 608 climatically homogeneous zones over the 1959–2007 period. Selected grid cells with high correlations are highlighted in red.

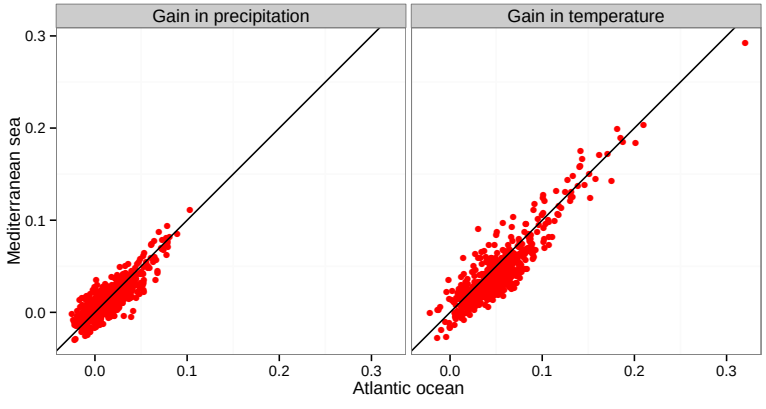

Figure B2. RMSE skill scores in precipitation and temperature for individual climatically homogeneous zones over the 1959–2007 period, calculated for a subselection of 25 analogue dates based on SST similarity on either the Atlantic and Mediterranean grid point, with the first domain selection as benchmark reference.

Appendix C: Optimization of the number of analogues for the calendar and stepwise selection

The number of analogues to retain for the calendar and stepwise methods described in Sect. 3.3 is optimized with respect to an RMSE skill score taking the first domain subselection as reference. RMSEs are calculated on monthly time series for each of the 608 climatically homogeneous zones (see Fig. 2) in relation to corresponding Safran data over the 1959–2007 period. They are furthermore calculated for individual realizations from each subselection method, and only the median is considered in the skill score. The skill score thus gives a map of the improvement of a subselection method with $N2$ (resp. $N1$–$N2$) analogues for the calendar (resp. stepwise) selection compared to the first domain selection with 25 analogues.

A detailed sensitivity analysis was performed with $N2$ varying from 125 – the number of analogues given by SANDHY, corresponding to no subselection – to 5. For the stepwise selection, $N1$ was considered to vary from 125 to $N2$. The aim here is then to find out what $N2$ (calendar selection) or $N1$–$N2$ (stepwise selection) leads to the highest improvement on average over France. For the calendar selection, each $N2$ was associated with a median rank for the skill score over the 608 climatically homogeneous zones. A standard deviation in rank was also computed to get an idea of the spatial variation in improvement over France. Similarly, for the stepwise selection, each $N1$–$N2$ was associated with a median rank and a standard deviation in rank for the skill score over France. Finally, skill scores and ranks are computed in parallel for precipitation and temperature.

Figure C1 summarizes the results, with dark grey rectangles highlighting $N2$ or $N1$–$N2$ with best ranks for both precipitation and temperature, i.e. where the best compromise for two variables is to be found. For the calendar selection, good results are obtained with $N2$ between 15 and 40, and for the stepwise selection, $N1$ should be between 40 and 80, and $N2$ between 10 and 25. As no specific optimal combination clearly arises within these domains, the choice was made to select $N2 = 25$ for the calendar selection and $N1 = 80$ and $N2 = 25$ for the stepwise selection. This choice has the advantage of generating an equal number of final analogue dates across the 3 subselection methods and the original single-domain SANDHY method, thus preventing potential biases in probabilistic scores like the CRPS (see Eq. 1 and Fig. 10).

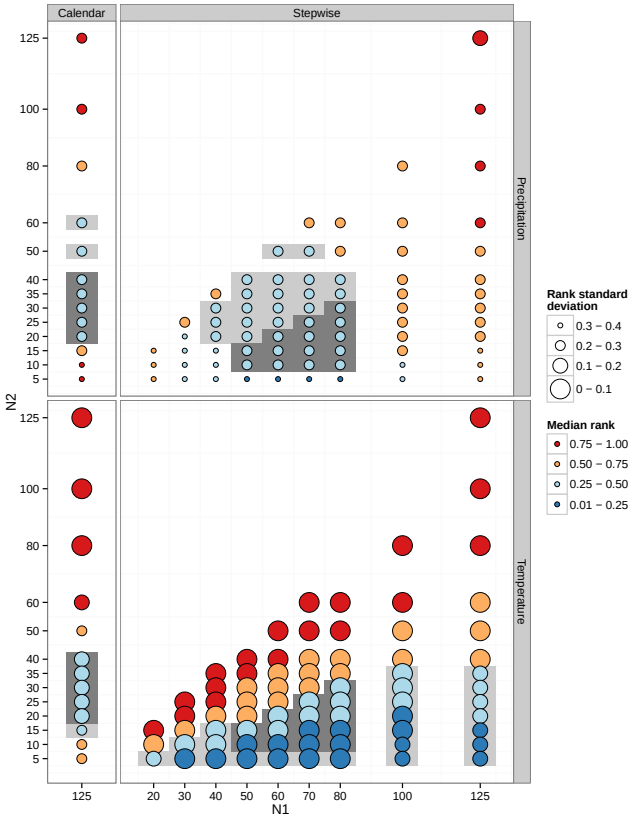

Figure C1. Normalized median ranks over France of RMSE skill scores and normalized standard deviations of these ranks for each $N2$ (calendar selection, left) and each $N1$–$N2$ (stepwise selection, right), and for precipitation (top) and temperature (bottom). Light grey rectangles highlight ranks lower than 0.5 independently for precipitation and temperature. Dark grey rectangles highlight ranks lower than 0.5 for both precipitation and temperature.

Acknowledgements. 20th Century Reanalysis V2 data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at http://www.esrl.noaa.gov/psd/. Support for the Twentieth Century Reanalysis Project data set is provided by the US Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, and Office of Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office. The authors would also like to thank Météo-France for providing access to the Safran database, the homogenized series as well as raw surface observations. The authors are grateful to Martin Boudou for providing the December 1909 precipitation map adapted from Angot (1911). L. Caillouet PhD thesis is funded by Irstea and CNR.

Edited by: S. Bronnimann

References

Allard, C.: Les récentes inondations à Besançon, Millot et Cie, Besançon, France, 1910.

Angot, A.: Annales du bureau central météorologique de France – Année 1879, vol. 2., in: Bulletin des Observations Françaises et Revue Climatologique, Chap. Revue Climatologique Mensuelle Pour la France et les Contrées Voisines Pendant l'Année 1879, Gauthiers-Villars, Paris, France, 1–50, 1881.

Angot, A. (Ed.): Annales du Bureau Central Météorologique de France – Année 1909, vol. III. Pluies, Chap. Planches, Paris, p. 148, 1911.

Auffray, A., Clavel, A., Jourdain, S., Ben Daoud, A., Sauquet, E., Lang, M., Obled, C., Panthou, G., Gautheron, A., Gottardi, F., and Garçon, R.: Reconstructing the hydrometeorological scenario of the 1859 flood of the Isere river, Houille Blanche, 44–50, doi:10.1051/lhb/2011005, 2011.

Ben Daoud, A., Sauquet, E., Lang, M., Bontron, G., and Obled, C.: Precipitation forecasting through an analog sorting technique: a comparative study, Adv. Geosci., 29, 103–107, doi:10.5194/adgeo-29-103-2011, 2011.

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, Atmos. Res., 169, 147–159, doi:10.1016/j.atmosres.2015.09.015, 2016.

Bichet, A., Folini, D., Wild, M., and Schär, C.: Enhanced Central European summer precipitation in the late 19th century: a link to the Tropics, Q. J. Roy. Meteor. Soc., 140, 111–123, doi:10.1002/qj.2111, 2014.

Boé, J. and Habets, F.: Multi-decadal river flow variations in France, Hydrol. Earth Syst. Sci., 18, 691–708, doi:10.5194/hess-18-691-2014, 2014.

Boé, J. and Terray, L.: A weather-type approach to analyzing winter precipitation in France: twentieth-century trends and the role of anthropogenic forcing, J. Climate, 21, 3118–3133, doi:10.1175/2007JCLI1796.1, 2008.

Bontron, G. and Obled, C.: A probabilistic adaptation of meteorological model outputs to hydrological forecasting, Houille Blanche, 23–28, doi:10.1051/lhb:200501002, 2005.

Boudou, M., Danière, B., and Lang, M.: Assessing changes in urban flood vulnerability through mapping land use from historical information, Hydrol. Earth Syst. Sci., 20, 161–173, doi:10.5194/hess-20-161-2016, 2016.

Brochet, P.: La sécheresse 1976 en France: aspects climatologiques et conséquences, Hydrological Sciences Bulletin, 22, 393–411, doi:10.1080/02626667709491733, 1977.

Brown, T. A.: Admissible Scoring Systems for Continuous Distributions, Tech. rep., The Rand Corporation, Santa Monica, CA, 1974.

Camus, P., Menéndez, M., Méndez, F. J., Izaguirre, C., Espejo, A., Cánovas, V., Pérez, J., Rueda, A., Losada, I. J., and Medina, R.: A weather-type statistical downscaling framework for ocean wave climate, J. Geophys. Res., 119, 7389–7405, doi:10.1002/2014JC010141, 2014.

Cassou, C., Terray, L., and Phillips, A. S.: Tropical Atlantic influence on European heat waves, J. Climate, 18, 2805–2811, doi:10.1175/JCLI3506.1, 2005.

Caussinus, H. and Mestre, O.: Detection and correction of artificial shifts in climate series, J. Roy. Stat. Soc. C-App., 53, 405–425, doi:10.1111/j.1467-9876.2004.05155.x, 2004.

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, J. Climate, 27, 5056-5074, doi:10.1175/JCLI-D-13-00464.1, 2014.

Colman, A. and Davey, M.: Prediction of summer temperature, rainfall and pressure in Europe from preceding winter North Atlantic Ocean temperature, Int. J. Climatol., 19, 513–536, doi:10.1002/(SICI)1097-0088(199904)19:5<513::AID-JOC370>3.0.CO;2-D, 1999.

Compo, G. P., Whitaker, J. S., and Sardeshmukh, P. D.: Feasibility of a 100 year reanalysis using only surface pressure data, B. Am. Meteorol. Soc., 87, 175–190, doi:10.1175/BAMS-87-2-175, 2006.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: International Surface Pressure Databank (ISPDv2), 2010.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The twentieth century reanalysis project, Q. J. Roy. Meteor. Soc., 137, 1–28, doi:10.1002/qj.776, 2011.

Dayon, G., Boé, J., and Martin, E.: Transferability in the future climate of a statistical downscaling method for precipitation in France, J. Geophys. Res., 120, 1023–1043, doi:10.1002/2014JD022236, 2015.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-

K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Dieppois, B., Durand, A., Fournier, M., and Massei, N.: Links between multidecadal and interdecadal climatic oscillations in the North Atlantic and regional climate variability of northern France and England since the 17th century, J. Geophys. Res., 118, 4359–4372, doi:10.1002/jgrd.50392, 2013.

DiNapoli, S. M. and Misra, V.: Reconstructing the 20th century high-resolution climate of the southeastern United States, J. Geophys. Res., 117, D19113, doi:10.1029/2012JD018303, 2012.

Duband, D., Schoeneich, P., and Stanescu, V. A.: The example of 1921 drought in Europe (Italy, France, Rumania, Swiss...): climatology and hydrology, Houille Blanche, 18–29, doi:10.1051/lhb:200405001, 2004.

Ferguson, C. R. and Villarini, G.: An evaluation of the statistical homogeneity of the Twentieth Century Reanalysis, Clim. Dynam., 42, 2841–2866, doi:10.1007/s00382-013-1996-1, 2014.

Gibelin, A.-L., Dubuisson, B., Corre, L., Deaux, N., Jourdain, S., Laval, L., Piquemal, J.-M., Mestre, O., Dennetière, D., Desmidt, S., and Tamburini, A.: Évolution de la température en France depuis les années 1950: constitution d'un nouveau jeu de séries homogénéisées de référence, La Météorologie, 87, 45–53, doi:10.4267/2042/54336, 2014.

Gimeno, L., Drumond, A., Nieto, R., Trigo, R. M., and Stohl, A.: On the origin of continental precipitation, Geophys. Res. Lett., 37, L13804, doi:10.1029/2010GL043712, 2010.

Giuntoli, I., Renard, B., Vidal, J.-P., and Bard, A.: Low flows in France and their relationship to large-scale climate indices, J. Hydrol., 482, 105–118, doi:10.1016/j.jhydrol.2012.12.038, 2013.

Greatbatch, R. J., Gollan, G., Jung, T., and Kunz, T.: Tropical origin of the severe European winter of 1962/1963, Q. J. Roy. Meteor. Soc., 141, 153–165, doi:10.1002/qj.2346, 2015.

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, Weather Forecast., 15, 559–570, doi:10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2, 2000.

Hoerling, M. and Kumar, A.: The perfect ocean for drought, Science, 299, 691–694, doi:10.1126/science.1079053, 2003.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, B. Am. Meteorol. Soc., 77, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Kington, J.: The weather of the 1780s over Europe, Cambridge University Press, Cambridge, UK, 1988.

Kirchner-Bossi, N., Prieto, L., García-Herrera, R., Carro-Calvo, L., and Salcedo-Sanz, S.: Multi-decadal variability in a centennial reconstruction of daily wind, Appl. Energ., 105, 30–46, doi:10.1016/j.apenergy.2012.11.072, 2013.

Krueger, O., Schenk, F., Feser, F., and Weisse, R.: Inconsistencies between long-term trends in storminess derived from the 20CR reanalysis and observations, J. Climate, 26, 868–874, doi:10.1175/JCLI-D-12-00309.1, 2013.

Kuentz, A., Mathevet, T., Gailhard, J., Perret, C., and Andréassian, V.: Over 100 years of climatic and hydrologic variability of a mediterranean and mountainous watershed: the Durance River, in: Cold and Mountain Region Hydrological Systems Under Climate Change: Towards Improved Projections, edited by: Gelfan, A., Yang, D., Gusev, Y., and Kunstmann, H., vol. 360 of IAHS Red Books, IAHS, 19–25, 2013.

Kuentz, A., Mathevet, T., Gailhard, J., and Hingray, B.: Building long-term and high spatio-temporal resolution precipitation and air temperature reanalyses by mixing local observations and global atmospheric reanalyses: the ANATEM model, Hydrol. Earth Syst. Sci., 19, 2717–2736, doi:10.5194/hess-19-2717-2015, 2015.

Lang, M., Coeur, D., Bard, A., Bacq, B., Becker, T., Bignon, E., Blanchard, R., Bruckmann, L., Delserieys, M., Edelblutte, C., and Merle, C.: Significant French flood events from the 2011 preliminary flood risk assessment of the European Flood Directive, Houille Blanche, 37–47, doi:10.1051/lhb/2013041, 2013.

Lee, D. E. and Biasutti, M.: Climatology and variability of precipitation in the Twentieth-Century Reanalysis, J. Climate, 27, 5964–5981, doi:10.1175/JCLI-D-13-00630.1, 2014.

Lorenz, E. N.: Atmospheric predictability as revealed by naturally occurring analogues, J. Atmos. Sci., 26, 636–646, doi:10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2, 1969.

Lüdecke, H.-J., Hempelmann, A., and Weiss, C. O.: Multi-periodic climate dynamics: spectral analysis of long-term instrumental and proxy temperature records, Clim. Past, 9, 447–452, doi:10.5194/cp-9-447-2013, 2013.

Marti, R. and Lepelletier, T.: L'hydrologie de la crue de 1910 et autres grandes crues du bassin de la Seine, Houille Blanche, 33–39, doi:10.1051/lhb/1997074, 1997.

Martin, B., Ansel, R., Guerrouah, O., Vitoux, M.-C., With, L., Drescher, A., Glaser, R., Himmelsbach, I., and Schönbein, J.: Geohistorical analysis of the flood of january 1910 in the Rhine Graben, Houille Blanche, 62–68, doi:10.1051/lhb/2011008, 2011.

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, J. Appl. Meteorol. Clim., 51, 505–520, doi:10.1175/JAMC-D-11-011.1, 2012.

Mascart, E.: Pluies en France; observations publiées avec la coopération du ministère des travaux publics, in: Annales du Bureau Central Météorologique de France – Année 1896, Chap. Introduction, vol. 2, 1–8, Gauthiers-Villars et Fils, Paris, France, 1898.

Matheson, J. E. and Winkler, R. L.: Scoring rules for continuous probability distributions, Manage. Sci., 22, 1087–1096, doi:10.1287/mnsc.22.10.1087, 1976.

Minvielle, M., Pagé, C., Céron, J.-P., and Besson, F.: Extension of the SIM reanalysis by combination of observations and statistical downscaling, Engineering Geology for Society and Territory, 1, 189–192, doi:10.1007/978-3-319-09300-0_36, 2015.

Misra, V., DiNapoli, S. M., and Bastola, S.: Dynamic downscaling of the twentieth-century reanalysis over the southeastern United States, Reg. Environ. Change, 13, 15–23, doi:10.1007/s10113-012-0372-8, 2013.

Moisselin, J.-M. and Jourdain, S.: Long-term data rescue and use, Houille Blanche, 33–38, doi:10.1051/lhb:2007002, 2007.

Moisselin, J.-M. and Schneider, M.: Homogenization of French rainfall long term data series, Houille Blanche, 126–130, doi:10.1051/lhb/2002096, 2002.

Moisselin, J.-M., Schneider, M., Canellas, C., and Mestre, O.: Les changements climatiques en France au XXème siecle, La Météorologie, 38, 45–56, doi:10.4267/2042/36233, 2002.

Obled, C., Bontron, G., and Garçon, R.: Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach, Atmos. Res., 63, 303–324, doi:10.1016/S0169-8095(02)00038-8, 2002.

Pardé, M.: Le régime du Rhône, Rev. Geogr. Alp., 13, 459–547, doi:10.3406/rga.1925.4941, 1925.

Plumandon, J.-R.: La sécheresse du printemps 1893, La Nature, 27–28, 1893.

Poli, P., Hersbach, H., Tan, D., Dee, D., Thépaut, J.-N., Simmons, A., Peubey, C., Laloyaux, P., Komori, T., Berrisford, P., Dragani, R., Trémolet, Y., Hólm, E., Bonavita, M., Isaksen, L., and Fisher, M.: The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C), Tech. Rep. 14, ECMWF, 2013.

Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchistéguy, L., and Morel, S.: Analysis of near-surface atmospheric variables: validation of the safran analysis over france, J. Appl. Meteorol. Clim., 47, 92–107, doi:10.1175/2007JAMC1636.1, 2008.

Quintana-Seguí, P., Habets, F., and Martin, E.: Comparison of past and future Mediterranean high and low extremes of precipitation and river flow projected using different statistical downscaling methods, Nat. Hazards Earth Syst. Sci., 11, 1411–1432, doi:10.5194/nhess-11-1411-2011, 2011.

Radanovics, S., Vidal, J.-P., Sauquet, E., Ben Daoud, A., and Bontron, G.: Optimising predictor domains for spatially coherent precipitation downscaling, Hydrol. Earth Syst. Sci., 17, 4189–4208, doi:10.5194/hess-17-4189-2013, 2013.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res., 108, 4407, doi:10.1029/2002JD002670, 2003.

Ribes, A., Azaïs, J.-M., and Planton, S.: A method for regional climate change detection using smooth temporal patterns, Clim. Dynam., 35, 391–340, doi:10.1007/s00382-009-0670-0, 2010.

Sanson, J.: La sécheresse remarquable de la période 1941–1949 en France, La Météorologie, 1–18, 1950.

Schneider, M.: Meteorological characterisation of the 1910 flood, Houille Blanche, 28–32, doi:10.1051/lhb/1997073, 1997.

Slonosky, V. C.: Wet winters, dry summers? Three centuries of precipitation data from Paris, Geophys. Res. Lett., 29, 1895, doi:10.1029/2001GL014302, 2002.

Smith, T. M. and Reynolds, R. W.: Extended reconstruction of global sea surface temperatures based on COADS data (1854–1997), J. Climate, 16, 1495–1510, doi:10.1175/1520-0442-16.10.1495, 2003.

Smith, T. M., Reynolds, R. W., Peterson, T. C., and Lawrimore, J.: Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006), J. Climate, 21, 2283–2296, doi:10.1175/2007JCLI2100.1, 2008.

Spagnoli, B., Planton, S., Déqué, M., Mestre, O., and Moisselin, J.-M.: Detecting climate change at a regional scale: the case of France, Geophys. Res. Lett., 29, 1450, doi:10.1029/2001GL014619, 2002.

Teweles, S. and Wobus, H. B.: Verification of prognosis charts, B. Am. Meteorol. Soc., 35, 455–463, 1954.

Timbal, B., Dufour, A., and McAvaney, B.: An estimate of future climate change for western France using a statistical downscaling technique, Clim. Dynam., 20, 807–823, doi:10.1007/s00382-002-0298-9, 2003.

Trigo, R. M., García-Herrera, R., Díaz, J., Trigo, I. F., and Valente, M. A.: How exceptional was the early August 2003 heatwave in France?, Geophys. Res. Lett., 32, 1–4, doi:10.1029/2005GL022410, 2005.

Turco, M., Marcos, R., Quintana-Seguí, P., and Llasat, M. C.: Testing instrumental and downscaled reanalysis time series for temperature trends in NE of Spain in the last century, Reg. Environ. Change, 14, 1811–1823, doi:10.1007/s10113-012-0363-9, 2014.

U., W.: The winter of 1879–80 in Europe, Science, 3, 485–486, doi:10.1126/science.ns-3.63.485, 1884.

Uppala, S. M., Kållberg, 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 re-analysis, Q. J. Roy. Meteor. Soc., 131, 2961–3012, doi:10.1256/qj.04.176, 2005.

van der Ent, R. J. and Savenije, H. H. G.: Oceanic sources of continental precipitation and the correlation with sea surface temperature, Water Resour. Res., 49, 3993–4004, doi:10.1002/wrcr.20296, 2013.

Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyroux, J.-M.: A 50-year high-resolution atmospheric reanalysis over France with the Safran system, Int. J. Climatol., 30, 1627–1644, doi:10.1002/joc.2003, 2010a.

Vidal, J.-P., Martin, E., Franchistéguy, L., Habets, F., Soubeyroux, J.-M., Blanchard, M., and Baillon, M.: Multilevel and multiscale drought reanalysis over France with the Safran-Isba-Modcou hydrometeorological suite, Hydrol. Earth Syst. Sci., 14, 459–478, doi:10.5194/hess-14-459-2010, 2010b.

Wang, X. L., Feng, Y., Compo, G. P., Swail, V. R., Zwiers, F. W., Allan, R. J., and Sardeshmukh, P. D.: Trends and low frequency variability of extra-tropcial cyclone activity in the ensemble of twentieth century reanalysis, Clim. Dynam., 40, 2775–2800, doi:10.1007/s00382-012-1450-9, 2013.

Wang, X. L., Feng, Y., Compo, G. P., Zwiers, F. W., Allan, R. J., Swail, V. R., and Sardeshmukh, P. D.: Is the storminess in the Twentieth Century Reanalysis really inconsistent with observations? A reply to the comment by Krueger et al. (2013b), Clim. Dynam., 42, 1113–1125, doi:10.1007/s00382-013-1828-3, 2014.

Wilby, R. L.: Downscaling summer rainfall in the UK from North Atlantic ocean temperatures, Hydrol. Earth Syst. Sci., 5, 245–257, doi:10.5194/hess-5-245-2001, 2001.

Woodruff, S. D., Diaz, H. F., Elms, J. D., and Worley, S. J.: COADS release 2 data and metadata enhancements for improvements of marine surface flux fields, Phys. Chem. Earth, 23, 517–526, doi:10.1016/S0079-1946(98)00064-0, 1998.

Yin, X., Gleason, B. E., Compo, G. P., Matsui, N., and Vose, R. S.: The International Surface Pressure Databank (ISPD) land component version 2.2, Tech. rep., National Climatic Data Center, Asheville, NC, 2008.

Yiou, P., Boichu, M., Vautard, R., Vrac, M., Jourdain, S., Garnier, E., Fluteau, F., and Menut, L.: Ensemble meteorological reconstruction using circulation analogues of 1781–1785, Clim. Past, 10, 797–809, doi:10.5194/cp-10-797-2014, 2014.