# Spatial analysis of precipitation in a high-mountain region:

# Exploring methods with multi-scale topographic predictors and circulation types

David Masson and Christoph Frei

MeteoSwiss

Federal Office of Meteorology and Climatology

Kraehbuehlstrasse 58, CH-8044 Zurich,Switzerland

Correspondence:  massond@phys.ethz.ch

# ABSTRACT

Statistical models of the relationship between precipitation and topography are key elements for the spatial interpolation of rain-gauge measurements in high-mountain regions. This study investigates several extensions of the classical precipitation-height model in a direct comparison and within two popular interpolation frameworks, namely linear regression and kriging with external drift. The models studied include predictors of topographic height and slope, eventually at several spatial scales, a stratification by types of a circulation classification, and a predictor for wind-aligned topographic gradients. The benefit of the modeling components is investigated for the interpolation of seasonal mean and daily precipitation using leave-one-out crossvalidation. The study domain is a north-south cross-section of the European Alps ($154 \times 187\ km^2$), which disposes of dense rain-gauge measurements (approx. 440 stations, 1971-2008).

The significance of the topographic predictors was found to strongly depend on the interpolation framework. In linear regression, predictors of slope and at multiple scales reduce interpolation errors substantially. But with as many as nine predictors the resulting interpolation still poorly replicates the across-ridge variation of climatological mean precipitation. Kriging with external drift (KED) leads to much smaller interpolation errors than linear regression. But this is achieved with a single predictor (local topographic height) already, whereas the incorporation of more extended predictor sets brings only marginal further improvement. Furthermore, the stratification by circulation types and the wind-aligned gradient predictor do not improve over the single predictor KED model. As for daily precipitation, the interpolation accuracy improves considerably with KED and the use of a single predictor field (the distribution of seasonal mean precipitation) as compared to ordinary kriging (i.e. without predictor at all). But, again, information from circulation types did not improve interpolation accuracy.

Our results confirm that the consideration of topography effects is important for spatial interpolation of precipitation in high-mountain regions. But a single predictor may be sufficient and taking appropriate account of the spatial autocorrelation (by kriging) can be more effective than the development of elaborate predictor sets within a regression model. Our results also question a popular practice of using linear regression for predictor selection in spatial interpolation. But they support the common practice of using a climatological mean field as a background in the interpolation of daily precipitation.

# 1. INTRODUCTION

High-mountain ranges contribute to the supply and storage of freshwater and river flow in many regions of the world (e.g., Viviroli et al., 2007). The role of mountains in extracting moisture from the atmosphere manifests in numerous regional anomalies and gradients in the distribution of the global precipitation climate (e.g., Basist et al., 1994; Schneider et al., 2013). Accurate knowledge of the distribution and variation of rain and snowfall is crucial for numerous planning tasks concerned, for example, with water resources, water power, agriculture, glaciology and natural hazards (e.g., Greminger, 2003; Holzkämper et al., 2012; Machguth et al., 2009; Yates et al., 2009). A convenient source of information are spatial analyses of observed precipitation, obtained by interpolation onto a regular grid, comprehensively over large areas. Such grid datasets have become of interest also for monitoring climate variations and for evaluating model-based re-analyses and climate models (e.g. (Alexander et al., 2006; Bukovsky and Karoly, 2007; Frei et al., 2003; Schmidli et al., 2002).

The construction of accurate precipitation grid datasets for high-mountain regions is confronted with the challenge of complex spatial variations. Even with idealized topographic settings and flow configurations (e.g. isolated hill or ridge, constant flow), situations can be distinguished where precipitation maxima occur over the windward slope, over the crest or the downwind slope of a topographic obstacle (e.g., Sinclair et al., 1997; Smith, 1979). Distributions depend on the height and scale of the obstacle, and the strength, static stability and moisture profile of the impinging flow. More complex topographic shapes, transient weather systems, convection and the drift of hydrometeors quickly complicate the picture (e.g., Cosma et al., 2002; Fuhrer and Schär, 2005; Houze et al., 2001; Roe, 2005; Sinclair et al., 1997; Steiner et al., 2003). Therefore, the distribution of long-term mean precipitation is, in many regions, a superposition of several distinct responses to topography, which act at different space scales, involve several characteristics of the topography (not just height) and pertain to different flow situations.

A further complication for spatial analysis in mountain regions is posed by the limited spatial density of rain gauges, the standard device for climatological inference on precipitation. Even in comparatively densely instrumented areas, such as the European Alps, the networks do not resolve contrasts between individual valleys and hills explicitly, and they miss out episodic fine-scale patterns familiar from radar observations and numerical models (e.g., Bergeron, 1961; Frei and Schär, 1998; Germann and Joss, 2001; Zangl et al., 2008). Moreover, the distribution of rain gauges in complex terrain is often biased, with a majority of measurements taken at valley floors, while steep slopes and high elevations are underrepresented (e.g., Frei and Schär, 1998; Sevruk, 1997). The sampling bias entails a risk of systematic errors in spatial interpolation, which can impinge upon estimates at larger scale, such as for averages over river catchments (e.g., Daly et al., 1994; Sinclair et al., 1997).

In this context, models of the relationship between precipitation and topography constitute an essential element of spatial interpolation methods. Their purpose is to enhance the methods' capabilities in describing variations not explicitly resolved by the observations, and to reduce the risk of systematic errors related to the non-representativity of the measurement network. Approaches for considering precipitation topography relationships in interpolation methods can roughly be grouped into *empirical statistical models* using more or less extensive sets of physiographic predictors (e.g., Benichou and Le Breton, 1986; Daly et al., 1994; Prudhomme and Reed, 1998) and simplified *physico-dynamical downscaling models* in combination with information on larger-scale circulation (e.g., Crochet et al., 2007; Sinclair, 1994).

In this study we explore and compare several ideas for the modeling of precipitation-topography relationships in the framework of empirical statistical models. Our specific focus is on models that (a) take account of the multi-scale nature of the relationship, (b) consider responses both to slope and elevation of the topography, (c) involve a dependency on the direction of the large-scale flow, and (d) examine the potential of a stratification by circulation types. The value of the different modeling components is assessed in terms of the skill of a geostatistical interpolation method, which has these models incorporated and is applied for the estimation of fields of seasonal mean and daily precipitation in a sub-region of the European Alps.

Systematic topography effects on precipitation are usually difficult to discern in observations at short time scales (e.g. for daily totals). Precipitation topography relationships are therefore mostly estimated from long-term averages, which are then used, via a climatological background field, for the interpolation of shorter duration totals (Haylock et al., 2008; Rauthe et al., 2013; Widmann and Bretherton, 2000).

A common model of topography effects is that of a linear relationship between climatological (seasonal or monthly) mean precipitation and in-situ topographic elevation. Precipitation-height gradients have been considered in various interpolation methodologies such as in linear regression by using height as a predictor (e.g., Gottardi et al., 2012; Rauthe et al., 2013; Sokol and Bližnák, 2009) in several variants of kriging by using a digital elevation model as secondary variable (Allamano et al., 2009; Goovaerts, 2000; Hevesi et al., 1992; Phillips et al., 1992; Tobin et al., 2011), in thin-plate splines interpolation by using height as a third regionalization variable (Haylock et al., 2008; Hutchinson, 1998) or in triangular interpolation by adopting height corrections  (Tveito et al., 2005). The assumption of these procedures is that local height is a key explanatory variable of the distribution of precipitation and that the relationship, commonly estimated over larger domains, is representative at the scale relevant for the interpolation, i.e. at and below the spacing of stations.

Three types of extensions of the aforementioned methodologies have been proposed: the first introduced a range of physiographic predictors (not just height) and/or predictors representing smoothed versions of the actual topography (e.g., Basist et al., 1994; Benichou and Le Breton, 1986; Gyalistras, 2003; Perry and Hollis, 2005; Prudhomme and Reed, 1998; Sharples et al., 2005). Additional predictors (e.g. slope, exposure) were found to significantly increase the explained variance compared to height only (e.g., Gyalistras, 2003; Prudhomme and Reed, 1998) and digital elevation models smoothed to resolutions of 5 to 50 kilometers (depending on region) were found to be more powerful predictors compared to high-resolution topography (e.g., Prudhomme and Reed, 1998; Sharples et al., 2005). Conversely, the second extension remains with univariate height dependencies, but considers the relationship to be spatially variable (Brunetti et al., 2012; Daly et al., 1994; Gottardi et al., 2012). The aim is to focus on dependencies at scales that are not explicitly resolved by the station network and, hence, are particularly relevant for interpolation. There are different emphases in the two extensions between robustness and local representativity of the precipitation-topography model used for interpolation.

The third type of extending traditional precipitation-height models is to incorporate information on atmospheric flow conditions into the interpolation: Kyriakidis et al. (2001) have constructed new rainfall predictors by combination of lower-atmosphere flow and moisture with local terrain height and slope. When used in kriging these dynamical predictors yielded more accurate interpolations of the seasonal mean precipitation compared to using elevation only. Hewitson and Crane (2005) have modified the weighting scheme of a daily interpolation method to depend on synoptic state (discrete types of daily low-level circulation) in order to account for the varying short-range representativity of station measurements. Gottardi et al. (2012) use the circulation regime of the day under consideration to estimate orographic effects specifically for different weather conditions. All these ideas are building on empirical evidence that the mesoscale precipitation distribution in complex terrain varies considerably between days with different large-scale flow conditions (Cortesi et al., 2013; Schiemann and Frei, 2010).

In this study we build on, extend and test ideas of all three extensions in a subregion of the European Alps. We compare several sets of physiographic predictors with regard to their relevance for high-resolution precipitation interpolation. Apart from including height and directional gradients, our set encompasses predictors at several spatial scales simultaneously in order to explicitly distinguish between patterns resolved and unresolved by the station network. We also compare the role of predictor setting between multivariate linear regression and kriging with external drift, to assess how a model of spatial autocorrelation (kriging) can compensate for extensive predictor sets. We further examine the prospect of stratifying seasonal means by independent analyses for composites of a circulation type classification and by including predictors of the pertinent circulation terrain effect. Most of our analyses focus on interpolations for seasonal mean precipitation, but we also assess the relevance of circulation-type dependent background fields for the interpolation of daily precipitation. Essential for all our comparisons is that interpolation errors will be examined as a function of topographic height and for both systematic and random error components. The main purpose of our study is to gain insight on the role of different approaches to precipitation-topography modelling, but some of our analyses also explore possibilities to improve an interpolation method previously developed for the generation of a precipitation grid dataset for the entire Alpine region (Isotta et al., 2013).

The region of the European Alps is an interesting example for studying interpolation procedures and pertinent models of the precipitation topography relationship. There is an exceptional density of long-term rain-gauge observations (see Fig. 1), which allows modeling approaches of larger complexity than in sparsely gauged mountain regions. Moreover, there is a broad range of topographic scales (from hundreds of kilometers for the main ridge down to few kilometers for individual massifs) and variations in ridge height (2000-3000 meters for the main ridge down to few hundred meters for adjacent hill ranges). Accordingly, the distribution of mean precipitation reveals several nested patterns of the precipitation response that is indicative of its multi-scale nature (see Fig. 1).

This study is part of the European project EURO4M (European Reanalysis and Observations for Monitoring). The outline of the study is organized as follows: in Section 2 we introduce the study domain and the data. The methods of spatial analysis and the procedure of evaluation are described in Section 3. The results of the evaluation are then presented and discussed in Section 4 and the conclusions of this study are drawn in section 5.

## 2. STUDY DOMAIN AND DATA

In this study we consider a sub-domain of the Alps (11°E-13°E / 46.85°N-48.5°N) that covers an area of $154 \times 187\ km^2$ and extends from the flatlands of Bavaria (Southern Germany), over the Northern slopes of the Alpine ridge (at the country border between Germany and Austria) towards the inner Alpine region of Tyrol (Inn and Salzach valleys, Austria and Northern Italy). The domain is indicated in Fig.1 (red frame) and a detailed topographic map is depicted in Fig.2a. Our choice is motivated by the comparatively simple large-scale pattern of the topography here, so that the domain can be considered as a cross-section through an elongated west-east oriented ridge, extending from flatlands over foothills to high mountains with major inner mountain valleys (from North to South). As opposed to a larger domain with more convoluted topography, the intermediate complexity eases the exploration of potential physiographic predictors but still comprises the challenges encountered with distinct and typical climates of the entire Alpine ridge. In addition, the selected domain disposes of a homogenous and, compared to other regions, very dense coverage with rain gauges (cf. Fig.1).

The rain-gauge data for this study (Fig.2a) was obtained from the German Weather Service (DWD, for Germany), from the Austrian Federal Ministry of Agriculture, Forestry, Environment and Water (for Austria) and from Servizio Meteorologico and Ufficio Idrografico Bolzano Alto Adige (for Italy). The dataset is a subset of 440 stations out of a pan-Alpine compilation of high-resolution daily rain-gauge time series extending over the period 1971-2008 (Isotta et al., 2013). On average the station density is 1 station per 70 km$^2$ corresponding to a typical inter-station distance of 8.5 km, a very dense coverage over a high-mountain region.

Like in other mountainous regions, the distribution of the stations in our study domain has a limited representativity with respect to terrain height (Fig.2b). High-elevation areas (>1500 mMSL) are significantly underrepresented. For example, elevations above 1500 mMSL contribute about 25% of the total area but are represented with only 6% of the stations. This setting involves a risk of precipitation estimates for high-elevation areas being biases due to inappropriate interpolation between valley stations. This will be given particular attention in the assessment of interpolation methods later.

The rain-gauge time series underwent different quality control procedures at the original data providers. In addition they were rigorously checked for raw errors, jointly after compilation, using criteria of temporal and spatial consistency and physical plausibility (for details see Isotta et al., 2013). One caveat of the quality of the data is, however, posed by the systematic measurement error emanating from wind-induced under-catch, wetting and evaporation losses (Groisman and Legates, 1994; Neff, 1977; Sevruk, 2005). Sevruk (1985) and Richter (1995) estimate the systematic measurement error in the Alps to range from about 7% (5%) over the flatland regions in winter (summer) to 30% (10%) above 1500 mMSL. The data used in this study is not corrected for these systematic errors. Indeed, water balance considerations in the Alps have challenged existing correction procedures (Schädler and Weingartner, 2002; Weingartner et al., 2007). The systematic errors may affect the strength and estimation of empirical precipitation-topography relationships. However, given that the spatial variability of mean precipitation across the domain (see the example in Fig.2a) is much larger than the range of expected systematic errors, we assume that these errors are not significantly affecting the conclusions of the present study.

Our statistical analyses are conducted with estimates of mean precipitation at the above stations, that is, with seasonal means over a multi-year period or with means over all days belonging to the same class of a daily circulation type classification. The fact that many rain-gauge series extend over a part of the full 38-year period only, requires care in establishing robust and comparable mean values. For this purpose quantitative tests have been carried out, aiming at determining the minimum number of days required to build a mean value of a given accuracy. The tests were conducted by bootstrap experiments (sampling across days) over the time series of the 20 most complete station records. The error metric is based on the relative mean root transformed error presented in the evaluation section. Our accuracy requirement was that the probability of a sampling error larger than 10% of the “full” mean (i.e. mean over the complete time series) should be smaller than 5%.  The error thresholds are somewhat arbitrary but are chosen to guarantee reliable climatic estimates compared to the spatial variations while retaining enough data.  The resulting minimum requirement on the available length of the time series varies between season and circulation class. Stations not fulfilling this minimum requirement are discarded from the analysis. As a result the station sample varies between analyses with different seasons and between seasonal and circulation-type stratifications. Typically, the selection procedure eliminates 5 to 15 % of the total number of stations, leaving between 317 and 420 time series, depending on stratification.

The circulation type classification chosen in this study is the PCACA classification (Philipp et al., 2010; Yarnal, 1993). It uses daily mean sea level pressure distributions as input for a hierarchical cluster analysis of principal components. The classification catalog used here was taken from an application of PCACA in the framework of COST-Action 733 over an extended Alpine domain, using sea level pressure fields from ERA40 and ERA-Interim (Dee et al., 2011; Uppala et al., 2005) and with a target number of 9 clusters (Weusthoff, 2011). The choice of the 9-types classification (PCACA9) is a compromise between differentiation of daily circulation patterns and robustness of mean values (i.e. enough days within a weather class). In a comprehensive intercomparison, PCACA9 was found to be particularly skillful in explaining the distribution of mesoscale daily precipitation in the Alpine region (Schiemann and Frei, 2010). The geostrophic wind fields for each of the clusters were calculated from sea level pressure composites based on ERA40 (Uppala et al., 2005).

# 3. METHODS AND EXPERIMENTS

Our study on the significance and utility of physiographic predictors for spatial interpolation is, in the first instance, dealing with seasonal mean precipitation, where topographic effects on the distribution are standing out more clearly from spatial variations of episodic nature. The methodological framework employed is that of kriging with external drift (KED, Schabenberger and Gotway, 2005), an interpolation model with a component for multi-linear dependence on pre-defined variables (external drift or trend, here a set of topographic predictors) and a component of spatial autocorrelation. Two limiting cases of KED will also be considered for comparison: multi-linear regression models (LM), which comprise the linear dependence on topographic predictors only (i.e. no spatial auto-correlation) and ordinary kriging (OK) with only the spatial autocorrelation component included (i.e. omitting dependence on predictors). As topographic predictors, a set of candidates will be considered, including elevation ('e'), gradients ('g') in two cardinal directions (across and along the main ridge), as well as the gradient in the direction of the geostrophic wind of circulation types ('v'). Various spatial scales of these predictors are considered, in combination, representing variations of the topography at and beyond scales of 1 km, 5 km, 10 km, 25 km and 75 km, respectively. The different method settings and predictor sets will be compared by means of leave-one-out cross-validation, examining statistics of the systematic and random errors of the interpolation and their dependence on elevation.

In a second step we will compare the quality of daily precipitation interpolations when using various climatologies (with different predictor sets, seasonal or circulation type stratification) as a background reference (Widmann and Bretherton, 2000). As in the seasonal experiments, KED will provide the methodological framework for the daily interpolation, but using the previously determined background reference fields as trend variables.

The following subsections describe in detail the methodological setup (section 3a), the derivation and usage of the topographic predictor sets (section 3b), the method for daily interpolation (section 3c) and the cross-validation procedure (section 3d). Table 1 lists the experiments conducted for seasonal precipitation with the different methods and predictor sets, using the acronyms just introduced. The experiments conducted for daily interpolation are listed in Table 2.

### 3.1. Interpolation methods

For the interpolation concept, the present study builds on kriging with external drift (KED, Schabenberger and Gotway, 2005) and two simplified limit cases of it. KED belongs to a broad class of geostatistical interpolation methods, which estimate values at target locations as the best linear unbiased combination of sample observations, under the assumption that the field of interest is a realization of a second order stationary Gaussian process (see e.g., Cressie, 1993; Diggle and Ribeiro, 2007). KED considers the observations Y at sample locations $\mathbf{s}$ as a random variable of the form (see e.g., Diggle and Ribeiro, 2007):

$$Y(\boldsymbol{s}) = \mu(\boldsymbol{s}) + Z(\boldsymbol{s}) \quad , \quad \mu(\boldsymbol{s}) = \beta_0 + \sum_{k=1}^{K} \beta_k \cdot x_k(\boldsymbol{s}) \qquad (1)$$

Here, $\mu(\mathbf{s})$ describes the deterministic component of the model (also termed *external drift* or *trend*), and is given as a linear combination of $K$ predictor fields $x_k(\mathbf{s})$ (also termed *trend variables*) plus an intercept $\beta_0$. The $\beta_k$ are denoted as trend coefficients. $Z(\mathbf{s})$ describes the stochastic part of the KED model and represents a random Gaussian field with a zero mean and a second order stationary covariance structure. The latter is conveniently modeled by an eligible parametric semi-variogram function, describing the dependence of semi-variance as a function of lag (eventually with a directional dependence).

In our application of KED for seasonal mean precipitation the trend variables $x_k(\mathbf{s})$ are specified as fields of topographic predictors (elevation and gradient) that have been pre-calculated from a high-resolution digital elevation model as further detailed in section 3.2. Several different sets of predictors will be considered and the accuracy of the pertinent interpolations will be compared by cross-validation.

In all our applications, the semi-variogram is assumed to be exponential with a nugget, sill and range as parameters. Despite the two-dimensional character of our study domain (i.e. ridge aligned in the east-west direction) we have chosen an isotropic variogram model in all our experiments. The reason for this is that the deterministic model component in KED comprises the angular asymmetry of the variations in precipitation implicitly via predictor fields that represent the orientation of the ridge. Predictors of height and slope, especially at larger space scales, vary in the north-south direction more than in the west-east direction. Introducing an anisotropy in the stochastic model part (variogram) is likely to compete with the significance of these predictors for interpolation. As a consequence, the results would become very specific to our study domain with its simple geography, where the missing of predictors can be compensated by variogram anisotropy. In a more complex domain – e.g. with a topography orientation changing across the region – such a compensation is far less effective and the incorporation of informative predictors more decisive. In this study, we are interested in predictor dependence in this more general setting, which is why we deliberately refrain from the added flexibility with anisotropic variograms. As for the choice of the exponential variogram, this is motivated by simplicity. Preliminary sensitivity experiments with a spherical variogram (again allowing for nugget) did show very minor differences in results compared to the exponential model.

All model parameters (trend coefficients and variogram parameters) are estimated jointly using the method of restricted maximum likelihood (Schabenberger and Gotway, 2005), which accounts for biases from limited sample size / large predictor sets. The utilization of a likelihood-based estimation procedure is central in our application. Estimating trend coefficients and variogram parameters jointly means that the procedure implicitly distinguishes between variations in the observations that are better explained by the predictors and variations that are better explained by spatial covariance (spatial continuity). This procedure ensures optimality of the parameter estimates and consistency of assumptions with the stochastic model of Eq. (1) (see also Diggle and Ribeiro, 2007). Prior estimation of predictor coefficients by linear regression followed by ordinary kriging of residuals, an estimation procedure frequently applied, has a risk of disturbing spatial autocorrelation when relationship to predictors is the sole source for explaining variance in the regression step.

A complication for adopting KED in the present study is posed by the assumption of a multivariate Gaussian with stationary variance in space for the stochastic component (the residuals of the trend). This condition is rarely met with precipitation data, whose distribution is bounded by zero, has positive skewness and shows larger variance in areas of high compared to low precipitation. Partial remedy of this can be made with a prior monotonic transformation of the data, the application of KED in transformed space, and subsequent back-transformation of the estimated kriging distribution. The procedure, commonly known as trans-Gaussian kriging (Schabenberger and Gotway, 2005), has been adopted in all KED experiments of the present study, using the Box-Cox power transformation (Box and Cox, 1964):

$$Y^* = \begin{cases} \frac{Y^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log(Y) & \lambda = 0 \end{cases} \qquad (2)$$

Here we prescribe the transformation parameter at $\lambda$=0.5, which corresponds to a square root transformation of the data. This choice is motivated by analyses of Erdin et al. (2012), showing that a formal estimation of $\lambda$ (by maximum likelihood) did not significantly alter the best estimates compared to when it was prescribed at 0.5. (The change was however significant for the kriging uncertainty.) Finally, the back-transformed results of KED were obtained, in the present study, following a numerical procedure described in Erdin et al. (2012).

It is worth noting here, that the Box-Cox transformation improves compliance with model assumptions only with respect to non-stationarity related to the skewness of precipitation amounts. Precipitation intermittency (the existence of contiguous dry/wet areas) is responsible for non-stationarities that the transformation does not eliminate. Note that, with $\lambda$=0.5, transformation (2) maps all dry measurements to –2. Methods have been proposed to deal with intermittency explicitly in the spatial modeling of precipitation (e.g. Fuentes et al., 2008; Schleiss et al., 2014; Seo, 1998). These were not considered in our application. While intermittency is violating model assumptions in the interpolation of daily precipitation, this is not an issue for the interpolation of seasonal climatological means.

The KED model of Eq. (1) comprises two simplifying special cases that will be considered in this study as alternative methods of spatial interpolation. The first is to assume that $Z(\mathbf{s})$ is a spatially uncorrelated Gaussian field with zero mean and constant variance. This corresponds to the classical linear regression model (hereafter denoted as LM) with estimates at location $\mathbf{s}$ determined by the linear combination of predictors only. As with KED we apply the linear regression case with square-root transformed data and appropriately back-transformed results. The LM method is used here for comparative purposes because it is often adopted as an exploratory tool to constitute suitable predictor sets for KED. It is important, however, to note that the best estimate of the linear model $\mu_{LM}(\boldsymbol{s})$ is not equal to the deterministic part of KED $\mu_{\mathrm{KED}}(\mathbf{s})$, because the estimates for the parameters $\beta_k$ differ without and with consideration of spatial autocorrelation.

The second special case of the KED model (1) is that when topographic predictors are omitted, i.e. presuming $\beta_k$=0 ($k$=1,..,$K$), and assuming the spatial variations in the observations are purely the result of a second order stationary process. This is the limit of Ordinary Kriging (denoted OK). As with the other methods, OK is used here with square-root transformed data. Differences in the performance between KED and OK describe the value added by topographic predictors. But, again, the best estimate fields of OK are not equal to the stochastic component of KED because the parameter estimates differ.

All computations are done in R (R Core Team, 2012) using the geostatistics package geoR (Diggle and Ribeiro, 2007).

### 3.2. Predictors for the interpolation of long-term mean precipitation

The topographic predictors used in this study are based on the digital elevation model (DEM) of the Shuttle Radar Topography Mission (SRTM, Farr et al., 2007). SRTM was obtained using both C- and X-band microwave radars and has, originally, a resolution of about 90m. In this study we use the SRTM elevation model on a 1 km grid of the Lambert Azimuthal Equal Area Coordinate Reference System (ETRS89-LAEA, Annoni et al., 2001).

The three main topographic predictors considered are fields of elevation and gradients in the two cardinal directions across the ridge (north-south) and along the ridge (east-west). Several predictors for each of these quantities will be considered, describing variations in elevation and gradients at different space scales. These were derived from smoothed versions of the original DEM, after applying a Gaussian kernel with window widths of 1 km, 5 km, 10 km, 25 km and 75 km, respectively. A predictor set that involves, for example, elevation and gradients at three space scales, comprises a total of 9 different predictor fields, 3 for elevation, 3 for the north-south gradient and 3 for the east-west gradient. Values of the predictors at the station locations were always taken from the nearest grid-cell of the predictor fields.

Care was required to avoid co-linearity between predictors when combining several of them for the various space scales. To this end, predictors for a scale were defined as the difference between the variable at that scale and the same variable at the next larger scale. For example, the 25 km elevation predictor in a set involving the scales 1, 25 and 75 km is obtained by calculating the difference between the 25 km and the 75 km smoothed versions of the DEM.

Apart from analyzing fields of seasonal mean precipitation directly from seasonal mean station observations, we also investigate the potential of recombining a seasonal mean field from several separate spatial analyses for average precipitation within the classes of a circulation type classification. Precipitation topography relationships may be more clearly established under conditions of similar large-scale circulation, and this could assist the derivation of a seasonal mean field through further stratification.

The consideration of circulation types permits the introduction of an additional circulation-guided topographic predictor. It is defined as

$$G_w(\boldsymbol{s},\lambda,k) = \boldsymbol{\nabla} e(\boldsymbol{s},\lambda) \cdot \frac{\boldsymbol{V}_{\boldsymbol{g}}^{(\boldsymbol{k})}(\boldsymbol{s})}{\left\|\boldsymbol{V}_{\boldsymbol{g}}^{(\boldsymbol{k})}(\boldsymbol{s})\right\|} \qquad (3)$$

where $\boldsymbol{\nabla} e(\boldsymbol{s},\lambda)$ denotes the gradient of the topographic elevation (valid for smoothing scale $\lambda$ at location $\mathbf{s}$). $\boldsymbol{V}_{\boldsymbol{g}}^{(\boldsymbol{k})}(\boldsymbol{s})$ the geostrophic wind of circulation class $k$ at location $\boldsymbol{s}$. $G_w$ describes the topographic gradient along the direction of the geostrophic wind and will be denoted as *wind-aligned gradient* for brevity. As with the topographic gradients along the cardinal directions, $G_w$ is considered to depend on spatial scale. The geostrophic wind was determined from the sea level pressure composites of the circulation type classification (PCACA9 see section 2), originally given on a 0.5 degree grid, by interpolation (Gaussian kernel) onto the 1 km grid of the DEM and subsequent calculation of the geostrophic wind. Note, that for $G_w$ the smoothing is applied to elevation $e(\mathbf{s})$ only because the geostrophic wind field is already smooth as a result of the coarse resolution of the underlying sea level pressure field and its smooth interpolation to the DEM grid.

Fig. 3 illustrates examples of the wind-aligned gradient $G_w$ obtained for two circulation types of the PCACA9 classification. The marked change of $G_w$ across topographic crests (and across valleys) is evident, as well as its distinct spatial distribution between the two circulation types with their distinct sea level pressure gradient (geostrophic wind) over the domain.

Consideration of $G_w$ as a candidate predictor is obviously motivated by ideas of upslope orographic rainfall enhancement and rain shadowing on the lee of mountains. Indeed at the scale of the entire ridge such flow related precipitation anomalies are clearly evident with the PCACA9 circulation type classification, at least in autumn, winter and spring (see Schiemann and Frei, 2010).

Apart from $G_w$ as defined in (3) we have also experimented with an alternative definition that has omitted the normalization of the geostrophic wind. Such a predictor was previously considered in Johansson and Chen (2003) and in Kyriakidis et al. (2001) for example. However, our experiments showed less explanatory power for precipitation in our study domain compared to $G_w$ as defined in (3). In the following, we consider $G_w$ simply as an alternative to the topographic gradients along the two cardinal axes and will examine how this replacement (together with the stratification of circulation types) affects interpolation quality for seasonal mean precipitation in the domain.

### 3.3. Interpolation of daily precipitation

Our experiments on the interpolation of daily precipitation are also making use of the concepts of kriging with external drift and ordinary kriging (section 3.1) as used for the interpolation of seasonal mean precipitation. However, rather than using the topographic predictors directly as trend variables, the daily interpolation adopts fields of seasonal mean or circulation-type mean precipitation as trend variables. Precipitation measurements at short time scales usually exhibit large spatial variations from which systematic topographic effects are difficult to estimate. The solution followed here is to inject this information via pre-calculated long-term averages. The approach is somewhat related to the common use of climatological mean fields as reference (e.g., New et al., 2000; Widmann and Bretherton, 2000), but instead of adopting the reference as scaling factor, uses it as trend variable in KED.

Following the main focus of our study on precipitation topography relationships, we conduct experiments with daily interpolations and shed light on the role of the climatological reference fields. To this end the interpolation errors are compared between different specifications of the trend variable (see Table 2 for a list of experiments). The trend settings include (a) a long-term seasonal mean built with topographic predictors (experiment *KED*(KED1e)), (b) the long-term mean of the day's pertinent circulation type (experiment *KED*(KED1e+)), and (c) a representation of the seasonal climatology that has not used topographic predictors (*KED*(OK)). Comparison of these settings with an ordinary kriging based direct interpolation (experiment $OK(\cdot)$) will clear up the benefit of using climatological reference fields in daily interpolation. (Note that in contrast to the interpolation of climatic average where most of the stations have non-zero precipitation values, daily measurements can sometimes report dry conditions everywhere. Since kriging cannot operate with zero variance, the precipitation field is set to zero in this particular case.)

Finally, we compare the results obtained in this study using KED over a small cross-section of the Alps with results obtained from a previously developed deterministic interpolation scheme that was applied for daily precipitation over the entire Alpine ridge (Isotta et al., 2013). The trans-Alpine method builds on a version of PRISM (Daly et al., 1994, 2002; Schwarb, 2001) for monthly long-term mean fields and on SYMAP (Frei et al., 1998; Shepard, 1984) for the daily relative anomalies from the mean. The experiment will be denoted as *SYMAP*(PRISM). Results from this method rely on a cross-validation table previously calculated and provided by Isotta et al. (2013).

### 3.4. Evaluation

Our comparison and discussion of the various interpolation experiments is based on systematic leave-one-out cross-validations, rejecting one-by-one all the stations of the domain and estimating pertinent interpolations at the location and with the predictors for that station.

Two error scores will be used to summarize the performance of the methods. The first is a measure of the relative bias and corresponds to the ratio of predicted ($p_i$) over observed $o_i$ precipitation totals, averaged over all (or a subset of n) rain gauges:

$$\boldsymbol{B} = \frac{\sum_{i=1}^{n} p_i}{\sum_{i=1}^{n} o_i} \qquad (4)$$

The second score is the relative mean root transformed error and defined as:

$$\boldsymbol{E} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\sqrt{p_i}-\sqrt{o_i}\right)^2}{\frac{1}{n}\sum_{i=1}^{n}\left(\sqrt{\bar{o}}-\sqrt{o_i}\right)^2} \qquad (5)$$

Here $\bar{o}$ is the spatial average of the observations over all (or a subset of n) stations. The numerator represents a sort of mean squared error, but with square-root transformed data. The transformation is introduced here to avoid excessive dependence on large precipitation values and hence to obtain a more balanced sensitivity on errors across the frequency distribution. The denominator is then representing some sort of spatial variance of the transformed values and this is used as a reference against which errors of the prediction are measured. Values of E are always greater than zero. Values smaller than 1 mean that typical errors are smaller than the spatial variations. Values larger than one mean that the prediction has larger errors compared to a simple prediction of the spatial mean and this can be considered a non-skillful prediction.

Depending on the data stratification and interpolation method, between 317 and 420 stations are available for estimation and interpolation. To ensure maximum comparability of the evaluation results, however, we use a fixed set of 317 stations to calculate the above error scores.

# 4. RESULTS

## 4.1. Interpolation of mean precipitation

### *4.1.1. Linear regression*

Linear regression is often considered an exploratory framework with which potential predictors for a trend model of KED can be compared. We therefore develop our discussion starting with results from the special case when spatial autocorrelation is neglected and then pursue the changes when introducing autocorrelation in combination with topographic predictors.

The number of possible regression models with three variables (elevation, north-south gradient, east-west gradient) and six different spatial scales is very large. We have selected three of them for our discussion because of their illustrative purposes. The simplest (LM1e, see Table 1) has only elevation at the finest spatial scale (1 km) as predictor. It is a traditional and wide spread model of topography effects on precipitation (see section 1). The second (LM3e, see Table 1) involves also elevation only, but at three different space scales (75 km, 25 km, 1 km). The third model (LM9eg, see Table 1) involves elevation and gradients (in both cardinal directions), again at the three space scales (75 km, 25 km, 1 km). Experiments with all five space scales (including also 5 km and 10 km) showed that the three selected scales led to the largest values in adjusted $R^2$. There were slight variations in the “optimal” model choice between seasons but the prescription of the three scales did not significantly lower the explanatory power. Note that a formal and automated model selection procedure (using step-wise linear regression) was not feasible in our application, because the predictors for one scale depend on those retained for other scales (elimination of co-linearity, see section 3b).

Table 3 lists values of adjusted $R^2$ for the three selected regression models. The overall pattern is very similar between the seasons. Topography at the finest scale only (LM1e) explains a very low proportion of the spatial variance in the observations. This is not too surprising, considering that the distribution of mean precipitation is mainly characterized by anomalous wet conditions along the northern foothills and dryer conditions in the high-elevation interior of the ridge (see e.g. Fig. 2a, results for other seasons are not much different). Local elevation does, obviously, not explain this larger-scale pattern well. The situation improves when involving elevation at three space scales (1 km, 25 km and 75 km): LM3e explains a considerable portion the precipitation variability across the domain. Finally, the largest explained variance is obtained when topographic gradient fields are included (LM9eg). Now, the predictor set involves a large-scale pattern (the north-south gradient at the coarsest scale) that distinguishes between flatland, foothills and inner Alps, i.e. the major large-scale contrasts in the precipitation field that was a major obstacle for the previous two models. Interestingly, the coefficient (and statistical significance) of the 1 km elevation predictor is much larger in this comprehensive model than in the simple model LM1e. This suggests that there is some dependence on local elevation in the distribution, but this was difficult to represent in the elevation-only models because it is superimposed by a larger-scale north-south profile that is, itself, poorly explained by elevation.

Despite its decent values in explained variance, the 9-predictor model LM9eg shows elementary deficiencies in reproducing the distribution of rain-gauge measurements in the domain. These are illustrated for the example of DJF mean precipitation in Fig. 4a. Precipitation is systematically overestimated over a wide flatland belt adjacent to the ridge (see e.g. full red square), underestimated along the foothills and, again, overestimated in interior parts of the ridge (see e.g. dashed red square). Apparently, the larger-scale topographic predictors provide, in linear combination, only a partial match to the observed north-south profile and the resulting prediction tends to smooth out some of the variations. Similar types of deficiencies (although differing in exact location) were evident with other combinations or the full set of space scales, and for the other seasons. There was always clear spatial clustering in the prediction errors (regression residuals). It seems that, even with quite comprehensive predictor sets, it is difficult to capture in a regression model all aspects of the precipitation field resolved by the station network. Surprisingly, this is even the case with the comparatively simple north-south profile of this study, for which the construction of a suitable predictor set may have looked easy at first.

#### *4.1.2. Kriging*

Ordinary kriging (OK) seeks to represent the precipitation distribution entirely without topographic predictors. The corresponding estimation (Fig. 4b) has a smooth appearance but reproduces the characteristic north-south contrasts between flatland, foothills and inner Alps. Hence, OK amends some of the regional deficiencies of the linear regression model of Fig. 4a (see red squares). However, in the inner Alpine region, several rain-gauges with anomalously wet conditions (mostly at mountain peak stations) are represented as isolated spots. It appears as if some elevation dependency that is not explicitly resolved by the station network is missed out because of the absence of predictors in OK.

Fig. 4c depicts the result obtained with KED, i.e. integrating predictors and spatial autocorrelation, using the comprehensive three-scale elevation and gradients model as trend (KED9eg). The distribution shows the superposition of a spatially smooth pattern (similar to OK, Fig. 4b) and a small-scale pattern with topographic features that are not explicitly resolved by the station network (similar to LM9eg). The consideration of spatial autocorrelation has amended for the deficiencies of LM9eg in representing the larger-scale north-south profile (e.g. red squares). Moreover, the strong contrasts between mountain stations (moist) and valley stations (dry) in the interior Alps are now integrated via an elevation (and gradient) dependence at small scales.

It is interesting to realize that the three just discussed interpolation methods yield markedly different estimates, not just regionally, but also when aggregated over larger scales. This is further illustrated in Fig. 5, which depicts the results of Fig. 4 when averaged over latitude bands (along the ridge). OK and KED9eg both represent a moist anomaly at the foothills, centered at an elevation of about 1200 mMSL. This anomaly is much less pronounced and more wide-spread in LM9eg. Towards the inner Alpine region the three methods yield markedly different areal estimates with OK being much dryer than the regression model and KED. OK and KED differ by between 5-25% in this area. In the inner Alpine region, it is not entirely clear, at this point, which of the methods are more realistic. Clearly, there is a risk of general underestimates by OK due to the missing out of topography dependence in conjunction with poor sampling of high-elevation areas. But there is also a risk that KED suffers from overestimates, if, for example, the elevation dependence estimated over the full domain is not representative for the inner Alps.

In the following we assess the relative performance of a range of interpolation models from the above three categories by means of a systematic leave-one-out cross-validation. Results are depicted for DJF mean precipitation in Figure 6. The two panels are for B (panel a, ratio) and for E (panel b, dimensionless, see section 2d for the definition of the scores). To better visualize the effects of the various interpolation schemes, both error scores are calculated separately for the stations within four elevation ranges. Here, we discuss the results more extensively for the case of DJF mean precipitation, but very similar results – and similar interpretations – were found for the other seasons. This is supported by Tables 4 and 5, which list a summary of the error scores for all seasons.

When averaged over all stations the values of bias are small, varying between 0.97 - 0.995 depending on method (Fig. 6a, dashed lines). The largest underestimate (three percent) is obtained for LM1e (the linear model with local elevation as single predictor). More significant biases are, however, found in individual elevation ranges. This is particularly so for the linear regression model LM1e and for ordinary kriging OK. The lack of topographic predictors in OK impinges upon the interpolation at high elevation. Here OK systematically underestimates by about 30%. This deficiency is mostly corrected with interpolation models that incorporate topographic predictors (LM9eg and KED9eg). The explicit modeling of topography allows for a compensation of the effects of non-representative vertical distribution of the station sample. In the framework of KED, this remedy is almost as good with only one predictor (KED1e) as with many predictors (KED9eg). In the linear model framework, however, in-situ elevation alone provides a poor model of the spatial distribution (see also Table 3), and this reflects in large and alternating biases between the elevation ranges. An interpretation of this difference may be seen in the fact that the estimated coefficient for the 1 km elevation predictor is quite different between LM1e and KED1e. It seems that the consideration of spatial autocorrelation in KED1e permitted for a much more realistic separation between small-scale elevation dependence (modelled by the predictor) and larger-scale precipitation variations (modelled by the autocorrelation part). In contrast, LM1e attempts to capture larger-scale and small-scale variations with one single linear dependence by construction. It is then likely that larger-scale variations (such as the north-south profile) disturb a realistic estimate of the small-scale elevation dependence.

The limited accuracy of linear regression models in predicting the spatial variations of seasonal mean precipitation is most evident in the relative error score E (Fig. 6b, Table 5). Values are close to the critical value of 1, where prediction errors are comparable to the magnitude of spatial variations (see section 3.4). There is improvement when including more predictors (e.g. LM9eg vs LM1e), but considerable errors remain even with comprehensive predictor sets. This reflects results previously seen in Fig. 4a. Note, that the inclusion of the gradient at the 75 km scale (the largest considered) yields the smallest errors. Obviously, this predictor is essential for a regression model to capture the characteristic north-south profile.

The OK model (no topographic predictors) has much smaller errors than the regression models, except for the highest elevation range (Fig. 6b). OK profits from its explicit account for spatial autocorrelation, which permits the reproduction of larger-scale variations (e.g. the north-south profile) from the information at neighboring stations (see also Fig. 4b). In our application, this methodological feature yields considerably smaller errors than a comprehensive predictor set in a regression model, at least for low and intermediate elevation ranges. At large elevations, however, the OK model suffers large E values (close to 1), which reflects the large bias there (see also Fig. 6a) and the poor reproduction of wet conditions at inner-Alpine mountain stations (see also Fig. 4b).

The family of KED models, which include both topographic predictors and spatial autocorrelation, yield the smallest interpolation errors of all models (E scores, Fig. 6b, Table 5). In comparison to OK the improvement is modest in the lower elevation classes, but substantial at higher elevation. The inclusion of topographic predictors seems to be central for reducing the caveats of OK in the inner-Alpine region (biases and over-smoothing of small-scale variations, see also Fig. 4). But the KED models also yield markedly smaller errors (at all elevations) compared to using the predictors in a linear regression.

Between the different KED models (with different predictor sets) there are only marginal differences in the scores (Fig. 6b, Table 5). Values of E are roughly the same for the model with only one predictor (elevation at the 1 km scale, KED1e) and models with elaborate predictor sets (e.g. KED3e, KED9eg). At first sight this is surprising, given that the scores for linear regression models showed to be sensitive to the predictor sets. Our explanation of this result is that the role of topographic predictors is distinct between linear models and KED. Linear models are in need of geographic predictors to capture the full distribution. The 25 km and 75 km predictors are therefore highly relevant. In KED, however, the part of the distribution that is well resolved by the station network can be represented by the spatial autocorrelation component (kriging) and topographic predictors are primarily used to describe smaller-scale variations not explicitly resolved by the station network. Here the 25 km and 75 km predictors may be virtually unnecessary. The distinct role of topographic predictors in the two model families also reflects in differences in the statistical significance and quantitative values of the predictor coefficients ($\beta_k$, see Eq. 1). In all the KED models, the 1 km elevation predictor is by far the most statistically significant, whereas in the linear models other predictors (notably the 75 km topography gradient) are occasionally more significant.

Experiment KED9eg (10 km, 5 km, 1 km) involves predictors at spatial scales all smaller than the station spacing. Still there seems to be little added value compared to the model with the 1 km elevation predictor only (KED1e, see Fig. 6b and Table 5). It is unclear if this result implies that the additional predictors (5 km and 10 km elevations and gradients) are, indeed, not very relevant (on top of the 1 km elevation) for describing small-scale precipitation variations in the Alps. There may be insufficient sampling of these predictors in the station sample, considering that most of the inner-Alpine stations are in valleys or on mountain tops.

Note that E shows a general U-shape for the more skillful interpolation models (Fig. 6b), implying that relative errors are larger (smaller) at low and high (intermediate) elevations. This pattern is also related to the definition of the score, which uses spatial variance within the elevation classes as a reference (see denominator in Eq. 3). Larger values of E at low elevations are primarily because of the small variance in precipitation measurements over the flatland. In fact the numerator of E increases monotonically with elevation.

### 4.2. Stratification by circulation types

In this section we examine the potential of considering circulation types for the derivation of interpolated mean seasonal precipitation fields. Two extensions will be considered. The first deals with a sub-stratification of the season. For this purpose, several KED interpolation models are adopted for each class of the circulation classification, separately. The resulting fields of mean precipitation for each class are subsequently re-combined into a seasonal mean field by weighting according to the classes' frequency. Experiments adopting this sub-stratification are labeled with a '+' sign (see Table 1). The second extension deals with the circulation-dependent predictor $G_w$ as outlined in section 3.2. The wind-aligned gradient is considered here as an alternative for the gradients in the two cardinal directions. The experiment involving this topographic predictor is labeled with the letter 'v' (KED6ev+, see Table 1). KED6ev+ uses three different components of the $G_w$ field, corresponding to three space scales (1 km, 25 km and 75 km). These were derived by the smoothing procedure and removal of co-linearities, just as with the previous predictor fields (see section 3b). Our results were derived with the 9-class PCACA9 classification as described in section 2.

Cross-validation results with these experiments are depicted in Fig. 7, again for B and E, using the same format as in Fig. 6. Note that these are scores for a mean seasonal (here DJF) precipitation field, not a field for the mean of a circulation class. Hence the scores include errors from the recombination over the classes. Results using circulation classification input are compared against a direct interpolation of seasonal means using the previously adopted model KED9eg. Results of the two scores for other seasons are listed in Tables 6 and 7.

With all tested interpolation methods, the biases are smaller than 2% (5%) below (above) 1000 mMSL (Fig. 7a). The interpolation with circulation classes (KED1e+, KED6ev+, KED9eg+) exhibits a slightly different bias pattern compared to that of seasonal the means directly (KED9eg), with a smaller underestimation at elevations between 1500-3500 m and a larger overestimation between 1000-1500 m. But these differences (and the bias values themselves) are small, much smaller than typical random errors, and there is not much meaning in using them for a relative assessment of the methods. The conclusion is that stratification by circulation class and usage of a wind-aligned gradient $G_w$ do not significantly change the bias pattern of the interpolation methods.

Comparison of the different methods in terms of E (Fig. 7b) reveals that all interpolation methods have a very similar error pattern. Neither the stratification by circulation class alone (with conventional predictors, KED1e+ and KED9eg+), nor the consideration of a wind-aligned gradient (KED6ev+) can significantly improve over the interpolation of mean seasonal values (KED9eg). The overall scores (dashed lines) are slightly better for the stratification methods with gradient (KED9eg+) and wind-aligned gradient (KED6ev+) predictors (see also Table 7), but the direct seasonal method (KED9eg) is superior at three of the four elevation classes.

We have tested several alternative definitions of a circulation dependent predictor, deviating from that in Eq 3. These included the introduction of an asymmetry between upslope and downslope gradients, truncating the $G_w$ field to only measure upslope gradients, including the wind speed (i.e. discarding the denominator in Eq. 3), and a simple model for an ageostrophic wind component. None of these alternative definitions led to significantly different results.

There are several possible reasons why circulation class information did not improve interpolation accuracy in our application: the region may be geographically too simple or too small to reveal the benefits of a predictor that builds on spatially variable wind directions. The large-scale wind field (derived from a coarse resolution sea level pressure field) may be of limited representativity for the true air flow in such a complex topography. The variability of airflows within a circulation class may be large, so that systematic topographic effects are not necessarily manifest at the small space scales addressed by the $G_w$ predictor. The station sample may not sample the $G_w$ predictor field adequately. And finally, there may be larger sampling errors involved, because less stations could be used in the estimation of means for circulation classes, due to the minimum constraint employed to ensure robustness in temporal sampling (see section 2).

### 4.3. Interpolation of daily precipitation

In this section we compare and evaluate several options for extending the KED interpolation framework for daily precipitation. The main purpose of this comparison is to investigate how sensitive the accuracy of a daily interpolation scheme is to various options of integrating small-scale topography-related information. Alongside, we also compare the KED-based daily models with results from a previously implemented deterministic daily interpolation scheme, that was calibrated over a much larger area (the entire Alpine region) and was used for a popular dataset of trans-Alpine daily precipitation (Isotta et al., 2013).

Table 2 lists the interpolation models compared here and Fig. 8 depicts results from some of these models for a day with widespread and intense precipitation in the study domain. All KED models considered adopt the stochastic concept of Eq. 1 but with one of the previously determined climatological mean fields as trend, rather than with the topographic predictors themselves. The trend field for *KED*(KED1e) is the mean seasonal field KED1e that was derived with the 1 km elevation predictor. Recall, that this version of the mean seasonal distribution showed cross-validation skills comparable to other versions with comprehensive predictor sets (Fig. 6). The precipitation for the example day (Fig. 8a) shows small-scale patterns along the foothills and in the interior of the ridge that reflect patterns of the trend field. For *KED*(KED1e+) the trend field is the mean precipitation for class 9 of the PCACA9 circulation classification. (The example day belongs to this class.) Again, the distribution for the example day (Fig. 8b) bares small-scale variations reflecting the trend field. There are only small differences to the result for *KED*(KED1e) (panel a), because the small-scale pattern (not the magnitude) is very similar between the mean over the class and the mean over the season. Our consideration of *KED*(KED1e+) in the subsequent evaluation will answer whether the sub-stratification by circulation classes can improve interpolation accuracy. As a reference we also consider the models *KED*(OK) and $OK(\cdot)$ which use, respectively, the OK-based seasonal climatology (Fig. 4b) as trend or a simple ordinary kriging of the (transformed) daily values (i.e. no trend). The distributions for the example day are very similar and, compared to the other models much smoother in appearance (see Fig. 8c ).

Fig. 8d depicts daily precipitation for the example day derived by the Alpine-wide *SYMAP*(PRISM) interpolation. This procedure uses, as background, a seasonal climatology derived from a local regression approach (PRISM, DALY et al., 1994; Daly et al., 2002; Schwarb, 2000). The result depicted comes from a 5 km grid interpolation (Isotta et al., 2013), hence, is coarser the results for the other models (1 km grid). It shows more variable and larger peak values than the other models. In contrast to the KED models with elevation as predictor, PRISM estimates precipitation-height gradients locally (considering the representativity of surrounding stations) and this results in more pronounced small-scale variations.

The daily interpolation methods have been quantitatively evaluated using cross-validation over all winter days of 1971-2008 (3400 days). For computational reasons, the cross-validation of the models was only calculated for the daily interpolation step, i.e. with the seasonal background field estimated from all the data, including the test station. Clearly, the daily interpolation step contributes the largest error component, but the errors calculated this simplified way should be considered as a lower bound of the true errors.

Fig. 9 depicts the bias B and the relative mean root transformed error E for daily interpolation in winter (DJF) using the same display format as with Figs. 6 and 7. Note that E values for daily interpolation are much smaller than for the climatological case, because the space-time variance in the observations (denominator in Eq. 3) is much larger.

The bias of the daily interpolation (Fig. 9a) reveals similar features like in the climatic case. Methods without consideration of topographic predictors in the climatological background field ($OK(\cdot)$ and *KED*(OK)) are prone to considerable underestimates at high elevations. The inclusion of topographic predictors in the climatology reduces this bias a lot (*KED*(KED1e) and *KED*(KED1e+)). The results differ only slightly between a seasonal and a circulation-class climatology as trend, the latter being slightly better. The *SYMAP*(PRISM) system is largely unbiased, except at the highest elevation class, where it underestimates by about 10%. Our results confirm that the use of a high-resolution climatology as a background, a widely used concept for the interpolation of daily precipitation (e.g., Haylock et al., 2008; Rauthe et al., 2013; Widmann and Bretherton, 2000), indeed contribute to reducing biases over complex terrain.

The relative ranking of methods in terms of E (Fig. 9b) is similar in all elevation classes, but the differences are largest at high elevations. The KED models that employ a climatology with topographic predictors score best (*KED(*KED1e) and *KED*(KED1e+)). There is no clear preference between the methods using a seasonal mean or a circulation-class mean as trend. Obviously, the categorical information on large-scale circulation did not improve daily interpolation. This may seem surprising considering that the classification utilized (PCACA9) distinguishes Alpine precipitation distributions better than others (Schiemann and Frei, 2010). A likely reason for this is that the circulation responses of precipitation in the study region are more clearly established at larger scales, but less so at scales below the station spacing which matter most for spatial interpolation.

The *KED*(KED1e) and *KED*(KED1e+) methods exhibit clearly better B and E scores than the Alpine-wide *SYMAP*(PRISM) interpolation in the highest elevation class (Fig. 9). Several reasons may contribute to these differences: Firstly, the distance-angular weighting scheme of *SYMAP* uses prescribed weighting functions, whereas the weighting in KED is optimized and flexibly estimated day-by-day (semi-variogram). Secondly, the local estimation of precipitation topography relationships in PRISM may be more prone to sampling errors (small local station sample) than the trend coefficients in *KED1e*/KED1e+. (See also the large small-scale variations in the example of Fig. 8d.) Finally, KED allows for a multiplicative adjustment of the background field and, hence, is more flexible to 'adjust' the background field to the concrete distribution of a day. In this comparison one should, however, take into account that *SYMAP*(PRISM) was designed and calibrated for a much larger area. The KED approach as used here for a subregion of the Alps might become inappropriate for the climatological diversity of the entire ridge given its assumption on stationarity in trend and variogram parameters (see e.g., Phillips et al., 1992).

# 5. CONCLUSION

Modeling the relationship between precipitation and topography is essential for the construction of accurate precipitation grid datasets by statistical interpolation. Here, we have investigated several extensions of the classical precipitation-height model, including predictors of slope in addition to elevation, a multi-scale decomposition of the predictors, a circulation-type dependence of the relationship and the inclusion of a wind-aligned gradient predictor. Variants of these extensions have been proposed previously but their effect on interpolation accuracy was not systematically evaluated and mutually compared so far. Station measurements in our study region (a cross-section of the European Alps) show imprints of slope effects and coarser scale topography in the distribution of mean seasonal precipitation. Intuitively one would therefore expect that the considered extensions could improve interpolation accuracy.

Our experiments illustrate that the benefit from complex predictor sets (elevation and slope, multiple scales) in the interpolation of seasonal mean precipitation depends strongly on the statistical modeling framework. In a linear regression framework there is a clear benefit in the sense that cross-validation errors (random and systematic) are reduced with more predictors included. However, even with nine predictors, the resulting interpolation is unsatisfactory. It poorly replicates the characteristic changes from the flatland over the foothills to the inner section of the ridge as revealed by the station measurements. Linear regression would require many more predictors for a decent reproduction of this pattern because all spatial variations need to be modeled with predictors.

For kriging with external drift (KED, predictors with spatially correlated residuals), however, the role of a complex predictor set was found to be much smaller. Local elevation (a 1 km digital elevation model) was found to be essential for reducing the systematic underestimates and large random errors observed at high elevations with ordinary kriging (OK, no predictors). In fact, the simple one-predictor KED model was substantially better than the linear regression model with nine predictors. But the inclusion of more complex physiographic predictor sets in KED did bring only marginal additional improvement. Neither topographic slopes nor a wind-aligned gradient could effectively reduce the cross-validation errors. Interpolation results with comprehensive multi-scale predictor sets in KED were very similar to those of the one-predictor model, and also the inclusion of circulation-type dependence had only small effects. It seems that a large portion of the spatial precipitation variation in our study region is captured by a model of spatial autocorrelation directly from the measurements (kriging), and that a simple digital elevation model was sufficient (but essential) to correct for interpolation errors emanating from the non-representative vertical distribution of stations.

Linear regression is often considered an exploratory framework in spatial interpolation to identify potential predictors for a trend model of KED. This practice is somewhat questioned by the results of our study. We find a strong contrast in sensitivity to predictor choice between the two methods. Linear regression tends to suggest larger predictor sets than are actually necessary in KED. Our results with KED were not measurably degraded by the inclusion of non-informative predictors. But this resistance is dependent on the estimation procedure. Our approach of estimating the trend coefficients and variogram parameters jointly by maximum likelihood (see section 3.1) permits the estimation process to distinguish between predictor dependence and spatial autocorrelation implicitly (Diggle and Ribeiro, 2007). This distinction is more restricted in an alternative estimation procedure, often referred to as residual kriging or detrended kriging (Martínez-Cob, 1996; Phillips et al., 1992; Prudhomme and Reed, 1999) where predictor coefficients and variogram parameters are estimated in disjoint steps (regression followed by simple kriging of residuals). This will make the method more prone to errors in predictor choice. Regression kriging, yet another estimation procedure (Hengl et al., 2007; Pebesma, 2004; Tadić Perčec, 2010) uses an iterative procedure and should be similarly robust to predictor choice like the likelihood-based estimation used in our study.

Our experiments for daily precipitation illustrate that the utilization of a climatological background field (seasonal climatology) reduces interpolation errors significantly, particularly systematic errors at high elevations in comparison to direct interpolation. The large spatial variability of daily precipitation complicates robust estimation of systematic topographic responses directly from the daily data, but a climatological background field can pick up some of these patterns, which translates into smaller interpolation errors. This result supports a practice widely used in the construction of short-term precipitation grid datasets, but rarely verified so far (Harris et al., 2013; Haylock et al., 2008; Isotta et al., 2013; Rauthe et al., 2013). Clearly, the topographic effects evident in mean precipitation are not necessarily representative for all weather conditions. Our results, however, suggest that estimating these effects separately for typical circulation types does not significantly improve the performance compared to that with a seasonal background. This result may depend on the region considered and the circulation-type classification chosen. At least, the classification we have experimented with here was previously shown to explain precipitation variations in the Alps better than other common classification schemes (Schiemann and Frei, 2010).

The daily KED interpolation method using a seasonal mean climatology as background has turned out to perform better in the Alpine cross-section compared to the method used for a grid dataset over the entire Alpine region (Isotta et al., 2013). This may hint to ways of methodological improvement, but it is premature to value the two methods with regard to their suitability over the entire Alpine region. On the one hand, the existing method makes compromises in order to meet very diverse conditions in climate and station density. On the other hand, extending the KED approach over the entire region raises questions about the representativity of 'globally' estimated trend coefficients and variogram parameters. Moreover, on a practical side, the KED approach may become computationally very demanding with several thousands of stations.

The results of our study are likely dependent on the setting of our study region, such as the density of the station network, the complexity of the topography and the diversity of weather patterns. In other regions where the station network is coarser and, hence, the nearest observations are less informative, extended predictor sets may become more relevant. Nevertheless, our results call for reluctance in our expectations into seemingly versatile topographic predictors for filling the information between in-situ measurements. Clearly, sensitivity experiments like those conducted can help to make a parsimonious choice and to ensure robustness of the final interpolation method.

# Acknowledgments

The research leading to these results has received funding from the European Union, Seventh Framework Programme (FP7/2007-2013) under grant agreement n242093.

# References


Alexander, L. V, Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Tank, A., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Kumar, K. R., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M. and Vazquez-Aguirre, J. L.: Global observed changes in daily climate extremes of temperature and precipitation, J. Geophys. Res., 111(D5), doi:10.1029/2005JD006290, 2006.

Allamano, P., Claps, P., Laio, F. and Thea, C.: A data-based assessment of the dependence of short-duration precipitation on elevation, Phys. Chem. Earth, Parts A/B/C, 34(10), 635–641, doi:10.1016/j.pce.2009.01.001, 2009.

Annoni, A., Luzet, C., Gubler, E. and Ihde, J.: Map projections for Europe, Inst. Environ. Sustain., 2001.

Basist, A., Bell, G. D. and Meentemeyer, V.: Statistical relationships between topography and precipitation patterns, J. Clim., 7(9), 1305–1315, doi:10.1175/1520-0442(1994)007<1305:SRBTAP>2.0.CO;2, 1994.

Benichou, P. and Le Breton, O.: Prise en compte de la topographie pour la cartographie des champs pluviométriques, Agrométéorologie des Régions Moy. Mont., 23–34, 1986.

Bergeron, T.: Preliminary results of Project Pluvius, Comm. L. Erosion, Publ, 53, 226–237, 1961.

Box, G. E. P. and Cox, D. R.: An analysis of transformations, J. R. Stat. Soc., 26A, 221–252, 1964.

Brunetti, M., Lentini, G., Maugeri, M., Nanni, T., Simolo, C. and Spinoni, J.: Projecting North Eastern Italy temperature and precipitation secular records onto a high-resolution grid, Phys. Chem. Earth, 40-41, 9–22, doi:10.1016/j.pce.2009.12.005, 2012.

Bukovsky, M. S. and Karoly, D. J.: A brief evaluation of precipitation from the North American Regional Reanalysis, J. Hydrometeorol., 8(4), 837–846, doi:10.1175/JHM595.1, 2007.

Cortesi, N., Trigo, R. M., Gonzalez-Hidalgo, J. C. and Ramos, A. M.: Modelling monthly precipitation with circulation weather types for a dense network of stations over Iberia, Hydrol. Earth Syst. Sci., 17(2), 665–678, doi:10.5194/hess-17-665-2013, 2013.

Cosma, S., Richard, E. and Miniscloux, F.: The role of small-scale orographic features in the spatial distribution of precipitation, Q. J. R. Meteorol. Soc., 128(579), 75–92, doi:10.1256/00359000260498798, 2002.

Cressie, N. A. and Cassie, N. A.: Statistics for spatial data, Wiley New York., 1993.

Crochet, P., Johannesson, T., Jonsson, T., Sigurdsson, O., Bjonsson, H., Palsson, F. and Barstad, I.: Estimating the spatial distribution of precipitation in Iceland using a linear model of orographic precipitation, J. Hydrometeorol., 8(6), 1285–1306, doi:10.1175/2007JHM795.1, 2007.

Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L. and Pasteris, P.: A knowledge-based approach to the statistical mapping of climate, Clim. Res., 22(2), 99–113, doi:10.3354/cr022099, 2002.

Daly, C., Neilson, R. P. and Phillips, D.: A statistical-topographic model for mapping climatological precipitation over mountainous terrain, J. Appl. Meteorol., 33(2), 140–158, doi:10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2, 1994.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Hólm, E., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A., Monge-Sanz, B., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137(656), 553–597, 2011.

Diggle, P. J. and Ribeiro, P. J.: Model-Based Geostatistics, Springer., Springer Series in Statistics., 2007.

Erdin, R., Frei, C. and Künsch, H. R.: Data Transformation and Uncertainty in Geostatistical Combination of Radar and Rain Gauges, J. Hydrometeorol., 13(4), 1332–1346, doi:10.1175/JHM-D-11-096.1, 2012.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D. and Alsdorf, D.: The shuttle radar topography mission, Rev. Geophys., 45(2), RG2004, doi:10.1029/2005RG000183, 2007.

Frei, C., Christensen, J. H., Deque, M., Jacob, D., Jones, R. G. and Vidale, P. L.: Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, J. Geophys. Res., 108(D3), 4124, doi:10.1029/2002JD002287, 2003.

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

Frei, C., Schär, C., Luthi, D. and Davies, H. C.: Heavy precipitation processes in a warmer climate, Geophys. Res. Lett., 25(9), 1431–1434, doi:10.1029/98GL51099, 1998.

Fuentes, M., Reich, B. and Lee, G.: Spatial-temporal mesoscale modeling of rainfall intensity using gage and radar data, Ann. Appl. Stat., 2, 1148–1169, doi:10.1214/08-AOAS166, 2008.

Fuhrer, O. and Schär, C.: Embedded cellular convection in moist flow past topography, J. Atmos. Sci., 62(8), 2810–2828, doi:10.1175/JAS3512.1, 2005.

Germann, U. and Joss, J.: Variograms of radar reflectivity to describe the spatial continuity of Alpine precipitation, J. Appl. Meteorol., 40(6), 1042–1059, doi:10.1175/1520-0450(2001)040<1042:VORRTD>2.0.CO;2, 2001.

Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, J. Hydrol., 228(1), 113–129, 2000.

Gottardi, F., Obled, C., Gailhard, J. and Paquet, E.: 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, 2012.

Greminger, P.: Natural hazards and the Alpine Convention: Event analysis and recommendations, Bern., 2003.

Groisman, P. Y. and Legates, D. R.: The accuracy of United States precipitation data, Bull. Am. Meteorol. Soc., 75(2), 215–227, doi:10.1175/1520-0477(1994)075<0215:TAOUSP>2.0.CO;2, 1994.

Gyalistras, D.: Development and validation of a high-resolution monthly gridded temperature and precipitation data set for Switzerland (1951-2000), Clim. Res., 25(1), 55–83, doi:10.3354/cr025055, 2003.

Harris, I., Jones, P. D., Osborn, T. J. and Lister, D. H.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, Int. J. Climatol., 34(3), 623–642, doi:10.1002/joc.3711, 2013.

Haylock, M. R., Hofstra, N., Tank, A., Klok, E. J., Jones, P. D. and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, J. Geophys. Res., 113(D20), D20119, doi:10.1029/2008JD010201, 2008.

Hengl, T., Heuvelink, G. and Rossiter, D. G.: About regression-kriging: from equations to case studies, Comput. Geosci., 33(10), 1301–1315, 2007.

Hevesi, J. A., Flint, A. L. and Istok, J. D.: Precipitation estimation in mountainous terrain using multivariate geostatistics. Part II: Isohyetal maps, J. Appl. Meteorol., 31(7), 677–688, doi:10.1175/1520-0450(1992)031<0677:PEIMTU>2.0.CO;2, 1992.

Hewitson, B. C. and Crane, R. G.: Gridded area-averaged daily precipitation via conditional interpolation, J. Clim., 18(1), 41–57, doi:10.1175/JCLI3246.1, 2005.

Holzkamper, A., Calanca, P. and Fuhrer, J.: Statistical crop models: predicting the effects of temperature and precipitation changes, Clim. Res., 51(1), 11–21, doi:10.3354/cr01057, 2012.

Houze, R. A., James, C. N. and Medina, S.: Radar observations of precipitation and airflow on the Mediterranean side of the Alps: Autumn 1998 and 1999, Q. J. R. Meteorol. Soc., 127(578), 2537–2558, doi:10.1002/qj.49712757804, 2001.

Hutchinson, M. F.: Interpolation of rainfall data with thin plate smoothing splines. Part I: Two dimensional smoothing of data with short range correlation, J. Geogr. Inf. Decis. Anal., 2(2), 139–151, 1998.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Gianni, M. and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, Int. J. Climatol., 34(5), 1657–1675, doi:10.1002/joc.3794, 2013.

Johansson, B. and Chen, D. L.: The influence of wind and topography on precipitation distribution in Sweden: Statistical analysis and modelling, Int. J. Climatol., 23(12), 1523–1535, doi:10.1002/joc.951, 2003.

Kyriakidis, P. C., Kim, J. and Miller, N. L.: Geostatistical mapping of precipitation from rain gauge data using atmospheric and terrain characteristics, J. Appl. Meteorol., 40(11), 1855–1877, doi:10.1175/1520-0450(2001)040<1855:GMOPFR>2.0.CO;2, 2001.

Machguth, H., Paul, F., Kotlarski, S. and Hoelzle, M.: Calculating distributed glacier mass balance for the Swiss Alps from regional climate model output: A methodical description and interpretation of the results, J. Geophys. Res., 114(D19), D19106, doi:10.1029/2009JD011775, 2009.

Martínez-Cob, A.: Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain, J. Hydrol., 174(1), 19–35, 1996.

Neff, E. L.: How much rain does a rain gage gage?, J. Hydrol., 35(3), 213–220, 1977.

New, M., Hulme, M. and Jones, P.: Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate, J. Clim., 13(13), 2217–2238, 2000.

Pebesma, E. J.: Multivariable geostatistics in S: the gstat package, Comput. Geosci., 30(7), 683–691, 2004.

Perry, M. and Hollis, D.: The development of a new set of long-term climate averages for the UK, Int. J. Climatol., 25(8), 1023–1039, doi:10.1002/joc.1160, 2005.

Philipp, A., Bartholy, J., Beck, C., Erpicum, M., Esteban, P., Fettweis, X., Huth, R., James, P., Jourdain, S., Kreienkamp, F., Krennert, T., Lykoudis, S., Michalides, S. C., Pianko-Kluczynska, K., Post, P., Alvarez, D. R., Schiemann, R., Spekat, A. and Tymvios, F. S.: Cost733cat-A database of weather and circulation type classifications, Phys. Chem. Earth, 35(9-12), 360–373, doi:10.1016/j.pce.2009.12.010, 2010.

Phillips, D. L., Dolph, J. and Marks, D.: A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain, Agric. For. Meteorol., 58(1-2), 119–141, doi:10.1016/0168-1923(92)90114-J, 1992.

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

Prudhomme, C. and Reed, D. W.: Mapping extreme rainfall in a mountainous region using geostatistical techniques: A case study in Scotland, Int. J. Climatol., 19(12), 1337–1356, doi:10.1002/(SICI)1097-0088(199910)19:12<1337::AID-JOC421>3.3.CO;2-7, 1999.

Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A. and Gratzki, A.: A Central European precipitation climatology—Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS), Meteorol. Zeitschrift, 22(3), 235–256, doi:10.1127/0941-2948/2013/0436, 2013.

R Development Core Team.: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN: 3-900051-07-0 http://www.R-project.org, 2012.

Richter, D.: Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Messfehlers des Hellmann-Niederschlagsmessers, Deutscher Wetterdienst, Offenbach am Main., 1995.

Roe, G. H.: Orographic precipitation, Annu. Rev. Earth Planet. Sci., 33, 645–671, doi:10.1146/annurev.earth.33.092203.122541, 2005.

Schabenberger, O. and Gotway, C. A.: Statistical Methods for Spatial Data Analysis, Chapman an., CRC Press., 2005.

Schädler, B. and Weingartner, R.: Ein detaillierter hydrologischer Blick auf die Wasserressourcen der Schweiz–Niederschlagskartierung im Gebirge als Herausforderung, Wasser, Energie, Luft, 94(7), 189–197, 2002.

Schiemann, R. and Frei, C.: How to quantify the resolution of surface climate by circulation types: An example for Alpine precipitation, Phys. Chem. Earth, 35(9-12), 403–410, doi:10.1016/j.pce.2009.09.005, 2010.

Schleiss, M., Chamoun, S. and Berne, A.: Non-stationarity in intermittent rainfall: the "dry drift," J. Hydrometeorol., 15, 1189–1204, doi:10.1175/JHM-D-13-095.1, 2014.

Schmidli, J., Schmutz, C., Frei, C., Wanner, H. and Schär, C.: Mesoscale precipitation variability in the region of the European Alps during the 20th century, Int. J. Climatol., 22(9), 1049–1074, doi:10.1002/joc.769, 2002.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M. and Rudolf, B.: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, Theor. Appl. Climatol., 115(1-2), 15–40, doi:10.1007/s00704-013-0860-x, 2013.

Schwarb, M.: The Alpine precipitation climate: Evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data., Zürcher Klima-Schriften, 80(Diss. ETHZ 13911), 1–138, 2001.

Seo, D.-J.: Real-time estimation of rainfall fields using rain gage data under fractional coverage conditions, J. Hydrol., 208(1), 25–36, 1998.

Sevruk, B.: Systematischer Niederschlagmessfehler in der Schweiz, Der Niederschlag der Schweiz, Beiträge zur Geol. Karte der Schweiz-Hydrologie, 31, 65–75, 1985.

Sevruk, B.: Regional dependency of precipitation-altitude relationship in the Swiss Alps, Clim. Change, 36(3-4), 355–369, doi:10.1023/A:1005302626066, 1997.

Sevruk, B.: Rainfall measurement: gauges, Encycl. Hydrol. Sci., doi:10.1002/0470848944.hsa038, 2005.

Sharples, J. J., Hutchinson, M. F. and Jellett, D. R.: On the horizontal scale of elevation dependence of Australian monthly precipitation, J. Appl. Meteorol., 44(12), 1850–1865, doi:10.1175/JAM2289.1, 2005.

Shepard, D. S.: Computer mapping: The SYMAP interpolation algorithm, in Spatial Statistics and Models, edited by G. Gayle and C. Willmott, pp. 133–145, Springer Netherlands., 1984.

Sinclair, M. R.: A diagnostic model for estimating orographic precipitation, J. Appl. Meteorol., 33(10), 1163–1175, doi:10.1175/1520-0450(1994)033<1163:ADMFEO>2.0.CO;2, 1994.

Sinclair, M. R., Wratt, D. S., Henderson, R. D. and Gray, W. R.: Factors affecting the distribution and spillover of precipitation in the Southern Alps of New Zealand - A case study, J. Appl. Meteorol., 36(5), 428–442, doi:10.1175/1520-0450(1997)036<0428:FATDAS>2.0.CO;2, 1997.

Smith, R. B.: The influence of mountains on the atmosphere, Adv. Geophys., 21, 87–230, 1979.

Sokol, Z. and Bližnák, V.: Areal distribution and precipitation--altitude relationship of heavy short-term precipitation in the Czech Republic in the warm part of the year, Atmos. Res., 94(4), 652–662, 2009.

Steiner, M., Bousquet, O., Houze, R. A., Smull, B. F. and Mancin, M.: Airflow within major Alpine river valleys under heavy rainfall, Q. J. R. Meteorol. Soc., 129(588), 411–431, doi:10.1256/qj.02.08, 2003.

Tadić Perčec, M.: Gridded Croatian climatology for 1961–1990, Theor. Appl. Climatol., 102(1-2), 87–103, 2010.

Tobin, C., Nicotina, L., Parlange, M. B., Berne, A. and Rinaldo, A.: Improved interpolation of meteorological forcings for hydrologic applications in a Swiss Alpine region, J. Hydrol., 401(1), 77–89, doi:10.1016/j.jhydrol.2011.02.010, 2011.

Tveito, O. E., Bjørdal, I., Skjelvåg, A. O. and Aune, B.: A GIS-based agro-ecological decision system based on gridded climatology, Meteorol. Appl., 12(1), 57–68, 2005.

Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B. J., Isaksen, L., Janssen, P., Jenne, R., McNally, A. P., Mahfouf, J. F., Morcrette, J. J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P. and Woollen, J.: The ERA-40 re-analysis, Q. J. R. Meteorol. Soc., 131(612), 2961–3012, doi:10.1256/qj.04.176, 2005.

Viviroli, D., Durr, H. H., Messerli, B., Meybeck, M. and Weingartner, R.: Mountains of the world, water towers for humanity: Typology, mapping, and global significance, Water Resour. Res., 43(7), W07447, doi:10.1029/2006WR005653, 2007.

Weingartner, R., Viviroli, D. and Schaedler, B.: Water resources in mountain regions: a methodological approach to assess the water balance in a highland-lowland-system, Hydrol. Process., 21(5), 578–585, 2007.

Weusthoff, T.: Weather Type Classification at MeteoSwiss—Introduction of New Automatic Classifications Schemes, Arbeitsberichte der MeteoSchweiz, 235, 1–47, 2011.

Widmann, M. and Bretherton, C. S.: Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States, J. Clim., 13(11), 1936–1950, doi:10.1175/1520-0442(2000)013<1936:VOMPIT>2.0.CO;2, 2000.

Yarnal, B.: Synoptic Climatology In Environmental Analysis, Behaven Press, London, UK., 1993.

Yates, D., Purkey, D., Sieber, J., Huber-Lee, A., Galbraith, H., West, J., Herrod-Julius, S., Young, C., Joyce, B. and Rayej, M.: Climate Driven Water Resources Model of the Sacramento Basin, California, J. Water Resour. Plan. Manag., 135(5), 303–313, doi:10.1061/(ASCE)0733-9496(2009)135:5(303), 2009.

Zangl, G., Aulehner, D., Wastl, C. and Pfeiffer, A.: Small-scale precipitation variability in the Alps: Climatology in comparison with semi-idealized numerical simulations, Q. J. R. Meteorol. Soc., 134(636), 1865–1880, doi:10.1002/qj.311, 2008.

# Tables and figures

| Acronym | Interpolation method | Predictors | Number of predictors |
|---|---|---|---|
| LM1e | Multi-linear regression. Topographic predictors only.<br><br>Spatial autocorrelation neglected. | Elevation only | 1 |
| LM3e | | • Elevation ('e') at 3 spatial scales (75 km,25 km,1 km). | 3 |
| LM9eg | | • Elevation ('e') at 3 spatial scales.<br>• Topographic gradient ('g') at 3 spatial scales.<br>• Two sets of scales:<br>i) 75 km,25 km,1 km.<br>ii) 10 km, 5 km, 1 km. | 9 |
| OK | Ordinary kriging (OK). Spatial autocorrelation only, no topographic predictors. | - | 0 |
| KED1e | Kriging with external drift (KED). Topographic predictors and spatial autocorrelation. Stratification by season. | Elevation ('e') only | 1 |
| KED3e | | • Elevation ('e') at 3 spatial scales (75 km,25 km,1 km). | 3 |
| KED9eg | | • Elevation ('e') at 3 spatial scales.<br>• Topographic gradient ('g') at 3 spatial scales .<br>• Two sets of scales:<br>i) 75 km,25 km,1 km.<br>ii) 10 km, 5 km, 1 km. | 9 |
| KED1e+ | Kriging with external drift (KED).<br><br>Season stratified by circulation types ('+'). | Elevation ('e') only | 1 |
| KED6ev+ | | • Elevation ('e') at 3 spatial scales.<br>• Wind-aligned topographic gradient ('v') at 3 spatial scales.<br>• Set of spatial scales:<br>75 km,25 km,1 km. | 6 |
| KED9eg+ | | • Elevation ('e') at 3 spatial scales.<br>• Topographic gradient ('g') at 3 spatial scales.<br>• Set of spatial scales: 75 km, 25 km, 1 km. | 9 |

Table 1: Interpolation experiments conducted for long-term seasonal mean precipitation. Interpolation method, predictors used and the total number of predictors included.

| Acronym | Interpolation method | Background field |
|---|---|---|
| $OK(\cdot)$ | Ordinary kriging (OK) of daily precipitation (square root transformed) | none |
| *KED*(KED1e) | Kriging with external drift (KED) | KED1e, long-term seasonal mean derived with elevation (1 km) as predictor |
| *KED*(KED1e+) | KED | KED1e+, long-term seasonal mean over days of circulation type, derived with elevation (1 km) as predictor |
| *SYMAP*(PRISM) | SYMAP | PRISM, long-term seasonal mean derived with PRISM |
| *KED*(OK) | KED | OK (long-term seasonal mean derived with OK, no topographic predictors) |

Table 2: Interpolation experiments conducted for daily precipitation. The name of a scheme is a combination of the name of the daily scheme and the background field used.

| | LM1e | LM3e | LM9eg |
|---|---|---|---|
| DJF | 0.01 | 0.42 | 0.59 |
| MAM | 0.05 | 0.52 | 0.66 |
| JJA | 0.1 | 0.51 | 0.73 |
| SON | 0.1 | 0.44 | 0.57 |

Table 3: Adjusted $R^2$ for three linear models (see Table 1) and for each season.

| | Winter | Spring | Summer | Fall |
|---|---|---|---|---|
| LM1e | 0.971 | 0.993 | 1.000 | 1.000 |
| LM9eg (10 km, 5 km, 1 km) | 0.981 | 0.997 | 1.004 | 1.003 |
| LM3e (75 km, 25 km, 1 km) | 0.976 | 0.996 | 1.002 | 1.002 |
| LM9e (75 km, 25 km, 1 km) | 0.979 | 0.997 | 1.003 | 1.001 |
| OK | 0.995 | 1.004 | 1.007 | 1.007 |
| KED1e | 0.989 | 1.002 | 1.006 | 1.005 |
| KED9eg (10 km, 5 km, 1 km) | 0.990 | 1.003 | 1.008 | 1.006 |
| KED3e (75 km, 25 km, 1 km) | 0.989 | 1.002 | 1.006 | 1.005 |
| KED9e (75 km, 25 km, 1 km) | 0.989 | 1.002 | 1.006 | 1.005 |

Table 4: Relative bias B calculated over all stations for different seasons using different interpolation models (see Table 1 for model acronyms).

| | Winter | Spring | Summer | Fall |
|---|---|---|---|---|
| LM1e | 1 | 0.972 | 0.931 | 0.929 |
| LM9eg (10 km, 5 km, 1 km) | 0.749 | 0.717 | 0.641 | 0.787 |
| LM3e (75 km, 25 km, 1 km) | 0.571 | 0.482 | 0.475 | 0.570 |
| LM9e (75 km, 25 km, 1 km) | 0.438 | 0.366 | 0.278 | 0.452 |
| OK | 0.217 | 0.237 | 0.104 | 0.173 |
| KED1e | 0.114 | 0.111 | 0.066 | 0.099 |
| KED9eg (10 km, 5 km, 1 km) | 0.109 | 0.105 | 0.062 | 0.098 |
| KED3e (75 km, 25 km, 1 km) | 0.114 | 0.111 | 0.066 | 0.099 |
| KED9e (75 km, 25 km, 1 km) | 0.109 | 0.101 | 0.063 | 0.095 |

Table 5: Relative mean root-transformed error E calculated over all stations for different seasons using different interpolation models (see Table 1 for model acronyms).

| | Winter | Spring | Summer | Fall |
|---|---|---|---|---|
| KED1e+ | 1 | 0.998 | 1.005 | 1 |
| KED6ev+ | 1 | 0.999 | 1.005 | 1 |
| KED9eg+ | 1 | 0.999 | 1.005 | 1 |
| KED9eg | 0.989 | 1.002 | 1.006 | 1.005 |

Table 6: Relative bias B calculated over all stations for different seasons using different interpolation models (see Table 1 for model acronyms).

| | Winter | Spring | Summer | Fall |
|---|---|---|---|---|
| KED1e+ | 0.113 | 0.104 | 0.062 | 0.092 |
| KED6ev+ | 0.105 | 0.095 | 0.061 | 0.089 |
| KED9eg+ | 0.106 | 0.095 | 0.059 | 0.090 |
| KED9eg | 0.109 | 0.101 | 0.063 | 0.096 |

Table 7: Relative mean root-transformed error E calculated over all stations for different seasons using different interpolation models (see Table 1 for model acronyms).

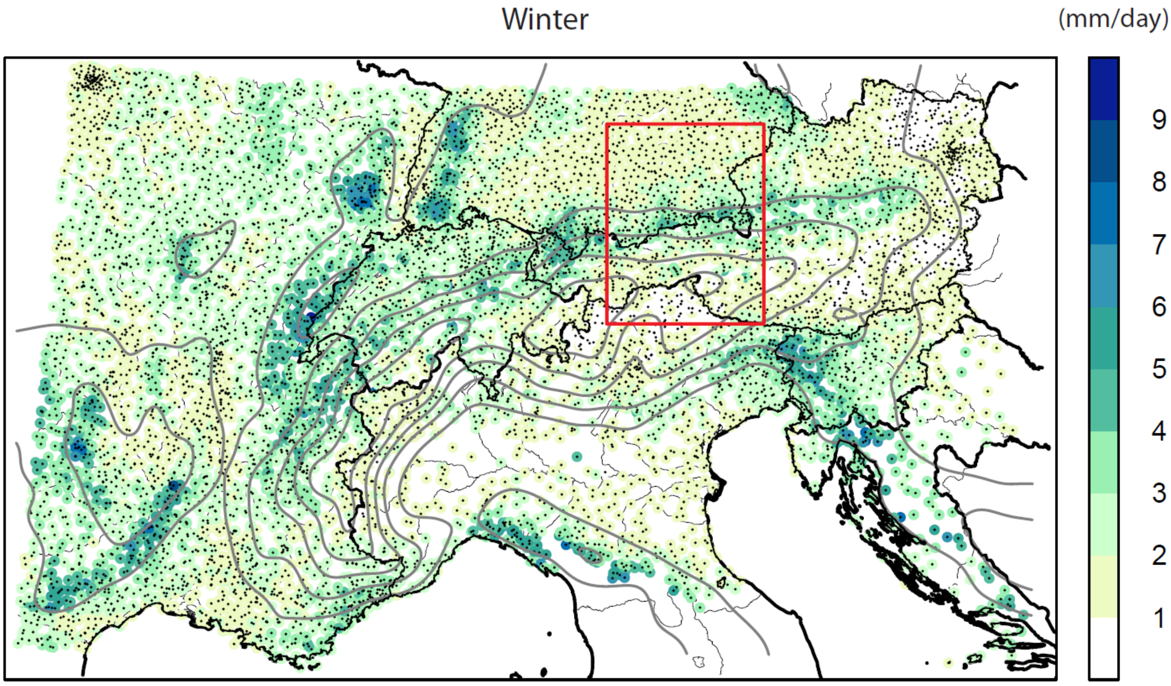


Figure 1: Map of long-term mean winter precipitation (mm/day) over the Alpine domain at station locations (dots) for the period 1971-2008. The grey contour lines indicate the Alpine relief (400 m levels) and the red frame delimits the region in which the interpolation methods are tested.

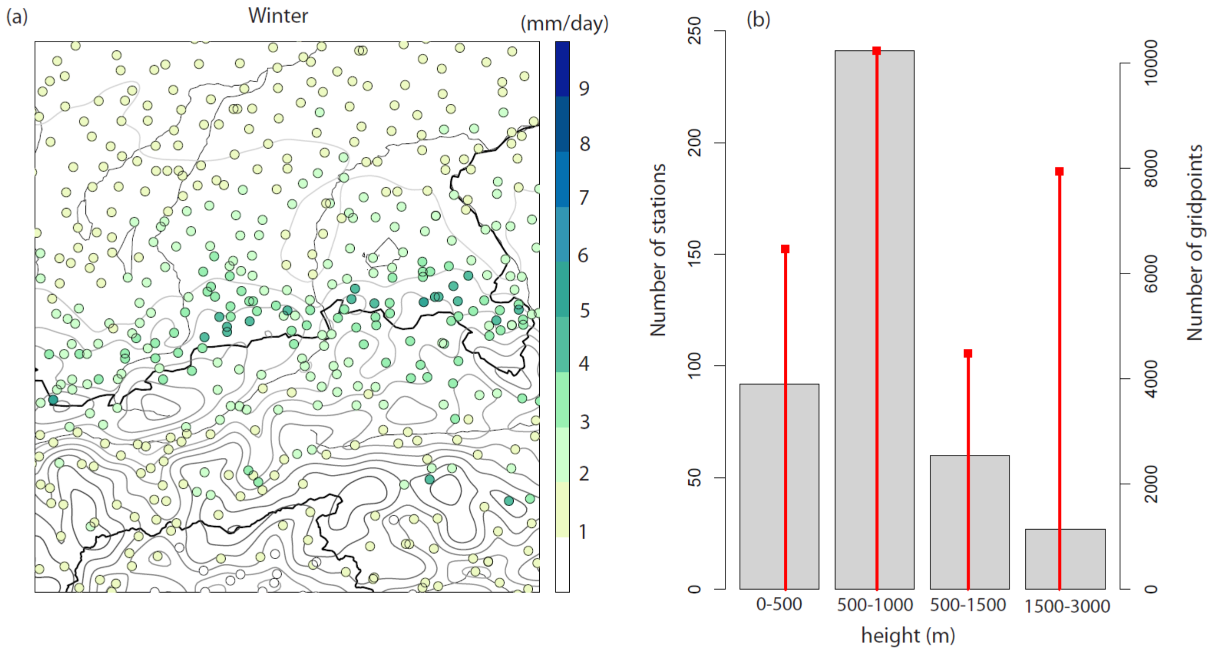


Figure 2: (a) Map of the study domain, a section of the Alpine ridge (see also Fig.1). The topography is indicated by grey-shaded contour lines (spacing 250 m). The station network is indicated by colored circles, representing long-term mean winter (DJF) precipitation in mm/day. The thick black line represents the national borders between Germany (top), Austria (middle) and Italy (bottom). (b) Barplot of the distribution with height (x-axis, mMSL) of the number of stations (grey, left y-axis) and the number of grid-points in a 1 km DEM (red, right y-axis).

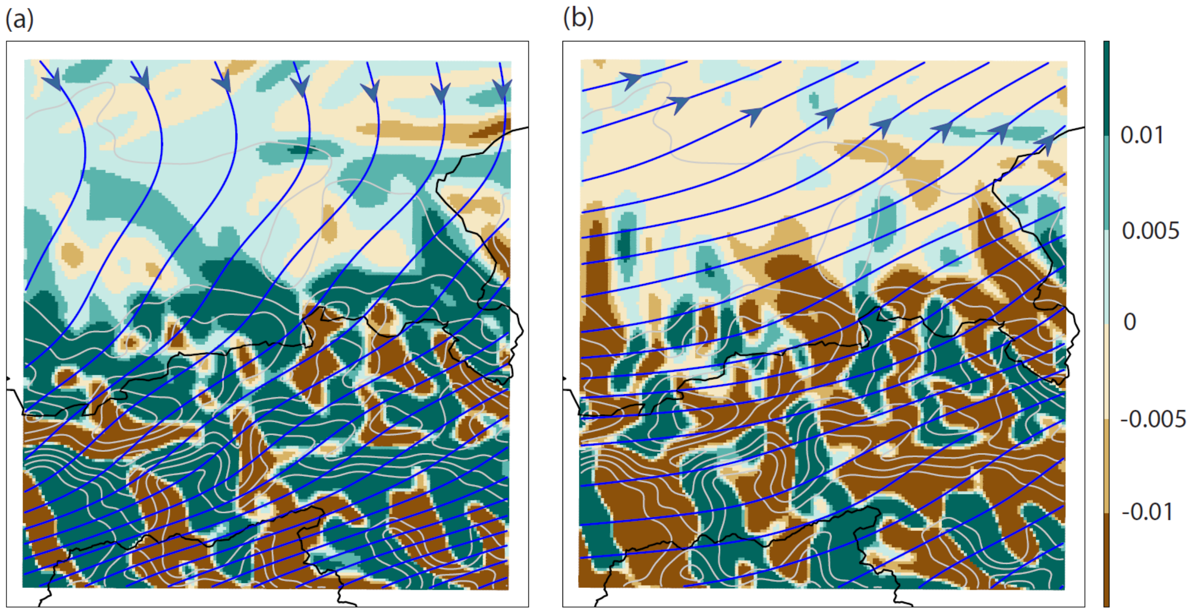


Figure 3: Illustration of $G_w$, the wind-aligned gradient, for two classes of the PCACA9 circulation type classification: (a) North-Easterly flow in the summer, and (b) South-Westerly flow in the autumn. The example fields are valid for a smoothing scale of 5 km. The topography is depicted in grey lines (spacing 250 m) and the streamlines of the geostrophic wind are shown by the blue curves.

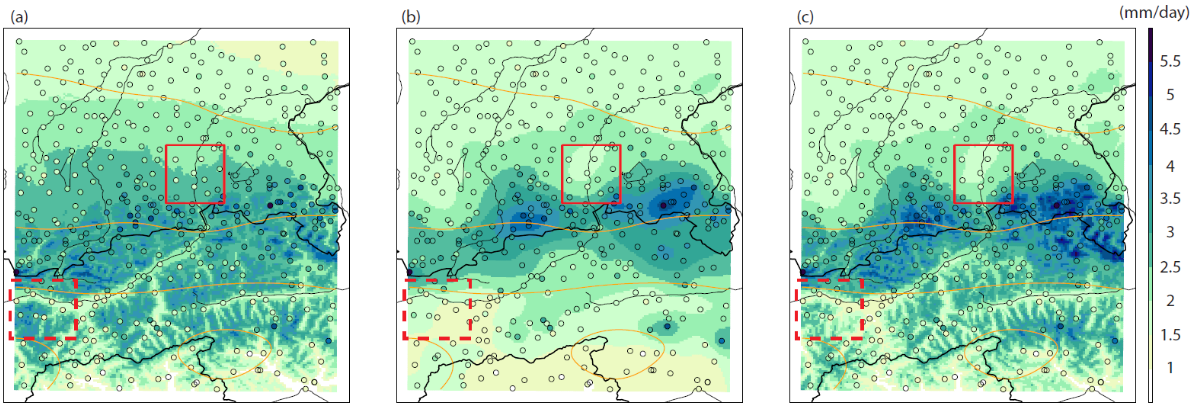


Figure 4: Distribution of DJF long-term mean precipitation (mm per day) as estimated by (a) a multi-linear regression using as predictors elevation and gradients at three spatial scales (75 km, 25 km and 1 km, LM9eg), (b) ordinary kriging (OK, no topographic predictors), (c) kriging with external drift using the same predictors an in (a). Color-filled circles represent observations at rain-gauge stations. Red squares denote areas mentioned in the text. The topography is depicted in orange lines (spacing 500 m).

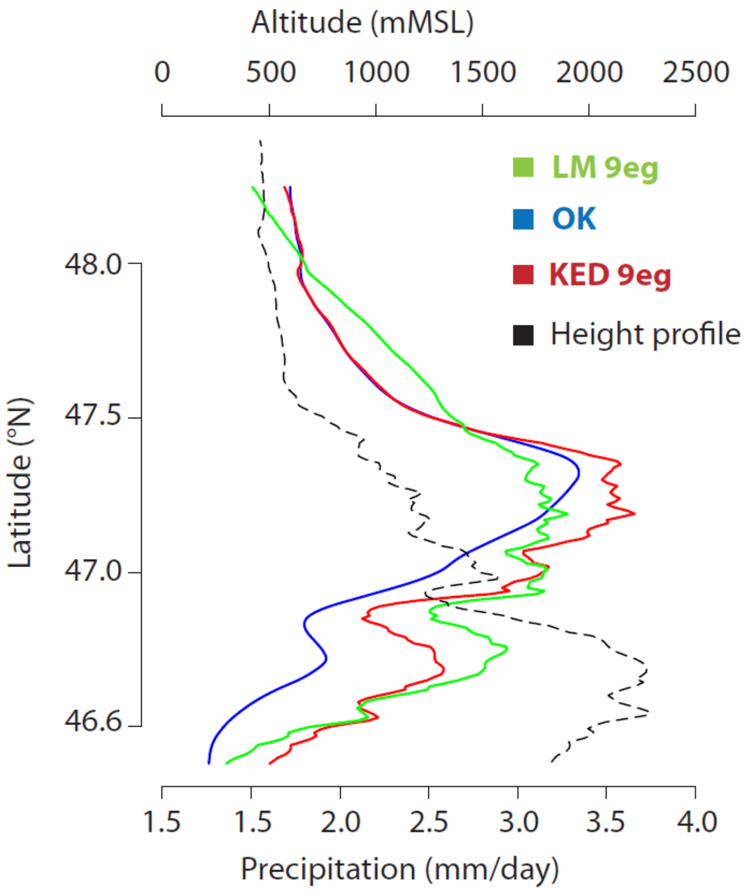


Figure 5: North-south precipitation profile as estimated by the three interpolation methods LM9eg, OK, KED9eg (see Table 1). DJF long-term mean precipitation (lower x-axis, mm per day) as a function of latitude (y-axis, degrees North). The dashed line indicates the height profile (upper x-axis, m) as function of the latitude.

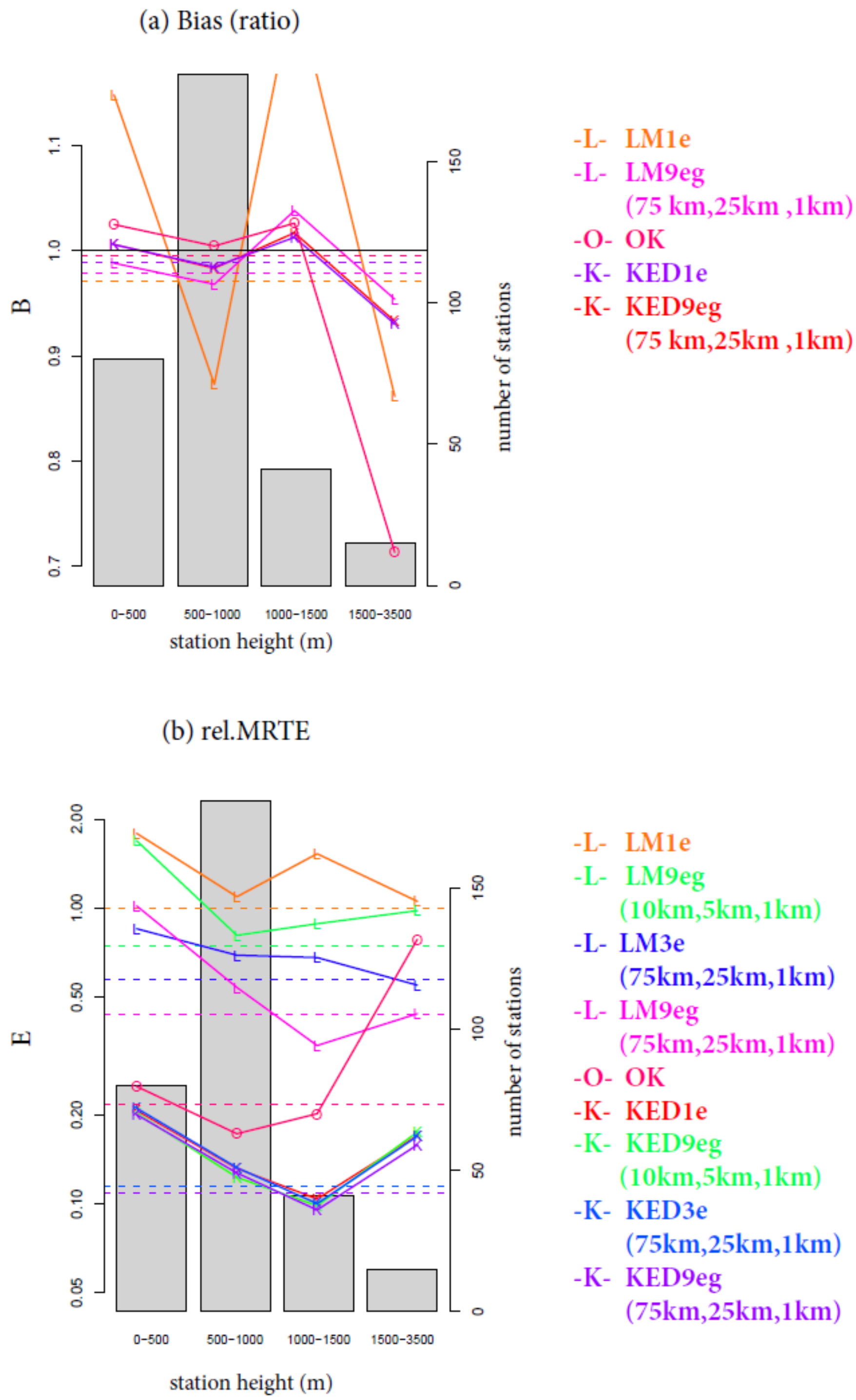


Figure 6: Error statistics for the interpolation of mean DJF precipitation using different interpolation models (see Table 1 for model acronyms). Relative bias B (dimensionless, Eq. 3, panel a) and relative mean root-transformed error E (dimensionless, Eq. 4, panel b, log-scale) of a leave-one-out cross-validation. Results are shown for four elevation classes. Horizontal dashed lines represent the scores over all stations. The vertical bars represent the number of stations per elevation class (right axes).

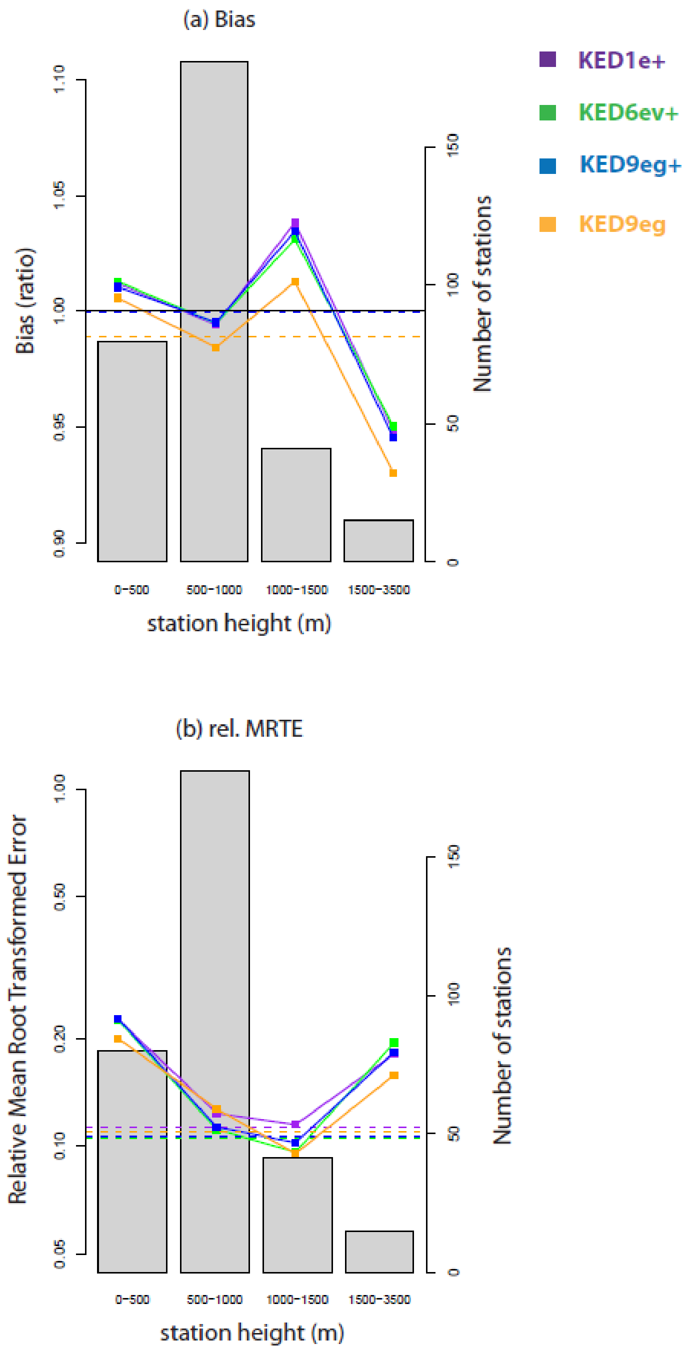


Figure 7: Error statistics for the interpolation of mean DJF precipitation using interpolation models that utilize information from a circulation classification (see Table 1 for model acronyms). Relative bias B (dimensionless, Eq. 3, panel a) and relative mean root-transformed error E (dimensionless, Eq. 4, panel b, log-scale) of a leave-one-out cross-validation. Results are shown for four elevation classes. Horizontal dashed lines represent the scores over all stations. The vertical bars represent the number of stations per elevation class (right axes).

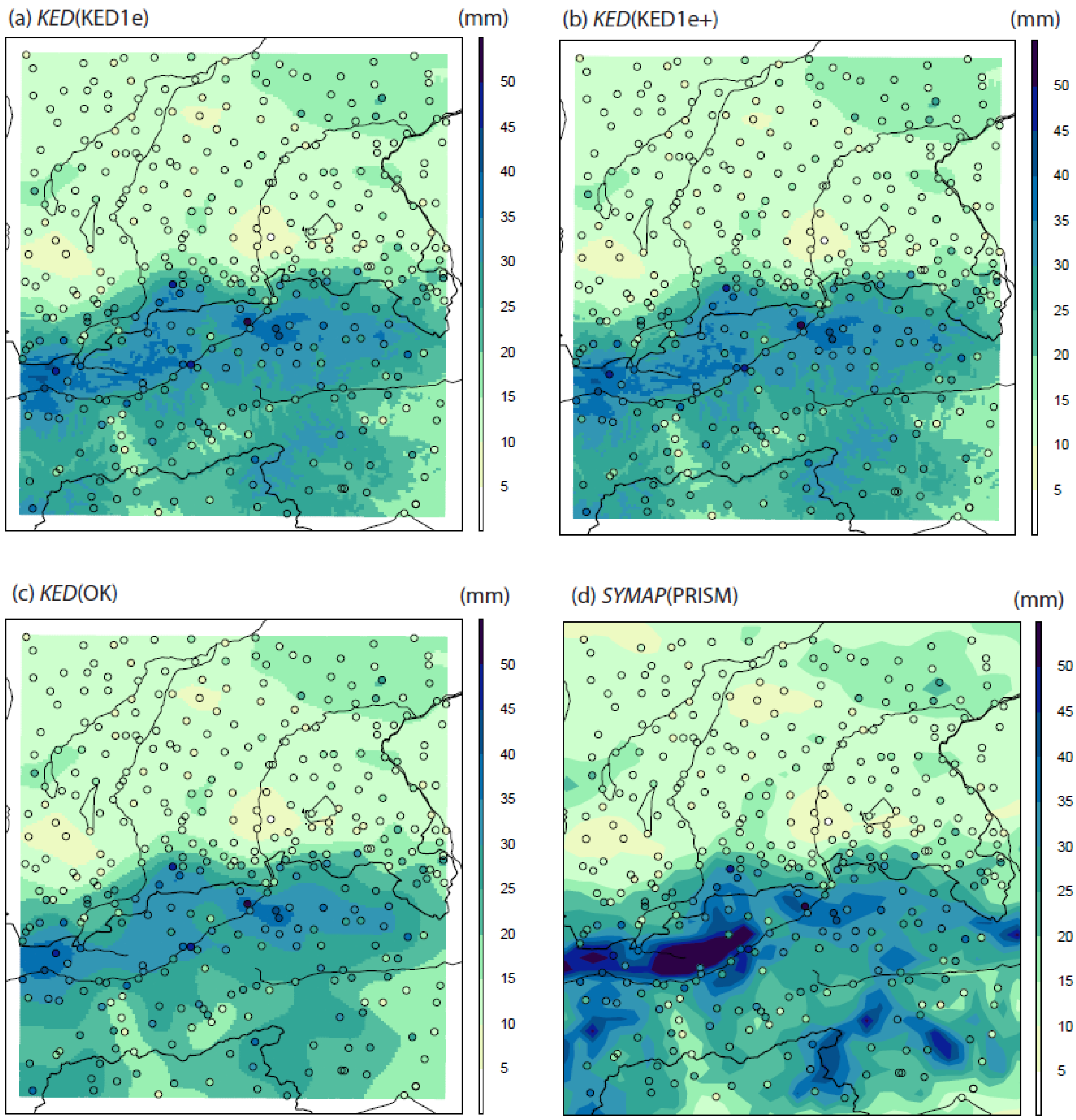


Figure 8: Daily precipitation total (mm) for February 13 1990, as derived by the daily interpolation methods investigated in this study. (a) *KED*(KED1e), (b) *KED*(KED1e+), (c) *KED*(OK), (d) *SYMAP*(PRISM), see Table 2 for a description of the method acronyms. The fields for panels (a-c) were produced on a 1 km grid, that of panel (d) on a 5 km grid.

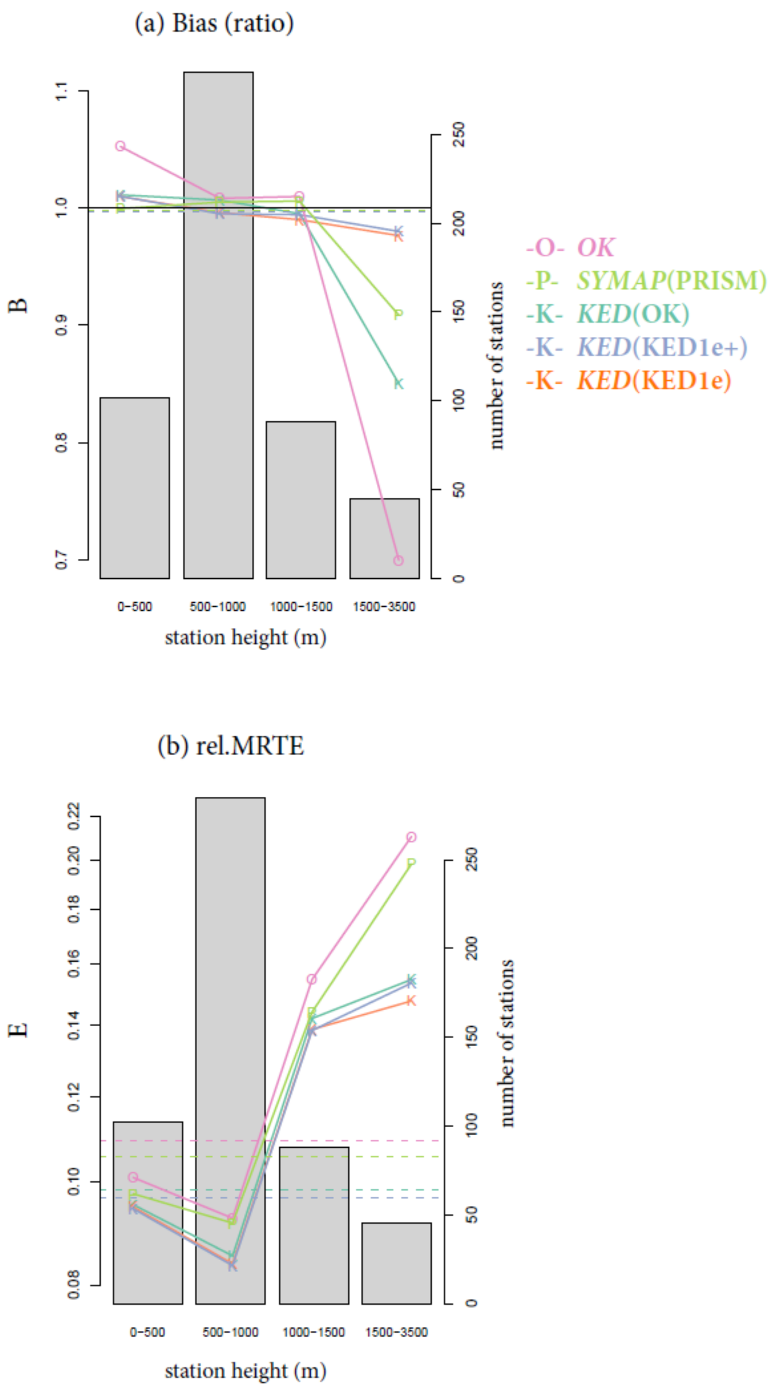


Figure 9: Error statistics for the interpolation of daily precipitation in winter (DJF, 1971-2008) using the interpolation models of Table 2 (see also section 3). Relative bias B (dimensionless, Eq. 3, panel a) and relative mean root-transformed error E (dimensionless, Eq. 4, panel b, log-scale) of a leave-one-out cross-validation. Results are shown for four elevation classes. Horizontal dashed lines represent the scores over all stations. The vertical bars represent the number of stations per elevation class (right axes).