Change points in predictors–predictand relationships within the scope of statistical downscaling

Elke Hertig,* Christian Merkenschlager and Jucundus Jacobeit

Institute of Geography, University of Augsburg, Augsburg, Germany

ABSTRACT: A statistical downscaling approach allowing for change points in the relationships between atmospheric predictors and local precipitation is introduced. Change point analysis within generalized linear models, and change points in the predictor characteristics were used to develop a change point statistical downscaling approach. The approach is illustrated for station-based winter precipitation in the Mediterranean area.

In this study, 94 stations were considered. The change point analysis yielded 37 stations with robust change points in the predictors–precipitation relationships. An added value regarding statistical model performance of the change point approach compared to the use of statistical models without change points was observed for 15 out of the 37 stations. In the projections under enhanced greenhouse gas forcing the application of the change point approach for the 15 stations affects the downscaled precipitation amounts, with significant differences compared to the application of downscaling models without change points for about one third of these stations.

Under RCP4.5 and RCP8.5 scenario conditions, mainly increases of winter precipitation are projected until the end of the 21st century for parts of the western and northern Mediterranean area, whereas the north-eastern and eastern Mediterranean areas are affected by decreases.

KEY WORDS   statistical downscaling; generalized linear models; change point analysis; climate variability; climate change; Mediterranean precipitation

1. Introduction

Different downscaling techniques exist to obtain regional to local climate change information from large-scale general circulation model (GCM) output. A widely used technique is Perfect-prog (PP) statistical downscaling, an approach which typically makes use of observation-based relationships of the large-scale atmospheric circulation (predictor) and regional climate variables (predictands) through time [for an overview of statistical downscaling approaches, see Maraun *et al.* (2010)]. In the scope of transferring these relationships to other time periods (validation and future projection periods) than the one used for model set-up (calibration period), it is assumed that the predictor–predictand relationships remain the same through time.

On the other hand, there is substantial evidence that circulation–climate relationships are varying with time. This characteristic emerges from substantial modifications of the atmospheric circulation, which lead to significant changes of the relationships with regional climate. The changes of the atmospheric circulation are measurable, for instance, as changes in the major large-scale modes of climate variability, such as the El Niño-Southern Oscillation

(ENSO) (Diaz *et al.*, 2001; Power and Smith, 2007) and the North Atlantic Oscillation (NAO) (Jung *et al.*, 2003; Hurrell and Deser, 2010). Prominent examples of abrupt shifts in climate are for instance the climate shift over the tropical Pacific at the end of the 1970s associated with the ENSO phenomenon (Miller *et al.*, 1994; Trenberth and Hoar, 1997) and the climate shifts over the North Pacific area (Rodionov, 2004; Overland *et al.*, 2008). Besides such abrupt changes also more gradual changes have been observed such as the pronounced positive trend of the NAO from the 1960s to the 1990s (Hurrell and Deser, 2010). Overall, changes have been observed in the frequencies and phases of atmospheric patterns, the strength and the spatial location of the pattern-related anomaly centres, as well as the within-pattern characteristics, such as the thermal and thermo-dynamic properties of the patterns. A comprehensive review of non-stationarities in climate variability of the last century is given in Hertig *et al.* (2015).

In the scope of future climate change due to anthropogenic increases of greenhouse gas forcing, it is in all probability that these increases change the intensity and geographical patterns of the atmospheric circulation influence on climate. For instance, Ullmann *et al.* (2013) identified in CMIP5 (fifth Coupled Model Intercomparison Project) future projections an eastward shift of the low-pressure system in the negative phase of the NAO and an eastward extension of the Azores High towards the Mediterranean area for the positive phase of the NAO

in winter (October–March). Hertig and Jacobeit (2014a) found that the specifications of the NAO negative phase in winter (December–February) are highly varying with time in the analysed CMIP5 historical and scenario runs, which is not the case in the reanalysis data. Ullmann *et al.* (2013) also detected significant intra-pattern baroclinic changes under rapid increases in the amount of atmospheric greenhouse gases. All these changes modify regional Mediterranean rainfall patterns (Ullmann *et al.*, 2013; Hertig and Jacobeit, 2014b).

In the context of PP statistical downscaling, it is of interest if and how the observed and modelled non-stationarities affect the relationships between the large-scale circulation and local/regional climate in specific downscaling set-ups. For instance, Hertig and Jacobeit (2013) used a combined circulation-based and generalized linear model (GLM)-based downscaling approach to assess daily precipitation in the Mediterranean area. They showed that observed changes in the frequency of circulation patterns are not the cause of non-stationarity in statistical model performance because changing frequencies are accommodated by the statistical downscaling approach. In contrast, changing positions of the centres of action such as it has been observed for instance for the NAO (Jung *et al.*, 2003; Moore *et al.*, 2013; Raible *et al.*, 2014) can be more challenging. The spatial changes can significantly alter correlations to regional climate (Goodess and Jones, 2002; Vicente-Serrano and López-Moreno, 2008). The physical changes underlying the changed predictor–predictand relationships due to changes in the associated dynamical and thermo-dynamical features form a major obstacle in the statistical modelling using linear methods. Even a careful predictor selection might not cope with such non-stationarities. But indeed, a careful analysis of non-stationarities and their origin is required (Wilby, 1998), as they could be due to either changes in the climate system structure(s) or simply be caused by an incomplete set of predictor variables or inadequate calibration periods. A practicable way to allow for non-stationarities in statistical downscaling is to adapt the downscaling models to observed non-stationarities and to selectively use these non-stationary models for the projections (Hertig and Jacobeit, 2014b). Under the assumption that during the 21st century climate will change in the range of observed natural variability (an assumption intrinsic to statistical downscaling and also to GCMs due to parameterizations, Zorita and von Storch, 1999) non-stationarities in observed and future climate are explicitly accounted for.

A number of methods have been developed to detect non-stationarities in the observational period. Common approaches to detect non-stationarities in the relationships between the large-scale circulation and regional climate characteristics include correlation or regression analysis over sliding intervals (e.g. Beranová and Huth, 2008; Vicente-Serrano and López-Moreno, 2008; Hertig and Jacobeit, 2013). Change point models are applied to detect abrupt shifts in climate time series. They are commonly based on the Student's or Mann–Kendall tests (e.g.

Rodionov, 2004) or on the use of information criterions within regression models (e.g. Beaulieu *et al.*, 2012). For an overview of different change point techniques for climate data, see for instance Reeves *et al.* (2007).

In the present contribution, we consider change points in particular regression models known as GLMs. For an introduction on the use of GLMs in the context of analysing climate variability, see Chandler (2005). We allow for multiple changes in the model's structure, determined by specific points in the time series, to take account of varying relationships between the large-scale atmospheric predictors and the local predictand. Furthermore, composites of the atmospheric characteristics before and after a change point are used within a cumulative sums approach to identify the change points in the predictor data only. Results of both approaches are considered to define robust change points and two different reanalysis data sets are applied to avoid spurious change points due to inhomogeneity in a data set. The two approaches are subsequently coupled for the statistical downscaling assessments. The change point statistical downscaling models are analysed regarding their relevance in terms of an improvement in statistical model performance. Thus, change points which can be identified in both approaches and which impact noticeably on the GLM model performance are considered as robust and relevant change points. Finally, the significance of using change point models in the projections until the end of the 21st century is assessed. Overall, the approach is used with the objective to enhance statistical downscaling model performance through the explicit consideration of change points in the predictors–predictand relationships.

This study focuses on station-based precipitation in the Mediterranean area as predictand. The Mediterranean area is chosen, because it is regarded as a 'climate change hot-spot' (Giorgi, 2006), being located in the transitional zone between the mid-latitude westerlies and the subtropical high-pressure belt. Section 2 describes the data used in this study and Section 3 presents the methodology. In Section 4, results are presented and discussed. Conclusions are drawn in Section 5.

2. Data

2.1. Predictand: station data in the Mediterranean area

Daily station data for the Mediterranean area have been collected from the European Climate Assessment & Dataset (Klein Tank *et al.*, 2002), from the GLOWA Jordan River Project (Global Change and the Hydrological Cycle, Kunstmann *et al.*, 2006), and from the EMULATE project (European and North Atlantic daily to multidecadal climate variability, Moberg *et al.*, 2006). Unfortunately, there are very little data available over the southern Mediterranean area (notably Northern Africa) as well as over the north-eastern Mediterranean area (particularly Turkey), but for the other regions an acceptable spatial coverage has been achieved.

After testing completeness (Moberg *et al.*, 2006) and homogeneity (Alexandersson, 1986; Wijngaard *et al.*,

2003) of the daily precipitation data, 94 stations could be retained for subsequent analyses. For each station, we calculated monthly precipitation sums from the daily data for the winter months December–February. Depending on the station, the time series start between 1950 and 1961 and end between 1990 and 2008. The mean time series length across all stations is 53 winters. The minimum (maximum) time series length is 40 (58) winters. This yields sample sizes of 120–174 winter months.

Statistical downscaling models were set up for each of the 94 stations separately, even though the use of local-scale precipitation may fit a variable amount of noise in the statistical models. This was done to capture local-scale specifics of this climatologically highly heterogeneous region. In the context of climate change impact studies, such as small-scale hydrological modelling, information is often required on the local scale. However, within the interpretation of results in Section 4 the inter-station variability arising from the local-scale models will be taken into consideration. A description of the precipitation characteristics in the Mediterranean area in the wet season and in general can be found for instance in Xoplaki *et al.* (2004) and Lionello *et al.* (2012), respectively.

2.2. Predictor data

In order to avoid potential influences from inhomogeneities in the reanalysis data, predictor data were taken from two independent reanalysis data sets – NCEP reanalysis (Kalnay *et al.*, 1996; Kistler *et al.*, 2001), which uses both, surface observations as well as satellite data, as well as ERA-20C reanalysis (Poli *et al.*, 2013; Stickler *et al.*, 2014), which assimilates surface observations only, i.e. surface pressure, atmospheric and oceanic temperatures and near-surface wind. ERA-20C was selected to include a reanalysis with no potential inhomogeneity due to the assimilation of satellite data. A first performance evaluation for lower to middle tropospheric pressure and temperature shows an adequate quality for the Northern Hemisphere extratropics (Poli *et al.*, 2013).

We selected monthly mean geopotential heights of the 700 hPa level for the period December 1950–February 2010. The geopotential heights can be regarded as representative predictors for the atmospheric circulation, which show a strong and physically meaningful relationship to Mediterranean precipitation during wintertime (Tatli *et al.*, 2004; Hertig and Jacobeit, 2013). We selected the domain 20°–70°N, 70°W–50°E to represent the North Atlantic-European-Mediterranean area, incorporating the major large-scale modes of variability such as the NAO. Also the more regional domain 30°–50°N, 10°W–40°E centred over the Mediterranean area was chosen in order to capture region-specific atmospheric conditions.

In general, the pressure-related variables are predictors, which are reliably represented in GCMs on the large scale and are strongly connected with regional climate and climate change. But also other variables describing the thermal and thermo-dynamic conditions of the atmosphere play an important role for precipitation in the

Mediterranean area (Hertig and Jacobeit, 2008; Hertig and Jacobeit, 2013; Hertig *et al.*, 2014). Generally, these variables have a lower quality in the GCMs, but their patterns of change can be very relevant in the scope of regional climate change. In addition to 700 hPa geopotential heights, specific and relative humidity, zonal and meridional wind components, vertical velocity and air temperature, all from the 850 hPa level, were chosen to serve as potential predictors for Mediterranean precipitation. The fields of these variables were cut to the smaller domain 30°–50°N, 10°W–40°E centred over the Mediterranean area.

Model data were taken from a three-member MPI-ESM-LR ensemble (Max Planck Institute Earth System Model running on low resolution grid). Historical, RCP4.5 scenario and RCP8.5 scenario (Van Vuuren *et al.*, 2011) runs performed for CMIP5 were downloaded from the CMIP5 archive (https://pcmdi.llnl.gov/projects/cmip5/). We used the period 1950–2005 of the historical runs and the period 2006–2100 of the scenario runs. The horizontal resolution of the model output data (1.875°) was fitted to that of the reanalysis data (NCEP reanalysis 2.5° × 2.5°, ERA-20C 2° × 2°) using ordinary kriging.

3. Methodology

Predictors were processed by means of principal component analysis (PCA, Section 3.1). Change point analysis was done using two different approaches. One approach is based on change point generalized linear models (CPGLMs, Section 3.2.1.). The CPGLM approach looks for changes in the predictors–precipitation relationships within the regression models. The other approach comprises change points in cumulative sums (CPCUMSUM) curves of composite correlations (Section 3.2.2.). The CPCUMSUM approach is applied to independently identify the change points in the predictor data only. Both approaches are applied using two independent reanalysis data sets (NCEP and ERA-20C). Thus, four configurations (two approaches multiplied by two data sets) are consulted to find change points. Only change points, which can be found in all four configurations, are considered as robust (further on referred to as robust change points). Subsequently, the change points are tested for their relevance within the framework of statistical downscaling using a specific cross-validation set-up (yielding robust and relevant change points, Section 3.3). Finally, the two approaches are coupled for local climate change projections and the impact of the change point statistical downscaling approach on the projections for the 21st century is assessed (Section 3.4). A schematic presentation of the methodology used in this study can be found in Figure 1.

3.1. Principal component analysis of predictor fields

In the scope of studying climate variability and climate change, large-scale climate phenomena are receiving considerable attention and are increasingly well simulated by climate models. Not surprisingly, a whole chapter of the
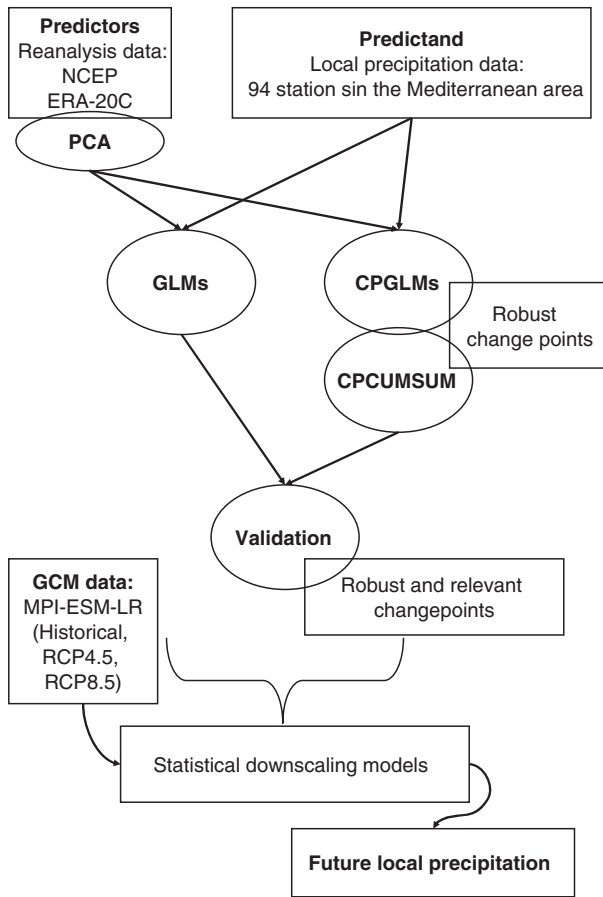
Figure 1. Schematic overview of the methodology used in this study.

latest IPCC report is devoted to major climate phenomena and their relevance for future regional climate change (IPCC AR5, chapter 14, Christensen *et al.*, 2013). Therefore we consider the well-known large-scale modes of variability in the North Atlantic sector [i.e. NAO, East Atlantic (EA) pattern, Eurasia-1 (EU1) pattern and Eurasia-2 (EU2) pattern, see Barnston and Livezey (1987)] which play a substantial role for Mediterranean precipitation in winter (Quadrelli *et al.*, 2001; Dünkeloh and Jacobeit, 2003; Hertig and Jacobeit, 2013). To obtain the large-scale modes of variability, we applied S-mode VARIMAX-rotated PCA (e.g. Preisendorfer, 1988; von Storch and Zwiers, 1999) to the area-weighted 700 hPa geopotential heights fields in the North Atlantic-European-Mediterranean domain. PCA is based on standardized monthly anomaly fields for winter in the period December 1950 to February 2010. Four PCs were kept with 66.3% explained variance (EV). PCA was also applied to the smaller Mediterranean domain in order to reduce the regional fields to specific centres of variation. A centre of variation depicts a region consisting of grid boxes with high PC loadings on a particular PC and thus having similar temporal variations represented by the PC score. The centres of variation in the small domain are used to describe regional atmospheric characteristics which are not adequately resolved by the large-scale modes. For the 700 hPa geopotential heights, five PCs were kept with 96.0% EV. 850 hPa specific humidity fields were reduced

to nine PCs (85.4% EV), 850 hPa relative humidity to eight PCs (80.9% EV), the 850 hPa zonal and meridional wind components to five PCs each with 86.9 and 86.8% EV, respectively. In case of the 850 hPa vertical velocity fields, nine PCs were obtained with 73.7% EV, and for 850 hPa air temperature, four PCs (88.9% EV) were extracted. The determination of the number of PCs to be extracted follows the approach of Philipp *et al.* (2007). It is based on the criterion that each PC has to be uniquely representative for at least one input variable.

3.2. Change point analysis

3.2.1. GLMs and CPGLMs

GLMs (e.g. Mc Cullagh and Nelder, 1989) were applied to detect relationships of the atmospheric circulation with local precipitation in the Mediterranean area. GLMs are a common choice within the context of precipitation modelling since precipitation is a variable, which is usually non-normally distributed. A maximum likelihood estimation is used to estimate model parameters for the expectation value $E(\cdot)$ for a variable Y_t at any given time t:

$$E\left(Y_t\right) = \mu_t = g^{-1}\left(\eta_t\right) = g^{-1}\sum_1^n \chi_{tj}\beta_j \qquad (1)$$

here μ_t is the mean for the probability density function (PDF) at time t, η_t is a linear predictor (i.e. a linear combination of the explanatory variables), g is the canonical link function, χ_{tj} is the value of the jth covariate for observation t, n is the total number of covariates and β_j are parameters whose values have to be estimated from the data.

We applied GLMs based on Tweedie exponential dispersion models (Dunn, 2004; Hertig and Jacobeit, 2014b). As canonical link function g the log link is used. Tweedie GLMs allow for the simultaneous modelling of the discrete (i.e. zero values, which occur in the Mediterranean precipitation time series even on a monthly time scale) and the continuous features of precipitation. The Tweedie family of distributions includes Poisson-gamma distributions and has three parameters with μ (the mean), $\varphi > 0$ (the dispersion parameter) and p (the Tweedie index parameter). The variance of the distribution is $var[Y] = \varphi\mu^p$. The index p specifies the particular distribution, special cases include the normal ($p = 0$), Poisson ($p = 1$) and gamma ($p = 2$) distributions. For $1 < p < 2$, the Tweedie family is suitable for modelling positive continuous data with exact zeros (Hasan and Dunn, 2011). Estimation of the index parameter p requires sophisticated numerical techniques and was done using a profile maximum likelihood estimate provided in the Tweedie package in R (Dunn, 2008). The index parameter p was estimated for each of the 94 precipitation time series.

As variables in the GLMs, we used the time series (PC scores) of the large-scale modes of variability and the time series of the regional centres of variation. Since our primary interest lies in the relationships of the atmospheric circulation with local precipitation, 700 hPa geopotential heights in the large-scale Atlantic-European domain and in

the smaller Mediterranean domain were always included as predictor variables in the GLMs. However, precipitation is not only governed by the atmospheric circulation, and thus further variables were considered for the modelling of local precipitation variability. For each additional variable (air temperature, relative and specific humidity, u- and v-wind components and vertical velocity), GLMs were derived using as predictors of the geopotential height PCs as well as the PCs of the additional variable. Exploratory analysis with predictor sets comprising more than two variables showed no further improvement in model skill. Within this approach, some collinearity between the predictor variables exist which might affect model performance. We applied a backward selection procedure within a cross-validation framework using 100 random samples to decide on the variables to be included in the downscaling models. The cross-validation involves the random selection of two third of all available observations at a station for calibration and the remaining one third for validation. The resulting station-specific predictor sets are subsequently used in the CPGLMs.

GLMs with multiple discontinuous change points were considered. Only changes in the linear parameters β_j were used, whereas all other quantities defining a GLM are the same for all segments. Following Zeileis *et al.* (2003) and Hofrichter (2007), GLMs were constructed with d segments recursively based on minimizing the deviance in each segment. The approach is based on the whole observational period for a station. To guarantee the estimation of the β_j, the smallest possible size for a segment is $q+1$, where q is the number of parameters in the model. We set the maximum of d to five, thus analysing models with one segment (no change point) up to five segments (four change points). In order to reduce the computational burden, a dynamic programming algorithm based on Bellman's principle of optimality (Bellman and Dreyfus, 1962) was applied. The function written in R (R Development Core Team, 2008) used to calculate the GLMs with change points can be found in the Supporting Information.

A GLM without a change point can be considered as nested in a GLM with a change point. Thus, a GLM with $d-1$ discontinuous change points $\tau = \tau_1, \ldots, \tau_{d-1}$ can be separated in $d = d_1, \ldots, d_m$ autonomous GLMs written as

$$E\left(Y_{t \leq \tau_1}\right) = \mu_{t \leq \tau_1} = g^{-1}\left(\eta_{t \leq \tau_1}\right) = g^{-1} \sum_{1}^{n} \chi_{t \leq \tau_1 j} \beta_{jd_1}$$

$$E\left(Y_{\tau_1 < t \leq \tau_2}\right) = \mu_{\tau_1 < t \leq \tau_2} = g^{-1}\left(\eta_{\tau_1 < t \leq \tau_2}\right)$$
$$= g^{-1} \sum_{1}^{n} \chi_{\tau_1 < t \leq \tau_2 j} \beta_{jd_2}$$
$$\vdots$$
$$E\left(Y_{\tau_{d_m-1} < t}\right) = \mu_{\tau_{d_m-1} < t} = g^{-1}\left(\eta_{\tau_{d_m-1} < t}\right)$$
$$= g^{-1} \sum_{1}^{n} \chi_{\tau_{d_m-1} < tj} \beta_{jd_m} \qquad (2)$$

As the deviance of a GLM with change points is additive, we analysed the difference between the deviances of the
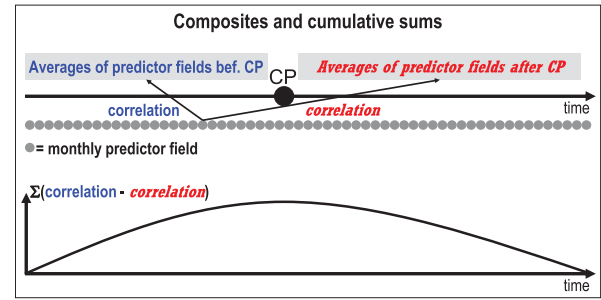

Figure 2. Schematic overview of the approach using change points in cumulative sums curves of composite correlations. CP: change point. The black arrows are exemplary and have to be thought for all monthly predictor fields (grey dots). For further details, see Section 3.2.2..

GLM without a change point and the GLMs with change points. The difference between the deviances of GLMs with different numbers of change points does not follow a χ^2-distribution (Hofrichter, 2007) and can therefore not be used to judge the significance of the change points. Thus, for each station, the model with the least deviance is selected.

3.2.2. Change point cumulative sums

Beside the CPGLMs an approach based on composites and cumulative sums was used. A schematic representation is given in Figure 2. Composites of the observed circulation characteristics before and after a change point are calculated. The location of the change point is taken from the CPGLM approach. For the positive (PC scores > 0) and the negative phases (PC scores < 0) of each predictor, composites are calculated as the mean of the PC scores-weighted original predictor fields before and after a change point, respectively. Thus, for each phase of a predictor (e.g. positive phase of the NAO), mean fields showing the characteristics before and after a change point are derived. In the next step, the individual monthly predictor fields are correlated with the corresponding composites. Cumulative sums of the difference of the correlation coefficients are computed. Positive (negative) values reflect that predictor characteristics at a specific time are better described by the conditions before (after) the change point. Maxima/minima in the curve progression of the cumulative sums are used to identify the change points in the predictor time series. For a description of the cumulative sums approach, see Page (1955, 1957).

Only change points, which can be identified in the CPGLMs as well as in the CPCUMSUM approach under the use of two different reanalysis data sets (NCEP and ERA-20C), are considered to be robust. We allowed for some tolerance regarding the time of the change points, since changes in the predictors (as identified with the CPCUMSUM approach) might lead a subsequent impact on precipitation (as captured by the CPGLMs). We decided on a maximum difference of 3 years between the CPCUMSUM and the CPGLM change points. It can be noted however that setting the tolerance time to smaller values or zero alters the results only marginally.

3.3. Validation

A common accuracy measure for continuous non-probabilistic forecasts is the mean squared error (MSE, e.g. Wilks, 2006):

$$\text{MSE} = \frac{1}{n} \sum_{k=1}^{n} (y_k - o_k)^2 \qquad (3)$$

MSE is the averaged squared difference between the model (y) and observation (o) pairs, or the reference and observation pairs, respectively. Then, the mean squared error skill score (MSESS) becomes:

$$\text{MSESS} = 1 - \frac{\text{MSE}_{\text{model}}}{\text{MSE}_{\text{reference}}} \qquad (4)$$

The MSESS was used to judge model performance with the long-term climatology serving as reference. For the presentation of the results (Figures 4 and 5), the MSESS has been translated according to Murphy (1988) into a measure of percentage improvement in accuracy by multiplying the right-hand side of Equation (4) by 100.

Change point analysis (Section 3.2) was done using the best-performing predictor combination and from the change point analysis robust change points were derived (see Section 3.2.2. for the definition of robust change points). However, the question remains whether these change points actually impact on the statistical model performance. To answer this question, we set up a specific cross-validation framework. For each station, the MSESS was assessed using conventional GLMs in a cross-validation set-up comprising 100 random calibration/validation periods sampled from the whole observational period. The resulting model performance sets the reference for the change point models, which take into account potential change points in the predictors–predictand relationships. To assess the performance of the change point models, we randomly selected from the period before the change point as well as from the period after the change point two third of the observations for calibration (with a minimum in each period set to 30 observations in view of robust calibration models), the remaining third of the observations for validation. GLMs are applied in the two different calibration periods and each model is validated with the corresponding validation set. The described procedure is also repeated 100 times and the mean MSESS of the two periods over the 100 samples is considered. In case a change point is located close to the start/end of the whole time series and thus leads to a calibration period shorter than 30 observations, only the other, longer period is evaluated. The MSESS of the conventional GLMs is compared to the MSESS of the CPGLMs in order to assess whether the change point approach yields an added value in terms of model performance compared to the application of conventional GLMs without change points.

3.4. Projections for the 21st century

CPGLMs can be used in the observational period to find change points and to describe the specific predictors–predictand relationships in the periods before and after the change points. However, for projection purposes, the question has to be answered which state of the circulation prevails in the projection periods, i.e. the question whether the circulation characteristics and the values of the additional predictors resemble more strongly the atmospheric conditions before or after a change point, and thus which GLM relationship is more appropriate. Within this context, the CPCUMSUM approach based on composites and cumulative sums is applied.

For the projections, the GCM data of the historical and future scenario runs were projected in each case onto the existing PCs of the observational period to obtain new predictor time series. Within this approach, the observational PC loadings and the standardized GCM data are used [for a detailed description of this approach, see von Storch and Zwiers (1999)]. The predictor PCs are the ones that showed the best GLM performance in the cross-validation set-up (Section 3.2.1.). Subsequently, the CPCUMSUM approach was applied under the use of the GCM predictors to assess change points in the predictor fields of the historical and future scenario runs. The GLM relationships associated with the specific predictor conditions in the segments before and after a change point were correspondingly applied to assess the response of local precipitation under scenario conditions.

The Mann–Whitney U test (Mann and Whitney, 1947) with a significance level of 95% has been used to assess the significance of the change point models in the local climate change projections. Local precipitation in the period 2070–2099 derived from the application of the change point statistical downscaling models is compared to the values derived from the use of conventional downscaling models.

4. Results

4.1. Change point analysis

For each of the 94 precipitation stations in the Mediterranean area, GLMs were constructed using the large-scale modes of variability (NAO, EA, EU1 and EU2) and the more regional 700 hPa geopotential heights centres of variation as predictors. GLMs were also derived under the inclusion of one additional predictor variable (air temperature, specific and relative humidity, u- and v-wind components and vertical velocity, all from the 850 hPa level). Under the use of 100 random calibration/validation periods, the best-performing predictor combination was assessed (Section 3.2.1.). The best-performing predictor combination for each station can be seen by means of the colour coding in Figure 3. In the eastern and northern Mediterranean area, circulation-type predictors (700 hPa geopotential heights and 850 hPa wind components) dominate, whereas for the Iberian Peninsula, particularly for the western, Atlantic oriented parts also humidity, and for the central parts air temperature play a role.

For each of the precipitation stations, GLMs with one (no change point) up to five (four change points) segments
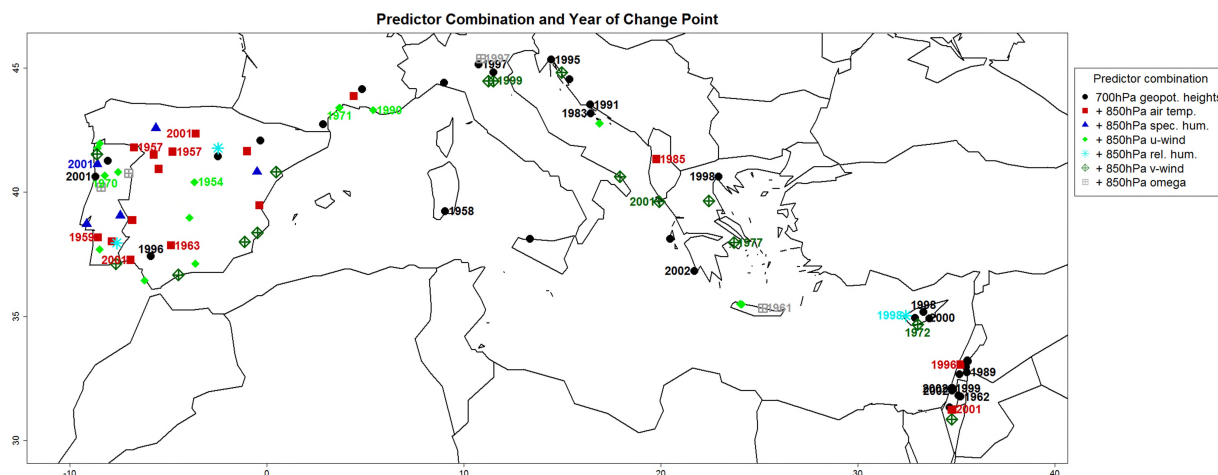
Figure 3. Best-performing predictor combination for each of the 94 precipitation stations in combination with the year of change point for the 37 stations with robust change points.

were constructed using the best-performing predictor combination. However, for all precipitation stations, the least GLM deviance was obtained when using two segments (one change point). Thus, all subsequent results are based on CPGLMs with one change point.

Only change point locations, which can be identified in the CPGLMs as well as in the CPCUMSUM approach under the use of two different reanalysis data sets, were considered to be robust (Section 3.2). This implies that the different climate characteristics before and after a change point, as derived from the CPGLMs, are caused by changes of the predictors, as evidenced in the CPCUMSUM approach. From the 94 stations considered in this study, for 54 stations consistent change points could be identified across the two reanalysis data sets using the CPGLM approach, and for 37 stations (approximately 40% of all stations), these change points can be confirmed with the CPCUMSUM approach (subsequently termed robust change points). The years of the robust change points are also indicated in Figure 3. Localizing the predictor sets and the dates of the change points reveals regional clusters. Thus, in the eastern Mediterranean area, change points occurred preferably with 700 hPa geopotential heights, and partially under the additional inclusion of 850 hPa air temperature, as predictors. The dates of the change points are mostly located in the 1990s and 2000s. For the majority of stations over northern Italy and Dalmatia, change points occurred in the 1990s using circulation-type predictors. For the Iberian Peninsula, robust change points can be seen mainly in the western and central parts of the region. They are predominantly related to the predictor combination 700 hPa geopotential heights and 850 hPa air temperature and occur preferably in the 1950s and at the beginning of the 2000s.

Depending on the station considered, the change points are associated with different non-stationarities in the predictor sets. As an example, Figure 4 shows for station Gafanha, Portugal (40.62° N, 8.70° W) the loading pattern of PC3, the large-scale mode resembling the EU1 (or Scandinavia) pattern. PC3 dominates the generation of the

change point in 2001 in the corresponding CPGLM. The southern centre of action is spatially more confined and located directly at the target station in the period before the change point (Figure 4(a)). In the period after the change point, there is a larger spatial spread and substantial shift of this centre of action towards the south-east (Figure 4(b)) resulting in significant changes of its relationship with precipitation at Gafanha. In a second example, Figure 5 shows the loading pattern for station Eilon, Israel (33.05° N, 35.21°E). Changes of PC1 (which resembles the NAO) play an important role for the change point at 1996 in the CPGLM. In the period before the change point, the eastern Mediterranean is essentially decoupled from the southern centre of action of NAO variability, whereas in the later period, there is a strong negative regression coefficient between the NAO and precipitation at Eilon. A second predictor, which contributes to the change point, is PC3 of 850 hPa air temperature (Figure 5(c) and (d)). In the earlier period, there is only a small influence of this temperature PC on precipitation, whereas in the period after the change point, the negative regression relationship (positive mode of the pattern associated with below normal precipitation, and vice versa) enhances considerably.

Overall, the occurrence of the change points in the various Mediterranean regions can be related to substantial circulation changes. The changes refer to modifications of the intensity, spatial position and spatial extent of the centres of action, as shown in the earlier examples. Non-stationarities are evident in all large-scale patterns considered in this study [NAO, EA, EU1 and EU2; for a detailed discussion see Hertig *et al.* (2015)]. Changes of the pattern-specific characteristics alter the predictors–precipitation relationships and are expressed as change points in the statistical transfer functions. In addition, changes in the thermal and thermo-dynamic characteristics of the atmosphere–climate relationships contribute to the generation of change points.

The predictor sets were obtained using the whole time period available for a station in a cross-validation framework based on 100 random samples. The CPGLM
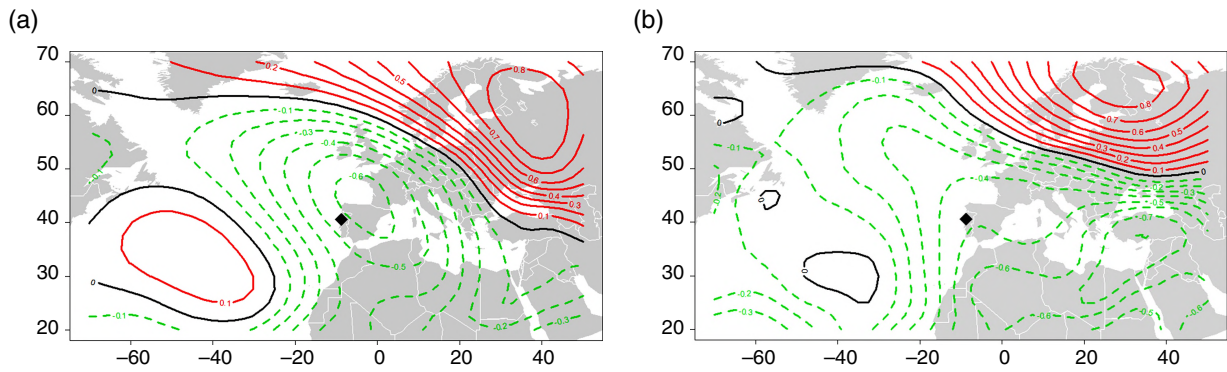
Figure 4. Loading pattern for PC3 of monthly 700 hPa geopotential heights December–February in the periods 1950–2001 (a) and 2002–2008 (b). Diamond symbol: location of station Gafanha, Portugal (40.62° N, 8.70° W).
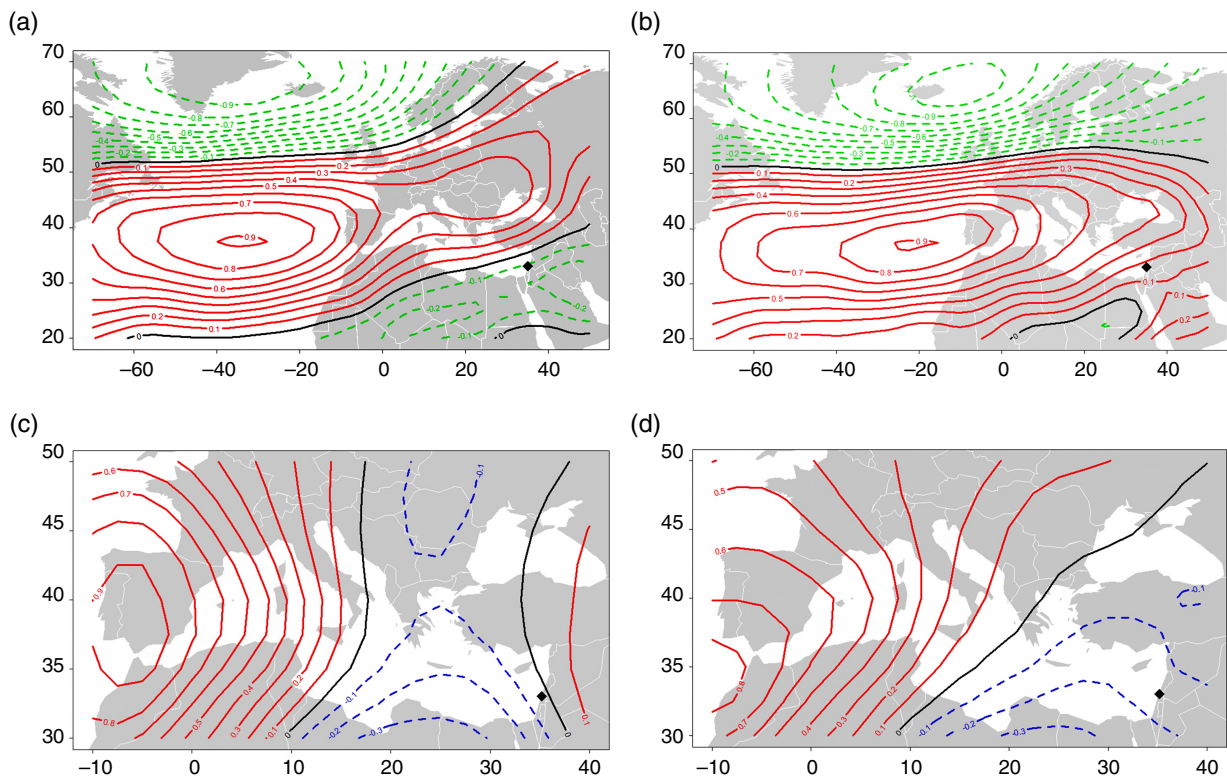

Figure 5. Loading pattern for PC1 of 700 hPa geopotential heights ((a) and (b)) and PC3 of 850 hPa air temperature ((c) and (d)) December–February in the periods 1958–1996 ((a) and (c)) and 1997–2003 ((b) and (d)). Diamond symbol: location of station Eilon, Israel (33.05° N, 35.21° E).

approach revealed that, even though the predictors are significant for the whole time period under consideration, there can be substantial temporal variations in the importance of a specific predictor. Thus, a specific predictor may not be very relevant in one period, but can gain high significance in another time slice, which leads to multiple GLMs to adequately capture the varying predictors–predictand relationships. The question whether the CPGLM and CPCUMSUM approach is more appropriate in a downscaling context compared to the use of conventional GLMs is pursued further in the following section.

4.2. Validation

The results of cross-validation using conventional GLMs can be seen in Figure 6. The MSESS ranges between 19 and 77% in calibration (Figure 6(a)) and between 10 and 66% in validation (Figure 6(b)). This indicates in general a good skill of the downscaling models under stationary conditions. Regionally the highest values of the MSESS can be observed over the western parts of the Iberian Peninsula, whereas the eastern parts of Spain and some stations over southern France and Greece show the lowest skill.

From the change point analysis robust change points were derived for 37 stations. However, within the framework of statistical downscaling, it is important to analyse whether these change points actually impact on the statistical model performance, and thus whether these change points are not only robust, but also relevant. Hence, we tested whether the use of the CPGLMs is generally justified, i.e. whether the different predictors–predictand
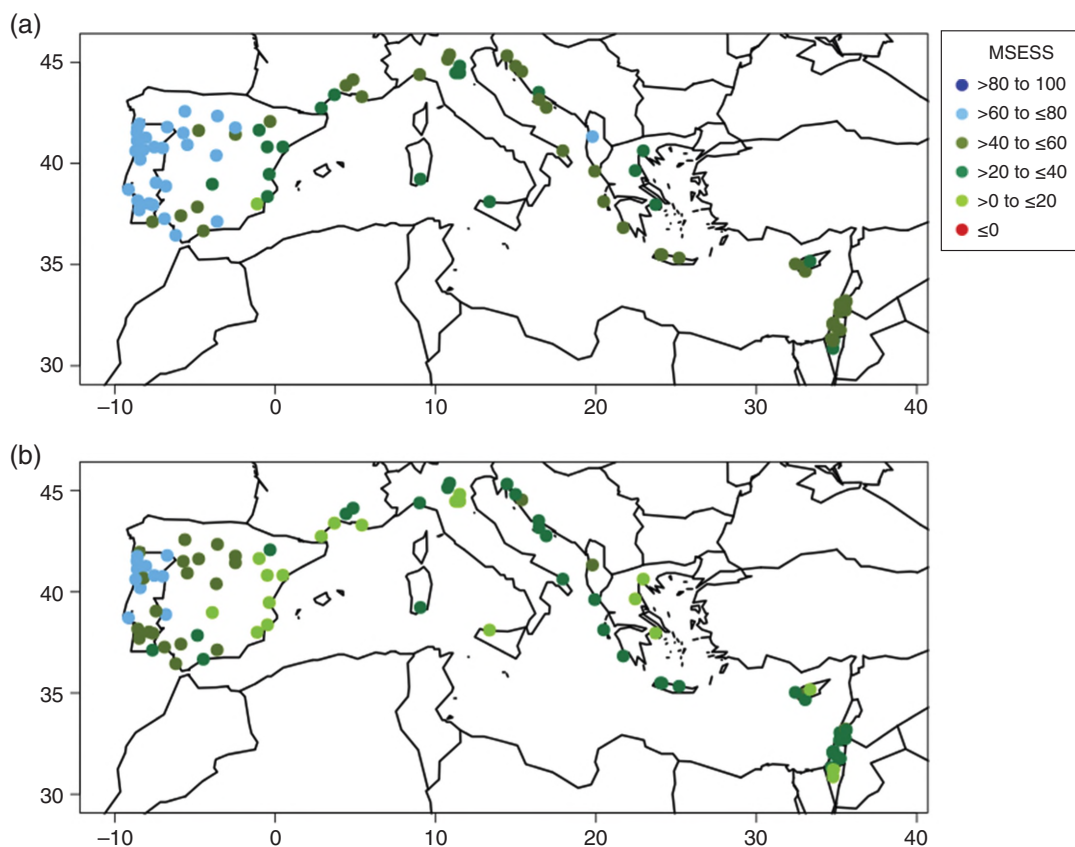
Figure 6. MSESS in calibration (a) and validation (b) using 100 random samples and the best-performing predictor combination for each station.

relationships can be described more accurately with the aid of CPGLMs compared to the use of conventional GLMs without change points. The difference of the mean MSESS from the CPGLMs and the MSESS of the GLMs without change points is shown in Figure 7. Note that for the sake of a complete picture we show the results for all 94 stations, the results for the 37 stations with robust change points are marked by the diamond symbol in Figure 7. Indicated by the green colours in Figure 7, model performance of the CPGLMs is higher in calibration for all stations, an expectable result arising from the better adaption of the CPGLMs in calibration. In validation, it can be seen that for 30 stations (thereof 22 with robust change points) the application of conventional GLMs yields better results (orange and red colours in Figure 7). For these stations, the predictors–predictand relationships are described better by a single GLM equation. Variations in the predictors–predictand relationships as identified for many of these stations do not substantially affect downscaling model performance in terms of the MSESS and thus the use of one model describing the relationships as a whole is more suitable. It should be noted that a different validation measure can alter this result. However, in this study, we are interested in the mean model performance, and validation of other aspects is not considered here. The GLM change point approach yields an added value regarding mean model performance in validation for 64 stations, thereof 15 stations with robust change points (Figure 7). The close-by location of stations with

different behaviour regarding stationarity/non-stationarity reflects the choice of local-scale models with site-specific predictor–predictand relationships. In order to insure both, the physical robustness and the statistical relevance, in the downscaling projections CPGLMs were considered only for the 15 stations featuring robust and relevant change points. For the remaining 79 stations, conventional GLMs were used in the downscaling assessments.

Overall, 37 (about 40%) of all stations show robust change points. Thus, change points in the predictors–precipitation relationships induced by changes of the characteristics of the atmospheric patterns and/or of the thermo-dynamic properties of specific patterns can be seen in many cases. Within the scope of statistical downscaling of future climate change, cross-validation is an important part. For 15 (approx. 16%) of all stations, the CPGLMs pass the specific cross-validation set-up yielding robust and relevant change points. It points to the necessity to allow for change points for the local precipitation projections at these stations.

4.3. Change points in model predictors

Within the change point statistical downscaling approach, the large-scale atmospheric state in the model data had also to be determined in order to decide whether the observational GLM relationships of the period before or the period after a change point are more appropriate to downscale local climate in the model periods. In principle, it is possible that there exist multiple change points
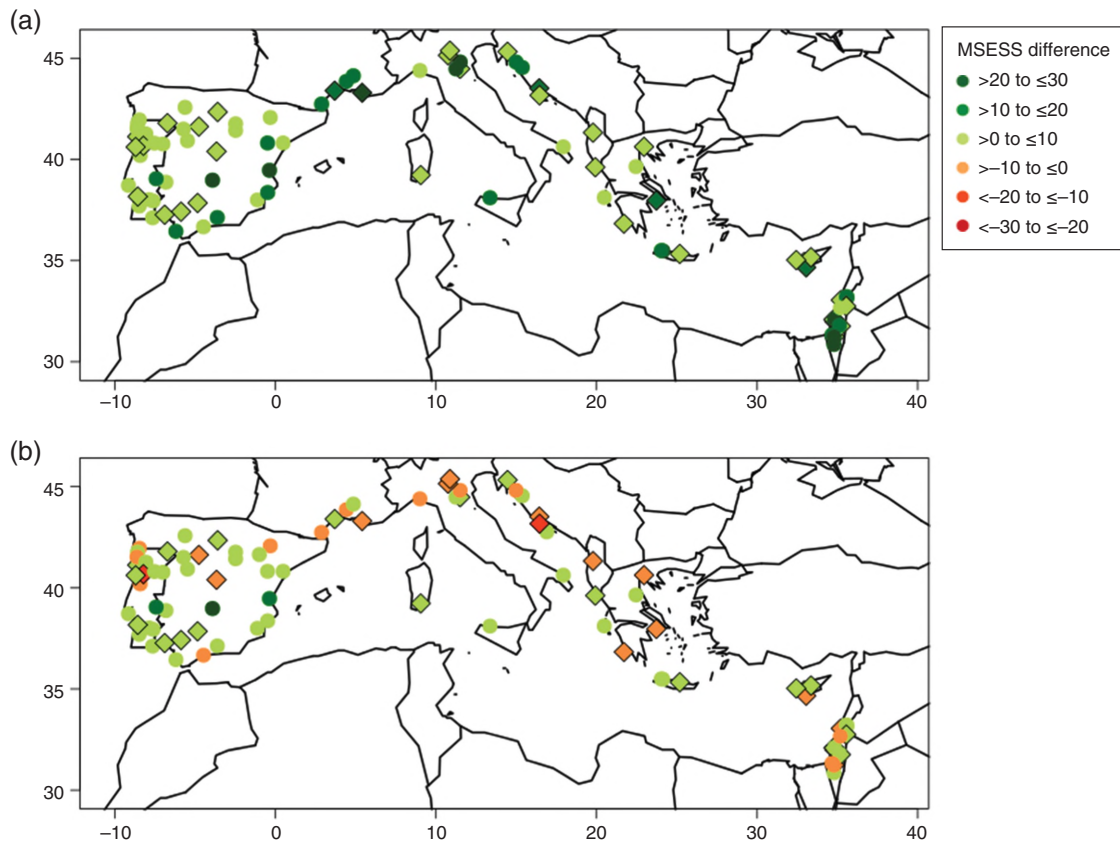
Figure 7. Difference in MSESS between CPGLM approach and GLM approach without considering change points in calibration (a) and validation (b). Stations with robust change point are shown as diamonds and those with no robust change points as dots.

in the model periods, with before change point predictor conditions alternating with after change point predictor conditions. Thus, not only maxima in the cumulative sum curves are relevant for the definition of change points, but also minima. The analysis was done for the station-specific predictor sets, i.e. the best-performing predictor combinations in the validation set-up.

However, the experiment using the three runs of MPI-ESM-LR with different initial conditions revealed no multiple change points in the cumulative sum plots of the historical runs in the period 1950–2005 and of the scenario runs in the period 2006–2100. For the 15 stations, which have robust and relevant change points in the observational period, the model predictors showed no or only one change point. The predictor combinations comprise circulation predictors for four stations, circulation and additionally air temperature for another four stations, circulation plus wind components for five stations and circulation plus specific humidity/relative vorticity at one station each. In the cases of no change point in the model data, the model predictor characteristics resembled more closely the observational atmospheric conditions after the change points. Thus, for the projection only the GLM of the segment after the change point is used. For the other cases, the cumulative sum curves showed one maximum. In the projections, the maximum was always located within the first 30 years (before 2035) of the scenario periods. The predictor characteristics in the projection period

before/after the maximum of the cumulative sum curve are closer related to the conditions before/after the change point and thus the GLM of the segment before/after the change point is used, respectively. Note that for the remaining 79 stations with no robust and relevant change points, the projections in the following Section 4.4 are based on conventional statistical downscaling models.

4.4. Local precipitation projections for the 21st century

Within the present change point statistical downscaling approach, it is of interest how the change point models impact on the projection results. Figure 8 shows for the 15 stations having robust and relevant change points the comparison of the downscaling results without considering change points (Figure 8(a)–(c)) and under consideration of change points (Figure 8(d)–(f)). Shown is the winter (December–February) precipitation mean for the period 2070–2099 under RCP8.5 scenario conditions for each of the three MPI-ESM-LR runs. Comparing the results reveals that the spatial precipitation pattern over the Mediterranean area is preserved. This is due to the fact that both approaches use the same predictor set and similar general relationships between predictors and precipitation. However, there are also some significant differences according to the U test (95% significance level), visualized by the square symbols in Figure 8. For the MPI-ESM-LR run 1, five stations show significant differences of the downscaling results between the model
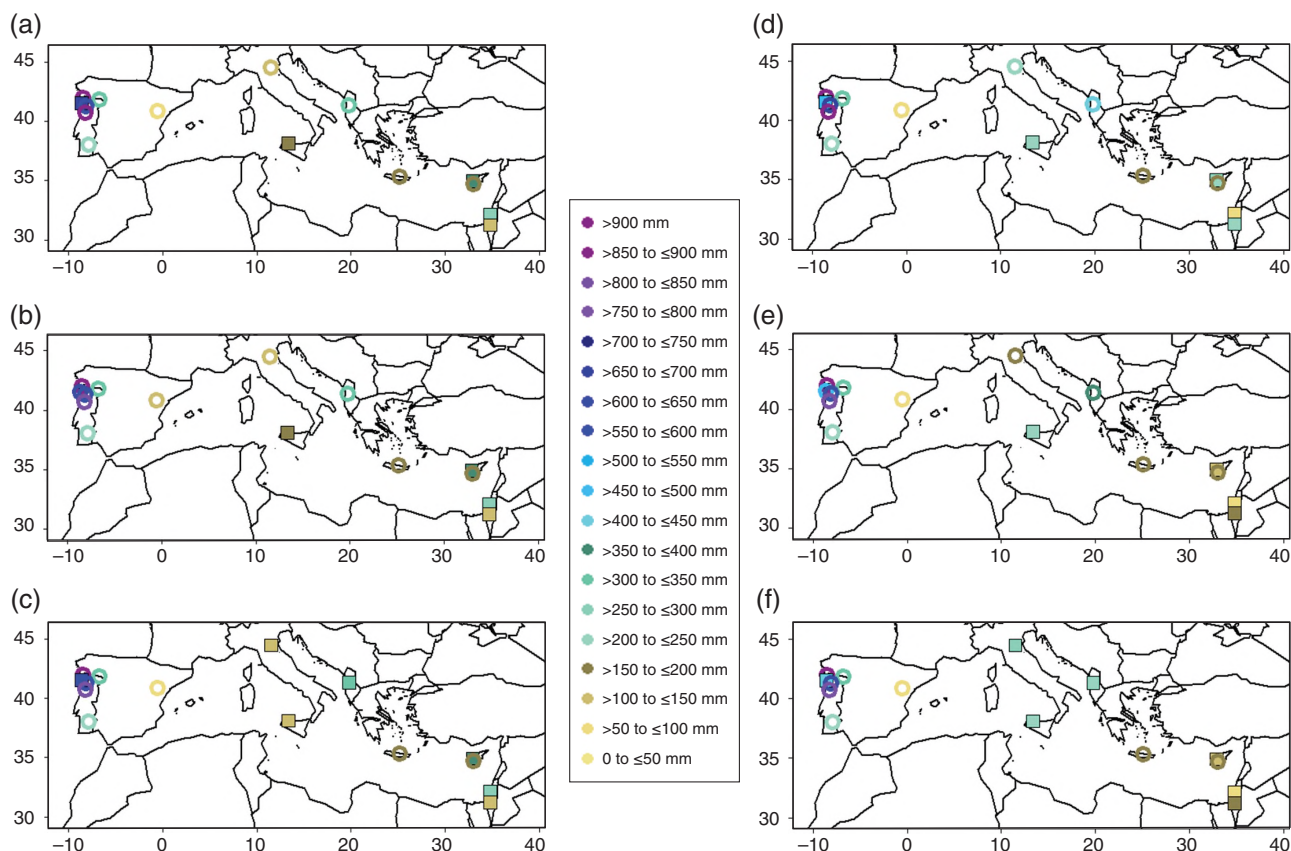
Figure 8. Precipitation mean in winter (December–February) of the period 2070–2099 for the 15 stations with robust and relevant change points. Shown are the results for each of the three MPI-ESM runs under RCP8.5 scenario. (a)–(c): no change point models, (d)–(f): change point models. Square denotes significant difference (*U* test, 95% significance level) of the winter precipitation mean between the change point model and the no change point model.

considering change points and the model without change points, four stations in case of run 2 and seven stations for run 3. In these cases, the projected 30-year winter precipitation mean differs significantly between the two downscaling approaches with maximum values of the difference of 250 mm.

The underlying reasons for these differences lie in the use of different GLM relationships of the large-scale and regional-scale circulation with local precipitation. Taking as an example station Palermo, Italy, there are significantly higher mean precipitation amounts assessed under the use of the change point model. The station's precipitation is connected mainly with the EA pattern as well as with two centres of variation over the Mediterranean area. Whereas the EA pattern contributes only little to the difference, the regression coefficients of the two regional centres of variation are significantly higher in the CPGLM compared to the conventional GLM, resulting in a greater impact on precipitation. Aside from circulation-related differences, for about half of the stations with significant differences temperature or vertical wind components play a role as well. For instance, station Beer Sheva, Israel, is connected to several circulation predictors (and in particular to the EU2 pattern and the regional centre of variation located directly over the station), but as well with temperature across the north-central and north-eastern Mediterranean

area. All predictors show somewhat higher absolute values of the regression coefficients in the CPGLM compared to the conventional GLM, jointly contributing to higher precipitation amounts in the CPGLM assessments.

Figure 9 shows the results of the precipitation downscaling in the winter season under RCP4.5 (Figure 9(a)) and RCP8.5 (Figure 9(b)) scenario assumptions. The ensemble mean changes from the three MPI-ESM-LR runs for the scenario period 2070–2099 minus the historical model period 1970–1999 are reproduced. The ensemble mean change is calculated from the differences between scenario period means and historical period means, averaged over the three runs. It should be noted that just one model from the CMIP5 database was selected to test the downscaling approach under consideration of change points in the predictor–predictand relationships. We do not consider variability arising from the use of different GCMs. The assessments were done using conventional statistical downscaling for 79 stations and the change point statistical downscaling approach for the 15 stations, which showed robust and relevant change points. Under the RCP4.5 scenario precipitation increases in winter up to about 80 mm are projected for the western and northern parts of the Mediterranean area, whereas the north-eastern and eastern Mediterranean regions are affected by precipitation reductions up to −40 mm. This spatial picture of precipitation
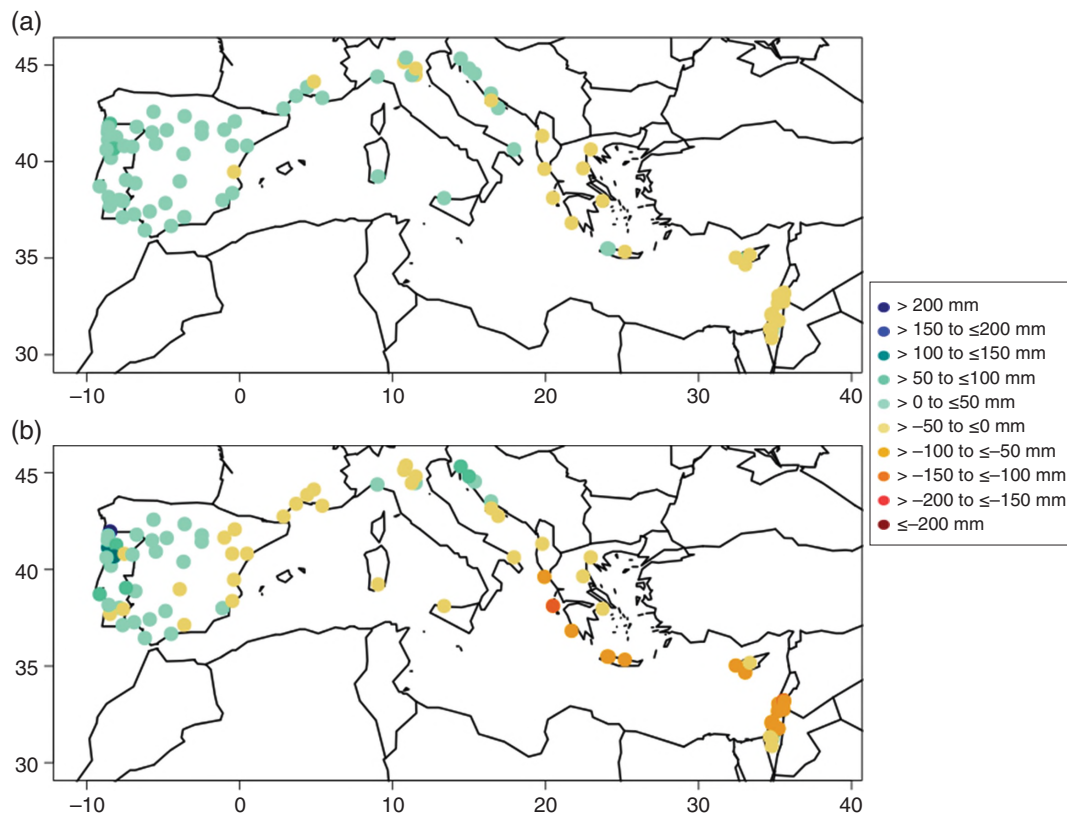
Figure 9. Precipitation changes in winter (December–February) as differences between the scenario period 2070–2099 and the historical period 1970–1999. Shown are the ensemble mean changes of the three MPI-ESM-LR runs for the RCP4.5 scenario (a) and for the RCP8.5 scenario (b).

change is partly consistent under the RCP8.5 scenario, however with more widespread reductions. Most noticeably, the eastern parts of Spain, southern France and northern Italy are affected by precipitation decreases. In addition, the magnitude of change is partly larger in the RCP8.5 scenario case. The strongest decreases occur over parts of the eastern Mediterranean area with values up to about −110 mm (−23%, station Kfar Sava, Israel). Strongest increases can be found for some stations over the western Iberian Peninsula with values up to 205 mm (27%, station Extremo, Portugal) and over the Dalmatian coast.

5. Discussion and conclusions

A multitude of methods has been developed to deal with change points in climate data depending on the application (e.g. data homogenization of a single time series up to the detection of structural change points in complex circulation–climate relationships), the variable of interest (e.g. temperature, precipitation and wind) and specific requirements (e.g. ability to detect single or multiple change points, parametric *vs* nonparametric approaches). In this work, change points in the relationships of large-scale atmospheric predictor variables and local climate conditions were taken into account through the application of CPGLMs. In combination with a change point method based on CPCUMSUM, a statistical downscaling approach considering change points was developed. This was done with the objective

to advance statistical downscaling approaches and to improve reliability of projections of regional/local climate change. In the present contribution, we focused on station-based mean precipitation. However, in the scope of future climate change, other aspects like changes in extreme precipitation are of particular importance as well. In order to specifically assess non-stationarities in extremes, a different approach should be chosen, for instance based on quantile regression or extreme value distributions like the generalized Pareto distribution.

Change points in the predictors–precipitation relationships for 94 stations in the Mediterranean area were assessed with the CPGLMs, and the CPCUMSUM method was used to independently find these change points in the predictor data. Two different reanalysis data sets were used in order to avoid artificial change point detections due to data inhomogeneity. Only change points, which could be found in all four configurations, were subsequently treated as reliable and were termed robust change points. About 40% (37 out of 94) of the analysed precipitation stations showed robust change points.

The change points in the predictors–precipitation relationships can be seen in the context of non-stationarities of the patterns of atmospheric variability and consequently in the regional climate affected by these patterns. Changes in the overall importance of a predictor, changes of the mean predictor state and/or intensity changes and spatial shifts of specific centres of variation have to be taken into account. Further in-depth analysis of causes and mechanisms

generating non-stationarities in recent climate variability can be found in Hertig *et al.* (2015). In general, any observed non-stationarity is related to multiple processes, which contribute in differing amounts. The decision on whether or not to include non-stationary relationships into the downscaling approach depends on its statistical evidence and its impact on the downscaling results.

Thus, subsequent to the change point analysis, the approach was tested with respect to whether the consideration of change points is of relevance within a statistical downscaling framework. The model performance in terms of the MSESS of the change point statistical downscaling approach was assessed in a cross-validation framework and compared to the model performance using conventional statistical downscaling models without allowing for change points. For 15 out of the 37 stations showing robust change points, the statistical model performance yielded an added value, i.e. model performance in validation was enhanced in the change point statistical downscaling models. It should be noted that also for other stations with no robust change points, the model performance could be enhanced by the change point statistical downscaling approach. Vice versa for some stations with robust change points, the consideration of change point models did not result in an enhanced MSESS. Therefore, the results from cross-validation were judged in combination with the outcome from the change point analysis. This resulted in the 15 stations, which featured robust and relevant change points in the predictors–precipitation relationships. Consequently, only for these stations, the change point statistical downscaling models were applied for the local climate change projections, whereas precipitation for the remaining 79 stations was downscaled using conventional GLMs without considering change points.

With regard to the GCM data, we found that the predictor characteristics of the model were more closely related to the observational predictor state of the periods after the change points. Several reasons could be responsible for this: (1) the GCM does not correctly reproduce the observed atmospheric variability. This is supported by the finding that this feature even emerged in the historical runs of the GCM; (2) changes of the atmospheric conditions driven by anthropogenic greenhouse gas forcing already occurred in the observational period and were mostly effective after the change points. Thus, in the GCM scenario runs, the circulation characteristics can be seen as a continuation of the conditions after the change points; (3) we have not sampled enough natural climate variability in our about 60 years long observation period. Thus, an extension of the analysis to longer historical periods may provide further insight into natural climate variability. Under the assumption the variations in an altered climate are roughly contained in the variability of the present climate it may give the possibility of further improvements of this statistical downscaling approach. Besides, it should be noted that in the scope of future climate change, it is possible that new atmospheric patterns may evolve. However, within the present approach, no entirely new predictor patterns are taken into account. The statistical downscaling approach can only accommodate to modifications of existing patterns.

For the 15 stations with robust and relevant change points, the significance of the change point models in the local climate change projections was assessed. Local precipitation in the period 2070–2099 derived from the application of the change point statistical downscaling models was compared to the values derived from the use of conventional downscaling models. The spatial precipitation pattern across the stations obtained from the change point approach conforms to the one resulting from the application of GLMs without change points due to common signals of the large-scale predictors on precipitation. However, the application of the change point approach affects the downscaled future precipitation amounts. Significant differences to the models without change points were found for about one third of the 15 stations. With respect to all 94 stations, the change point statistical downscaling approach resulted in significantly different local climate change projection results for about 6% of the stations. A very strict detection and selection procedure of the CPGLMs was used to qualify them as downscaling models. Thus, the final number of stations with robust, relevant and significant change points can be seen as a lower boundary. For these stations, future precipitation change depends critically on the specifics of the large-scale circulation changes. Since different GLMs were selected according to the prevailing predictor characteristics, the change point statistical downscaling approach allows for a flexible adaption of future precipitation change to these predictor changes.

Different approaches have been developed to consider non-stationarities in connection with statistical downscaling. Hertig and Jacobeit (2013) assessed non-stationarities based on sliding calibration periods. A direct comparison of the results obtained with the two different approaches is not possible due to differences in the predictors and differences in the GLM set-up. However, the projected changes for the RCP4.5 scenario case presented in Figure 7 resemble in large part the downscaling results from Hertig and Jacobeit (2013). This gives improved confidence in the robustness of these easy to apply and computationally inexpensive non-stationary approaches to downscaling. A further approach in the scope of statistical downscaling considering non-stationarities could be adopted from Zeileis *et al.* (2008). They developed a model-based recursive partitioning algorithm for parameter estimation and partitioning. Thus, different statistical models, defined by the combination of breakpoints and stratified predictor–predictand equations, can be used in a downscaling context. Furthermore, other statistical downscaling approaches like synoptic downscaling can be consulted. In this context, changes in the occurrence frequencies of circulation types could be used for partitioning. Subsequently, the specific dynamical (e.g. vorticity and intensity) and climatological (temperature and precipitation) characteristics of the circulation classes can be analysed in the corresponding sub-periods and subsequently

be used for the development of a non-stationary synoptic downscaling model.

Finally, with the present approach under RCP4.5 and RCP8.5 scenario conditions mainly increases of winter precipitation were assessed for parts of the western and northern Mediterranean area, whereas decreases dominate in the north-eastern and eastern Mediterranean area. However, only three runs of one GCM were considered in this study. Thus, it remains to analyse the relevance of the change point statistical downscaling approach under consideration of several different GCMs.

Acknowledgement

Supporting information

The following supporting information is available as part of the online article:
R function to find change points in generalized linear models.

References

Alexandersson H. 1986. A homogeneity test applied to precipitation data. *Int. J. Climatol.* **6**: 661–675.

Barnston AG, Livezey RE. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Weather Rev.* **115**: 1083–1126.

Beaulieu C, Chen J, Sarmiento JL. 2012. Change-point analysis as a tool to detect abrupt climate variations. *Phil. Trans. R. Soc. A* **370**: 1228–1249.

Bellman R, Dreyfus S. 1962. *Applied Dynamic Programming* (Tech. Rep.). Princeton University Press: Princeton, NJ.

Beranová R, Huth R. 2008. Time variations of the effects of circulation variability modes on European temperature and precipitation in winter. *Int. J. Climatol.* **28**: 139–158.

Chandler RE. 2005. On the use of generalized linear models for interpreting climate variability. *Environmetrics* **16**: 699–715.

Christensen JH, Krishna Kumar K, Aldrian E, An S-I, Cavalcanti IFA, de Castro M, Dong W, Goswami P, Hall A, Kanyanga JK, Kitoh A, Kossin J, Lau N-C, Renwick J, Stephenson DB, Xie S-P, Zhou T. 2013. Climate phenomena and their relevance for future regional climate change. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds). Cambridge University Press: Cambridge, UK and New York, NY.

Diaz HF, Hoerling MP, Eischeid JK. 2001. ENSO variability, teleconnections and climate change. *Int. J. Climatol.* **21**: 1845–1862.

Dünkeloh A, Jacobeit J. 2003. Circulation dynamics of Mediterranean precipitation variability 1948–1998. *Int. J. Climatol.* **23**: 1843–1866.

Dunn PK. 2004. Occurrence and quantity of precipitation can be modelled simultaneously. *Int. J. Climatol.* **24**: 1231–1239.

Dunn PK. 2008. Tweedie: Tweedie exponential family models. R package, R package version 1.5.1. Vienna, Austria.

Giorgi F. 2006. Climate change hot-spots. *Geophys. Res. Lett.* **33**: L08707, doi: 10.1029/2006GL025734.

Goodess CM, Jones PD. 2002. Links between circulation and changes in the characteristics of Iberian rainfall. *Int. J. Climatol.* **22**: 1593–1615.

Hasan MM, Dunn PK. 2011. Two Tweedie distributions that are near-optimal for modelling monthly rainfall in Australia. *Int. J. Climatol.* **31**: 1389–1397.

Hertig E, Jacobeit J. 2008. Assessments of Mediterranean precipitation changes for the 21st century using statistical downscaling techniques. *Int. J. Climatol.* **28**: 1025–1045.

Hertig E, Jacobeit J. 2013. A novel approach to statistical downscaling considering non-stationarities: application to daily precipitation in the Mediterranean area. *J. Geophys. Res. Atmos.* **118**: 520–533.

Hertig E, Jacobeit J. 2014a. Variability of weather regimes in the North Atlantic-European area: past and future. *Atmos. Sci. Lett.* **15**: 314–320.

Hertig E, Jacobeit J. 2014b. Considering observed and future nonstationarities in statistical downscaling of Mediterranean precipitation. *Theor. Appl. Climatol.* **122**: 667–683.

Hertig E, Seubert S, Paxian A, Vogt G, Paeth H, Jacobeit J. 2014. Statistical modeling of extreme precipitation for the Mediterranean area under future climate change. *Int. J. Climatol.* **34**: 1132–1156.

Hertig E, Beck C, Wanner H, Jacobeit J. 2015. A review of non-stationarities in climate variability of the last century with focus on the North Atlantic-European sector. *Earth Sci. Rev.* **147**: 1–17.

Hofrichter J. 2007. *Change Point Detection in Generalized Linear Models*. PhD thesis, Graz University of Technology, Graz.

Hurrell J, Deser C. 2010. North Atlantic climate variability: the role of the North Atlantic Oscillation. *J. Mar. Syst.* **79**: 231–244.

Jung T, Hilmer M, Ruprecht E, Kleppek S, Gulev S, Zolina O. 2003. Characteristics of the recent eastward shift of interannual NAO variability. *J. Clim.* **16**: 3371–3381.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**: 437–471.

Kistler R, Kalnay E, Collins W, Saha S, White G, Woollen J, Chelliah M, Ebisuzaki W, Kanamitsu M, Kousky V, van den Dool H, Jenne R, Fiorino M. 2001. The NCEP/NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bull. Am. Meteorol. Soc.* **82**: 247–268.

Klein Tank AMG, Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A, Mileta M, Pashiardis S, Hejkrlik L, Kern-Hansen C, Heino R, Bessemoulin P, Müller-Westermeier G, Tzanakou M, Szalai S, Pálsdóttir T, Fitzgerald D, Rubin S, Capaldo M, Maugeri M, Leitass A, Bukantis A, Aberfeld R, van Engelen AFV, Forland E, Mietus M, Coelho F, Mares C, Razuvaev V, Nieplova E, Cegnar T, Antonio López J, Dahlström B, Moberg A, Kirchhofer W, Ceylan A, Pachaliuk O, Alexander LV, Petrovic P. 2002. Daily dataset of the 20th-century surface air temperature and precipitation series for the European climate assessment. *Int. J. Climatol.* **22**: 1441–1453.

Kunstmann H, Heckl A, Rimmer A. 2006. Physically based distributed hydrological modeling of the Upper Jordan catchment and investigation of effective model equations. *Adv. Geosci.* **9**: 123–130.

Lionello P, Abrantes F, Congedi L, Dulac F, Gacic M, Gomis D, Goodess C, Hoff H, Kutiel H, Luterbacher J, Planton S, Reale M, Schröder K, Struglia MV, Toreti A, Tsimplis M, Ulbrich U, Xoplaki E. 2012. Introduction: Mediterranean climate – background information. In *The Climate of the Mediterranean Region. From the Past to the Future*, Lionello P (ed). Elsevier: London.

Mann B, Whitney DR. 1947. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **18**(1): 50–60.

Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themeßl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I. 2010. Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.* **48**(3): RG3003, doi: 10.1029/2009RG000314.

Mc Cullagh P, Nelder JA. 1989. *Generalized Linear Models. Monographs on Statistics and Applied Probability*, Vol. 37. Chapman & Hall: London.

Miller AJ, Cayan DR, Barnett TP, Graham NE, Oberhuber JM. 1994. The 1976–77 climate shift of the Pacific Ocean. *Oceanography* **7**: 21–26.

Moberg A, Jones PD, Lister D, Walther A, Brunet M, Jacobeit J, Alexander LV, Della-Marta P, Luterbacher J, Yiou P, Chen D, Klein Tank AMG, Saladié O, Sigró J, Aguilar E, Alexandersson H, Almarza C, Auer I, Barriendos M, Begert M, Bergström H, Böhm R, Butler CJ, Caesar J, Drebs A, Founda D, Gerstengarbe F-W, Micela G, Maugeri M, Österle H, Pandzic K, Petrakis M, Srnec L, Tolasz R, Tuomenvirta H, Werner PC, Linderholm H, Philipp A, Wanner H, Xoplaki E. 2006. Indices for daily temperature and precipitation extremes in Europe analyzed for the

period 1901–2000. *J. Geophys. Res. Atmos.* **111**: D22106, doi: 10.1029/2006JD007103.

Moore GWK, Renfrew IA, Pickart RS. 2013. Multidecadal mobility of the North Atlantic Oscillation. *J. Clim.* **26**: 2453–2466.

Murphy AH. 1988. Skill scores based on the mean squared error and their relationship to the correlation coefficient. *Mon. Weather Rev.* **116**: 2417–2424.

Overland J, Rodionov S, Minobe S, Bond N. 2008. North Pacific regime shifts: definitions, issues and recent transitions. *Prog. Oceanogr.* **77**: 92–102.

Page ES. 1955. A test for a change in a parameter occurring at an unknown point. *Biometrika* **42**: 523–527.

Page ES. 1957. On problems in which a change in parameter occurs at an unknown point. *Biometrika* **44**: 248–252.

Philipp A, Della-Marta PM, Jacobeit J, Fereday DR, Jones PD, Moberg A, Wanner H. 2007. Long term variability of daily North Atlantic-European pressure patterns since 1850 classified by simulated annealing clustering. *J. Clim.* **20**(16): 4065–4095.

Poli P, Hersbach H, Tan D, Dee D, Thépaut J-N, Simmons A, Peubey C, Laloyaux P, Komori T, Berrisford P, Dragani R, Trémolet Y, Hólm E, Bonavita M, Isaksen L, Fisher M. 2013. The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface obersvations only (ERA-20C). ECMWF ERA Rep. 14, 59 pp.

Power SB, Smith IN. 2007. Weakening of the Walker Circulation and apparent dominance of El Niño both reach record levels, but has ENSO really changed? *Geophys. Res. Lett.* **34**: L18702, doi: 10.1029/2007GL030854.

Preisendorfer RW. 1988. *Principal Component Analysis in Meteorology and Oceanography. Developments in Atmospheric Science 17*. Elsevier: Amsterdam.

Quadrelli R, Pavan V, Molteni F. 2001. Wintertime variability of Mediterranean precipitation and its links with large-scale circulation anomalies. *Clim. Dyn.* **17**: 457–466.

R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna. ISBN 3-900051-07-0. http://www.R-project.org (accessed 18 January 2016).

Raible CC, Lehner F, González-Rouco JF, Fernández-Donado L. 2014. Changing correlation structures of the Northern Hemisphere atmospheric circulation from 1000 to 2100 AD. *Clim. Past* **10**: 537–550.

Reeves J, Chen J, Wang XL, Lund R, Lu Q. 2007. A review and comparison of changepoint detection techniques for climate data. *J. Appl. Meteorol. Climatol.* **46**: 900–915.

Rodionov SN. 2004. A sequential algorithm for testing climate regime shifts. *Geophys. Res. Lett.* **31**: L09204, doi: 10.1029/2004GL019448.

Stickler A, Brönnimann S, Valente MA, Bethke J, Sterin A, Jourdain S, Roucaute E, Vasquez MV, Reyes DA, Allan R, Dee D. 2014. ERA-CLIM. Historical surface and upper-air data for future reanalysis. *Bull. Am. Meteorol. Soc.* **95**: 1419–1430.

Tatli H, Nüzhet Dalfes H, Sibel MS. 2004. A statistical downscaling method for monthly total precipitation over Turkey. *Int. J. Climatol.* **24**: 161–180.

Trenberth KE, Hoar TJ. 1997. El Niño and climate change. *Geophys. Res. Lett.* **24**(23): 3057–3060.

Ullmann A, Fontaine B, Roucou P. 2013. Euro-Atlantic weather regimes and Mediterranean rainfall patterns: present-day variability and expected changes under CMIP5 projections. *Int. J. Climatol.* **34**: 2634–2650.

Van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque J-F, Masui T, Meinshausen M, Nakicenovic N, Smith SJ, Rose SK. 2011. The representative concentration pathways: an overview. *Clim. Change* **109**: 5–31.

Vicente-Serrano SM, López-Moreno JI. 2008. Nonstationary influence of the North Atlantic Oscillation on European precipitation. *J. Geophys. Res. Atmos.* **113**: D20120.

Von Storch H, Zwiers FW. 1999. *Statistical Analysis in Climate Research*. Cambridge University Press: Cambridge, UK, 484.

Wijngaard JB, Klein Tank AMG, Können GP. 2003. Homogeneity of 20th century European daily temperature and precipitation series. *Int. J. Climatol.* **23**: 679–692.

Wilby RL. 1998. Statistical downscaling of daily precipitation using daily airflow and seasonal teleconnection indices. *Clim. Res.* **10**: 163–178.

Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences*. Elsevier: Oxford, UK, 704.

Xoplaki E, González-Rouco JF, Luterbacher J, Wanner H. 2004. Wet season Mediterranean precipitation variability: influence of large-scale dynamics and trends. *Clim. Dyn.* **23**: 63–78.

Zeileis A, Kleiber C, Krämer W, Hornik K. 2003. Testing and dating of structural changes in practice. *Comput. Stat. Data Anal.* **44**: 109–123.

Zeileis A, Hothorn T, Hornik K. 2008. Model-based recursive partitioning. *J. Comput. Graph. Stat.* **17**(2): 492–514.

Zorita E, von Storch H. 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *J. Clim.* **12**: 2474–2489.