



Optimising predictor domains for spatially coherent precipitation downscaling

S. Radanovics¹, J.-P. Vidal¹, E. Sauquet¹, A. Ben Daoud², and G. Bontron²

¹Irstea, UR HHLY, 5 rue de La Doua, 69100 Villeurbanne, France

²Compagnie Nationale du Rhône (CNR), 2 rue André Bonin, 69316 Lyon Cedex 04, France

Correspondence to: S. Radanovics (sabine.radanovics@irstea.fr)

Received: 18 March 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 2 April 2013

Revised: 29 August 2013 – Accepted: 9 September 2013 – Published: 28 October 2013

Abstract. Statistical downscaling is widely used to overcome the scale gap between predictors from numerical weather prediction models or global circulation models and predictands like local precipitation, required for example for medium-term operational forecasts or climate change impact studies. The predictors are considered over a given spatial domain which is rarely optimised with respect to the target predictand location. In this study, an extended version of the growing rectangular domain algorithm is proposed to provide an ensemble of near-optimum predictor domains for a statistical downscaling method. This algorithm is applied to find five-member ensembles of near-optimum geopotential predictor domains for an analogue downscaling method for 608 individual target zones covering France. Results first show that very similar downscaling performances based on the continuous ranked probability score (CRPS) can be achieved by different predictor domains for any specific target zone, demonstrating the need for considering alternative domains in this context of high equifinality. A second result is the large diversity of optimised predictor domains over the country that questions the commonly made hypothesis of a common predictor domain for large areas. The domain centres are mainly distributed following the geographical location of the target location, but there are apparent differences between the windward and the lee side of mountain ridges. Moreover, domains for target zones located in southeastern France are centred more east and south than the ones for target locations on the same longitude. The size of the optimised domains tends to be larger in the southeastern part of the country, while domains with a very small meridional extent can be found in an east–west band around 47° N. Sensitivity experiments finally show that results are rather insensitive to the starting point

of the optimisation algorithm except for zones located in the transition area north of this east–west band. Results also appear generally robust with respect to the archive length considered for the analogue method, except for zones with high interannual variability like in the Cévennes area. This study paves the way for defining regions with homogeneous geopotential predictor domains for precipitation downscaling over France, and therefore de facto ensuring the spatial coherence required for hydrological applications.

1 Introduction

For both, climate change impact studies and operational hydrological forecasts, precipitation information on the scale of small subcatchments is needed. Numerical weather prediction (NWP) models and general circulation models (GCMs) provide relevant information about the atmospheric large-scale circulation but have too coarse a resolution to be directly used in impact models like hydrological models or for precipitation forecasts on the scale of small subcatchments. A downscaling step is therefore required, and this can be done dynamically using regional climate models and limited-area models or using statistical methods that make use of statistical relationships between large-scale predictors and local-scale predictands.

Requirements for hydrological use of predictands specifically include the spatial coherence of precipitation fields – i.e. a realistic spatial distribution of precipitation at any time step – over potentially large basins. Indeed, the generation of floods is, for example, particularly sensitive to the spatial distribution of precipitation over the catchment

considered. While dynamical downscaling methods naturally provide such a sought-after spatial coherence, this is not necessarily the case for statistical methods.

This paper proposes some development on how to ensure such a spatial coherence in precipitation by optimising the predictor domains of an analogue downscaling method for different individual target locations over France and analysing the spatial variability of results. This work will help in identifying regions with homogeneous geopotential predictor domains for precipitation, over which the spatial coherence would be de facto ensured by the selection of common analogue dates.

1.1 Statistical downscaling methods

In statistical downscaling, a relationship between large-scale predictors provided by GCMs and local-scale predictands is established. There are three major groups of statistical downscaling methods used in a climate change context: model output statistics (MOS) (e.g. Chandler, 2002; Friederichs and Hense, 2007; Vidal and Wade, 2008; Lavaysse et al., 2012), perfect prognosis (PP) (e.g. Timbal et al., 2003; Boé et al., 2006; Hertig et al., 2012) and weather generators (e.g. Vrac et al., 2007; Chen et al., 2010; Bellone et al., 2000). A review of methods and their strengths and weaknesses to produce relevant input for impact models can be found in Maraun et al. (2010). Both PP and MOS methods are also applied for operational precipitation forecast (e.g. Marty et al., 2008, 2012, 2013; Voisin et al., 2010; Nam et al., 2011; Liu and Coulibaly, 2011; Muluye, 2011).

A number of statistical downscaling studies with various methods have been performed over France over the last few years, but mainly for specific regions like the French Mediterranean (e.g. Quintana Seguí et al., 2010, 2011; Lavaysse et al., 2012; Kallache et al., 2011; Carreau and Vrac, 2011; Nuissier et al., 2011; Beaulant et al., 2011), western France (e.g. Timbal et al., 2003), the French Alps (e.g. Martin et al., 1997) or the Seine Basin (e.g. Boé et al., 2006). Until the present study, only a few of them had been performed at the country scale (Boé and Terray, 2008; Boé et al., 2009).

The downscaling method used in this work follows an analogue approach. It belongs to the PP methods, and is based on the idea introduced by Lorenz (1969) in weather forecasting that similar causes have similar effects; that is, similar predictor fields lead to similar predictand values. Nowadays numerous variants using different types of predictor fields and distance measures are in use. They range from weather-typing-based methods based on principal components of mean sea level pressure fields (Boé et al., 2006) to MOS-like techniques based on precipitation field analogues (Hamill and Whitaker, 2006; Turco et al., 2011). A description of the theory of probabilistic forecasts with analogues can be found in Hamill and Whitaker (2006). Analogue methods have been applied in different regions of the world with

very diverse climates, e.g. Switzerland (Horton et al., 2012), Australia (Timbal and McAvaney, 2001), central Sweden (Wetterhall et al., 2005), Punjab (India) (Raje and Mujumdar, 2011), southeast USA (Zhang and Georgakakos, 2012), the Alpine region (Thiemeßl et al., 2011), and northeast Spain (Ibarra-Berastegi et al., 2011).

1.2 Predictor domains: optimisation

The predictor variables used for statistical downscaling and the predictor domains have to be chosen carefully. The predictor variables should have predictive skill for the quantity to predict – in this case, precipitation. These predictors should be quantities that are reliably simulated by NWP and GCMs, and ideally they should be related to the processes leading to precipitation, and for climate change applications this relationship should persist in a changing climate (Wilby et al., 1998).

In most downscaling studies, no optimisation of the predictor domains has been performed, and only a few of them have tested even a handful of different domains (Timbal and McAvaney, 2001; Timbal et al., 2003; Gutiérrez et al., 2013). Timbal and McAvaney (2001) especially found that choosing an informative predictor domain is an important issue for the analogue selection. Ben Daoud (2010) found that some predictors like temperature or moisture variables have their main influence close to the target location, and therefore a small predictor domain close to the target location is likely to be sufficient. The predictor domains for the shape of the geopotential field are usually larger, and their optimum location depends on the meteorological situations that lead to precipitation at the target location.

Various algorithms may be used to optimise predictor domains. Ideally all predictor variables, predictor domains and other parameters should be optimised together and predictor domains of any size and shape should be possible. This was done by Sauter and Venema (2011) for an artificial neural network downscaling method and one target location in the Rhineland (Germany). Large computer resources were needed to do so because the search space is huge. A global optimisation of the analogue method was done by Horton (2012) for some stations in the Swiss Alps using genetic algorithms, with substantial computational costs as well.

This work focuses on optimising the predictor domain of one variable: the geopotential height. Restricting the parameters to be optimised allows for optimising domains for a large number of target zones separately and for exploration of the near-optimum domains for each target zone rather than searching for a unique optimum following the equifinality thesis (Beven, 2006).

1.3 Predictor domains: spatial variability

When analogue methods are applied, only one predictor domain for all target locations is generally used (see, e.g.

Timbal et al., 2003), because this ensures that the same analogue date will be selected for the whole region, and this naturally leads to spatial coherence of the precipitation field as long as individual fields and no summary measures are used. But for large target regions like France or large catchments with diverse precipitation climates like the Rhône Basin, this will likely result in a lower skill for smaller subcatchments. Bontron (2004) optimised the geopotential predictor domains for individual groups of precipitation stations located in France and northern Italy and compared the performance with those optimised for all groups together. For the groups near the barycentre of all groups, the difference in skill was small, but for groups far from the barycentre the skill was clearly better using the individually optimised predictor domains. Furthermore, he suggested to use the same predictor domain for groups situated not more than 250 km apart from each other and not separated by a major “climatological barrier”.

This work considers a large number of individual target locations over France in order to assess the spatial variability in locally optimised domains, in terms of both location and shape. The use of several near-optimum domains furthermore allows for assessment of the diversity of domains associated with very similar performance for single target locations.

1.4 Objectives and outline of the paper

The first objective of this paper is to present an extended version of the growing rectangular domain algorithm for optimising the predictor domains used by a statistical downscaling method. Such an algorithm may be used to find an ensemble of near-optimum predictor domains for any statistical downscaling method and consequently to address the issue of equifinality raised when trying to perform a numerical optimisation based on some summary skill score.

The second objective of the paper is to answer the following question: is the assumption made for example by Timbal et al. (2003) and Boé and Terray (2008) of a common predictor domain for large regions in France actually valid? To this aim, the extended version of the growing rectangular domain algorithm is applied to derive an ensemble of near-optimum geopotential predictor domains for 608 target zones covering the whole of France. The downscaling method considered for this application is an analogue method that has a long history of development with various applications, like hydrological forecasts (Ben Daoud et al., 2011b) or historical flood reconstruction (Auffray et al., 2011). However, it is applied here for the first time to the whole of France.

The methods used for downscaling and optimisation are described in Sect. 2, and results are described in Sect. 3. Section 4 presents some sensitivity tests that have been performed to check the robustness of findings with respect to (1) the archive length, (2) the version of the optimisation algorithm and (3) the optimisation starting point. Critical

methodological choices are discussed in Sect. 5, and conclusions are given in Sect. 6.

2 Data and methods

2.1 Data

2.1.1 Reanalyses

The predictor domain optimisation is done using two archives of reanalysis data. ERA-40 data (Uppala et al., 2005) at 2.5° resolution were selected as the large-scale archive against other global reanalyses because of the trade-off between archive length and data assimilation technique, following Ben Daoud et al. (2011b, a). The archive length is critical (1) for including as many diverse analogue situations as possible, (2) for studying the sensitivity on the archive length (see Sect. 4.3) and (3) for having a completely independent (i.e. not used either for optimising domains or as an archive) time period left for validation using a rigorous split-sample approach as defined by Klemesš (1986). The NCEP/NCAR reanalysis (Kalnay et al., 1996) has a longer archive, but ERA-40 made use of the more advanced three-dimensional variational data assimilation. Ben Daoud et al. (2009) compared ERA-40 and NCEP/NCAR reanalysis as sources for large-scale predictors for the downscaling method used here and found a slightly higher skill using ERA-40. ERA-Interim (Dee et al., 2011) uses an even more advanced data assimilation technique and has a higher spatial resolution, leading to a better temporal consistency and a better representation of the hydrological cycle (see Dee et al., 2011), but has still a shorter archive than ERA-40. Preliminary tests by Ben Daoud (2010) with a 1.125° version of ERA-40 and a simpler variant of the downscaling method (see Bontron and Obled, 2005) showed only very small improvements in skill and quite similar optimised domains with the higher resolution archive. Moreover, a higher resolution large-scale archive would increase both the equifinality issue and the computation time. Furthermore, some hypotheses, notably on the predictor domains for temperature, vertical velocity and humidity (see Sect. 2.2), may not be appropriate anymore. Lastly, using ERA-40 data ensures consistency with the local-scale archive as this global reanalysis has been used as a first guess by the Safran system for computing vertical profiles of near-surface variables.

Safran (French near-surface reanalysis) data (Vidal et al., 2010) are used as predictands for the local daily precipitation, which is the target variable addressed by the downscaling. The Safran reanalysis data are defined on 608 climatologically homogeneous zones covering France. Inside these zones the meteorological variables are supposed to depend only on altitude. These zones are used as elementary units in this work and are shown in Fig. 1. The algorithm used for the Safran analysis as well as its validation and application over

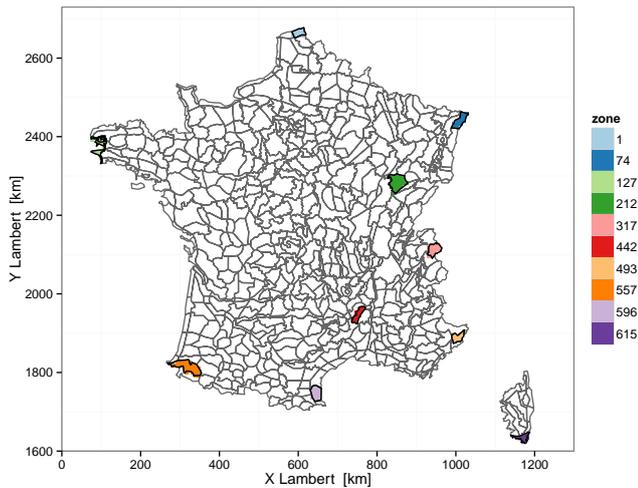


Fig. 1. The 608 climatologically homogeneous zones defined in the Safran data. The case study zones used in this study are coloured. The three zones situated in the Rhône catchment are Saône (212, dark green), Arve (317, light red) and Ardèche (442, dark red). The seven zones located at the geographical limits of the country are shown with other colours.

France are described by Quintana-Seguí et al. (2008). A detailed validation of this 50 yr atmospheric reanalysis over France has been carried out by Vidal et al. (2010). Of particular interest to this study, they found that the reanalysis uncertainty on precipitation is both very low and relatively constant over the 1958–2008 period when considering both dependent and independent validation data. The bias calculated with 83 high-quality independent validation stations is smaller than 0.1 mm day^{-1} , and the root-mean-square error is around 2.5 mm day^{-1} (Vidal et al., 2010).

The common archive period for the two reanalysis data archives is from 1 August 1958 to 31 July 2002. The period 1 August 1982–31 July 2002 is used to optimise the geopotential predictor domains except for in the sensitivity test on archive length, where the whole common archive is used. This is discussed later in Sect. 5.

2.1.2 Case study zones

The domain optimisation was performed for all 608 zones in the Safran data set, but detailed sensitivity tests focused on three case study zones. All three selected zones are part of the Rhône catchment, but have different precipitation climates as shown in Fig. 2. This has implications on the spatial coherence, since different parts of the catchment receive precipitation in different meteorological situations, and this may lead to different informative spatial predictor domains and therefore different analogue dates. Furthermore, in Sect. 3.1 results are shown for zones located at the geographical limits of the country. Maps showing the skill of the downscaling method using a unitary-sized domain at all possible locations, so-called relevance maps, for these zones were used in

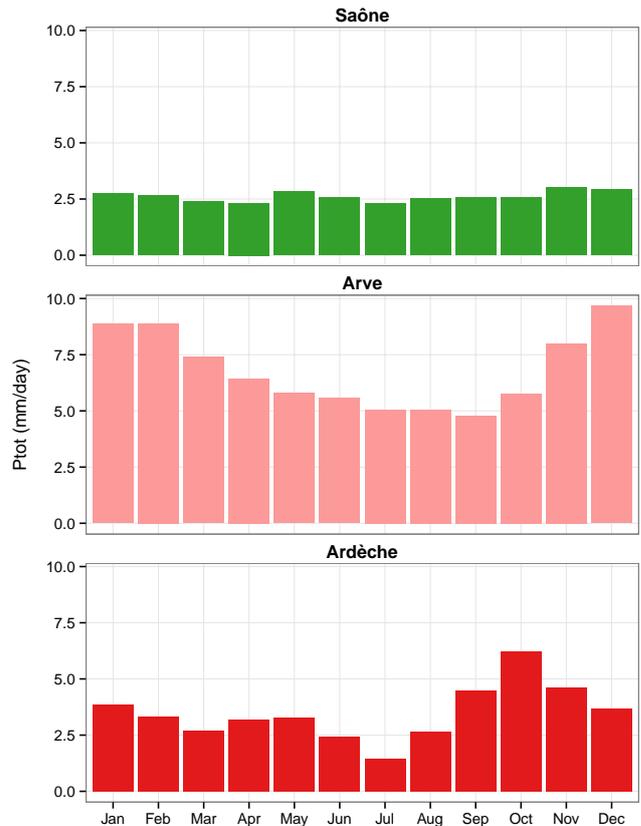


Fig. 2. Monthly mean precipitation, 1 August 1958–31 July 2002, from the Safran data for the three case study zones Saône, Arve and Ardèche.

a preliminary analysis to define the edges of the search domain. These zones at the geographical limits of the country as well as the case study zones are coloured in Fig. 1.

The case study zone called *Saône* (212) is located in the Burgundy region, in the Saône River valley. The terrain is rather flat and the zone is mainly influenced by the westerlies. The precipitation is uniformly distributed over the whole year. The second zone, named *Arve* (317), is located in the upper Arve catchment near Mont Blanc. The precipitation has a yearly cycle with a maximum in winter and a minimum in summer and early autumn. The third case study zone, named *Ardèche* (442), is located in the upper Ardèche catchment in the Cévennes area, and has a precipitation maximum in October with a high inter-annual variability (see Fig. 4 in Vidal et al., 2010). The precipitation maximum in autumn results from heavy precipitation events that are frequently observed in the Cévennes region during this season (see, e.g. Ricard et al., 2012).

2.2 Downscaling method

The downscaling method used here is an analogue approach that has already a long history of development in weather forecasting context, and some developments are underway to

Table 1. Predictors and their corresponding pressure levels, times, similarity criterion and number of situations selected in the given step used by the downscaling method.

Predictor	Pressure level [hPa] and time [UTC]	Similarity criterion	Number of analogues
Temperature	925 at 12:00 $D + 1$, 600 at 12:00 D	E.d. ¹	2000*
Geopotential	1000 at 12:00 D , 500 at 00:00 $D + 1$	TWS ²	170
Vertical velocity	850 at 06:00, 12:00, 18:00 D and 00:00 $D + 1$	E.d. ¹	70
Humidity (TCW \times Rh)	850 at 12:00 D and 00:00 $D + 1$	E.d. ¹	25

* depending on the length of the archive: $100 \times$ number of years in the archive (44 yr \rightarrow 4400, 20 yr \rightarrow 2000, ...). ¹ E.d., Euclidean distance; ² TWS, the Teweles and Wobus shape criteria (Teweles and Wobus, 1954).

apply it in a climate change context. Duband (1981) was the first who applied the analogue method in France. Guilbaud and Obled (1998) introduced the analogue selection on gridded geopotential fields with the Teweles and Wobus shape criteria (Teweles and Wobus, 1954). Obled et al. (2002) calibrated the method for 50 French, Spanish and Italian catchments. Bontron and Obled (2005) introduced the use of re-analysis data as a historical archive – the NCEP/NCAR re-analysis (Kalnay et al., 1996) – instead of interpolated radiosonde data, and added local humidity to the predictor variables. Ben Daoud et al. (2011a, b) then introduced the temperature and the vertical velocity predictor variables.

This downscaling algorithm developed by Ben Daoud et al. (2011a) and applied here performs a four-step selection on temperature, geopotential heights, vertical velocity and humidity, respectively, to identify analogue dates in the archive. The predictor variables, similarity criterion and the number of analogue situations selected after each step are summarised in Table 1. The main characteristics of the method are summarised below; for details, see Ben Daoud (2010), who identified the optimum combinations of variables and times for the Seine and Saône river basins. The number of analogue dates retained after step 2, 3 and 4 were again taken from Ben Daoud (2010). The precipitation value for day D corresponds to the precipitation accumulated between 06:00 UTC day D and 06:00 UTC day $D + 1$.

The first step is a selection on temperature at 925 hPa at 12:00 UTC day $D + 1$ and 600 hPa at 12:00 UTC day D . The pressure levels and corresponding times were optimised by Ben Daoud et al. (2011a). The predictor domain is the ERA-40 grid point closest to the target location, which is reasonable as temperature can be seen as a proxy for the thermodynamical properties of the air on the local scale. A similar choice was made by Hanssen-Bauer et al. (2003). The similarity criteria is the Euclidean distance with equal weights for the two pressure levels. As shown by Timbal et al. (2008) and Hanssen-Bauer et al. (2003), including a temperature variable as a predictor is especially important in a climate change context since different temperatures may occur in a given season and the amount of water the atmosphere can hold depends on temperature. The number of analogue

situations selected in step 1 depends on the length of the archive; it is $100 \times$ number of years in the archive – for example 2000 analogue situations for a 20 yr archive, as is used for the optimisation of the predictor domains for geopotential. This approximates the 4-month season length used by Bontron (2004) and the 2900 days Ben Daoud (2010) used with a 30 yr archive. The four days before and after the target date are excluded to avoid the selection of days within possibly the same low-pressure system.

The second step is a selection on geopotential at 1000 hPa at 12:00 UTC day D and 500 hPa at 00:00 UTC day $D + 1$. The similarity criteria used is the Teweles and Wobus criteria S1 (Teweles and Wobus, 1954), called TWS in the following, which measures the similarity between the zonal- and meridional gradients expressed as the difference between each point of the predictor domain and all other points with the same longitude or latitude. Therefore the TWS measures the similarity of the shape of the fields. Guilbaud and Obled (1998) found that the TWS leads to better downscaling performance than the Euclidean distance for the geopotential predictor. This criterion has been widely used in various analogue methods (e.g. Wetterhall et al., 2005; Wetterhall et al., 2007; Teutschbein et al., 2011; Horton et al., 2012; Brigode et al., 2012) and weather-type classification (e.g. Garavaglia et al., 2010). Again equal weights are given for the two pressure levels. The same predictor domain is used for the two pressure levels. For this step the predictor domains are optimised using the method described later in Sect. 2.5. The 170 most similar days regarding geopotential shape out of the 2000 with the most similar temperature are selected. Geopotential- or pressure fields are often used as predictors because they are well simulated by the GCMs and contain information about the atmospheric dynamics like flow strength and direction or divergence (Wilby and Wigley, 2000).

The third step is a selection on vertical velocity at 850 hPa at 06:00, 12:00 and 18:00 UTC day D and 00:00 UTC day $D + 1$. The similarity criterion is Euclidean distance, and the predictor domain is the nearest ERA-40 grid cell. Equal weights are given to the different times. Upward motion is necessary for the formation of clouds and precipitation. With a model resolution of 2.5° this predictor can only account

for large-scale upward motion due to dynamical reasons, and not for upward motion due to local convection or orography. Ben Daoud et al. (2011a) found some additional skill for the first two forecast days and fewer false alarms with the vertical velocity added as a predictor. The most similar 70 days out of the 170 days remaining after step 2 are selected.

The fourth step is a selection on humidity, more precisely the product of the total column water (TCW) and relative humidity at 850 hPa (RH) at 12:00 UTC day D and 00:00 UTC day $D + 1$. This compound variable was found to be more informative than other simple indicators by Bontron (2004). The similarity criteria is Euclidean distance and the predictor domain is the nearest ERA-40 grid cell. The most similar 25 days out of the 70 days remaining after step three are selected.

The predictor variables, their pressure levels and hours, and the number of analogues to select after each step were taken from Ben Daoud (2010), where they were selected for the Seine and Saône basins. It has also to be noted that identical combinations of variables, pressure levels and hours in steps 2 and 4 had also been selected by Bontron and Obled (2005) for application at various locations in southeastern France.

2.3 Performance criterion

The skill of the downscaling method is assessed with the continuous ranked probability score (CRPS) (Brown, 1974; Matheson and Winkler, 1976). The CRPS is widely used for the verification of probabilistic atmospheric or hydrological forecasts (see, e.g. Hagedorn et al., 2008; Demargne et al., 2010; Aspelien et al., 2011). It is defined as follows:

$$\text{CRPS} = \int_{-\infty}^{\infty} \left[F(x) - H_{x_{\text{obs}}}^0(x) \right]^2 dx, \quad (1)$$

where $F(x)$ is the forecasted cumulative distribution function of the variable x , x_{obs}^0 the observed value and $H_{x_{\text{obs}}}^0(x)$ the Heaviside function of $x - x_{\text{obs}}^0$. The properties of the CRPS are as described in Hersbach (2000). The CRPS is sensitive to the entire range of the parameter, and no predefined classes are required; it is equal to the mean absolute error (MAE) in the case of a deterministic forecast, and it can be interpreted as an integral over all possible Brier scores. In order to compare results from different zones, the continuous ranked probability skill score (CRPSS) with the climatology as a reference forecast is used:

$$\text{CRPSS} = 1 - \frac{\langle \text{CRPS} \rangle}{\langle \text{CRPS}_{\text{clim}} \rangle}, \quad (2)$$

where $\langle \rangle$ denotes the time average and the $\text{CRPS}_{\text{clim}}$ is calculated over the 1 August 1982–31 July 2002 period – except for the 44 yr experiments, where the whole archive period is considered – using precipitation data from ± 60 days around

the target day from different years in order to take seasonality into account.

2.4 Relevance maps

A relevance map represents the forecast skill for each grid cell of the predictor data set. Relevance maps were used, for example, by Bontron (2004) and Horton et al. (2012) to select the most predictive pressure level and time step for the geopotential predictor. Relevance maps are obtained by fixing every parameter except the location of a unitary-sized spatial domain (2×2 ERA-40 grid points, 2.5° resolution) that moves across the whole map (Horton et al., 2012). Using the TWS 2×2 grid points is the smallest possible domain since the TWS is based on the calculation of gradients, i.e. differences between two grid points. By iterating the position of this small domain, the CRPS score corresponding to every location is obtained. Relevance maps thus allow for one to see where the synoptic circulation information is relevant to explain observed or analysed precipitation time series. It is expected that the best predictor locations are consistent with the meteorological characteristics that are responsible for the region's weather. The predictor's best locations are therefore expected to be different for sub-catchments or stations influenced by different meteorological phenomena (Horton et al., 2012).

Relevance maps are used in this study (1) to illustrate the different atmospheric influences for zones in different parts of the country (Sect. 3.1), (2) to compare the optimised domains with the regions of high skill in the relevance maps (Sect. 3.2.2) and (3) to have an additional starting point for experiments on the sensitivity of the optimisation algorithm (Sect. 4.1).

2.5 Optimisation method

Geopotential was chosen for optimisation because it is the most important predictor in the downscaling method used, and the size and location of the predictor domain is supposed to depend more strongly on the typical weather pattern causing precipitation in the target area than for the other predictor variables. The predictor domain optimised here is a domain common to both geopotential levels described in Table 1.

The selected optimisation method is based on the idea of growing rectangular domains as applied by Bontron (2004) and Ben Daoud (2010). The basic version starts from a given 2×2 grid point domain (here the nearest one to the target zone), calculates a score (here the CRPS) and then expands the domain in four directions by adding one grid point east, west, north or south. For these four resulting domains the CRPS is calculated and the domain with the smallest CRPS is selected. This selected domain is then used as a starting domain in the next step. This is done until the score is not improved during four consecutive steps or the edge of the search domain is reached. This method is very fast, but explores

only a very small subset of the space of predictor domains, and therefore it is likely that some relevant domains are not tested.

For this work an extended version of this method was developed: five domains, instead of a single one, are retained and expanded in each step. With this method a larger number of domains are explored, and the five best domains found in this procedure are returned, thus providing an indication of variability around optimal domains. The equifinality thesis (Beven, 2006) postulates that very similar skill score values can be obtained with different parameter sets, and that it is therefore beneficial to search for a number of very good performing parameter sets rather than for the best one.

Like in the basic version, the search is started from a given 2×2 grid point domain, here the nearest one to the target zone where not indicated differently. After calculating the CRPS this domain is expanded in four directions by adding one grid point east, west, north or south, and calculating the CRPS for each of them. For the second step all four domains from the previous step are expanded. This gives 16 domains, but only 10 actually different ones, so 10 new domains are explored. From these 10 domains in the second step, the 5 best are selected to be expanded in the next step. Theoretically there are up to 20 domains (5×4) to explore from step 3 on, but there is some redundancy or some domains have already been explored in a previous step, and as such between 13 and 18 actually new ones were found. In the end the five best domains found during the whole procedure, in general stemming from different steps, are returned.

Brigode et al. (2012) optimised the predictor domains for a rainfall-based weather pattern classification and therefore tested domains of three different sizes for every possible location similar to the relevance map calculation. This is a complementary approach to the one used in this work since they assumed a domain size and shape and then tested where to centre it. In the present work, a starting point is fixed, which defines a location that has to be included in the final domain, and then domains of different size and shapes but all containing the starting point are tested. An approach similar to the one adopted by Brigode et al. (2012) was chosen by Obled et al. (2002) for a previous version of the analogue method used here. They first tested domains of six different sizes centred over the target location and then shifted the best one to find the best location. Sauter and Venema (2011) optimised the predictor domains for an artificial neural network for all predictor variables together, allowing for three-dimensional and disjointed predictor domains. In contrast to our study, they did the optimisation for only one target location, and even for this one location they stated that the computational costs were very high. It needs to be noted that the approach selected here does not allow for exploration of non-rectangular domains suggested by the examination of relevance maps (see Sect. 3.1) and that may lead to better skill values. Allowing for this type of domains would, however, involve much higher computational costs.

3 Results

In the following sections, results are shown on the different regions of influence as mapped with relevance maps in Sect. 3.1, the size and location of the optimised predictor domains (Sect. 3.2) and the downscaling skill of the method using the optimised domains during the optimisation period (Sect. 3.2.1). In Sect. 4 sensitivity experiments on the choice of the starting point of the optimisation, the choice of the basic or extended optimisation method and the archive length are shown.

The downscaling and optimisation methods are implemented in Fortran 2003 using the NetCDF Fortran90 library (Pincus and Rew, 2011) for data input and output. The subsequent analysis and figures are done using the R software environment (R Development Core Team, 2012) with packages *ncdf* (Pierce, 2011), *ggplot2* (Wickham, 2009), *reshape2* (Wickham, 2007), *RColorBrewer* (Neuwirth, 2011; Harrower and Brewer, 2003), *sp* (Pebesma and Bivand, 2005; Bivand et al., 2008), *zoo* (Zeileis and Grothendieck, 2005) and *gridExtra* (Auguie, 2012).

3.1 Different regions of influence

The relevance maps for different Safran zones located at the geographical limits of France and in the Rhône Basin (cf. Fig. 1) and calculated from the 20 yr archive are compared in Fig. 3. This figure also contains the corresponding optimised domains that will be discussed later in Sect. 3.2.2. First the magnitude of the skill differs between the relevance maps for different zones. The highest skill is found for zones that are mainly exposed to the westerlies (127, 557, 317). Furthermore, there is a clear difference in the spatial pattern between different zones. The zones in western, northern and northeastern France (001, 074, 127, 557, 317) have their region of maximum skill located west or southwest of the zone. Their regions of high skill are larger in zonal direction than in meridional direction and are cyclonically curved. They are exposed to the westerlies and receive precipitation mainly from frontal systems. A similar shape was found by Horton et al. (2012) for the Marécottes station in Switzerland, located close to the Arve zone (317). The zones in southeastern France have their region of maximum skill located south or southeast of the zone (493, 596, 442, 615). Indeed, the heavy precipitation events in this region are associated with southerly or southeasterly flow (e.g. Ricard et al., 2012). Their regions of highest skill are more north–south oriented, with high-skill regions extending westward at the southern end and eastward and northwestward at the northern end. What all relevance maps have in common is a local minimum of skill surrounded by regions with higher skill. This is due to the use of the TWS criterion that is sensitive to the gradients of the geopotential fields and their anomalies on days with precipitation. The region of low skill corresponds to the location of a minimum in the mean geopotential

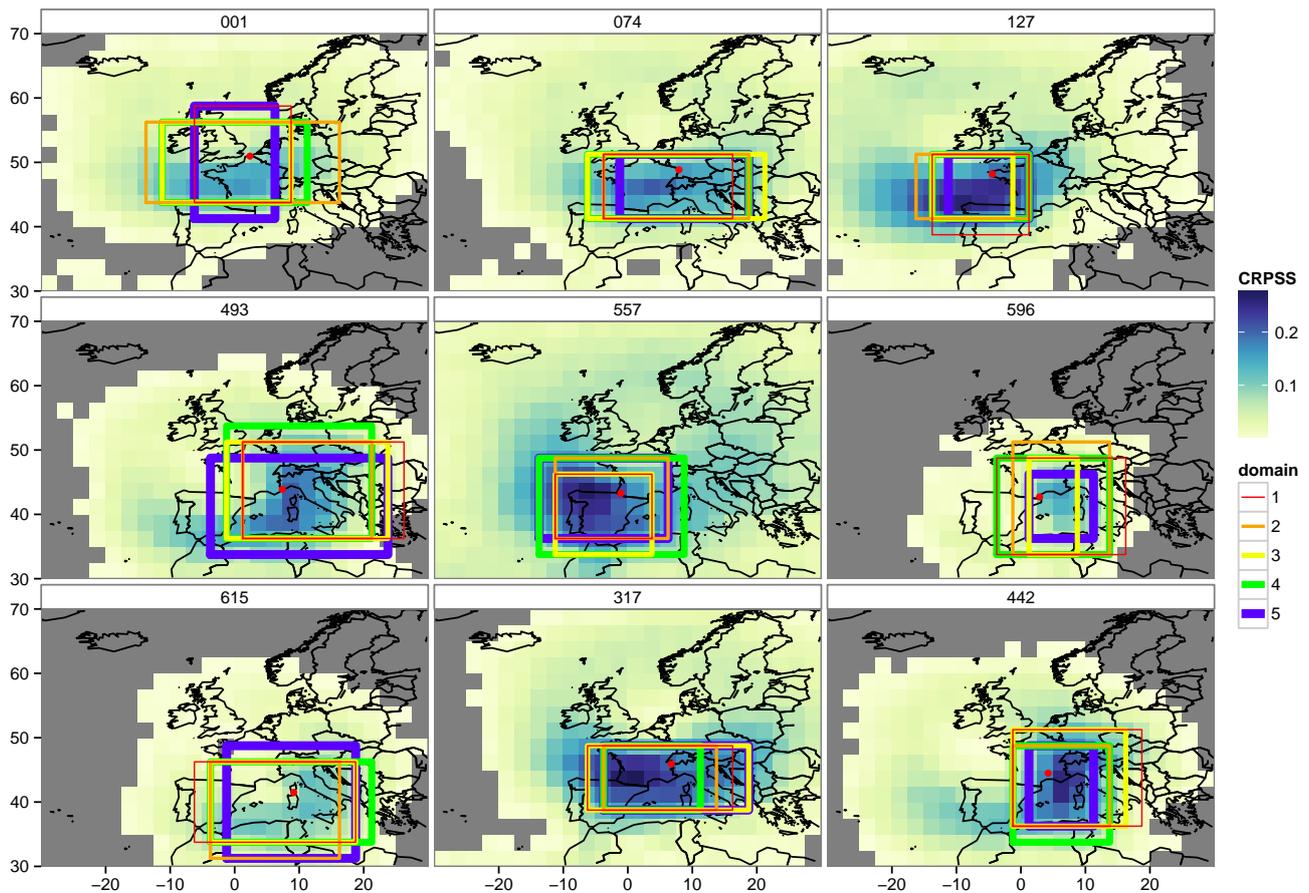


Fig. 3. Relevance maps truncated at CRPSS = 0 (areas with higher skill than the climatology) and optimised domains for nine zones identified in Fig. 1 using the 20 yr archive. The best domain found is drawn in red, followed then by those in orange, yellow, green and blue. The zone location is indicated by a red dot.

anomaly fields for rainy days (not shown). The largest gradients are situated around this minimum, which makes these regions more relevant using a similarity measure based on gradients.

Given the high seasonality with the precipitation maximum in autumn for the Ardèche case study zone (442) (cf. Fig. 2) due to specific atmospheric flow conditions, we investigated the score variations for different seasons. Seasonal relevance maps were obtained by averaging the CRPSS over different seasons instead of the whole year. In order to have enough data for each season, the score was calculated for the whole 44 yr archive. Relevance maps for the Ardèche case study zone for different seasons are shown in Fig. 4. The highest skill can be found for the winter season, followed by autumn. The location of the maximum of skill southeast to south-southeast from the Ardèche target zone corresponds well with the south-southeasterly flow found by Duffourg and Ducrocq (2011) for heavy precipitation events in the Cévennes region. In spring and summer the skill is lower due to convective precipitation, which is more difficult to predict based only on large-scale fields. This is a common feature

for the three case study zones (not shown). Interestingly the shape of the region with high skill is very similar between the seasons, which was not expected for the Ardèche zone due to the specific flow condition that leads to the autumn precipitation maximum in this zone (cf. Fig. 2). Further investigation could look at relevance maps for days with different precipitation thresholds, but this is beyond the scope of this paper.

3.2 Optimised predictor domains

In this section we will show results on the optimised domains. The downscaling skill measured with the CRPSS for all 608 zones in France is shown first. The near-optimum domains found for the case study zones are then presented, before looking at summary characteristics for all 608 zones in France.

3.2.1 Downscaling skill

Figure 5 (left panel) shows the CRPSS calculated over the 20 yr optimisation period (1 August 1982–31 July 2002) for the best domain found for each of the 608 climatologically

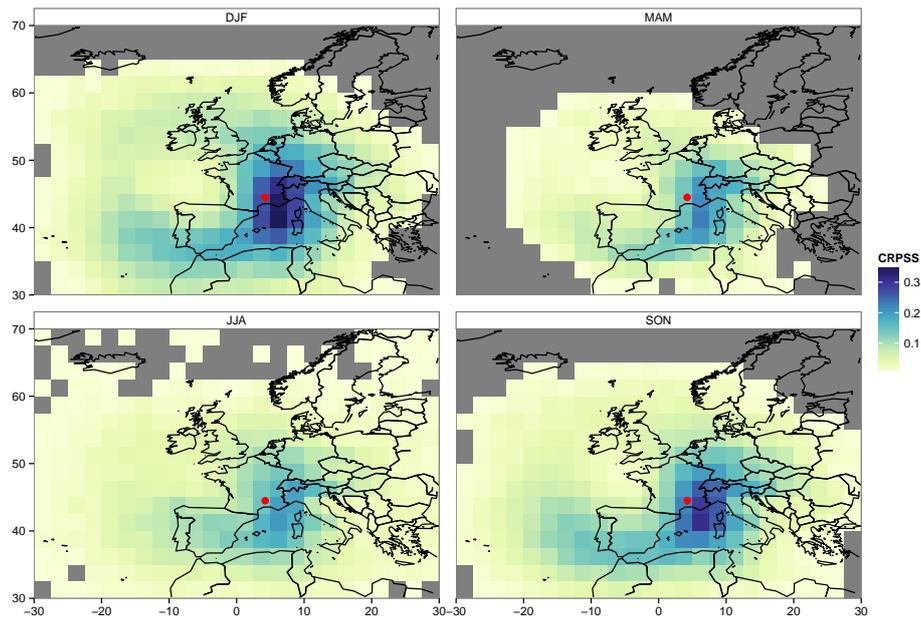


Fig. 4. Seasonal relevance maps for the Ardèche case study zone.

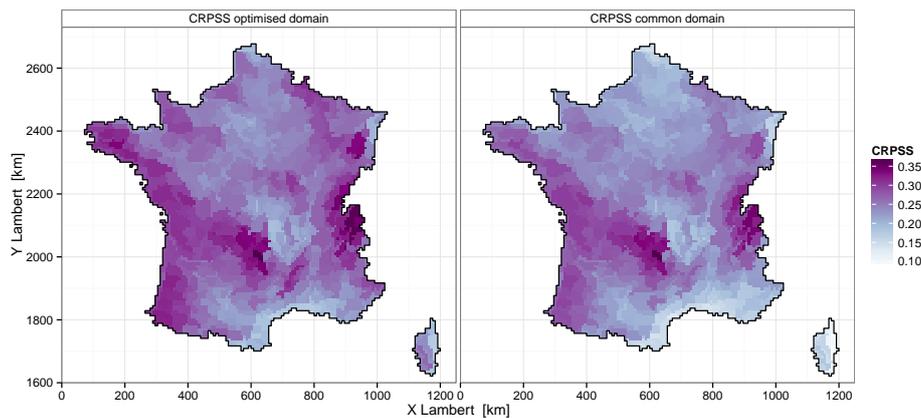


Fig. 5. Left panel: CRPSS obtained with the best domain found during optimisation. Right panel: CRPSS obtained with a domain optimised for average precipitation over France. Dark purple corresponds to a higher (that is, better) skill score, while light blue corresponds to lower skill.

homogeneous zones in France. The zones are coloured according to the CRPSS value obtained. Unsurprisingly, the CRPSS shows a spatial distribution similar to the one of the mean precipitation (Vidal et al., 2010). The more precipitation a region receives, the higher the CRPSS. The highest skill, between 0.30 and 0.35, is found on the windward side (west side in this case) of the Alps, the Massif Central and the Vosges, and along the Atlantic coast. Poorer skill, around 0.2, can be found on the lee side of mountains and around the Mediterranean coast. Quite interestingly, the difference in skill measured with the CRPSS between the best and the fifth-best domain found is never larger than 0.01. So the difference in skill between different optimised domains for the same zone are about one order of magnitude smaller

than the differences in skill between different zones, which makes all five domains equally plausible. Additionally, the skill difference does not show any apparent spatial structure.

In order to compare these CRPSS values with some reference values, a set of common geopotential predictor domains were optimised using the average precipitation time series over France. The starting point as well as the predictor domains for the other predictor variables were chosen to be close to the centroid of the country. The right-hand side of Fig. 5 shows the CRPSS obtained for each zone with the best of the common predictor domains found. The mean CRPSS over the whole country is 0.24, compared to 0.26 for the individually optimised domains. Optimising the domains locally corresponds to improvements ranging from 0.45 to 77 % for

specific locations. The largest differences can be seen close to the country borders, in southeastern France and especially on Corsica, i.e. in areas with very specific regions of influence (see Fig. 3).

3.2.2 Case study zones

As shown in Fig. 3, the optimised domains tend to include the most relevant area depicted by the relevance maps for the case study zones. They differ reasonably between different locations and inside the ensemble for a given zone. For the majority of the zones (074, 127, 493, 557, 615, 317) the aspect ratio of the optimised domains varies little inside the five-domain ensemble, while there are larger variations of this property for the zones 001, 596 and 442. These large differences in aspect ratio do not lead to larger differences in skill as mentioned above. This exemplifies the equifinality issue mentioned in Sect. 2.5.

3.2.3 Domain characteristics at the scale of France

Figure 6a shows the mean location of the centre of the optimised domains for each of the 608 zones using a 2-D colour scheme for bivariate maps introduced by Teuling et al. (2011). Here the two variables that are combined are the longitude and the latitude of the domain centre. Thus the colours correspond to the mean location of the domain centres of the five best domains for each target zone. The domain centres for the best domains are mainly distributed following the location of the target zone, but in general the mean domain centre is situated south of the target zone. Nevertheless there are some deviations from this general pattern. For zones on the east side of the Massif Central, the centres of the optimised domains are located clearly more east than the ones for zones on the west side of the massif. The same feature can be seen at other mountain ridges, for example the Vosges mountain range. Furthermore the domain centres for the zones in southeastern France are located more east than north of this area. In some regions such as, for example, Champagne in the northeast of the country (approximate Lambert coordinates $X = 700$, $Y = 2400$), we can see that many zones have their average optimised domain centre at approximately the same location. In contrast, for the Cévennes and the southern Alps regions, the average domain centres differ more often between neighbouring zones.

Figure 6b shows the maximum difference in domain centre location between two domains in the five best domain ensemble in degrees longitude and latitude. For the majority of the zones the domain centre location is a very stable property (green colour), especially in latitude direction, where differences of more than 2° are rare. So, in general, the centre points of the five near-optimum domains for a zone are close to each other. Zones with larger differences, up to 8° in longitude, are located in the southeastern part of the country at the slopes of the Alps and the Massif Central.

Figure 7a shows the mean size in degrees longitude and latitude of the optimised domains for each zone. Again the 2-D colour scheme is used, with the mean domain length in zonal direction and the mean domain length in meridional direction being the two variables. Small optimised domains (green) can be found in Brittany (150, 2400), Champagne (750, 2500), Lorraine (850, 2450), Poitou–Charentes (300, 2200) and in some parts of Normandy (300, 2500). Optimised domains with small extent in longitude direction but somewhat larger extent in latitude direction (blue) can be found along the Mediterranean coast and in the northernmost part of the country. Domains with small extent in the latitude direction and larger extent in the longitude direction (yellow) form an east–west-oriented band in the middle of the country (around 2250 km Y Lambert). Medium-sized domains (grey, brown) are found north of this band (500–900, 2350), in the southwest of the country and on the west side of Corsica (1150, 1700). The largest domains (purple, red, dark blue) tend to be situated in the southeastern part of France, except near the coast. The most prominent feature in this map is the area in the middle, where the optimised domains are very small in the meridional direction, while being reasonably stretched zonally.

The domain sizes used in other downscaling studies were compared to the domain sizes found in this study. Bontron (2004), Ben Daoud (2010), Timbal and McAvaney (2001), Boé et al. (2006) and Guilbaud and Obled (1998) used predictor domains with sizes of 20 – 25° longitude and 10 – 15° latitude, which corresponds to upper-medium-sized domains found in this study. Timbal and McAvaney (2001) (for daily minimum and maximum temperatures) tested somewhat smaller and much larger domains as well, but found the one of $20 \times 12^\circ$ to perform best. The domains tested by Brigode et al. (2012) correspond to small- to medium-sized ones found in this study. Timbal et al. (2003) (for daily minimum and maximum temperatures) used a domain somewhat larger in north–south direction. Larger domains were used by Boé and Terray (2008), Hanssen-Bauer et al. (2003), Matulla et al. (2008) and Obled et al. (2002).

Figure 7b shows the ratio of domain size range in the five-domain ensemble, defined as follows:

$$\text{ratio} = \frac{\max(X) - \min(X)}{\text{mean}(X)}, \quad (3)$$

where X is the extent of the domains in degrees longitude or latitude. A size ratio of 0 means that all five domains have equal extent. A size ratio of 1 means that the difference in extent between the largest and the smallest domain is equal to the mean extent. On average the size ratio is larger in the longitude direction than in the latitude direction. The figure is quite patchy, with individual zones showing large ratios in one or both dimensions. In the north of the country and along the Mediterranean coast, these individual zones have large ratios in longitude or both dimensions. The zones 001 and 596 shown in Fig. 3 are examples of such zones. The zones

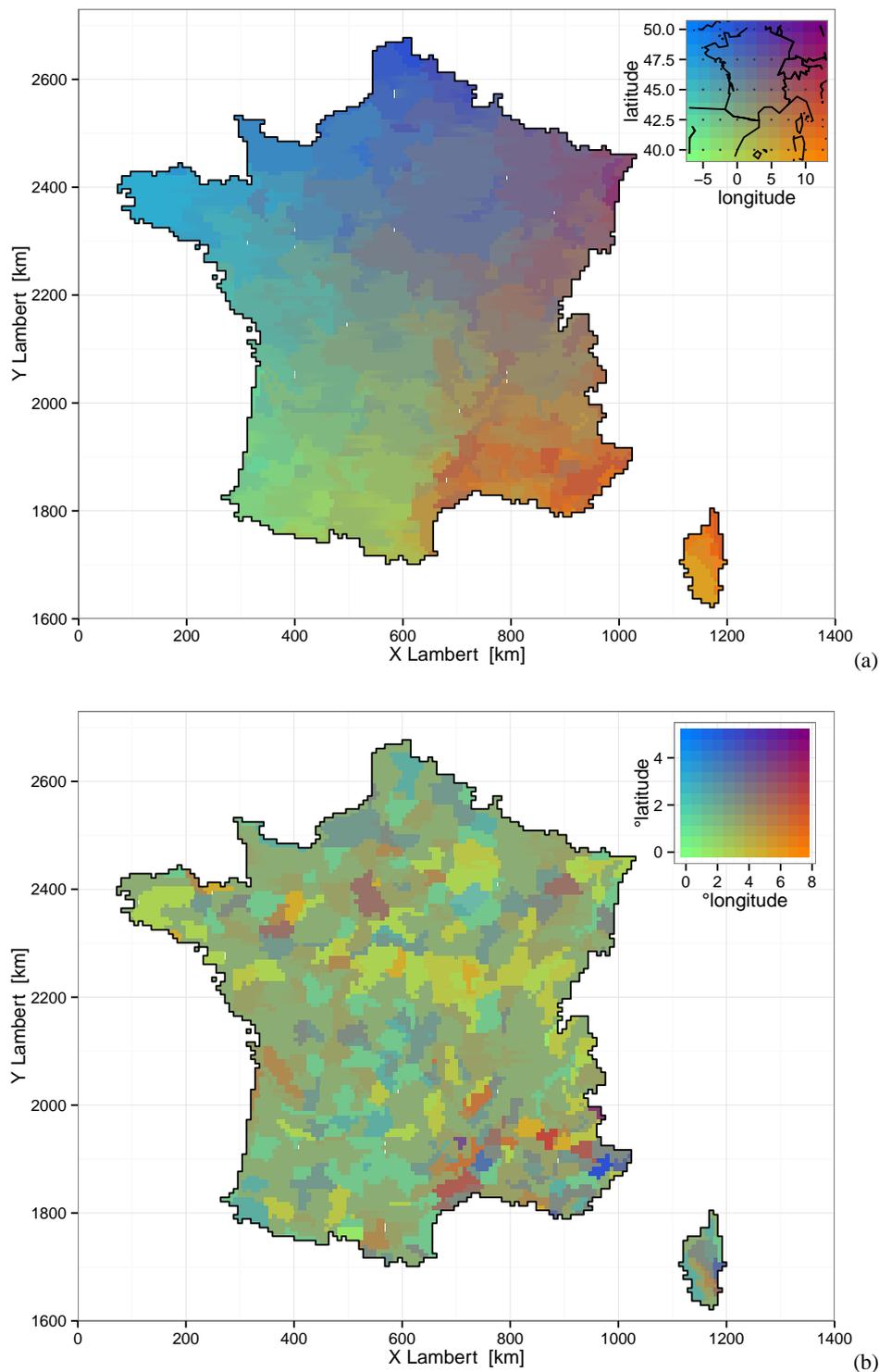


Fig. 6. (a) Mean domain centre of the five best domains found during the domain optimisation. The colours correspond to the mean location of the domain centres of the five best domains for each target zone. (b) Maximum difference in domain centre location between two domains in the five-best-domain ensemble. The colours correspond to the maximum distance of centre points in degrees longitude and latitude. Green, the centres are very close to each other; purple, the centres are very far from each other; blue, the centres are close in longitude but comparatively far in latitude direction; and orange, far in longitude and close in latitude direction.

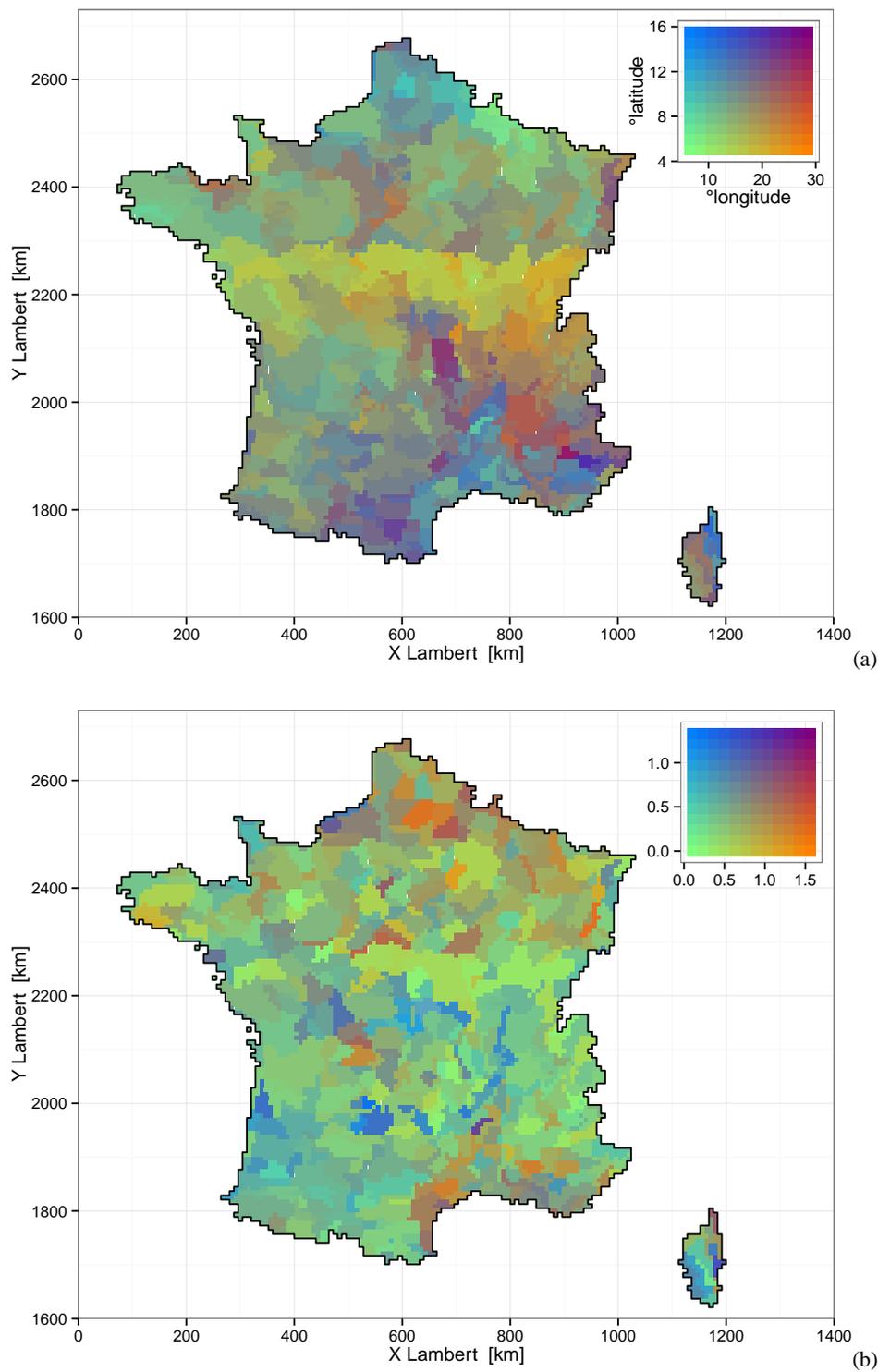


Fig. 7. (a) Mean domain size of the five best domains found during the optimisation procedure for each zone. The size is given in degrees longitude and latitude. Green, small domains; purple, large domains; blue, small in longitude direction and comparatively large in latitude direction; orange, large in longitude direction and small in latitude direction. (b) Ratio of domain size range in the five-domain ensemble (range/mean size). 0, no difference in domain extent; 1, the difference in extent between the largest and the smallest equals the medium extent.

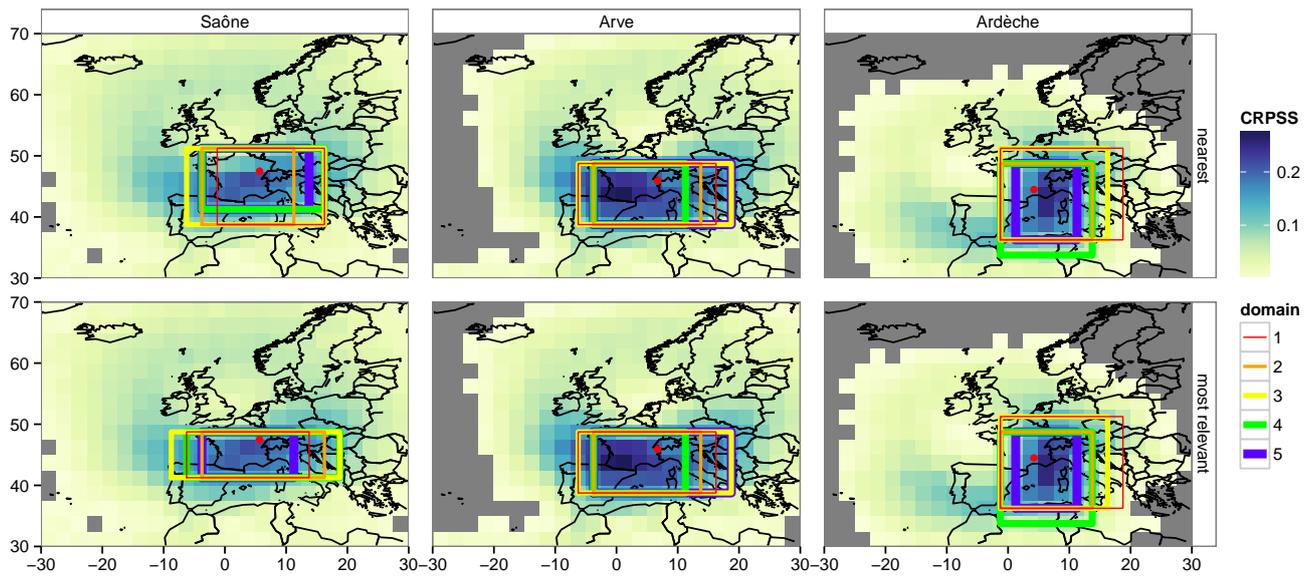


Fig. 8. Optimised predictor domains for three case study zones and different starting domains for the optimisation. First row, start at the nearest elementary domain; second row, start from the most relevant elementary domain from the relevance map. The relevance maps for each zone are shown with a colour scale underneath the predictor domains.

with the largest domain size range in the latitude direction are situated in the southern half of the country, except near the Mediterranean coast.

4 Sensitivity experiments

For the optimisation study, options were selected concerning the choice of the algorithm, the starting point and the archive length. In this section we take a detailed look at the impact of these choices on the optimised domains for the three case study zones by comparing with results for alternative choices.

4.1 Starting domain for optimisation

The growing rectangular domain algorithm requires the definition of a starting domain for the optimisation. This is true for other algorithms as well, but since the growing rectangular domain algorithm only adds grid cell rows or columns in each step and never subtracts any, the starting domain will automatically be included in the final domain (see Sect. 2.5). Therefore the choice of the starting domain can influence the predictor domains found and a poorly chosen starting point may lead to less skillful predictor domains.

One reasonable assumption is that the best predictor domain will comprise the large-scale grid cell closest to the target location, as is done here or in Obled et al. (2002). Another possibility is to start at the most relevant elementary domain, as obtained through a relevance map as done by Bontron (2004) and Ben Daoud (2010), to make sure that the most relevant location is included in the final predictor domain.

The drawback of the second approach is that the computational costs for the relevance maps are high if performed for over 600 target locations. Roughly 2.8 million CRPS calculations per zone are, for example, needed for a $40^\circ \times 60^\circ$ sized relevance map with a 20 yr archive. Therefore the relevance maps were computed only for the case study zones, and for these zones the optimised domains obtained with the two different starting domains are compared.

The first line of Fig. 8 shows the five best domains found with the optimisation procedure starting at the nearest elementary domain, with a 20 yr archive. In the second line the same procedure is used, but the optimisation was started from the most relevant elementary domain as found with the relevance map. Comparing them we can see for the Arve zone and the Ardèche zone that exactly the same five domains are found even if the two starting domains are different. For the Saône zone, five different domains are found, with lower meridional extent and systematically higher zonal extent when starting from the most relevant elementary domain. The domains found starting from the most relevant elementary domain have higher CRPS.

4.2 Optimisation method

Results obtained with the basic growing rectangular domain algorithm and the extended one developed here are compared for the case study zones. Figure 9 shows the optimised domain found with the extended algorithm (first row) and the ones found with the basic algorithm. For the Arve zone the best domain is the same for the two algorithms. For the Saône zone and the Ardèche zone the domain found with the basic algorithm is the second-best found with the extended version.

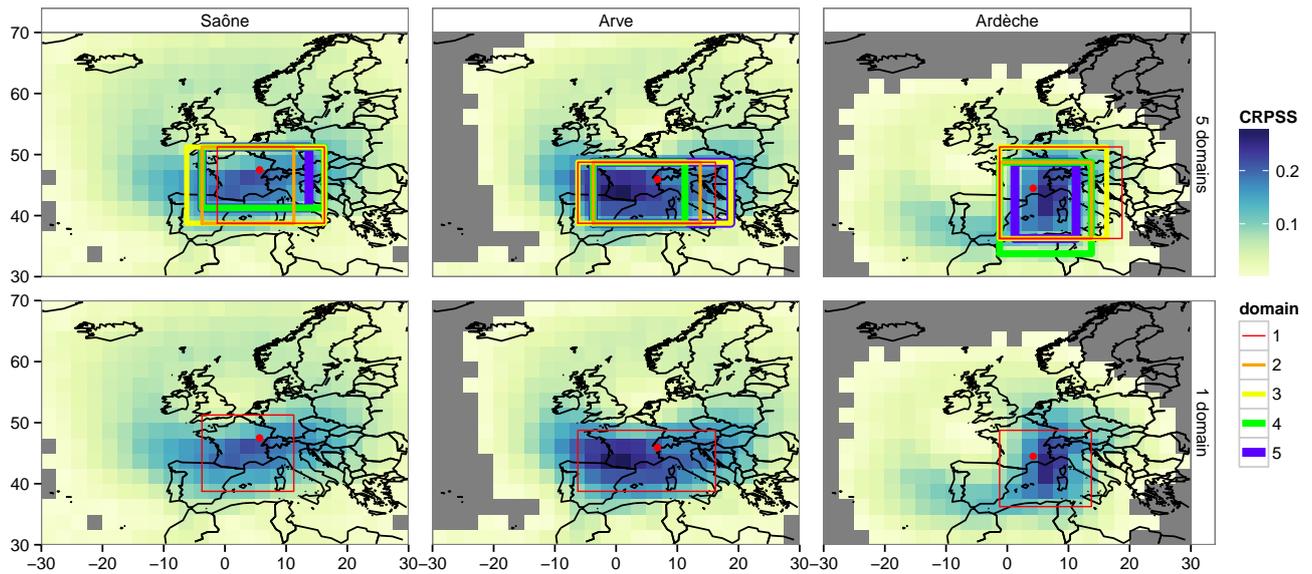


Fig. 9. Optimised predictor domains for three case study zones using the extended optimisation method with five domains (first row) or the basic optimisation method with one domain (second row). The relevance maps for each zone are shown with a colour scale underneath the predictor domains.

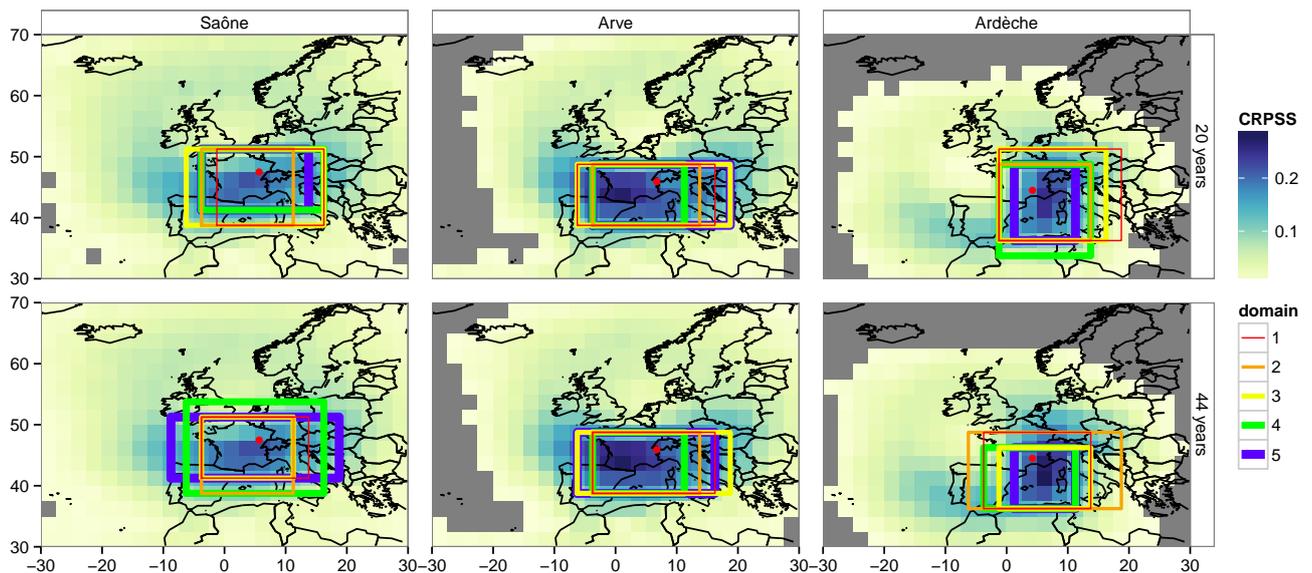


Fig. 10. Optimised predictor domains for three case study zones using 20 yr (first row) and 44 yr (second row) archives for optimisation. The relevance maps for each zone and each archive length are shown with a colour scale underneath the predictor domains.

This shows that for some zones the extended algorithm finds domains with slightly better CRPSS, together with an indication of variability between near-optimum ones.

4.3 Archive length

Figure 10 shows the optimised domains found with the 20 yr archive (1 August 1982–31 July 2002, first row) and 44 yr archive (1 August 1958–31 July 2002, second row). For the Saône zone (first column) the domains found with different

archive lengths differ, but the second-best domains found are the same, and the best domain found with the 44 yr archive is the same as the fifth best found with the 20 yr archive. For the Arve zone (second column) the differences between the best domains found with the two archive lengths are small. The best domain found with the 20 yr archive is one grid cell larger in the west than the one found with the 44 yr archive, and was found to be fifth best with the 44 yr archive. All top-five domains found have the same extent in meridional

direction. For the Ardèche zone, five completely different domains are found with the different archive lengths. This is probably related to the high inter-annual variability in the target region (see Fig. 4 in Vidal et al., 2010). A slight reduction of skill over these 20 yr (0.282 to 0.275) can be observed for the Ardèche zone when considering the domains optimised over 44 yr. Additionally, the skill computed over the 44 yr is slightly lower (0.305 to 0.310) for the domains optimised over 20 yr compared to the ones optimised over 44 yr.

Additionally, the relevance maps obtained with different archive lengths show the same structure and the same location of maximum values, but the absolute values are slightly higher with the longer archive. On relevance maps obtained with a 10 yr archive (not shown) the same overall structure is still visible, but with a decrease in CRPSS of approximately one-third.

5 Discussion

5.1 Choice of the archive period

For successful statistical downscaling it is necessary to have long data sets of predictors and predictands for building and validating the model (Timbal and McAvaney, 2001). The archive length and optimisation period chosen for statistical downscaling development depend strongly on the data that are available and the validation strategy. Bontron (2004) and Ben Daoud (2010) left only five years of their archive for validation. Ben Daoud (2010) for example excluded the years 1978, 1983, 1988, 1993 and 1998 from the 1972–2002 optimisation period. The specific years were chosen to resemble the 1972–2002 climate as closely as possible in order to validate the method for forecast purposes, i.e. in the same climate. Timbal et al. (2003) found that using more than 20 yr of the reanalysis archive does not further reduce the error in the reconstructed time series of minimum and maximum temperature as long as the more recent part of the data is used, indicating that the quality of the observation data in terms of homogeneity and the reliability of the reanalysis plays an important role too. A 20 yr recent period has been considered here for optimising the predictor domains in order to (1) leave out enough data for future validation and (2) retain a period with the highest number of observations entering the ERA-40 reanalysis system (Uppala et al., 2005). Section 4.3 above provides some preliminary analysis of the sensitivity to the archive length.

5.2 Optimisation starting point

The starting point for optimisation was chosen to be the nearest elementary domain to the target zone. The optimisation method used requires that this elementary domain is included in the final domains. As seen above, using an alternative starting point at the most relevant elementary domain instead of the nearest one results in the same domains for two of the

case study zones but in different ones for the third one, where more skillful domains, 3 % higher CRPSS, were found starting at the most relevant elementary domain. In Fig. 7a a sudden change in domain size can be seen around 47.5° N, with the domains north of this line having slightly larger domains in meridional direction. The Saône case study zone happens to be situated north of this line, and the experiment with the most relevant elementary domain as a starting point, more southwest in this case, showed that the optimised domains differ for this case study zone. The domains found starting the optimisation from the most relevant elementary domain are indeed very similar to those found for zones south of the Saône case study zone (not shown). Thus the sudden change in the domain sizes in Fig. 7a is likely to be a result of the starting point choice.

6 Conclusions

6.1 An algorithm to provide near-optimum predictor domains

An extended version of the growing rectangular domain algorithm has been described and applied for deriving ensembles of five near-optimum geopotential predictor domains for 608 individual target zones covering France. This algorithm allowed for us to find that different predictor domains may lead to very similar performances for the analogue downscaling method considered here. It exemplifies the equifinality issue in statistical downscaling that has been recognised in many other research domains (Beven, 2006). The equifinality is a consequence of a single-objective optimisation approach; that is, the use of a single-valued objective function. Indeed, the CRPSS used in this study as the objective function gives only an overall skill of the method. Consequently, for a given target location, some near-optimum domains may perform better than others – for example, for days with specific circulation patterns. This algorithm is potentially applicable in other contexts. This study has already shown that it could be applied at different target locations, but one may also think of considering another predictand, such as minimum or maximum temperature (Gutiérrez et al., 2013), or optimising the spatial domain of other predictors. As a result, this algorithm could be perfectly applied to another type of statistical downscaling method.

This first application of the downscaling procedure by Ben Daoud (2010) to the whole of France together with the use of the optimisation algorithm led here to a country-wide assessment of predictor domains. The domains resulting from an optimisation with the presented algorithm include the most relevant area depicted by the relevance maps for all three case study zones. The domains differ moderately between different locations and inside the ensemble for a given zone. In some regions, such as Brittany, we found a larger region with the same optimised domain, while

especially in the Rhône catchment we found high variability in the location and even more in the size of the optimised predictor domains. For the majority of the zones the aspect ratio of their five domains is rather similar, but for some zones, equally skillful domains with very different aspect ratios are found. The centres of the optimised domains are mainly distributed following the geographical location of the target zone but with clear differences between eastern and western slopes of mountain ridges. The domain centres for zones in southeastern France are located more east than north of this area. The domain centre location is a stable property in the five-domain ensemble except for isolated zones at the slopes of the Alps and the Massif Central. The domain sizes vary considerably between the zones with ensemble mean zonal extents between 6.5 and 28.5° and meridional extents between 5 and 15.5°.

6.2 On the assumption of a common predictor domain

This work addressed the hypothesis of a common predictor domain for different regions of France for statistical downscaling of precipitation. This assumption has been indeed made implicitly by all previous studies over France, e.g. Timbal et al. (2003) for western and southern France separately, and Boé and Terray (2008) for the whole of France. Results from the optimisation of geopotential predictor domains showed a large diversity of near-optimum domains for the set of 608 climatically homogeneous zones covering France, and therefore suggest that this assumption is questionable, at least when one seeks to obtain the most skillful method for each individual zone. However, relatively large zones have been found to share similar near-optimum predictor domains, and making this assumption within each of them may lead to limited loss of skill compared to domains optimised for individual locations. This is seemingly the case for the Seine Basin, where only minor variations in the optimised domains can be found (see Figs. 6 and 7), supporting the hypothesis made by Boé et al. (2006, 2007) for a common predictor domain over this basin.

Conversely, large river basins like the Rhône Basin include zones with very diverse influence as exemplified by the three case study zones located in the Saône, upper Arve and upper Ardèche catchments (see Fig. 3). The present work suggests that the performance of any perfect prognosis downscaling method using a common predictor domain is far from optimal for individual locations in France as a consequence of the assumption of a common predictor domain, as shown in Fig. 5 for the analogue downscaling method used here. This may be specifically the case for the method developed by Boé et al. (2006), which was later extended to the whole of France by Boé and Terray (2008), Pagé et al. (2008) and Pagé and Terray (2010). This method has been used in many subsequent national-scale climate change impact studies on hydrology (see, e.g. Boé et al., 2009; Vidal et al., 2012), and downscaled products are now disseminated through a

national climate service platform built in the DRIAS project (Lémond et al., 2011). This issue of a common predictor domain thus provides some explanation for the identified biases (Boé, 2007) and weak correlations (Boé and Terray, 2008) in downscaled precipitation outputs for regions around the Mediterranean coast. Indeed, as shown in Figs. 6 and 7, the optimum geopotential predictor domains for these regions are quite different from the rest of the country.

6.3 Towards predictand areas with homogeneous predictors

The spatial coherence of the downscaled precipitation is often taken as a given when using the analogue method, but this is only true if the same analogue dates are found for the whole target region, which is not guaranteed if different sub-target regions are using different predictor domains. On the other hand, if the target region is large, like a large river basin, a common predictor domain is likely to be suboptimal on the local scale as the best domains differ for the subcatchments, as shown in this study, for example, for the Rhône catchment.

Despite the simplicity of the concept, the analogue method has a large number of parameters: the predictor variables and their spatial and temporal domains, the similarity criteria and the number of analogues. A global optimisation of all these parameters together is desirable but involves high computational costs. In this work the optimisation was restricted to the horizontal domains of the geopotential predictor but was performed for a large number of predictand zones.

Using individual predictor domains for each zone will in general result in different analogue dates, thereby not ensuring systematically the spatial coherence it has if a common domain is used for all predictand locations. Therefore it will be beneficial to group zones together that can use the same parameters, i.e. the same geopotential predictor domain. The presented analysis will help to this end by building on the idea of grouping zones by equal domains in the five-domain ensemble, as equal optimised predictor domains reflect proximity and similar flow exposure. To this end, for each zone, one predictor domain from the five-domain ensemble has to be selected such that contiguous areas with the same predictor domain are formed. The smooth distribution of the domain centre locations together with their rather small range should facilitate this. The domain size has a higher spatial variability that could hamper the attempt of aggregating zones by same domain, but this goes along with a larger range that may compensate it to a certain degree.

Acknowledgements. The authors would like to thank Météo-France for providing access to the Safran database and ECMWF for providing access to the ERA-40 data. We would like to thank Pascal Horton, Tobias Sauter and an anonymous reviewer for their constructive comments on the discussion paper that helped to significantly improve the manuscript.

Edited by: F. Pappenberger

References

- Aspelien, T., Iversen, T., Bremnes, J. B., and Frogner, I.-L.: Short-range probabilistic forecasts from the Norwegian limited-area EPS: long-term validation and a polar low study, *Tellus A*, 63, 564–584, doi:10.1111/j.1600-0870.2010.00502.x, 2011.
- Auffray, A., Clavel, A., Jourdain, S., Ben Daoud, A., Sauquet, E., Lang, M., Obled, C., Panthou, G., Gautheron, A., Gottardi, F., and Garçon, R.: Reconstitution hydrométéorologique de la crue de l'Isère de 1859 (Reconstructing the hydrometeorological scenario of the 1859 flood of the Isère river), *La Houille Blanche*, 44–50, doi:10.1051/lhb/2011005, 2011.
- Auguie, B.: gridExtra: functions in Grid graphics, r package version 0.9.1, <http://CRAN.R-project.org/package=gridExtra>, last access: 24 October 2013, 2012.
- Beaulant, A.-L., Joly, B., Nuissier, O., Somot, S., Ducrocq, V., Joly, A., Sevault, F., Deque, M., and Ricard, D.: Statistico-dynamical downscaling for Mediterranean heavy precipitation, *Q. J. Roy. Meteorol. Soc.*, 137, 736–748, doi:10.1002/qj.796, 2011.
- Bellone, E., Hughes, J. P., and Guttorp, P.: A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts, *Clim. Res.*, 15, 1–12, 2000.
- Ben Daoud, A.: Améliorations et développements d'une méthode de prévision probabiliste des pluies par analogie, Application à la prévision hydrologique sur les grands bassins fluviaux de la Saône et de la Seine, Thèse de doctorat, Université Joseph Fourier, Grenoble, 2010.
- Ben Daoud, A., Sauquet, E., Lang, M., Obled, C., and Bontron, G.: La prévision des précipitations par recherche d'analogues: état de l'art et perspectives (Precipitation forecasting through an analog sorting technique: state of the art and further investigations), *La Houille Blanche*, 6, 60–65, doi:10.1051/lhb/2009079, 2009.
- 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, 2011a.
- Ben Daoud, A., Sauquet, E., Lang, M., and Ramos, M.-H.: Peut-on étendre l'échéance de prévision des crues en optimisant la prévision de pluies par recherche d'analogues? Application au bassin de la Seine à Paris (Can we extend flood forecasting lead-time by optimising precipitation forecasting based on analogs? Application to the Seine river basin), *La Houille Blanche*, 37–43, doi:10.1051/lhb/2011004, 2011b.
- Beven, K.: A manifesto for the equifinality thesis, *J. Hydrol.*, 320, 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006.
- Bivand, R. S., Pebesma, E. J., and Gomez-Rubio, V.: Applied spatial data analysis with R, <http://www.asdar-book.org/>, Springer, New York, NY, 2008.
- Boé, J.: Changement global et cycle hydrologique : Une étude de régionalisation sur la France, Thèse de doctorat, 3, <http://thesesups.ups-tlse.fr/227/>, Université Toulouse, 2007.
- 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.
- Boé, J., Terray, L., Habets, F., and Martin, E.: A simple statistical-dynamical downscaling scheme based on weather types and conditional resampling, *J. Geophys. Res.*, 111, D23106, doi:10.1029/2005JD006889, 2006.
- Boé, J., Terray, L., Habets, F., and Martin, E.: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies, *Int. J. Climatol.*, 27, 1643–1655, doi:10.1002/joc.1602, 2007.
- Boé, J., Terray, L., Martin, E., and Habets, F.: Projected changes in components of the hydrological cycle in French river basins during the 21st century, *Water Resour. Res.*, 45, W08426, doi:10.1029/2008WR007437, 2009.
- Bontron, G.: Prévision quantitative des précipitations: adaptation probabiliste par recherche d'analogues, Utilisation des réanalyses NCEP/NCAR et application aux précipitations du Sud-Est de la France, Thèse de doctorat, Institut National Polytechnique de Grenoble, 2004.
- Bontron, G. and Obled, C.: L'adaptation probabiliste des prévisions météorologiques pour la prévision hydrologique (A probabilistic adaptation of meteorological model outputs to hydrological forecasting), *La Houille Blanche*, 23–28, doi:10.1051/lhb:200501002, 2005.
- Brigode, P., Bernardara, P., Gailhard, J., and Ribstein, P.: Optimization of the geopotential heights informatio used in a rainfall based weather patterns classification over Austria, *Int. J. Climatol.*, 33, 1563–1573, doi:10.1002/joc.3535, 2013.
- Brown, T. A.: Admissible scoring systems for continuous distributions, The Rand Paper Series P-5235, The Rand Corporation, Santa Monica, California, 1974.
- Carreau, J. and Vrac, M.: Stochastic downscaling of precipitation with neural network conditional mixture models, *Water Resour. Res.*, 47, W10502, doi:10.1029/2010WR010128, 2011.
- Chandler, R. E.: On the use of generalized linear models for interpreting climate variability, *Environmetrics*, 16, 699–715, doi:10.1002/env.731, 2002.
- Chen, J., Brissette, F. P., and Leconte, R.: A daily stochastic weather generator for preserving low-frequency of climate variability, *J. Hydrol.*, 388, 480–490, doi:10.1016/j.jhydrol.2010.05.032, 2010.
- 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, I., 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. Meteorol. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011.
- Demargne, J., Brown, J., Liu, Y., Seo, D.-J., Wu, L., Toth, Z., and Zhu, Y.: Diagnostic verification of hydrometeorological and hydrologic ensembles, *Atmos. Sci. Lett.*, 11, 114–122, doi:10.1002/asl.261, 2010.
- Duband, D.: Prévision spatiale des hauteurs de précipitations journalières (A spatial forecast of daily precipitation heights), *La Houille Blanche*, 497–512, doi:10.1051/lhb/1981046, 1981.

- Duffourg, F. and Ducrocq, V.: Origin of the moisture feeding the Heavy Precipitating Systems over Southeastern France, *Nat. Hazards Earth Syst. Sci.*, 11, 1163–1178, doi:10.5194/nhess-11-1163-2011, 2011.
- Friederichs, P. and Hense, A.: Statistical downscaling of extreme precipitation events using censored quantile regression, *Mon. Weather Rev.*, 135, 2365–2378, doi:10.1175/MWR3403.1, 2007.
- Garavaglia, F., Gailhard, J., Paquet, E., Lang, M., Garçon, R., and Bernardara, P.: Introducing a rainfall compound distribution model based on weather patterns sub-sampling, *Hydrol. Earth Syst. Sci.*, 14, 951–964, doi:10.5194/hess-14-951-2010, 2010.
- Guilbaud, S. and Obled, C.: Prévision quantitative des précipitations journalières par une technique de recherche de journées antérieures analogues: optimisation du critère d'analogie (Daily quantitative precipitation forecast by an analogue technique: optimisation of the analogy criterion), *Comptes Rendus de l'Académie des Sciences – Series IIA, Earth Planet. Sci.*, 327, 181–188, doi:10.1016/S1251-8050(98)80006-2, 1998.
- Gutiérrez, J. M., San-Martín, D., Brands, S., Manzanas, R., and Herrera, S.: Reassessing statistical downscaling techniques for their robust application under climate change conditions, *J. Climate*, 26, 171–188, doi:10.1175/JCLI-D-11-00687.1, 2013.
- Hagedorn, R., Hamill, T. M., and Whitaker, J. S.: Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts, Part I: Two-meter temperatures, *Mon. Weather Rev.*, 136, 2608–2619, doi:10.1175/2007MWR2410.1, 2008.
- Hamill, T. M. and Whitaker, J. S.: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application, *Mon. Weather Rev.*, 134, 3209–3229, doi:10.1175/MWR3237.1, 2006.
- Hanssen-Bauer, I., Førland, E. J., Haugen, J. E., and Tveito, O. E.: Temperature and precipitation scenarios for Norway: comparison of results from dynamical and empirical downscaling, *Clim. Res.*, 25, 15–27, doi:10.3354/cr025015, 2003.
- Harrower, M. and Brewer, C. A.: ColorBrewer.org: An online tool for selecting colour schemes for maps, *Cartogr. J.*, 40, 27–37, doi:10.1179/000870403235002042, 2003.
- Hersbach, H.: Decomposition of the Continuous Ranked Probability Score for ensemble prediction systems, *Weather Forecast.*, 15, 559–570, 2000.
- Hertig, E., Paxian, A., Vogt, G., Seubert, S., Paeth, H., and Jacobeit, J.: Statistical and dynamical downscaling assessments of precipitation extremes in the Mediterranean area, *Meteorol. Z.*, 21, 61–77, doi:10.1127/0941-2948/2012/0271, 2012.
- Horton, P.: Améliorations et optimisation globale de la méthode des analogues pour la prévision statistique des précipitations, Développement d'un outil de prévision et application opérationnelle au bassin du Rhône à l'amont du Léman, Thèse de doctorat, Université de Lausanne, Switzerland, 2012.
- Horton, P., Jaboyedoff, M., Metzger, R., Obled, C., and Marty, R.: Spatial relationship between the atmospheric circulation and the precipitation measured in the western Swiss Alps by means of the analogue method, *Nat. Hazards Earth Syst. Sci.*, 12, 777–784, doi:10.5194/nhess-12-777-2012, 2012.
- Ibarra-Berastegi, G., Saéñz, J., Ezcurra, A., Elías, A., Diaz Argandoña, J., and Errasti, I.: Downscaling of surface moisture flux and precipitation in the Ebro Valley (Spain) using analogues and analogues followed by random forests and multiple linear regression, *Hydrol. Earth Syst. Sci.*, 15, 1895–1907, doi:10.5194/hess-15-1895-2011, 2011.
- Kallache, M., Vrac, M., Naveau, P., and Michelangeli, P.-A.: Non-stationary probabilistic downscaling of extreme precipitation, *J. Geophys. Res.*, 116, D05113, doi:10.1029/2010JD014892, 2011.
- Kalnay, E., Kanamitsu, M., Kistler, R. W. C., 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.
- Klemeš, V.: Operational testing of hydrological simulation models, *Hydrolog. Sci. J.*, 31, 13–24, 1986.
- Lavaysse, C., Vrac, M., Drobinski, P., Lengaigne, M., and Vischel, T.: Statistical downscaling of the French Mediterranean climate: assessment for present and projection in an anthropogenic scenario, *Nat. Hazards Earth Syst. Sci.*, 12, 651–670, doi:10.5194/nhess-12-651-2012, 2012.
- Lémond, J., Dandin, P., Planton, S., Vautard, R., Pagé, C., Déqué, M., Franchistéguy, L., Geindre, S., Kerdoncuff, M., Li, L., Moisselin, J.-M., Noël, T., and Tourre, Y. M.: DRIAS: a step toward Climate Services in France, *Adv. Sci. Res.*, 6, 179–186, doi:10.5194/asr-6-179-2011, 2011.
- Liu, X. and Coulibaly, P.: Downscaling ensemble weather predictions for improved week-2 hydrologic forecasting, *J. Hydrometeorol.*, 12, 1564–1580, doi:10.1175/2011JHM1366.1, 2011.
- 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.
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themessl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change. Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, RG3003, doi:10.1029/2009RG000314, 2010.
- Martin, E., Timbal, B., and Brun, E.: Downscaling of general circulation model outputs: simulation of the snow climatology of the French Alps and sensitivity to climate change, *Clim. Dynam.*, 13, 45–56, doi:10.1007/s003820050152, 1997.
- Marty, R., Zin, I., and Obled, C.: On adapting QPPFs to fit hydrological needs: the case of flash flood forecasting, *Atmos. Sci. Lett.*, 9, 73–79, 2008.
- Marty, R., Zin, I., Obled, C., Bontron, G., and Djerboua, A.: Toward real-time daily QPPF 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.
- Marty, R., Zin, I., and Obled, C.: Sensitivity of hydrological ensemble forecasts to different sources and temporal resolutions of probabilistic quantitative precipitation forecasts: flash flood case studies in the Cévennes-Vivarais region (Southern France), *Hydrol. Process.*, 27, 33–44, doi:10.1002/hyp.9543, 2013.
- Matheson, J. E. and Winkler, R. L.: Scoring rules for continuous probability distributions, *Manage. Sci.*, 22, 1087–1095, 1976.

- Matulla, C., Zhang, X., Wang, X. L., Wang, J., Zorita, E., Wagner, S., and von Storch, H.: Influence of similarity measures on the performance of the analog method for downscaling daily precipitation, *Clim. Dynam.*, 30, 133–144, doi:10.1007/s00382-007-0277-2, 2008.
- Muluye, G. Y.: Deriving meteorological variables from numerical weather prediction model output: A nearest neighbor approach, *Water Resour. Res.*, 47, W07509, doi:10.1029/2010WR009750, 2011.
- Nam, D. H., Udo, K., and Mano, A.: Downscaling global weather forecast outputs using ANN for flood prediction, *J. Appl. Math.*, 2011, 246–286, doi:10.1155/2011/246286, 2011.
- Neuwirth, E.: RColorBrewer: ColorBrewer palettes, <http://CRAN.R-project.org/package=RColorBrewer> (last access: 24 October 2013), 2011.
- Nuissier, O., Joly, B., Joly, A., Ducrocq, V., and Arbogast, P.: A statistical downscaling to identify the large-scale circulation patterns associated with heavy precipitation events over southern France, *Q. J. Roy. Meteorol. Soc.*, 137, 1812–1827, doi:10.1002/qj.866, 2011.
- 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.
- Pagé, C. and Terray, L.: Nouvelles projections climatiques à échelle fine sur la France pour le 21ème siècle : les scénarii SCRATCH2010, Technical Report TR/CMGC/10/58, CERFACS, Toulouse Cedex, France, 2010.
- Pagé, C., Terray, L., and Boé, J.: Projections climatiques à échelle fine sur la France pour le 21ème siècle: les scénarii SCRATCH08, Technical note, Climate Modelling and Global Change TR/CMGC/08/64, CERFACS, Toulouse, France, 2008.
- Pebesma, E. and Bivand, R.: Classes and methods for spatial data in R, *r news* 5th Edn., <http://cran.r-project.org/doc/Rnews/> (last access: 24 October 2013), 2005.
- Pierce, D.: ncd: Interface to Unidata netCDF data files, *r* package version 1.6.6, <http://CRAN.R-project.org/package=ncdf> (last access: 24 October 2013), 2011.
- Pincus, R. and Rew, R.: NetCDF Fortran 90 Guide, netcdf 4.1.3 Edn., UCAR Unidata, Boulder, Colorado, 2011.
- 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., Ribes, A., Martin, E., Habets, F., and Boé, J.: Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins, *J. Hydrol.*, 383, 111–124, doi:10.1016/j.jhydrol.2009.09.050, 2010.
- 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.
- Raje, D. and Mujumdar, P. P.: A comparison of three methods for downscaling daily precipitation in the Punjab region, *Hydrol. Process.*, 25, 3575–3589, doi:10.1002/hyp.8083, 2011.
- R Development Core Team: R: A Language and Environment for Statistical Computing, <http://www.R-project.org/>, R Foundation for Statistical Computing, Vienna, Austria, 2012.
- Ricard, D., Ducrocq, V., and Auger, L.: A climatology of the mesoscale environment associated with Heavily Precipitating Events over a Northwestern Mediterranean Area, *J. Appl. Meteorol. Clim.*, 51, 468–488, doi:10.1175/JAMC-D-11-017.1, 2012.
- Sauter, T. and Venema, V.: Natural three-dimensional predictor domains for statistical precipitation downscaling, *J. Climate*, 24, 6132–6145, doi:10.1175/2011JCLI4155.1, 2011.
- Teuling, A. J., Stöckli, R., and Seneviratne, S. I.: Bivariate colour maps for visualizing climate data, *Int. J. Climatol.*, 31, 1408–1412, doi:10.1002/joc.2153, 2011.
- Teutschbein, C., Wetterhall, F., and Seibert, J.: Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale, *Clim. Dynam.*, 37, 2087–2105, doi:10.1007/s00382-010-0979-8, 2011.
- Teweles, S. and Wobus, H. B.: Verification of prognostic charts, *B. Am. Meteorol. Soc.*, 35, 455–463, 1954.
- Thiemeßl, M. J., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, *Int. J. Climatol.*, 31, 1530–1544, doi:10.1002/joc.2168, 2011.
- Timbal, B. and McAvaney, B. J.: An analogue-based method to downscale surface air temperature: application for Australia, *Clim. Dynam.*, 17, 947–963, doi:10.1007/s003820100156, 2001.
- 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.
- Timbal, B., Hope, P., and Charles, S.: Evaluating the consistency between statistically downscaled and global dynamical model climate change projections, *J. Climate*, 21, 6052–6059, doi:10.1175/2008JCLI2379.1, 2008.
- Turco, M., Quintana Seguí, P., Llasat, M. C., Herrera, S., and Gutiérrez, J. M.: Testing MOS precipitation downscaling for ENSEMBLES regional climate models over Spain, *J. Geophys. Res.*, 116, D18109, doi:10.1029/2011JD016166, 2011.
- 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, I., 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. Meteorol. Soc.*, 131, 2961–3012, doi:10.1256/qj.04.176, 2005.
- Vidal, J.-P. and Wade, S. D.: A framework for developing high-resolution multi-model climate projections: 21st century scenarios for the UK, *Int. J. Climatol.*, 28, 843–858, doi:10.1002/joc.1593, 2008.
- Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyrou, 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, 2010.

- Vidal, J.-P., Martin, E., Kitova, N., Najac, J., and Soubeyroux, J.-M.: Evolution of spatio-temporal drought characteristics: validation, projections and effect of adaptation scenarios, *Hydrol. Earth Syst. Sci.*, 16, 2935–2955, doi:10.5194/hess-16-2935-2012, 2012.
- Voisin, N., Schaake, J. C., and Lettenmaier, D. P.: Calibration and downscaling methods for quantitative ensemble precipitation forecasts, *Weather Forecast.*, 25, 1603–1627, doi:10.1175/2010WAF2222367.1, 2010.
- Vrac, M., Stein, M., and Hayhoe, K.: Statistical downscaling of precipitation through nonhomogeneous stochastic weather typing, *Clim. Res.*, 34, 169–184, doi:10.3354/cr00696, 2007.
- Wetterhall, F., Halldin, S., and Xu, C.-y.: Statistical precipitation downscaling in central Sweden with the analogue method, *J. Hydrol.*, 306, 174–190, doi:10.1016/j.jhydrol.2004.09.008, 2005.
- Wetterhall, F., Halldin, S., and Xu, C.-Y.: Seasonality properties of four statistical-downscaling methods in central Sweden, *Theor. Appl. Climatol.*, 87, 123–137, doi:10.1007/s00704-005-0223-3, 2007.
- Wickham, H.: Reshaping data with the reshape package, *J. Stat. Softw.*, 21, 1–20, 2007.
- Wickham, H.: *ggplot2: elegant graphics for data analysis*, Use R!, Springer, New York, doi:10.1007/978-0-387-98141-3, 2009.
- Wilby, R. L. and Wigley, T. M. L.: Precipitation predictors for downscaling: observed and general circulation model relationships, *Int. J. Climatol.*, 20, 641–661, 2000.
- Wilby, R. L., Hassan, H., and Hanaki, K.: Statistical downscaling of hydrometeorological variables using general circulation model output, *J. Hydrol.*, 205, 1–19, doi:10.1016/S0022-1694(97)00130-3, 1998.
- Zeileis, A. and Grothendieck, G.: zoo: S3 infrastructure for regular and irregular time series, *J. Stat. Softw.*, 14, 1–27, 2005.
- Zhang, F. and Georgakakos, A. P.: Joint variable spatial downscaling, *Climatic Change*, 111, 945–972, doi:10.1007/s10584-011-0167-9, 2012.