Statistical modelling of extreme precipitation indices for the Mediterranean area under future climate change

Elke Hertig, Stefanie Seubert, A. Paxian, G. Vogt, H. Paeth, Jucundus Jacobeit

Angaben zur Veröffentlichung / Publication details:

Hertig, Elke, Stefanie Seubert, A. Paxian, G. Vogt, H. Paeth, and Jucundus Jacobeit. 2014. "Statistical modelling of extreme precipitation indices for the Mediterranean area under future climate change." *International Journal of Climatology* 34 (4): 1132–56. https://doi.org/10.1002/joc.3751.

Statistical modelling of extreme precipitation indices for the Mediterranean area under future climate change

E. Hertig,[a]* S. Seubert,[a] A. Paxian,[b] G. Vogt,[b] H. Paeth[b] and J. Jacobeit[a]

[a] *Institute of Geography, University of Augsburg, Germany*

[b] *Institute of Geography and Geology, University of Wuerzburg, Germany*

ABSTRACT: Projected changes of extreme precipitation in the Mediterranean area up until the end of the 21st century are analysed by means of statistical downscaling. Generalized linear models are used as downscaling technique to assess different percentile-based indices of extreme precipitation on a fine-scale spatial resolution. In the region under consideration extreme precipitation is related to anomalies of the large-scale circulation as well as to convective conditions. To account for this, predictor selection encompasses variables describing the large-scale circulation (geopotential heights of the 700 hPa and 500 hPa levels, u- and v-wind components of the 850 hPa level) as well as thermo-dynamic parameters (specific humidity of the 850 hPa and 700 hPa levels, Showalter-Index, convective inhibition). In the scope of the statistical downscaling approach a specific statistical ensemble technique is applied in order to allow for non-stationarities in the predictors–predictand relationships. Consequently, the statistical ensembles include a range of possible future evolutions of extreme precipitation. Two different emission scenarios (A1B and B1), multiple runs for each scenario, and output of two different general circulation models (ECHAM5 and HadCM3) are applied to assess extreme precipitation under enhanced greenhouse warming conditions. The results yield mainly decreases over many parts of the Mediterranean area in spring. In summer increases are assessed around the Tyrrhenian Sea, the Ionian Sea, and the Aegean Sea, whereas decreases are projected for most of the western and northern Mediterranean regions. In autumn reductions of heavy rainfall occur over many parts of the western and central areas. In winter distinct increases are widespread in the Mediterranean area. Beyond the assessments using all predictors it is shown in the present contribution that different predictor variables can lead to varying statistical downscaling results. It points to distinct impacts of the change of specific atmospheric conditions on local extreme precipitation.

KEY WORDS climate change; statistical downscaling; generalized linear models; extreme precipitation; Mediterranean area

1. Introduction

The Mediterranean area shows a wide range of different climatic characteristics, from humid conditions in the western, northern and north-eastern regions in the wet season from approximately September to May, to arid conditions in the southern and eastern regions in summer. This is due to the location of the region in the transitional zone between tropical and extra-tropical circulation dynamics. Furthermore, the Mediterranean area is characterized by a complex topography and by high climatic variability.

In the scope of climate variability and climate change the variations and changes of extremes are of special importance. Extreme events are not only of scientific interest but also have a profound impact on society. Moreover, the Mediterranean area represents a region where the extreme values make up a large proportion of total precipitation. Toreti *et al.* (2010) note in a study of extreme precipitation during the extended winter season at twenty Mediterranean coastal sites that at all stations precipitation extremes have an important contribution (around 60%) to make seasonal totals.

In the scope of future climate change due to anthropogenic forcing, several climate models are consistent with the projection that precipitation extremes will decrease over most regions of the Mediterranean area. Tebaldi *et al.* (2006) identify, under the consideration of nine different Atmosphere–Ocean General Circulation Models (AOGCMs), only the northern Mediterranean regions as areas with increases of precipitation extremes whereas the southern and eastern Mediterranean regions will mostly be affected by decreases up until the end of the 21st century. Emori and Brown (2005) split the projected annual extreme precipitation change from six GCMs up into a dynamic and a thermodynamic component. The overall increases over the northern Mediterranean area are due to increases of the thermodynamic component caused by increased atmospheric moisture content (whereas decreases of extreme precipitation would result from the dynamic component alone).

Beniston *et al.* (2007) compare the results from multiple runs of four regional climate models (RCMs)

* Correspondence to: E. Hertig, Institute of Geography, University of Augsburg, Germany. E-mail: elke.hertig@geo.uni-augsburg.de

regarding the 5-year return levels of precipitation and found a consistent decrease in winter precipitation south of about 45°N that is similar to the large-scale pattern of mean precipitation change (Giorgi and Lionello, 2008; Gao *et al.*, 2006; Räisänen *et al.*, 2004). In contrast, Kundzewicz *et al.* (2006), using the HadRM3-P model under SRES A2-scenario assumptions (SRES: Special Report on Emission Scenarios, Nakicenovic and Swart, 2000), find an increase in the annual maximum daily precipitation for most parts of the Mediterranean area despite decreases of the annual mean values. Using 10 RCMs forced with A1B assumptions, Rajczak *et al.* (2013) also project decreases in mean precipitation during spring and autumn, in addition to increases in heavy events. Fowler *et al.* (2007) note a large uncertainty amongst RCMs concerning the spatial pattern of extreme precipitation change, which is attributed to the influence of the driving GCMs. From the analysis of the RCM ensemble used by Fowler *et al.* (2007) smaller increases, or potentially decreases, in longer duration extremes (10 day precipitation intensities) are determined for southern Europe. Goubanova and Li (2007) use a variable-grid AGCM and A2-scenario assumptions to show that precipitation extremes will increase in the Mediterranean region in all seasons except summer. Kyselý *et al.* (2012) obtain a similar pattern of extreme precipitation change from a RCM ensemble and note a higher consistency of the projected changes for winter and in general for short-term rather than multi-day extremes. However, Flaounas *et al.* (2013) note from dynamical downscaling with the Weather Research and Forecasting (WRF) model an overestimation of extreme rainfall by 50% and more in most parts of the Mediterranean area in winter, and a smaller bias (±20%) in summer. The authors find that the subsequent application of statistical downscaling improves the spatial correlation of the rainfall patterns and can be used to correct temporal biases. Tolika *et al.* (2008) apply two statistical downscaling techniques for the simulation of future extreme rainfall changes over Greece. They note a general underestimation of the natural variability and a spatial incoherent picture of change. In a study by Hertig *et al.* (2011), which used RCM output and results from statistical downscaling, mainly slight decreases are found in the number of extreme precipitation events of up to about one day for the Mediterranean area during autumn and winter up until mid-21st century. Regarding the intensity of the extremes, a small-scale pattern of change arises with wintertime increases most likely at topographical elevations exposed to the West, where the uplift of humid air profits by the increase of atmospheric moisture under climate change conditions. Also in a statistical downscaling assessment of extreme precipitation changes in southern France, Tramblay *et al.* (2011) find an upward trend in the humidity flux towards the end of the 21st century. This in turn leads to an increase in the seasonal number of extreme precipitation events, although does not impact greatly on the magnitude of the extremes.

Despite considerable advances regarding the physical realism and the spatial resolution of AOGCMs, downscaling of the large-scale dynamical model output is very important for assessing regional climate change. This includes, in particular, extreme events, which are given focus in this study. In this context, statistical downscaling presents a computationally inexpensive technique which can be adapted for a wide range of applications. A recent evaluation of statistical downscaling models for precipitation can be found in Maraun *et al.* (2010). For the statistical downscaling of extremes, promising approaches have been introduced and/or developed further in the last few years. For instance, Benestad (2010) re-calibrates the output of analogue models of daily precipitation using the probability density function from downscaling monthly precipitation in order to produce a more reliable statistical distribution of the precipitation amounts, particularly near the upper tail. Vrac and Naveau (2007) apply a probability mixture model of Gamma and Generalized Pareto distributions within a stochastic weather typing approach to improve the representation of local extreme precipitation. A further approach to model extremes are based on generalized linear models (GLMs), which are used for example by Yang *et al.* (2005) to model station-based daily rainfall data in southern England and by Benestad (2007) to infer extreme rainfall over Northern Europe. A range of further applications of GLM-based analyses within climate research is given by Chandler (2005).

In this study, changes of extreme precipitation events in the Mediterranean area up until the end of the 21st century are analysed with a statistical downscaling approach using GLMs. In the region under consideration, extreme precipitation is connected to variations of the large-scale circulation as well as to convective activity. Thus, commonly used predictor variables as well as novel ones, particularly describing convection, are considered. A detailed description of the predictors used can be found in Section 2. As already mentioned, GLMs are chosen to downscale different percentile-based indices of extreme precipitation. Furthermore, a specific statistical ensemble technique is used for the assessments to account for non-stationarities in the predictors–predictand relationships. It yields a range of possible future evolutions of extreme precipitation. A detailed description of the methodology is given in Section 3. Overall, the analyses focus on the improvement of the statistical techniques to downscale precipitation extremes. In addition, other sources of uncertainty are taken into account in the following ways: uncertainties arising from the future emission scenarios by using two different scenarios, the A1B- and the B1-scenario; uncertainties due to the particular GCM response to a given forcing by applying multiple runs for each scenario; and uncertainties related to the structure of climate models and their parameterizations by using output of two different GCMs (ECHAM5 and HadCM3). The results of the assessment of extreme precipitation in the Mediterranean area up until the end of the 21st century are discussed in Section 4 and some conclusions are drawn in Section 5.

2. Data

2.1. Predictand

High-resolution precipitation data for the Mediterranean land areas are taken from the E-OBS gridded dataset (Haylock *et al.*, 2008). This data set contains daily values of precipitation on a $0.25^\circ \times 0.25^\circ$ grid for the European land areas. In a first step all grid boxes within the area from 27°N to 46.5°N and from the North Atlantic Ocean to 40.5°E are considered which feature a characteristic Mediterranean precipitation regime, i.e. show a summer minimum of precipitation. Further specifications of the Mediterranean rainfall regime, like its spatial variability, are described in Hertig *et al.* (2011). In a next step the time series at each grid box are checked whether they contain less than 20% missing values and less than 80% zero-rainfall amounts. This mainly affects grid boxes in the area of northern Africa, which are excluded from further analysis. In this context it has to be emphasized that the utility of the data set is generally reduced over northern Africa due to the very poor station coverage over this area (Haylock *et al.*, 2008). The data is used in a seasonal context and thus the selection procedure results in 2843 grid boxes in spring (March to May), 2520 grid boxes in summer (June to August), 2857 grid boxes in autumn (September to November), and in 2825 grid boxes in winter (December to February). The observational time period used in this study covers the years from 1950 to 2006.

2.2. Predictors

Essential within the scope of statistical downscaling is the choice of the predictor variables. For example Hertig and Jacobeit (2008) show for mean precipitation in the Mediterranean area that the downscaling results of future precipitation change are significantly influenced by whether or not humidity is included as a predictor. In this context, it is necessary to consider that while a predictor may or may not appear as important when developing the downscaling model under present climate conditions, the changes in that predictor under a future climate may be decisive for determining the character and amount of climate change. Furthermore, it is vital to select an optimal predictor set which exhibits physically meaningful relationships to the local predictand and which describes a significant part of the variations of the predictand variable. Thus, the main criteria for large-scale variables to be qualified as predictors, in accordance with Wilby *et al.* (1999) is that they should carry climate change information and strongly correlate with the surface variables of interest. In addition, they should be reliably simulated by GCMs. In this context GCM uncertainties have to be mentioned which are associated with the spatial and temporal resolutions, the issues of discretisation and parameterizations, the dependence on initial conditions, and the degree of reproduction of several important variables. As a result CMIP3 (Coupled Model Intercomparison Project phase 3, Meehl *et al.*, 2007a) models commonly overestimate the meridional pressure gradient in the North Atlantic/European sector (van Ulden and van Oldenborgh, 2006). But in general the pressure-related variables are regarded as reliable and useful predictors to assess regional climate change. Regarding specific humidity, Willett *et al.* (2010) find that the range of CMIP3 seasonal climatology and variance encompasses the observations. The models also reproduce the magnitude of observed inter-annual variability over all large regions.

In the Mediterranean area extreme precipitation is induced by advection and/or it is a matter of convective precipitation. In this context various predictors are chosen, including more 'classical ones' as well as novel ones, especially for the description of convective processes: To give equal weight to predictors describing the large-scale circulation and to variables characterizing convection four variables are selected in each case. Atmospheric factors describing the large-scale circulation are geopotential heights at the 700 hPa and 500 hPa levels as well as the u- (zonal) and the v- (meridional) wind components at the 850 hPa level. The choice is based on own analyses (Hertig *et al.*, 2011; Hertig and Jacobeit, 2008) in addition to various publications on climate variability and statistical downscaling. Sea level pressure and/or geopotential height anomalies of different levels are widely used in the context of studying precipitation extremes in the Mediterranean area, e.g. to characterize extreme precipitation at Mediterranean coastal sites in winter (Toreti *et al.*, 2010), for the explanation of the variability of extreme winter rainfall in Europe, including the northern Mediterranean areas (Haylock and Goodess, 2004), in the scope of a statistical downscaling study for southern France (Vrac and Yiou (2010), and for a statistical downscaling assessment of extreme winter precipitation in the Emilia-Romagna region of Italy (Busuioc *et al.* (2008). Concerning the wind components, Cavazos and Hewitson (2005) show in a study about the performance of reanalysis variables in statistical downscaling of daily precipitation that the meridional wind component appears in the list of the top variables, suggesting an influence from surface meridional synoptic systems on precipitation.

According to Doswell (1987), a triad of ingredients is required for deep convection: moisture, conditional instability, and a source of lift. In this study, the information about ample moisture in the lower troposphere is provided by taking specific humidity data of the 850 hPa and 700 hPa levels as predictors. Convective instability with a steep enough lapse rate may be assessed by examining convective available potential energy (CAPE) or other proxies like the Showalter-Index (Showalter, 1953), which is selected in the study being on hand. The third requisite for deep convection, a source of lift, can be restated as a requirement of sufficiently small convective inhibition (CIN, Myoung and Nielsen-Gammon, 2010). The mentioned authors assess that CIN is particularly important for precipitation variability over land areas and that the existence of a large amount of CIN tends to inhibit the initiation of convection despite substantial

conditional instability and moisture availability. Therefore, a proxy for CIN calculated from reanalysis data is also selected as a predictor. Details of the calculation procedure of the convective indices (Showalter-Index and CIN) are given in Section 3.2.

It should be noted that the list of predictor variables is not complete and could be extended by further potential predictor variables, like for example GCM-simulated precipitation as a predictor for small-scale precipitation. This is done for example by Schmidli *et al.* (2006) for the European Alps or by Benestad *et al.* (2007) for Fennoscandia. Also other variables like sea surface temperatures, aerosols, and cyclonic activity play a major role for extreme precipitation in the Mediterranean area. But since these variables mostly still suffer from large uncertainties in the reanalysis data and in the GCMs, they are not included in this study.

The predictor variables are obtained from the NCEP/NCAR reanalysis project (Kalnay *et al.*, 1996; Kistler *et al.* 2001). Geopotential heights of the 700 hPa and 500 hPa levels are taken in the area 20°N–70°N and 40°W–50°E. The other predictor variables (850 hPa and 700 hPa specific humidity, 850 hPa U-wind and V-wind components; also temperature, humidity, and pressure of various levels for the calculation of convective indices, for details see Section 3.2..) are selected for the area 27.5°N–52.5°N and 12.5°W–42.5°E (horizontal resolution of $2.5^\circ \times 2.5^\circ$ in each case). The limits of the larger area for the geopotential heights incorporate the major influences on the Mediterranean area, especially the mid-latitude westerlies in the upstream area. Tests with different sizes of the domain show that for the other predictor variables a smaller size, concentrated over the target area itself, works better within the statistical downscaling procedure.

Model output of the ECHAM5/MPI-OM-AOGCM (Roeckner *et al.*, 2003, 2006) and of the UKMO-HadCM3-AOGCM (Gordon *et al.*, 2000; Pope *et al.*, 2000) is taken for the representation of the model predictors within the scope of the statistical downscaling approach. The forcing uses observed greenhouse-gas emissions for 1950–2000 and SRES A1B- as well as B1-scenario emissions from 2001 to 2100. Three 20th century-, three A1B-, and three B1-scenario simulations of the ECHAM5 model and one simulation in each case of the HadCM3 model are used as boundary conditions. General data pre-processing includes the fitting of the horizontal resolution of the model output data (T63, 1.875°) to those of the observed predictor data ($2.5^\circ \times 2.5^\circ$).

3. Methodology

3.1. Calculation of extreme precipitation indices

Different percentile-based indices of extreme precipitation, characterizing the frequency and the seasonal amount of heavy rainfall events, are calculated for each grid box with a seasonal resolution: the number of events exceeding the 95th percentile of daily precipitation from the reference period 1961–1990 (R95N) and the total amount of precipitation from these events (R95AM) (Moberg *et al.*, 2006). Only days with a minimum rainfall amount of 0.1 mm are included in the calculation of the 95th percentile; dry days are omitted. In the context of using the 95th percentile one has to keep in mind that this threshold is rather moderate to define the area of extremes within the whole precipitation distribution. However, it is chosen in this study because it includes a larger number of events, easing a sound statistical analysis. Furthermore, one has to take into account that the spatial interpolation, inherent to the gridded dataset, has a large impact on the magnitude of extremes. Thus the interpolation methodology reduces the intensities of the extremes, with a reduction in all extremes higher than the annual 75th percentile for precipitation (Haylock *et al.*, 2008).

3.2. Calculation of convective indices

The Showalter-Index (Showalter, 1953) is used to describe convective instability and is calculated as follows:

$$\text{Showalter-Index} = T_{500} - \text{TLCL}_{500} \tag{1}$$

T_{500} is the temperature at 500 hPa and TLCL_{500} is the temperature that a parcel will achieve if it is lifted dry adiabatically from 850 hPa to its condensation level and then moist adiabatically to 500 hPa. Thus the difference between the 500 hPa temperature and the air parcel lifted from the lifting condensation level to 500 hPa provides an estimate of the instability. It should be noted that the Showalter-Index does not take into account the atmospheric conditions below 850 hPa. To calculate the Showalter-Index from the reanalysis data and the GCM data, 850 hPa air temperature and 850 hPa relative humidity are used to calculate the dew point temperature and from this the height of the condensation level at each grid box (253 grid boxes in the area 27.5°N–52.5°N and 12.5°W–42.5°E, horizontal resolution of $2.5^\circ \times 2.5^\circ$). The difference of the geopotential heights of the 500 hPa and 850 hPa levels is used to obtain the absolute height of the air parcel to be lifted. The air parcel is lifted dry adiabatically to its condensation level and then moist adiabatically. In doing so the value of the moist adiabatic temperature gradient is readjusted in steps of 100 metres uplift.

CIN characterizing the presence of large- or small-scale lifting mechanisms, is represented by a proxy, following Myoung and Nielsen-Gammon (2010):

$$\text{CIN} = T\,(\text{inv}) - T_\text{d}\,(\text{s}) \tag{2}$$

Thus CIN depends on the dewpoint temperature at the surface (T_d(s)) and the virtual temperature at some level just above the mixed layer or within a capping inversion (T(inv)). According to Myoung and Nielsen-Gammon (2010) surface pressure values can be taken to obtain the associated best proxy levels for the calculation of

Table 1. Surface pressure values and associated best proxy levels for the calculation of CIN, adopted from Myoung and Nielsen-Gammon (2010).

Surface pressure (hPa)	Assumed best proxy level (hPa)
<870	600
870–980	700
980–1013.25	850
>1013.25	925

the proxy of CIN. In this study for each grid box of the predictor domain the assumed best proxy level is determined by taking Table 1 as reference (Table I in Myoung and Nielsen-Gammon, 2010). Thus for the calculation of CIN from the reanalysis data and the GCM output surface dewpoint temperatures are required, as well as surface pressure (to obtain the best proxy level), and also specific humidity and air temperature of the best proxy level (for the calculation of the virtual temperature). Like the other predictors (except geopotential heights) CIN is used in the area 27.5°N–52.5°N and 12.5°W–42.5°E.

3.3. Principal component analysis

S-mode, Varimax-rotated principal component analysis (PCA; Preisendorfer, 1988) is applied to each predictor field to reduce dimensions of the data and to define independent (orthogonal) spatial centres of variation within each data set. PCAs are carried out for each of the five calibration periods (for details on calibration/verification see Section 3.4.2.) in the observational period from 1950 to 2006. The determination of the number of PCs to be extracted follows the approach of Philipp *et al.* (2007) and is based on the criterion that each PC has to be uniquely representative for at least one input variable. Representativeness is assumed when the maximum loading of a variable on a particular PC is at least one standard deviation greater than the other loadings of this variable on the remaining PCs; additionally, this maximum loading has to be statistically significant at the 95% level.

Depending on the particular season, 8 (winter) up to 12 (summer) PCs with overall explained variances (EVs) between 90% and 96% are obtained from PCA of the geopotential height fields. The wind components (note the smaller domain of these variables compared to the geopotential height fields) yield 8 up to 16 PCs, the specific humidity fields 10 up to 18 PCs, and the convective indices 7 up to 15 PCs, all of them with EVs of about 90%. Generally, in the summer season more PCs are necessary to describe the large-scale variation within a predictor field compared to the other seasons.

Then, the time series of the extreme indices at each individual grid box are linked to the large-scale atmospheric circulation represented by their corresponding PC time series. The predictand–predictor relationships are subsequently used to assess the response of extreme precipitation to changes of the large-scale predictors. For this purpose the reanalysis data in the verification periods as well as the GCM model data of the control runs, the A1B- and B1-scenario runs are projected in each case onto the existing PCs of the observational period to obtain new predictor time series.

3.4. Statistical downscaling approach

3.4.1. Generalized linear models

For deriving relationships of the large-scale circulation with R95N (the number of events exceeding the 95th percentile of daily precipitation, thus a description of the frequency of heavy precipitation events) and with R95AM (precipitation total from events exceeding the 95th percentile of daily precipitation, an index which characterizes the total seasonal amount of heavy precipitation events) at each individual grid box, GLMs are applied. Deriving a GLM involves three decisions: the distribution of the data, the function of the mean modelled as linear in the predictors, and the predictors (McCulloch and Searle, 2001).

R95N is a matter of counted data and therefore the analysis can be based on the idea of a log-linear model. In such a model the systematic effects are multiplicative and the nominal error distribution is Poisson (in contrast to the Gaussian model where the effects are additive and the error distribution normal, Mc Cullagh and Nelder, 1989). A maximum likelihood estimation is used to derive a model for the expectation value $E(\cdot)$ for a variable Y_t at any given time t:

$$E\,(Y_t) = \eta_t = g\,(\mu_t) = g\sum_{1}^{p} \chi_{tj}\beta_j \qquad (3)$$

Here, η_t is a linear predictor (i.e. a linear combination of the explanatory variables), μ_t is the mean for the PDF at time t (GLM predicts the mean μ of the distribution rather than the realization itself), g is the canonical link function (one of the main ideas of GLMs is to get away from the idea of transforming the data. The strategy is to apply a link function to the mean of the response and fit the resulting model by the method of maximum likelihood), χ_{tj} is the value of the jth covariate for observation t, and β_j are parameters whose values have to be estimated from the data. The variance function used by the GLM is $V(\mu) = \mu$, assuming a Poisson process with the canonical link log ($\eta = \log \mu$).

The modelling of R95AM is a twofold challenge: first there is a presence-absence problem in such a sense that there are years with zero values in the time series (years when no events exceed the 95th percentile). Second a correct distribution of R95AM and an appropriate link function has to be determined. In order to solve these issues the zero values in the time series are disregarded and a GLM is established for the resulting continuous data. In the calibration periods the information about the occurrence of an extreme event is retrieved from the original time series. So as to get the complete time series in the verification periods and in the model periods the information about zero values is taken from the modelling of R95N.

To test the continuous time series (zero values left out) of R95AM regarding the correct distributional form, the data is fitted to possible theoretical distributions and chi-square and Kolmogoroff and Smirnov goodness-of-fit tests are applied on the fit. It is found that the variance function is best described by a Gamma distribution with $V(\mu) = \mu^2$. The associated canonical link for this distribution is reciprocal ($\eta = \mu^{-1}$). Although this canonical link leads to desirable statistical properties of the GLM (Mc Cullagh and Nelder, 1989), it is found that the systematic effects in the models are additive not on a reciprocal scale, but given by the link log. Therefore a maximum quasi-likelihood-model is established with the canonical link log ($\eta = \log \mu$).

In order to ensure reliable results from the GLMs, 100 iterations of bootstrapping are performed for each model. Only predictor variables which have a significant regression coefficient (t-test with a significance level of 0.1) are included in the models. As a measure of the overall ability of a model the discrepancy of the fit is examined by the generalized Pearson χ^2 statistic (Mc Cullagh and Nelder, 1989). Further diagnostic procedures are applied to examine the residuals by looking at the deviance residual and the Pearson residual (Mc Cullagh and Nelder, 1989). Additional measures of the model performance are used within the context of the calibration and verification procedure which is described in the next section.

3.4.2. Calibration and verification

A statistical ensemble approach is applied for the assessment of future changes of the extremes: the whole study period 1950–2006 is divided into five calibration periods, each comprising 47 years (46 years in winter, because December of the last year has no subsequent January and February values to calculate the winter value). Each of the five verification periods comprises 10 years which are not included in the corresponding calibration periods: 1956–1965, 1966–1975, 1976–1985, 1986–1995, 1996–2005. The predictors–predictand relationships are established in the different specified calibration periods, resulting in multiple models which cover a range of the observed natural variability and which give a particular quantification of uncertainties.

The ability of a model to describe a stable connection between the large-scale circulation and Mediterranean extreme rainfall is determined by means of the correlation coefficients between modelled and observed extremes indices in the calibration as well as in the verification periods. Besides from the correlation coefficients inferences can be made regarding phase errors. Also, mean error and bias are consulted to judge model performance. Additionally, the reduction of variance (RV) is calculated, being similar to the root mean squared skill score.

$$\mathrm{RV} = \left(1 - \left(\frac{\mathrm{rmse}_{\mathrm{model}}}{\mathrm{rmse}_{\mathrm{reference}}}\right)^2\right) \times 100\,[\%] \qquad (4)$$

with rmse being the root mean squared error. The 'reference' in this study is the mean of the observations from the verification sample. $\mathrm{RV} = 100\%$ would mean a perfect model, $\mathrm{RV} = 0\%$ implies no improvement compared to the simple use of the sample climatology, and $\mathrm{RV} > 0\%$ indicates some improvement by the model results. In this study all statistical models with $\mathrm{RV} > 0$ in calibration and verification pass the model check and are subsequently used for further analyses. The statistical model performance in the scope of the assessment of R95N is given in Section 4.1.1., for R95AM in Section 4.2.1..

At first the downscaling procedure is carried out for each predictor type separately. Subsequently a joint analysis is conducted with the PCs selected in the single predictor analyses serving as potential predictors in the multi-type predictor analysis. Then for each grid box and calibration/verification period the decision is made which model to use for the assessment of future extreme precipitation changes: the model with the best performance is chosen in each case. This implies that for a certain grid box a statistical model ensemble becomes available which includes different versions of the observed relationships. Thus, the statistical models of an ensemble comprise a larger range of the observed natural variability including differing influences from the various predictor types. A discussion on this subject is given in Section 4.1.2.

In order to distinguish between the role of advection and convection further assessments are performed, once under the sole inclusion of the predictors describing the large-scale circulation and once again under the sole application of the convective predictors. The outcome of this analysis is presented in Section 4.1.3. for R95N and in Section 4.2.3. for R95AM.

Subsequently, results are examined under the consideration of all predictor variables. The assessment results are presented as the changes in the future time slice 2070–2099 in relation to the control run period 1961–1990. In the future period A1B- and B1-scenario conditions are considered under the use of multiple ECHAM5 runs as well as HadCM3 output. Also, for each grid box the upper and lower limits of the 95% bootstrap confidence interval (DiCiccio and Efron, 1996) for the assessed change in the two 30-year time periods are calculated. The bootstrap confidence intervals are calculated taking 1000 iterations. In addition, the ensemble mean changes for the future time period 2070–2099 in relation to the reference period 1961–1990 are tested for significant differences (95% level) using the non-parametric U-test (Mann and Whitney, 1947).

4. Results

4.1. Number of extreme precipitation events per season (R95N)

In this section the results of the assessment of R95N using the combined information of each single-predictor

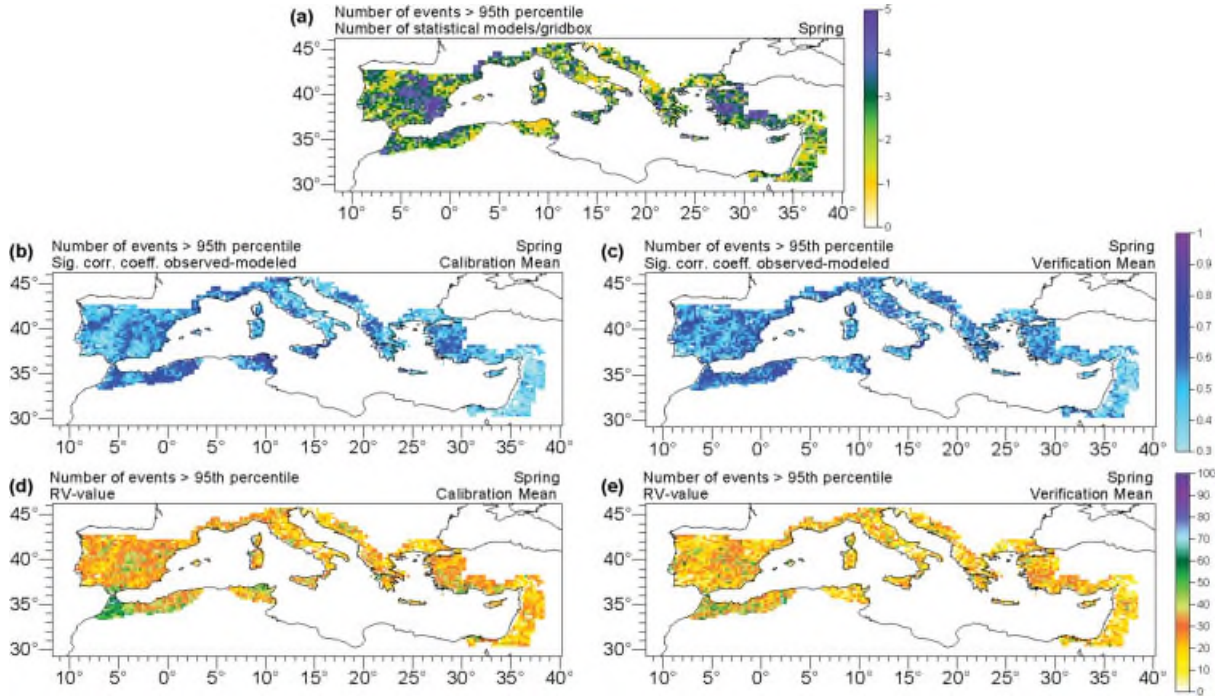

Figure 1. Statistical model performance for R95N in spring (March to May). (a) Size of the statistical model ensemble per grid box (maximum: 5). (b) Correlation coefficients between modelled and observed extremes indices in the calibration (mean value over the statistical model ensemble). (c) Correlation coefficients between modelled and observed extremes indices in the verification (ensemble mean). (d) RV for the calibration (mean value over the statistical ensemble). (e) RV for the verification (ensemble mean).

analysis (700 hPa, 500 hPa geopotential heights, 850 hPa u- and v-wind components, 850 hPa, 700 hPa specific humidity, CIN, Showalter-Index) and of the multi-type predictor analysis of all predictor types are discussed (for details on the generation of this assessment see Section 3.4.2.)

4.1.1. Model performance

Under the use of the five calibration/verification periods for the total of 2728 grid boxes, 7501 downscaling models can be established in spring. This implies that on average over all grid boxes for almost three (2.75, resulting from 7501 divided by 2728) out of five calibration/verification periods a downscaling model is available. Yet the size of the statistical model ensemble per grid box is not equally distributed over the Mediterranean area, but varies over space. This can be seen for spring in Figure 1(a), where the central and south-eastern Iberian Peninsula as well as eastern and southern Turkey represent regions with three up to five statistical models per grid box, whereas for example around Mediterranean Tunisia and central Italy often only one statistical model is available. The total number of models and the number of models per grid box for summer, autumn, and winter can be seen in Table 2 and in Figures 2(a), 3(a), and 4(a), respectively. In summer the total number of grid boxes entering the analyses is lowest compared to the other seasons, due to the completely dry character of the south-eastern Mediterranean area. In autumn the number of downscaling models which can be established for R95N traces the areas where mean precipitation is high, with high numbers over the western Iberian Peninsula and the northern Mediterranean area (Figure 3(a)). In winter, along with the Iberian Peninsula, the eastern Mediterranean areas depict regions of high modelling power (Figure 4(a)). Overall the size of the statistical model ensembles shows a clear connection to the total precipitation amounts, pinpointing to the fact that the statistical models perform best in regions and seasons where and when the maximum of mean precipitation can be observed.

To give further insight on the quality of the assessments the correlation coefficients between statistically modelled and observation-based values of R95N in the calibration and verification periods are shown in panels (b) and (c) of Figures 1–4. The coefficients are the mean over the statistical ensemble. In addition, panels (d) and (e) of Figures 1–4 illustrate the ensemble mean of the RV-value for the calibration and verification. From these figures it becomes evident that for the Mediterranean area as a whole the best performance is achieved in winter (Figure 4). In the other seasons the quality of the assessments varies in a very small-scale spatial pattern, only the eastern Mediterranean area emerges as a more problematic region regarding the establishment of robust downscaling models. Overall, it becomes apparent that the assessments of R95N can be seen as very positive in terms of the number of good models, the correlation coefficients, and the RV-values. However, some drawbacks of the assessments which do

Table 2. Frequency of the predictors selected as independent variables within the statistical downscaling models for the number of events exceeding the 95th percentile of daily precipitation (R59N). The total number of models consists of the number of grid boxes times the number of calibration/verification periods where a downscaling model can be established. A multi-type predictor assessment in terms of a specific combination of individual predictor variables is used where it yields better results than an assessment with a single predictor. From top to bottom: results for spring, summer, autumn, and winter.

R95N/Spring Predictor	Number of models 7501	Percentage 100%
700 hPa geopotential heights	1177	15.69
multi-type predictor combination	1166	15.54
700 hPa specific humidity	899	11.99
850 hPa meridional wind	885	11.80
850 hPa zonal wind	857	11.43
500 hPa geopotential heights	669	8.92
850 hPa specific humidity	661	8.81
Convective inhibition	625	8.33
Showalter-Index	562	7.49
R95N/Summer **Predictor**	**Number of models** **6928**	**Percentage** **100%**
850 hPa specific humidity	985	14.22
multi-type predictor combination	984	14.20
700 hPa geopotential heights	914	13.19
500 hPa geopotential heights	861	12.43
850 hPa zonal wind	766	11.06
Convective inhibition	726	10.48
700 hPa specific humidity	630	9.09
Showalter-Index	541	7.81
850 hPa meridional wind	521	7.52
R95N/Autumn **Predictor**	**Number of models** **7717**	**Percentage** **100%**
700 hPa geopotential heights	1126	14.59
multi-type predictor combination	1083	14.04
850 hPa meridional wind	984	12.75
500 hPa geopotential heights	913	11.83
850 hPa zonal wind	831	10.77
Showalter-Index	760	9.85
850 hPa specific humidity	733	9.50
700 hPa specific humidity	711	9.21
Convective inhibition	576	7.46
R95N/Winter **Predictor**	**Number of models** **8815**	**Percentage** **100%**
850 hPa specific humidity	1431	16.23
700 hPa geopotential heights	1415	16.05
850 hPa zonal wind	1087	12.33
700 hPa specific humidity	906	10.28
500 hPa geopotential heights	884	10.03
multi-type predictor combination	872	9.89
Convective inhibition	802	9.10
850 hPa meridional wind	777	8.81
Showalter-Index	641	7.28

not become apparent directly have also to be mentioned: no systematic bias is found in the assessments with the exception of the almost completely dry regions in the summer season. In these cases there is a positive bias towards too large values of R95N in those years, when the predictor exerts a strong signal on the predictand.

4.1.2. Predictors

Table 2 gives the information on how often a specific predictor is selected in the downscaling models in terms of absolute numbers and in relation to the total number of models per season. It becomes evident that the geopotential heights of the 700 hPa level as well as a multi-type predictor set are most frequently selected in the scope of the assessment of the number of extreme precipitation events in spring and in autumn. Thus, in the transitional seasons the large-scale circulation contributes strongly to the generation of extreme precipitation events (e.g. through large differences of the characteristics of different air masses within meridional circulation types). This is also supported by the very frequent selection of the 500 hPa geopotential heights and the wind components in these seasons. However, during spring a source of lift and moisture availability seem to be more essential compared to autumn. During the summer season the predictor selection shows the important requisite of moisture existence for the generation of precipitation extremes, but also the requirement of large-scale circulation anomalies to induce extreme precipitation events. In the winter season R95N is governed by the moisture availability in the lower troposphere and by the large-scale circulation. Hence, even in high winter, when the cyclonic west wind drift is at its southernmost position, the large-scale circulation is not the only influencing factor for extreme precipitation events in the Mediterranean area, but the thermo-dynamic conditions have also to be taken into account. Overall it becomes clear that the distinction between different seasons is important within the context of the predictor choice in order to account for the varying influence of the predictors on extreme precipitation during the course of the year.

In order to demonstrate the effects which the predictor selection exerts on the temporal evolution of R95N under scenario conditions, for each season the statistically modelled time series of one arbitrarily chosen (the availability of a complete statistical ensemble with five ensemble members is set as the only prerequisite) grid box are illustrated in Figure 5. Shown are the modelled values of each of the five statistical ensemble members, the ensemble mean, and the polynomial fit of each model for the control run period 1950–2000 and the scenario period 2001–2100 of the ECHAM5-A1B run1. The predictors used in each of the five models are depicted in the legend of Figure 5. Figure 5(a) shows the assessed number of extreme precipitation events in spring for a grid box located in Central Iberia. It can be seen that the inter-annual variation of the extreme precipitation frequency differs from model to model depending on

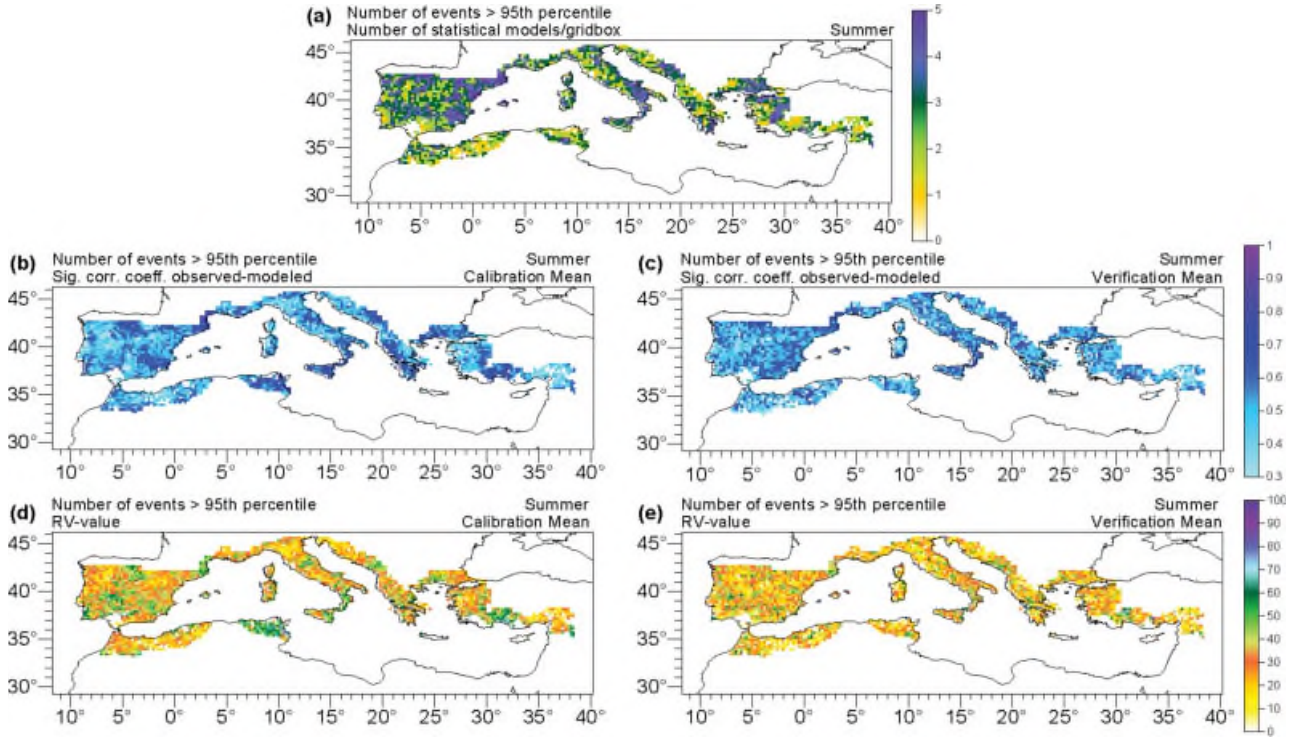

Figure 2. Statistical model performance for R95N in summer (June to August). (a) Size of the statistical model ensemble per grid box (maximum: 5). (b) Correlation coefficients between modelled and observed extremes indices in the calibration (mean value over the statistical model ensemble). (c) Correlation coefficients between modelled and observed extremes indices in the verification (ensemble mean). (d) RV for the calibration (mean value over the statistical ensemble). (e) RV for the verification (ensemble mean).

the signals given by a specific predictor. Nevertheless, when looking at the polynomial fit of the models, it becomes visible, that these are mainly in accordance. Only the assessment which uses geopotential heights of the 700 hPa level as predictor shows an overall downward trend until the end of the 21st century, in contrast to the other assessments where almost no change of R95N is visible. The predictors which are chosen in the GLMs indicate that for this area in spring the configuration of the large-scale circulation is of prominent importance. But also CIN is selected as predictor in one of the models, suggesting the question whether or not high pressure within the context of the Azores High is already established at this time of the year and therefore convective events are suppressed. Looking at the results in Figure 5(b) for a grid box located in Sicily reveals a larger range between the different models with almost no change (predictor: v-wind component) up to a strong increase of R95N from around the year 2010 until the end of the 21st century (predictor: 850 hPa specific humidity). Also, a generally very high inter-annual variability of extreme precipitation events becomes visible in this example for the summer season. The selected predictor variables humidity, wind components, and CIN indicate that advective moisture transport in the lower troposphere and local lifting mechanisms plays a major role for R95N at this central Mediterranean island in summer. Figure 5(c) shows the temporal evolution of R95N in autumn for a grid box located in Croatia. In this example all statistical ensemble members are in a good agreement, yielding an overall decrease of R95N of about 1 day until the end of the 21st century. The selection of humidity and CIN as predictors points to the importance of moisture availability and unhindered conditions within orographically induced air mass uplift in the area of the Dinaric Alps. Figure 5(d) presents the results for a grid box in Israel in the winter season. Regarding the overall trend the statistical models are in good accordance with each other. As predictors geopotential heights of the 700 hPa level and specific humidity of the 850 hPa level are primarily selected, indicating that moist circulation anomalies (e.g. in the context of low-pressure systems which are generated or are travelling over the eastern Mediterranean Sea in winter) are responsible for the generation of extreme precipitation events. Additionally, the selection of the Showalter-Index in one of the statistical models points to the relevance of conditional instability of the atmosphere.

4.1.3. Thermo-dynamic factors versus large-scale circulation

The aforementioned results and the overall assessment which will be described in Section 4.1.4. are based on the use of the whole predictor set which describes the large-scale circulation as well as convective conditions. But at first the question about the specific contributions of the different predictor types to the assessed changes will be addressed in this subsection

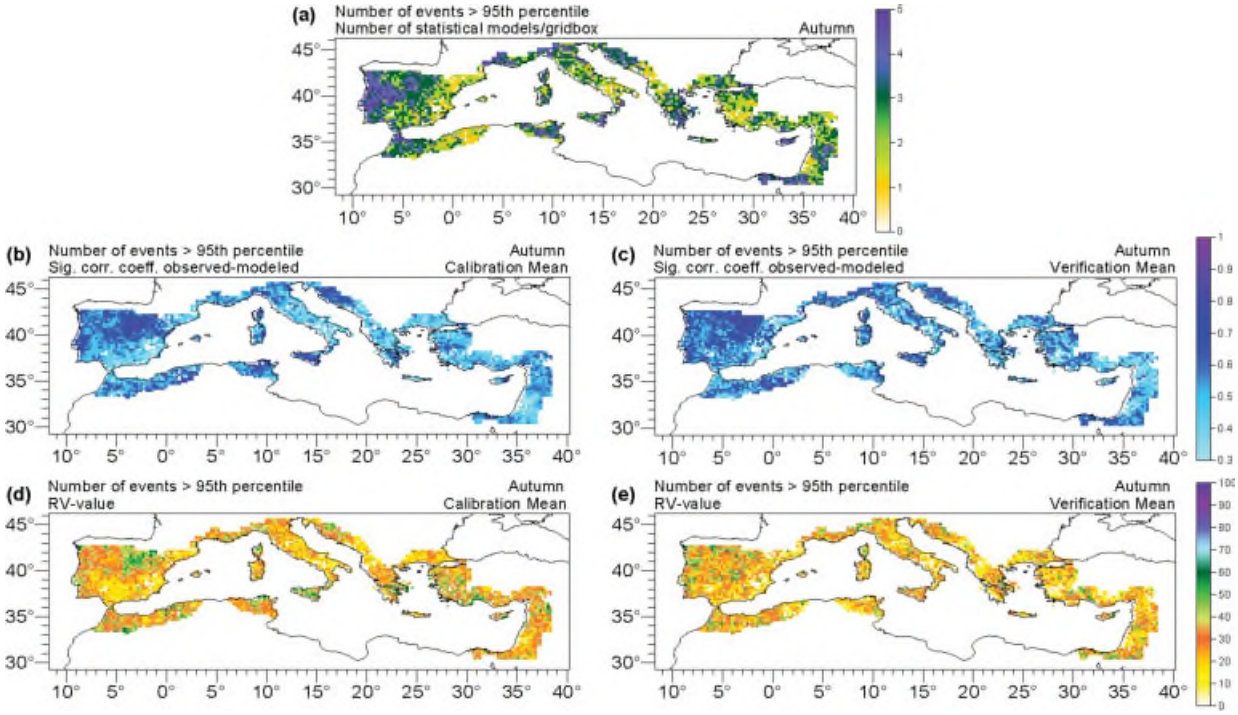

Figure 3. Statistical model performance for R95N in autumn (September to November). (a) Size of the statistical model ensemble per grid box (maximum: 5). (b) Correlation coefficients between modelled and observed extremes indices in the calibration (mean value over the statistical model ensemble). (c) Correlation coefficients between modelled and observed extremes indices in the verification (ensemble mean). (d) RV for the calibration (mean value over the statistical ensemble). (e) RV for the verification (ensemble mean).

in more detail. In addition to the overall assessment considering all predictors, two further types of statistical downscaling models are realized. In one analysis just circulation-type predictors (700 hPa/500 hPa geopotential heights, 850 hPa u-,v-wind components) are used, and in the other one solely predictors describing thermodynamic conditions (850 hPa/700 hPa specific humidity, CIN, Showalter-Index).

Regarding the performance of the two different types of assessments in terms of the number of statistical models per grid box, correlation coefficients, and RV-values in calibration and verification, there is no systematic preference of one predictor set (not shown). Overall, the modelling power of the two predictor sets is of the same value, but it is reduced in relation to the model performance of the assessment using all predictors. For example, the number of statistical models which can be established is reduced to about 3/4 compared to the number resulting from the modelling with all predictors (e.g. in winter (summer) 7051 (4793) circulation-type models, 7268 (5181) thermo-dynamic models, and 8815 (6928) all-predictors models). Between the two predictor sets differences can be found on a seasonal and regional view, e.g. the circulation-type predictor set performs in general better over the western Mediterranean area in autumn and over the eastern Mediterranean area in winter compared to the thermo-dynamic predictor set. In contrast, widespread in the Mediterranean area the thermodynamic predictor set exerts a higher modelling power in summer. Comprehensively, under the cost of a lower modelling power compared to the overall assessment, the Mediterranean area can be assessed in all seasons through the sole use of one predictor set, due to the indirect containment of information from the other predictor set (for instance does a certain circulation pattern imply a characteristic humidity distribution). Nevertheless, considerable differences might appear under future climate conditions, because the relationship among different predictor variables may not be preserved. This leads to different assessment results depending on the particular predictor variables being included into the downscaling models.

The assessed changes of R95N resulting from the two different downscaling exercises are shown in Figure 6 as the difference of the future time slice 2070–2099 to the control run period 1961–1990 under the use of the ensemble mean of the three A1B-scenario runs of ECHAM5. On the left of Figure 6 the results of the downscaling using the circulation-type predictors are presented, on the right of Figure 6 the assessment based on the thermo-dynamic predictors is illustrated. It becomes visible that there are several differences between the two assessments. The most noticeable features become apparent in autumn and in the winter season.

Summing up the results clearly show that different predictor types lead to a different outcome of the statistical downscaling and point to the importance to include different predictor variables in the statistical downscaling models. This is necessary to get a comprehensive description of the observed predictand–predictor relationships

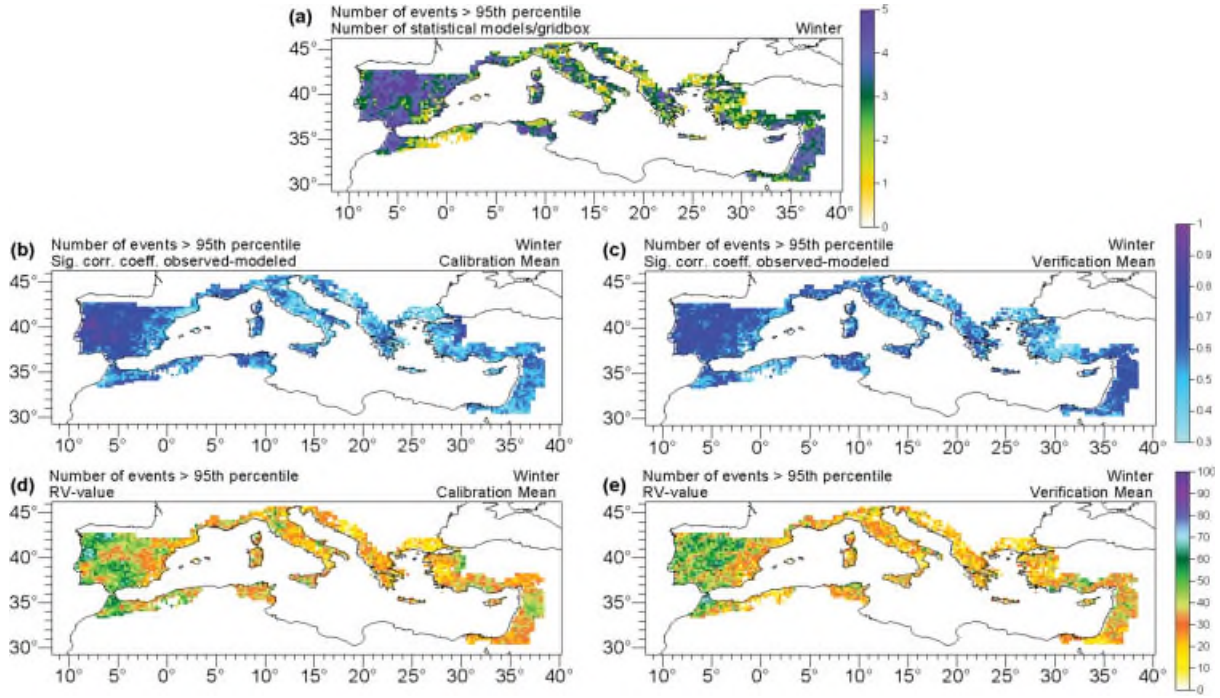

Figure 4. Statistical model performance for R95N in winter (December to February). (a) Size of the statistical model ensemble per grid box (maximum: 5). (b) Correlation coefficients between modelled and observed extremes indices in the calibration (mean value over the statistical model ensemble). (c) Correlation coefficients between modelled and observed extremes indices in the verification (ensemble mean). (d) RV for the calibration (mean value over the statistical ensemble). (e) RV for the verification (ensemble mean).

and to obtain a realistic picture of possible future climatic change on a regional basis. The assessment using all predictor variables is described in next Section 4.1.4.

4.1.4. Assessment of R95N for the 21st century

Figure 7 illustrates for the four seasons the assessment of the frequency of extreme precipitation events in the Mediterranean area up until the end of the 21st century as the change in the future time slice 2070–2099 in relation to the control run period 1961–1990. Presented is the ensemble mean which consists of 3 up to 15 statistical assessments, resulting from the use of the three ECHAM5-A1B runs, each of them multiplied by the size of the statistical ensemble (one to five ensemble members). Also shown are the significance of the changes (U-test, 95%-level) and the upper and lower limits of the 95% bootstrap confidence interval. The results are based on the assessment incorporating all predictor variables.

As an overall result in spring (Figure 7(a)) mainly decreases of up to about one day emerge over many parts of the Mediterranean area. In contrast, increases of R95N can be found over some western exposed mountainous areas of the central Mediterranean area with maximum values of about two days increase. These increases can be attributed to changes of the large-scale circulation (see Figure 6). Concerning the significance (95%-level) of the described changes, the results from the U-test can be seen in the lower part of Figure 7(a). It becomes apparent that the significance pattern is rather scattered across the Mediterranean area, but for many grid boxes the change turns out to be significant. Looking at the upper and lower limit of the 95% bootstrap confidence interval on the right side of Figure 7(a) reveals that the uncertainty of the change of R95N between the periods 2070–2099 and 1961–1990 results in possible decreases (lower limit) as well as potential increases (upper limit) for most parts of the Mediterranean area. In the summer season increases of up to about two extreme precipitation days are visible around the Tyrrhenian Sea, the Ionian Sea, and the Aegean Sea, whereas decreases of up to about one day arise for most of the western and northern Mediterranean regions (Figure 7(b)). In autumn reductions of R95N occur over many parts of the western and central Mediterranean area with values of partly more than one day decrease (Figure 7(c)). In the eastern Mediterranean area a widespread rise of the frequency is assessed. For Greece mainly thermo-dynamic factors control the increases, whereas for southern Turkey circulation changes can be attributed to the rise (Figure 6). In winter distinct increases of R95N can be seen for many parts of the Mediterranean area, most pronounced over the north-western Iberian Peninsula, the northern Dinaric Alps, around the northern Aegean Sea, and the southern coast of Turkey with a maximum of about four days (Figure 7(d)). The increases over the western Mediterranean area are mainly related to changed thermo-dynamic conditions, whereas over the central and north-eastern Mediterranean area changes of the large-scale circulation play a major role (Figure 6). Overall, the pattern of significance in winter (lower part of Figure 7(d)) is less pronounced compared to the other

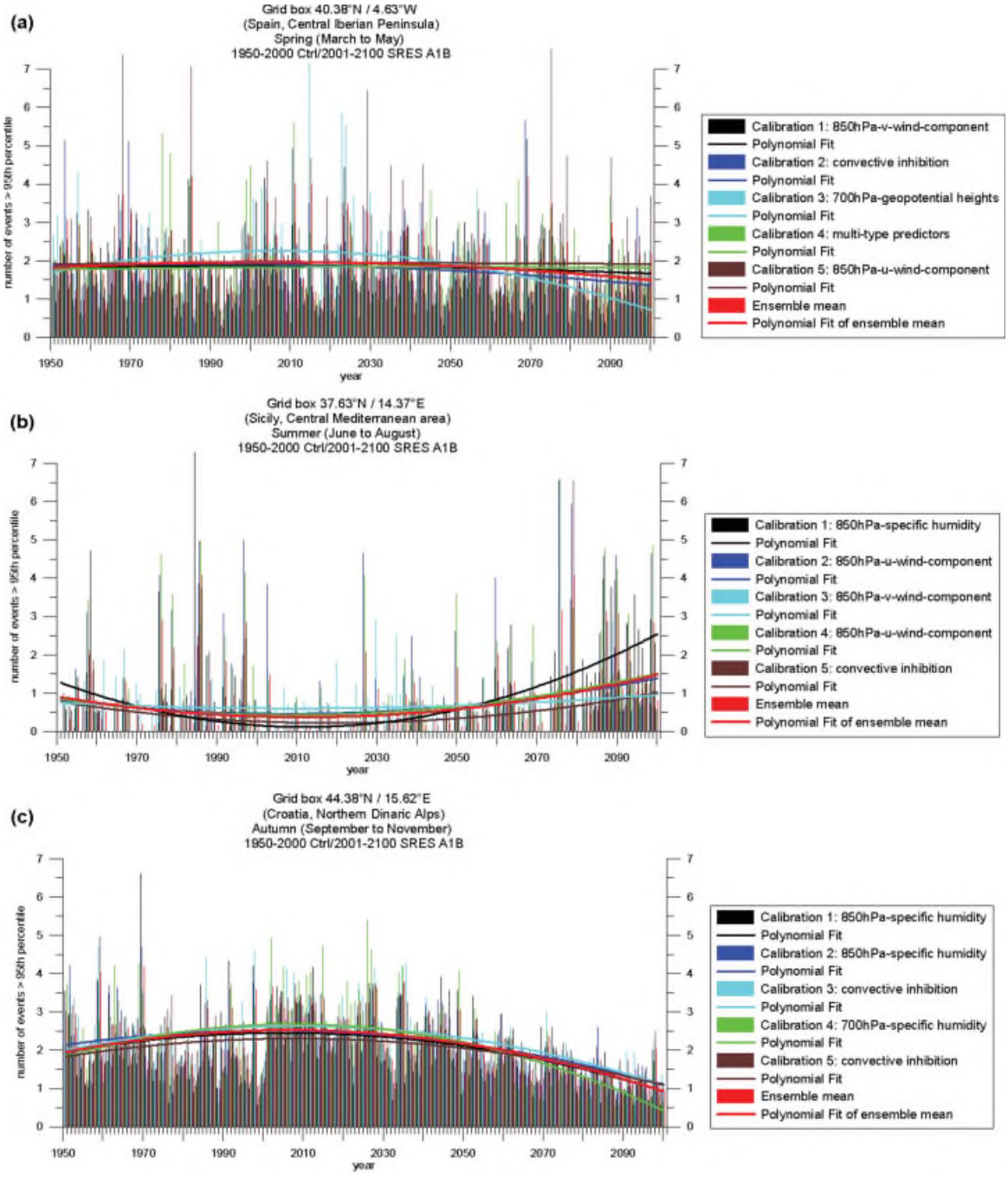

Figure 5. Temporal evolution of the number of extreme precipitation events for the time period 1950–2100 resulting from five statistical downscaling models for four example grid boxes in different seasons and regions of the Mediterranean area. Output of the ECHAM5-run1 is used for the assessment (1950–2000 control run conditions, 2001–2100 SRES A1B scenario assumptions). Shown are the values of R95N for each individual year as well as the polynomial fit of each time series.

seasons despite considerable changes of R95N over some Mediterranean areas. This is due to the generally higher values of R95N in the winter season.

The changes described above refer to the statistical downscaling of ECHAM5-A1B model output. However, the assessment using HadCM3-A1B data yields very similar results. Only small differences become apparent like for example a larger area of small increases over the Iberian Peninsula and over Syria and the Lebanon in spring and generally less pronounced but more widespread increases in the winter season. Also the statistical downscaling under the use of the B1-scenario instead of A1B-scenario leads to results comparable to the changes visible in Figure 7. Using the B1-scenario causes generally a slightly smaller amount of change, leaving the overall pattern of change unmodified. Since these further assessments do not yield differing results they are not explicitly presented here. Nevertheless, to give a comprehensive picture of the downscaled evolution of R95N, Figure 8 shows the time series of the number of extreme

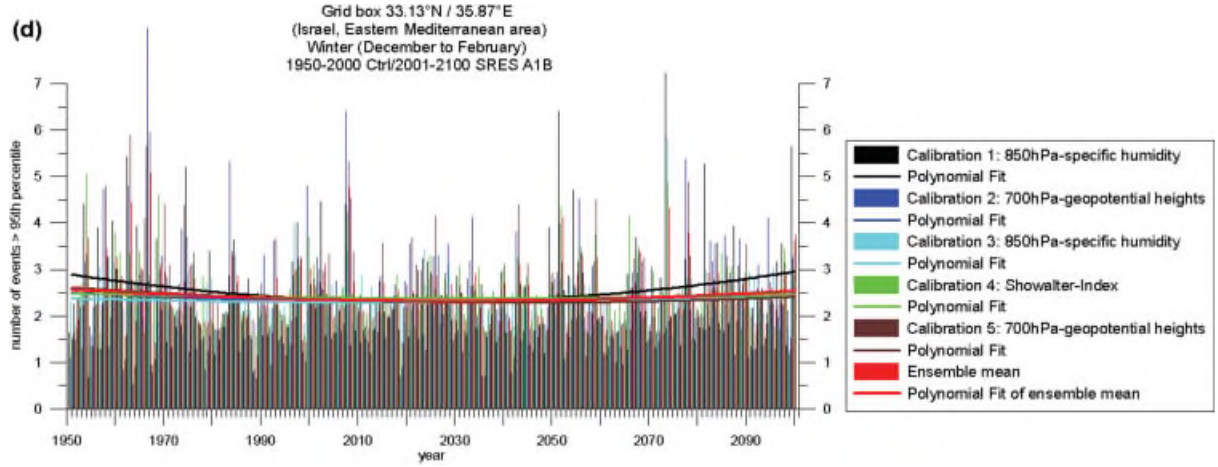

Figure 5. *Continued*.

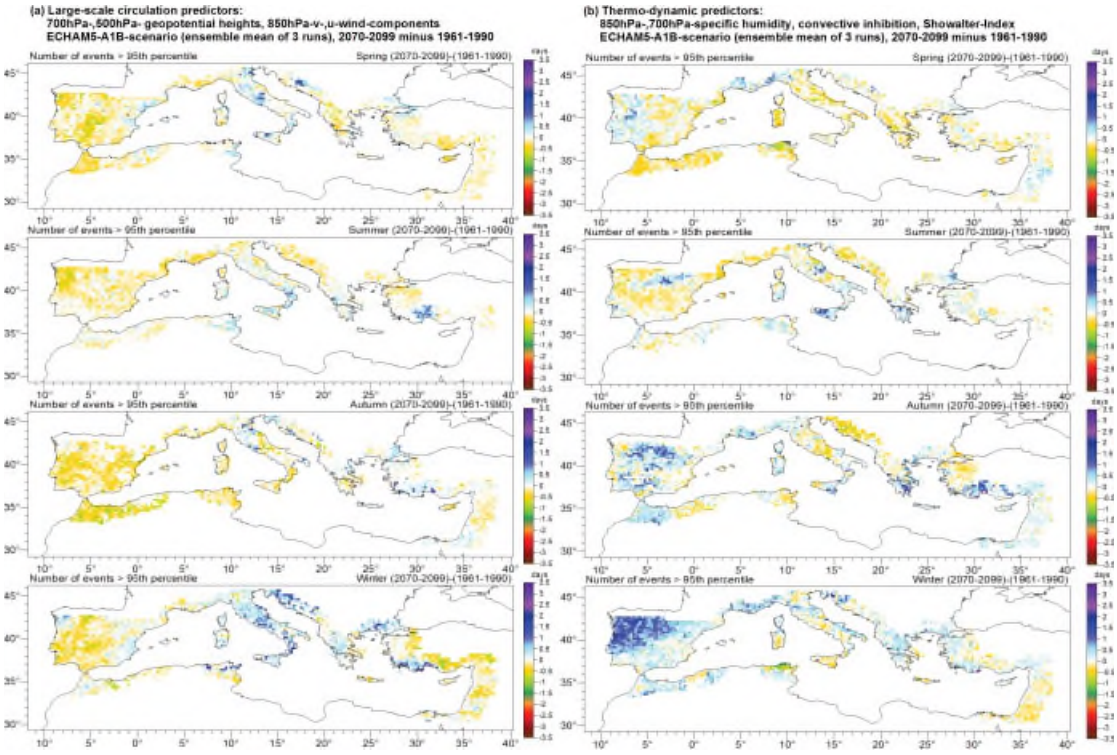

Figure 6. Changes of the number of events exceeding the 95th percentile of daily precipitation for the future time period 2070–2099 in relation to the control run period 1961–1990. Left side: Assessment taking only predictors describing the large-scale circulation (700hPa, 500 hPa geopotential heights, 850 hPa u-, v-wind components). Right side: Assessment under the sole inclusion of convective predictors (850hPa, 700 hPa specific humidity, CIN, Showalter-Index). From top to bottom: spring, summer, autumn, winter.

precipitation events for the time period 1950–2100 for the four example grid boxes which were already used in Figure 5. Each time series depicts the statistical ensemble mean using output from three ECHAM5 model runs and one HadCM3 run, respectively. For the time period 1950–2000 control run conditions, for 2001–2100 SRES A1B scenario conditions (left side) and SRES B1 scenario assumptions (right side) are applied. The time series with the black signature in the left figures of Figure 8 (ECHAM5-SRESA1B run1) corresponds to the ensemble mean illustrated in red colour in Figure 5. Overall it becomes visible that the assessments are not in agreement for individual years, but the general temporal evolution indicated by the polynomial fits is largely concurrent. Summing up, all different downscaling assessments generally produce similar results regarding the change of the number of extreme precipitation events in the Mediterranean area up until the end of the 21st century under enhanced greenhouse warming conditions. Yet one has to keep in mind the uncertainties related to the results which are reflected for example by the 95% bootstrap confidence intervals.

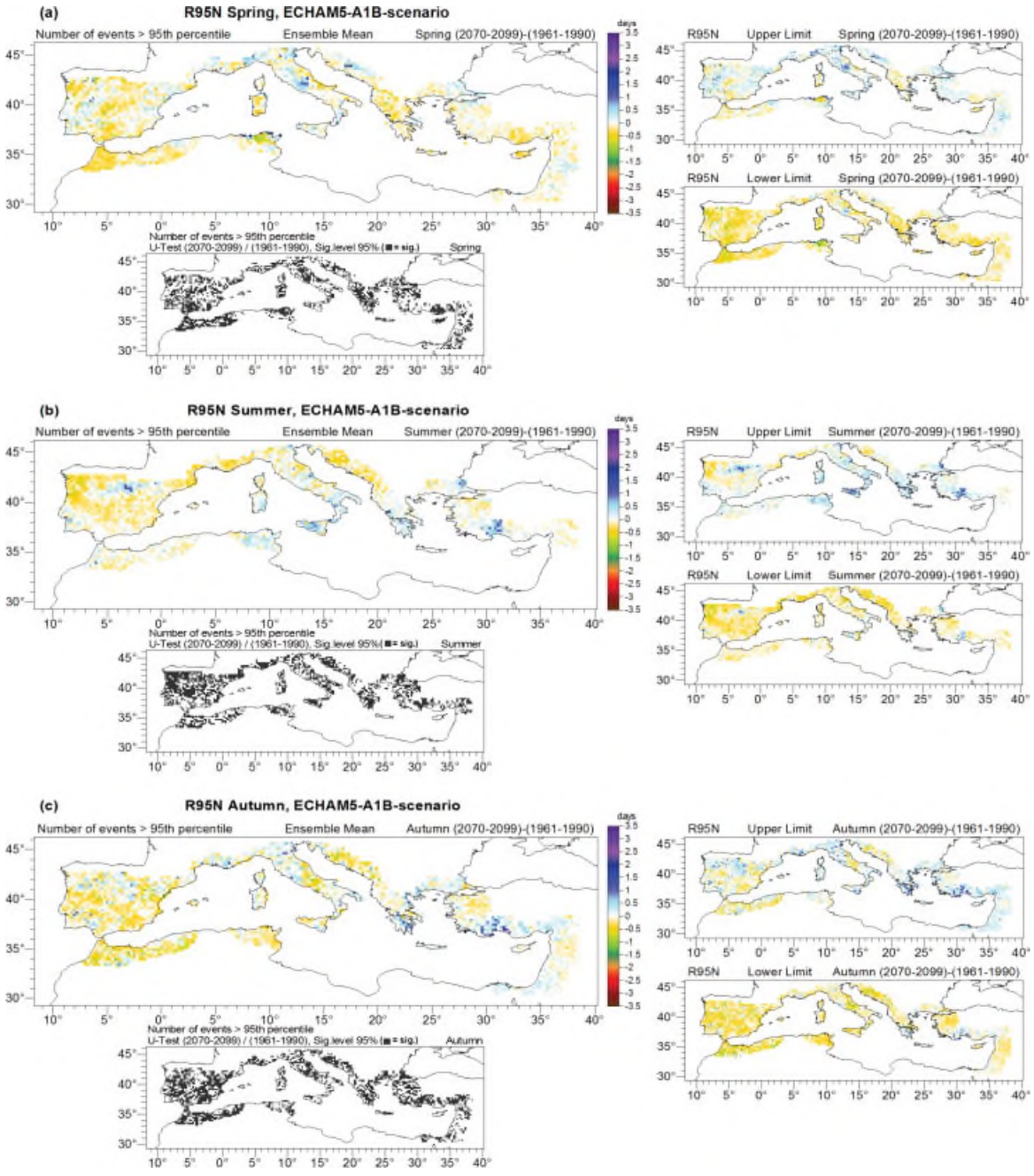

Figure 7. Ensemble mean changes of the number of events exceeding the 95th percentile of daily precipitation for the future time period 2070–2099 in relation to the control run period 1961–1990 under SRES A1B scenario assumptions. In addition the significance of the changes (95% level, U-test) and the upper and lower limit of the 95% bootstrap confidence intervals are shown. Results are based on statistical downscaling assessments using GLMs as downscaling technique and predictors from three ECHAM5/MPI-OM runs. From top to bottom: spring, summer, autumn, winter.

4.2. Total amount of precipitation from extreme precipitation events per season (R95AM)

This section deals with the assessment of R95AM, analogue to the previous Section 4.1. where the results of the statistical downscaling of R95N were discussed.

4.2.1. Model performance

Models for the total amount of precipitation from events exceeding the 95th percentile can only be established for those grid boxes and calibration/verification periods, where a model for R95N is available, since the information about the occurrence of an extreme precipitation event in the verification and model periods is taken from the assessment of R95N. Under the use of the various calibration/verification periods 6265 downscaling models can be established for 2611 grid boxes in spring. In the summer season 5065 statistical models are successfully built for 2150 grid boxes, in autumn 6373 models are available for 2605 grid boxes, and in winter 7566 models for 2618 grid boxes. The number of statistical models

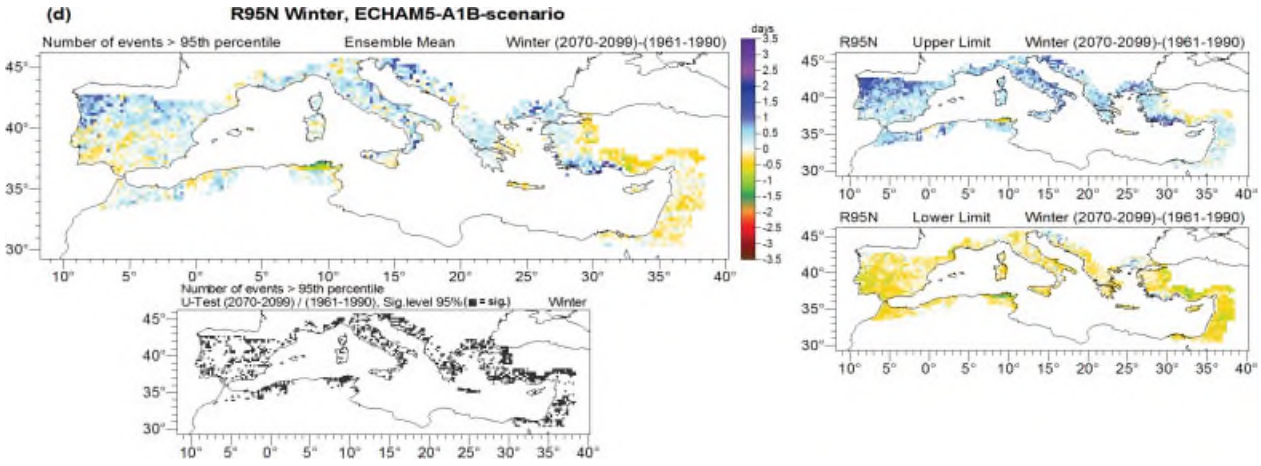

Figure 7. *Continued*.

per grid box can be seen in Figures 9–12 (panel a) for the four seasons.

Looking at the statistical ensemble mean correlation coefficients between statistically modelled and observation-based values of R95AM in the calibration and verification periods (Figures 9–12, panels (b) and (c)) and at the RV-values (Figures 9–12, panels (d) and (e)) reveals that the modelling of the seasonal amount of extreme precipitation events is quite successful when the occurrence of these extremes is perfectly known. Thus, the performance of the statistical models is quite good in the calibration, because the information on the frequency of the extremes is taken from the original time series, whereas the quality of the models drops noticeably in the verification due to the dependence of the information about zero values on the modelling power of R95N. An exploratory analysis of the model performance of R95AM in the verification periods reveals that the models remain robust in terms of correlation coefficients and RV-values when taking the information about zero values from the original time series instead of the modelled ones (not shown). Consequently, the limits of the modelling potential are primarily determined by the presence-absence problem of the extremes, whereas the distributional form of R95AM without the zero values can be described very successfully with the GLMs. In summary, the statistical model performance for the seasonal precipitation amount of the extremes features ensemble mean correlation coefficients of about 0.4 up to 0.8 and ensemble mean RV-values of up to 90% in the calibration and correlation coefficients of about 0.3 up to 0.7 and RV-values of up to about 60% in the verification.

4.2.2. Predictors

The information on how often a specific predictor is selected in the downscaling models in terms of absolute numbers and in relation to the total number of models per season is given in Table 3. In comparison to the predictor selection for R95N it becomes visible that the top three predictor variables are the same for R95AM. An exception occurs in autumn, where there is a stronger connection of the 500 hPa geopotential heights to the frequency of the extremes (R95N), whereas specific humidity of the 850 hPa level is of greater importance for the seasonal rainfall amount of the extremes (R95AM). Overall, another apparent feature is the increased relevance of the Showalter-Index and CIN for R95AM in comparison to R95N in the winter season. Besides, there is a greater modelling power of the multi-type predictor assessments for R95N compared to R95AM in all seasons. But altogether there is a high correspondence regarding the predictor selection in the statistical downscaling models for the frequency and the seasonal amount of extreme precipitation events in the four seasons.

4.2.3. Thermo-dynamic factors versus large-scale circulation

Prior to discussing the downscaling results of R95AM based on the application of all predictor variables in Section 4.2.4. the specific signals from the two different predictor sets (large-scale circulation and thermo-dynamic factors) are addressed in this subsection in more detail. As for the analysis of R95N two further types of statistical downscaling models are realized for R95AM. In one analysis only the circulation-type predictors (700 hPa/500 hPa geopotential heights, 850 hPa u-, v-wind components) are regarded, and in the other analysis only the thermo-dynamic predictors (850 hPa/700 hPa specific humidity, CIN, Showalter-Index).

The statistical model performance of the two different types of assessments is lower compared to the model performance of the assessment using all predictors. Thus the number of statistical models which can be established is reduced to 59.4% (3010 models, summer season) to 74.1% (5606 models, winter) when using only thermo-dynamic variables and to 54.1% (2741 models, summer) to 67.5% (5106 models, winter) in case of the circulation-type predictor set compared to the downscaling with all predictors (the number of models of the all-predictors assessment is specified in Section 4.2.1.). For the downscaling using thermo-dynamic predictors more statistical

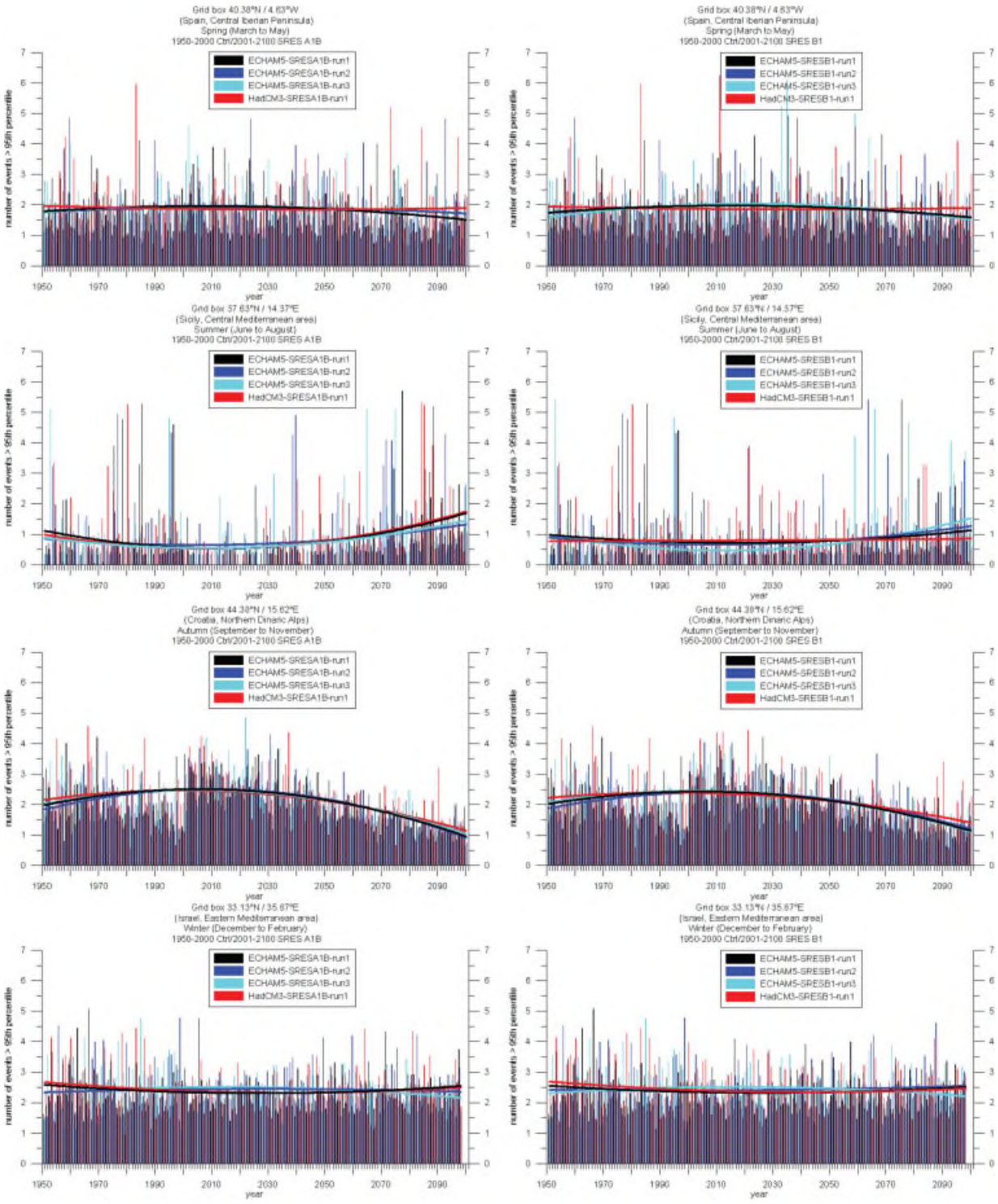

Figure 8. Temporal evolution of the number of extreme precipitation events for the time period 1950–2100 for four example grid boxes in the Mediterranean area (same grid boxes like in Figure 5). Each time series presents the statistical ensemble mean using output from three ECHAM5 model runs and one HadCM3 run, respectively. For the time period 1950–2000 control run conditions, for 2001–2100 SRES A1B scenario conditions (left side) and SRES B1 scenario assumptions (right side) are applied. Shown are the values of R95N for each individual year as well as the polynomial fit of each time series.

models can be established compared to the downscaling with circulation-type predictors. The differences of the number of statistical models which can be established is greatest in winter and summer, and least in the transitional seasons autumn and spring. But with respect to the correlation coefficients and RV-values in calibration and verification no systematic higher modelling power of one predictor set is found. Thus, the differences are related to the basic circumstance whether or not a statistical model can be successfully built, but in case of

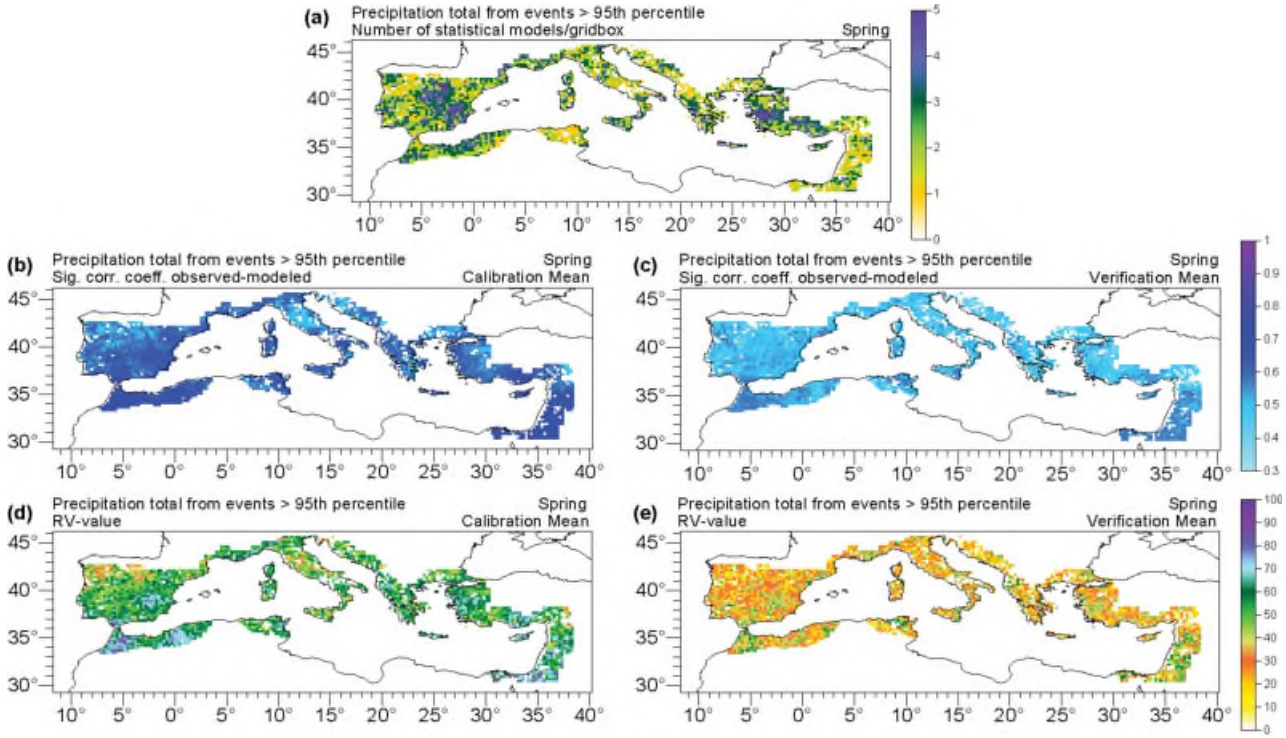

Figure 9. Statistical model performance for R95AM in spring (March to May). (a) Size of the statistical model ensemble per grid box (maximum: 5). (b) Correlation coefficients between modelled and observed extremes indices in the calibration (mean value over the statistical model ensemble). (c) Correlation coefficients between modelled and observed extremes indices in the verification (ensemble mean). (d) RV for the calibration (mean value over the statistical ensemble). (e) RV for the verification (ensemble mean).

model availability the model performance is equally high. In summary, statistical downscaling can be performed under the sole inclusion of only one predictor set at the expense of a smaller number of statistical models compared to the assessment using all predictors. Model performance in terms of correlation coefficients and RV-values is not influenced as much as it is the case in the similar study of R95N described in Section 4.1.3.. This can be attributed to the lesser importance of the multi-type predictor assessments for R95AM compared to R95N. That way model performance of R95AM is not frequently enhanced by providing different predictor types, but often one predictor variable plays a decisive role and therefore determines the performance of a specific GLM.

In spite of a lower modelling power compared to the overall assessment, R95AM can be assessed in the observational period through the sole use of one predictor set. However, substantial differences become apparent when comparing the downscaling results under future climate change conditions. Figure 13 shows the changes of R95AM for the future time period 2070–2099 in relation to the control run period 1961–1990. On the left side the assessment results taking only predictors describing the large-scale circulation (700 hPa, 500 hPa geopotential heights, 850 hPa u-, v-wind components) are shown, on the right side the projection under the sole inclusion of the convective predictors (850 hPa, 700 hPa specific humidity, CIN, Showalter-Index). In spring pronounced decreases of R95AM are visible over the Iberian Peninsula in the downscaling with circulation-type predictors (upper left figure of Figure 13), whereas a mixed spatial pattern with some decreases and some increases emerges for this area in the thermo-dynamic projection (upper right figure of Figure 13). Besides, stronger decreases in spring can be found over western North Africa in the thermo-dynamic assessment, which are not evident in the modelling result taking large-scale circulation predictors. In the summer season main differences are visible over the Iberian Peninsula. In autumn widespread differences become apparent, only the eastern coast of the Adriatic Sea and the eastern Mediterranean area stand out as regions of coherent decreases. In winter a rise of R95AM over the central-northern Mediterranean area and mainly reductions of R95AM over the western Mediterranean area results from circulation changes (lowest left figure of Figure 13), whereas thermo-dynamic forcing induces rather decreases over the central-northern region and strong increases over the western Mediterranean area (lowest right part of Figure 13). Around the north-eastern Mediterranean region both assessments yield comparable results with increases around the Aegean Sea and decreases over parts of southern Turkey.

Overall, it is shown that different predictor variables can lead to varying statistical downscaling results pointing to distinct impacts of the change of specific atmospheric conditions on local extreme precipitation.

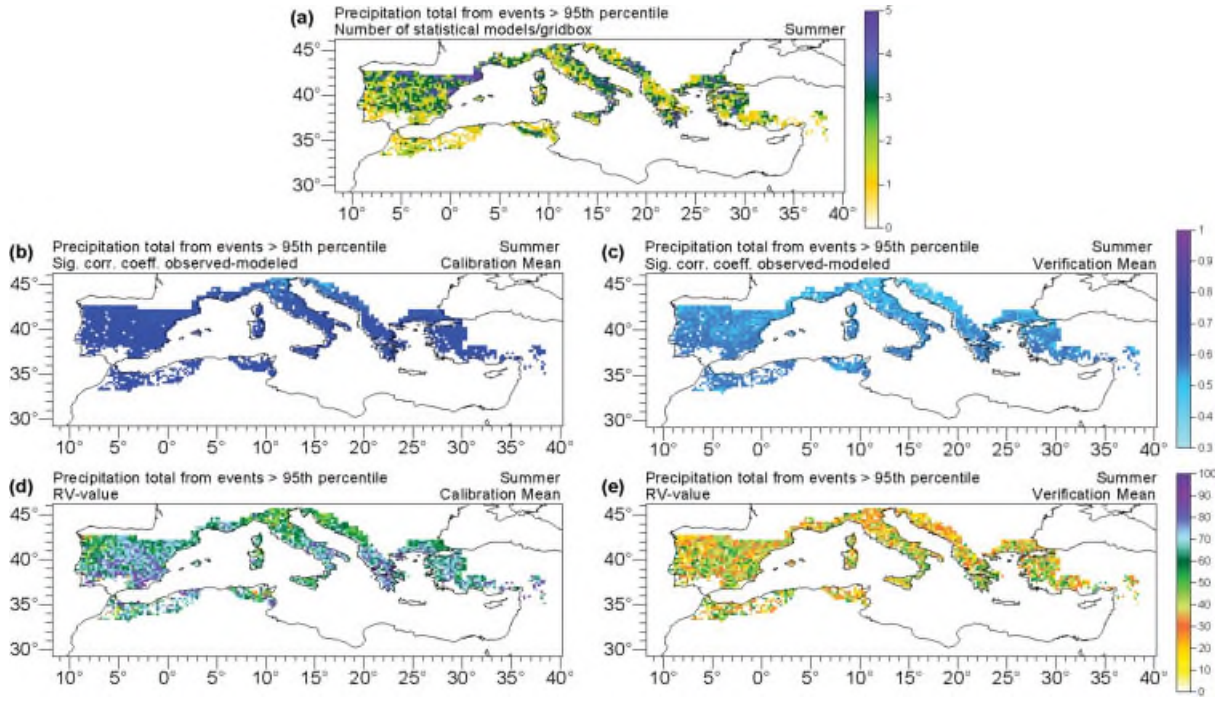

Figure 10. Statistical model performance for R95AM in summer (June to August). (a) Size of the statistical model ensemble per grid box (maximum: 5). (b) Correlation coefficients between modelled and observed extremes indices in the calibration (mean value over the statistical model ensemble). (c) Correlation coefficients between modelled and observed extremes indices in the verification (ensemble mean). (d) RV for the calibration (mean value over the statistical ensemble). (e) RV for the verification (ensemble mean).

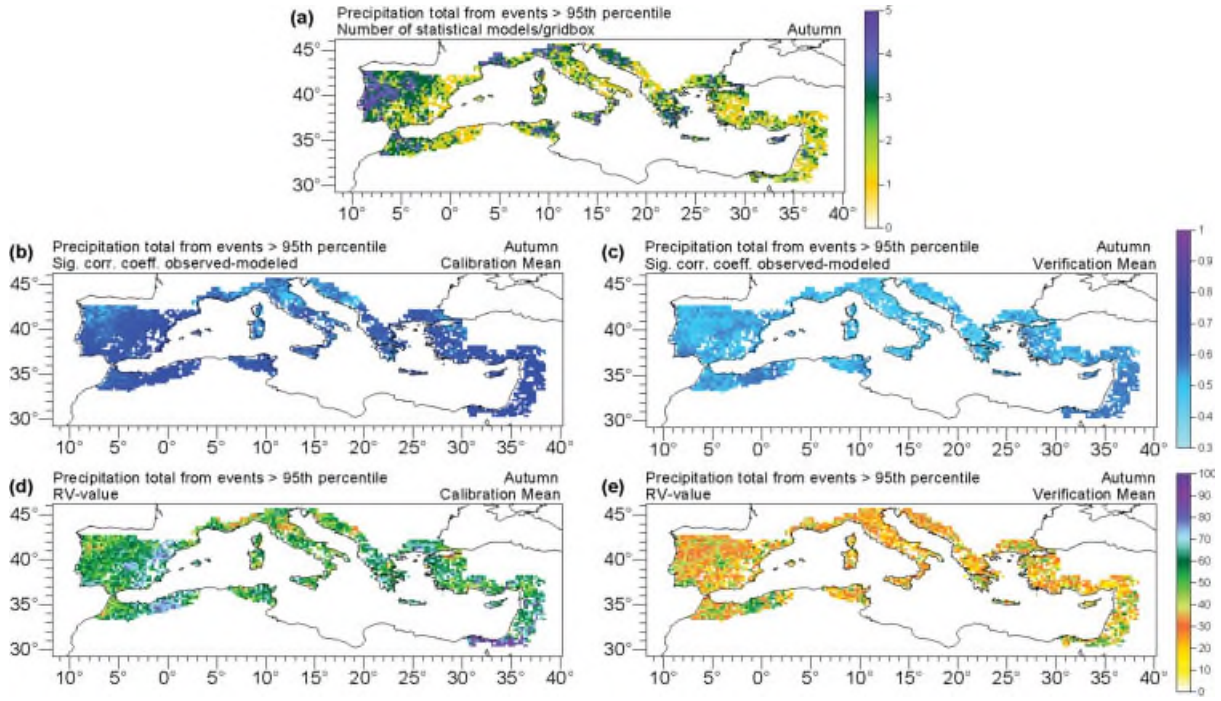

Figure 11. Statistical model performance for R95AM in autumn (September to November). (a) Size of the statistical model ensemble per grid box (maximum: 5). (b) Correlation coefficients between modelled and observed extremes indices in the calibration (mean value over the statistical model ensemble). (c) Correlation coefficients between modelled and observed extremes indices in the verification (ensemble mean). (d) RV for the calibration (mean value over the statistical ensemble). (e) RV for the verification (ensemble mean).

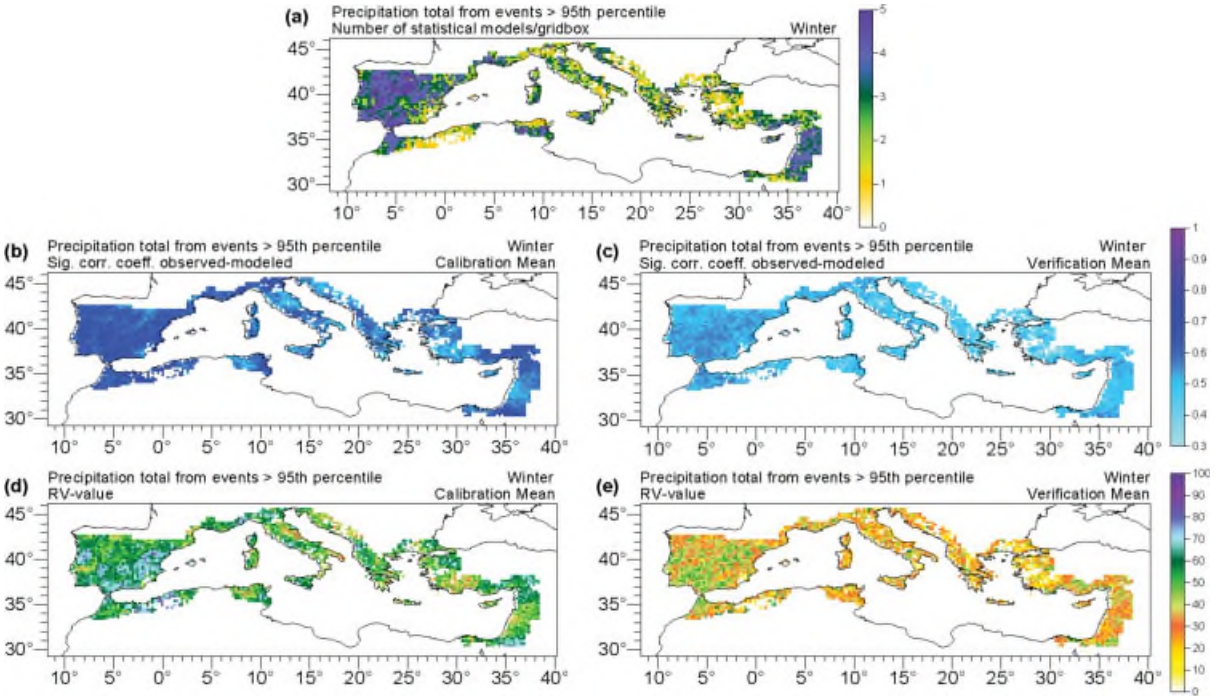

Figure 12. Statistical model performance for R95AM in winter (December to February). (a) Size of the statistical model ensemble per grid box (maximum: 5). (b) Correlation coefficients between modelled and observed extremes indices in the calibration (mean value over the statistical model ensemble). (c) Correlation coefficients between modelled and observed extremes indices in the verification (ensemble mean). (d) RV for the calibration (mean value over the statistical ensemble). (e) RV for the verification (ensemble mean).

Consequently, as for R95N, this subsection points to the importance to include different predictor variables in the statistical downscaling models to obtain a comprehensive picture of regional extreme precipitation changes. Subsequently, the analysis presented in this section can be used in the scope of the downscaling using all predictor variables which is described in the next Section 4.2.4. to give additional information on the specific sources of the assessed changes.

4.2.4. Assessment for the 21st century

Due to substantial similarities concerning the assessed general tendencies of R95AM under future climate change conditions from different scenarios (A1B and B1), multiple runs, and different GCMs (ECHAM5 and HadCM3), the statistical downscaling results discussed in this section are primarily based on the ensemble of the three ECHAM5 runs under A1B-scenario conditions. The results using B1-scenario conditions feature in general the same patterns of change. Only the absolute amounts of change are reduced to some extent compared to the A1B-scenario case. Figure 14 shows for the four seasons the downscaled projection of the precipitation total from events exceeding the 95th percentile in the Mediterranean area up until the end of the 21st century as the change in the future time slice 2070–2099 in relation to the control run period 1961–1990. Presented is the mean of the ensemble which incorporates for each grid box up to five statistical downscaling models each with the input of the three ECHAM5-A1B runs. Also shown are the significance of the changes (U-test, 95%-level) and the upper and lower limits of the 95% bootstrap confidence interval. The results are based on the statistical downscaling taking all predictor variables into account.

The pattern of change of R95AM in spring follows closely the assessed changes of R95N (see Figure 7(a)), indicating that the total amount of precipitation from events exceeding the 95th percentile is closely connected to the frequency of these events. In the summer season mainly reductions of R95AM of up to about minus 40 mm become visible over the western and northern Mediterranean area, whereas increases of up to about 30 mm are visible around the Tyrrhenian Sea, the Ionian Sea, and the Aegean Sea (Figure 14(b)). Modifications of the large-scale circulation and of the thermo-dynamic conditions contribute in an analogous manner to the assessed changes. The overall pattern of change of R95AM is in close correspondence with the change assessed for R95N, with some interesting features like the increase of R95N over parts of southern Turkey with values of more than one day (Figure 7(b)), but increases of R95AM of only about 10 mm, which points to more frequent but less intense extreme precipitation events over this area. In autumn a decrease of the seasonal extreme precipitation amount up until the end of the 21st century becomes visible for many parts of the Iberian Peninsula with strongest reductions of partly more than 70 mm over the eastern Mediterranean coast of Spain and parts of Mediterranean Algeria (Figure 14(c)). Decreases are also projected for Tunisia, parts of Italy, the eastern coast of

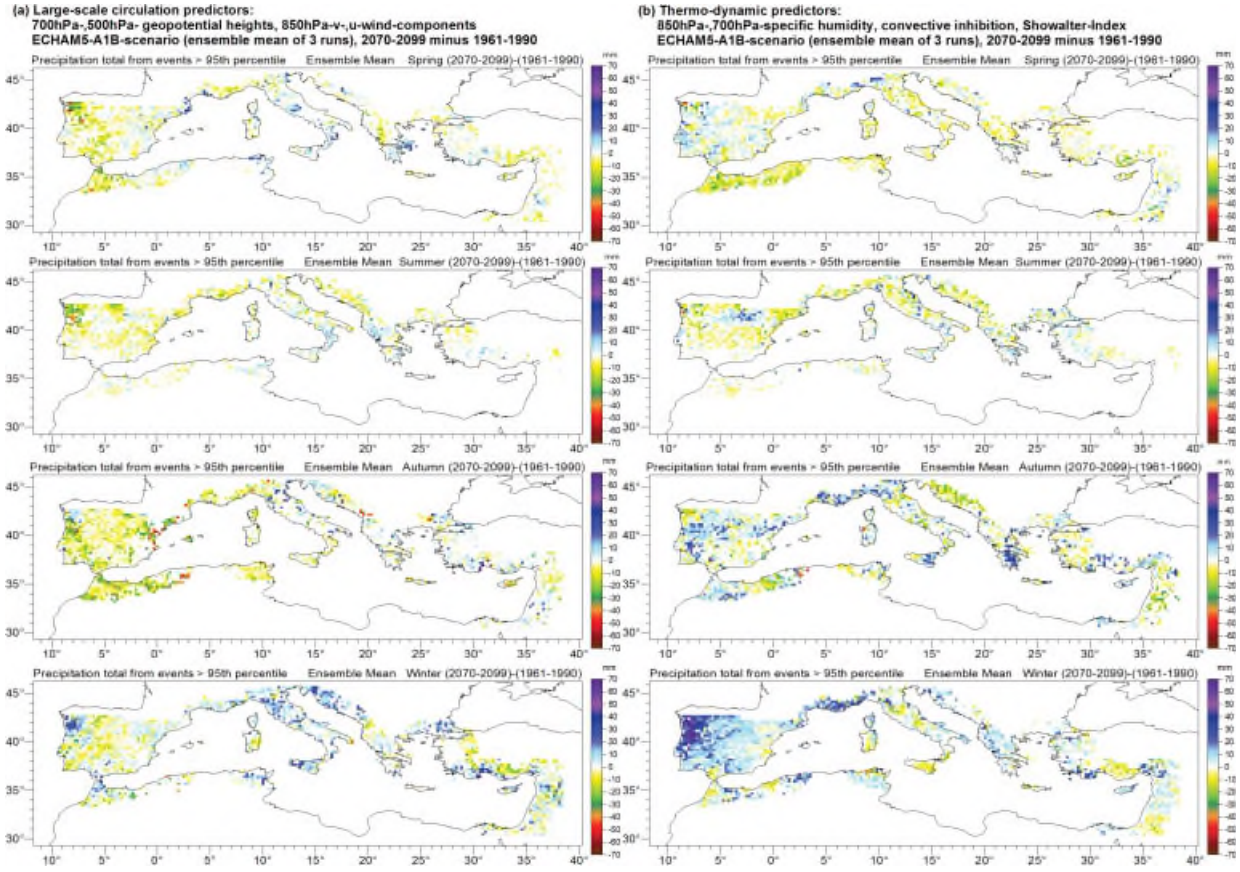

Figure 13. Changes of the total amount of precipitation from events exceeding the 95th percentile of daily precipitation for the future time period 2070–2099 in relation to the control run period 1961–1990. Left side: Assessment taking only predictors describing the large-scale circulation (700hPa, 500 hPa geopotential heights, 850 hPa u-, v-wind components). Right side: Assessment under the sole inclusion of convective predictors (850hPa, 700 hPa specific humidity, CIN, Showalter-Index). From top to bottom: spring, summer, autumn, winter.

the Adriatic Sea and some eastern Mediterranean regions. The decreases over the western Mediterranean area are mainly due to changes of the large-scale circulation, whereas the reductions over the northern and eastern Mediterranean areas are connected to thermo-dynamic modifications (see Figure 13). In contrast, increases of R95AM over the northern coast of the Ligurian Sea, Greece, and the southern coast of Turkey are primarily due to thermo-dynamic changes. In winter increases of R95AM up until the end of the 21st century dominate the pattern of change in Figure 14(d). In this context, the increases over the western Mediterranean area are mainly related to changes of thermo-dynamic conditions whereas the increases over the central-northern Mediterranean area are induced by circulation changes (see Figure 13). The increases over the north-eastern and eastern Mediterranean area as well as the decreases over parts of Turkey are connected to similar effects from both thermo-dynamic as well as circulation-type forcing. The overall pattern of change in winter is largely correspondent to the assessed changes of R95N. A comparison of the results indicates some special features, like, e.g. that in the north-eastern Mediterranean area, especially over Turkey the assessed frequency decreases of partly more than one day are stronger in relation to the overall amount decrease of up to about 20 mm, indicating that this region might experience less frequent but more intense extreme precipitation events.

5. Discussion and conclusions

For the Mediterranean area possible future changes of the frequency (R95N, represented by the number of events exceeding the 95th percentile of daily precipitation) and the total seasonal amount (R95AM, total amount of precipitation from these events) of extreme precipitation events are analysed on a seasonal basis. The results are obtained by means of statistical downscaling. Uncertainties due to future greenhouse-gas emission scenarios, as well as to the structure of climate models and their parameterizations are considered by the inclusion of different emission scenarios (A1B and B1), two different GCMs (ECHAM5 and HadCM3), and multiple runs of a specific GCM (three A1B and three B1 runs of ECHAM5).

A statistical downscaling technique based on GLMs is adopted to assess extreme precipitation changes. For R95N a Poisson process is assumed with the canonical link log. For the modelling of R95AM the zero values in the time series are disregarded and GLMs are established for the resulting continuous data. In this context the

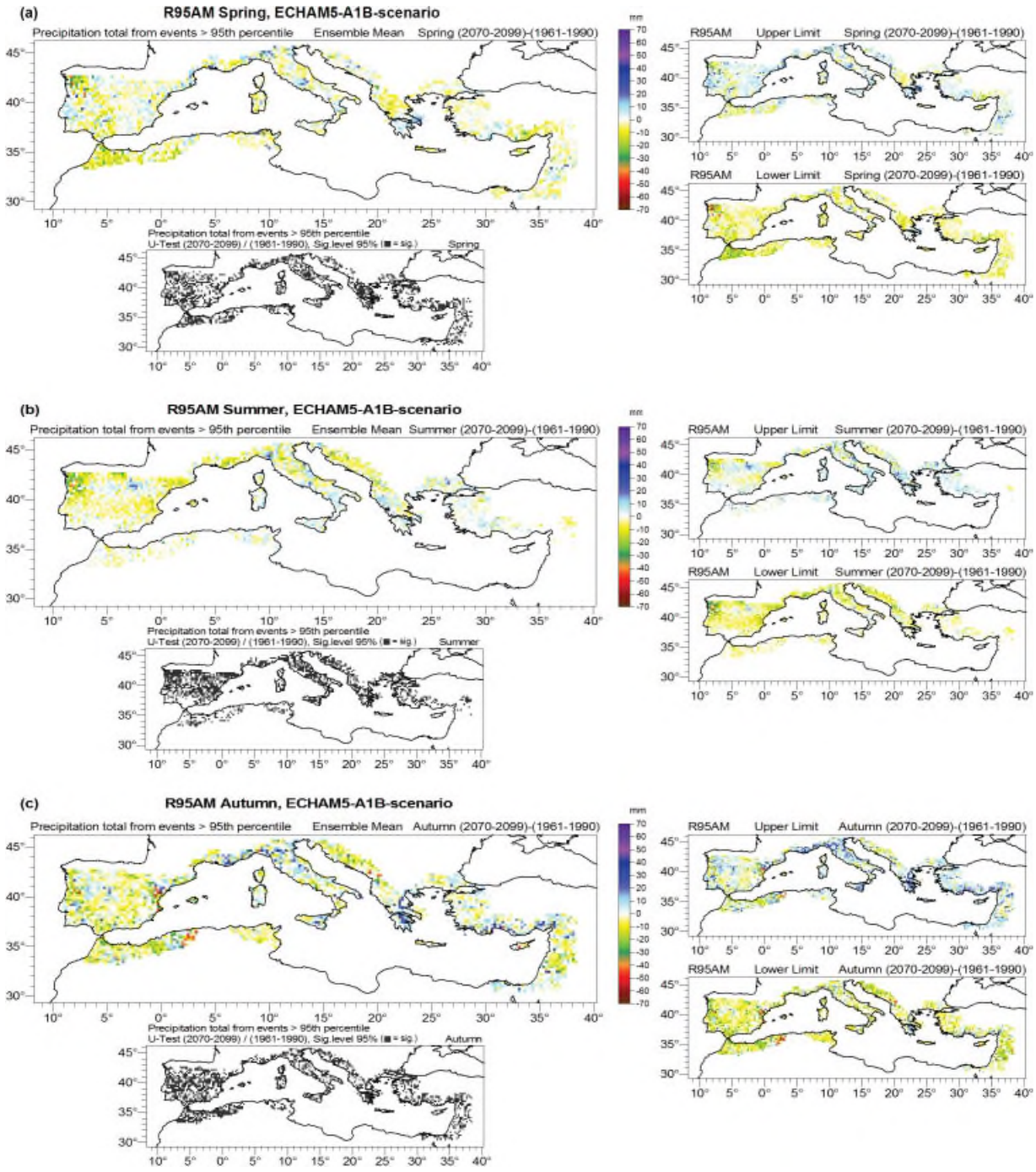

Figure 14. Ensemble mean changes of the total amount of precipitation from events exceeding the 95th percentile of daily precipitation for the future time period 2070–2099 in relation to the control run period 1961–1990 under SRES A1B scenario assumptions. In addition the significance of the changes (95% level, *U*-test) and the 95% bootstrap confidence intervals are shown. Results are based on statistical downscaling assessments using GLMs as downscaling technique and predictors from three ECHAM5/MPI-OM runs. From top to bottom: spring, summer, autumn, winter.

variance function is described by a Gamma distribution and the canonical link used is also log. The information about zero values is taken from the modelling of R95N. The particular predictand–predictor relationships are derived in different calibration and verification periods to account for non-stationarities in the relationships. The consideration of various predictor–predictand relationships helps to cover a larger range of determining factors for the generation of extreme precipitation events and different possible evolutions of the predictand are allowed for. Consequently, statistical model ensembles are available which cover a range of the observed natural variability and which give a particular quantification of uncertainties.

In the Mediterranean area extreme precipitation is related to advective processes and/or is a matter of convective precipitation. To account for this, various predictors are chosen, including more ‘classical ones’

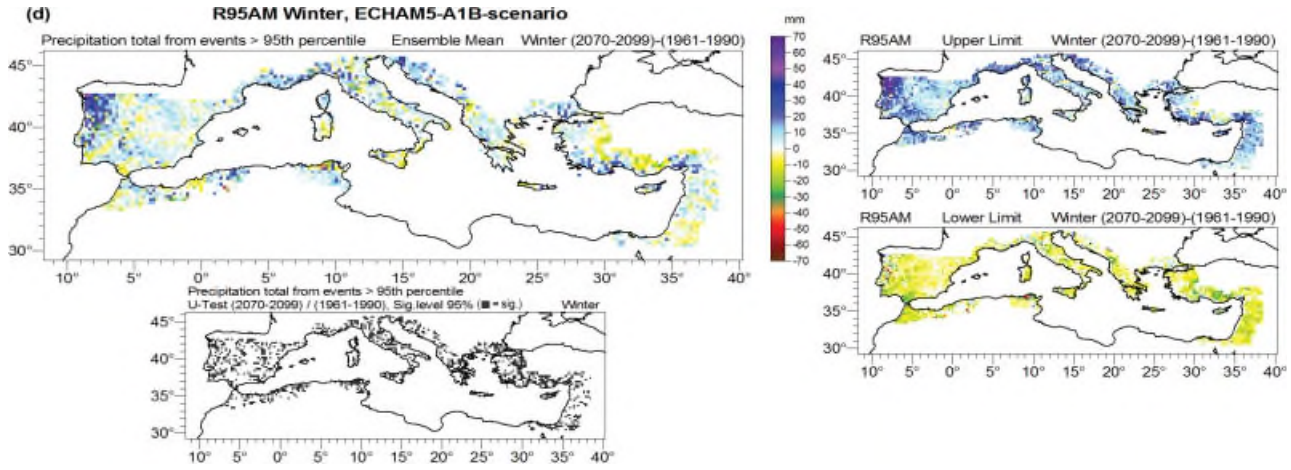

Figure 14. *Continued*.

as well as novel ones, especially for the description of convective processes. Thus, for the assessments of R95N and of R95AM, 700 hPa and 500 hPa geopotential heights and u-wind and v-wind components of the 850 hPa level are used to describe the large-scale circulation, 850 hPa and 700 hPa specific humidity, CIN, and the Showalter-Index are taken as thermo-dynamic predictors.

It is shown that the model performance for R95N and R95AM by means of GLMs as a statistical downscaling technique fluctuates depending on the region and season. The best results can generally be obtained in areas and seasons where and when total precipitation amounts are high and the inter-annual precipitation variability is comparatively low. Thus, summer represents the season and the eastern and southern Mediterranean areas the regions which are more difficult to assess.

Regarding the predictor selection within the scope of the modelling of R95N in spring and autumn, geopotential heights of the 700 hPa level as well as a multi-type predictor set are most frequently selected. Besides, the wind components are very relevant, indicating that the large-scale circulation plays an important role for the generation of extreme precipitation events in the transitional seasons. Additionally, a source of lift and moisture availability seem to be more essential in spring, whereas convective instability plays a more important role in autumn. The most frequent selection of 850 hPa specific humidity in summer points to the basic requirement of moisture existence for the generation of precipitation extremes in this season. The inclusion of 700 hPa and 500 hPa geopotential heights and the zonal wind component in the GLMs indicates that large-scale circulation anomalies are also necessary to induce extreme precipitation events in summer. In the winter season R95N is governed by the moisture availability in the lower troposphere and by the large-scale circulation, represented by the geopotential heights of the 700 hPa level and the zonal wind component of the 850 hPa level. Thus, even in high winter, when the Mediterranean area is under the increased influence of the cyclonic west wind drift, extreme precipitation events in the Mediterranean area are not solely determined by the large-scale circulation, but the moisture content of the atmosphere has also to be taken into account. Concerning the predictor selection for the statistical assessment of R95AM, it appears that the top three predictor variables are the same as those for R95N. An exception is during autumn when 500 hPa geopotential heights are more relevant for R95N, and 850 hPa specific humidity is of greater importance for R95AM. Beyond that, in the winter season an apparent feature is the increased relevance of the Showalter-Index and CIN for R95AM in comparison to R95N. Besides, there is a greater modelling power of the multi-type predictor assessments for R95N compared to R95AM in all seasons. Yet in summary there is a high correspondence regarding the predictor selection in the statistical downscaling models for R95N and for R95AM.

In addition to the statistical downscaling based on the application of all predictor variables, the specific signals from two different predictor sets (large-scale circulation represented by 700 hPa/500 hPa geopotential heights, 850 hPa u-, v-wind components and thermo-dynamic factors described by 850 hPa/700 hPa specific humidity, CIN, Showalter-Index) are addressed in this study. From this analysis considerable differences appear under conditions of future climate change. For example, in winter the increases of the frequency and seasonal amount of extreme precipitation events over the western Mediterranean area are primarily related to changes of thermo-dynamic conditions whereas mainly decreases arise from circulation changes alone. In contrast, the increases over the central-northern Mediterranean area in winter are more strongly connected to circulation changes, while the change of the thermo-dynamic factors rather results in decreases over this area. Overall it is shown that different predictor variables can lead to varying statistical downscaling results pointing to distinct impacts of the change of specific atmospheric conditions on local extreme precipitation. It points to the importance to include different predictor variables in the

Table 3. Frequency of the predictors selected as independent variables within the statistical downscaling models for the precipitation total from events exceeding the 95th percentile of daily precipitation (R95AM). The total number of models consists of the number of grid boxes times the number of calibration/verification periods where a downscaling model can be established. A multi-type predictor assessment in terms of a specific combination of individual predictor variables is used where it yields better results than an assessment with a single predictor. From top to bottom: results for spring, summer, autumn, and winter.

R95AM/Spring Predictor	Number of models 6265	percentage 100%
700 hPa geopotential heights	897	14.32
700 hPa specific humidity	762	12.16
850 hPa meridional wind	727	11.60
850 hPa specific humidity	720	11.49
850 hPa zonal wind	699	11.16
500 hPa geopotential heights	633	10.10
Showalter-Index	632	10.09
Multi-type predictor combination	612	9.77
Convective inhibition	583	9.31

R95AM/Summer Predictor	Number of models 5065	percentage 100%
700 hPa geopotential heights	658	12.99
850 hPa specific humidity	657	12.98
500 hPa geopotential heights	635	12.54
850 hPa zonal wind	572	11.29
Multi-type predictor combination	568	11.21
Convective inhibition	538	10.62
700 hPa specific humidity	496	9.79
Showalter-Index	476	9.40
850 hPa meridional wind	465	9.18

R95AM/Autumn Predictor	Number of models 6373	Percentage 100%
850 hPa meridional wind	837	13.13
700 hPa geopotential heights	812	12.74
850 hPa specific humidity	768	12.05
850 hPa zonal wind	752	11.80
700 hPa specific humidity	712	11.17
Showalter-Index	649	10.19
Multi-type predictor combination	625	9.81
500 hPa geopotential heights	612	9.60
Convective inhibition	606	9.51

R95AM/Winter Predictor	Number of models 7566	Percentage 100%
850 hPa specific humidity	1146	15.16
700 hPa geopotential heights	1049	13.86
850 hPa zonal wind	906	11.97
Convective inhibition	874	11.55
Showalter-Index	788	10.42
700 hPa specific humidity	775	10.24
850 hPa meridional wind	747	9.87
500 hPa geopotential heights	725	9.58
Multi-type predictor combination	556	7.35

statistical downscaling models to obtain a comprehensive picture of regional extreme precipitation changes.

The change of R95N in the Mediterranean area in the future time slice 2070–2099 in relation to the control run period 1961–1990 from the assessment using all predictor variables shows in spring mainly decreases over many parts of the Mediterranean area. In the summer season increases of up to about two extreme precipitation days up until the end of the 21st century are visible around the Tyrrhenian Sea, the Ionian Sea, and the Aegean Sea, whereas decreases are assessed for most of the western and northern Mediterranean regions. In autumn reductions of R95N occur over many parts of the western and central Mediterranean area. In the eastern Mediterranean area a widespread rise of the frequency is assessed, in particular over southern Turkey and southern Greece with values of up to about three days. In winter distinct increases of R95N can be seen for many parts of the Mediterranean area, most pronounced over the north-western Iberian Peninsula, the northern Dinaric Alps, around the northern Aegean Sea, and the southern coast of Turkey with a maximum of about four days. Also for northern Morocco, central and southern Italy increases are assessed. In contrast for the Mediterranean coast of Tunisia, the eastern Mediterranean area, and parts of Turkey decreases of the number of precipitation extremes are modelled. Overall, the described changes are significant at many grid boxes, but the limits of the 95% bootstrap confidence intervals indicate that for most regions the change can range from possible decreases up to potential increases.

The overall pattern of change of R95AM in the Mediterranean area is largely correspondent to the assessed changes of R95N. Thus, up until the end of the 21st century widespread decreases of R95AM are visible for the spring season. The decreases are strongest over the north-western Iberian Peninsula, north-western Africa, parts of Greece and Albania as well as over south-eastern Turkey with values of up to about minus 50 mm. Noticeable increases of R95AM in spring can only be found for parts of northern and western Italy, southern France and southern Greece. In the summer season mainly reductions of R95AM occur over the western and northern Mediterranean area, whereas increases up until the end of the 21st century emerge around the Tyrrhenian Sea, the Ionian Sea, and the Aegean Sea. In autumn a decrease in the amount of seasonal extreme precipitation becomes visible for many parts of the Iberian Peninsula with strongest reductions of partly more than 70 mm over the eastern Mediterranean coast of Spain and parts of Mediterranean Algeria. Decreases are also projected for Tunisia, parts of Italy, the eastern coast of the Adriatic Sea and some eastern Mediterranean regions. In contrast, increases of R95AM can be found over the northern coast of the Ligurian Sea, Greece, and the southern coast of Turkey. In winter mainly increases of R95AM are noted up until the end of the 21st century.

The described changes refer to the statistical downscaling of ECHAM5-A1B model output. However, the

assessment using HadCM3-A1B data yields a very similar outcome. Also the statistical downscaling under the use of the B1-scenario is consistent to the results obtained with the A1B-scenario. Using the B1-scenario results in a slightly smaller change, leaving the overall pattern of change unmodified.

Overall, the statistical assessment of extreme precipitation in the Mediterranean area up until the end of the 21st century yields diverse seasonal and regional patterns of change. In general there are consistencies to the projected large-scale pattern of mean precipitation change as it is, e.g. statistically assessed by Hertig and Jacobeit (2008) and dynamically by Giorgi and Lionello (2008). However, considerable variations occur which are primarily related to the projected increases of extreme precipitation over some regions in the transitional seasons, which are not evident for mean precipitation. A further pursuit of the question about the relation between mean and extreme precipitation change in the Mediterranean area will be addressed in future work.

Acknowledgements

Financial support is provided by the DFG (German Research Foundation). We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu-metoffice.com) and the data providers in the ECA&D project (http://eca.knmi.nl). The authors thank two anonymous reviewers for providing useful comments.

References

Benestad RE. 2010. Downscaling precipitation extremes. Correction of analog models through PDF predictions. *Theoretical and Applied Climatology* **100**: 1–21.

Benestad RE. 2007. Novel methods for inferring future changes in extreme rainfall over Northern Europe. *Climate Research* **34**: 195–210.

Benestad RE, Hanssen-Bauer I, Førland EJ. 2007. An evaluation of statistical models for downscaling precipitation and their ability to capture long-term trends. *International Journal of Climatology* **27**: 649–665.

Beniston M, Stephenson DB, Christensen OB, Ferro C, Frei C, Goyette S, Halsnaes K, Holt T, Jylhä K, Koffi B, Palutikof J, Schöll R, Semmler T, Woth K. 2007. Future extreme events in European climate: an exploration of regional climate model projections. *Climatic Change* **81**: 71–95.

Busuioc A, Tomozeiu R, Cacciamani C. 2008. Statistical downscaling model based on canonical correlation analysis for winter extreme precipitation events in the Emilia-Romagna region. *International Journal of Climatology* **28**: 449–464.

Cavazos T, Hewitson BC. 2005. Performance of NCEP-NCAR reanalysis variables in statistical downscaling of daily precipitation. *Climate Research* **28**: 95–107.

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

DiCiccio TJ, Efron B. 1996. Bootstrap confidence intervals. *Statistical Science* **11**: 189–228.

Doswell CA III. 1987. The distinction between large-scale and mesoscale contribution to severe convection: A case study example. *Weather and Forecasting* **2**: 3–16.

Emori S, Brown SJ. 2005. Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate. *Geophysical Research Letters* **32**: L17706.

Flaounas E, Drobinski P, Vrac M, Bastin S, Lebeaupin-Brossier C, Stefanon M, Borga M, Calvet JC. 2013. Precipitation and temperature space–time variability and extremes in the Mediterranean region: evaluation of dynamical and statistical downscaling methods. *Climate Dynamics* **40**: 2687–2705. DOI: 10.1007/s00382-012-1558-y

Fowler HJ, Ekström M, Blenkinsop S, Smith AP. 2007. Estimating change in extreme European precipitation using a multimodel ensemble. *Journal of Geophysical Research: Atmospheres* **112**: D18104.

Gao X, Pal JS, Giorgi F. 2006. Projected changes in mean and extreme precipitation over the Mediterranean region from high resolution double nested RCM simulations. *Geophysical Research Letters* **33**: L03706.

Giorgi F, Lionello P. 2008. Climate change projections for the Mediterranean area. *Global and Planetary Change* **63**: 90–104.

Gordon C, Cooper C, Senior CA, Banks HT, Gregory JM, Johns TC, Mitchell JFB, Wood RA. 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics* **16**: 147–168.

Goubanova K, Li L. 2007. Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations. *Global and Planetary Change* **57**: 27–42.

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded dataset of surface temperature and precipitation. *Journal of Geophysical Research: Atmospheres* **113**: D20119. DOI: 10.1029/2008JD10201

Haylock MR, Goodess CM. 2004. Interannual variability of European winter rainfall and links with mean large-scale circulation. *International Journal of Climatology* **24**: 759–776.

Hertig E, Paxian A, Vogt G, Seubert S, Paeth H, Jacobeit J. 2011. Statistical and dynamical downscaling assessments of precipitation extremes in the Mediterranean area. *Meteorologische Zeitschrift* **21**: 61–77.

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

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. *Bulletin of the American Meteorological Society* **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. *Bulletin of the American Meteorological Society* **82**: 247–268.

Kundzewicz ZW, Radziejewski M, Pinskwar I. 2006. Precipitation extremes in the changing climate of Europe. *Climate Research* **31**: 51–58.

Kyselý J, Beguerías S, Beranová R, Gaála L, López-Moreno JI. 2012. Different patterns of climate change scenarios for short-term and multi-day precipitation extremes in the Mediterranean. *Global and Planetary Change* **98–99**: 63–72.

Mann HB, Whitney DR. 1947. On a test whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics* **18**: 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. *Reviews of Geophysics* **48**(3): RG3003.

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

McCulloch CE, Searle SR. 2001. *Generalized, Linear, and Mixed Models*. Wiley Series in Probability and Statistics. Wiley: New York.

Meehl GA, Covey C, Taylor KE, Delworth T, Stouffer RJ. 2007a. The WCRP CMIP3 Multimodel Dataset: a new era in climate change research. *Bulletin of the American Meteorological Society* **88**: 1383–1394. DOI: 10.1175/BAMS-88-9-1383

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 A, 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 analysed for the period 1901–2000. *Journal of Geophysical Research-Atmospheres* **111**: D22106. DOI: 10.1029/2006JD007103

Myoung B, Nielsen-Gammon JW. 2010. Sensitivity of monthly convective precipitation to environmental conditions. *Journal of Climate* **23**: 166–188.

Nakicenovic N, Swart R (eds). 2000. *Emissions Scenarios 2000. Special Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press: Cambridge.

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. *Journal of Climate* **20**(16): 4065–4095.

Pope V, Gallani ML, Rowntree PR, Stratton RA. 2000. The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics* **16**: 123–146.

Preisendorfer RW. 1988. *Principal Component Analysis in Meteorology and Oceanography*. Developments in Atmospheric Sciences, Vol. **17**, Mobly CD (ed). Elsevier: Amsterdam.

Räisänen J, Hansson U, Ullerstig A, Döscher R, Graham LP, Jones C, Meier M, Samuelsson P, Willén U. 2004. European climate in the late 21st century: regional simulations with two driving global models and two forcing scenarios. *Climate Dynamics* **22**(1): 13–31.

Rajczak J, Pall P, Schär C. 2013. Projections of extreme precipitation events in regional climate simulations for Europe and the Alpine Region. *Journal of Geophysical Research* **118**: 3610–3626. DOI: 10.1002/jgrd.50297

Roeckner E, Baeuml G, Bonaventura L, Brokopf R, Esch M, Giorgetta M, Hagemann S, Kirchner I, Kornblueh L, Manzani E, Rhodin A, Schlese U, Schulzweida U, Tompkins A. 2003. The atmospheric general circulation model ECHAM5. Part I: Model description. Report No. 349, Max Planck Institute for Meteorology. Hamburg.

Roeckner E, Brasseur GP, Giorgetta M, Jacob D, Jungclaus J, Reick C, Sillmann J. 2006. Climate projections for the 21st century. Max Planck Institute for Meteorology. Hamburg.

Schmidli J, Frei C, Vidale PL. 2006. Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology* **26**: 679–689.

Showalter AK. 1953. A convective index as an indicator of cumulonimbus development. *Journal of Applied Meteorology* **5**: 839–846.

Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA. 2006. Going to the extremes. An intercomparison of model-simulated historical and future changes in extreme events. *Climatic Change* **79**: 185–211.

Tolika K, Anagnostopoulou C, Maheras P, Vafiadis M. 2008. Simulation of future changes in extreme rainfall and temperature conditions over the Greek area: a comparison of two statistical downscaling approaches. *Global and Planetary Change* **63**: 132–151.

Toreti A, Xoplaki E, Maraun D, Kuglitsch FG, Wanner H, Luterbacher J. 2010. Characterisation of extreme winter precipitation in Mediterranean coastal sites and associated anomalous atmospheric circulation patterns. *Natural Hazards and Earth System Sciences* **10**: 1037–1050.

Tramblay Y, Neppel L, Carreau J, Sanchez-Gomez E. 2011. Extreme value modelling of daily areal rainfall over Mediterranean catchments in a changing climate. *Hydrological Processes* **26**: 3934–3944. DOI: 10.1002/hyp.8417

Van Ulden A, van Oldenborgh G. 2006. Large-scale atmospheric circulation biases and changes in global climate model simulations and their importance for climate change in Central Europe. *Atmospheric Chemistry and Physics* **6**: 863–881.

Vrac M, Yiou P. 2010. Weather regimes designed for local precipitation modelling: Application to the Mediterranean basin. *Journal of Geophysical Research* **115**: D12103.

Vrac M, Naveau P. 2007. Stochastic downscaling of precipitation: From dry events to heavy rainfalls. *Water Resources Research* **43**: W07402.

Wilby RL, Hay L, Leavesley G. 1999. A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado. *Journal of Hydrology* **225**(1–2): 67–91.

Willett KM, Jones PD, Thorne PW, Gillett NP. 2010. A comparison of large scale changes in surface humidity over land in observations and CMIP3 general circulation models. *Environmental Research Letters* **5**: 025210. DOI: 10.1088/1748-9326/5/2/025210

Yang C, Chandler RE, Isham VS, Wheater HS. 2005. Spatial-temporal rainfall simulation using generalized linear models. *Water Resources Research* **41**: W11415.