CrossMark

ORIGINAL PAPER

Daily and climatological fields of precipitation over the western Alps with a high density network for the period of 1990–2012

Pierre Lassegues[1]

Abstract There is still considerable uncertainty about precipitation at high elevation in mountain terrain due to the relatively few in situ measurements available and to the particular variability of the parameter. In this study, several spatialization techniques were tested, some for climatological time scale and others for daily fields, for precipitation over the western Alps for the period of 1990–2012. The study domain and period were chosen for the quality of available in situ observations and density of the network. First, a weather-type classification was established with a technique based on canonical correlation analysis combining large- and regional-scale data. The spatialization techniques applied for the climatological time scale were adapted from the Aurelhy method which uses elevation and principal components of the topography as predictors. The spatialization techniques applied to daily fields were based on kriging of daily rain gauges and used the climatological fields as predictors. This study aims to validate the advantage of using the climatology of the weather type of the day as predictor for daily fields over a monthly climatology. The climatology of the weather type of the day seems to demonstrate some small improvement.

Finally, annual means over the period of 1990–2012 were produced using several methods, including some from accumulation of daily fields and others from the spatialization of in situ station means. Precipitation at high elevations and vertical climatological gradients were particularly scrutinized. Annual means based on sums of daily fields seem to have better performances.

This paper only presents results for precipitation but temperature was also analysed.

1 Introduction

Knowledge of meteorological conditions over mountainous areas is of great interest, particularly in terms of precipitation which has important economic consequences for agriculture, water resources, hydro-power, tourism and transport. The study of meteorological and climatological conditions of rainfall is important in its own right but also in terms of its close links with other sciences including hydrology, biology, glaciology, distribution and diversity of vegetation and wild life (Whiteman 2000; Barry 2008). Precipitation is also a key factor for natural hazards, of which the most important for mountainous zones are flash floods and avalanches.

Today, understanding climate change mechanism is critical. In this context, chronological long-term time series of spatial analysis of meteorological variables is a valuable tool for a present and past reference climate database.

This is particularly important for mountainous areas, for which the vertical stratification is expected to show enhanced response to climate change (Beniston 2006).

Analyses of the meteorological parameters at short (hourly, daily) and long time steps (monthly, yearly, decadal, 30-year climatology) show very different spatial variability and physical patterns. They also differ in terms of the spatial analysis techniques applied.

The basis of knowledge of precipitation still relies on in situ measurement networks.

✉ Pierre Lassegues
  pierre.lassegues@meteo.fr

1 Météo-France Direction de la Climatologie (DCSC), 42 Av. Coriolis, 31057 Toulouse, France

Large networks of in situ measurements of precipitation have existed for over a hundred years in Europe, even in mountainous areas, but are constantly evolving. Measurement networks are still heterogeneous and it is difficult to manage measurement condition information and build homogeneous long-term time series for climatological purposes. Finally, the spatial density of good quality long-term time series without gaps is very limited due to the difficulty of preserving stable networks. These features are enhanced in mountain regions. Networks of in situ measurement for precipitation suffer from the lack of high elevation stations.

The harsh conditions (wind, snow, freezing temperatures) have deplorable consequences on the quality and representativeness of data collected, particularly for precipitation (Goodison et al.1998; Sevruk 2005; Sevruk et al. 2009).

Even with a high quality and very dense network of in situ measurements, sophisticated spatial analysis techniques are required to deliver a valid estimate of the meteorological variables.

There are many spatial analysis techniques available (COST-79 1997; Hofstra et al. 2008, 2010; Li and Heap 2008). Three major families of techniques are commonly distinguished: deterministic methods, statistical methods and geostatistical methods. Common examples of the first family are inverse distance weighting methods (or other weighting functions like Epanechnikov). Examples of the second family are linear models (Ripley 1981) and also splines with a smoothing parameter adjustment (here thin plate splines) as presented in Hutchinson (1998). Today, Kriging is a very popular technique in the third family (Matheron 1963; Cressie and Cassie 1993; Goovaert 1997). In this case, the variable of interest must show a spatial dependence and the latter must verify a hypothesis of stationarity. The analysis of spatial dependence is realized through the spatial covariance function or the variogram.

Statistical and geostatistical methods produce unbiased estimates with the lowest error possible.

Another important difference among spatial analysis techniques is between those using only the in situ measurement sample and those also using ancillary variables. The latter techniques are highly effective when the sampled values are of poor quality and density and when the associated variables (one or several) are closely correlated to the parameter of interest and available at high spatial density. This is the case for analysis of hourly and daily precipitation fields with the help of radar data (De Gaetano and Wilks 2009; Wuest et al. 2009). Very often, the orography parameters are introduced here (not only elevation but often elaborate parameters) (Prudhomme 1998).

For climatological analysis, two early important studies must be briefly described here as they illustrate an interesting use of orography parameters.

The first one is PRISM proposed by Christopher Daly (1994). It is based on a local linear model of precipitation with elevation as predictor. The characteristic of this method is an elaborate selection and weighting of the close neighbour sampled values. This technique was applied by Schwarb for the Alps, Gottardi for French mountains, adapted by Brunetti for temperature in Italy and again by Mergili for precipitation in Tyrol (Schwarb 2001; Gottardi et al. 2012, Brunetti et al. 2013; Mergili and Keschner 2015).

The second one is the Benichou and Le Breton (1987) Aurelhy method (Benichou 1994). This technique is based on a principal component analysis of the orography. For each point of the target grid, representative points of the neighbourhood are selected, here called landscapes. A principal component analysis (hereafter PCA) is applied on the matrix of landscapes. The first components are representative of slope vectors, curvature, convexity/concavity and other characteristics of the relief. They are not correlated and have turned out to be good parameters for a linear model for a first estimate of the meteorological variable. An improvement in the result is then added with a kriging of the residuals.

A new software for the implementation of the Aurelhy method was developed in 2014 at Météo-France and this spatial analysis technique plays a central role in the present study.

The two methods, PRISM and Aurelhy, have in common the idea that a statistical relationship between orography and precipitation can be found at the climatological time scale. In this study, as most of the time in the literature, statistical models using orography parameters for the spatialization of precipitation are considered suitable only at the climatological time scale and not at the daily or hourly time scale.

Another technique to improve the quality of spatial analysis for precipitation is to discriminate between the daily data with the help of a weather type or a weather regime classification.

We can confidently suppose that there are typical spatial patterns of the meteorological variable at the local scale associated with the general large-scale meteorological situation. This idea has been developed and applied in many studies as, for example, Courault and Monestiez (1999) for temperatures over southeast of France, Tveito (2007) for temperature in Norway and Esteban et al. (2008) for precipitation over the central Pyrenees. Gottardi (2012) has produced an analysis of precipitation and temperature over the French mountains based on the PRISM technique with the help of a weather-type stratification. Masson and Frei (2014) have also tested the capacity of circulation types to improve spatial analysis of precipitation over the Alps.

The present study is devoted to spatial analysis of precipitation over the western Alps. The available data, the period, the study area and the main goals of the study are first presented (Section 2). A weather-type classification with an original method is then proposed (Section 3). A climatological spatial analysis of precipitation is produced within each weather type with the help of the Aurelhy method and using

topology parameters as predictors (Section 4). The climatology of the weather type of the day is then used as guess pattern for the spatialization of daily data to the production of long-term time series of daily fields (Section 5). One of the goals of the study is to analyse the advantage of a weather-type climatology as predictor to the spatialization of daily data.

Finally, four climatological annual means are produced, two from spatialization of in situ station climatological means and two from accumulation of daily grids (Section 6). These four final results are then compared in a validation process with the aim of improving precipitation estimations at high elevation.

2 Section 1: preliminary steps

2.1 Study area and data

The area studied (Fig. 1) covers the western Alps: the French part of the Alps, western Switzerland and western Italy. This area includes a large part of the high mountain regions of the ridge but also lower regions in France in the southwest of the zone, the west of the Pô valley in Italy, the Jura in the northwest between France and Switzerland and a small maritime area in the southeast. This is a rectangular zone 210 × 360km. The longitude and latitude coordinates of the corners of the domain are NW 5.35–46.76°, NE 8.09–46.66°, SW 5.18–43.52°, SE 7.77–43.42°. The geographic data in this study is processed in Lambert II projection.

The total surface is 75,600 km^2. The mean elevation is 1202 m. More information on the vertical profile of the elevation is given in Fig. 3.
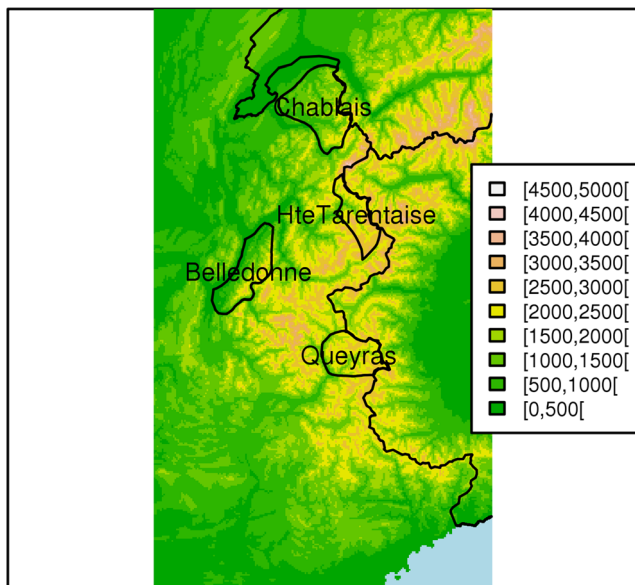

Fig. 1 Study domain over western Alps (210 × 360 km) with topography and mention of the four local massifs Chablais, Belledonne, Haute-Tarentaise and Queyras (see Section 6)

The digital elevation model (hereafter DEM) comes from the NASA Shuttle Radar Topography Mission (http://www2.jpl.nasa.gov/srtm/). The original data come from the CGIAR-CSI web site at resolution of 1/1200th degree (corresponding to about 90 m) on a lon/lat projection. For our study, this DEM was first interpolated on a 1/120th degree grid with a cubic splines method and then re-projected on a 1-km LambertII grid with the library gdal (http://www.gdal.org/).

The in situ measurements are daily precipitation (06hUTC-06hUTC), over a 23-year period from 1990 to 2012. The French data come from the Météo-France DClim data-base, the Swiss data from the Meteo-Swiss IDAWEB database and the Italian data from ARPA Piemonte and ARPA Val d'Aosta.

For the weather-type classification, daily fields of mean sea level pressure (here after MSLP) from the ERA-Interim reanalysis were also needed. The ERA-Interim reanalysis (Dee et al. 2011) is a European Centre Medium-Range Weather Forecast global reanalysis produced with a 3D-Var assimilation model with a horizontal resolution of about 80 km covering a period beginning in 1979 and continuously updated to real time. The MSLP fields are on the domain 30° W, 30° E, 30° N, 70°N on a 2.5° projection (daily fields at 12 h UTC).

2.2 Different processes and critical issues analysed

The different steps are shown in the flowchart in Fig. 2.

The most important steps are identified from number 1 to 6 on the flowchart as follows:

1. Production of a spatial analysis of monthly climate normals from 1990 to 2012 with Aurelhy method
2. Production of daily fields using the previous monthly climatological fields as predictor
3. Production of a weather-type classification based on a canonical correlation analysis of the daily fields at the local scale and corresponding daily fields of the large-scale parameter MSLP
4. Production of a climatology for each weather type with Aurelhy method
5. Production of daily fields using the previous climatological field corresponding to the weather type of the class of the day as predictor.
6. Production of four different 1990–2012 annual means from aggregation of the fields of steps 1, 2, 4, 5 for the purpose of inter-comparison and validation of the different methods.

The analysis presented here was applied separately for precipitation, minimum temperature and maximum temperature. However, for the sake of brevity, only the results for precipitation are presented in this paper.
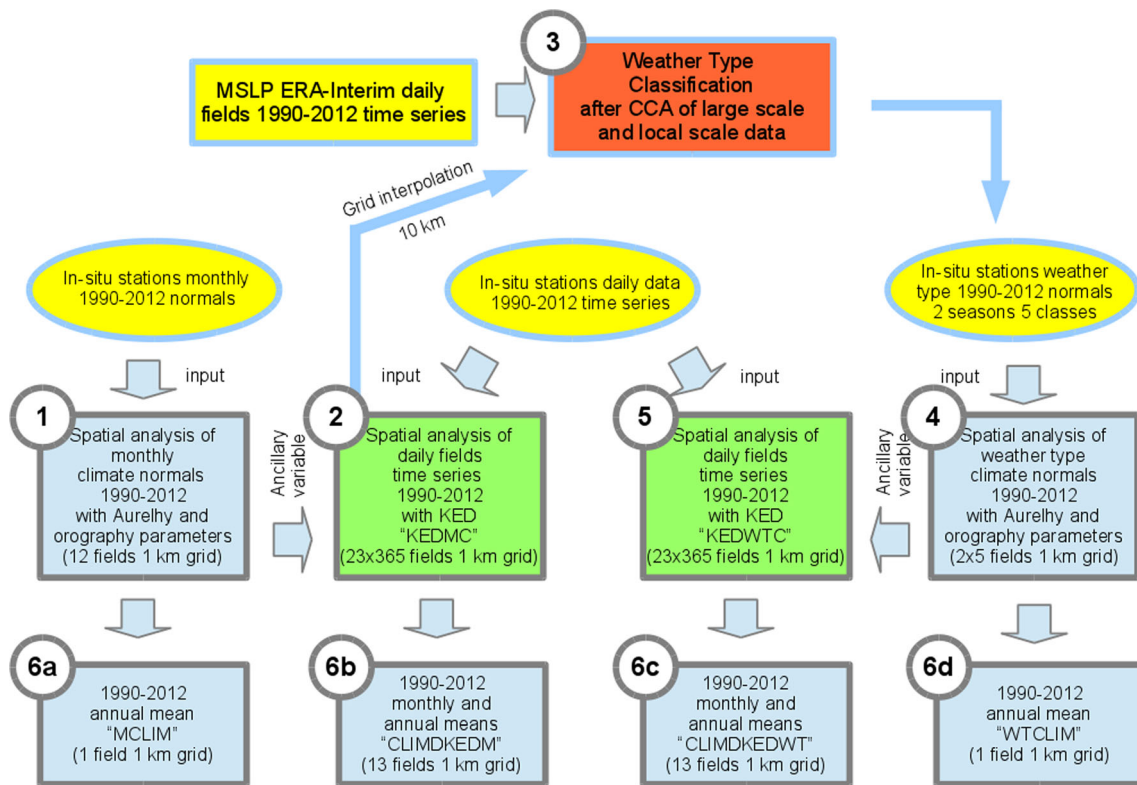
Fig. 2 Flowchart of the study

The different steps of the study are described in detail in the following sections, except for the two preliminary proceedings (steps 1 and 2).

Step 1 is a monthly climatological spatial analysis for precipitation realised with the Aurelhy method. The details for the implementation of a climatological spatial analysis with the Aurelhy method are fully described in Section 4 of the study, after a weather-type classification (step 4 of the flowchart).

Similarly, step 2 is the production of daily fields using the previous monthly climatological fields as predictor. This technique is fully presented for the daily fields of step 5 (see Section 5).

One of the most important aims of this study is to analyse the benefit of using a climatological field as predictor in the production of daily fields. The idea is that the climatological field produced with Aurelhy involves the statistical relation between topography and precipitation. In this study, the climatological field is introduced as predictor in the spatial analysis of the daily fields in order to enhance the quality of the daily analysis. To this end, two candidate climatological fields for predictor were tested. The first one is a common monthly climatological field (step 2) as applied by Isotta et al. (2013). The second one is the climatological field of the weather type of the day (step 5).

Another important objective of this study is to compare different methods used to produce climatological means. Some of the latter are based on spatial analysis of the climatological statistics of observation stations (steps 1, 4, 6a, 6d)

while others are based on accumulation of daily fields (steps 2, 5, 6b, 6c). The different methods are sometimes called aggregation-integration for the first category and integration-aggregation for the second category (Journée 2015).The difficulty with the first category is that climatological statistics from observation stations are often of low density and preserving constant quality and stable measurement conditions is difficult. Many stations have large gaps during the climatological period, they are often displaced or closed and their equipment can change. The second category, with spatial analysis of daily fields, benefits from a higher spatial density of observation stations for each day and the final mean or accumulation of these daily fields may prove to be less dependent on the low quality or occasional failure of a particular observation station (Frei and Schär 1998; Isotta et al. 2013).

2.3 Density of the data, quality control and variable transformation

This study required a very dense observation network. Table 1 shows the mean density for each country. Unfortunately, it was not possible to obtain dense enough data before 1990.

The number of daily observations for each year during the period studied is not absolutely stable and has a slight tendency to increase progressively over the period of 1990–2012. However, this difference in density is small and meets the requirements of the study.

Table 1 Mean number and density of daily rain gauges after quality control

Study domain	French part	Italian part	Swiss part
691	509	107	79
1 station for 106.8 km^2	1 station for 94.3 km^2	1 station for 150.0 km^2	1 station for 123.5 km^2
9.4 stations/10,000 km^2	10.6 stations/10,000 km^2	6.7 stations/10,000 km^2	8.1 stations/10,000 km^2

After having checked that the density requirements were satisfied, all the data available were submitted to a quality control (QC) process. This automatic quality control process is derived from the operational one used at Météo-France.

All controls were automatic. It should be noted that there is necessarily a proportion of errors in this QC system (erroneous qualification doubtful or undetected wrong values). However, for the operational QC process of Météo-France, it was checked that this proportion was small.

In this study, the share of values qualified as doubtful was 1.6 %.

It was checked that the PDF of the data after rejection of doubtful values was nearly the same as original data.

A final pre-processing of observation data was to check how the high elevation was sampled. To this end, histograms of the mean daily density of observations by elevation steps of 250 m are shown in Fig. 3. The number and proportion of observation stations are shown with, in parallel, the surface of the corresponding elevation step.

This histogram shows an important problem: the lack of observation stations at high elevation. For elevations above 2000 m, there are very few rain gauges. For this reason, spatial analysis of meteorological parameters in mountainous regions requires sophisticated techniques and is very challenging.

The question of variable transformation must be discussed here. Daily precipitation has a particularly skewed distribution. In this study, the techniques applied were regression and kriging which are linear techniques designed for Gaussian variables. In many studies, these techniques are applied for precipitation after variable transformation. The Box and Cox transformation is a technique adapted for all kinds of variables because it can be applied with different degrees of intensity (Erdin et al. 2012). In this study, log transformation and square root transformation were tested (which are equivalent to two commonly used Box and Cox transformations). The difficulty with these techniques when they are applied with regression-kriging and kriging with external drift is that back transformation is not direct and needs to use the variance of error. The particularities of this transformation and back transformation, so-called Trans-Gaussian Kriging, are presented in Cressie and Cassie (1993).

3 Section 2: weather-type classification based on canonical correlation analysis of large- and local-scale data

Classification is an analysis method that has been used for decades in meteorology. Results from Hess and Brezowsky (1952) Grosswetterlagen for central Europe and Lamb (1972) for the British Isles were very popular at the time. Today, the first methods based on human expertise have been replaced by automatic statistical methods.

The recent Cost 733 European project (2004–2010) has produced interesting results on the topic of classifications for Europe (Huth et al. 2008; Beck and Philipp 2010; Philipp et al. 2010).

The idea behind weather-type classifications for mid-latitude climate and particularly for Europe is that there must be a small number of characteristic meteorological situations usually lasting a few days and then changing rapidly from one to another.

There are many specific applications of weather-type classification for mountain climate including the French Safran model precipitation analysis by Durand et al. (1993) and the Swiss Alpine snow pack analysis by Scherrer and Appenzeller (2006). Other references were previously cited (Gottardi 2012; Esteban et al. 2008; Masson and Frei 2014).

In the vocabulary used in this study, large meteorological scale (or synoptic scale) is for meteorological analysis covering an area of several thousands of kilometres, while local scale covers an area from ten to several hundreds of kilometres. In this study, the large-scale area covers western
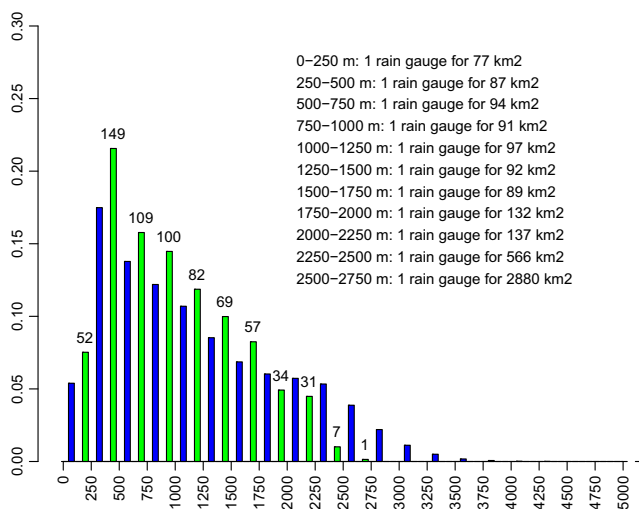
0–250 m: 1 rain gauge for 77 km2
250–500 m: 1 rain gauge for 87 km2
500–750 m: 1 rain gauge for 94 km2
750–1000 m: 1 rain gauge for 91 km2
1000–1250 m: 1 rain gauge for 97 km2
1250–1500 m: 1 rain gauge for 92 km2
1500–1750 m: 1 rain gauge for 89 km2
1750–2000 m: 1 rain gauge for 132 km2
2000–2250 m: 1 rain gauge for 137 km2
2250–2500 m: 1 rain gauge for 566 km2
2500–2750 m: 1 rain gauge for 2880 km2

Fig. 3 *Blue*: relative surface of layers by step of elevation of 250 m over the study domain. *Green*: mean daily number of rain gauge observations for each layer

Europe and the eastern Atlantic zone and the local scale is the domain over the western Alps. Furthermore, a weather regime classification (sometimes also called circulation patterns) is based only on large-scale analysis, while a weather-type classification is based on the merging of large-scale and local-scale analyses. In this context, the technique applied in this study must be called a weather-type classification and not a weather regime.

A preliminary question needed to be addressed is should the classifications be applied to the data with a seasonal division or not? Two seasonal subdivisions were tested: a subdivision in two seasons of 6 months (ONDJFM and AMJJAS) and a subdivision in four seasons of 3 months (DJF, MAM, JJA, SON).

A flowchart of the classification is presented in Fig. 4.

The weather-type classification proposed in this study is based on a canonical correlation analysis (here after CCA) of the large- and local-scale parameters. A presentation of the theory and advantages of CCA for this kind of coupling is proposed by Bretherton et al. (1992). Von Storch and Zorita (1993) have also presented an interesting analysis of downscaling of large-scale signals to regional monthly precipitation over the Iberian Peninsula.

The CCA is applied here to the following two sets of data: the large-scale parameter matrix with the points of a grid over western Europe and east Atlantic in columns and daily index in rows and an equivalent matrix for local-scale parameter (precipitation) with points of the grid over the western Alps in columns and daily index in rows. The CCA produces a set of two groups of new variables, the canonical variables, with maximum correlation by pairs but orthogonal otherwise. These canonical variables are sorted by decreasing correlation by pairs and decreasing correlation between them and the original variables. A limited number of these new variables is then selected and is considered to be representative of the most informative correlation of the original variables.

In our study, several parameters were tested for the representation of the large scale. In the same way, different grid sizes and ranges of the large-scale domain were tested. This analysis resulted in the choice of MSLP over a northern Atlantic-Western Europe domain 30° W, 30° E, 30° N, 70° N on a 2.5° lon/lat grid (425 points). As the original data are on a regular lon/lat grid, a weighting by the cosinus of the latitude is applied. The annual cycle of the large-scale data is removed with the help of a fast Fourier transform function.

The regional-scale data for the CCA are daily analysis of precipitation over the western Alps shown in step 2 in Fig. 2 but finally interpolated on a 10-km grid (790 points).

For precipitation, a preliminary log variable transformation of the daily data is applied, due to the highly skewed distribution, as discussed in the previous section.

The final stage of the CCA is to select a reduced number of canonical variables with the help of two techniques: the Bartlett test (Bartlett 1941) and an implementation of a forecast model of the local variable with the large-scale variable as predictor and analysis of explained variance (Barnston 1999).
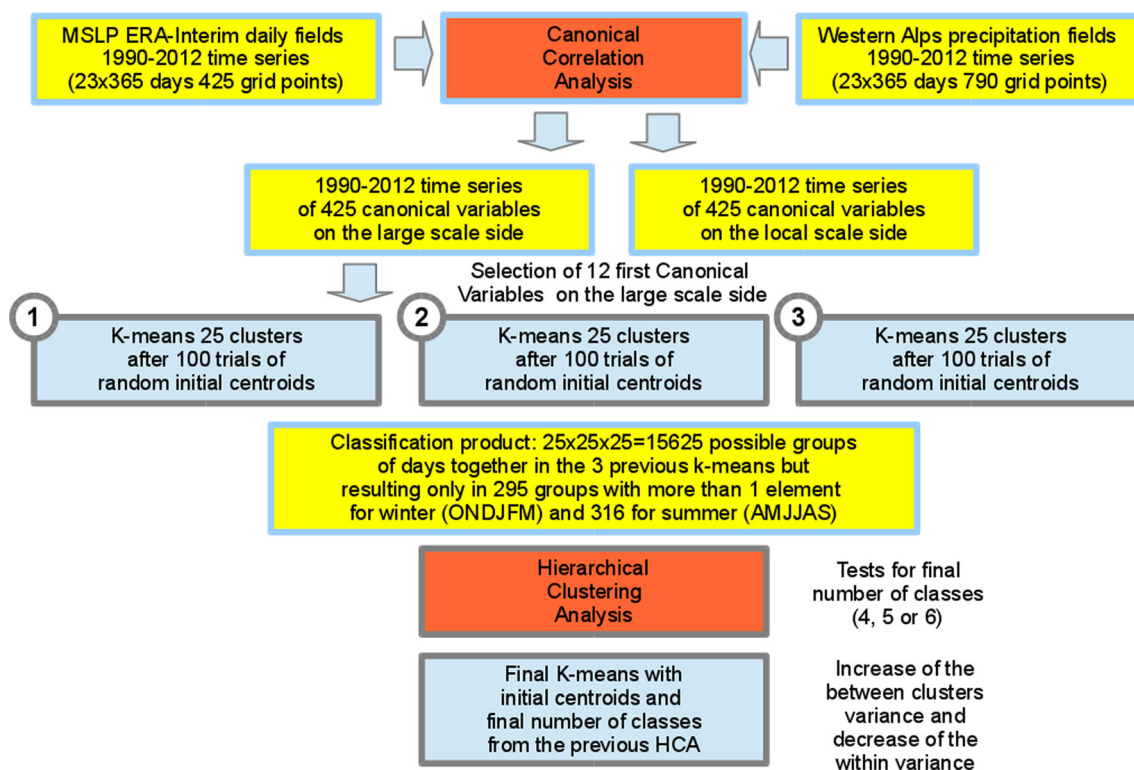

Fig. 4 Graph of the classification

This resulted in the selection of 12 canonical variables with an explained variance for precipitation of 62 % for ONDJFM and 60 % for AMJJAS.

Each of the 365 × 23 days of the period of 1990–2012 is then represented by 12 canonical variables and a classification algorithm is applied to these variables. Only the canonical variables on the side of the large-scale original variables are used (this technique preserves the possibility of further developments, see conclusion).

Many different techniques for automatic classifications are available today. However, the most frequently used methods in data mining are still hierarchical clustering, hereafter HCA, (Ward 1963; Hastie et al. 2009) and K-means centroid-based clustering (Forgy 1965; MacQueen 1967).

The method applied here belongs to the family of hybrid methods (Wong 1980) based successively on K-means, HCA and again K-means (see work flow Fig. 4).

The first level is based on K-means multiple trials and is designed to produce a large number of clusters. Three K-means partitions with 25 clusters are produced here. Each one is selected after 100 trials of random initial centroids, with the help of the similarity criterion proposed by Michelangeli (1995). Groups of days always together in the 25 × 25 × 25 = 15,625 possible groups of the partition product are then searched for. Only a few of the 15,625 possible groups had more than 1 day: 295 groups for the winter season (ONDJFM) and 316 groups for summer (AMJJAS). This is considered as a first level of partitioning.

After this first step, the HCA is applied on the 295 clusters for winter and 316 clusters for summer. This HCA is on Euclidean distance with Ward's linkage criterion. This HCA is designed to help choose the final number of clusters.

Finally, a new K-means is applied with the centres of classes coming from the HCA as initial centroids. This final K-means again improves the between-clusters variance.

Objective measures were also tested to determine the best number of clusters: Rousseuw's silhouette index (Rousseuw 1987), classifiability and reproducibility of Michelangeli et al. (1995) and others proposed by Huth et al. (2008). In our study, the classifications were tested for 4, 5 and 6 classes.

A climatological validation of the classifications and selection of the final experiment is then realised with the help of tercile anomalies for the local scale, a validation technique proposed by Plaut (2001) and Simonnet (2001): the climatological tercile thresholds for light, medium and heavy precipitations are calculated for each grid point and anomalies of frequencies in each class according to these thresholds are analysed. See Fig. 5 for an illustration of such result. This analysis resulted in the selection of a two-season, five-cluster classification.

The patterns emerging from the classification, at the local scale, are expected to show important contrasts between the different sides of the alpine ridge, depending on the dominant
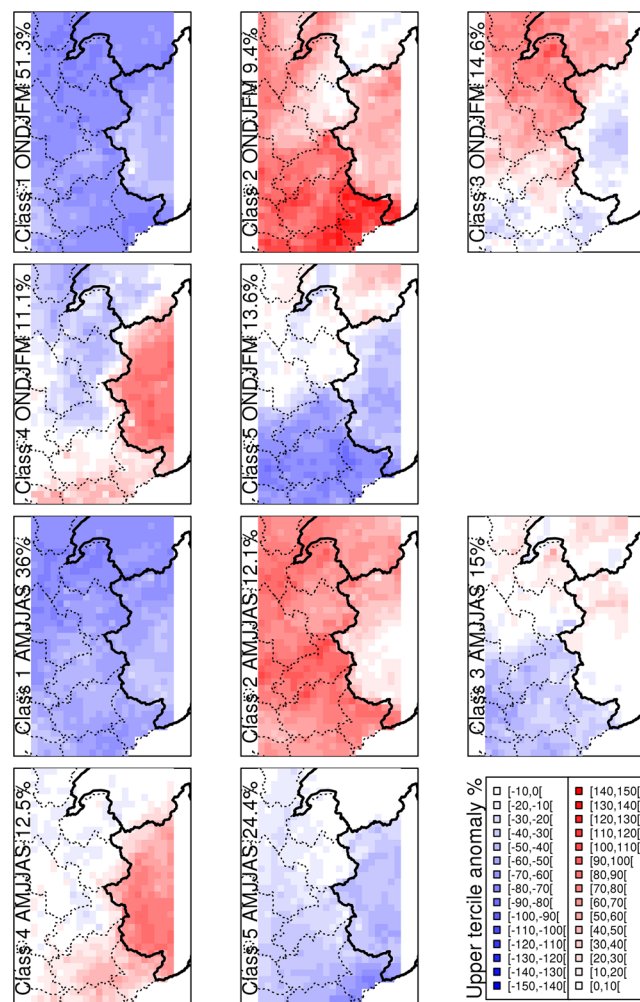

Fig. 5 Upper tercile anomalies (%) for weather-type classification for precipitation (five classes, two seasons) and frequencies of the classes

meteorological fluxes in each weather type and associated with the large-scale synoptic patterns.

To illustrate the ability of the classification to discriminate between very different situations, the anomalies for the upper terciles for precipitation for each class and each season are shown in Fig. 5. The frequencies of the classes are also shown in this figure.

For the sake of brevity, it is not possible to present in this paper too many maps for large- and local-scale patterns resulting from the classification.

4 Section 3: spatial analysis of weather-type climatological fields with the Aurelhy method

In this section the Aurelhy method is applied for the spatialization of climatologies for each weather type (climate normals for the period of 1990–2012). This spatialization is shown in step 4 of the flowchart in Fig. 2. The implementation of the Aurelhy method is presented here in detail.

The input data for this spatialization are the climatological precipitation for the in situ stations within the weather type (mean daily intensity over all the days of the weather type). The number of stations for each weather type is shown in Table 2. The DEM is also necessary as input data for the Aurelhy method. As the classification is in two seasons and five clusters, ten fields are finally produced here. The grid size of these fields is 1 km.

The Aurelhy method was first implemented by Benichou and Le Breton in 1986 (Benichou 1987; Benichou 1994). A new implementation based on R language was developed in 2014, introducing important improvements. The method is based on a linear model with the principal components of the orography as predictors followed by a kriging of the residuals. The PCA has the interesting advantage of orthogonality, hierarchy and reduced noise. The principal components of the terrain model are produced on a matrix with each line corresponding to the elevations of a selection of points in the vicinity of each point of the target grid. They are called the landscapes in the specific Aurelhy vocabulary. The most important parameters of this selection are range, density and shape. These parameters can be fitted after testing. One usual choice for the landscapes is a square matrix of 11 × 11 points with a 5-km lag. But, other parameters are often selected. The importance of the spatial scale to the relation between the meteorological variable and the topography is discussed in Gyalistras (2003) and in Masson and Frei (2014). Most of the time, elevation itself is also introduced as candidate predictors along with the principal components.

An automatic selection of the candidate predictors is then applied. The methods available are increase of R-squared, Akaike information criterion (AIC), Bayesian information criterion (BIC) and Mallows's CP.

The second step of the Aurelhy method is then a kriging of the residuals and addition of the two contributions. The new version offers extensive possibilities in terms of choosing the kriging parameters. Different values for the range and the nugget of the variogram can be tested as well as the choice of the model of variogram.

Aurelhy's first implementation was a spatialization technique of the family Regression-Kriging. Today's version also offers the possibility of choosing the true Kriging with external drift (KED). For a presentation of the differences between regression-kriging and KED, see the publications of Goovaert (1997), Hengl et al. (2003) and Li and Heap (2008).

Finally, a possibility of cross-validation is also implemented in the new software.

Two different issues are addressed in this section: the selection of input data (climatological means of stations) and the methodological choices for the parameterization of the Aurelhy method.

Before applying the spatialization with Aurelhy, we have to compute the climatological means for each station and each weather type. In this process, we have to deal with stations with missing daily values. To preserve the highest spatial density possible, stations with a maximum ratio of missing value of 30 % are selected (however, only a small number of these stations had a high ratio of absent values, more than half had less than 10 %).

To try to improve the quality of the spatialization, two different techniques often applied to prepare the input data for climatological means are also tested. The first technique consists in replacing the missing daily values with a local estimate coming from a spatialization of the day (Journée et al. 2015). The second technique consists in a bias correction of the mean of the stations not available during the entire reference period of 23 years (Isotta et al. 2013). If the station

Table 2 Summary of the Aurelhy spatialization of weather types climatological fields (five classes, two seasons) with number of input in situ stations means, number of predictors selected (principal component of the orography), elevation selected as predictor or not, adjusted R-squared of the linear model, standard deviation of the residuals of the linear model, final model (linear model + kriging of residuals) RMSE after cross-validation and R-squared after cross-validation (unit is mm/day for ResSD and RmseCV)

Class	ONDJFM1	ONDJFM2	ONDJFM3	ONDJFM4	ONDJFM5
NbStations	573	555	566	561	559
NbPred	6	11	7	7	7
Elev	Y	Y	N	N	N
AdjR2	0.41	0.72	0.56	0.62	0.56
ResSD	0.16	2.84	2.68	1.57	2.01
RmseCV	0.1	1.43	1.14	1.07	0.79
FinR2CV	0.77	0.93	0.92	0.82	0.93
Class	AMJJAS1	AMJJAS2	AMJJAS3	AMJJAS4	AMJJAS5
NbStations	554	562	575	543	574
NbPred	6	8	9	9	10
Elev	Y	Y	N	Y	Y
AdjR2	0.46	0.51	0.39	0.62	0.54
ResSD	0.18	1.84	1.9	1.72	0.85
RmseCV	0.1	0.93	0.67	0.93	0.25
FinR2CV	0.83	0.88	0.92	0.88	0.96

is present during *n* days, the bias correction is realized according to the ratio of the mean during these *n* days for the stations of the vicinity without gaps over the mean of the same stations during the whole climatological period.

A leave-one-out cross-validation (each input station mean is put aside one by one, and an independent estimate is produced with the Aurelhy spatialization) was realized with different datasets and showed that the technique of completing the missing values by means of daily spatialization is the best method.

For the second question of optimization of the Aurelhy method, nine different parameterizations were tested. A first version was produced. Then, several new versions were tested, modifying only one parameter each time and objectively evaluating the improvement with a cross-validation (similar to the previous one, for each input station mean). Finally, the different versions showing an improvement were combined and tested again.

In the final Aurelhy parameterization, the selection of points for the PCA (landscapes) is with a range of 50 km and grid density of 10 km. The predictor selection method is BIC exhaustive. Kriging with external drift is applied instead of regression-kriging. A square root preliminary data transformation of precipitation is applied. The variogram is with free nugget, range of 80 km and the fitting method is maximum likelihood.

A summary of the results of this spatialization is presented in Table 2.

An interesting point to note is that, for precipitation, the automatic selection of the predictors in the linear model applied by Aurelhy very often does not select elevation, and the principal components are considered better and sufficient predictors.

The final weather-type climatology for precipitation is presented in Fig. 6. The visual analysis of these maps shows spatial patterns that are very different from one weather type to another in localization and intensity. These patterns are consistent with Fig. 5 of Section 5.

5 Section 4: spatial analysis of daily fields of precipitation with weather-type climatologies as predictor

At this stage of the study, daily in situ observations of precipitation as well as results of the spatial analysis of climatologies for each weather type are available. A production of daily fields is then implemented. In this production, the input data are daily precipitation of all the in situ stations. The climatology of the class of the weather type of the day produced in the previous section is introduced as predictor. This production corresponds to step 5 of the flowchart in Fig. 2. A field on a 1-km grid is produced for each of the 23 × 365 days of the period of 1990–2012.
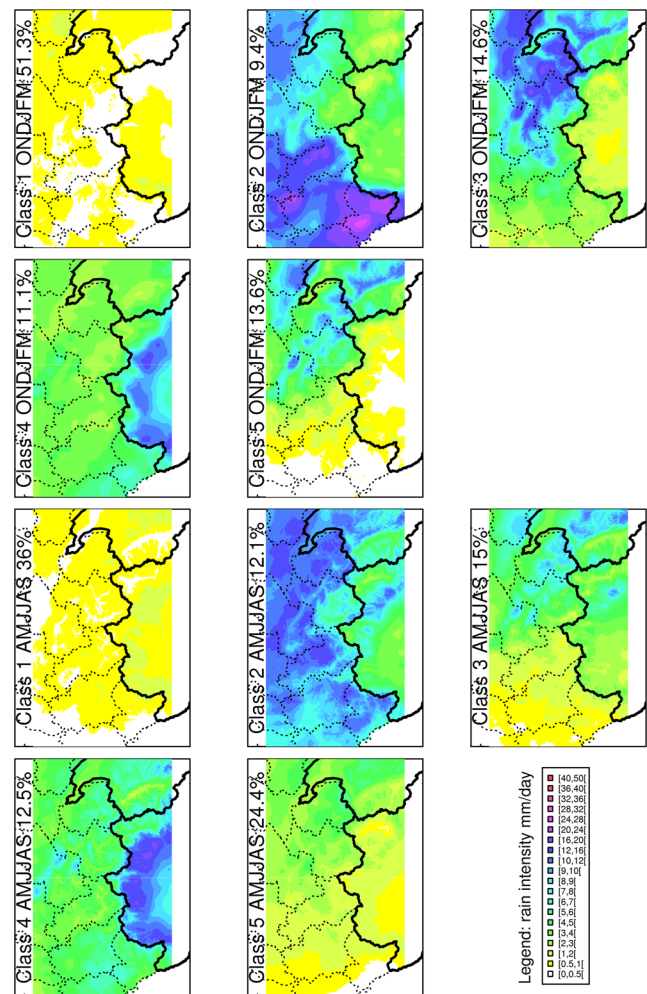

Fig. 6 Climatology of mean daily rainfall intensity (mm/day) for each class of weather type (five weather types, two seasons). Frequencies of the classes

The spatialization technique applied here is very similar to the one applied in step 2 of the study, presented in the flowchart in Fig. 2, except that weather-type climatology is used as predictor here whereas a monthly climatology was used in step 2. Then, the two time series of daily fields are compared.

The first question addressed in this section is to determine if there is any advantage to using weather-type climatology as predictor rather than a monthly climatology in the spatialization of daily data.

The second question developed in this section regards which technique should be used to integrate the climatology as predictor.

Seven distinct methods are compared (five methods with the help of weather-type climatologies as predictor and two methods for the purpose of comparison):

– KEDWTC5c2s: kriging with external drift with the climatology of the weather type of the day (five classes, two seasons) as predictor.

- KEDMC: kriging with external drift with the monthly climatology as predictor (step 2 of Fig. 2).
- RatiosWTC5c2s: ordinary kriging of the ratios of the observations of the day over the corresponding values of the climatology of the weather type of the day.
- LMKResWTC5c2s: regression-kriging with the climatology of the weather type of the day as predictor.
- KEDLogWTC5c2s: same as KEDWTC5c2s but with log transformation of the daily precipitation.
- KEDsqrtWTC5c2s: same as KEDWTC5c2s but with square root transformation of the daily precipitation.
- OKObs: ordinary kriging of observations only. This method is here as a naive reference.

All the variograms applied in the seven methods are with free nugget, range of 80 km and fitting method maximum likelihood.

To decide between the seven methods, different scores against observations are produced.

A leave-one-out cross-validation is realized. This cross-validation consists in putting aside one by one each daily value of in situ stations and producing an independent estimate of this value. It was decided to apply this cross-validation only to days with significant rain (more than 2 mm for the mean value of the observations of the day). However, this resulted in a very large sample of about 10^6 observation-estimation couples.

The RMSE and SEEP score by steps of elevation of 500 m are presented in Table 3.

The SEEP score (Rodwell et al. 2010) is adapted for precipitation, based on a 3×3 contingency table with two thresholds for light, median and heavy precipitation (probabilities 0.33, 0.44, 0.22 on the sample of observations).

It should be noted that the score of RMSE is strongly influenced by heavy precipitation.

The comparison between KEDWTC5c2s and KEDMC is interesting in terms of choosing the best predictor. KEDWTC5c2s shows a small advantage over KEDMC (better results for RMSE and SEEPS for all the classes).

The comparison between KEDWTC5c2s, RatiosWTC5c2s, LMKResWTC5c2s, KEDLogWTC5c2s and KEDsqrtWTC5c2s is interesting in terms of choosing the best spatialization technique. It can be seen that the kriging with external drift after square root transformation shows the best results for RMSE score, and log transformation has the best result for the SEEP score.

An apparently disappointing result is that the six methods with climatology as predictor show only a very small increase in performance compared to the ordinary kriging of observations. However, a visual expertise of the daily maps shows immediately that ordinary kriging of observations results in highly smoothed fields. The true spatial resolution in that case is only the density of the observation network, and the lack of ancillary information coming from the climatology is clearly a weakness.

More scores are produced to improve the analysis. Figure 7 shows the RMSE by classes of precipitation intensity (top panels), the Peirce skill score (middle panels) and the Bias score (lower panels). The Peirce skill score and Bias score (also called frequency bias) are based on 2×2 contingency tables over different thresholds covering the distribution (here 1, 2, 5, 10, 20, 30, 40 mm).

Table 3 RMSE and SEEP score (perfect score 1) by step of elevation for seven spatialization methods of daily precipitation (five methods using weather-type climatology as predictor and two methods as reference for comparison)

RMSE	All	[0–500]	[500–1000]	[1000–1500]	[1500–2000]	[2000–2500]	
nbobs	983,853	284,402	314,469	234,772	119,000	30,033	
KED5c2s	5.507	5.196	5.238	5.560	6.238	7.211	
KEDMC	5.634	5.225	5.336	5.749	6.456	7.555	
Ratios5c2s	6.425	5.004	5.478	6.128	6.752	7.997	
LMKRes5c2s	5.476	5.075	5.191	5.567	6.279	7.432	
KEDLog5c2s	5.577	5.112	5.352	5.649	6.461	7.381	
KEDsqrt5c2s	**5.364**	**4.944**	**5.122**	**5.452**	**6.165**	**7.154**	
OKobs	5.700	5.229	5.406	5.871	6.481	7.826	
SEEPS	**All**	**[0–500]**	**[500–1000]**	**[1000–1500]**	**[1500–2000]**	**[2000–2500]**	**[2500–**
KED5c2s	0.750	0.772	0.764	0.756	0.660	0.559	0.707
KEDMC	0.744	0.772	0.759	0.749	0.645	0.523	0.702
Ratios5c2s	0.765	0.783	0.774	0.777	0.688	0.581	0.719
LMKRes5c2s	0.762	0.783	0.772	0.770	0.676	0.563	0.708
KEDLog5c2s	**0.774**	**0.790**	**0.782**	**0.781**	**0.703**	**0.605**	**0.770**
KEDsqrt5c2s	0.766	0.783	0.774	0.772	0.694	0.599	0.749
OKobs	0.758	0.783	0.765	0.765	0.679	0.567	0.738

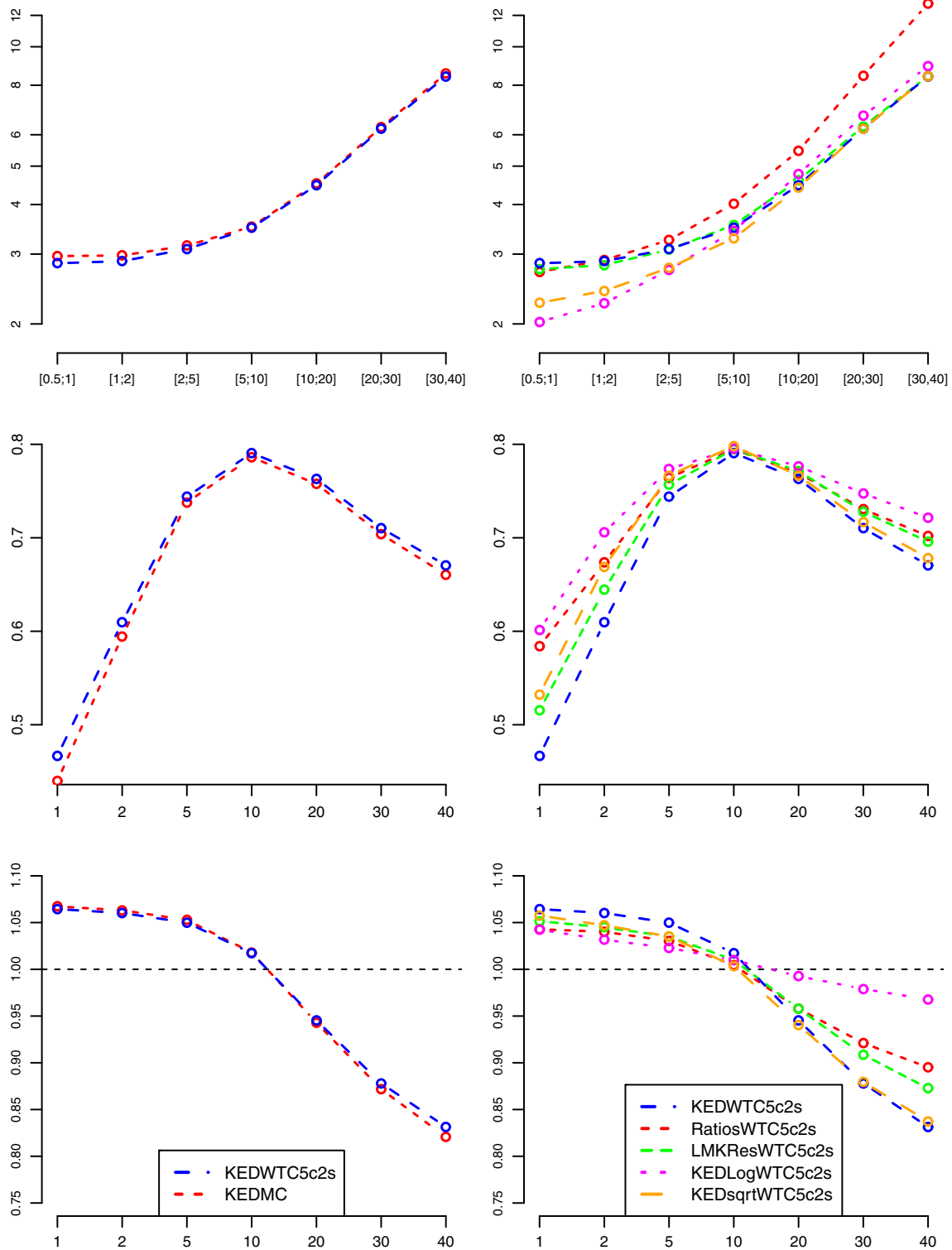Scores for the methods performing the best are in bold characters

Fig. 7 Performances of daily precipitation spatialization (see Section 5). *Top panels*: root mean square error (mm, log scale) for classes of precipitation. *Middle*: Peirce skill score (perfect score = 1) for different thresholds. *Bottom*: Bias score (perfect score = 1) for different thresholds.

Left panels: results for different predictors. *Right panels*: results for different spatialization techniques. Precipitation classes and thresholds are in millimetres

The Peirce skill score is the difference between probability of detection (POD) and probability of false detection (POFD). Figure 7 shows, on the left panels, results for choosing the

predictor and, on the right panels, results for choosing the spatialization technique. The left panels show that KEDWTC5c2s always presents slightly better performances.

But, this advantage is very small. On the right panels, we can see that KEDLogWTC5c2s and KEDsqrtWTC5c2s show the best results for RMSE by classes of intensity and KEDLogWTC5c2s is the best one for Peirce skill score and Bias score.

The conclusion of the comparison presented here is that all the methods show only small differences. However, KEDsqrtWTC5c2s and KEDLogWTC5c2s show the best results among the seven versions. Other elements for indirect validation are added in the following section.

The analysis of the scores for all the spatialization methods for daily precipitation tested in this study shows a smoothing effect (over estimation of low values and under estimation of high values). It should be noted that all the kriging techniques applied were with free nugget value for the fitting of the variograms and a non-zero nugget can be a part of the cause of smoothing effect. However, non-zero nugget value is justified when sampled observations are weakened by a non-negligible uncertainty and we consider that this is the case in this study.

6 Section 5: annual mean of precipitation over the period 1990–2012 and vertical profiles

In this section, several climatological products are compared (annual means over the 1990–2012 period):

–   MCLIM: climatological annual mean produced from the monthly Aurelhy climatology (step 6a of the flowchart Fig. 2).
–   CLIMDKEDM: climatological annual mean produced from daily fields, with these daily fields coming from kriging with external drift with the monthly climatology as predictor (step 6b of the flowchart Fig. 2).
–   WTCLIM: climatological annual mean produced from the weather-type climatology (five classes, two seasons) after a weighting according to the frequencies of each class (step 6d of the flowchart Fig. 2).
–   CLIMDKEDWT: climatological annual mean produced from the daily fields presented in the previous Section 5, with these daily fields coming from kriging with external
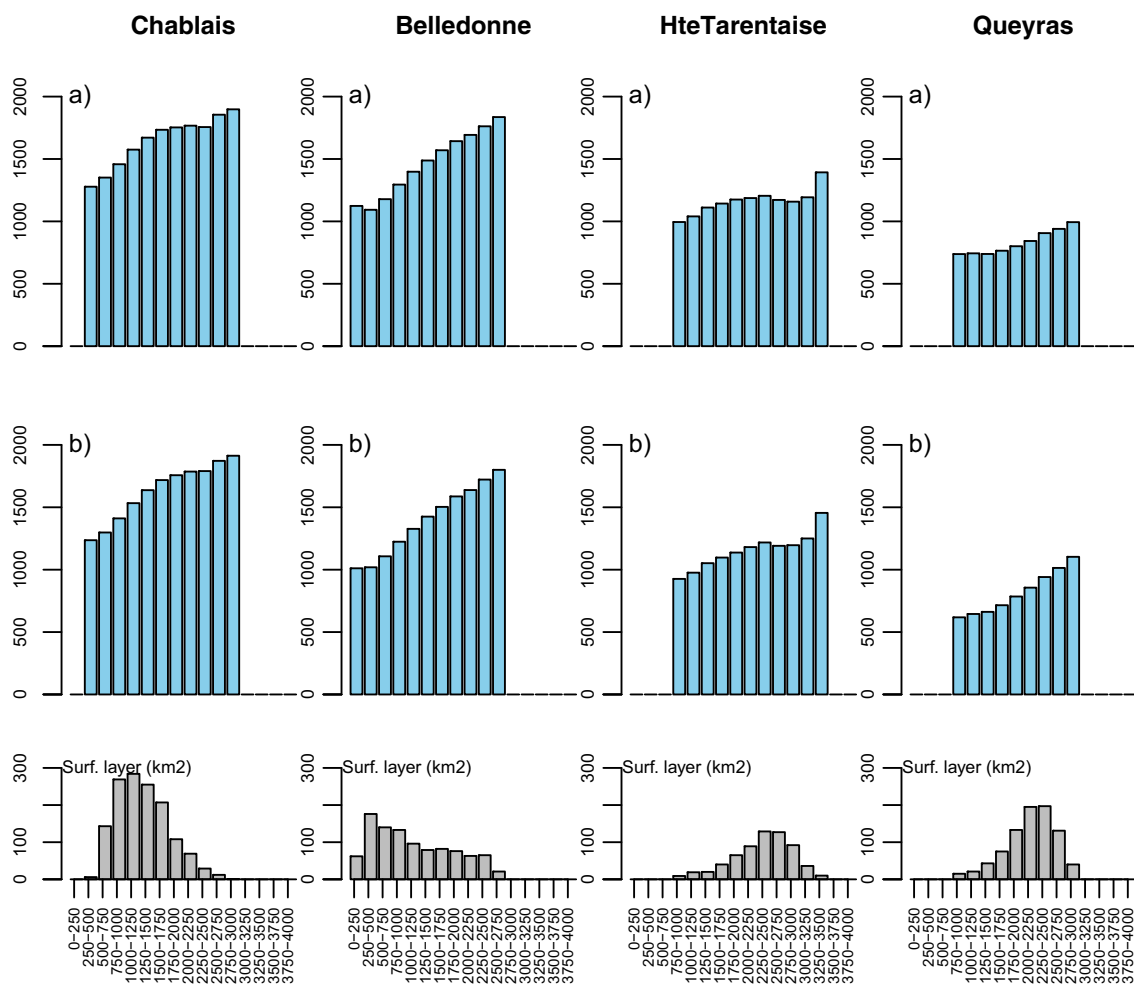

Fig. 8 Vertical profiles of precipitation (mm) for four zones by step of elevation of 250 m for two climatologies produced from stations long-term time series: a WTCLIM: yearly mean from weighting of weather-type climatologies with Aurelhy, b MCLIM: yearly mean from classical monthly climatologies with Aurelhy, *bottom*: surface of the layers for each zone

drift with weather-type climatologies as predictor (step 6c of the flowchart Fig. 2).

– CLIMDOK: climatological annual mean produced from daily fields, with these daily fields coming from ordinary kriging of observations only (OKObs presented in the previous section). This method is here only for a naive reference in the inter-comparison.

All the fields in this section are on a 1-km grid.

The climatologies are compared through local analysis of the vertical profiles and through a k-fold cross-validation experiment.

The capacity of a spatial analysis to reproduce good estimates of precipitation at high elevation is critical. This point was particularly discussed in the context of snow modelling

and hydrology (Vidal et al. 2010; Tobin et al. 2011; Gottardi et al. 2012).

To implement the local analysis of vertical profiles, a suitable spatial scale has to be chosen.

It was decided to rely on a subdivision of the French Alps which has been applied for the Safran model of Météo-France for many years (Durand 1993). The French Alps are divided into 23 zones which are considered physically and climatologically homogeneous. The results for all the 23 French zones were produced but the conclusions for only 4 of them are presented here. The four most accurately analysed zones are Chablais (1385 km^2, mean elev. 1291 m), Belledonne (993 km^2, mean elev. 1121 m), Haute-Tarentaise (642 km^2, mean elev. 2332 m) and Queyras (843 km^2, mean elev. 2127 m). They are shown in Fig. 1, map of the study domain.
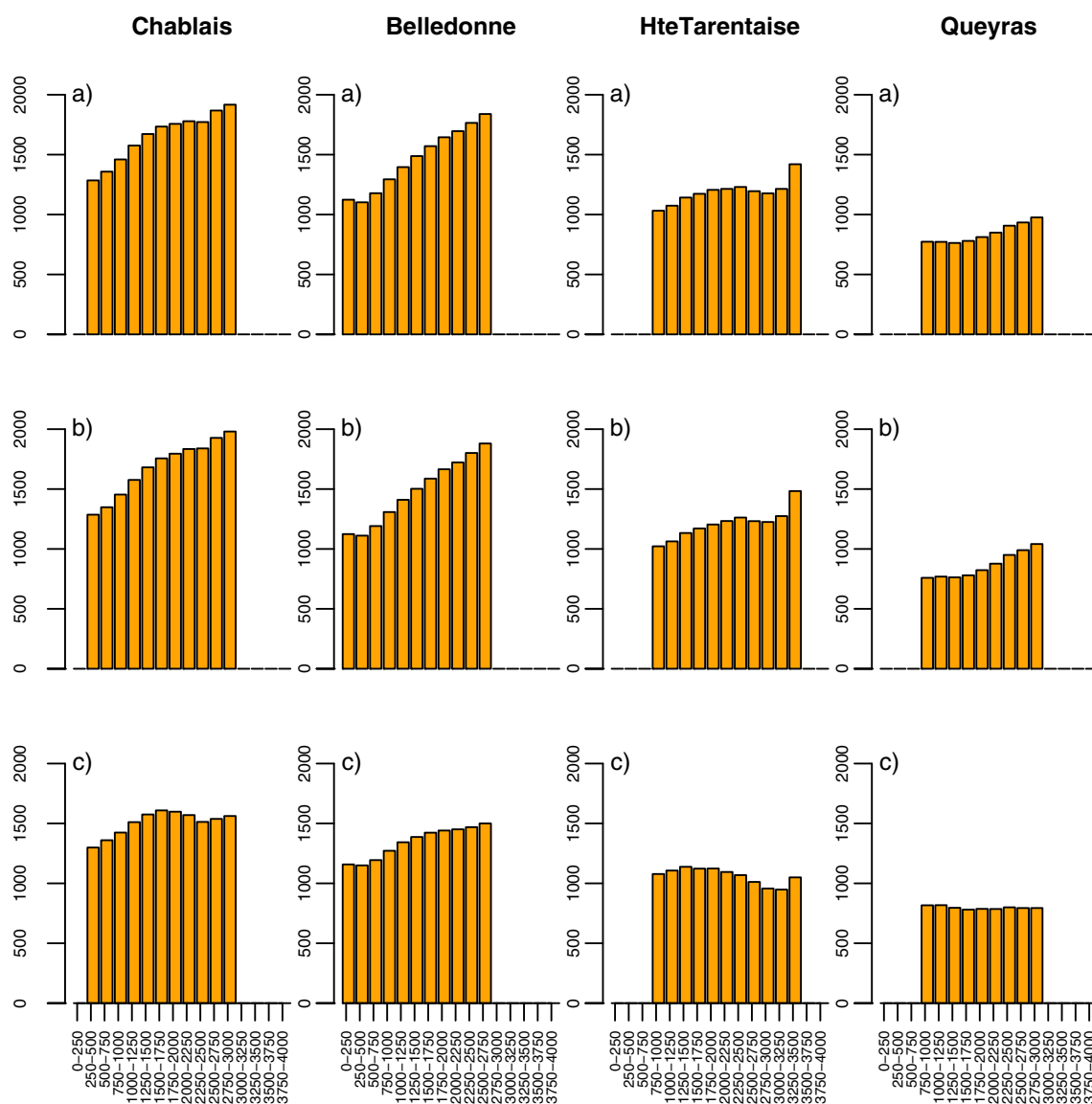

Fig. 9 Vertical profiles of precipitation (mm) for four zones by step of elevation of 250 m for three climatologies produced from daily fields: a CLIMDKEDWT: yearly mean from daily fields produced with KED with weather-type clim. b CLIMDKEDM yearly mean from daily fields produced with KED with monthly clim. c CLIMDOK: yearly mean from daily fields produced with ordinary kriging of observations only

Chablais and Belledonne are in forward position, fully exposed to the western and northwestern meteorological fluxes. Haute-Tarentaise and Queyras are more embedded in the alpine ridge with higher summits and mean elevation. Queyras, more southern, shows some influence of the Mediterranean climate.

For the implementation of a cross-validation of the climatological products, a leave-one-out cross-validation was not conceivable for computational reasons. Consequently, a k-fold technique is applied. Five sub-samples of 50 stations are randomly selected among the stations with less than 10 % missing values (each station is selected once and can be only in one sub-sample). Then, the different climatological products are calculated five times, with one of the validation samples put aside each time. It must be noted that 50 stations are about 9 % of the total number of stations available for the climatological fields. With such a low percentage, the density of stations is not seriously damaged. Finally, an independent estimate of the climatological value is available for 247 precipitation stations (for 3 stations, the process could not be applied). This final cross-validation showed some interesting results but also has certain limits: the sample is not very large and lacks high elevation stations.

Figure 8 shows the vertical profiles of precipitation for Chablais, Belledonne, Hte-Tarentaise and Queyras for WTCLIM and MCLIM. The steps of elevation are by 250 m. It can be seen through the histograms at the bottom of the figure that the four zones have very different distribution of elevation. The climatological conditions are very different for Chablais and Belledonne than for Hte-Tarentaise and Queyras with more precipitation for the first ones and a strong slope of precipitation with elevation, in agreement with the idea of more rain for the first slopes of a mountain ridge and less rain for the inner zones, though higher in elevation.

It is interesting to note that the amounts of precipitation for each step of elevation are very close for WTCLIM and for MCLIM.

Figure 9 shows the vertical profiles of precipitation for the same zones for CLIMDKEDWT, CLIMDKEDM and CLIMDOK. The profiles for CLIMDKEDWT and CLIMDKEDM are very similar and absolutely consistent with those presented before in Fig. 8. This is not the case for CLIMDOK. CLIMDOK is not able to reproduce realistic quantities of precipitation at high elevation. CLIMDOK is now disqualified compared to the other methods presented in this study.

Figure 10 shows the results for RMSE after cross-validation for 247 stations for MCLIM (blue), WTCLIM (red) and CLIMDKEDWT (yellow). Because of the small sample of stations, a bootstrap is applied to produce an estimate of the uncertainty. One thousand samples were drawn from re-sampling with replacement of the original 247 couples observation-estimation, and the RMSE is computed each time. The results are presented for all the stations but also by
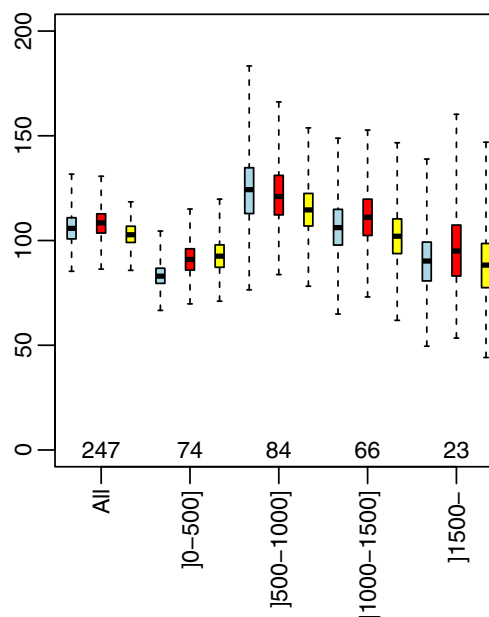

Fig. 10 Cross-validation for annual mean climatology 1990–2012 for precipitation. *Boxplot* of RMSE after bootstrap (for all the stations available and by step of elevation, with size of the sample). *Blue*: annual mean from monthly climatologies (MCLIM). *Red*: annual mean from weighting of weather type climatologies (WTCLIM). *Yellow*: annual mean from daily fields with weather-type climatologies as predictor (CLIMDKEDWT)

steps of elevation of 500 m. However, the number of stations above 1500 m is small. The sample size is given for each step of elevation.

This figure shows that the performance of the climatology realized from the daily fields (CLIMDKEDWT) is slightly higher than the two spatialization of long-term time series from stations (MCLIM and WTCLIM).

7 Discussion and conclusion

In this study, several different problems were analysed. These include the following:

1. The advantage of a weather-type classification as predictor for the spatialization of daily fields of precipitation in mountain terrain
2. The parameterization of the spatialization techniques
3. The advantage of producing climatology from the average of daily fields compared to climatology produced by the spatialization of climatological means of in situ stations time series
4. The difficulty to estimate and to validate precipitation at high elevation

The weather-type classification developed in this study highlights clearly different meteorological situations for the

study domain throughout the year. The annual climatology over the period of 1990–2012 produced from the weighting of the weather-type climatologies seems to be of good quality and shows realistic vertical gradients. Using weather-type climatology of the class of the day as predictor for the spatialization of daily fields has proven to be a good choice after a cross-validation against the observations. However, and this is probably one of the most important conclusions of the study, using weather-type climatology as predictor shows only a small advantage compared to using simple monthly climatology despite the efforts in this study to produce a classification with a methodology closely adapted to the needs.

The Aurelhy method remains an interesting technique for spatialization at the climatological time scale. The advantages of the parameters of the topography as predictors at this time scale are recognized. Aurelhy is able to estimate the most important topography predictors after the elevation for a particular region with the advantages of PCA (parameters sorted by importance, without correlation, noise reduced). The question of the best horizontal spatial scale to build the relation between the topography and the meteorological variable is an important one. However, the tests produced in this study may not be sufficient (only three possibilities were tested here for the Aurelhy landscapes). Hence, it is worth performing further tests in the future. With the new Aurelhy software, it is possible to combine two different spatial scales in the analysis of the topography.

Differences between the climatologies produced from long-term time series stations and the climatologies produced from accumulation or average of daily fields have been analysed in this paper. It seems that the latter take advantage of a better density of input data. For climatologies produced from long-term time series stations, the most important problem is that the density of stations without gaps is low. A compromise must be found between density and length of the series selected. Some techniques like a preliminary completion of the station gaps with daily estimates can improve the quality of the results. This was proven previously by Journée et al. (2015).

Conversely, working from daily fields to build climatology can be a response to the problem of stations with incomplete daily data. This can be considered as a kind of weighting of the stations according to the length of their presence. However, in this case, the density must be as stable as possible throughout the period of the study and quality control is an important preliminary step.

Estimating precipitation at high elevation, though analysed in many studies, is still a difficult question. Methods based on the climatological relation between orography parameters and the meteorological variables are a solution. In this study, four different methods analysed in Section 6 show very close climatological vertical gradients of precipitation with elevation.

These gradients seem realistic because they are strong enough and they comply with the regional variability expected. It is not possible to validate the precipitation estimation at high elevation through comparison with in situ measurements. Rain gauge data are too scarce at high elevation and with high uncertainty.

Consequently, we must look for more elements from other fields of investigation. These elements can be found in meteorological modelling, snow modelling and hydrology.

Meteorological models are now a promising field with the development of regional re-analysis. The dynamic methods for the downscaling of global coarse resolution re-analysis to regional and then high resolution non-hydrostatic models are available. The European project UERRA is expected to deliver such results for Europe (http://www.uerra.eu/). But today, the reference re-analysis for precipitation with a high resolution meteorological model over the Alps is still awaited.

The snow models are another source of indirect validation of the precipitation (and temperature) estimates. The most interesting information coming from the snow models is at the seasonal time scale. The difficulty is that precipitation and temperature are only two of the numerous input elements of snow modelling, and the complex evolution of snow along the winter season results in great uncertainty. This uncertainty is too large for an accurate validation of precipitation input. However, we can say that snow modelling has proven that long-term vertical gradient of precipitation must be strong. Very often, the input estimate of precipitation at high elevation is considered too low in regard with the snow depth (Durand et al. 1993; Gottardi et al. 2012).

Hydrology is a third indirect mean of validation but, as for snow models, with a lot of difficulties. The final runoff of rivers is the consequence of many different phenomena, each of which is accompanied by its own particular uncertainty in the models. Precipitation is only one of the contributions of the sum of uncertainties of the hydrological models. The problem is particularly challenging at high elevation because river basins are small and measurements are scarce and realized under difficult conditions (Tobin et al. 2011; Vidal et al. 2010).

The results presented in this study are limited to the period of 1990–2012 because of the lack of sufficient high density data covering the study domain for a longer period. But, we expect to produce 30-year climatology for 1981–2010 for the French side of the Alps. To this end, we need first to determine the weather type classes for the decade 1981–1990. This is possible despite the lack of in situ data covering all the domains for 1981–1990. The canonical variables for the latter period can be computed with only the ERA-Interim data covering 1981–1990 and the canonical vectors previously computed on 1990–2012. With this new data set, we plan to produce a comparison over a long period with two reference climatologies: the first one is the Safran re-analysis (Durand

et al. 2009) and the second one is the Alpine Precipitation Gridded Dataset (Isotta et al. 2013).

The results presented in this study are only for precipitation, but all the developments were also realized for minimum and maximum temperature. The results for temperature are similar to the ones for precipitation. The most suitable classification for temperatures was with four classes and four seasons.

References

Barnston AG, Glantz MH, He Y (1999) Predictive skill of statistical and dynamical climate models in SST forecast during the 1999-1998 el Nino episode the 1998 Nina onset. Bull Am Meteor Soc 80:217–243

Barry RG (2008) Mountain weather and climate, 3rd edn. Cambridge University Press, Cambridge, p. 506

Bartlett MS (1941) The statistical significance of canonical correlations. Biometrika 32(1):29–37

Beck C, Philipp A (2010) Evaluation and comparison of circulation type classifications for the European domain. Phys Chem Earth 35:374–387. doi:10.1016/j.pce.2010.01.001

Benichou P (1994) Cartography of statistical pluviometric fields with automatic allowance for topography. Global Precipitations and Climate Change NATO ASI Series 26(1994):187–199

Benichou P, Le Breton O (1987) Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. La Météorologie 7:23–34

Beniston M (2006) Mountain weather and climate : a general overview and a focus on climate change in the alps. Hydrobiologia 562:3–16. doi:10.1007/s10750-005-1802-0

Bretherton CS, Smith C, Wallace JM (1992) An inter-comparison of methods for finding coupled patterns in climate data. J Clim 5:541–560

Brunetti M, Maugeri M, Nanni T, Simolo C, Spinoni J (2013) High-resolution temperature climatology for Italy: interpolation method inter-comparison. Int J Climatol 34(4):1278–1296. doi:10.1002/joc.3764

COST-79 (1997) Proc. of the EU/COST-79 Seminar on Data Spatial Distribution in Meteorology and Climatology, Volterra, (28 Sep.-3 Oct. 1997), EU, Luxembourg,. EUR18472, 226 pp

Courault D, Monestiez P (1999) Spatial interpolation of air temperature according to atmospheric circulation patterns in south-east France. Int J Climatol 19:365–378. doi:10.1002/(SICI)1097-0088(19990330)19:4<365::AID-JOC369>3.0.CO;2-E

Cressie NA, Cassie NA (1993) Statistics for spatial data. Wiley, New-York

Daly C, Neilson RP, Phillips D (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J Appl Meteorol 33:140–158

Dee D et al (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. J R Meteorol Soc 137:553–597 . doi:10.1002/qj.828April 2011

DeGaetano AT, Wilks DS (2009) Radar-guided interpolation of climatological precipitation data. Int J Climatol 29(2):185–196. doi:10.1175/2007JAMC1536.1

Durand Y et al (1993) A meteorological estimation of relevant parameters for snow models. Ann Glaciol 18:65–71

Durand Y et al (2009) Reanalysis of 44 Yr of climate in the French alps (1958-2002): methodology, model validation, climatology and trends for air temperature and precipitation. J Of Appl Meteorol And Climatol 48:429–449. doi:10.1175/2008JAMC1808.1

Erdin R, Frei C, Kunsch HR (2012) Data transformation and uncertainty in geostatistic combination of radar and rain gauges. J Hydrometeorol 13:1332–1346. doi:10.1175/JHM-D-11-096.1

Esteban P, Ninyerola M, Prohom M (2008) Spatial modelling of air temperature and precipitation for Andorra (Pyrenees) from daily circulation patterns. Theor Appl Climatol 26:1501–1015. doi:10.1007/s00704-008-0035-3

Forgy EW (1965) Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics 21:768–769

Frei C, Schär C (1998) A precipitation climatology of the alps from high-resolution rain-gauges observations. Int J Climatol 18:873–900. doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9

Goodison B., Louie P., Yang D. (1998) Solid precipitation measurement inter-comparison. WMO Instrum Observing Methods. report no67

Goovaert P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York

Gottardi F, Obled C, Gailhard J, Paquet E (2012) Statistical reanalysis of precipitation fields based on ground network data and weather patterns: application over French mountains. J Hydrol 432:154–167. doi:10.1016/j.jhydrol.2012.02.014

Gyalistras D (2003) Development and validation of a high-resolution monthly gridded temperature and precipitation data set for Switzerland (1951-2000). Clim Res 25:55–83. doi:10.3354/cr025055

Hastie T, Tibshirani R, Friedman J (2009) Hierarchical clustering. The elements of statistical learning, 2nd edn. Springer, New York, pp. 520–528

Hengl T, Geuvelink G.B.M., Stein A. (2003) Comparison of kriging with external drift and regression kriging. Tech Note ITC. http://www.itc.nl/library/Academic_output/ Accessed 26 june 2015

Hess P. and Brezowsky H. (1952) Katalog der Großwetterlagen Europas (Catalog of the European 9 Large Scale Weather Types). Ber. Dt. Wetterd. in der US-Zone 33, Bad Kissingen, 10 Germany

Hofstra N, Haylock M, New M, Jones PD, Frei C (2008) Comparison of six methods for the interpolation of daily european climate data. J Geophys Res 113:D21110. doi:10.1029/2008JD010100

Hofstra N, New M, McSweeney C (2010) The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. Clim Dyn 35:841. doi:10.1007/s00382-009-0698-1

Hutchinson MF (1998) Interpolation of rainfall data with thin plates smoothing splines. Part I : two dimensional smoothing of data with short range correlation. Part II : analysis of topographic dependence. J Geogr Inf Decis Anal 2:139–151 and 152-167

Huth R, Beck C, Philipp A, Demuzere M, Untstrul Z, Cahynova M, Kysely J, Tveito OE (2008) Classification of atmospheric circulation patterns: recent advances and applications. Ann N Y Acad Sci 1146:105–152. doi:10.1196/annals.1446.019

Isotta FA, Frei C et al (2013) The climate of daily precipitation in the alps : development and analysis of high-resolution grid dataset from pan-alpine rain-gauges data. Int J Climatol 34:1657–1675. doi:10.1002/joc.3794

Journée M, Delvaux C, Bertrand C (2015) Precipitation climate maps of Belgium. Adv Sci Res 12(73–78):2015. doi:10.5194/asr-12-73-2015

Lamb HH (1972) British Isles weather types and a register of daily sequence of circulation patterns 1861–1971. Geophysical memoir, vol 116. HMSO, London, p. 85

Li J. and Heap A.D. (2008) A review of spatial interpolation methods for environmental scientists. *Geoscience Australia* Record 2008/23 137 pp ISBN 978 1 921498 28 2

MacQueen J. B. (1967) Some methods for classification and analysis of multivariate observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability 1. *University of California Press*. pp. 281–297

Masson D, Frei C (2014) Spatial analysis of precipitation in a high-mountain region: exploring methods with multi-scale topographic predictors and circulation types. Hydrol Earth Syst Sci Discussions 11:4639–4694. doi:10.5194/hessd-11-4639-2014

Matheron G. (1963) Traité de Géostatistique appliquée, *coll. Mémoires du Bureau de recherches géologiques et minières*. Technip, Paris (no 14, 24), 1962–1963, 2 volumes, 333–171 p

Mergili M, Kerschner H (2015) Gridded precipitation mapping in mountainous terrain combining GRASS and R. *Norvegian*. J Geogr 69(1): 2–17. doi:10.1080/0029151.2014.992807

Michelangeli PA, Vautard R, Legras B (1995) Weather regimes : recurrence and quasi-stationarity. J Atmos Sci 52:–1256. doi:10.1175 /1520-0469(1995)052<1237:WRRAQS>2.0.CO;2

Philipp A et al (2010) COST733CAT—a database of weather and circulation type classifications. Phys Chem Earth 35:360–373. doi:10.1016/j.pce.2009.12.010

Plaut G, Simonnet E (2001) Large-scale circulation classification, weather regimes, and local climate over France, the alps and Western Europe. *Clim*. Research 17:303–324

Prudhomme C, Reed DW (1998) Relationships between extreme daily precipitation and topography in a mountainous region : a case study in Scotland. Int J Climatol 18:1439–1453. doi:10.1002/(SICI)1097-0088(19981115)18:13<1439::AID-JOC320>3.0.CO;2-7

Ripley BD (1981) Spatial statistics. Wiley, New York

Rodwell MJ, Richardson DS, Hewson TD, Haiden T (2010) A new equitable score suitable for verifying precipitation in numerical weather prediction. Q J R Meteorol Soc 136:1344–1363

Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Comput Appl Math 20:53–65

Schwarb M. (2001) The Alpine precipitation climate: evaluation of a high- resolution analysis scheme using comprehensive rain-gauge data. Zucher Klima-Schriften, 80 Dissertation ETHZ 13911

Sevruk B (2005) Rainfall measurement: gauges. Encyclopedia of hydrological sciences. In: Anderson MG (ed) Part 2, hydrometeorology, chapter 40, vol. 1. Wiley&Sons Ltd, Chichester UK

Sevruk B, Ondras M, Chvila B (2009) The WMO precipitation measurement inter-comparisons. *Journal of*. Atmos Res 92(No. 3):376–380. doi:10.1016/j.atmosres.2009.01.016

Sherrer SC, Appenzeller C (2006) Swiss alpine snow pack variability : majors patterns and link to local climate and large scale flow. Clim Res 32:187–199

Simonnet E, Plaut G (2001) Space-time analysis of geopotential height and SLP, intra-seasonal oscillations, weather regimes, and local climates over the North Atlantic and Europe. Clim Research 17:325–342

Tobin C et al (2011) Improved interpolation of meteorological forcings for hydrologic applications in a Swiss alpine region. J Hydrol 401: 77–89. doi:10.1016/j.jhydrol.2011.02.10

Tveito OE (2007) Spatial distribution of winter temperatures in Norway related to topography and large-scale atmospheric circulation. Predictions in Ungauged basins kick-off meeting Brasilia 20-22 november 2002. IAHS Publ 309:2007

Vidal JP et al (2010) A 50-year high-resolution atmospheric reanalysis over France with Safran system. Int J of Climatol 30:1627–1644. doi:10.1002/joc.2003

Von Storch H, Zorita E, Cubasch U (1993) Downscaling of global climate change estimates to regional scales : an application to Iberian rainfall in wintertime. J Clim 6:1161–1171

Ward JH (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58(301):236–244

Whiteman CD (2000) Mountain meteorology: fundamentals and applications. Oxford University Press, Oxford, p. 355

Wong A (1980) A hybrid clustering algorithm for identifying high density clusters. J of American Statistical Ass 77 – 380:841–847

Wuest M. et al. (2009) A gridded hourly precipitation dataset for Switzerland using rain-gauge analysis and radar-based disaggregation. Int. J. Climatol. Royal Meteorological Society Published on line in Wiley InterScience, doi:10.1002/joc.2025