Adv. Sci. Res., 12, 73–78, 2015
www.adv-sci-res.net/12/73/2015/
doi:10.5194/asr-12-73-2015

Advances in Science & Research
Open Access Proceedings

14th EMS Annual Meeting & 10th European Conference on Applied Climatology (ECAC)

Precipitation climate maps of Belgium

M. Journée, C. Delvaux, and C. Bertrand

Royal Meteorological Institute of Belgium, Brussels, Belgium

Correspondence to: M. Journée (michel.journee@meteo.be)

Received: 18 November 2014 – Revised: 9 February 2015 – Accepted: 15 April 2015 – Published: 29 April 2015

Abstract. Investigations are conducted to best estimate precipitation climate maps over Belgium from daily observations available for the period 1981–2010. Several mapping approaches are compared in a cross-validation exercise. These approaches differ by several aspects and in particular by the order in which the temporal aggregation (i.e. computation of climate mean values from daily data) and spatial interpolation steps are performed, and by the integration of ancillary information in the spatial interpolation method. The selected approach is used to derive a large panel of climate maps. In particular, the main spatio-temporal features of the annual cycle of rainfall in Belgium are extracted by principal component analysis (PCA).

1 Introduction

The Royal Meteorological Institute of Belgium (RMI) has recently updated the climate maps of Belgium for various meteorological parameters. These maps represent the spatial distribution over Belgium of 30-years climate mean values based on the reference period 1981–2010, in accordance with the recommendation of the World Meteorological Organization (WMO, 2014). Such climate maps have already been derived for several regions of the world, e.g. The Netherlands (Sluiter, 2012), the Alps (Frei and Schär, 1998), the USA (Daly et al., 2008) and New Zealand (Tait and Zheng , 2007) to name a few examples.

The precipitation climate maps of Belgium rely on daily rain gauges observations from the Belgian network of climatological stations. In the past, such maps were derived manually on the basis of climate mean values derived from continuous time series of observations (Dupriez and Sneyers, 1979). Nowadays, climate maps are processed with reproducible methods without any subjective manual intervention (Daly, 2006).

This paper summarizes the investigations conducted to update the precipitation climate maps of Belgium. These investigations have two objectives: first, to determine the best method to generate climate maps from daily observations and second, to highlight the main characteristics of precipitation regimes along the year in Belgium. To reach the first objective, various methods have been considered and evaluated by cross-validation. For the second objective, a spatio-temporal analysis of rainfall by principal component analysis (PCA) was performed.

2 Materials and methods

2.1 Precipitation data

Monitoring of the Belgian climate is a key responsability of RMI, which implies the regular observation of several climate variables. RMI operates since the end of the 19th century a network of climatological stations providing observations of the daily precipitation quantities and the daily extreme air temperatures. Regarding precipitation, the number of these climatological stations has varied during the period 1981–2010 between 333 sites in 1981 and 247 sites in 1997 (see Fig. 1, left panel). The locations of the climatological stations (with at least 50 % of daily data available for the period 1981–2010) are illustrated in Fig. 2.

Although the climate maps are mainly derived from data acquired by the network of climatological stations, it is worth to mention that, since the begin of the 2000s, additional networks of automatic rain gauges have been deployed by the regional Belgian hydrological services. In particular, a network of 90 well-maintained and regularly calibrated tipping bucket rain gauges is operated by the Walloon hydrological services (WHS) since 2005 (see Fig. 2 for the location of these rain gauges). This data may be valuable ancillary information to improve the spatial representation of rainfall in

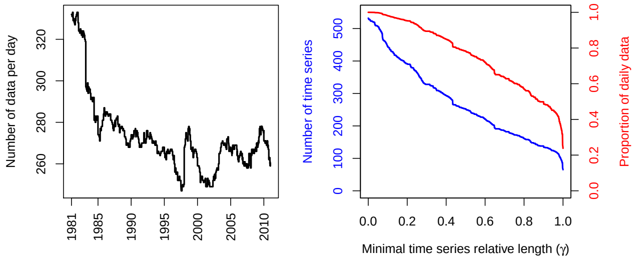

Figure 1. Evolution of the number of observations from the climatological stations available per day from 1981 to 2010 (left panel) and number of time series with a relative length greater or equal to a level γ (right panel, blue line and left axis). The corresponding fraction of daily observations is given by the red line (right axis).

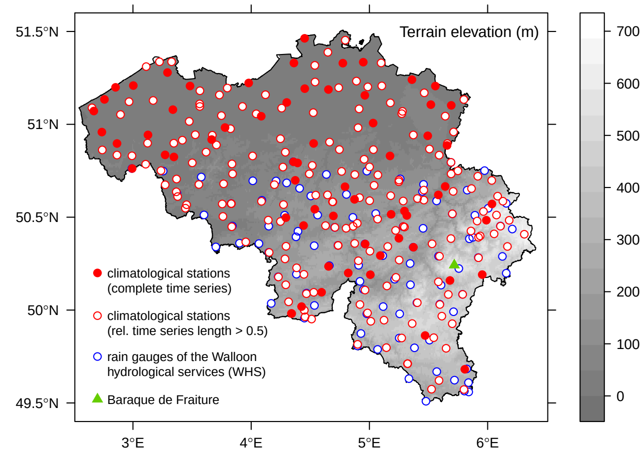

Figure 2. Elevation map of Belgium and spatial distribution of the climatological stations with a relative observations time series length larger or equal to 50 % (empty red circles) and 100 % (filled red circles) for 1981–2010 and the WHS rain gauges (blue circles). The green triangle indicates the location of "Baraque de Fraiture".

the region covered by the WHS network. This point will be discussed in the following.

All measurements made within the networks operated by RMI and WHS are routinely quality controlled by RMI.

2.2 Spatial interpolation methods

Climate maps typically result from the spatial interpolation of irregularly distributed station data to a regular high-resolution grid (Daly, 2006). The commonly used methods for the spatial analysis of climate data include inverse-distance weighting (IDW), thin plate smoothing splines, local regression models and geostatistical interpolation techniques (Sluiter, 2012). These methods can be divided into exact or approximate interpolation, depending on whether the interpolate equals or not the observation at the stations' locations. In this study, only exact interpolation methods are considered namely IDW and two geostatistical methods.

To derive climate maps from daily observations, two approaches are possible. First, the observations can be spatially interpolated to generate a set of daily precipitation grids, which are then temporally aggregated to provide the climate mean at each grid point (i.e. "interpolation-aggregation" approach). Second, climate mean values estimated from the observations time series can be spatially interpolated (i.e. "aggregation-interpolation" approach). These two approaches distinguish themselves by several aspects.

First, the "aggregation-interpolation" approach (denoted AI in the following) requires observations time series that are continuous on the entire 30-years period to allow a proper computation of the climate mean values, while the "interpolation-aggregation" approach (denoted IA in the following) can include all observations time series, even the short ones. As illustrated in Fig. 1 (right panel), only 64 time series are complete from 1981 to 2010, which represents 24 % of all daily data available for the 1981–2010 period. This drawback of the AI approach can be mitigated if one allows gaps in the time series used to compute the climate mean values. These gaps can be filled by estimations (obtained, for instance, by spatial interpolation of neighboring daily data) before temporal aggregation. This enables to significantly increase the number of involved time series, as illustrated in Fig. 1 (right panel). For instance, 138 time series are complete at 90 % and they represent half of the database of daily precipitation data.

Second, the spatial interpolation of climate mean values is generally less prone to uncertainties when compared to daily observations, as the spatial variability of precipitation fields reduces when the temporal aggregation horizon increases. Furthermore, since precipitation quantities are, on climate average, affected by physiographic features of the Earth's surface such as the terrain elevation (Goovaerts, 2000; Daly, 2006), the spatial interpolation of the climate mean values can be improved with the knowledge of densely sampled covariate data. As an example, Fig. 3 (left panel) illustrates the strong correlation between mean annual precipitation amount and terrain elevation.

Finally, the daily precipitation data acquired within the ancillary network of tipping bucket rain gauges operated by WHS can contribute to improve the representation of the climate patterns of rainfall in a region where they are the most complex. In the IA approach, these data can improve the estimation of the daily precipitation fields from 2005 to 2010, i.e. for only 6 years of the 30-years reference period. In case of the AI approach, a covariate of the 1981–2010 mean climate values can be derived if one assumes that the climate patterns of rainfall in Belgium can already be identified from a few years long time series. Such a covariate is defined by the spatial interpolation at high resolution of the 2005–2013 (9 years) mean values derived from the WHS data. This interpolation integrates terrain elevation data in a kriging with external drift approach (KED, Wackernagel, 1995). Figure 3

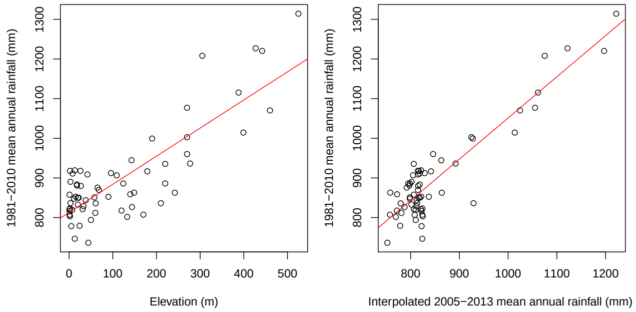

Figure 3. Scatter plots of the 1981–2010 mean annual rainfall computed for the 64 complete time series versus the terrain elevation (left panel) and the interpolated 2005–2013 mean annual rainfall based on the WHS rain gauges (right panel). The regression line is represented in red. The corresponding correlation coefficients are 0.825 and 0.915, respectively.

(right panel) illustrates the correlation between the 1981–2010 mean annual rainfall and the covariate based on the WHS network of rain gauges.

To sum up, the following methods to derive precipitation climate maps for Belgium are considered and evaluated in this study:

– Estimation of daily precipitation fields based on the maximum number of observations available each day, followed by temporal aggregation at each grid point (method "IA"). The spatial interpolation at the daily time scale is performed by inverse distance weighting (IDW) based on a maximum number of 20 neighbors located within a radius of 50 km.

– Completion of the gaps by IDW in all time series exhibiting a relative length equal or larger than a level $\gamma \in [0, 1]$, computation of the corresponding climate mean values and finally spatial interpolation. This spatial interpolation of climate mean values can be done by IDW, ordinary kriging (OK, Wackernagel, 1995), KED with the terrain elevation as drift or KED with the covariate based on the WHS network of rain gauges as drift. These methods will be respectively denoted by $\mathrm{AI_{IDW}}$, $\mathrm{AI_{OK}}$, $\mathrm{AI_{KED1}}$ and $\mathrm{AI_{KED2}}$. For the geostatistical methods, a variogram model is systematically fitted to the data. A zero nugget is imposed to estimate a precipitation field without any discontinuity at the stations' locations.

2.3 Validation methodology

The considered climate mapping methods are evaluated and compared by cross-validation. The validation relies on the $n = 64$ time series that are complete on the entire 1981–2010 period (filled red circles in Fig. 2). Because some of these stations are very close from each other (i.e. less than 5 km), the classical leave-one-out cross-validation (i.e. one station is systematically used to validate the estimation derived from the remaining stations' values) may lead to overly optimistic results. Therefore, a larger validation data set of size $k = 20$ was considered. The set of all daily observations is thus repeatedly splitted in separate training and validation data sets, with the validation data set being a random selection of k out of the n complete time series. The training data contains the remaining complete and incomplete time series. For the AI approaches, a condition on the minimal time series length (i.e. γ) is furthermore imposed. All methods are then applied on this training data set to estimate the mean annual and monthly precipitation quantities at the k validation locations. This process is restarted several times until each of the n complete time series has been used 10 times for validation. For each validation location $i \in (1, \ldots, n)$, the average estimation over the various reinitializations, $\hat{Y}_i$, is compared against the actual climate mean value, Y_i. The performance of the various methods is then summarized by the following indices:

– the mean bias error $\mathrm{MBE} = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)$,

– the mean absolute error $\mathrm{MAE} = \frac{1}{n} \sum_{i=1}^{n} |\hat{Y}_i - Y_i|$,

– the root mean square error $\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2}$ and

– the maximum absolute bias $\max_{i=1,\ldots,n} |\hat{Y}_i - Y_i|$.

MBE provides information about the average bias of the methods, while both MAE and RMSE characterize the average error magnitude. The maximum absolute bias indicates the largest deviation that is locally reached. These indices are assumed to provide sufficient evidence to select the best method.

2.4 Spatio-temporal analysis method

Principal component analysis (PCA) has been successfully used in several studies in order to improve the understanding of precipitation regimes at the scale of a country (Sneyers et al., 1989; Baeriswyl and Rebetez, 1997; Benzi et al., 1997). PCA is in general directly applied on stations's values, e.g. daily values (Benzi et al., 1997), monthly values (Baeriswyl and Rebetez, 1997) or monthly climate mean values (Sneyers et al., 1989). In contrast, in this study, PCA is used after spatial interpolation in order to more specifically highlight spatial patterns related to the precipitation regimes in Belgium. To reach that objective, we compiled a set of 365 maps of the mean rainfall per periods of 60-days in a sliding window approach by steps of one day. These maps were derived by the selected spatial interpolation method. This data set characterizes the spatial variability of the mean

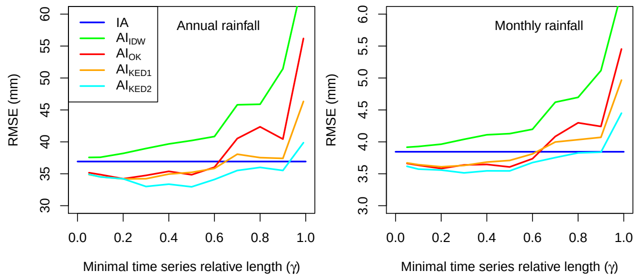

Figure 4. Cross-validation root mean square error (RMSE) for the considered mapping methods in the case of mean annual rainfall (left panel) and mean monthly rainfall (right panel). In case of the monthly rainfall, the results are averaged over the 12 months of the year. For the AI methods, the RMSE is illustrated as a function of the minimal time series length considered to define the training set (γ).

seasonal cycle of rainfall in Belgium. A PCA analysis of this data set was performed to extract its main spatio-temporal features. The PCA decomposition was evaluated after centering (i.e. substraction of the mean value for each map) in order to remove the spatial average of the annual cycle and highlight relative differences between different areas. The dominant components can highlight specific spatial patterns, while the corresponding loadings represent the seasonal variability of these patterns.

3 Results

3.1 Cross-validation analysis

The cross-validation exercise described in Sect. 2.3 enabled to evaluate the ability of the considered mapping methods to estimate the mean annual rainfall and the 12 mean monthly rainfall. Figure 4 provides the RMSE of the various methods for the mean annual rainfall and on average for the mean monthly rainfall (i.e. average over the 12 months of the year). Several observations can be made from these results.

First, the AI methods, and especially AI_{IDW} and AI_{OK}, are very sensitive to the minimal time series length γ. Globally, the performance of these methods increases when γ decreases, which highlights the importance of filling gaps in incomplete time series. The high sensitivity to γ between 0.9 and 1 is related to the large number of time series that are actually almost complete (see Fig. 1, right panel), i.e. 64 time series are complete (24 % of all daily data) while 138 time series are complete at 90 % (50 % of all daily data). Although the RMSE of AI_{IDW} continuously decreases towards the one of IA when γ goes to zero, for the other AI methods, there is an optimal γ below which the performance does improve anymore and even slightly degrades. In the case of AI_{KED1} and AI_{KED2}, this means that the ancillary data contribute more than short time series to the estimation of the precipitation fields.

Table 1. Cross-validation scores (in mm) for the considered mapping methods in case of the mean annual rainfall and the mean monthly rainfall (average over the 12 months of the year). The AI approaches rely on time series with at least 30 % of data available over the period 1981–2010 (i.e. $\gamma = 0.3$).

Mean annual rainfall				
	MBE	MAE	RMSE	max. abs. bias
IA	−2.30	31.20	36.90	100.00
AI_{IDW}	−3.00	33.20	39.00	98.10
AI_{OK}	−2.00	29.30	34.70	82.40
AI_{KED1}	−1.40	29.00	34.20	74.60
AI_{KED2}	−0.00	27.70	33.00	70.90
Mean monthly rainfall				
	MBE	MAE	RMSE	max. abs. bias
IA	−0.20	3.10	3.80	13.00
AI_{IDW}	−0.30	3.30	4.00	12.80
AI_{OK}	−0.20	2.90	3.60	13.60
AI_{KED1}	−0.10	2.90	3.60	12.70
AI_{KED2}	−0.00	2.80	3.50	11.70

Second, the geostatistical methods (AI_{OK}, AI_{KED1} and AI_{KED2}) outperform the simple IDW method (AI_{IDW}). There is furthermore a clear benefit in using ancillary data (topography and WHS data). Overall, the best performance is obtained with AI_{KED2} and γ between 0.3 and 0.5. This observation is confirmed by the other scores in Table 1. In particular, the maximum absolute bias, which provides an idea of the maximum error that can be locally reached, is significantly reduced by AI_{KED2} in the case of the mean annual rainfall. The sensitivity with respect to the method is however higher in the annual case than in the monthly one.

3.2 Qualitative validation

In parallel to this cross-validation analysis, a qualitative evaluation was performed by comparing the spatial distribution of the 1981–2010 average spatial distribution of rainfall estimated by the various methods against the climate maps related to the period 1833–1975 previously published by Dupriez and Sneyers (1979). As expected, the climate patterns are globally similar between the new and the former climate maps. However, several small-scale differences can be highlighted, which are related to the evolution of the climatological stations network. As an example, the 1833–1975 climate map indicates a local increase of the mean annual rainfall above 1200 mm per year in the area of the location called "Baraque de Fraiture" (indicated by the green triangle in Fig. 2). This local pattern is however not reproduced by the methods IA, AI_{IDW} and AI_{OK} because no station has been in operation in this area during a significant period between 1981 and 2010. Nevertheless, it can be estimated by

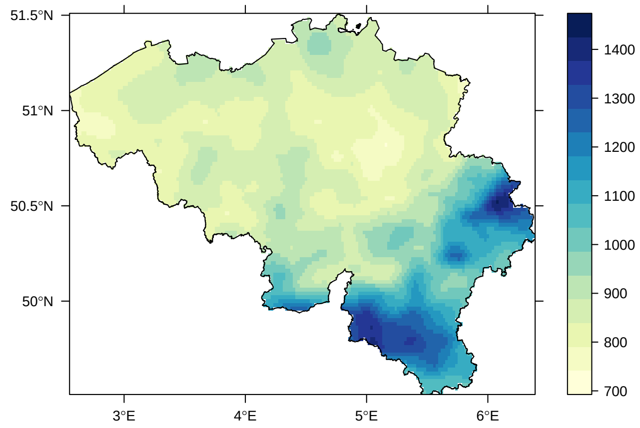

Figure 5. Map of the mean annual rainfall in Belgium for the period 1981–2010 (in mm).

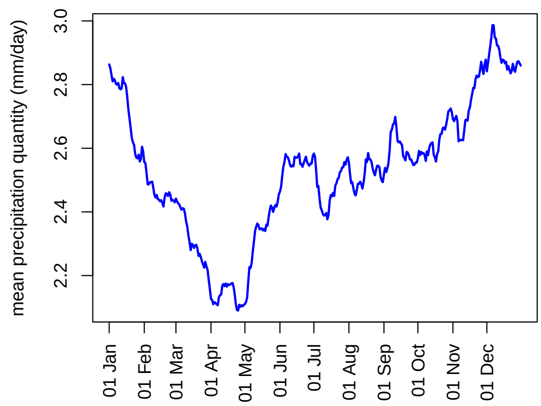

Figure 6. Annual cycle of the Belgian spatial average of the 60-days mean rainfall.

the method AI_{KED1} thanks to the correlation with the topography (Baraque de Fraiture is located at 652 m a.s.l. (above sea level) and is the third highest point of Belgium). The same is true for AI_{KED2}, which integrates topographic data as well as measurements from a rain gauge operated by WHS since 2005 in Baraque de Fraiture. This comparative analysis against the 1833–1975 climate maps favors thus the methods that integrate ancillary information to estimate the spatial distribution of the mean annual rainfall.

4 Discussion

The results presented in the previous section indicates that the best mapping performance for both the annual and the monthly average precipitation quantities is obtained with the method AI_{KED2} based on all time series with a minimal relative length γ between 0.3 and 0.5. This method with $\gamma = 0.3$ is therefore selected to derive a large panel of 1981–2010 climate maps of Belgium related to precipitation: maps of the mean rainfall per year (Fig. 5), season and month, as well as maps of related climate indices such as the average number of days per year, season or month with daily precipitation accumulation exceeding a given threshold (e.g. 1 and 10 mm).

As illustrated in Fig. 5, the mean annual rainfall varies in Belgium from 700 to 1400 mm. The largest values are observed in the east and south parts of Belgium. As expected, a local maximum is present in the area of Baraque de Fraiture. Figure 6, which represents the Belgian spatial average computed for each of the 60-days mean rainfall map derived in a sliding window approach, provides an idea of the average seasonal cycle of rainfall in Belgium. The mean rainfall reaches its minimum in April and its maximum in December. The amplitude of this annual cycle is of the order of 14 % with respect to its mean value.

Finally, the three dominant components resulting from the PCA decomposition of the set of 60-days mean rainfall maps are illustrated in Fig. 7 with the corresponding loadings (normalized to unit norm vectors) and fractions of explained variance. These results have to be considered in a relative manner only, both spatially and temporally, e.g. locations and periods of time can be compared against each other but the absolute level can not be directly interpreted. The first principal component indicates that the southern part of Belgium receives more precipitation in winter and less in summer when compared to the rest of the country. Similarly, the North-West (coastal area) is subject to lower precipitation quantities in spring and larger ones in autumn, as highlighted by the second principal component. Since the first two components account for more than 90 % of the variability, the interpretation of the other components is less significant.

5 Conclusions

In order to update the precipitation climate maps of Belgium with respect to the reference period 1981–2010, a study was conducted to evaluate several mapping approaches. The "interpolation–aggregation" method is able to integrate the entire database of daily observations but requires a spatial interpolation at the daily time scale. On the other hand, the "aggregation-interpolation" approaches perform a spatial interpolation of climate mean values, which is less prone to uncertainties and can furthermore integrate ancillary information, such as terrain elevation data. The benefit of these ancillary topographic information has been highlighted in the study. The observations made within the ancillary network of rain gauges operated by WHS furthermore improves the mapping performance, although this network is operational since 2005 only. However, the integration of incomplete time series has the strongest impact on the performance of the "aggregation-interpolation" methods. The approach selected to derive the precipitation climate maps of Belgium is thus a geostatistical "aggregation-interpolation" method involving all time series complete at 30 % with ancillary terrain elevation and WHS data. In addition to the classical maps about

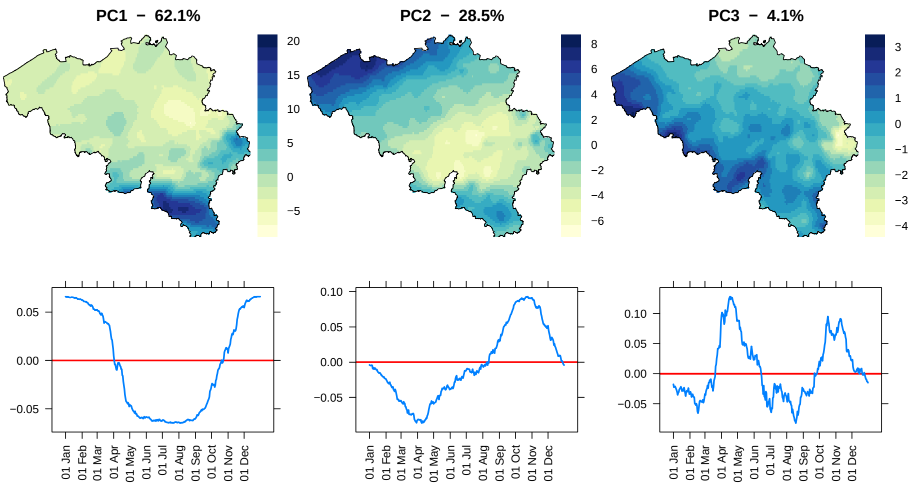

Figure 7. Illustation of the three dominant principal components of the set of 60-days mean rainfall maps with the corresponding fractions of explained variance and loadings normalized to unit norm vectors.

mean rainfall and average number of precipitation days, a set of maps charaterizing the spatial variability of the mean seasonal cycle of rainfall in Belgium was compiled and analyzed by PCA to highlight the main spatio-temporal patterns of the precipitation regimes in Belgium.

Acknowledgements. The present study was conducted in R (3.1.2) and the gstat package was used for geostatistical interpolation (Pebesma, 2004).

Edited by: M. Dolinar
Reviewed by: two anonymous referees

References

Baeriswyl, P. A. and Rebetez, M.: Regionalization of precipitation in Switzerland by means of principal component analysis, Theor. Appl. Climatol., 58, 31–41, 1997.

Benzi, R., Deidda, R., and Marrocu, M.: Characterization of temperature and precipitation fields over Sardinia with principal component analysis and singular spectrum analysis, Int. J. Climatol., 17, 1231–1262, 1997.

Daly, C.: Guidelines for assessing the suitability of spatial climate data sets, Int. J. Climatol., 26, 707–721, 2006.

Daly, C., Halbleib, M., Smith, J., Gibson, W., Doggett, M., Taylor, G., Curtis, J., and Pasteris, P.: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, Int. J. Climatol., 28, 2031–2064, 2008.

Dupriez, G. L. and Sneyers, R.: Les nouvelles cartes pluviométriques de la Belgique, Publications de l'Institut Royal Météorologique de Belgique, Brussels, Belgium, 1979.

Frei, C. and Schär, C.: A precipitation climatology of the Alps from high-resolution rain-gauge observations, Int. J. Climatol., 18, 873–900, 1998.

Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, J. Hydrol., 228, 113–129, 2000.

Pebesma, E.J.: Multivariable geostatistics in S: the gstat package, Comput. Geosci., 30, 683–691, 2004.

Sluiter, R.: Interpolation Methods for the Climate Atlas, KNMI technical rapport TR–335, Royal Netherlands Meteorological Institute, De Bilt, 1–71, 2012.

Sneyers, R., Vandiepenbeeck, M., and Vanlierde, R.: Principal component analysis of Belgian rainfall, Theor. Appl. Climatol., 39, 199–204, 1989.

Tait, A. and Zheng, X.: Analysis of the Spatial Interpolation Error associated with Maps of Median Annual Climate Variables, National Institute of Water & Atmospheric Research (NIWA), Wellington, New Zealand, 1–21, 2007.

Wackernagel, H.: Multivariate geostatistics: an introduction with applications, Springer-Verlag, Berlin, 1995.

WMO – World Meteorological Organization: CCI-16 Session report, WMO No. 1137, Geneva, Switzerland, 2014.