In this response we include the comments of the reviewers (in bold), reply from the open review ("Reply") and our final response ("Change in manuscript") describing what was done. The line numbers of changes are related to the marked-up version of the manuscript, which is attached to this response.

## Comments by prof. Pegram

**1. Comment: "*Do you mean the medians of r(5,11) in the 2 periods when you give the values as precisely 0.90 and 0.73?*"**

**Reply:** No, the values 0.90 and 0.73 are the sample correlations $r_{5,11}$ from the control and future period, respectively, calculated from the original data with outliers. We suggest changing the sentence as follows: "In the simulation of the model 2A, *the sample correlation $r_{5,11}$* decreased from…"

**Change in manuscript:** The sentence was changed as proposed (line 167).

**2. Comment: "*Please mark the item with an arrow, as I have done - it was very difficult to find in the figure. Please explain what you have done in more detail in Figure 4's caption, noting my comment on Figure 2.*"**

**Reply:** We will add the arrow to Figure 4 and we will complete the Figure 4's caption accordingly. Together with this we suggest changing Figure 2 so that it depicts the numbering of *all* pairs of grid-boxes:



Figure 2. The numbering of individual pairs of grid-boxes. The figure depicts the correlation matrix, the orders of rows/columns correspond to the grid-box labels from Fig. 1. The sub-diagonal part of the (symmetrical) matrix was used for numbering of individual pairs of grid-boxes – the numbers inside of the matrix represent the identifiers used in Fig. 4.

**Change in manuscript:** The arrow was added to the Figure 4 and the caption was changed (line 154, see also the next comment). Figure 2 and it´s caption were changed as we proposed (line 105).

**3. Comment: "*To my understanding, row 5 of the correlation matrix in Figure 2 (extrapolated) goes from 1 to 4, not 1 to 11. If you are referring to the item which I've marked with an arrow in Figure 4, then in the caption you should note that the divisions inserted in the figure, as I previously requested, are for columns 2 to 12, not rows as in the caption! I wasted a lot of time trying to sort it out.*"**

**Reply:** The identifiers placed between two neighbouring blue lines in Fig. 4 are located in one row of the correlation matrix, or more precisely they are located in the sub-diagonal part of this row. To state this clearly, we suggest changing the caption of Figure 4 as follows:

*"Figure 4. The 95% confidence intervals of the cross-correlation for overlapping wet periods for all models. The identifiers of grid-box pairs (ID) are explained in Fig. 2. The blue lines separate identifiers located in successive rows in the correlation matrix (see Fig. 2). The arrow marks the confidence intervals around the $r_{5, 11}$ (ID 50) of the model 2A, discussed in more detail in section 4.2."*

We hope that this caption together with revised Figure 2 will prevent misunderstanding.

**Change in manuscript:** The caption of Figure 4 and the complete Figure 2 were changed as proposed (lines 154 and 105, respectively).

**4. Line 87, comment: "*I do not understand this sentence; please reword it*"**

**Reply:** The sentence "*The joints of adjacent blocks were not included in the calculations*" is only a supplementary comment related to the calculation of serial correlation in the block bootstrap. Due to random selection of the blocks the beginning part of blocks is independent on the end of the previous blocks. To minimize bias introduced by block resampling, data that are potentially influenced were not considered for the calculation of the serial correlation. The sentence is unnecessary, we suggest it's removing.

**Change in manuscript:** The problem of bootstrap-derived confidence intervals of serial correlation was mentioned by prof. Sharma in his review. Therefore the sentence was extended to describe the issue more clearly (line 93):

*"Due to random selection of the blocks the beginning part of blocks is independent on the end of the previous blocks. To minimize bias introduced by block resampling, data that are potentially influenced (joints of the adjacent blocks) were not considered for the calculation of the serial correlation."*

**5. Line 99, comment: "*of the medians?*"**

**Reply:** No, the interval $\langle -0.02, 0.03 \rangle$ represents the overall extent of changes of binary cross-correlations; see Fig. 3(a).

**6. Line 108, comment: "*Surely these values are 0.08 and 0.023?? see my estimates in Figures 3(b) and 3(c)*"**

**Reply:** The values 0.8 and 0.23 are the *averages* of cross-correlation and lag-1 auto-correlation, respectively, calculated across all models. Figures 3(b) and 3(c) depict the *changes* of cross-correlation and lag-1 auto-correlation, respectively. The information about averages is related to the

previous sentence and explains why the *relative changes* of cross-correlations are higher than in the case of auto-correlations. We suggest rewording the sentences as follows:

"The maximal relative changes in cross-correlation reach up to 18% of the value from the control period, in the case of auto-correlation it is almost 45%. This is because the auto-correlations are in general markedly lower than cross-correlations (the mean cross-correlation of individual models exceeds 0.8; the mean lag-1 auto-correlation is around 0.23)."

**Change in manuscript:** The sentence was reworded as we suggested (line 126).


**7. Lines 116 – 118:**

**Reply:** We agree, the formulations *"…except models 1A and 2A"* and *"…but the overall trend is a drop in the future"* complement the description of Figures 4 and 5 properly, we will add them to the text, thank you.

**Change in manuscript:** The formulations were added to the text (lines 140 and 142).


**8. Lines 132 – 133, comment: „*there is no mention of Fig. 5 in the paper*"**

**Reply:** Figure 5 is described in lines 116-118.


**9. Line 156, comment:** *"Here are 2 more cases where the authors have not heeded my previous corrections, which I've had to repeat; there are others which is an irritation for the reviewer"*

**Reply:** We apologize for this omission.


Finally we note that all grammatical corrections suggested by prof. Pegram have been included into the text.


## Comments by prof. Sharma

**1. Comment: "*My request to the authors is to use the multivariate bias correction software now publically available and described in [Mehrotra, R., F. Johnson, and A. Sharma (2018), A software toolkit for correcting systematic biases in climate model simulations, Environmental Modelling and Software, 104, 130-152, doi:10.1016/j.envsoft.2018.02.010.] to show the impact these robust correlation metrics have on results.*"**

**Reply:** The comprehensive reply to this comment (comprising of 5 pages) constitutes the main point of our reply to RC2, we do not repeat it here.

**Change in manuscript:** The reply was partly reworked and attached as a supplementary material to the manuscript. We also refer to these results in the manuscript (line 254).


**2. Comment: "*l79 - Using the block approach will alter the lag-one correlation at the end of year boundaries. I presume the impact will not be much but should be stated by the authors. On the same point, I would expect the cross dependence to remain unchanged, and the lag 1 correlation to only slightly be changed.*"**

**Reply:** We are aware of this problem; the calculations were designed such that the lag-1 autocorrelations were not affected in this manner. This point is briefly mentioned in the methodological part of the submitted paper (line 87), where the following sentence can be found "*The joints of the adjacent blocks were not included in the calculation*". Nevertheless, prof. Pegram in his review mentioned that this sentence is not intelligible. As a reply to his comment we suggested it´s removing, but we can reformulate the sentence, for example as follows:

"*Due to random selection of the blocks the beginning part of blocks is independent on the end of the previous blocks. To minimize bias introduced by block resampling, data that are potentially influenced (joints of the adjacent blocks) were not considered for the calculation of the serial correlation*".

**Change in manuscript:** The sentence was reformulated as we suggested (line 93).

**3. Comment: "*And I am unable to figure out how these confidence intervals are finally used? Were all the correlations from the raw data and the resampled opnes pooled in deriving the results in Figs 3 and 4? Usually one does bootstrap tests to assess the significance of correlation from zero - here it seems the idea is to assess the significance of correlation from what it would be if the year to year dependence is made null. Some clarification is needed.*"**

**Reply:** The procedure for estimation of confidence intervals is standard [see e.g. Davison and Hinkley 1997] and relies on sampling with replacement of annual blocks of data. The resampling of blocks is done to preserve seasonal variation of rainfall. It is true that year-to-year dependence of rainfall is ignored, but we do not expect that this would significantly affect the estimates of confidence intervals of dependence indicators. In the original manuscript we did not perform real test on the significance of changes in auto- and cross dependence, instead we only visually assessed the overlap of estimated confidence intervals for correlations and autocorrelations in figs. 4 a 5 (boxplots in fig. 3 represent distribution of indicators between grid boxes). If the confidence intervals overlap, then the changes are not significant. Standard test would be easily performed e.g. by subtracting the estimated correlations for control period from those for the future period. If the confidence interval of this difference contains zero, then the change is not statistically significant. We will modify the description if we are invited to revise the manuscript. We can also quantify the significance explicitly.

**Change in manuscript:** The methodology was extended with the description of significance test (line 97). The significance tests were performed for each assessed coefficient. The results of tests (corresponding well with the visual assessment presented in Figs. 4 and 5) were presented in a new paragraph in Section 4.1 (line 144).

**4. Comment*: "l98 - it would be nice to know what is the fraction of zeroes and non-zeroes in the data used, and how that might be impacting the binary cross-correlation results. here. From my experience, storms in warmer climates are getting smaller in size, hence the fraction of zeroes is increasing.*"**

**Reply:** We calculated the fractions of zeroes in all time series. Two different thresholds were applied to determine dry days: 0 mm and 0.1 mm. The results for the model 1A are presented in the following table:

| grid-box | threshold 0.0 mm | | | threshold 0.1 mm | | |
|---|---|---|---|---|---|---|
| | historical | Future | Difference | historical | future | Difference |
| 1 | 0.267 | 0.300 | 0.033 | 0.379 | 0.406 | 0.027 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2 | 0.259 | 0.293 | 0.034 | 0.367 | 0.396 | 0.029 |
| 3 | 0.252 | 0.287 | 0.036 | 0.359 | 0.387 | 0.028 |
| 4 | 0.248 | 0.287 | 0.039 | 0.358 | 0.387 | 0.028 |
| 5 | 0.267 | 0.302 | 0.035 | 0.388 | 0.413 | 0.025 |
| 6 | 0.261 | 0.298 | 0.037 | 0.378 | 0.407 | 0.028 |
| 7 | 0.255 | 0.292 | 0.037 | 0.371 | 0.397 | 0.026 |
| 8 | 0.250 | 0.290 | 0.040 | 0.368 | 0.396 | 0.028 |
| 9 | 0.271 | 0.310 | 0.039 | 0.398 | 0.425 | 0.027 |
| 10 | 0.268 | 0.306 | 0.038 | 0.390 | 0.416 | 0.026 |
| 11 | 0.261 | 0.299 | 0.037 | 0.377 | 0.407 | 0.030 |
| 12 | 0.257 | 0.296 | 0.039 | 0.376 | 0.405 | 0.029 |

As seen from the table, the fraction of zeroes slightly increases regardless the applied threshold, which is in accordance with your experience. The results for the other models were similar (except the model 2B, where the fraction of zeroes remains almost constant). We can add this information (in a reduced form) to the paper.

**Change in manuscript:** The information about the fractions of zeroes and their changes were added to Section 4.1 (line 119).

**5. Comment:** *"What I think the authors are doing is to estimate sample correlations of the current and the future independently (i.e. taking their respective sample means and standard deviations). As a result of which they may be finding the change is insignificant, whereas the change with respect to a faxed reference (sat the historical climate) may be more. At the very least, some clarification on how the correlations are estimated as well as the change in the first order statistics that are used in its estimation is needed."*

**Reply:** Indeed, the sample correlations are estimated independently, and we agree that the results may change slightly if a fixed reference is considered. We will extend the whole description of the bootstrap procedure and the details on estimation of cross- and auto- correlation will be given in detail in the revised manuscript, if we are invited to submit revised version.

**Change in manuscript:** We improved the description of the estimation of changes in correlation and autocorrelation as well as we mention the changes in mean and the fact that this could have some influence on the significance of changes (lines 77 and 148).


**6. Comment:** *"l110 - the negative change in autocorrelations is consistent with my experience. If one were to consider changes in the associated means and standard deviations this becomes even greater. Additionally, these changes manifest themselves at longer time scales as well, as a AR1 model structure is not a great characterisation of the system. This has formed the argument for the range of "nesting" approaches in the literature for rainfall generation and bias correction. This needs to be discussed somewhere in the paper at the very least."*

**Reply:** We agree that the AR1 model is not sufficient characterization of the system. We will discuss this explicitly in the revised version of the manuscript if we are invited.

**Change in manuscript:** We added a comment on this in the revised version (line 129).


**7. Comment:** *"l115 - The figure title states 95% confidence of correlations. Does this mean 95% of the 66 correlations, or all the resampled correlation estimates as well?"*

**Reply:** The confidence intervals presented in Fig. 4 are derived for each correlation coefficient separately. The title of the figure should state this more clearly, therefore we suggest changing the first sentence of the title as follows: "*The 95% confidence intervals of the individual cross-correlation coefficients for overlapping wet periods for all models…*". The first sentence of the Figure 5's title will be changed in the same way.

**Change in manuscript:** The titles were changed as suggested (lines 154 and 160).

**8. Comment:** *"l180 - I believe the authors need to write a simple equation to show how they will ascertain their dependance outlier, and give us results of some tests that help argue these are genuine outliers and not examples of real extremes that would be of interest in hydrology. This is kind of important as this seems to be the key contribution the paper is making."*

**Reply:** We do not propose any new dependence measure; rather we are offering a procedure how to obtain robust estimates of correlation and autocorrelation by removing few data points that have large influence on the estimates. Therefore, there is no straightforward formula. We agree that real extremes are of special interest in hydrology, but our point is that dependence outliers (no matter if genuine or real extremes) affect data transformation within some bias correction methods and may possibly distort the dependence structure of the corrected data. In our concept, the dependence outlier is <u>any</u> value deviating from the correlation structure (as demonstrated in Fig. 7), regardless of the origin of this value. The distinguishing between real extremes and "true" outliers (say for example measurement errors) can be only hardly based clearly on statistics. This would involve an expert assessment of particular event, considering local conditions and the data from surrounding locations. Therefore we cannot design a simple equation for such purposes. Nevertheless, the real extremes as well as genuine outliers affect the correlation structures in the same way, which subsequently affects the bias corrections (or stochastic generators). Therefore the dependence outliers, regardless of their origin, can be detected and should be removed from the calibration data. Note that it is possible to insert these data back later.

**Change in manuscript:** The issue of the origin of the outliers was included into the Conclusions of the paper (line 256).

# Technical note: Changes of cross- and auto-dependence structures in climate projections of daily precipitation and their sensitivity to outliers

Jan Hnilica[1,2], Martin Hanel[2] and Vladimír Puš[2]

[1] The Czech Academy of Sciences, Institute of Hydrodynamics, Pod Paťankou 5, Prague 6, 166 12, Czech Republic

5  [2] Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Prague 6 – Suchdol, 165 21, Czech Republic

*Correspondence to*: Jan Hnilica (hnilica@ih.cas.cz)

**Abstract.** Simulations of regional or global climate models are often used for climate change impact assessment. To eliminate systematic errors, which are inherent to all climate model simulations, a number of post processing (statistical

10  downscaling) methods have been proposed recently. In addition to basic statistical properties of simulated variables, some of these methods consider also a dependence structure between or within variables. In the present paper we assess the changes in cross- and auto-correlation structures of daily precipitation in six regional climate model simulations. In addition the effect of outliers is explored making a distinction between ordinary outliers (i.e. values exceptionally small or large) and dependence outliers (values deviating from dependence structures). It is demonstrated that correlation estimates can be

15  strongly influenced by a few outliers even in large data sets. In turn, any statistical downscaling method relying on sample correlation can therefore provide misleading results. An exploratory procedure is proposed to detect the dependence outliers in multi-variate data and to quantify their impact on correlation structures.

## 1 Introduction

The investigation of climate change impact on the hydrological cycle is one of the crucial topics in the field of water

20  resources management and planning (Mehrotra and Saharma, 2015). Simulations of regional and global climate models (RCMs and GCMs) represent a fundamental data source for climate change impact studies. It is well known that raw climate model outputs cannot be used directly in impact studies due to inherent biases which are found even for basic statistical properties (Chen et al., 2015). The bias is caused primarily by simplified representation of important physical processes (Solomon et al., 2007), which is often resulting from low spatial resolution of the RCMs.

25  Therefore, many methods have been developed to post-process the climate model outputs in order to move their statistical indicators closer to observations. The An overview of these methods is presented e.g. by Maraun et al. (2010). Precipitation is a key input into the hydrological climate change impact studies and at the same time it belongs to meteorological variables that are most affected by bias. The comparison of correction methods commonly used for precipitation data is provided by Teutschbein and Seibert (2012). Nevertheless, these standard methods correct only the bias in statistical indicators (mean,

30  variance, distribution function) of individual variables. The bias in persistence parameters of time series as well as the bias in

cross-dependence structures between variables is often neglected. However, the dependence structures of the meteorological variables affect the hydrological response of a catchment (Bárdossy and Pegram, 2012), thus their inadequate representation in the data can impair hydrological impact studies (Teng et al., 2015; Hanel et al., 2017).

In recent years several studies attempted to overcome this limitation. Hoffmann and Rath (2012) and Piani and Haerter (2012) focused on the relationship between precipitation and temperature data from a single location. Bárdossy and Pegram (2012) developed two procedures correcting a spatial correlation structure of RCM precipitation. Mao et al. (2015) proposed a stochastic multivariate procedure based on copulas. Johnson and Sharma (2012) developed a procedure correcting common statistics (mean, variance) together with lag-1 autocorrelation in multiple-time scales. The procedure was later extended with a recursive approach by Mehrotra and Sharma (2015) and subsequently with a non-parametric quantile mapping by Mehrotra and Sharma (2016) to correct the bias in auto- and cross-dependence structures across multiple time scales. An approach based on the principal components was presented by Hnilica et al. (2017), correcting bias in cross-covariance and cross-correlation structures.

This study is focused on a temporal stability of dependence structures. We evaluate the temporal changes in cross-and auto-correlation structures in multivariate precipitation data simulated by an ensemble of climate models. We further investigate whether the magnitude of the changes exceeds considerably the natural variability. ~~The attention~~ Attention is finally paid to the effect of outlying values, which can significantly affect the correlations and can thus lead to artefacts in bias-corrected time series.

The paper is organised as follows. In Sect. 2 the data used in this study are presented and Sect. 3 describes the methodology. In Sect. 4 the results are reported and in Sect. 5 their consequences for climate changes impact studies are discussed.

## 2 Data and study area

The daily precipitation ~~sums~~ data from six EURO-CORDEX (Giorgi et al., 2009) regional climate models were considered. The ensemble of models was composed of two RCMs (CCLM, RCA) driven by three GCMs (EC_EARTH, HadGEM2-ES and MPI-ESM-LR); see Table 1 for the overview. The simulations with 0.11 degree spatial resolution forced by RCP8.5 were used. The data from twelve model grid-boxes located in the western part of the Czech Republic were analysed, see Fig. 1 for the details of the area. The control period spans the years from 1971 to 2000, the future period the years from 2051 to 2080.

2

**Table 1. Global and regional climate models used in the present study.**

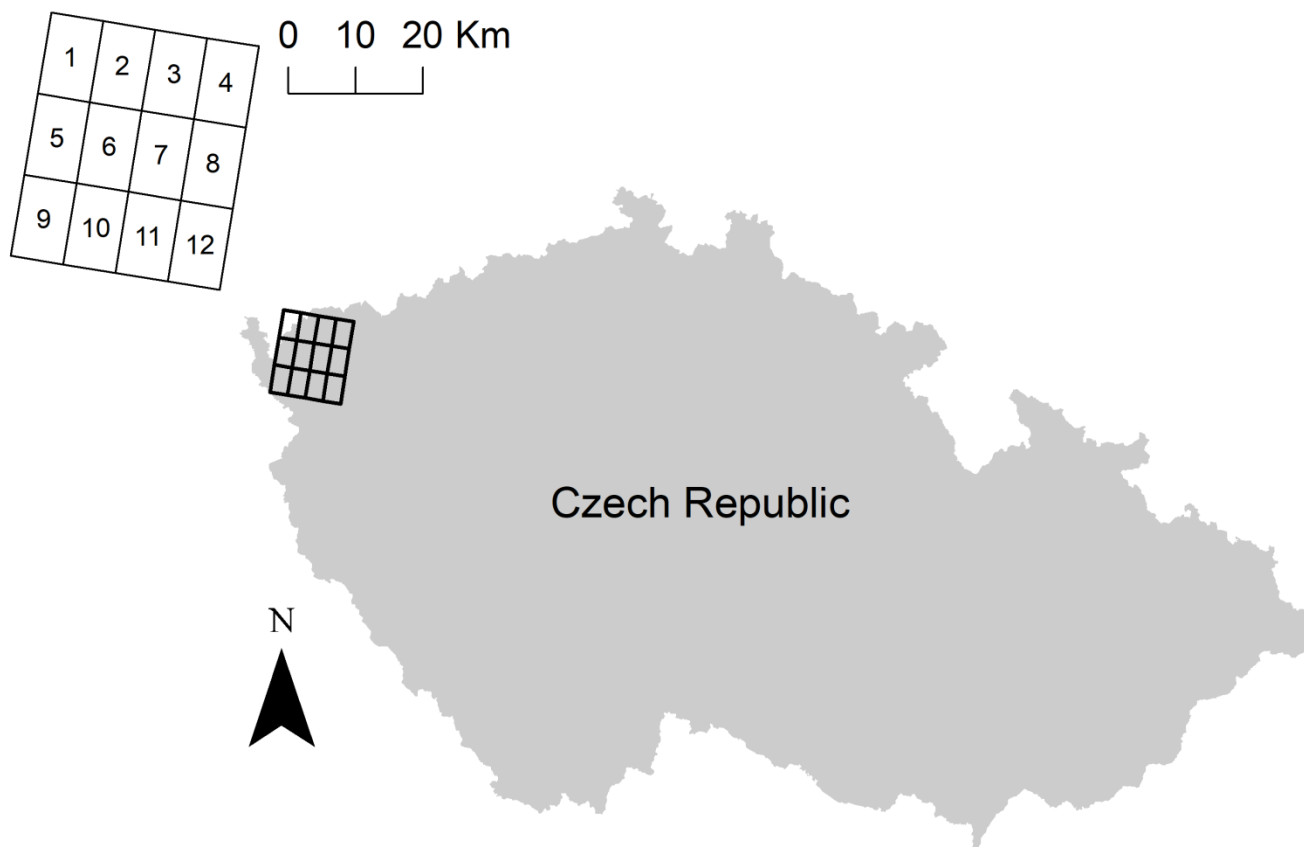| GCM | RCM | ID |
|---|---|---|
| EC-EARTH | CCLM-4-8-17 | 1A |
| | RCA4 | 1B |
| HadGEM2-ES | CCLM-4-8-17 | 2A |
| | RCA4 | 2B |
| MPI-ESM-LR | CCLM-4-8-17 | 3A |
| | RCA4 | 3B |

60



**Figure 1. Location of the considered grid-boxes in the Czech Republic.**

## 3 Methods

65    The wet and dry periods were treated separately in this study. The cross-correlations were calculated in two stages. Firstly the binary cross-correlations were calculated to assess the correspondence of wet/dry periods, using the time series with the values replaced by 0 (dry day) or by 1 (wet day). In the second stage the cross-correlations of overlapping wet periods were calculated. The auto correlations were analysed through the lag-1 auto-correlation coefficient, where only the non-zero pairs of neighbouring values $x_i$ and $x_{i+1}$ were considered.

70    The individual grid-boxes were labelled by numbers 1-12, as shown by labels in Fig. 1. The cross-correlation between the grid-boxes $i$ and $j$ is denoted as $r_{i,j}$. The symbol **R** denotes the correlation matrix (i.e. the square matrix with elements $r_{i,j}$). The lag-1 auto-correlation from grid-box $i$ is denoted as $r_i^1$. If appropriate, the subscripts denoting the grid-boxes are omitted for clarity.

The changes of correlation coefficients were calculated as

75    $$\Delta r = r_{\mathrm{F}} - r_{\mathrm{C}} \tag{1}$$

where $\Delta r$ denotes the change of $r$ (cross- or auto-correlation), subscripts F and C denote the future and control periods, respectively. Note that the first order moments needed for calculation of $r_{\mathrm{F}}$ and $r_{\mathrm{C}}$ are calculated independently for future and control period. Another option, leading potentially to larger changes is to consider fixed reference, e.g. the mean for the control period. However, the changes in the mean are relatively small, the average change is 0.14 mm across all considered

80    simulations, which represents approximately 5% of the value from the control period. Therefore the difference in the calculated $\Delta r$ and their significance are not expected to be large.

The sampling variability of individual cross- and auto-correlation was investigated to assess the statistical significance of their changes. The confidence intervals were derived using the block bootstrap approach (Davison and Hinkley, 1997). Specifically, the confidence interval around the correlation $r_{i,j}$ was obtained as follows:

85    1.    one-year blocks from the time series for basins $i$ and $j$ were randomly selected with replacement (30-times to obtain the same sample size as the original data), subsequently the correlation of the 30-year sample was calculated

2.    step 1 was repeated 1000 times

3.    the 95% confidence interval was derived as a range between the 0.025 and 0.975 quantiles of the resampled correlations.

90    The block approach was chosen to preserve seasonal variability in the bootstrap samples. For the presentation of confidence intervals, the unique identifier (ID) was assigned to each pair of grid-boxes, the numbering was done according to rows of

4

correlation matrix; the scheme is depicted in Fig. 2. The confidence intervals for auto-correlation were derived in the same way using one-year blocks of time series. ~~The joints of the adjacent blocks were not included in the calculation.~~ Due to random selection of the blocks the beginning part of blocks is independent on the end of the previous blocks. To minimize bias introduced by block resampling, data that are potentially influenced (joints of the adjacent blocks) were not considered for the calculation of the serial correlation.

The confidence intervals around the correlations from control and future period were used to visually assess their overlap. In addition, the real bootstrap-based tests of significance of individual changes were performed. In each of thousand steps the correlation of resampled control data was subtracted from the correlation of resampled future data. The change was found significant if the confidence interval of these differences contains zero.

grid-box number

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | | | |
| 2 | 1 | | | | | | | | | | | |
| 3 | 2 | 3 | | | | | | | | | | |
| 4 | 4 | 5 | 6 | | | | | | | | | |
| 5 | 7 | 8 | 9 | 10 | | | | | | | | |
| 6 | 11 | 12 | 13 | 14 | 15 | | | | | | | |
| 7 | 16 | 17 | 18 | 19 | 20 | 21 | | | | | | |
| 8 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | | | | | |
| 9 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | | | | |
| 10 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | | | |
| 11 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | | |
| 12 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | |

grid-box number

105    **Figure 2. The numbering of individual pairs of grid-boxes. The figure depicts the correlation matrix, the orders of rows/columns correspond to the grid-box labels from Fig. 1. The sub-diagonal part of the (symmetrical) matrix was used for numbering of individual pairs of grid-boxes – the numbers inside of the matrix represent the identifiers used in Fig. 4.**

## 4 Results

110    ### 4.1 Changes in correlation structures

In the case of 12-dimensional data, the change of the cross-correlation structure consists of changes in $r_{i,j}$ coefficients for 66 pairs of grid-boxes (corresponding to the sub-diagonal part of the correlation matrix). For clarity, these 66 changes are presented in the form of box-plots for individual models.

Figures 3a and 3b present the changes in the binary cross-correlations and in the cross-correlations of wet periods, respectively. As seen from the figures, the binary correlations are relatively stable; their changes range approximately from -0.02 to 0.03. Therefore, the correspondence of wet/dry periods between individual grid-boxes remains similar in the control and future periods. The correlations of wet periods change more substantially, the changes range from -0.16 to 0.05. Nevertheless, there are strong differences between individual models, the models 1A and 2A reach noticeably higher changes than other models. In addition, the changes in fractions of dry days were calculated. In general, the fraction of zeroes fluctuates around 0.25 in time series across all models and it was found that it slightly increases in most cases. The difference can be found between the regional models A and B. While in the simulations of the model A the fractions of zeroes increases on average by 0.05, for the model B the average increment is only 0.009.

Figure 3c presents the changes of lag-1 auto-correlations; the box-plots for individual models are compiled from 12 changes in time series from individual grid-boxes. The changes range from -0.1 to 0.025, the widest range of changes is reached by the model 2A. The maximal relative changes in cross-correlation reach up to 18% of the value from the control period, in the case of auto-correlation it is almost 45%. Note that the auto-correlations are in general markedly lower than cross-correlations (the mean cross-correlation of individual models exceeded 0.8; the mean lag-1 auto-correlation is around 0.23). This is because the auto-correlations are in general markedly lower than cross-correlations (the mean cross-correlation of individual models exceeds 0.8; the mean lag-1 auto-correlation is around 0.23). In addition, it is worth noting that the reported changes do not completely characterise the changes in auto-dependence structure. Substantial changes in dependence at longer temporal scales have been reported in several studies (Mehrotra and Sharma, 2015, 2016; Hanel et al., 2017).
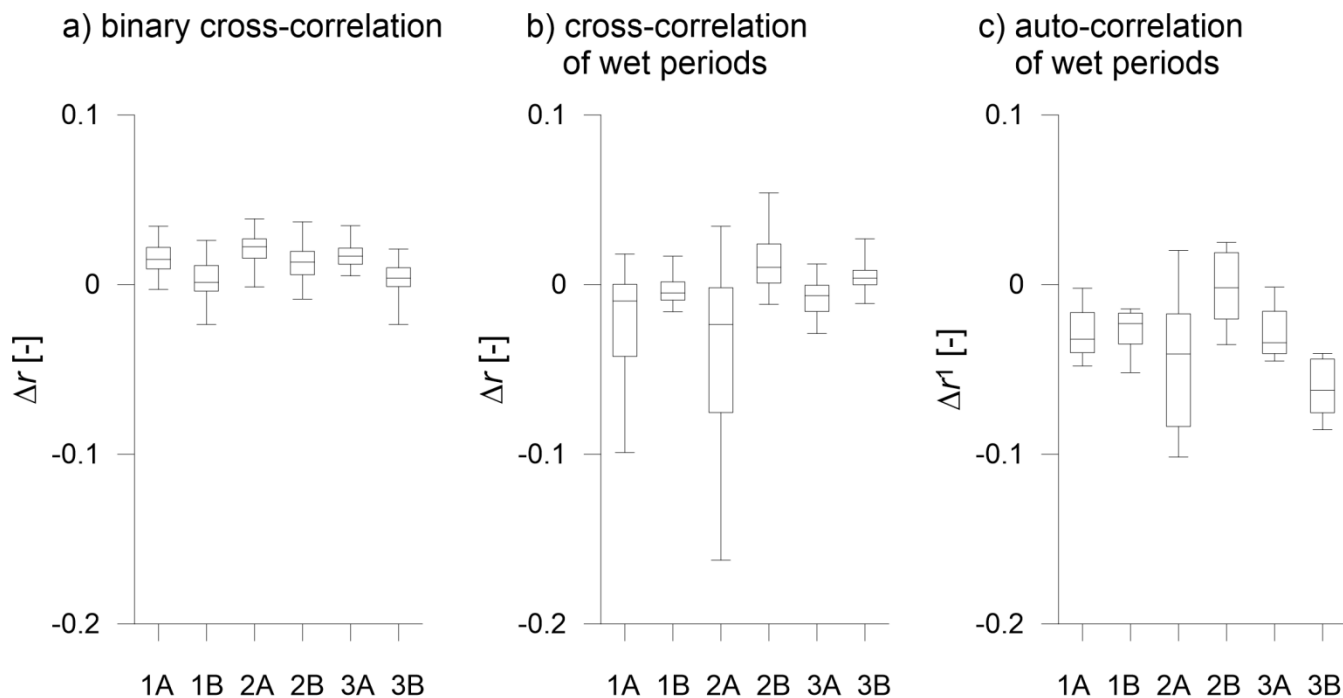
**Figure 3. Overview of the changes in correlation structures for all models: (a) the changes of binary cross-correlations, (b) the changes of cross-correlations of overlapping wet periods, (c) the changes of lag-1 auto-correlations.**

The significance of the changes in wet-periods correlations was assessed using a block bootstrap. Figure 4 presents the 95% confidence intervals of individual cross-correlations for all models. The blue dividers identify the successive rows below the diagonal in the correlation matrix. In general, the majority of changes show a little significance, the intervals from control and future periods overlap considerably (except models 1A and 2A, which show exceptionally wide intervals for the future period in many cases). Figure 5 shows the same for lag-1 auto-correlations of individual grid-boxes. Also in this case the majority of changes do not exceed the sampling variability; the most significant changes are reached by the model 3B, but the overall trend is a drop in the future.

To verify these results, the significance of individual changes was tested using the bootstrap approach. The results of tests correspond well with the visual assessment presented in Fig. 4 and Fig 5. In case of cross-correlations, only four changes were found significant for the model 1A, no one for the model 1B, two for the model 2A, eight for the model 2B, two for the model 3A and no one for the model 3B. In case of auto-correlations the significant changes were found only for the model 3B. Note, however, that the fraction of significant changes might be larger in the case that fixed reference is used for calculating correlations and autocorrelations (see Section 3).
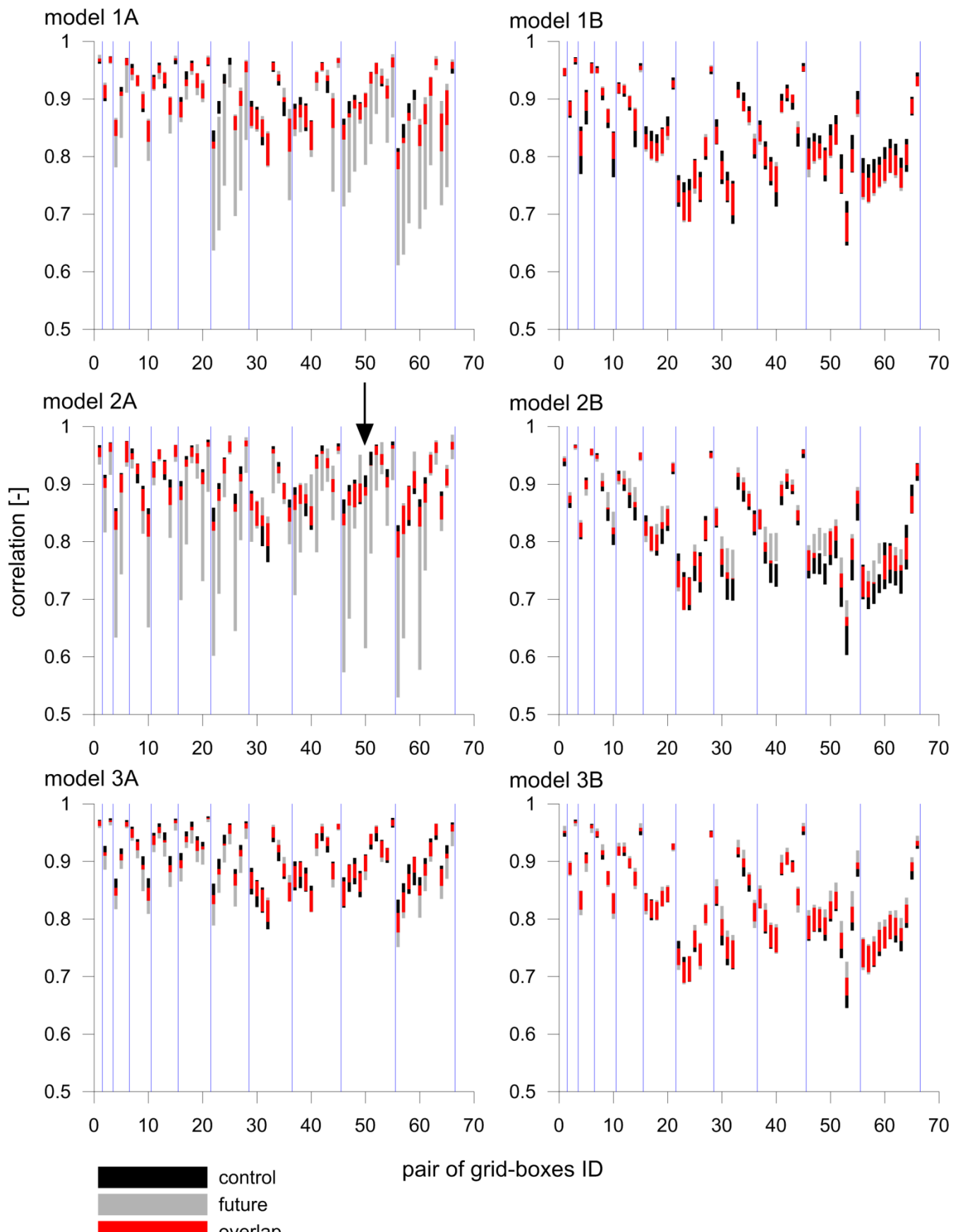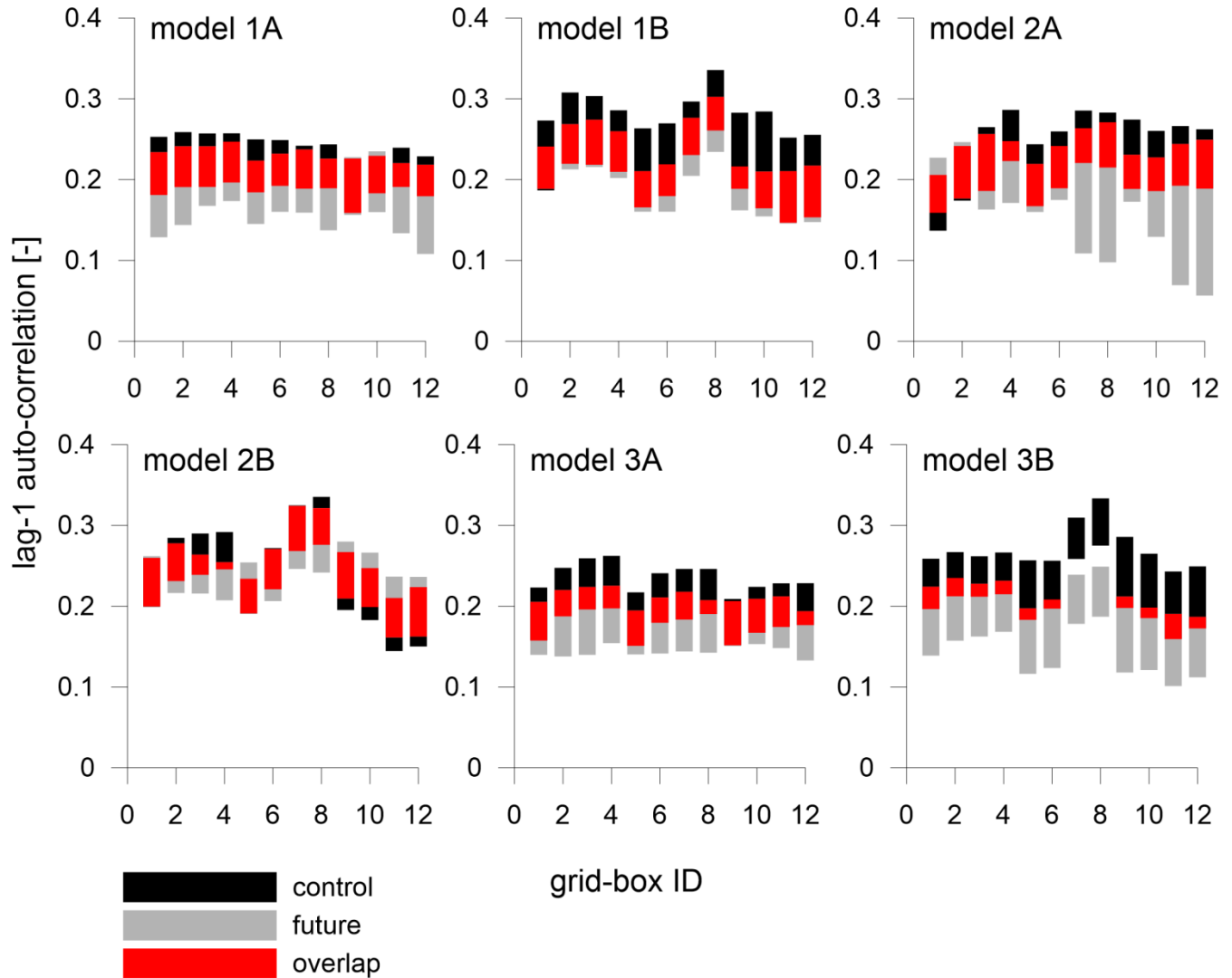
7

model 1A  model 1B

model 2A  model 2B

model 3A  model 3B

correlation [-]

pair of grid-boxes ID

control
future
overlap

Figure 4. The 95% confidence intervals of the individual cross-correlation coefficients for overlapping wet periods for all models. The identifiers of grid-box pairs (ID) are explained in Fig. 2. The blue lines separate identifiers located in successive rows in the correlation matrix (see Fig. 2). The arrow marks the confidence intervals around the $r_{5, 11}$ (ID 50) of the model 2A, discussed in more detail in section 4.2.

160 **Figure 5. The 95% confidence intervals of the individual lag-1 auto-correlation coefficients for all models.**

## 4.2 Effect of outliers

The previous section demonstrated that in some cases the changes in cross-correlation show a little significance despite of their high absolute values, which is particularly related to models 1A and 2A. At the same time, it can be seen in Fig. 4 that

165 some confidence intervals for these models are exceptionally wide. Further analyses showed that this instability of correlation estimates is introduced by outlying values, which causes seeming changes of the correlation structures.

In the simulation of the model 2A, the sample correlation $r_{5,11}$ decreased from 0.90 in the control period to 0.73 in the future period. Figure 6a depicts the data from the future period (values from the grid-box 5 plotted against values from the grid-box 11, the data with any zero values are excluded). The decrease is in large part caused by one outlying point, which is

170 circled in the plot. Its removal from the data increases the correlation in the future period to 0.86, which markedly reduces the change. On the other hand, high values do not necessarily affect the correlation, as seen in Fig. 6b, where the data from grid-boxes 11 and 12 are plotted (again the model 2A, future period). The circled outlier does not affect the correlation in this case, since the location of the point is in accordance with the configuration of data – the point lies approximately in a direction of a potential regression line.

175 Outlying values affect also the auto-correlation. The largest change of the auto-correlation was achieved by the model 2A, where $r_{12}^1$ decreased from 0.23 in the control period to 0.12 in the future period. This decrease in caused by the outlier 349.4 mm in the future data; this extraordinary value was simulated by the model 2A for the 8$^{th}$ May 2080. Figure 6c depicts the data for the calculation of $r_{12}^1$, i.e. the values $x_i$ plotted against the values $x_{i+1}$, where $i$ denotes the order of the value $x$ in the time series. The outlier is employed twice within the calculation (as $x_i$ and as $x_{i+1}$, circled values in Fig. 6c) which markedly

180 affects the result. If the outlier is removed from the time series, $r_{12}^1$ increases from 0.12 to 0.22, which reduces the change almost to zero. The calculation of other members of the auto-correlation function is affected by the outlier in the same way. We note, that the effect of an outlier to the auto-correlation strongly depends on values, by which the outlier is surrounded in the time series. The presence of a noticeable outlier thus makes the calculation of the auto-correlation very unstable.
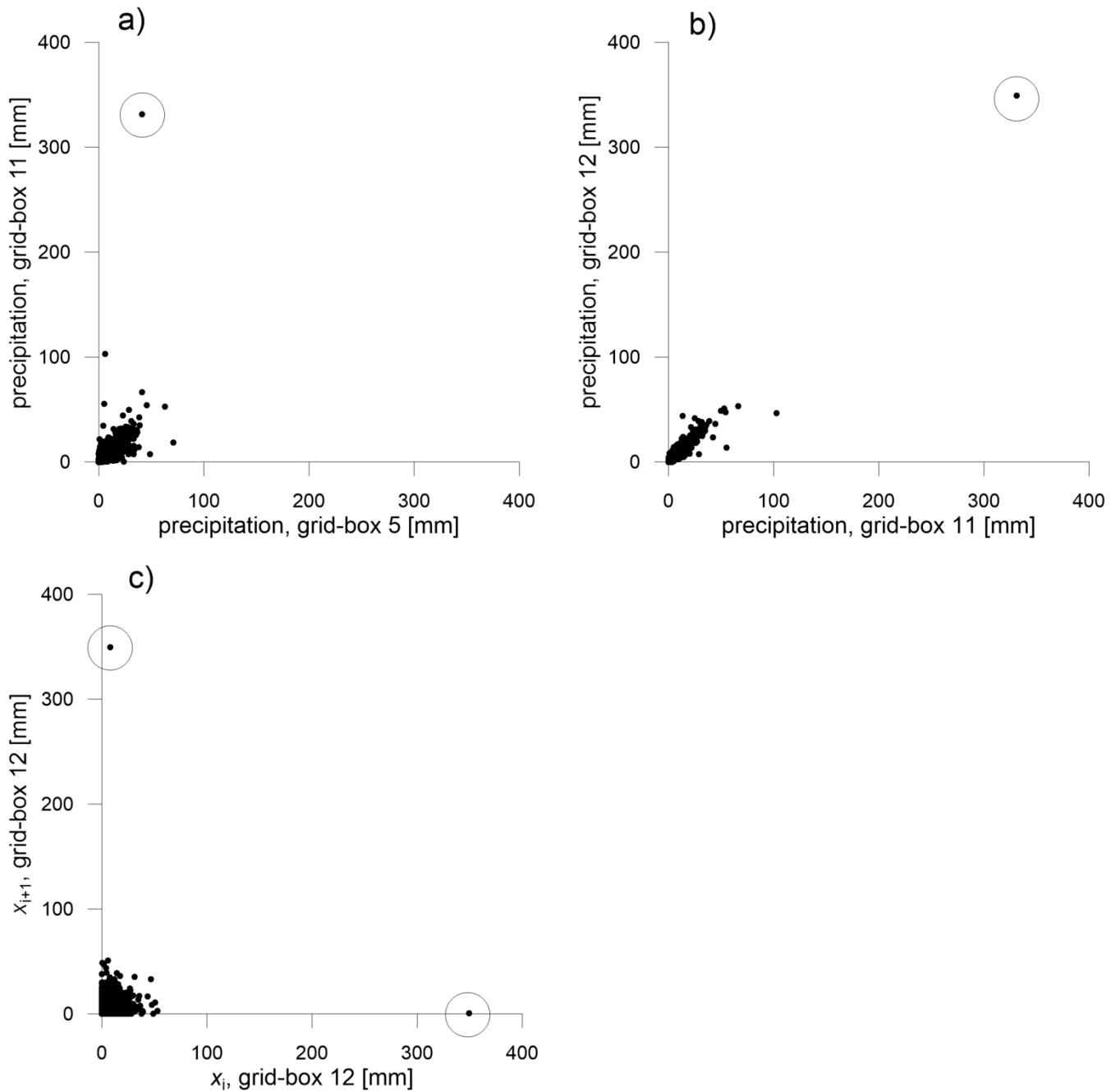
Figure 6. The effect of outliers on correlation structures of model 2A in the future period, the outliers are circled: (a) the outlier strongly affecting the cross-correlation, (b) the outlier with no effect on the correlation, (c) the outlier affecting the calculation of the serial correlation $r_{12}^1$.

**4.3 Detection of outliers**

190    The examples showed that outliers can distort cross- and auto-correlation structures of a large dataset comprising many thousands of values. Nevertheless~~;~~, it should be realized that not each extreme value necessarily affects the correlation (as seen in Fig. 6b). Therefore, a more specific concept of outliers is presented in this study. Values deviating from the correlation structure are denoted as *dependence* outliers. As well as ordinary outliers, the dependence outliers are values at a long distance from the origin; nevertheless, the difference between them and ordinary outliers ~~consists in a~~depends on the

195    coordinate system in which the distance is measured. Figure 7 illustrates this by an example of synthetic 2-dimensional data. The dashed lines and coordinates in square brackets define the standard (canonical) coordinate system. The ordinary outliers are points in a long distance from the origin [0, 0], measured in the standard coordinates; the point A represents an example. The solid lines and coordinates in round brackets define an alternative coordinate system, which reflects the intensity of linear dependence between the variables X and Y. The dependence outliers are points in a long distance from the origin (0,

200    0), measured in the alternative coordinates; the point B represents an example. Let us remark that the point B does not represent an extreme value neither in X nor Y data, in contrast to the point A.  Nevertheless, the point B deviates from the dependence structure, which comes out when its distance from (0, 0) in alternative coordinates is calculated. The alternative coordinate system is constructed through the covariance matrix of the data. The directions of the axes are given by the eigenvectors of the matrix, the lengths of unit vectors are given by the square root of the corresponding eigenvalues and the

205    origin is located in the mean of the data. The construction of the system is related to the principal component analysis; see for example Wilks (2011) for details.
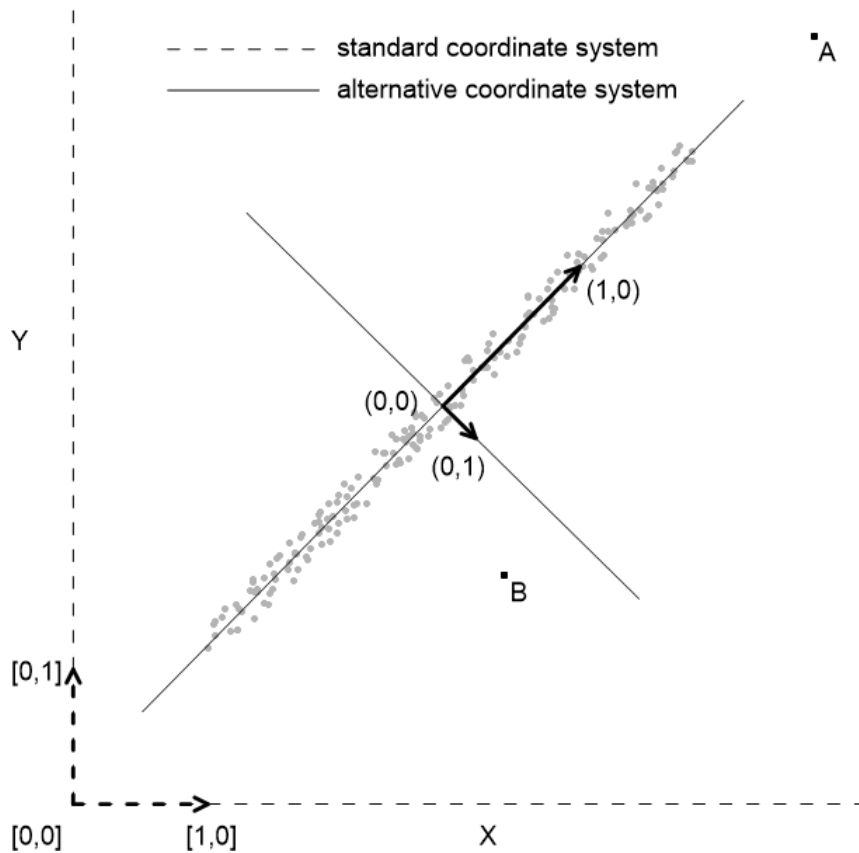
12

**Figure 7. The difference between the ordinary and dependence outliers. The dashed lines define the standard coordinate system; the solid lines define an alternative coordinate system. The points outlying in the standard coordinates are ordinary outliers (point A); the points outlying in the alternative coordinates are denoted as dependence outliers (point deviating from the dependence structure, point B). The construction of alternative coordinates is explained in the text.**

The problem is that the presence of outliers is not simply detectable from the changes of dependence structures. It can be indicated ~~either~~ indirectly from the analysis of sampling variability; nevertheless the wide confidence intervals do not necessarily imply the presence of outliers. Or alternatively, it can be found ~~out~~ when the individual pairs of datasets are visually checked. We propose a procedure allowing for identification of significant dependence outliers and assessment of their effect on correlation structure. The procedure consists of three steps:

1. the most outlying (multi-variate) value is found in the data (in alternative coordinates)

2. the value is removed from the data and a new correlation matrix is calculated

3. a difference between the new and the previous correlation matrix is calculated and recorded.

13

220 These three steps are repeated. The difference in step 3 is quantified through

$$\delta \mathbf{R}_m = \|\mathbf{R}_m - \mathbf{R}_{m-1}\| \tag{2}$$

where $\mathbf{R}_m$ denotes the correlation matrix of the data from which $m$ largest outliers were removed, the symbol $\|\cdot\|$ denotes the Frobenius matrix norm. The most outlying value in the step 1 is simply defined as the value with the highest distance from origin (measured in the alternative coordinates). A result of the proposed exploratory procedure is a sequence of $\delta \mathbf{R}_m$, which

225 clearly indicates the presence of noticeable outliers. We note that the alternative coordinate system in which the dependence outliers are searched is data-dependent (in contrast to the standard coordinates). It means that after each outlier removal the alternative coordinates slightly change and must be recalculated to correspond to the ~~actual~~ remaining data.

The procedure is demonstrated on two simple 2-dimensional examples. Figure 8a depicts the sequence of $\delta \mathbf{R}_m$ for the data from Fig. 6a. A massive impact of the first outlier is clearly visible; the removal of next outliers already does not affect the

230 correlation matrix substantially (the first member $\delta \mathbf{R}_1$ corresponds to the circled outlier in Fig. 6a). Figure 8b depicts the same for the data from Fig. 6b; a gradual evolution of $\delta \mathbf{R}_m$ indicates that the data do not contain noticeable (dependence) outliers.
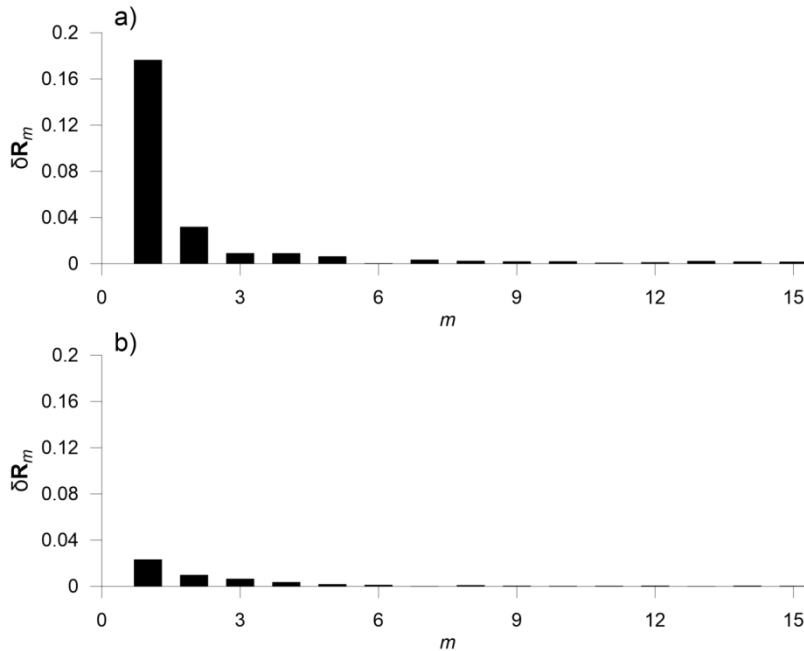


**Figure 8. The demonstration of the exploratory procedure: a) the detection of dependence outliers for 2-dimensional data from**
235 **Fig. 6a – the plot of $\delta \mathrm{R}_m$ indicates a noticeable outlier in the data, b) the same for the data from Fig. 6b – a gradual evolution of**
**$\delta \mathrm{R}_m$ indicates that data do not contain dependence outliers.**

A real ~~utility of the procedure consists in a fact,~~ usefulness of the procedure is that a large set of multi-variate data can be explored as a whole. The $n$-dimensional outliers can be searched in the same way as the 2-dimensional outliers in the examples presented above. A result of the procedure is always a one-dimensional plot of $\delta R_m$, regardless of a dimension of input dataset. Figure 9 shows the plots of $\delta R_m$ for the complete 12-dimensional data from the future period for all models. The strong outliers in data from 1A and 2A are simply detectable from the plots. Generally, a plot of $\delta R_m$ enables a simple assessment of the internal structure of the data and a direct evaluation of an importance of individual outliers.
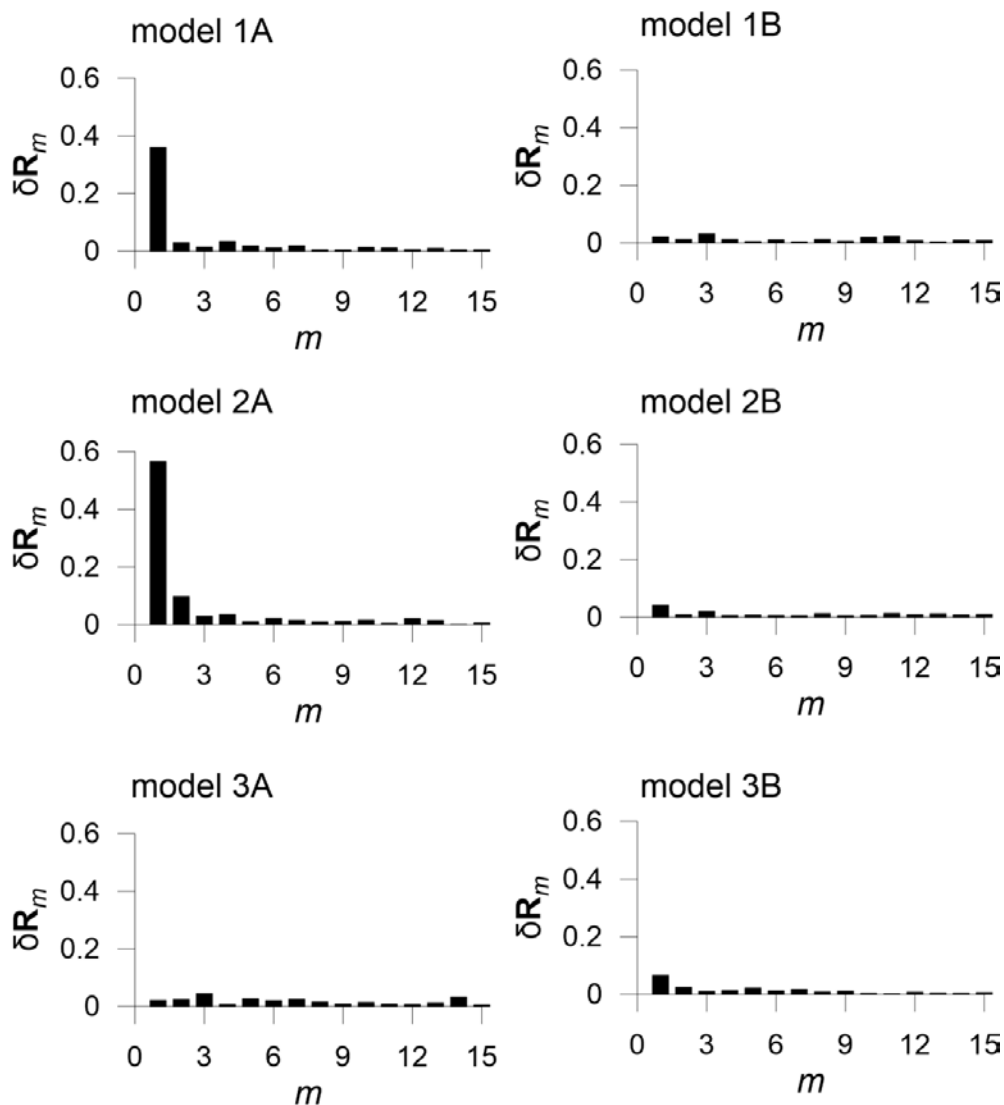


**Figure 9. The detection of dependence outliers for complete 12-dimensional data from all models in the future period. The strong outliers in data from the models 1A and 2A are clearly distinguishable.**

15

**5 Conclusions**

The examples demonstrated that outliers can strongly affect the cross- and auto-correlation structures of the data comprising many thousands of values. In general, it must be stressed that the presence of outliers cannot be considered as a bias. The extreme precipitation values as well as the dependence outliers naturally occur. Nevertheless, although the dependence structures are markedly influenced by a small number of outliers, they characterize the data as a whole. Therefore a substantial bias can arise when the data with noticeable outliers are used to assess the dependence structures or when their dependence structures are used e.g. for calibration of the bias correction functions. The cross- and auto-correlation structures are the key ingredients in several multi-variate bias correction methods, for example in Mehrotra and Sharma (2015) and Mehrotra and Sharma (2016). The results based on these methods can be devalued by outliers, see the supplement to this paper.

From this point of view there is no need to distinguish between real extremes and "genuine" outliers (for example measurement errors). The real extremes as well as genuine outliers affect the correlation structures in the same way, which subsequently affects the bias corrections (or stochastic generators). Therefore the dependence outliers, regardless of their origin, should be removed from the calibration data. The appropriate tool for testing the presence of outliers is the analysis of the difference $\delta \mathbf{R}_m$ between the new and previous correlation matrix (Eq. 2) presented above; the exploratory procedure can be automatized and included in the modelling chain as a pre-processing step to automatically remove at least the most noticeable outliers.

The analysis of significance showed, that in most cases the correlations are stable in time, their changes are insignificant and caused by outlying values. Therefore the climate projection can be interpreted as a linear transformation of an initial state, because a nonlinear transformation would change the correlations substantially. From this point of view a reasonable scenario of future precipitation can be obtained by the corresponding linear transformation of observations, i.e. by the multiplicative delta method (Déqué, 2007). Such an approach avoids the problems of complex bias correction methods (e.g. their increasing complexity and unclear effect on climate change signal), which are recently the subjects of serious criticism, for example by Ehret et al. (2012) or Maraun et al. (2017).

16

## References

Bárdossy, A. and Pegram, G.: Multiscale spatial recorrelation of RCM precipitation to produce unbiased climate change scenarios over large areas and small, Water Resour. Res., 48, W09502, doi:10.1029/2011WR011524, 2012.

Chen. J., Brissette. F. P. and Lucas-Picher, P.: Assessing the limits of bias-correcting climate model outputs for climate change impact studies, J. Geophys. Res.-Atmos, 120, 1123–1136, doi:10.1002/2014JD022635, 2015.

Davison, A. C. and Hinkley, D. V.: Bootstrap methods and their application, Cambridge University Press, Cambridge, United Kingdom, 1997.

Déqué, M.: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, Global Planet. Change, 57, 16–26, doi:10.1016/j.gloplacha.2006.11.030, 2007.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K. and Liebert, J.: Should we apply bias correction to global and regional climate model data?, Hydrol. Earth Syst. Sc., 16, 3391–3404, doi:10.5194/hess-16-3391-2012, 2012.

Giorgi, F., Jones, C. and Ghassem, R.: Addressing climate information needs at the regional level: the CORDEX framework, World Meteorological Organization (WMO) Bulletin, 58, 175–183, 2009.

Hanel, M., Kožín, R., Heřmanovský, M. and Roub, R.: An R package for assessment of statistical downscaling methods for hydrological climate change impact studies, Environ. Modell. Softw., 95, 22–28, 2017.

Hnilica, J., Hanel, M. and Puš, V.: Multisite bias correction of precipitation data from regional climate models, Int. J. Climatol., 37, 2934–2946, doi:10.1002/joc.4890, 2017.

Hoffmann, H. and Rath, T.: Meteorologically consistent bias correction of climate time series for agricultural models, Theor. Appl. Climatol., 110, 129–141, doi:10.1007/s00704-012-0618-x, 2012.

Johnson, F. and Sharma A.: A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations, Water Resour. Res., 48, W01504, doi:10.1029/2011WR010464, 2012.

305    Mao, G., Vogl, S., Laux, P., Wagner, S. and Kunstmann, H.: Stochastic bias correction of dynamically downscaled precipitation fields for Germany through copula-based integration of gridded observation data, Hydrol. Earth Syst. Sc., 19, 1787–1806, doi:l0.5194/hess-19-1787-2015, 2015.

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., Hagemann, S., Richter, I., Soares, P. M. M., Hall, A. and Mearns, L. O.: Towards process-informed bias correction of climate change simulations, Nat. Clim. Change, 7(11), 764-773, doi: 10.1038/NCLIMATE3418, 2017.

310    Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J.,Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M. and Thiele-Eich, I.: Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, RG3003, doi:10.1029/2009RG000314, 2010.

Mehrotra, R. and Sharma, A.: Correcting for systematic biases in multiple raw GCM variables across a range of timescales, 315    J. Hydrol., 520, 214–223, doi:10.1016/j.jhydrol.2014.11.037, 2015.

Mehrotra, R. and Sharma, A.: A multivariate quantile-matching bias correction approach with auto-and cross-dependence across multiple time scales: Implications for downscaling. J. Climate, 29, 3519–3539, doi:10.1175/JCLI-D-15-0356.1, 2016.

Piani, C. and Haerter, J. O.: Two dimensional bias correction of temperature and precipitation copulas in climate models, Geophys. Res. Lett., 39, L20401, doi:10.1029/2012GL053839, 2012.

320    Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L. (Eds): Climate Change 2007: The Physical Science Basis Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Teng, J., Potter, N. J., Chiew, F. H. S., Zhang, L., Wang, B., Vaze, J. and Evans, J. P.: How does bias correction of regional 325    climate model precipitation affect modelled runoff?, Hydrol. Earth Syst. Sc., 19, 711–728, doi:10.5194/hess-19-711-2015, 2015.

Teutschbein, C. and Seibert, J.: Bias correction of regional climatemodel simulations for hydrological climate-change impact studies: Review and evaluation of different methods, J. Hydrol., 456, 12–29, doi:10.1016/j.jhydrol.2012.05.052, 2012.

Wilks, D. S.: Statistical methods in the atmospheric science, 3rd edn. Academic, Amsterdam, 2011.