

## Assessment of robustness and significance of climate change signals for an ensemble of distribution-based scaled climate projections

Seaby, Lauren Paige; Refsgaard, Jens Christian; Sonnenborg, Torben; Stisen, Simon; Christensen, Jens Hesselbjerg; Jensen, Karsten Høgh

*Published in:*  
Journal of Hydrology

*DOI:*  
[10.1016/j.jhydrol.2013.02.015](https://doi.org/10.1016/j.jhydrol.2013.02.015)

*Publication date:*  
2013

*Document Version*  
Peer reviewed version

### *Citation for published version (APA):*

Seaby, L. P., Refsgaard, J. C., Sonnenborg, T., Stisen, S., Christensen, J. H., & Jensen, K. H. (2013). Assessment of robustness and significance of climate change signals for an ensemble of distribution-based scaled climate projections. *Journal of Hydrology*, 486, 479-493. <https://doi.org/10.1016/j.jhydrol.2013.02.015>

### **General rights**

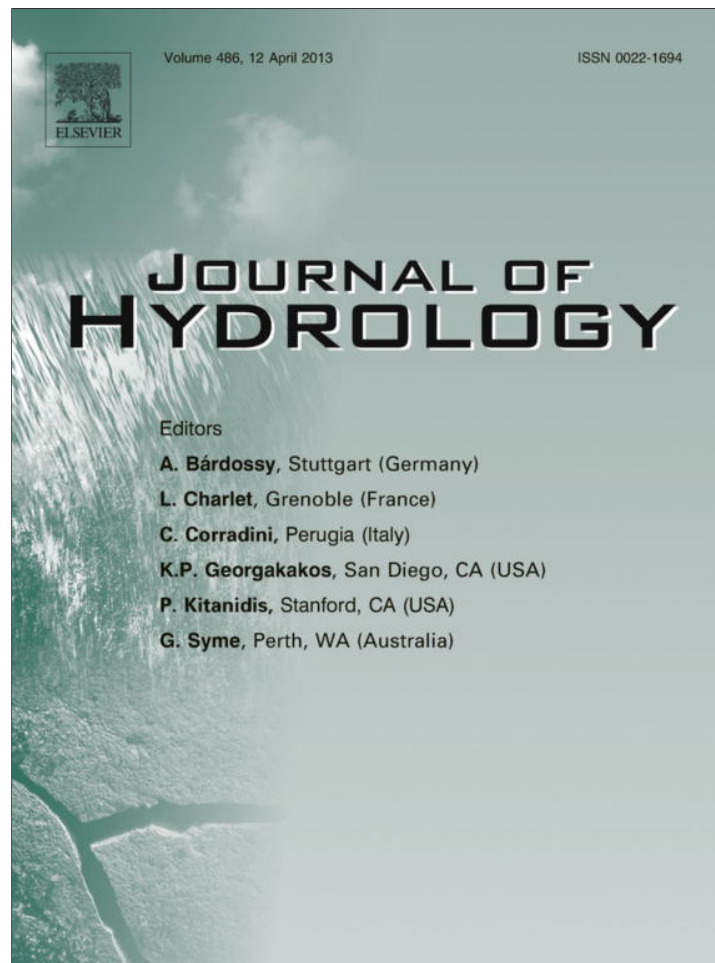
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

### **Take down policy**

If you believe that this document breaches copyright please contact [rucforsk@ruc.dk](mailto:rucforsk@ruc.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>



Contents lists available at SciVerse ScienceDirect

## Journal of Hydrology

journal homepage: [www.elsevier.com/locate/jhydrol](http://www.elsevier.com/locate/jhydrol)

## Assessment of robustness and significance of climate change signals for an ensemble of distribution-based scaled climate projections

L.P. Seaby<sup>a,b,\*</sup>, J.C. Refsgaard<sup>a</sup>, T.O. Sonnenborg<sup>a</sup>, S. Stisen<sup>a</sup>, J.H. Christensen<sup>c</sup>, K.H. Jensen<sup>b</sup>

<sup>a</sup> Geological Survey of Denmark and Greenland, Department of Hydrology, Øster Voldgade 10, 1350 Copenhagen K, Denmark

<sup>b</sup> University of Copenhagen, Institute of Geology and Geography, Øster Voldgade 10, 1350 Copenhagen K, Denmark

<sup>c</sup> Danish Meteorological Institute, Lyngbyvej 100, 2100 Copenhagen Ø, Denmark

## ARTICLE INFO

## Article history:

Received 19 June 2012

Received in revised form 7 February 2013

Accepted 10 February 2013

Available online 20 February 2013

This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Ashish Sharma, Associate Editor

## Keywords:

Climate change

Bias correction

ENSEMBLES

Delta change

Distribution-based scaling

## SUMMARY

An ensemble of 11 regional climate model (RCM) projections are analysed for Denmark from a hydrological modelling inputs perspective. Two bias correction approaches are applied: a relatively simple monthly delta change (DC) method and a more complex daily distribution-based scaling (DBS) method. Differences in the strength and direction of climate change signals are compared across models and between bias correction methods, the statistical significance of climate change is tested as it evolves over the 21st century, and the impact of choice of reference and change period lengths is analysed as it relates to assumptions of stationary in current climate and change signals. Both DC and DBS methods are able to capture mean monthly and seasonal climate characteristics in temperature ( $T$ ), precipitation ( $P$ ), and potential evapotranspiration ( $ET_{pot}$ ). For  $P$ , which is comparatively more variable by day, the DC approach is insufficient at recreating projected regimes while the DBS correction method can transfer changes in the mean as well as the variance, improving the characterisation of temporal dynamics as well as heavy precipitation events. Climate change signals in the near-future (2011–2040) are hidden by natural variability and are therefore not significant, in the mid-future (2041–2070) the significance of climate change signals depend on the choice of climate model, and in the far-future (2071–2100) climate change signals are strong across all models and variables. Some models already display significant differences in climate variables within the past timeframe for Denmark. Current climate characteristics are not necessarily stationary and the temporal positioning of a reference period might impact the magnitude of relative climate change. Reference and change period lengths over 15 years are adequate in size to overcome natural variability and still have stationarity in the climate change signal within the periods.

© 2013 Elsevier B.V. All rights reserved.

### 1. Introduction

Most climate models predict increases in annual precipitation and temperature for Denmark. Projections generally show winter precipitation increasing in mean and frequency, and reduced mean but higher intensity precipitation and increased evapotranspiration rates in summer. Globally, future changes in climate are expected to result in an intensification of hydrological conditions (Huntington, 2006; IPCC, 2007; Loaiciga et al., 1996). Future precipitation regimes will be sensitive to changes in atmospheric temperature and radiation balance, which influence the mechanisms and conditions for precipitation generation (i.e. available atmospheric moisture, atmospheric uplift and condensation) (Trenberth et al., 2003; Wild and Liepert, 2010). Temperature increases will

result in increased availability of atmospheric water vapour content from surface evaporation (increased specific humidity), which can decrease water availability but increase heavy precipitation events (Bates et al., 2008). These future climate dynamics lead to increased variability, namely more intense precipitation and more droughts. In Denmark, changes in the quantity, timing, and delivery of precipitation is expected to result in higher rates of groundwater recharge, flooding, and wetland inundation in the winter months, and decreased water tables, dry root zones, and depletion of low flows in the summer months (van Roosmalen et al., 2009). However, there is variability between climate models on the direction and strength of the climate change signals and characteristics of future climate dynamics.

General Circulation Models (GCMs) are numerical coupled models representing ocean–atmosphere circulation and can be used to project changes in atmospheric variables under scenarios of climate change (IPCC, 2007). GCMs effectively capture large-scale climate features and sub-continental patterns (i.e. temperature and precipitation), but the coarse resolution and inability of GCMs to

\* Corresponding author at: Geological Survey of Denmark and Greenland, Department of Hydrology, Øster Voldgade 10, 1350 Copenhagen K, Denmark. Tel.: +45 3814 2768; fax: +45 3814 2050.

E-mail address: [lps@geus.dk](mailto:lps@geus.dk) (L.P. Seaby).

resolve sub-grid processes make their outputs unrepresentative of local climate change dynamics (Rummukainen, 2010) and therefore not suitable for direct forcing of a hydrological model. Regional climate models (RCMs) are simulations using lateral boundary conditions from GCMs over a limited area to produce higher resolution outputs (i.e. dynamically downscaled). Because of the GCM–RCM nesting, the overall quality of dynamically downscaled RCM output is tied to the realistic large-scale forcing of the underlying GCM and likewise affected by its biases (Xu et al., 2005). The PRUDENCE ensemble project (Christensen and Christensen, 2007) found the uncertainty introduced by the driving GCM to be greater than from the RCM or emissions scenario (Déqué et al., 2007). The proceeding EU project ENSEMBLES pairs multiple GCMs and RCMs to generate a matrix of transient climate change simulations for the European region (van der Linden and Mitchell, 2009). Compared to previous regional climate modelling projects, ENSEMBLES has a more extensive set of GCM–RCM pairings, longer transient runs to the middle or end of the 21st century, and they are resolved at a higher resolution from 25 to 50 km. A common RCM domain, model resolution, and set of output variables are defined making the results between climate models convenient to compare and ideal for use in impact studies (Christensen et al., 2009). Dynamically downscaled RCMs outputs are still subject to systematic errors and biases (Fowler et al., 2007; Jones et al., 2004b) and precipitation especially requires further downscaling and/or bias correction prior to use in hydrological simulations (Hansen et al., 2006; Sharma et al., 2007). A variety of downscaling and bias correction methods have been developed to estimate or scale future climate variables.

In this paper, we classify bias correction methods as direct or indirect in terms of their use of RCM outputs and linear or nonlinear in terms of their scaling procedures, where the nonlinear direct methods are sub-classified as parametric or nonparametric. A relatively straightforward and therefore commonly applied statistical downscaling approach is the delta change (DC) method, where mean change factors derived by comparing GCM or RCM past and future climate are perturbed onto a reference climate series (e.g. Fowler et al., 2007; Graham et al., 2007; IPCC, 2007). DC methods can be implemented at various temporal (e.g. daily, monthly, seasonal, annual) and spatial (e.g. grid, basin, national) scales, can be formulated to transfer relative or absolute changes (e.g. additive, multiplicative factors), and can be calculated as a single factor per variable or multiple magnitude dependent factors (Anandhi et al., 2011; Hay et al., 2000). DC methods are indirect since RCM outputs are not used directly and the perturbed change is usually linear, an exception being e.g. magnitude dependent factors which perturb change in a nonlinear way. By definition, DC methods preserve the climate dynamics of the observed reference period, producing locally realistic climate variables reflecting mean changes simulated by the climate models, but do not utilise or retain information on changes in precipitation dynamics simulated by RCMs in the future period (Lenderink et al., 2007). Since change factors are perturbed evenly over daily data, changes in variability and regime are not captured (i.e. frequency of wet and dry days, more frequent precipitation events of high intensities) (Fowler et al., 2007). Bias removal (BR), or linear scaling (Teutschbein and Seibert, 2012b), operates similar to DC methods except mean correction factors are found between observed and simulated values in a reference period and are perturbed onto future RCM simulated climate outputs, making it a direct, linear method. BR methods can be expanded to correct for variance bias with nonlinear power transformations (Leander and Buishand, 2007; Sunyer et al., 2012), in this case making a direct, nonlinear method. By definition, BR methods will generate corrected RCM outputs in the reference period that match the mean monthly values in the observations, and this relationship is assumed to hold under future conditions.

Distribution correction methods adjust the function of RCM outputs to match the function of observed climate variables by deriving a transfer function from observed and simulated cumulative distribution functions (CDFs). These methods fully utilise the RCM simulation's projected changes in precipitation regimes (i.e. mean, variability, frequency, and intensity) by bias correcting based on daily precipitation intensity, producing internally consistent time series that have the same statistical intensity distribution as the observations (Piani et al., 2010). After Yang et al. (2010), we refer to this method as distribution-based scaling (DBS), but these methods go by various names in the literature: distribution mapping (Teutschbein and Seibert, 2012b), statistical bias correction (Piani et al., 2010), quantile mapping (Johnson and Sharma, 2011; Li et al., 2010; Wood et al., 2004), and histogram equalisation (Rojas et al., 2011). DBS methods are direct, nonlinear, and parametric, and assume stationarity in climate model error over time. Precipitation ( $P$ ) intensity is well characterised by the asymmetrical gamma distribution (Piani et al., 2010; Yang et al., 2010) and temperature ( $T$ ) by a Gaussian (normal) distribution (Teutschbein and Seibert, 2012b; Yang et al., 2010) or a beta distribution (Li et al., 2010). van Roosmalen et al. (2011) fit potential evapotranspiration ( $ET_{pot}$ ) to a gamma distribution, but goodness of fit criteria were not reported.

The DBS method has been developed and documented for annual  $P$  over Europe (Piani et al., 2010), seasonal  $P$  and  $T$  over Sweden (Teutschbein and Seibert, 2012b; Yang et al., 2010), and seasonal  $P$ ,  $T$ , and  $ET_{pot}$  in Denmark, where previous work by van Roosmalen et al. (2011) made limited comparisons between the DBS and DC methods. We know of no other studies that have successfully fitted  $ET_{pot}$  to a distribution for bias correction purposes, so it remains unclear how well  $ET_{pot}$  can be bias corrected with a DBS approach. So far no studies have compared both methods over an ensemble of climate models through the end of the 21st century considering all three variables  $P$ ,  $T$ , and  $ET_{pot}$ , which are needed for many hydrological models. There is a need to benchmark the DBS method against the more simplistic DC method in terms of relative climate change and explore the uncertainty associated with choosing a bias correction method that directly uses RCM output series compared with the DC method of perturbing change onto observation series. It is important to continue to evaluate the DBS method in different regimes, on variables beyond  $P$  and  $T$ , and to develop methods for bias correcting  $P$ ,  $T$ , and  $ET_{pot}$  that retain daily covariation between the variables.

More sophisticated indirect statistical downscaling methods have been developed that utilise more information from the RCM simulations. For example, weather typing schemes use various classification techniques to link large-scale atmospheric variables to surface weather variables. Conditional probability distributions for observed statistics are associated with classes of atmospheric circulation patterns, and climate change is estimated by the change in frequency of the weather classes as simulated in the climate model (e.g. Bárdossy and Pegram, 2011). In contrast to weather classification techniques based on patterns, weather generators are stochastic models based on the occurrence of events, where daily climate is dictated by climate on previous days based on transition probabilities (e.g. Fowler et al., 2007). Weather typing schemes are parametric and weather generating methods are non-parametric, but both methods are indirect and nonlinear.

In a hydrologic modelling context, stationarity is usually discussed in terms of natural systems variability (e.g. annual streamflow, flood peaks) (Milley et al., 2008). Climate variables, both observed and simulated, show significant natural variability both on interannual and decadal time scales (especially precipitation), where climatological reference periods of 30 years are generally accepted to be adequate in length to encompass these natural cycles. Under RCM simulated climate the resultant change signal vs.

climate noise issue is often discussed because it can overwhelm long term trends of climate change (Bates et al., 2008). Though hydrologic change studies assume a non-stationary climate, the associated bias correction methods still rely on an assumption of stationarity in the climate change signal itself. The balance between climate change and climate variability (signal to noise ratio) projected by the RCMs governs the appropriate change period length. There is a need to characterise the impact of natural variability in RCM variables on the robustness of climate change signals and the ramifications on bias correction methods.

Historical data are often analysed for significance of climate change in the 20th century, whereas projections of future climates are often discussed in terms of magnitude and timing but not in terms of statistical significance (IPCC, 2007). It could be useful to evaluate climate models for statistically significant changes over the 21st century for purposes of model comparison and/or selection for impact modelling. There is a need to quantify variability and uncertainty on climate model inputs and between bias correction methods prior to use in an impact modelling methodology. Hydrological outputs simulated under climate change might compound the uncertainties within individual climate model predictions, between various climate models, and in the choice of bias correction and scaling methods.

The objectives of the present paper are (i) to assess the accuracy and robustness of the DBS bias correction method by evaluating it for an ensemble of climate models over multiple model domains; (ii) to analyse to which extent projected climate changes are actually significantly different from the current climate; and (iii) to evaluate reference and change period lengths to ensure distinguishing climate change from natural variability.

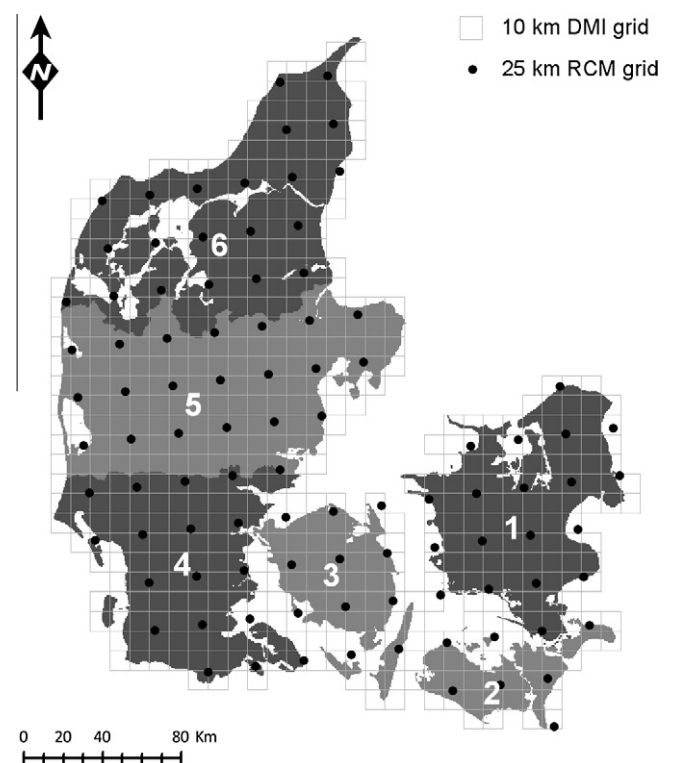
## 2. Data and methods

### 2.1. Study area

Multiple projections of future climate change are bias corrected for Denmark, spatially distributed according to six hydrologically distinct model domains as defined by the Danish national water resource model (DK-model) (Højberg et al., 2013; Henriksen et al., 2003). The DK-model delineates seven model domains (DK1–7) covering all of Denmark (approx. 43,000 km<sup>2</sup>), with boundaries defined by the sea or topographical divides coinciding with natural hydraulic conditions (Fig. 1). DK1 is a single domain over the island Sjælland (7163 km<sup>2</sup>) and including Copenhagen, DK2 covers the Southern Islands (2042 km<sup>2</sup>), DK3 covers the island Fyn (3473 km<sup>2</sup>), and DK4–6 cover the Jylland peninsula divided into South (7897 km<sup>2</sup>), Central (11,578 km<sup>2</sup>), and North (9934 km<sup>2</sup>) respectively. We exclude the 590 km<sup>2</sup> island Bornholm (DK7) from this climate model bias correction effort due to the high proportion of ocean to land in the grid cells covering the small island, making the climate model outputs more representative for the surrounding ocean than for the island itself.

### 2.2. Observed climate data and processing

The Danish Meteorological Institute (DMI) provide the daily gridded climate variables  $P$  [mm/day],  $T$  [°C], and  $ET_{pot}$  [mm/day] for the 20 year period 1991–2010. Observed  $P$  data is gridded to 10 km through an interpolation method of the approximately 500 rain gauges distributed evenly throughout Denmark, while  $T$  and  $ET_{pot}$  are gridded to 20 km based on a sparser network of climate stations (Scharling, 2000).  $ET_{pot}$  was estimated following the Makink equation (Makink, 1957) and negative values (i.e. dew) were adjusted to zero by DMI (Scharling, 2001). It is important to correct  $P$  prior to use in hydrological impact studies to



**Fig. 1.** The study area (Denmark) as divided by the DK-model into six hydrological domains; the 10 km DMI grid and centroids from the 25 km ENSEMBLES RCM grid are shown.

compensate for gauge under catch due to aerodynamic effects and wetting losses. We use corrected precipitation data calculated by Stisen et al. (2012) with daily correction factors for each 10 km grid based on daily observations of wind speed and air temperature (Allerup et al., 1997). To achieve a uniform grid for all climate variables, we further interpolate  $T$  and  $ET_{pot}$  to the same 10 km grid as  $P$  using an inverse weighted distance interpolation method (Shepard, 1967). Fig. 1 illustrates the DMI 10 km grid in comparison to the ENSEMBLES RCM 25 km grid (as grid centroid points) in relation to the six DK-model domains.

### 2.3. Climate model data and processing

A subset of climate models (GCM–RCM pairings) from the ENSEMBLES matrix are selected based on the following criteria for maximum comparability: highest resolution (25 km), longest simulations to the end of the 21st century covering approximately 1951–2100 (some start in 1961 and/or end in 2099), and consistent climate sensitivity to atmospheric CO<sub>2</sub> (excludes the low and high sensitivity Hadley GCMs) (Table 1). To explore climate model uncertainty as fully as possible, we include all eligible models, and the resultant subset of the ENSEMBLES matrix represents four GCMs and eight RCMs in 11 model pairings (Table 2).

Direct RCM output variables were downloaded from the ENSEMBLES data portal, including  $P$  [mm/day] and  $T$  [K] at 2 m above ground, and the variables needed for calculating  $ET_{pot}$  (i.e. temperature minimum and maximum, incoming long and short wave solar radiation, relative humidity, and wind speed) on a 25 km grid over a common European region. Post-processing of RCM climate outputs was as follows: (1) estimate  $ET_{pot}$  from RCM outputs, (2) correct  $P$  for RCM wet day bias, and (3) interpolate  $P$ ,  $T$ , and  $ET_{pot}$  to 10 km DMI grid.

Actual evapotranspiration ( $ET_{act}$ ) is a direct RCM output, but the simplified representation of land–surface processes makes  $ET_{act}$



**Table 1**  
Climate models from the ENSEMBLES project for which projections have been used in the present study.

Climate model	Organisation	Reference
<i>GCM</i>		
HadCM3	Met Office Hadley Center (METO-HC), UK	Collins et al. (2006)
ECHAM5	Max Planck Institute for Meteorology (MPI), Germany	Roeckner et al. (2003)
ARPEGE	National Centre of Meteorological Research (CNRM), France	Déqué et al. (1994)
BCM2	Bjerknes Centre for Climate Research (BCCR), Norway	Otterå et al. (2009)
<i>RCM</i>		
HadRM3	Met Office Hadley Center (METO-HC), UK	Jones et al. (2004)
REMO	Max Planck Institute for Meteorology (MPI)	Jacob et al. (2001)
RM5.1	National Centre of Meteorological Research (CNRM), France	Radu et al. (2008)
HIRHAM5	Danish Meteorological Institute (DMI)	Christensen and Christensen (2007)
CLM	Swiss Federal Institute of Technology Zurich (ETHZ)	Böhm et al. (2006)
RACMO2	Royal Netherlands Meteorological Institute (KNMI)	van Meijgaard et al. (2008)
RegCM3	International Centre for Theoretical Physics (ICTP), Italy	Pal et al. (2007)
RCA3	Swedish Meteorological and Hydrological Institute (SMHI)	Kjellström et al. (2005)

**Table 2**  
Matrix of ENSEMBLES climate models shown as GCM–RCM pairings.

RCM	GCM			
	HadCM3	ECHAM5	ARPEGE	BCM2
HadRM3	X			
REMO		X		
RM5.1			X	
HIRHAM5		X	X	X
CLM	X			
RAC MO2		X		
RegCM3		X		
RCA3		X		X

values inadequate as hydrological modelling inputs. The processes effecting  $ET_{act}$  (i.e. relative humidity, temperature, solar radiation, wind speed) are simulated at more realistic scales, therefore, it is common practice to estimate  $ET_{pot}$  using empirical formulas and output variables from the RCMs (van Roosmalen et al., 2009; Ekström et al., 2007). We calculate  $ET_{pot}$  following an adapted Penman–Monteith equation developed by the Food and Agriculture Organization of the United Nations (Allen et al., 1998):

$$ET_{pot} = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273.15} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

where  $ET_{pot}$  is potential evapotranspiration (mm/day),  $R_n$  is net radiation at the crop surface ( $\text{MJ m}^{-2} \text{d}^{-1}$ ),  $G$  is soil heat flux density ( $\text{MJ m}^{-2} \text{d}^{-1}$ ),  $u_2$  is wind speed at 2 m height ( $\text{m s}^{-1}$ ),  $e_s - e_a$  is saturation vapour pressure deficit (kPa),  $\Delta$  is slope vapour pressure curve ( $\text{kPa } ^\circ\text{C}^{-1}$ ), and  $\gamma$  is the psychrometric constant ( $\text{kPa } ^\circ\text{C}^{-1}$ ). The meteorological variables are assumed to be measured at 2 m, which is the approximate height of RCM variables with the exception of wind speed at 10 m. A conversion factor based on the logarithmic wind profile is applied to adjust wind speed to 2 m (Allen et al., 1998).  $R_n$  is calculated using RCM predictions of net incoming short and long wave radiation. The mean saturation vapour pressure ( $e_s$ ) and the slope of the vapour pressure curve are calculated using daily mean temperature, which is calculated as the average of daily maximum and minimum temperature. The actual vapour pressure ( $e_a$ ) is derived from relative humidity. Soil heat flux ( $G$ ) is small compared to  $R_n$  (Allen et al., 1998) and is therefore set to zero. For consistency with DMI's calculated  $ET_{pot}$ , all negative values (i.e. dew) of RCM estimated  $ET_{pot}$  are set to zero.

It is well established that RCMs have a systematic wet day bias, or a so-called drizzle effect, resulting in excessive low intensity precipitation on a high number of days (e.g. Gutowski et al., 2007). In the reference period 1991–2010, 46% of days are dry in Denmark annually, with winter having the lowest frequency of

39% and spring having the highest frequency at 54%. The 11 ENSEMBLES RCMs average 15% dry days annually. While these small daily amounts of drizzle may have negligible effects on mean monthly and total annual precipitation using e.g. DC, bias correction methods based on the distribution and frequency of wet days (e.g. DBS) would be affected by this climate model error. It is common practice to correct this wet bias so the frequency of dry days in the climate model reference period is equivalent to the frequency in the observations (e.g. Yang et al., 2010). Future data are corrected using the cut-off value obtained in the reference period (Table 3), which can be considered the value above which modelled precipitation is realistic, and below which values are erroneous drizzle and therefore set to zero. This allows for correcting the wet bias while preserving potential changes in the frequency of dry days in modelled future precipitation regimes. Across all climate models, domains, and seasons, dry day cut-off values range from 0.08 to 1.63 mm and average 0.67 mm. The correction is applied at a temporal scale (i.e. seasonal) and spatial scale (i.e. regional) that is consistent with the application of the DBS bias correction to ensure the functions can train properly in the reference period.

It is essential to have all climate variables on a uniform grid for the forthcoming scaling methods, and we choose to interpolate down to 10 km to preserve the higher resolution observation data. The same inverse weighted distance interpolation method is applied to the RCM climate variables ( $P$ ,  $T$ , and  $ET_{pot}$ ). This effectively subsets the European dataset to the grids over Denmark while interpolating the 25 km ENSEMBLES grid to the 10 km DMI grid.

#### 2.4. Bias correction methods

Assuming RCMs reliably simulate relative changes and dynamics rather than absolute values, we explore two methods of bias correction to overcome the inadequacy of the direct data for use in assessing hydrological impacts: a monthly DC approach and a seasonal DBS method. The RCM transient simulations are divided into past (1951–2010) and future (2011–2100) periods, from which we use a reference period (1991–2010) to coincide with the observations. Most of the climate models (8 out of 11) show a significant positive trend in precipitation over the 90 year future period. At 30 year time scales this trend was no longer significant, therefore, we assume stationarity in the climate change signal within the following designated future periods: near (2011–2040), mid (2041–2070), and far (2071–2100). For the DC approach, change factors for each future period are calculated. The DBS method is applied to the entire transient simulation (i.e. past and future periods) at once, but the three future time periods are still used in analysing the results to see how climate change

**Table 3**

Seasonal dry day correction cut-off values (mm/day) estimated from the reference period (1991–2010) and applied prior to downscaling RCM precipitation data (1951–2100).

Season	DK-domain	Obs freq. dry days (%)	ARPEGE-RM5.1	ARPEGE-HIRHAM5	BCM2-HIRHAM5	BCM2-RCA3	ECHAM5-HIRHAM5	ECHAM5-RegCM3	ECHAM5-RACMO2	ECHAM5-REMO	ECHAM5-RCA3	HadCM3-CLM	HadCM3-HadRM3
Winter	1	42	0.53	0.76	0.98	1.02	1.06	0.55	0.62	0.80	0.69	0.58	0.39
	2	42	0.60	0.80	1.04	1.00	1.01	0.54	0.64	0.80	0.76	0.67	0.49
	3	40	0.49	0.84	0.99	0.91	1.08	0.52	0.62	0.81	0.68	0.62	0.42
	4	35	0.38	0.55	0.69	0.77	0.83	0.44	0.53	0.55	0.57	0.27	0.24
	5	37	0.41	0.55	0.69	0.79	0.86	0.48	0.55	0.54	0.61	0.32	0.25
	6	36	0.42	0.56	0.71	0.77	0.85	0.38	0.52	0.66	0.49	0.35	0.22
Spring	1	56	1.30	0.61	0.67	0.92	0.90	0.84	0.57	0.53	0.92	0.40	0.52
	2	56	0.99	0.58	0.68	0.90	0.85	0.78	0.55	0.49	0.91	0.40	0.46
	3	55	1.25	0.62	0.71	1.00	1.01	0.72	0.57	0.57	0.94	0.41	0.52
	4	51	1.45	0.48	0.56	0.98	0.80	0.68	0.58	0.58	0.86	0.35	0.51
	5	53	1.60	0.53	0.60	1.10	0.81	0.73	0.64	0.59	0.92	0.42	0.58
	6	54	1.31	0.62	0.66	1.05	0.94	0.67	0.60	0.57	0.76	0.42	0.51
Summer	1	51	0.74	0.12	0.26	0.46	1.30	0.78	0.57	0.90	1.40	0.09	0.22
	2	53	0.56	0.11	0.28	0.55	1.51	0.92	0.69	0.81	1.53	0.08	0.21
	3	51	0.69	0.18	0.38	0.56	1.63	0.80	0.59	0.90	1.54	0.09	0.24
	4	47	0.78	0.09	0.21	0.50	1.11	0.77	0.66	0.84	1.39	0.10	0.24
	5	49	0.96	0.08	0.17	0.54	1.01	0.75	0.64	0.84	1.39	0.14	0.27
	6	52	0.90	0.15	0.28	0.60	1.31	0.67	0.74	0.70	1.31	0.16	0.30
Autumn	1	44	0.53	0.55	0.92	1.03	1.43	0.57	0.71	1.08	1.15	0.36	0.25
	2	45	0.58	0.56	0.98	0.99	1.57	0.57	0.75	1.26	1.24	0.50	0.32
	3	42	0.48	0.65	1.02	0.96	1.60	0.57	0.62	0.99	1.03	0.36	0.23
	4	36	0.36	0.25	0.52	0.89	0.94	0.53	0.64	0.58	0.91	0.16	0.16
	5	38	0.37	0.22	0.43	0.94	0.81	0.54	0.58	0.52	0.91	0.17	0.17
	6	38	0.41	0.31	0.50	0.98	0.96	0.42	0.62	0.71	0.88	0.18	0.16

evolves throughout the 21st century. The DC and DBS methods are used mainly for bias correction purposes, though there are elements of downscaling as we go from 25 km RCM output to 10 km grid scales.

2.4.1. Delta change (DC)

The DC method consists of altering an observed (reference) climate series with change factors to obtain a new series representative of future change. For state variables (e.g.  $T$ ) absolute change is applied, whereas for flux variables (e.g.  $P$  and  $ET_{pot}$ ) relative change factors are applied. Monthly change factors are derived and perturbed as follows for day and month ( $i, j$ ), where  $i = 1, 2, \dots, 31$  and  $j = 1, 2, \dots, 12$ :

$$P_A(i, j) = \Delta P(j) * P_{obs}(i, j); \quad \Delta P(j) = \frac{\bar{P}_{fut}(j)}{\bar{P}_{ref}(j)} \quad (2)$$

$$ET_{potA}(i, j) = \Delta ET_{pot,obs}(j) * ET_{pot,obs}(i, j); \quad \Delta ET_{pot}(j) = \frac{\overline{ET}_{pot,fut}(j)}{\overline{ET}_{pot,ref}(j)} \quad (3)$$

$$T_A(i, j) = T_{obs}(i, j) + \Delta T(j); \quad \Delta T(j) = \bar{T}_{fut}(j) - \bar{T}_{ref}(j) \quad (4)$$

where  $P_A, ET_{potA}, T_A$  are DC perturbed daily climate change variables,  $P_{obs}, ET_{pot,obs}, T_{obs}$  are observed climate variables in the reference period,  $\Delta P, \Delta ET_{pot}, \Delta T$  are the changes in climate as simulated by the RCMs,  $\bar{P}, \overline{ET}_{pot}, \bar{T}(j)$  are daily climate means by month, the index  $ref$  indicates the reference period, and  $fut$  indicates a future period. Here, DC values for  $P, T$ , and  $ET_{pot}$  are determined from the 11 RCM simulations' output, for the three future time periods, with unique sets of monthly (12) factors spatially averaged for each of the six DK-model domains. These factors are then used to adjust the observed daily climate variables within the individual months to a climate input series representing mean change of the respective future period (near, mid, far).

2.4.2. Distribution based scaling (DBS)

Unlike the DC approach, the DBS bias correction method directly operates on RCM output in a manner that preserves the

statistical distribution of the observed precipitation in the reference period. The application follows these steps: (1) group and prepare data for scaling, (2) fit probability distributions to observed and RCM precipitation in the reference period, (3) scale RCM precipitation for entire simulation period, and (4) perform a separate error bias method on  $ET_{pot}$  and  $T$ .

2.4.2.1. DBS step 1: Group and prepare data for scaling. The DBS method can be applied at various temporal scales (e.g. annually, seasonally, monthly) and spatial scales (e.g. nationwide, regional, single grid) which decide how the data should be grouped for fitting distribution parameters and the subsequent scaling. Ideally, the temporal scale should be short enough to preserve intra-annual climate characteristics (e.g. seasonality) and long enough to allow for potential shifts in future regimes (e.g. monthly changes). Likewise, the spatial scale should be fine enough to retain spatial heterogeneities in the RCM projections (e.g. local climate changes) but large enough to be robust and meaningful when fit to a distribution (e.g. more grids pooled). To retain spatial variation in the RCM outputs at the basin scale, DBS parameters are fit for each DK-model domain covering Denmark (DK1–6), and to maintain the seasonal characteristics, we further divide RCM and observed data into four seasons: winter (December–February), spring (March–May), summer (June–August), and autumn (September–November). For  $P$ , a subset is made of the wet days per season which are then sorted in ascending order and divided into two groups split the 95th percentile. From here on we refer to these groups as  $95^{upper}$  and  $95^{lower}$ , and the reasoning behind this grouping is explained below.

2.4.2.2. DBS step 2: Fit a probability distribution to observed and RCM precipitation for rainy days. Given the physical constraint of precipitation to be nonnegative, the statistical distribution of daily intensities is characteristically asymmetric and positively skewed. Gamma distributions are commonly used to theoretically represent precipitation intensity distributions as they are bound on the left by zero and skewed to the right (Wilks, 2006). The gamma distribution is defined by two parameters, the shape parameter al-

pha ( $\alpha$ ) and the scale parameter beta ( $\beta$ ). The probability density function (PDF) of daily  $P$  (mm/day) is as follows:

$$f(P) = \frac{(P/\beta)^{\alpha-1} \exp(-P/\beta)}{\beta \Gamma(\alpha)}; \quad P, \alpha, \beta > 0 \quad (5)$$

where  $\Gamma(\alpha)$  is the gamma function:

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt \quad (6)$$

The parameters ( $\alpha, \beta$ ) are fit using the method of maximum likelihood for the shape parameter ( $\beta$ ) after which the scale parameter ( $\alpha$ ) is obtained (Wilks, 2006). The parameter set that maximises the log likelihood function is found through an iterative process using the multidimensional generalisation of the Newton–Raphson method (Press et al., 2007). To minimise the number of parameters in the scaling method and associated uncertainty, we initially attempt to fit a single gamma distribution to observed  $P$  in the reference period, per season and per DK-domain. A good fit between the observations and a gamma distribution is prerequisite, as the RCM scaling will be trained on this relationship, ultimately transferring the difference onto the RCM data outside of the reference period. Therefore, the match between the observations and the fitted gamma distribution was checked prior to moving onto the RCM fitting. With a single gamma approach mean seasonal  $P$  was estimated well (i.e. zero error) but daily intensity was not as indicated by consistently high error in variance (20–40%). In general, precipitation regimes are characterised by a high number of low intensity events, which can dominate the fitting of the distribution parameters. The high error in variance suggests that extreme high values (i.e. upper tail) could not be represented within a single gamma distribution. These findings are consistent with Yang et al. (2010), who also expanded to a double gamma distribution partitioned at the 95th percentile, effectively isolating normal precipitation from heavy precipitation. With a double gamma approach, distributions are fit to the 95<sup>upper</sup> and 95<sup>lower</sup> groups separately, resulting in two parameter sets per season and per DK-domain. Finally, the same approach is applied to  $P$  from the 11 RCM's outputs in the reference period. Table 4 provides the seasonal 95<sup>upper</sup> and 95<sup>lower</sup> gamma distribution parameters for the 11 RCMs spatially averaged over the six DK-domains.

**2.4.2.3. DBS step 3: Scale RCM data.** Though DBS parameters are found per basin and season, DBS scaling is done per grid on a daily basis. When scaling (=correcting) a daily  $P$  value simulated by a RCM,  $F_{RCM}$ , the probability of the precipitation not exceeding  $P_{RCM}$  in the RCM simulation in the reference period is found from the CDF for the RCM in the reference period. Locating this probability in the CDF for the observations provides the corresponding (i.e. scaled) daily  $P$  value following this method:

$$P_{corr} = F^{-1}(\alpha_{obs}, \beta_{obs}, F(\alpha_{ctrl}, \beta_{ctrl}, P_{RCM})) \quad (7)$$

where  $P_{corr}$  is the bias corrected RCM daily  $P$ ,  $F$  is the CDF of the gamma distribution (Eq. (5)), and  $F(\alpha_{ctrl}, \beta_{ctrl}, P_{RCM})$  is the probability of not exceeding the value  $P_{RCM}$ . We assume that model biases are the same in the reference, past, and future periods, therefore,  $P_{corr}$  can be any value in the entire transient RCM output. Fig. 2 graphically depicts the correction in Eq. (7) with  $y_1 = P_{corr}$  and  $x_1 = P_{RCM}$ .

To evaluate the effectiveness of the scaling method, a comparison of mean and variance between observed and RCM scaled  $P$  in the reference period is made. Across all domains in Denmark and for all climate models, error bias on the mean is within 2% and most are within 10% on the variance. The exceptions being high error variance in the summer months for ARPEGE-HIRHAM5 (+4% to +17%) and ECHAM5-HIRHAM5 (–4% to +11%).

**2.4.2.4. DBS step 4: Error bias correction for potential evapotranspiration and temperature.** In the DBS application outlined by Yang et al. (2010) a double gamma distribution is fit to  $P$  and a normal distribution is fit to  $T$ , with emphasis on how the method retains the daily covariation between the variables. However,  $ET_{pot}$  was not included in the Yang et al. (2010) study, and DBS scaling of  $ET_{pot}$  remains a challenge. We attempt to fit observed and simulated  $ET_{pot}$  on a seasonal basis and in each DK-domain initially to a single gamma distribution and then a double gamma distribution split at various percentiles (i.e. isolating the upper and lower tails), but find that  $ET_{pot}$  does not conform to a gamma distribution. The motivation of this bias correction exercise is to prepare climate inputs for hydrological modelling, therefore, it is important that  $P, T$  and  $ET_{pot}$  are realistic as a set of daily inputs. Though  $P$  and  $T$  covary on a daily basis,  $ET_{pot}$  and  $T$  are more strongly tied given that daily  $T$  is one of the variables used in the estimation of  $ET_{pot}$ . Therefore, it is important that  $ET_{pot}$  and  $T$  are bias corrected in an alternate way, consistent with each other to persevere their dependence, and with a direct method so the transient series can be used alongside scaled  $P$ . We use a standard monthly BR method to correct daily values, where bias factors are calculated between the RCM reference period and the observed data for each day  $i$  in month  $j$ :

$$\varepsilon_{ET_{pot}}(j) = \frac{\overline{ET_{pot,ref}}(j)}{\overline{ET_{pot,obs}}(j)}; \quad j = 1, 2, \dots, 12 \quad (8)$$

$$\varepsilon_T(j) = \overline{T}_{ref}(j) - \overline{T}_{obs}(j); \quad j = 1, 2, \dots, 12 \quad (9)$$

where  $\overline{ET_{pot,ref}}$  and  $\overline{T}_{ref}$  are mean daily RCM simulated outputs and  $\overline{ET_{pot,obs}}$  and  $\overline{T}_{obs}$  are mean daily observed values, each calculated as the mean of all days per month and per DK-domain, and  $\varepsilon_{ET_{pot}}$  and  $\varepsilon_T$  are the monthly biases, which are then applied to the RCM outputs to remove mean monthly biases:

$$ET_{pot,corr}(i,j) = ET_{pot,RCM}(i,j) * \varepsilon_{ET_{pot}}(j); \quad i = 1, 2, \dots, 31; j = 1, 2, \dots, 12 \quad (10)$$

$$T_{corr}(i,j) = T_{RCM}(i,j) - \varepsilon_T(j); \quad i = 1, 2, \dots, 31; j = 1, 2, \dots, 12 \quad (11)$$

where  $ET_{pot,corr}$  and  $T_{corr}$  are corrected RCM outputs and  $ET_{pot,RCM}$  and  $T_{RCM}$  are simulated RCM outputs, and  $\varepsilon_{ET_{pot}}$  is the monthly bias factor representing relative change for the flux variable  $ET_{pot}$  and  $\varepsilon_T$  is the absolute change for the state variable  $T$ .

### 3. Results

#### 3.1. DC factors for Denmark

Across Europe, the ENSEMBLES RCMs project increases in  $T$  in all seasons and positive  $P$  changes in northern Europe. For Denmark, we calculate DC factors for the climate variables  $P, T$ , and  $ET_{pot}$  on a monthly basis for the near, mid, and far-future, and spatially averaged for the six DK-domains. DC values for net potential precipitation  $Net P_{pot}$  ( $P - ET_{pot}$ ) are calculated as an indicator of changes in both radiation balance and precipitation. Fig. 3 shows plots of monthly DC values for all of Denmark (average of six domains) in the far-future period per climate model. We include the ensemble mean DC factors calculated as the mean of all 11 factors from each RCM. Table 5 shows annual DC factors for all climate models calculated as the mean of all (12) monthly DC factors.

##### 3.1.1. Temperature

Across all models and months,  $T$  is projected to increase. The models driven by the HadCM3 GCM project the highest annual increases in  $T$  (more than +3.3 °C), while the smallest increases (less than +2 °C) are projected by the BCM2-RCA3 and ECHAM5-HIR-



**Table 4**  
Gamma distribution parameters with median value shown in bold and range across all six DK-model domains shown in parenthesis.

Climate series	Winter			Spring			Summer			Autumn						
	95 <sup>upper</sup>		95 <sup>lower</sup>	95 <sup>upper</sup>		95 <sup>lower</sup>	95 <sup>upper</sup>		95 <sup>lower</sup>	95 <sup>upper</sup>		95 <sup>lower</sup>				
	Alpha	Beta	Alpha	Beta	Alpha	Beta	Alpha	Beta	Alpha	Beta	Alpha	Beta				
Observed	<b>2.4</b> (2.5–2.5)	<b>7.8</b> (7.5–8.5)	<b>1.0</b> (1–1.1)	<b>3.1</b> (2.8–3.6)	<b>2.4</b> (2.5–2.5)	<b>7.3</b> (6.9–7.4)	<b>1.1</b> (1.2–1.2)	<b>2.7</b> (2.4–2.9)	<b>2.4</b> (2.5–2.5)	<b>10.3</b> (10.1–11.1)	<b>1.0</b> (1–1.1)	<b>3.7</b> (3.4–4)	<b>2.4</b> (2.5–2.5)	<b>9.0</b> (7.9–9.6)	<b>1.0</b> (1–1.1)	<b>3.5</b> (3–4.2)
ARPEGE-RM5.1	<b>4.8</b> (2.5–14.4)	<b>4.1</b> (1.2–7)	<b>1.5</b> (1.3–1.9)	<b>2.3</b> (1.8–2.7)	<b>2.4</b> (2.5–10.9)	<b>6.2</b> (1.6–6.8)	<b>2.2</b> (2.2–2.3)	<b>1.8</b> (1.6–2)	<b>2.4</b> (2.4–2.5)	<b>8.2</b> (8.2–9.1)	<b>2.1</b> (2–2.2)	<b>1.6</b> (1.6–1.7)	<b>2.4</b> (2.5–2.5)	<b>7.5</b> (7.4–8.1)	<b>1.4</b> (1.3–1.8)	<b>2.2</b> (2.1–2.8)
ARPEGE-HIRHAM5	<b>2.4</b> (2.5–8.9)	<b>7.6</b> (2.1–8.4)	<b>1.6</b> (1.4–2)	<b>2.4</b> (2.1–3.2)	<b>2.4</b> (2.5–2.5)	<b>7.2</b> (6.9–7.6)	<b>1.7</b> (1.5–1.8)	<b>2.0</b> (1.9–2.4)	<b>2.4</b> (2.4–2.5)	<b>7.5</b> (7.1–8.1)	<b>0.9</b> (0.9–1)	<b>2.0</b> (1.9–2.5)	<b>2.4</b> (2.5–2.5)	<b>9.4</b> (9.2–10.1)	<b>1.2</b> (1–1.6)	<b>3.3</b> (2.9–4)
BCM2-HIRHAM5	<b>2.4</b> (2.5–6.3)	<b>6.3</b> (2.5–6.9)	<b>2.0</b> (1.9–2.2)	<b>1.8</b> (1.8–2)	<b>2.4</b> (2.5–2.5)	<b>7.0</b> (6.5–7.3)	<b>1.8</b> (1.7–2)	<b>1.8</b> (1.7–2.1)	<b>2.4</b> (2.5–2.5)	<b>8.3</b> (8.1–8.5)	<b>1.0</b> (1–1.2)	<b>2.7</b> (2.6–3.2)	<b>2.4</b> (2.5–2.5)	<b>8.9</b> (8.2–9)	<b>1.7</b> (1.2–2.1)	<b>2.7</b> (2.3–3.2)
BCM2-RCA3	<b>5.9</b> (2.5–15.2)	<b>3.0</b> (1–5.7)	<b>2.1</b> (2.1–2.3)	<b>1.8</b> (1.6–2)	<b>2.4</b> (2.5–5.6)	<b>6.7</b> (2.9–6.9)	<b>2.2</b> (2.2–2.2)	<b>1.8</b> (1.6–1.9)	<b>2.4</b> (2.5–2.5)	<b>9.4</b> (9.1–9.9)	<b>1.4</b> (1.3–1.5)	<b>2.4</b> (2.3–2.8)	<b>2.4</b> (2.5–2.5)	<b>7.5</b> (7–8.4)	<b>2.1</b> (2.1–2.2)	<b>2.0</b> (1.8–2.3)
ECHAM5-HIRHAM5	<b>2.4</b> (2.5–2.5)	<b>7.4</b> (6.9–8)	<b>2.0</b> (1.9–2.2)	<b>2.1</b> (2–2.4)	<b>2.4</b> (2.5–3.9)	<b>7.8</b> (4.9–8.8)	<b>2.0</b> (1.9–2.1)	<b>2.1</b> (2–2.4)	<b>2.4</b> (2.5–2.5)	<b>10.0</b> (9.7–11.1)	<b>2.0</b> (2–2.3)	<b>2.7</b> (2.7–3.1)	<b>2.4</b> (2.5–4.5)	<b>10.1</b> (5.5–10.6)	<b>2.0</b> (1.8–2.2)	<b>3.0</b> (2.9–3.3)
ECHAM5-RegCM3	<b>11.5</b> (2.5–14.8)	<b>1.4</b> (1.1–7.5)	<b>1.7</b> (1.5–1.9)	<b>2.2</b> (2–2.8)	<b>2.4</b> (2.5–2.5)	<b>7.4</b> (7.3–7.8)	<b>1.9</b> (1.9–2.1)	<b>2.0</b> (1.9–2.3)	<b>2.4</b> (2.4–2.5)	<b>9.8</b> (9–10.3)	<b>1.9</b> (1.8–2.1)	<b>2.5</b> (2.2–2.8)	<b>5.1</b> (2.5–10.5)	<b>3.9</b> (2.1–9)	<b>1.5</b> (1.4–1.7)	<b>2.9</b> (2.6–3.4)
ECHAM5-RACMO2	<b>2.4</b> (2.5–14.1)	<b>6.3</b> (1.2–7.5)	<b>1.7</b> (1.5–2)	<b>2.1</b> (1.7–2.6)	<b>2.4</b> (2.5–2.5)	<b>6.6</b> (6.1–7)	<b>1.8</b> (1.8–1.9)	<b>1.8</b> (1.8–2.1)	<b>2.4</b> (2.5–2.5)	<b>8.8</b> (8.7–9.1)	<b>1.8</b> (1.7–1.9)	<b>2.5</b> (2.3–2.9)	<b>2.5</b> (2.5–12.8)	<b>8.0</b> (1.8–8.2)	<b>1.6</b> (1.5–1.9)	<b>2.7</b> (2.4–3.4)
ECHAM5-REMO	<b>4.8</b> (2.5–14.7)	<b>3.7</b> (1.2–7.7)	<b>1.9</b> (1.5–2.1)	<b>1.9</b> (1.6–2.5)	<b>2.4</b> (2.5–3.7)	<b>6.2</b> (3.9–6.6)	<b>1.8</b> (1.7–1.9)	<b>1.7</b> (1.6–2)	<b>2.4</b> (2.5–2.5)	<b>9.5</b> (8.7–10)	<b>1.9</b> (1.8–2)	<b>2.5</b> (2.5–2.8)	<b>2.5</b> (2.5–9.7)	<b>8.7</b> (2.3–10)	<b>1.8</b> (1.4–2.2)	<b>2.7</b> (2.5–3.6)
ECHAM5-RCA3	<b>2.5</b> (2.5–7.5)	<b>6.1</b> (1.9–7.1)	<b>1.9</b> (1.7–2.1)	<b>1.9</b> (1.7–2.2)	<b>2.4</b> (2.5–2.5)	<b>6.7</b> (6.3–7.1)	<b>2.1</b> (2.1–2.2)	<b>1.7</b> (1.6–1.9)	<b>2.4</b> (2.4–2.5)	<b>9.4</b> (8.8–9.7)	<b>2.2</b> (2.2–2.3)	<b>2.3</b> (2.1–2.7)	<b>3.9</b> (2.5–9.3)	<b>5.5</b> (2.1–8.3)	<b>2.1</b> (2–2.2)	<b>2.2</b> (2.1–2.4)
HadCM3-CLM	<b>2.4</b> (2.5–2.5)	<b>6.3</b> (6–6.8)	<b>1.3</b> (1.2–1.9)	<b>2.2</b> (1.6–2.6)	<b>2.4</b> (2.5–2.5)	<b>7.5</b> (7.2–8.3)	<b>1.7</b> (1.6–1.8)	<b>1.9</b> (1.8–2.2)	<b>2.4</b> (2.4–2.5)	<b>10.5</b> (9.4–12.7)	<b>1.0</b> (1–1.2)	<b>3.0</b> (3–3.2)	<b>2.4</b> (2.5–2.5)	<b>7.8</b> (7.6–8.3)	<b>1.0</b> (1–1.2)	<b>2.8</b> (2.6–3)
HadCM3-HadRM3	<b>2.4</b> (2.5–2.6)	<b>6.6</b> (6–7.2)	<b>1.6</b> (1.7–2.1)	<b>2.1</b> (1.7–2.8)	<b>2.4</b> (2.5–2.5)	<b>6.9</b> (6.6–7.4)	<b>1.5</b> (1.4–1.7)	<b>2.1</b> (1.9–2.4)	<b>2.4</b> (2.4–2.5)	<b>10.1</b> (10–11.7)	<b>0.9</b> (0.9–1)	<b>3.1</b> (2.8–3.5)	<b>2.4</b> (2.5–2.5)	<b>8.8</b> (8.5–9.2)	<b>1.1</b> (1–1.7)	<b>3.2</b> (2.4–3.7)

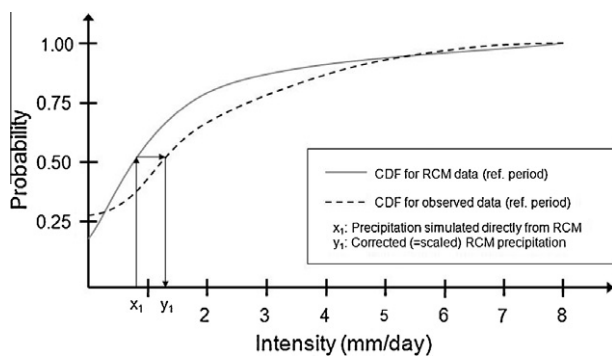


Fig. 2. Graphical illustration of the DBS correction method on daily precipitation.

HAM5 models. The rest of the models show relatively small differences among GCM–RCM pairings, with annual increases between +2.02 and +2.29 °C by 2100. For nine of the 11 models annual average  $T$  increases are below the ensemble average increase of +2.34 °C. The ensemble shows the highest  $T$  increases and largest spread (i.e. variability among models) in November, December, and January, and the lowest increases in  $T$  and least spread in May, June, and July.

### 3.1.2. Precipitation

Most models project increases in  $P$  ranging between +6% and +19% annually. The two ARPEGE GCM driven models show an opposite change signal, projecting decreases between –11% and –16%. The HadCM3–HadRM3 model projects relatively no change in annual  $P$  (–1%). The ensemble mean trend is of relative decreases in  $P$  from June to October (–2% to –15%) and increases from November to May (+4% to +24%). The two BCM2 GCM driven models deviate from the ensemble mean trend with relative increases in  $P$  projected nearly every month throughout the year. The two ARPEGE GCM driven models also deviate from the ensemble mean trend with most months projecting decreases in  $P$ . Most ECHAM5 GCM driven models project relatively moderate increases in annual  $P$ , and monthly changes follow the ensemble mean pattern of  $P$  increase in winter and decrease in summer. The ECHAM5–HIRHAM5 model projects the highest annual increase in  $P$  (+19%) and the ARPEGE–HIRHAM5 model projects the highest annual decrease in  $P$  (–16%). Though these models are paired with the same RCM (HIRHAM5), they represent the largest spread in the ensemble for relative changes in  $P$ , reiterating the importance of the GCM.

### 3.1.3. Potential evapotranspiration

All models project an annual increases in  $ET_{pot}$  ranging between +7% and +37%. The ensemble mean trend projects an increase in every month throughout the year. The greatest relative increases as well as variability in change factors (spread) within the ensemble is in the winter months, while the smallest increases and least variability in the change factors is in the summer months. The model pairings with the highest annual increases of  $ET_{pot}$ , HadCM3–HadRM3 (+37%) and ARPEGE–HIRHAM5 (+25%), are also at the higher end of the ensemble for projected increases in  $T$ . The ECHAM5 GCM driven models project the least annual increases in  $ET_{pot}$  (+7% to +14%), yet these models are varied in terms of projected  $P$  and  $T$  changes.

### 3.1.4. Net potential precipitation (mm/day)

Though  $Net P_{pot}$  represents a flux, we present it in absolute change to de-emphasise relatively large changes on small quantities (e.g. summer  $Net P_{pot}$ ). Most models in the ensemble project annual increases in  $Net P_{pot}$ , with a pattern of increase from

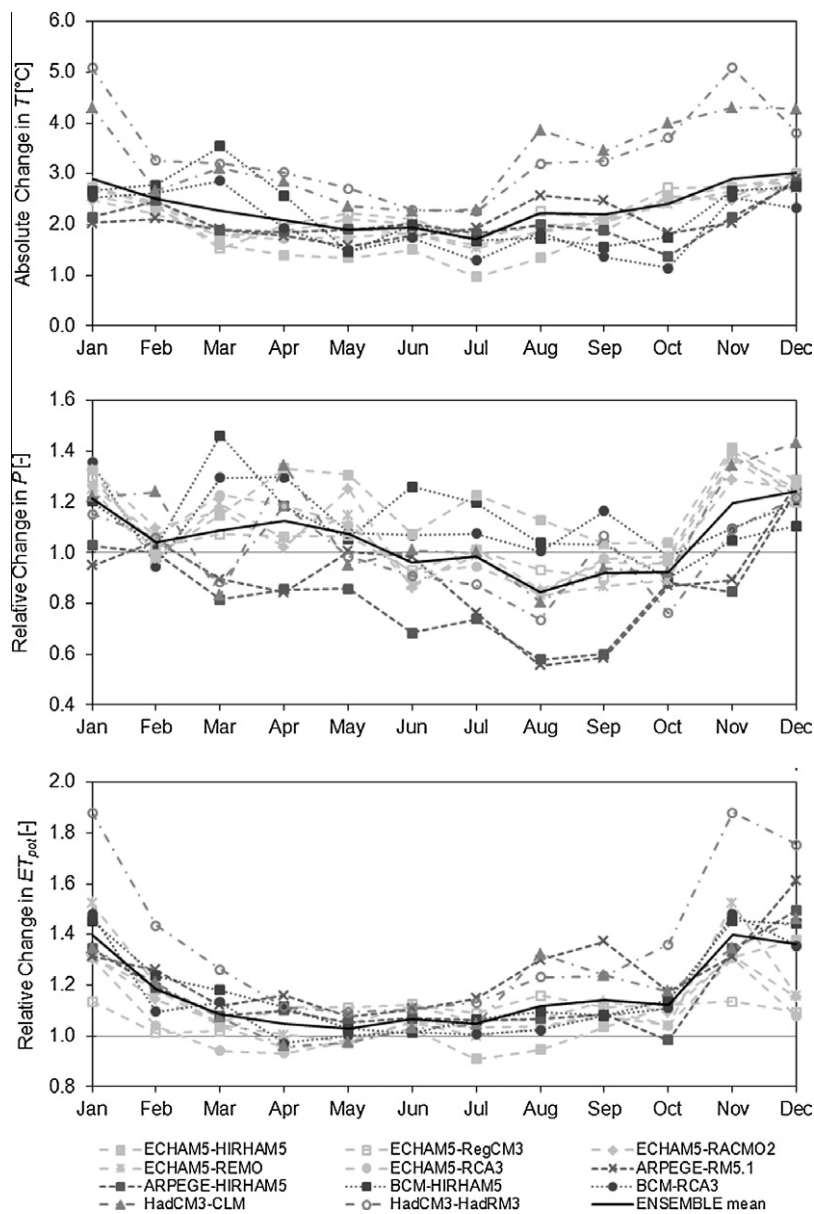
November to May and decrease from June to October, which closely follows projected patterns of  $P$  change. These patterns are especially clear in Fig. 4, which shows monthly far-future DC factors for  $Net P_{pot}$  for GCM ensembles within the climate model pairings (i.e. ARPEGE, BCM2, ECHAM5, and HadCM3 paired climate models) and the ensemble mean. The ARPEGE GCM driven models are clear outliers within the ensemble, projecting decreases in  $Net P_{pot}$  for every month except December. The BCM2 GCM driven models also deviate from the ensemble mean trend with increases in  $Net P_{pot}$  projected throughout most of the year. The HadCM2 and ECHAM5 GCM driven models follow the ensemble  $Net P_{pot}$  mean trend more closely, but there are still months with high variability among the models.

### 3.2. Validation of the DBS method

The DBS method has been calibrated to fit the probability distribution of the observed  $P$  for a 20 year reference period for four seasons over six domains. We calculate monthly DC factors with DBS corrected  $P$  to compare with DC factors derived from direct RCM output to evaluate how well the seasonal scaling retains the projected monthly change characteristics. The DC factors were essentially the same between the DBS corrected and RCM direct  $P$ . Despite being trained on seasonal data, monthly change characteristics are not compromised in the DBS correction. We further evaluate the DBS correction by testing to which extent it is able to reproduce other statistics of the observed data for which it has not directly been calibrated. We have made two tests with focus on annual mean  $P$  and extreme  $P$ , respectively.

In the first test we compare annual  $P$  for the entire country for the 20 year reference period and test to which extent the DBS method is able to reproduce the mean and the variance of the 20 annual values as in the observation data. To perform such validation test we first confirm by a chi-square test that the 20 annual mean  $P$  values, for each of the 11 models, can be assumed to be normally distributed. Next, we test the hypothesis that DBS scaled  $P$  (mm/year) is not significantly different than the observations in mean ( $t$ -test) and variance ( $f$ -test). Table 6 shows the significance level results, where the  $p$ -values indicate the level of confidence that mean and variance are not from different populations, along with annual statistics on the reference period. The results show that the hypothesis cannot be rejected for any of the models (i.e.  $p$ -values below 0.05) and that DBS scaled  $P$  from all 11 RCMs reproduce the annual mean and variance of the observations. The confidence levels, however, differ between models. For most models there is very high confidence that the means are from the same population (i.e.  $p$ -values above 0.95), while for the two HadCM3 GCM driven models there is lower confidence ( $p$ -values below 0.80). The significance levels for the test in variance are generally lower, which is not surprising as the calibration of the double gamma distributions focused on preserving the mean and variance over the entire 20 years reference period (Table 6) but did not consider the interannual variation.

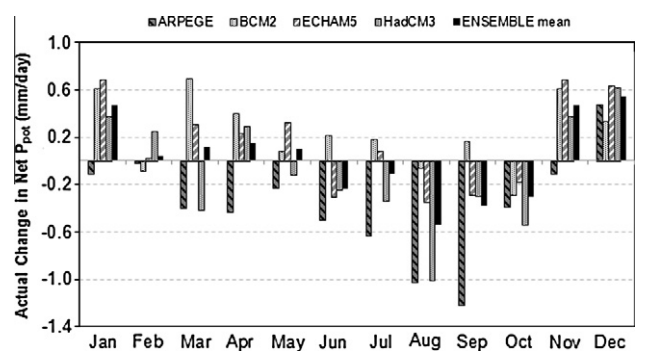
In the second test we analyse selected percentiles in the upper 10% tail of  $P$  to evaluate how well the DBS method reproduces the high intensity  $P$  in the observed dataset. Fig. 5 shows observed cumulative daily  $P$  for high intensity events at selected percentiles during the reference period with error bars indicating the range of values estimated by DBS  $P$  in all 11 models. These selected percentiles allows for analysis in the upper tail of the 95<sup>lower</sup> distribution (i.e. 90th percentile) and the full range of the 95<sup>upper</sup> distribution (i.e. 95–99.9th percentile). DBS scaled  $P$  from all 11 RCM models replicates the upper tail well, with very narrow RCM ranges which in all cases encapsulate the observed values.



**Fig. 3.** Monthly DC factors for the variables  $T$ ,  $P$ , and  $ET_{pot}$  in the far-future period (2071–2100) for all 11 RCMs. Climate models originating from the same GCM are depicted with the same line colour and climate models from the same RCM are depicted with the same marker symbol, along with the ensemble mean. A 1:1 line is shown on the  $P$  and  $ET_{pot}$  plots dividing positive and negative change.

**Table 5**  
Annual DC values for all variables ( $T$ ,  $P$ ,  $ET_{pot}$ , and  $Net P_{pot}$ ) in the far-future period (2071–2100) for all 11 RCMs.

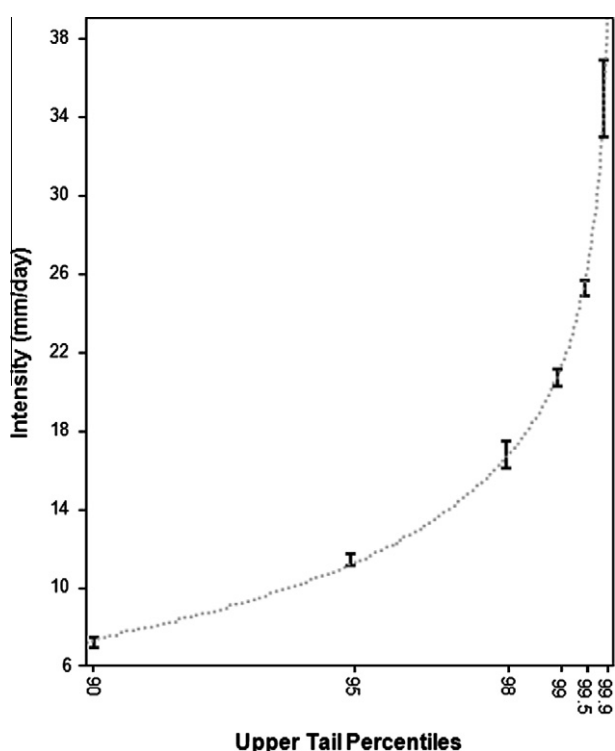
Climate model	Annual $\Delta$			
	$T$ (°C)	$P$ (-)	$ET_{pot}$ (-)	$Net P_{pot}$ (mm/day)
ARPEGE-RM5.1	2.07	0.89	1.25	-0.44
ARPEGE-HIRHAM5	2.02	0.84	1.16	-0.33
BCM2-HIRHAM5	2.26	1.13	1.19	0.18
BCM2-RCA3	1.97	1.13	1.15	0.30
ECHAM5-HIRHAM5	1.92	1.19	1.10	0.51
ECHAM5-RegCM3	2.29	1.07	1.10	0.06
ECHAM5-RACMO2	2.17	1.09	1.13	0.08
ECHAM5-REMO	2.14	1.06	1.14	-0.01
ECHAM5-RCA3	2.14	1.10	1.07	0.14
HadCM3-CLM	3.30	1.09	1.18	0.05
HadCM3-HadRM3	3.40	0.99	1.37	-0.23
ENSEMBLE mean	2.34	1.05	1.17	0.03



**Fig. 4.** Monthly DC values averaged for all RCM grids over Denmark in the far-future period 2071–2100, and averaged across the four GCM ensembles (i.e. all RCM pairings with each GCM) for the indicator variable potential net precipitation  $Net P_{pot}$  (mm) (i.e.  $P - ET_{pot}$ ) characterising overall climate change by compounding changes in radiation balance and precipitation.

**Table 6**  
Summary statistics and significance levels (*p*-values) for the hypotheses that DBS scaled precipitation is able to preserve the mean (*t*-test) and variance (*f*-test) of annual precipitation during the reference period (1991–2010).

Precipitation	Mean (mm/year)	Std. dev. (mm/year)	Max (mm/year)	Min (mm/year)	<i>t</i> -Test ( <i>p</i> -value)	<i>f</i> -Test ( <i>p</i> -value)
Observed	859	112	1016	611	–	–
<i>Climate models</i>						
ARPEGE-RM5.1	856	151	1120	611	0.93	0.20
ARPEGE-HIRHAM5	858	122	1120	604	0.97	0.71
BCM2-HIRHAM5	861	108	1112	610	0.97	0.88
BCM2-RCA3	857	102	1066	675	0.95	0.70
ECHAM5-HIRHAM5	861	88	1053	671	0.96	0.32
ECHAM5-RegCM3	860	108	1056	729	0.98	0.90
ECHAM5-RACMO2	862	115	1087	604	0.95	0.89
ECHAM5-REMO	862	101	1008	678	0.94	0.68
ECHAM5-RCA3	858	107	1023	690	0.97	0.86
HadCM3-CLM	846	102	1022	580	0.71	0.71
HadCM3-HadRM3	848	134	1053	680	0.77	0.44



**Fig. 5.** Observed cumulative daily precipitation for high intensity rainfall events during the reference period (1991–2010) vs. the range of DBS scaled precipitation projected by the 11 RCMs (error bars).

### 3.3. Difference between DC and DBC for future climates

To assess potential differences in mean *P* under both bias correction methods, we calculate DC factors between the DBS reference period and future periods for comparison with the DC factors from direct RCM output used in the DC correction method. Annual and seasonal DC factors for all DK-domains were approximately the same, indicating no difference in mean *P* changes under both methods. This result is unsurprising, given that the DC method perturbs mean changes and DBS methods was applied to preserve mean and variance.

The frequency of dry days in future periods under the DC method is identical to the observed reference period because mean changes in *P* are perturbed onto wet days only. For DBS far-future *P*, we calculate the frequency of dry days per season for comparison with the observed/DC method frequencies. For the two ARPEGE GCM driven models, which had negative *P* change signals, the fre-

quency of dry days increases by +5% to +8% in winter and spring and by +10% to +15% in summer and autumn. The two BCM GCM driven models project a reduction in dry days by –1% to –4% in spring and summer and an increase by +1% to +4% in autumn, with negligible winter changes in dry day frequency. With five RCMs paired with the ECHAM GCM, there is more variation between these models, though they all project an increase in spring and autumn dry days up to +7%. The HIRHAM5 and RCA3 RCMs project a reduction in dry days by –3% to –5% in the summer, while the RegCM3, RACMO2, and REMO RCMs project up to +2% increase in summer dry days. In winter, the HIRHAM5 and RACMO2 RCMs project a reduction in dry days by –2%, while the RCM3, RegCM3, and REMO RCMs project an increase in dry days up to +3%. The two HadCM3 GCM driven models project an increase by +8% to +11% in winter dry days and +6% in autumn dry days, with negligible differences in spring and summer dry days.

We analyse selected percentiles in the lower 75% tail and the upper 5% tail of *P* from DC and DBS in the far-future to evaluate how the methods compare in their generation of low and high intensity *P*. Because 30 year (2071–2100) DC factors are perturbed onto a 20 year observed dataset, we select 20 years of DBS *P* (2080–2099) for this comparison. In the reference period, the observed 25th percentile of *P* is 0.78 mm/day and the 75th percentile of *P* is 5.65 mm/day. Across all climate models and methods the 25th percentile of *P* in the far-future increases slightly, up to 0.87 mm/day with the DC method and 0.82 mm/day with DBS correction. At the 75th percentile of *P*, the two ARPEGE GCM driven models deviate from the rest of the ensemble, decreasing to around 4.7 mm/day under DC and 5.4 mm/day under DBS in the far-future. Across the rest of the models, the 75th percentile of *P* increases to a mean value of 6.09 mm/day under DC and 6.33 under DBS in the far-future.

To look at the upper tail of *P*, we focus on three models representing the spread of ensemble in terms of mean annual and seasonal *P* changes. Fig. 6 shows daily cumulative *P* curves for the far-future time period above the 95th percentile for three models' DC (perturbed on 20 year series) and DBS (20 years from 2080 to 2090) result, with observed *P* (20 years from 1991 to 2010) shown for reference. The ARPEGE-HIRHAM5 model projects a negative change signal, with decreasing mean annual, seasonal, and monthly *P*. With the DC method, this reduction is evenly transferred to every wet day (by month), which is why the upper percentiles of ARPEGE-HIRHAM5 *P* are all below the observations. The DBS method transfers the same mean decrease in *P*, but in a nonlinear way, by adjusting the distribution and intensity of wet days. The upper tail percentiles of ARPEGE-HIRHAM5 under the DBS method are estimated to be quite close to the upper tail of the observations (i.e. no change). The BCM2-RCA3 model projects a relatively high increase in *P*, and in this case the DC and DBS



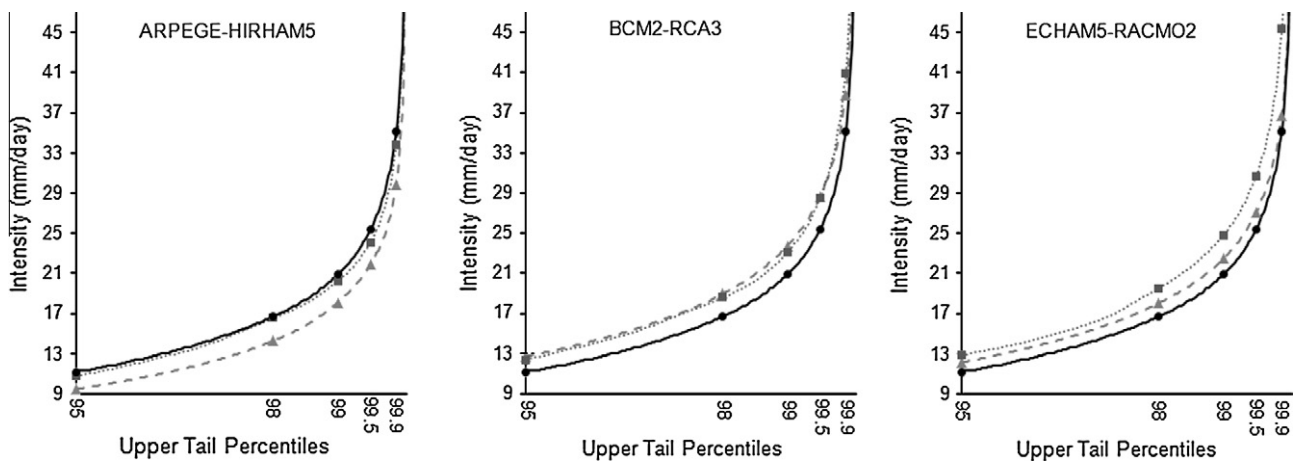


Fig. 6. Upper tail percentiles for DBS and DC precipitation in the far-future (2071–2100) displayed with observed precipitation in the reference period (1991–2010).

Table 7

*p*-Values for the hypotheses of unchanged mean between the reference period (1990–2010) and four other periods (1961–1990; 2011–2040; 2041–2070; 2071–2100) for the variables *T*, *P*, and *ET<sub>pot</sub>*. Significance above the 95th confidence level is shown in bold and above the 90th confidence level is shown in italic.

Climate Models	past 1961-1990	near 2011-2040	mid 2041-2070	far 2071-2100
<b>Precipitation (<i>p</i> -values)</b>				
ARPEGE-RM5.1	0.32	0.42	0.22	<b>0.05</b>
ARPEGE-HIRHAM5	<b>0.09</b>	0.50	<b>0.08</b>	<b>0.00</b>
BCM2-HIRHAM5	<b>0.08</b>	0.26	<b>0.02</b>	<b>0.00</b>
BCM2-RCA3	0.17	0.05	<b>0.01</b>	<b>0.00</b>
ECHAM5-HIRHAM5	<b>0.01</b>	0.28	<b>0.00</b>	<b>0.00</b>
ECHAM5-RegCM3	0.43	0.93	0.31	<b>0.01</b>
ECHAM5-RACMO2	0.75	0.66	0.13	<b>0.01</b>
ECHAM5-REMO	0.23	0.88	0.46	0.13
ECHAM5-RCA3	0.21	0.50	0.17	<b>0.01</b>
HadCM3-CLM	0.96	0.17	<b>0.01</b>	<b>0.03</b>
HadCM3-HadRM3	0.69	0.48	0.23	0.56
<b>Temperature (<i>p</i> -values)</b>				
ARPEGE-RM5.1	<b>0.04</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
ARPEGE-HIRHAM5	0.16	<b>0.01</b>	<b>0.00</b>	<b>0.00</b>
BCM2-HIRHAM5	<b>0.03</b>	0.65	<b>0.00</b>	<b>0.00</b>
BCM2-RCA3	<b>0.05</b>	0.28	<b>0.00</b>	<b>0.00</b>
ECHAM5-HIRHAM5	0.13	0.26	<b>0.00</b>	<b>0.00</b>
ECHAM5-RegCM3	<b>0.07</b>	<b>0.08</b>	<b>0.00</b>	<b>0.00</b>
ECHAM5-RACMO2	<b>0.06</b>	<b>0.10</b>	<b>0.00</b>	<b>0.00</b>
ECHAM5-REMO	<b>0.04</b>	0.14	<b>0.00</b>	<b>0.00</b>
ECHAM5-RCA3	<b>0.04</b>	0.23	<b>0.00</b>	<b>0.00</b>
HadCM3-CLM	0.18	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
HadCM3-HadRM3	<b>0.03</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
<b>Potential Evapotranspiration (<i>p</i> -values)</b>				
ARPEGE-RM5.1	0.77	0.19	<b>0.00</b>	<b>0.00</b>
ARPEGE-HIRHAM5	0.31	0.49	<b>0.00</b>	<b>0.00</b>
BCM2-HIRHAM5	<b>0.06</b>	0.37	<b>0.03</b>	<b>0.00</b>
BCM2-RCA3	0.92	0.97	<b>0.02</b>	<b>0.00</b>
ECHAM5-HIRHAM5	0.63	0.28	0.88	0.83
ECHAM5-RegCM3	0.99	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
ECHAM5-RACMO2	0.38	<b>0.06</b>	<b>0.01</b>	<b>0.00</b>
ECHAM5-REMO	0.28	0.31	0.68	0.13
ECHAM5-RCA3	0.38	0.31	0.37	0.21
HadCM3-CLM	0.73	0.22	0.22	<b>0.00</b>
HadCM3-HadRM3	<b>0.04</b>	0.24	<b>0.00</b>	<b>0.00</b>

upper tails are very similar to each other, with each percentile just a few mm above observed *P*. The method by which the mean

increase in *P* was transferred to each day, for this model, has negligible impacts (i.e. differences between DC and DBS) to the high intensity *P*. The ECHAM5-RACMO2 model is a median model in terms of annual and monthly *P* increase within the ensemble. The DC *P* above the 95th percentile is slightly higher than the observed *P*, while the DBS method allows for much higher intensity values in the upper tail. Most notably, the DBS method allows the upper tail of this median model to surpass the model with higher relative increase (BCM-RCA3).

### 3.4. Statistical significance of annual mean projected climate change

To analyse the evolution of climate change over time, including past and future changes as simulated by the RCMs, we test for difference of annual means between the past and future periods for DBS corrected data. We look at two past periods, 1961–1990 and 1991–2010, to see if significant climate changes have already occurred. We compare the reference period 1991–2010 to the near, mid, and far-future periods to see at what point in the 21st century projected climate changes are significantly different from the current RCM climate, if at all. With annual mean values following normal distributions in all past and future periods (with the exception of *P* from the BCM2-RCA3 model, which is not normally distributed in the mid-future), we can make tests of change in means between periods as *t*-tests. Table 7 reports *p*-values for the hypothesis of no change in annual mean, which is rejected when *P* < 0.05. *P* = 0 indicates a 100% probability that the mean annual climate is different in the two periods, 0.05 indicates a 95% probability, etc. Comparing the two past periods, *T* is significantly different above 95% in six models, while *P* and *ET<sub>pot</sub>* are significantly different above 95% in only one model. Above 90% probability, two more models are significantly different in *T* and *P*, and one more model in terms of *ET<sub>pot</sub>*. For some models, simulated climate is already significantly different within the past time period, and more pronounced in variables like *T* which are not subject to high interannual variability like *P* and *ET<sub>pot</sub>*.

In the future periods, *T* is the first variable to show strong significant differences at 100% probability for the mid and far-future in every model, as well in the near-future for four models. *P* and *ET<sub>pot</sub>* are more similar, with eight and nine models significantly different above 95% probability, respectively. However, in the mid-future period, the *P* signal is not as strong, with only four models significantly different above 95% probability, while *ET<sub>pot</sub>* is significantly different above 95% probability in seven models. For the future periods, it is clear that by the end of the 21st century, mean annual climate for all variables and in most models is significantly different. For the variables *T* and *ET<sub>pot</sub>*, changes are significant in the

mid-future period for more models compared to  $P$ . For the near-future, changes are logically not as strong and most models do not have a significant difference to the reference period.

### 3.5. Length of reference and change period

To evaluate the stability of climate change signals in different length reference periods, we calculate monthly DC factors of  $P$  for 5, 10, 15, 20, and 30 year periods for the six model pairings with the longest transient simulation (1951–2100). The analyses are based on data for the past period 1951–2010 and the far-future period 2071–2100 for which 72 combinations of non-overlapping 5 year periods, 18 combinations of 10 year periods, etc. were sampled. Reference periods are only compared to future periods of the same length. The underlying assumption that the past and far-future periods do not have trends within the respective periods is tested for all six models used in further analyses. For example, all models demonstrate a significant positive or negative trend in annual  $P$  over the 21st century (90 years), but the respective near, mid, and far-future 30 year periods have no significant trend in annual  $P$ . Though magnitude of the DC change factor varies by month and model, the standard deviations are consistent between months within a respective reference period. For each period length the standard deviations among the monthly DC factors are calculated as an indicator of the robustness of using a period of this length. Fig. 7 shows a plot of the mean standard deviations from annual DC factors for all reference period lengths, with error bars indicating the range of standard deviations across the six models. The results show that reference periods of 10 years and below have high variability between DC factors, implying that 10 year periods are too short to robustly assess the strength in the climate change signal. On the other hand period lengths over 15 years appear suitable.

## 4. Discussion

### 4.1. Distribution based scaling (DBS)

We applied and evaluated the DBS method on a seasonal basis over six domains covering Denmark and compared it with results from the DC approach using 11 climate models from the ENSEMBLES project (Christensen et al., 2009; van der Linden and Mitchell, 2009). A good fit of climate data, both observed and simulated, to a gamma distribution was crucial for obtaining an accurate scaling. A double gamma distribution split at the 95th percentile was necessary to properly fit the precipitation data. We noted a considerable spread of dry day correction values and gamma distribution parameters across models, domains and seasons.

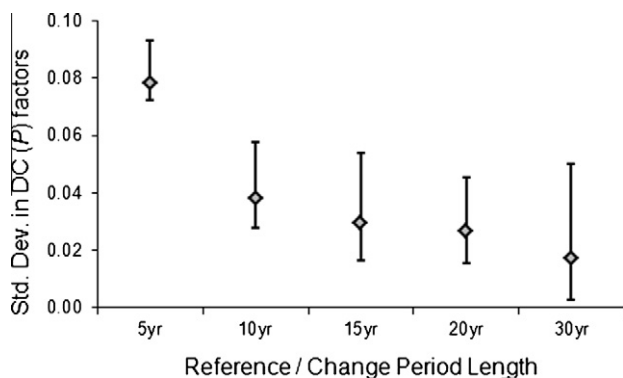


Fig. 7. Mean std. dev. of annual DC values for precipitation in six climate models from 5, 10, 15, 20, and 30 year reference periods compared to far-future periods of the same length. Error bars show the range in std. dev. across the six models.

To assess the accuracy and robustness of the DBS bias correction method, we tested how well it was able to reproduce the statistical properties of precipitation data for the 20 year reference period for which the DBS parameters were fitted. DBS corrected  $P$  from all 11 RCMs accurately reproduces high intensity rainfall and mean annual precipitation sampled from all six domains and all four seasons. While hypotheses that the interannual variations were preserved could not be rejected, they were simulated with varying accuracy. Previous studies on the DBS methods have not reported such tests.

We further benchmarked the DBS method against the DC method. While there is no difference in the mean monthly or annual climate change signals for a given model under the two methods, for daily precipitation intensity the strength of the signal varied between methods. The DC approach perturbs mean changes accurately, but this often results in the precipitation regime being mischaracterised, especially in the upper tail. There are limited studies (e.g. Teutschbein and Seibert, 2012b) that have compared multiple bias correction methods on an ensemble of climate models, which provides an opportunity for both inter and intra model comparisons, and this is the first study to do so for Denmark. Interestingly, some models considered average (median of the ensemble) for projected precipitation increase with DC, after DBS correction, showed higher frequency and intensity of wet days than the models that had been characterised by relatively higher increases. Focus of previous work has been on developing the DBS method and evaluating future precipitation, especially extremes, whereas we thoroughly evaluate the DBS bias correction performance in the reference period the method was trained on. A comparison of mean, variance and daily percentiles in the upper tail between observed precipitation and the DBS corrected data validate the method's ability to reproduce local climate dynamics, though mean precipitation patterns are more accurately reproduced than the variance. Clearly both methods are capable of preserving mean changes, but the DBS method is suited for reproducing extreme precipitation changes.

Previous DBS studies have mostly focussed on precipitation and temperature (Piani et al., 2010; Yang et al., 2010), while van Roosmalen et al. (2011) also used DBS to scale potential evapotranspiration. We found that accurate scaling of potential evapotranspiration with the DBS method as applied to precipitation is not possible because the variable does not follow a gamma distribution. Gamma distribution parameters fit to preserve the mean do not also capture the variance in potential evapotranspiration. This variable is challenging to scale as it is the product of many atmospheric variables, some normally distributed (e.g. temperature) and others that are not (e.g. relative humidity). Temperature and potential evapotranspiration are closely tied and both covary with precipitation on a daily basis. With DBS scaling only performing well on precipitation, we applied an error bias correction on direct RCM outputs of temperature and potential evapotranspiration (calculated from RCM model outputs) to preserve the daily covariance between all three variables.

### 4.2. Significance of projected climate change

Historical data are often analysed for significance of climate change (Karlsson et al., 2013). However, we know of no other studies that have analysed the extent projected climate changes are actually significantly different from the current climate, and if changes simulated within the past climate are significant. The driving GCM is most influential in the climate change signal strength, magnitude, and direction when considering the variables precipitation, temperature, and potential evapotranspiration. Over the historical period (1961–2010) climate models do simulate climate changes already in temperature. Natural variability in precipitation

and potential evapotranspiration is stronger than any simulated changes, resulting in no significant changes over the historical period. By the end of the 21st century, across all models and variables, climate is significantly different from the reference periods. In the far-future, climate change signals are clear, robust, and can be distinguished from natural climate variability. In the near and mid-future periods, the signal to noise ratio make changes less detectable and therefore less significant statistically.

#### 4.3. Choice of reference period

As projected climate change in the 11 RCMs evolves throughout the 21st century, there are clear trends of either increasing or decreasing precipitation. Therefore, a stationary climate change signal cannot be assumed in a 90 year time period. We evaluate DC factors from periods ranging from 5 to 30 years to determine the length of time needed to ensure distinguishing climate change from natural variability. Time periods under 10 years exhibited low signal to noise ratio (variability) evidenced by their high standard deviations, at 15–20 year time periods there is a drastic decrease in the standard deviations of DC factors. At time periods of 20–30 years the climate change signals are stable and distinguishable from natural variability, and should result in robust DC factors.

#### 4.4. Robustness of findings

Compared to the many bias correction methods reported in literature (e.g. Johnson and Sharma, 2011; Rojas et al., 2011; Li et al., 2010; Piani et al., 2010; Yang et al., 2010; Graham et al., 2007; Lenderink et al., 2007; Leander and Buishand, 2007), we have applied two methods for precipitation located at opposite ends of the spectrum in terms of complexity. The DC method is the most simplistic (i.e. indirect, linear) and the DBS method is the most sophisticated (i.e. direct, nonlinear, parametric). For  $T$  and  $ET_{pot}$  we have applied DC method and the BR method which is in the lower-middle range of this spectrum (i.e. direct, linear).

Results from other studies show that the robustness of these downscaling methods with respect to precipitation tend to follow the spectrum of complexity, with DBS methods being the most robust both in terms of ability to reproduce the historical data (Berg et al., 2012; Sunyer et al., 2012; Teutschbein and Seibert, 2012b; Watanabe et al., 2012; Anandhi et al., 2011) and performance under changed conditions (Teutschbein and Seibert, 2012a). DC methods cannot be evaluated in the reference period but have the least reliable corrections for changed (future) conditions, especially in dry years, and generally cannot capture variance in the RCMs (Teutschbein and Seibert, 2012a). BR and linear scaling methods often underestimate heavy precipitation events (Leander and Buishand, 2007), and can be somewhat inconsistent between models, with some corrected RCM outputs exhibiting biases of the same magnitude as uncorrected RCM output (Teutschbein and Seibert, 2012a). Both quantile mapping and DBS methods outperform these methods in terms of ability to reproduce dry day frequency and wet day intensity, but DBS methods have higher correction skills under changed conditions as they are less parameterised and therefore more robust. Hence our findings based on the DBS method appear to produce results that are robust compared to other possible choices of downscaling methods.

Johnson and Sharma, 2011 found increased uncertainty among direct bias correction methods in terms of variability, noting that future variability is constrained to that of the historical period using DC methods. To our knowledge, none of the methods have been tested explicitly for their ability to reproduce the interannual variability, which is a key factor in the significance tests, though it appears to be equally well reproduced by the various direct methods. With respect to our findings on the significance of climate

change throughout the 21st century, we argue that the results are not affected by which downscaling method is used as long as we keep within direct methods.

We evaluated the stability of reference periods is by calculating change factors on DBS corrected data for systematically increasing period lengths. We expect less sophisticated methods, like linear scaling and the most simplistic quantile mapping methods, to show more stability at shorter change periods lengths due to their inability to capture variance and tendency to underestimate extreme values.

## 5. Conclusions

This analysis of multiple climate models, bias correction methods, and time periods elucidates the nature of climate model uncertainty. Hydrological modelling studies of climate change impacts must consider these additional sources of uncertainty from the climate model inputs. The ensemble of 11 RCMs considered in this study vary in strength, significance, and sometimes in direction of the climate change signal. This study confirms previous studies showing that it is unreliable to base an impact study on just one model's projection of climate change. Incorporating multiple climate models, ideally from different GCMs, begins to address climate model uncertainty by understanding the range of projections that are possible from different models. The choice of bias correction method introduces further uncertainty in climate change data.

We demonstrate that both DC and DBS methods equally retain mean monthly change characteristics. The simplistic monthly DC approach is adequate for capturing the smooth temporal characteristics of temperature changes, but is insufficient at recreating projected precipitation regimes, which vary day to day and grid to grid. The more complex daily DBS correction method is accurate and robust, transferring changes in the mean as well as the variance, and improving the characterisation of day to day variation as well as heavy precipitation events. Since the DBS method corrects direct RCM output, the climate series is not limited to the reference period length and variations in the evolution of climate change over the time are preserved. Compared to the DC method, the DBS method could carry possible spatial biases in the RCM onto the impact model.

We know of no studies that have tested the temporal significance of past and future climate change, or systematically analysed the robustness of change periods within RCMs. The analyses of projected climate change significance over time provide a different perspective on climate change and assumptions about current and future climate. With some models having significant differences in climate variables within the past timeframe, we show that simulated current climate characteristics are not necessarily stationary within Denmark at a basin scale. The temporal positioning of a reference period (e.g. 1961–1990 vs. 1991–2010) might impact the magnitude of relative climate change (e.g. DC factors) from the RCMs. The length of the reference and change periods was also analysed to determine what size time period was adequate to overcome natural variability and still have stationarity in the climate change signal. A drastic decrease in the standard deviations of DC factors was seen in periods over 15 years, with periods over 20 years continuing to decrease and level off.

While 30 year time periods prove adequate in length for robust climate change signals, the significance tests in the future periods show where these change signals overcame the noise of natural variability. Overall, climate change signals in the near-future are hidden by natural variability and are therefore not significant, in the mid-future the choice of climate model mostly influences the significance of climate change signals, and in the far-future the climate change signals are strong across all models and variables.



## Acknowledgements

The present study was funded by a grant from the Danish Strategic Research Council for the project HYdrological Modelling for Assessing Climate Change Impacts at different Scales (HYACINTS – [www.hyacints.dk](http://www.hyacints.dk)) under Contract No: DSF-EnMi 2104-07-0008. We acknowledge the ENSEMBLES data sets (<http://ensembles-eu.metoffice.com>) and would like to thank Ole Bøssing Christensen DMI for assistance in retrieving and understanding these data.

## References

- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. Irr. & Drain. Paper 56. UN-FAO, Rome, Italy.
- Allerup, P., Madsen, H., Vejen, F., 1997. A comprehensive model for correcting point precipitation. *Nord. Hydrol.* 28, 1–20.
- Anandhi, A., Frei, A., Pierson, D.C., Schneiderman, E.M., Zion, M.S., Lounsbury, D., Matonse, A.H., 2011. Examination of change factor methodologies for climate change impact assessment. *Water Resour. Res.* 47(W030501), 1–10.
- Bárdossy, A., Pegram, G., 2011. Downscaling precipitation using regional climate models and circulation patterns towards hydrology. *Water Resour. Res.* 47(W04505), 1–18.
- Bates, B.C., Kundzewicz, Z.W., Wu, S., Palutikof, J.P. (Eds.), 2008. Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change. IPCC Secretariat, Geneva, pp. 1–210.
- Berg, P., Feldmann, H., Panitz, H.-J., 2012. Bias correction of high resolution regional climate data. *J. Hydrol.* 448–449, 80–92.
- Böhm, U., Kücken, M., Ahrens, W., Block, A., Hauffe, D., Keuler, K., Rockel, B., Will, A., 2006. C1m – The Climate Version of Im: Brief Description and Long-Term Applications. COSMO Newsletter, 6.
- Christensen, J.H., Rummukainen, M., Lenderink, G., 2009. Formulation of Very-High-Resolution Regional Climate Model Ensembles for Europe [Research Theme 3]. ENSEMBLES: Climate Change and Its Impacts: Summary of research and Results from the ENSEMBLES Project. Met Office Hadley Centre, Exeter, UK, pp. 47–58.
- Christensen, J.H., Christensen, O.B., 2007. A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Clim. Change* 81, 7–30.
- Collins, M., Booth, B.B.B., Harris, G.R., Murphy, J.M., Sexton, D.M.H., Webb, M.J., 2006. Towards quantifying uncertainty in transient climate change. *Clim. Dynam.* 27 (2–3), 127–147.
- Déqué, M., Rowell, D.P., Lüthi, D., Giorgi, F., Christensen, J.H., Rockel, B., Jacob, D., Kjellström, E., de Castro, M., van den Hurk, B., 2007. An Intercomparison of Regional Climate Model Simulations for Europe: Assessing Uncertainties in Model Predictions.
- Déqué, M., Dreveton, C., Braun, A., Cariolle, D., 1994. The ARPEGEIFS atmosphere model: a contribution to the French community climate modelling. *Clim. Dynam.* 10, 249–266.
- Ekström, M., Jones, P.D., Fowler, H.J., Lenderink, G., Buishand, T.A., Conway, D., 2007. Regional climate model data used within the SWURVE project 1: projected changes in seasonal patterns and estimation of PET. *Hydrol. Earth Syst. Sci.* 11 (3), 1069–1083.
- Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modeling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.* 27, 1547–1578.
- Graham, L.P., Andréasson, J., Carlsson, B., 2007. Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods – a case study on the Lule River Basin. *Clim. Change* 81, 293–307.
- Gutowski, W.J., Kozak, K.A., Arritt, R.W., Christensen, J.H., Patton, J.C., Takle, E.S., 2007. A possible constraint on regional precipitation intensity changes under global warming. *J. Hydrometeorol.* 8, 1382–1396.
- Hansen, J.W., Challion, A., Ines, A., Wheeler, T., Moron, V., 2006. Translating climate forecasts into agricultural terms: advances and challenges. *Clim. Res.* 33, 27–41.
- Hay, L.E., Wilby, R.J.L., Leavesley, G.H., 2000. A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States. *J. Am. Water Resour. Assoc.* 36 (2), 387–397.
- Henriksen, H.J., Troldborg, L., Nyegaard, P., Sonnenborg, T.O., Refsgaard, J.C., Madsen, B., 2003. Methodology for construction, calibration and validation of a national hydrological model for Denmark. *J. Hydrol.* 280 (1–4), 52–71.
- Højberg, A.L., Troldborg, L., Stisen, S., Christensen, B., Henriksen, H.J., 2013. Stakeholder driven update and improvement of a national water resources model. *Environ. Modell. Softw.* 40, 202–213.
- Huntington, T.G., 2006. Evidence for intensification of the global water cycle: review and synthesis. *J. Hydrol.* 319, 83–95.
- IPCC, 2007. Climate Change 2007. Synthesis Report. Contribution of Working Groups I, II, and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland.
- Jacob, D., Andrae, U., Elgered, G., Fortelius, C., Graham, L.P., Jackson, S.D., Karstens, U., Chr. Koepken Lindau, R., Podzun, R., Rockel, B., Rubel, F., Sass, H.B., Smith, R.N.D., Van den Hurk, B.J.J.M., Yang, X., 2001. A comprehensive model intercomparison study investigating the water budget during the BALTEX-PIDCAP period. *Meteorol. Atmos. Phys.* 77 (1–4), 19–43.
- Johnson, F., Sharma, A., 2011. Accounting for interannual variability: a comparison of options for water resources climate change impact assessments. *Water Resour. Res.* 47(W04508), 1–20.
- Jones, C.G., Willén, W., Ullerstig, A., Hannsson, U., 2004. The Rossby centre regional atmospheric climate model: part I: model climatology and performance for the present climate over Europe. *Ambio* 33 (4–5), 211–220.
- Karlsson, I.B., Sonnenborg, T.O., Jensen, K.H., Refsgaard, J.C., 2013. Evaluating the influence of long term historical climate change on catchment hydrology – using drought and flood indices. *Hydrol. Earth Syst. Sci. Discuss.* 10, 1–56.
- Kjellström, E., Bärring, L., Gollvik, S., Hansson, U., Jones, C., Samuelsson, P., Rummukainen, M., Ullerstig, A., Willén, U., Wyser, K., 2005. A 140-year Simulation of European Climate with the New Version of the Rossby Centre Regional Atmospheric Climate Model (RCA3). Reports Meteorology and Climatology, 108, SMHI, SE-60176 Norrköping, Sweden, pp. 1–54.
- Leander, R., Buishand, T.A., 2007. Resampling of regional climate model output for the simulation of extreme river flows. *J. Hydrol.* 332, 487–496.
- Lenderink, G., Buishand, T.A., van Deursen, W.P.A., 2007. Estimates of future discharges of the river Rhine using two climate scenario methodologies: direct versus delta approach. *Hydrol. Earth Syst. Sci.* 11 (3), 1145–1159.
- Li, H., Sheffield, J., Wood, E.F., 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *J. Geophys. Res.* 115(D10101), 1–20.
- Loaiciga, H.A., Valdes, J.B., Vogel, R., Garvey, J., Schwarz, H., 1996. Global warming and the hydrologic cycle. *J. Hydrol.* 174, 83–127.
- Makkink, G.F., 1957. Testing the Penman formula by means of lysimeters. *J. Inst. Water Eng.* 11, 277–288.
- Milley, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Stationarity is dead: whither water management? *Science* 319, 573–574.
- Otterå, O.H., Bentsen, M., Bethke, I., Kvamstø, N.G., 2009. Simulated pre-industrial climate in Bergen Climate Model (version 2): model description and large-scale circulation features. *Geosci. Model Dev.* 2, 197–212.
- Pal, J.S., Giorgi, F., Bi, X., Elguindi, N., Solmon, F., Gao, X.J., Francisco, R., Zakey, A., Winter, J., Ashfaq, M., Syed, F., Bell, J., Diffenbaugh, N., Karmacharya, J., Konare, A., Martinez-Castro, D., Porfiro da Rocha, R., Sloan, L., Steiner, A., 2007. Regional climate modeling for the developing world: the ICTP RegCM3 and RegCM3. *Bull. Am. Meteorol. Soc.* 88, 1395–1409.
- Piani, C., Haerter, J.O., Coppola, E., 2010. Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.* 99, 187–192.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes: The Art of Scientific Computing, third ed. Cambridge University Press, Cambridge, UK.
- Rummukainen, M., 2010. State-of-the-art with regional climate models. *Clim. Change* 1 (1), 82–96.
- Radu, R., Déqué, M., Somot, S., 2008. Spectral nudging in a spectral regional climate model. *Tellus* 60A, 898–910.
- Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblüeh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., Tompkins, A., 2003. The Atmospheric General Circulation Model ECHAM5. Part I: Model Description. Max Planck Institute for Meteorology Rep. 349. MPI for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany.
- Rojas, R., Feyen, L., Dosio, A., Bavera, D., 2011. Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations. *Hydrol. Earth Syst. Sci.* 15, 2599–2620.
- Scharling, M., 2001. KLIMAGRID DANMARK. Sammenligning af potentiel fordampning beregnet ud fra Makkinks formel og den modificerede Penman formel. Danish Meteorological Institute, Technical Report 01-19. Danish Meteorological Institute, Copenhagen (in Danish).
- Scharling, M., 2000. KLIMAGRID DANMARK. Klimagrid - Normaler 1961–90, måned- og årsværdier. Nedbør 10\*10, 20\*20 & 40\*40 km, temperatur og potentiel fordampning 20\*20 & 40\*40 km. Danish Meteorological Institute, Technical Report 00–11. Danish Meteorological Institute, Copenhagen (in Danish).
- Sharma, D., Das Gupta, A., Babel, M.S., 2007. Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand. *Hydrol. Earth Syst. Sci.* 11, 1373–1390.
- Shepard, D., 1967. A two-dimensional interpolation function for irregularly spaced data. In: Proc. 23rd Nat. Conf. ACM.
- Stisen, S., Højberg, A.L., Troldborg, L., Refsgaard, J.C., Christensen, B.S.B., Olsen, M., Henriksen, H.J., 2012. On the importance of appropriate rain-gauge catch correction for hydrological modelling at mid to high latitudes. *Hydrol. Earth Syst. Sci.* 16, 4157–4176.
- Sunyer, M.A., Madsen, H., Ang, P.H., 2012. A comparison of different regional climate models and statistical downscaling methods for extreme rainfall estimation under climate change. *Atmos. Res.* 103, 119–128.
- Teutschbein, C., Seibert, J., 2012a. Is bias correction of Regional Climate Model (RCM) simulations possible in non-stationary conditions? *Hydrol. Earth Syst. Sci. Discuss.* 9, 12765–12795.
- Teutschbein, C., Seibert, J., 2012b. Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. *J. Hydrol.* 456–457, 12–29.



- Trenberth, K.E., Dai, A.G., Rasmussen, R.M., Parsons, D.B., 2003. The changing character of precipitation. *Bull. Am. Meteorol. Soc.* 84, 1205–1217.
- van der Linden, P., Mitchell, J.F.B., 2009. ENSEMBLES: Climate Change and its Impacts: Summary of Research and Results from the ENSEMBLES Project. Tech. Rep., Met Office Hadley Centre, Exeter, UK, pp. 1–160.
- van Meijgaard, E., van Ulft, L.H., van de Berg, W.J., Bosveld, F.C., van den Hurk, B.J.J.M., Lenderink, G., Siebesma, A.P., 2008. The KNMI Regional Atmospheric Climate Model RACMO, Version 2.1. KNMI Technical Report 302. KNMI, De Bilt, The Netherlands.
- van Roosmalen, L., Sonnenborg, T.O., Jensen, K.H., Christensen, J.H., 2011. Comparison of hydrological simulations of climate change using perturbations of observations and distribution-based scaling. *Vadose Zone J.* 10, 136–150.
- van Roosmalen, L., Sonnenborg, T.O., Jensen, K.H., 2009. Impact of climate and land use change on the hydrology of a large-scale agricultural catchment. *Water Resour. Res.* 45(W00A15), 1–18.
- Wild, M., Liepert, B., 2010. The Earth radiation balance as driver of the global hydrological cycle. *Environ. Res. Lett.* 5, 1–8.
- Wilks, D.S., 2006. *Statistical Methods in the Atmospheric Sciences*, second ed. Academic Press, New York, pp. 95–102.
- Wood, A.W., Leung, L.R., Sridhar, V., Lettenmaier, D.P., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim. Change* 62, 189–216.
- Xu, C., Widén, E., Halldin, S., 2005. Modelling hydrological consequences of climate change—progress and challenges. *Adv. Atmos. Sci.* 22 (6), 789–797.
- Watanabe, S., Kanae, S., Seto, S., Yeh, P.J.-F., Hirabayashi, Y., Oki, T., 2012. Intercomparison of bias-correction methods for monthly temperature and precipitation simulated by multiple climate models. *J. Geophys. Res.* 114, D23114.
- Yang, W., Andréasson, J., Graham, L.P., Olsson, J., Rosberg, J., Wetterhall, F., 2010. Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *Hydrol. Res.* 41 (3–4), 211–229.