Projection of future daily precipitation series and extreme events by using a multi-site statistical downscaling model over the great Montréal area, Québec, Canada

D. I. Jeong, A. St-Hilaire, T. B. M. J. Ouarda and P. Gachon

ABSTRACT

This study suggested strategies to project future precipitation series based on a multi-site hybrid SDM (statistical downscaling model), which can downscale precipitation series at multiple observation sites simultaneously by combining the multivariate multiple linear regression (MMLR) model and the stochastic randomization procedure. The hybrid SDM and future projection methodologies applied to 10 observation sites located in the great area of Montréal, Québec, Canada. Six future independent precipitation series were projected from six sets of future atmospheric predictors using three AOGCMs (Atmosphere-Ocean Global Climate Models, i.e. CGCM2, CGCM3, HadCM3) and three IPCC SRES emission scenarios (B2, A1B and A2). Downscaled climate change signals on wet/dry sequences and extreme indices of precipitation time series were evaluated over the future period from 2060 to 2099 with respect to the historical period from 1961 to 2000. The future scenarios of all three AOGCMs showed a consistent increase of 7.9–44.6% in winter while only those of HadCM3 and CGCM3 showed a decrease of 2.3–23.0% in summer compared to their historical values. Precipitation series of CGCM2 A2 and CGCM3 A2 scenarios yielded the largest increase in winter, while those of HadCM3 B2 and A2 scenarios yielded the largest decrease in summer for all statistics indices.

Key words | climate change, extreme event, multi-site, precipitation, statistical downscaling model

D. I. Jeong (corresponding author)
Centre ESCER (Étude et Simulation du Climat à l'Échelle Régionale),
UQAM (Université du Québec à Montréal),
201 Ave. President-Kennedy,
Montreal,
Quebec H3A 2K6,
Canada
E-mail: *jeong@sca.uqam.ca*

A. St-Hilaire
T. B. M. J. Ouarda
INRS-ETE,
University of Québec,
490 de la Couronne,
Québec G1K 9A9,
Canada

P. Gachon
Canadian Centre for Climate Modelling and Analysis (CCCma),
Climate Research Division,
Environment Canada,
Montréal,
Québec,
Canada

INTRODUCTION

Climate change impact studies related to hydrologic regimes often require realistic future scenarios of climate variables such as daily precipitation and temperatures for a targeted region or basin. Atmosphere-Ocean Global Climate Models (AOGCMs) are the predominant sources of climate information to provide present and future atmospheric variables (e.g. air pressure, humidity, temperature, wind, etc.) under the constraints of increases in greenhouse gases and aerosols (GHGA) concentration scenarios in the atmosphere. However, outputs of AOGCMs cannot be employed directly in regional or basin scale hydrologic models because of their coarse global-scale grids (generally larger than 2° latitude by 2° longitude). Downscaling techniques can be applied to deal with the scales mismatch between outputs of AOGCMs and inputs of regional impact models.

doi: 10.2166/nh.2012.183

Downscaling techniques are broadly categorized into two approaches: dynamic downscaling and statistical downscaling. Dynamic downscaling uses regional climate models (RCMs) physically to drive finer-scale regional climate variables from boundary variables driven by a host AOGCM. However, direct application of RCMs for a regional impact assessment is often restricted because of their complicated design, high computational cost and/or large bias partly originating from coarse AOGCM outputs (Wilby *et al.* 2002; Anandhi *et al.* 2008). Statistical downscaling models (SDMs) use statistical relationships between global scale AOGCM outputs and at site surface variables of interest (Wilby *et al.* 2002; Tripathi *et al.* 2006). SDMs are both computationally inexpensive with respect to RCMs and flexible. They provide downscaled information at the equivalent of

climatic stations. SDMs are generally categorized into three groups: weather typing approaches, stochastic weather generators and regression-based approaches (see further details in Wilby *et al*. 2004). Weather typing approaches generate local climate variable by resampling from observation atmospheric data (Buishand & Brandsma 1997; Zorita & von Storch 1999; Hewitson & Crane 2002). Stochastic weather generator approaches reproduce the statistical characteristics of local climate variables based on random number generator, Markov chain model and/or parametric (or non-parametric) distributions (e.g. Semenov & Barrow 1997; Wilks 1998 and 1999; Qian *et al*. 2002; Mehrotra & Sharma 2007; Khalili *et al*. 2009). Regression-based approaches derive empirical relationships between the reanalysis atmospheric variables (predictors) and local climate variables (predictands) by linear or non-linear transfer functions (e.g. Buishand & Brandsma 2001; Trigo & Palutikof 2001; Huth 2002; Huth 2004; Busuioc *et al*. 2006; Tolika *et al*. 2007; Hessami *et al*. 2008; Huth *et al*. 2008). In this process, the reanalysis variables are used as surrogate AOGCM outputs for the calibration.

Among surface climate variables, daily precipitation is one of the most important pieces of information for a range of socioeconomic activities (e.g. agriculture, forestry and tourism), as well as human health and ecology (Schmidli *et al*. 2007).

Weather typing and weather generator approaches are frequently employed for present precipitation simulation among the three SDM approaches, because they can reproduce realistic precipitation occurrence and amount series easily by using classification, resampling technique, Markov chain process and conventional statistical distributions (Bardossy & Plate 1992; Wilks 1999; Bellone *et al*. 2000; Mehrotra & Sharma 2007). Yang *et al*. (2005) adopted a generalized linear modelling (GLM) approach to rainfall weather generator. They developed rainfall occurrence and amount models using logistic regression and gamma distribution in the GLM framework. This GLM approach includes a form of weather generator based on the Markov chain approach by using previous wet or dry day condition as input variable for the occurrence model and uses the gamma distribution for the amount model. However, future climate scenarios generated by weather typing are hardly consistent with the scenarios of a selected AOGCM. In the stochastic weather generator approaches, adjusted parameters to project future climate scenarios often produce unrealistic outputs (Wilby *et al*. 2002). Regression-based approaches can produce consistent future climate variable from a host AOGCM by using empirical relationships between climate variable and atmospheric predictors. However, the AOGCM predictors of regression-based SDM usually explain only a part of the observed variability of the local precipitation amounts (Wilby *et al*. 2002). Therefore, some regression-based SDMs employ stochastic processes to reproduce the unexplained natural variability (Wilby *et al*. 2003; Bürger & Chen 2005; Hessami *et al*. 2008).

Projection of reliable climate series for future climate change is essential to prepare appropriate adaptation strategies for impact studies on a targeted region (Bell *et al*. 2004). In particular, simultaneous generation of surface precipitation series at multiple observation sites that account for their spatial coherence in a basin are often required in water resources management (e.g. flood control, water supply and drought management; Mehrotra & Sharma 2007). Several studies have suggested multi-site simulation methodologies for present precipitation series based on stochastic weather generator approaches (Wilks 1998; Wilks 1999; Qian *et al*. 2002; Mehrotra & Sharma 2007), weather typing approaches (Palutikof *et al*. 2002; Fowler *et al*. 2005), regression-based approaches (Harpham & Wilby 2005) and hybrid approaches combining regression-based and stochastic generator methods (Wilby *et al*. 2003; Harpham & Wilby 2005; Hessami *et al*. 2008). However, the multi-site SDMs have been scantily applied for projection of future precipitation scenarios although they evaluated the reliability for present precipitation series. Although an SDM is well calibrated for generating present climate variables, it often requires additional strategies to produce plausible future climate scenarios from a host AOGCM future simulation.

In the present study, a multi-site hybrid SDM was employed to project future daily precipitation series. The SDM has been developed by combining the regression-based approach and the stochastic weather generator approach. The methods use multivariate multiple linear regression (MMLR), a stochastic randomization procedure, a Markov chain model and a gamma distribution to generate spatially correlated multi-site surface precipitation on a local region from global-scale reanalysis predictors. Hence, the

ultimate goal of this study is to project realistic daily precipitation series at multiple observation sites on a region based on the multi-site hybrid SDM and future atmospheric predictors driven by three AOGCMs. The multi-site hybrid SDM was applied on the observation sites located in the greater Montréal area, Québec, Canada. The three AOGCMs employed are the second and third generation of the Canadian coupled Global Climate Models (CGCM2 and CGCM3; e.g. Flato & Boer 2001; Kim *et al*. 2002; Kim *et al*. 2003, respectively) of the Canadian Centre for Climate Modeling and Analysis (CCCma), and the third version of Hadley Centre coupled Global Climate Model (HadCM3; e.g. Gordon *et al*. 2000; Pope *et al*. 2000) developed at the Hadley Centre in the United Kingdom. Impacts of climate change on extreme precipitation events are also analysed using the projected future precipitation daily time series.

This paper is organized as follows. The next section describes the methodologies. The third section describes the study area and the AOGCM predictors used. The fourth section summarizes the results of the study and the final section provides the discussion and conclusion.

METHODOLOGIES

Statistical downscaling model

The multi-site hybrid SDM generates precipitation occurrence and amount time series separately at multiple observation sites. Deterministic series for precipitation occurrence and amount, which are explained by global-scale atmospheric predictors, are simultaneously predicted by using a MMLR model for multiple-site application. Unexplained natural variability and cross-correlations among multiple regional sites are then modelled using a randomization procedure based on a multivariate normal distribution. For single site, several studies (e.g. von Storch 1999; Wilby *et al*. 2002; Hessami *et al*. 2008) coupled linear regression approaches and the randomization procedure to downscale daily precipitation from AOGCM atmospheric predictors. However, the present model additionally employs the stochastic weather generation technique (i.e. first-order Markov chain model and gamma distribution) to generate more realistic daily precipitation time series.

The deterministic series of daily probabilities of precipitation occurrence at multiple observation sites can be simultaneously predicted from reanalysis predictors (e.g. the National Centers for Environmental Prediction/National Center for Atmospheric Research, i.e. NCEP/NCAR, NC hereafter; e.g. Kalnay *et al*. 1996; Kistler *et al*. 2001), used to calibrate the SDM model, using the following MMLR equation:

$$\hat{\mathbf{O}}^{\mathrm{NC}} = \hat{\boldsymbol{\alpha}}_0 + \mathbf{X}^{\mathrm{NC}}\hat{\boldsymbol{\alpha}} \tag{1}$$

where $\hat{O}_{ij}^{\mathrm{NC}}$, an element of the matrix $\hat{\mathbf{O}}^{\mathrm{NC}}$, is the downscaled probability of precipitation occurrence on day i (= 1, 2, …, n) at site j (= 1, 2, …, m). $\mathbf{X}^{\mathrm{NC}}[n \times k]$ is the standardized NC predictor matrix, where each predictor was standardized with a mean of zero and a standard deviation of one. The constant term matrix $\hat{\boldsymbol{\alpha}}_0[n \times m]$ and the parameter matrix $\hat{\boldsymbol{\alpha}}[k \times m]$ can be estimated by the ordinary least squares (OLS) estimation method. Logistic regression is an option to model the occurrence of wet and dry days directly from the atmospheric predictors (e.g. Yang *et al*. 2005). However, the present model employed the MMLR to predict deterministic series of precipitation occurrence probability. The deterministic series were then transformed to [0 or 1] series by applying first-order Markov chain after adding random noise in the hybrid model. Several previous studies (e.g. Wilby *et al*. 2002; Wilby *et al*. 2003; Harpham & Wilby 2005; Haylock *et al*. 2006; Dibike *et al*. 2008; Hessami *et al*. 2008) employed linear regression approaches to predict precipitation occurrence probability from large-scale atmospheric predictors.

Residual matrix $\mathbf{E}_{\mathrm{O}}[n \times m]$ of the MMLR model for daily probability of precipitation occurrence is as below:

$$\mathbf{E}_{\mathrm{O}} = (\mathbf{O} - \hat{\mathbf{O}}^{\mathrm{NC}}) \tag{2}$$

where $\mathbf{O}[n \times m]$ is the observed binary (0 for dry day and 1 for wet day matrix of precipitation occurrence, i.e. defined as observed precipitation $\geq$1 mm/day). Generally, the probability matrix $\hat{\mathbf{O}}^{\mathrm{NC}}$ represents only a part of the at-site temporal variance and cross-site correlation of the observed matrix $\mathbf{O}$. To reproduce completely temporal variability and spatial dependency of the observed precipitation occurrence series, the random noise matrix $\tilde{\mathbf{E}}_{\mathrm{O}}[n \times m]$ is generated from a multivariate normal distribution having error means that

are equal to zero, and error covariance matrix [$\mathbf{\Sigma}_{\mathrm{O}}^{\mathrm{NC}} = \mathbf{S}_{\mathrm{O}}^{\mathrm{NC}}\mathbf{\Omega}_{\mathrm{O}}^{\mathrm{NC}}\mathbf{S}_{\mathrm{O}}^{\mathrm{NC}}$] equal to that of the residual matrix $\mathbf{E}_{\mathrm{O}}(\tilde{\mathbf{E}}_{\mathrm{O}} \sim \mathrm{N_m}(0, \mathbf{\Sigma}_{\mathrm{O}}^{\mathrm{NC}}))$, where $\mathbf{S}_{\mathrm{O}}^{\mathrm{NC}}$ is a diagonal matrix of standard deviations and $\mathbf{\Omega}_{\mathrm{O}}^{\mathrm{NC}}$ is a correlation matrix of the residual matrix $\mathbf{E}_{\mathrm{O}}$. Generated random noises are added to the predicted probability matrix of the precipitation occurrence $\hat{\mathbf{O}}^{\mathrm{NC}}$ as below:

$$\tilde{\mathbf{O}}^{\mathrm{NC}} = \hat{\mathbf{O}}^{\mathrm{NC}} + \tilde{\mathbf{E}}_{\mathrm{O}} \tag{3}$$

The continuous probability series of precipitation occurrence at each station in the matrix $\tilde{\mathbf{O}}^{\mathrm{NC}}$ have normal distribution; however, they convert to the [0 or 1] binary series by selecting an appropriate threshold. The hybrid SDM employed first-order Markov chain model for the transformation, where the probability of precipitation occurrence depends only on whether precipitation occurred or not on the previous day. The model employs two transition probabilities: (1) the probability of a wet day following a dry day (p_{01}) and (2) the probability of a wet day following a wet day (p_{11}). $\dot{\mathbf{O}}^{\mathrm{NC}}$ denotes the binary matrix of precipitation occurrence and $\dot{O}_t^{\mathrm{NC}}(k)$ is a [0 or 1] binary value of rainfall occurrence by the first-order Markov chain model at location k on a day t. The $\dot{O}_t^{\mathrm{NC}}(k)$ is then determined as:

$$\dot{O}_t^{\mathrm{NC}}(k)\begin{cases} 1, & \text{if } \tilde{O}_t^{\mathrm{NC}}(k) \geq \Phi^{-1}(k)[1 - p_{01}(k)] \text{ and } \dot{O}_{t-1}^{\mathrm{NC}}(k) = 0 \\ 1, & \text{if } \tilde{O}_t^{\mathrm{NC}}(k) \geq \Phi^{-1}(k)[1 - p_{11}(k)] \text{ and } \dot{O}_{t-1}^{\mathrm{NC}}(k) = 1 \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where $\Phi^{-1}(k)$ indicates the normal cumulative distribution function, which uses mean and standard deviation parameter estimates from the time series of $\tilde{O}^{\mathrm{NC}}(k)$. One problem during the [0, 1] transformation is that the transferred binary series $\dot{\mathbf{O}}^{\mathrm{NC}}$ cannot represent the original multi-site cross-correlation in the observed matrix $\mathbf{O}$ (Wilks 1998). To deal with this problem, the empirical relationship of cross-correlation between the binary series and the continuous series was derived by a simple power function and adjusted $\mathbf{\Omega}_{\mathrm{O}}^{\mathrm{NC}}$ matrices were used to generate the random noise matrix $\tilde{\mathbf{E}}_{\mathbf{O}}$.

Before calibrating the MMLR precipitation amount model, appropriate transformation should be performed because the precipitation amount vector $\mathbf{Y}_j$ for a site j is not normally distributed. If the vector $\mathbf{Y}_j$ has a gamma distribution, the distribution of $R_{ij} = {Y_{ij}}^{1/3}$ on a day i at a site j (where the R values are called Anscombe residuals) is normal as described by Terrell (2003) and Yang *et al.* (2005). The transformed precipitation amount matrix $\mathbf{R}$ $[n \times m]$ at multiple observation sites can be simultaneously predicted from NC reanalysis predictors using the following MMLR equation:

$$\hat{\mathbf{R}}^{\mathrm{NC}} = \hat{\boldsymbol{\beta}}_0 + \mathbf{X}^{\mathrm{NC}}\hat{\boldsymbol{\beta}} \tag{5}$$

where $\hat{\mathbf{R}}^{\mathrm{NC}}[n \times m]$ is the downscaled deterministic series matrix of transformed precipitation amount. Again, the $\hat{\boldsymbol{\beta}}_0[n \times m]$ and $\hat{\boldsymbol{\beta}}[k \times m]$ can be estimated by the OLS estimation method. The residual matrix $\mathbf{E}_{\mathrm{R}}[n \times m]$ of the MMLR model for daily precipitation amount can be calculated by $\mathbf{E}_{\mathrm{R}} = (\mathbf{R} - \hat{\mathbf{R}}^{\mathrm{NC}})$, where $\mathbf{R}[n \times m]$ is the Anscombe residuals matrix of the observed precipitation amount $\mathbf{Y}[n \times m]$ and $\hat{\mathbf{R}}^{\mathrm{NC}}[n \times m]$ is the predicted deterministic series of transformed precipitation amount. Again, the random noise matrix $\tilde{\mathbf{E}}_{\mathrm{R}}[n \times m]$ can be generated from the multivariate normal distribution having error mean equal to zero and error covariance matrix [$\mathbf{\Sigma}_{\mathrm{R}}^{\mathrm{NC}} = \mathbf{S}_{\mathrm{R}}^{\mathrm{NC}}\mathbf{\Omega}_{\mathrm{R}}^{\mathrm{NC}}\mathbf{S}_{\mathrm{R}}^{\mathrm{NC}}$] equal to that of the residual matrix $\mathbf{E}_{\mathrm{R}}(\tilde{\mathbf{E}}_{\mathrm{R}}\tilde{\mathrm{N}}_m(0, \mathbf{\Sigma}_{\mathrm{R}}^{\mathrm{NC}}))$ where $\mathbf{S}_{\mathrm{R}}^{\mathrm{NC}}$ is a diagonal matrix of standard deviations and $\mathbf{\Omega}_{\mathrm{R}}^{\mathrm{NC}}$ is a correlation matrix of the residual matrix $\mathbf{E}_{\mathrm{R}}$ to reproduce temporal variability and spatial dependency of the $\mathbf{R}$, and then is added to the modelled matrix $\hat{\mathbf{R}}^{\mathrm{NC}}$ as below:

$$\tilde{\mathbf{R}}^{\mathrm{NC}} = \hat{\mathbf{R}}^{\mathrm{NC}} + \tilde{\mathbf{E}}_{\mathrm{R}} \tag{6}$$

The generated precipitation amount series are inversely transformed by the equation $\tilde{Y}_{ij}^{\mathrm{NC}} = \tilde{R}_{ij}^{\mathrm{NC}^3}$, where $\tilde{Y}_{ij}^{\mathrm{NC}}$ is counted as a non-zero precipitation amount only when the occurrence model simulates a wet day ($\dot{Y}_{ij}^{\mathrm{NC}} = \dot{O}_{ij}^{\mathrm{NC}} \times \tilde{Y}_{ij}^{\mathrm{NC}}$). Generally, downscaled precipitation series in the matrix $\dot{\mathbf{Y}}^{\mathrm{NC}}$ have different statistical properties (e.g. mean and standard deviation) from the observed precipitation series for each observation site, because the Anscombe residuals $\mathbf{R}$ from the observed precipitation amount are not exactly normally distributed and, also, the residual matrix $\mathbf{E_R}$ of each

site is not normally distributed and has skew. Therefore, the downscaled precipitation amounts on wet days for each site in the $\dot{\mathbf{Y}}^{NC}$ are fitted to empirical distributions and their cumulative probabilities are calculated. Afterwards, the final precipitation amounts on wet days are recalculated using the calculated cumulative probabilities from the $\dot{\mathbf{Y}}^{NC}$ with the gamma distributions fitted to observation data for each site (this procedure is called the probability mapping approach in this study). Note that the final downscaled precipitation series ($\dot{\mathbf{Y}}^{NC}$) should be distinguished from the precipitation amount series ($\tilde{\mathbf{Y}}^{NC}$).

The gamma distribution is a special case of the Pearson (3) distribution in which the shape parameter is equal to zero and skewness (γ) is equal to twice the coefficient of variation (CV) (i.e. $\gamma = 2CV$). The probability density function of a gamma distribution is:

$$f(y) = \frac{(y/\beta)^{\alpha-1} \exp(-y/\beta)}{\beta\,\Gamma(\alpha)};\ \alpha,\ \beta > 0; \tag{7}$$

where α is the shape parameter, β is the scale parameter and $\Gamma(\cdot)$ is the gamma function. The mean (μ) and variance (σ^2) of the daily precipitation amounts on wet days are as below:

$$\mu = \alpha\beta \quad \text{and} \quad \sigma^2 = \alpha\beta^2 \tag{8}$$

Future projection methodology

MMLR models calibrated using NC reanalysis predictors are subsequently applied to predictors simulated by an AOGCM to predict deterministic series of precipitation occurrence and amount. However, AOGCM models do not reproduce perfectly observed atmospheric variables and their temporal distribution (i.e. as for NC reanalysis predictors), and there are often systematic differences between NC and AOGCM predictors series during the calibration or current period.

Standardization of reanalysis and AOGCM predictors is a widely used pre-processing technique in the regression-based statistical downscaling approach to reduce the systematic biases in the mean and variances of AOGCM outputs relative to the NC reanalysis data (see Wilby *et al.* 2004). Standardization may alleviate some of the problem, but systematic differences might still exist between the NC and AOGCM predictors' series. These systematic differences over the current period eventually propagate to the projected future precipitation occurrence and amount generated from future AOGCM predictor variables. Therefore, this study employed a bias adjustment factor (BAF) and a variance adjustment factor (VAF) to adjust differences of deterministic series of precipitation occurrence and amount between NC predictors and hosted AOGCM predictors during the calibration period. These two factors estimated from the calibration period are also applied to generate deterministic series of precipitation series for a future period to minimize the risk of propagating the systematic differences.

Deterministic series of precipitation occurrence probability $\mathbf{O}^{H}$ can be predicted by the calibrated MMLR (Equation (1)) and the standardized AOGCM historical predictors $\mathbf{X}^{H}[n \times k]$. From the $\hat{\mathbf{O}}^{NC}$ and $\mathbf{O}^{H}$, the BAF and VAF of each observation site can be estimated as below:

$$\text{BAF} = M^{NC} - M^{H} \tag{9a}$$

$$\text{VAF} = \frac{S^{NC}}{S^{H}} \tag{9b}$$

where M^{NC} and M^{H} are the mean values and S^{NC} and S^{H} are the standard deviations on a site in the predicted deterministic matrices $\hat{\mathbf{O}}^{NC}$ and $\mathbf{O}^{H}$. Adjusted precipitation occurrence deterministic series from AOGCM historical predictors can be produced using the two adjustment factors as below:

$$\hat{\mathbf{O}}_j^{H} = \mathbf{O}_j^{H} \times \text{VAF} + \text{BAF} \tag{10}$$

where $\mathbf{O}_j^{H}$ and $\hat{\mathbf{O}}_j^{H}$ are unadjusted and adjusted deterministic probability vectors of precipitation occurrence at site j, respectively. The continuous probability matrix of precipitation occurrence ($\tilde{\mathbf{O}}^{H}$) of an AOGCM historical predictor is produced by adding the multi-site random noise $\tilde{\mathbf{E}}_O$ to the adjusted deterministic probability matrix of precipitation occurrence $\hat{\mathbf{O}}^{H}$ ($\tilde{\mathbf{O}}^{H} = \hat{\mathbf{O}}^{H} + \tilde{\mathbf{E}}_O$). Again the continuous probability matrix of precipitation occurrence ($\tilde{\mathbf{O}}^{H}$) can be transformed to a binary matrix of precipitation occurrence ($\dot{\mathbf{O}}^{H}$) by Equation (4).

Precipitation occurrence scenarios for a targeted future period ($\dot{\mathbf{O}}^{F}$) can be projected using the following procedure. Firstly, deterministic probability series of precipitation

occurrence ($\mathbf{O}^{\mathrm{F}}$) are predicted by the calibrated MMLR (Equation (1)) using the biased future AOGCM predictors $\mathbf{X}^{\mathrm{F}}[n \times k]$ to the historical predictors $\mathbf{X}^{\mathrm{H}}$. Secondly, adjusted deterministic probability series matrix of precipitation occurrence ($\hat{\mathbf{O}}^{\mathrm{F}}$) is produced by applying the estimated BAFs and VAFs to the matrix $\mathbf{O}^{\mathrm{F}}$. Thirdly, the continuous probability matrix of precipitation occurrence ($\tilde{\mathbf{O}}^{\mathrm{F}}$) is produced by adding the multi-site random noise matrix $\tilde{\mathbf{E}}_{\mathrm{O}}$ to the matrix $\hat{\mathbf{O}}^{\mathrm{F}}$ ($\tilde{\mathbf{O}}^{\mathrm{F}} = \hat{\mathbf{O}}^{\mathrm{F}} + \tilde{\mathbf{E}}_{\mathrm{O}}$), which is conducted under the assumption that unexplained natural variability on each site and spatial coherence among multiple regional sites would remain constant in the future. Finally, the continuous probability matrix $\tilde{\mathbf{O}}^{\mathrm{F}}$ is transformed to the binary matrix ($\dot{\mathbf{O}}^{\mathrm{F}}$) by substituting $\tilde{O}^{\mathrm{NC}}(k)$ to $\tilde{O}_t^{\mathrm{F}}(k)$ in Equation (4). In the equation, the two transition probabilities p_{01} and p_{11} and their two thresholds $\Phi^{-1}(k)[1 - p_{01}(k)]$ and $\Phi^{-1}(k)[1 - p_{11}(k)]$ are fixed, therefore increased or decreased future probabilities of precipitation occurrence $\tilde{O}^{\mathrm{F}}(k)$ from $\tilde{O}^{\mathrm{NC}}(k)$ produce increased or decreased numbers of future wet and dry days, and also increased or decreased future transition probabilities.

Procedures for projecting historical and future scenarios of precipitation amount can be produced by following similar procedures as for precipitation occurrence. Firstly, the deterministic series matrices $\mathbf{R}^{\mathrm{H}}$ and $\mathbf{R}^{\mathrm{F}}$ are predicted by the calibrated MMLR (Equation (5)) and the AOGCM historical and future predictors $\mathbf{X}^{\mathrm{H}}$ and $\mathbf{X}^{\mathrm{F}}$. Secondly, BAF and VAF at each site are estimated from the $\hat{\mathbf{R}}^{\mathrm{NC}}$ and $\mathbf{R}^{\mathrm{H}}$, and then adjusted deterministic series matrices $\hat{\mathbf{R}}^{\mathrm{H}}$ and $\hat{\mathbf{R}}^{\mathrm{F}}$ are produced by applying the estimated BAFs and VAFs to the $\mathbf{R}^{\mathrm{H}}$ and $\mathbf{R}^{\mathrm{F}}$. Thirdly, $\tilde{\mathbf{R}}^{\mathrm{H}}$ and $\tilde{\mathbf{R}}^{\mathrm{F}}$ are produced by adding random noise matrix $\tilde{\mathbf{E}}_{\mathrm{R}}$ to the $\hat{\mathbf{R}}^{\mathrm{H}}$ and $\hat{\mathbf{R}}^{\mathrm{F}}$ (e.g. $\tilde{\mathbf{R}}^{\mathrm{F}} = \hat{\mathbf{R}}^{\mathrm{F}} + \tilde{\mathbf{E}}_{\mathrm{R}}$). Again, this is conducted under the assumption that the proportion of unexplained natural variability by the model at each site, as well as the spatial coherence among multiple sites would remain constant in the future. Fourthly, the generated precipitation amount matrices $\tilde{\mathbf{Y}}^{\mathrm{H}}$ and $\tilde{\mathbf{Y}}^{\mathrm{F}}$ are back-transformed from the $\tilde{\mathbf{R}}^{\mathrm{H}}$ and $\tilde{\mathbf{R}}^{\mathrm{F}}$ (e.g. $\tilde{Y}_{ij}^{\mathrm{F}} = \tilde{R}_{ij}^{\mathrm{F}^3}$), where the $\tilde{Y}_{ij}^{\mathrm{H}}$ and $\tilde{Y}_{ij}^{\mathrm{F}}$ are eventually counted as non-zero precipitation amounts when the precipitation occurrence model simulates the ith day to be wet (e.g. $\dot{Y}_{ij}^{\mathrm{F}} = \dot{O}_{ij}^{\mathrm{F}} \times \tilde{Y}_{ij}^{\mathrm{F}}$). Finally, to project more realistic future precipitation scenarios, the probability mapping technique is applied again, where predicted future gamma distribution is employed for the probability mapping technique. Downscaled precipitation amounts on wet days for each site in the matrix $\dot{\mathbf{Y}}^{\mathrm{F}}$ are fitted to gamma distributions and their cumulative probabilities are calculated. Afterwards, the final precipitation amounts on wet days are recalculated from the calculated cumulative probabilities from the matrix $\dot{\mathbf{Y}}^{\mathrm{F}}$ with the future gamma distributions, where the shape and scale parameters of the future gamma distributions are predicted based on the observation matrix ($\mathbf{Y}$) and downscaled historical and future precipitation amount matrices ($\dot{\mathbf{Y}}^{\mathrm{H}}$and $\dot{\mathbf{Y}}^{\mathrm{F}}$). The mean and standard deviation of the final future precipitation amounts on wet days for a site (μ_{F} and σ_{F}) and the shape and scale parameters of the future gamma distribution of future precipitation amounts (α_{F} and β_{F}) are predicted as below:

$$\mu_{\mathrm{F}} = \mu_{\mathrm{O}} \frac{M^{\mathrm{F}}}{M^{\mathrm{H}}},\ \sigma_{\mathrm{F}} = \sigma_{\mathrm{O}} \frac{S^{\mathrm{F}}}{S^{\mathrm{H}}} \tag{11a}$$

$$\alpha_{\mathrm{F}} = \frac{\mu_{\mathrm{F}}^2}{\sigma_{\mathrm{F}}^2},\ \beta_{\mathrm{F}} = \frac{\sigma_{\mathrm{F}}^2}{\mu_{\mathrm{F}}} \tag{11b}$$

where μ_{F} and σ_{F} are the predicted mean and standard deviation of the final future precipitation amounts, and α_{F} and β_{F} are the predicted shape and scale parameters of future gamma distribution on a site. μ_{O} and σ_{O}, M^{H} and S^{H}, and M^{F} and S^{F}, are the estimated mean and standard deviation of the precipitation amounts of the observation matrix ($\mathbf{Y}$), and downscaled historical and future precipitation amount matrices ($\dot{\mathbf{Y}}^{\mathrm{H}}$and $\dot{\mathbf{Y}}^{\mathrm{F}}$) on the same site, respectively. Figure 1 presents the procedure to calibrate the multi-site hybrid SDM and to generate precipitation series with NC reanalysis and with AOGCM historical and future predictors.

MODEL APPLICATION

Study area and predictand

This study employs daily precipitation series of the 10 observation sites of Environment Canada in the greater Montréal area, located in southern Québec (Canada). The selected area is the most populous region of Québec;

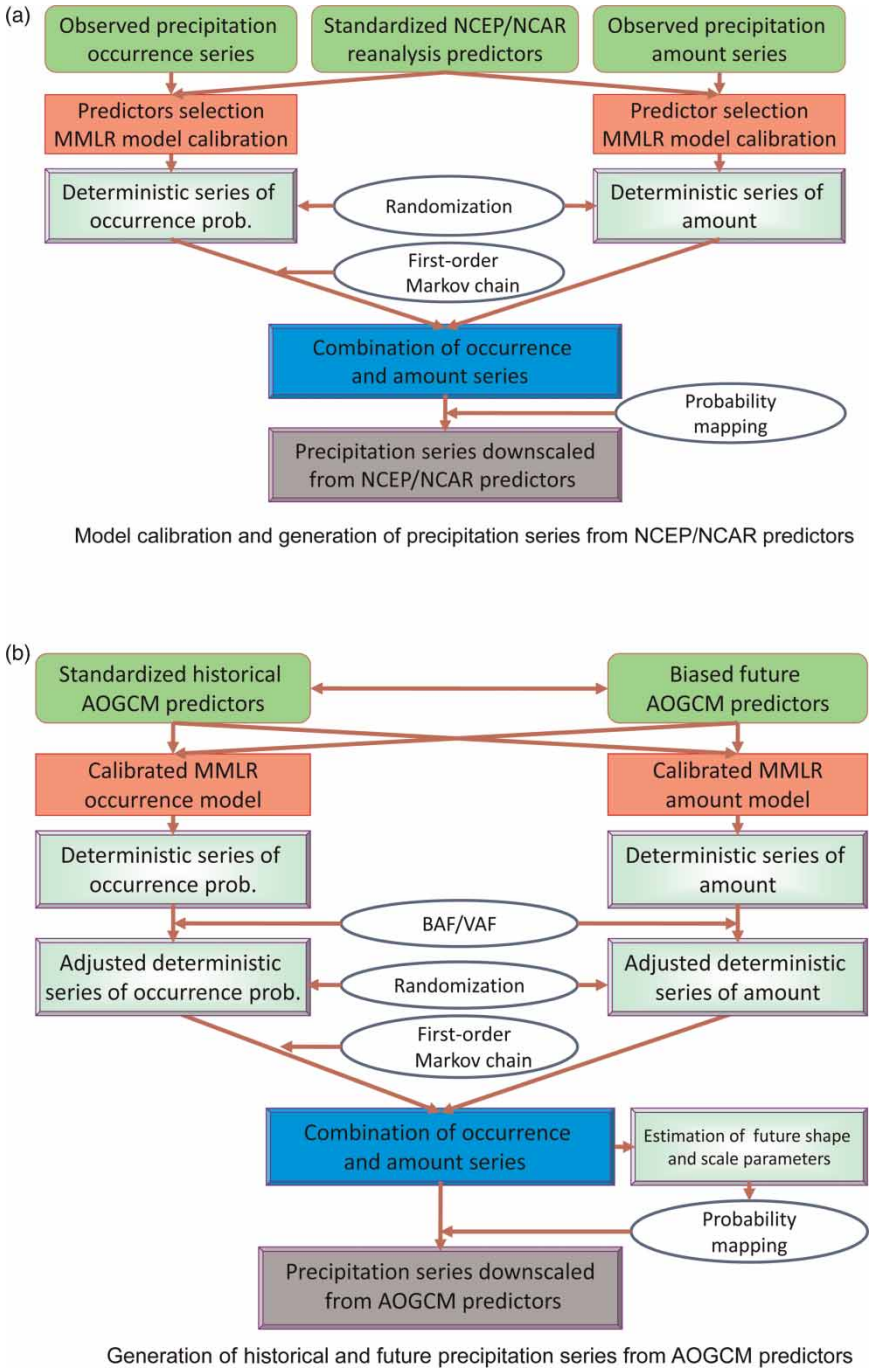

Figure 1 | Procedures of model calibration and generation of downscaled precipitation series.

hence, adaptation to climate change impact is of the utmost importance, especially as it relates to health, water resources and urban infrastructure issues. Figure 2 shows the study area with the meteorological stations near Montréal and grid points of the three AOGCMs (CGCM2, CGCM3 and HadCM3) located near the observation sites. Four grid points (GC1–GC4) of CGCM2 and CGCM3 models, which have the same horizontal

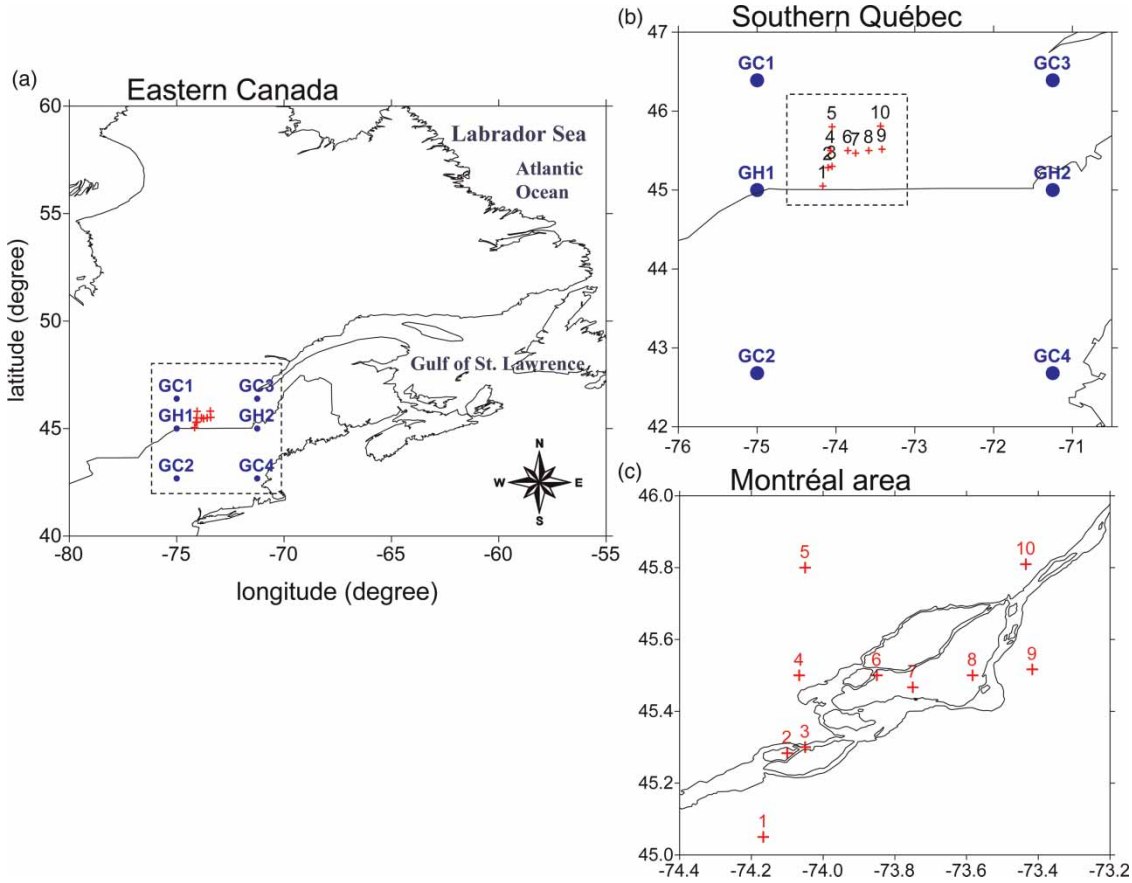

Figure 2 | Meteorological observation sites and grid pointes of three AOGCMs located near the observation sites. Numbers from 1 to 10 represent the meteorological observation sites (see the respective name of each station in Table 1). GC1–GC4 are grid points of CGCM2 and CGCM3 (i.e. same grid for both models), and GH1 and GH2 are grid points of HadCM3.

resolution or grid, are employed, while only two grid points (GH1 and GH2) of HadCM3 are in the vicinity of the observed stations. This last model has a higher resolution than the CGCM2/3 model (see details in the following section). Table 1 reports the names and locations (longitudes and latitudes) of the meteorological stations and of the AOGCMs' grid points. The distance between sites 1 and 10, the longest distance combination among the 10 observation sites, is 102 km.

AOGCM predictors

This study employed historical and future atmospheric predictors simulated from CGCM2, CGCM3 and HadCM3 models as inputs for the hybrid SDM to downscale historical and future daily precipitation series. The CGCM2 (the second-generation Coupled Global Climate Model) of the CCCma is an improved version of the previous CGCM1 (Flato *et al.* 2000). In the CGCM2, ocean mixing parameterization scheme, ocean spin up and flux adjustment procedures have been improved (Flato & Boer 2001). CGCM2 provides simulated results at 10 vertical levels on a surface grid resolution of 3.75° latitude × 3.75° longitude under the IPCC SRES (Special Report on Emissions Scenarios) GHGA emission scenarios (IPCC 2000; Nakicenovic *et al.* 2000). This study employed future atmospheric scenarios of CGCM2 under the B2 and A2 IPCC SRES scenarios.

CGCM3.1 is a subsequent version of the previous CGCM2 AOGCM. CGCM3.1 is run at T47 (which provides the same horizontal grid resolution as the CGCM2). It provides simulated outputs at 31 vertical levels. This study used future CGCM3 with the A1B and A2 IPCC SRES scenarios.

Table 1 | Observation sites and grid points of AOGCMs in the Montréal area (see Figure 2)

	No.	Name of stations or grid points	Latitude (°N)	Longitude (°W)
Observation	1	Huntingdon	45.05	74.17
	2	Valleyfield	45.28	74.10
	3	Les Cedres	45.30	74.05
	4	Oka	45.50	74.07
	5	St Jerome	45.80	74.05
	6	Ste Genevieve	45.50	73.85
	7	Montréal Airport	45.47	73.75
	8	Montréal McGill	45.50	73.58
	9	St-Hubert Airport	45.52	73.42
	10	L'assomption	45.81	73.43
CGCM2 & CGCM3	1	GC1	46.39	75.00
	2	GC2	42.68	75.00
	3	GC3	46.39	71.25
	4	GC4	42.68	71.25
HadCM3	1	GH1	45.00	75.00
	2	GH2	45.00	71.25

HadCM3 was developed in the United Kingdom by the Hadley Centre and is described in Gordon *et al*. (2000) and Pope *et al*. (2000). Unlike the previous CCCma models, it does not use surface flux adjustment procedures. The model provides an atmospheric component at 19 levels on a regular latitude-longitude grid of $2.5^{\circ} \times 3.75^{\circ}$. This study employed future HadCM3 atmospheric scenarios under the B2 and A2 IPCC SRES scenarios.

Historical predictors of the three AOGCMs cover the period from 1961 to 2000 and future predictors cover the period from 2001 to 2099 for HadCM3 or until 2100 for CGCM2/3. Near observed global-scale predictors are issued from the NC reanalysis data (e.g. Kalnay *et al*. 1996; Kistler *et al*. 2001) for the period from 1961 to 2000 to calibrate the SDM model. As the NC reanalysis data have been produced on a different grid from the AOGCMs, i.e. over a 2.5° latitude $\times 2.5^{\circ}$ longitude regular grid, an interpolation procedure of the predictor data series onto the respective AOGCM grid resolutions has been completed (see the DAI CGCM3 predictors 2008). Further details on the preparation of the NC reanalysis predictors and the equivalent AOGCMs predictors are available from the CCCSN (Canadian Climate Change Scenarios Network) project of Environment Canada (www.cccsn.ca) and the Data Access and Integration (DAI) website: http://loki.qc.ec.gc.ca/DAI/predictors-e.html.

In SDMs, a comprehensive procedure is necessary to select appropriate global-scale atmospheric variables as projected surface predictands are highly sensitive to the combination of predictor variables. The optimum combination of predictors among the potential NC predictor variables for the downscaling of daily precipitation occurrence and amount over the 10 observation sites is determined by the backward stepwise regression method (as in Hessami *et al*. 2008). The MMLR occurrence and amount models were calibrated for each month and the best combination of predictors was identified individually for each month. Table 2 reports the selected NC predictors for the two CGCM2 models, where numbers in parentheses refer to the number of times that each predictor has been selected among the 48 possibilities (i.e. four grid points from GC1 to GC4 and 12 months). For each month, 6–11 and 7–17 variables were selected as predictors for precipitation occurrence and amount, respectively. The frequently selected predictors for both precipitation occurrence and

Table 2 | Selected NC predictors for CGCM2 model by backward stepwise approach under the assumption that the predictors are independent. Numbers in parentheses refer to the number of times that each predictor has been selected among the 48 possibilities (the four grid points (GC1–GC4) and twelve months)

Precipitation occurrence	Precipitation amount
Mean sea level pressure (15)	500 hPa specific humidity (20)
500 hPa specific humidity (15)	Mean sea level pressure (15)
850 hPa geopotential (9)	1,000 hPa U-component (10)
1,000 hPa U-component (8)	500 hPa U-component (10)
1,000 hPa V-component (8)	1,000 hPa V-component (9)
1,000 hPa wind speed (6)	1,000 hPa wind speed (8)
850 hPa U-component (6)	850 hPa U-component (7)
1,000 hPa specific humidity (6)	500 hPa geopotential (6)
500 hPa U-component (5)	850 hPa wind speed (6)
850 hPa V-component (5)	500 hPa V-component (5)
500 hPa V-component (4)	500 hPa divergence (5)
850 hPa divergence (4)	850 hPa geopotential (5)
500 hPa geopotential (3)	850 hPa divergence (5)
850 hPa wind speed (3)	1,000 hPh specific humidity (5)
500 hPa divergence (2)	850 hPa V-component (4)
850 hPa specific humidity (2)	850 hPa specific humidity (3)
1,000 hPa divergence (1)	1,000 hPa divergence (1)
	500 hPa wind speed (1)

amount were 500-hPa specific humidity, mean sea level pressure and 1,000-hPa atmospheric circulation variables (U- and V-components, and wind speed). For CGCM3 and HadCM3 models, similar predictor sets, not provided in this text, were selected by the backward stepwise approach. IPCC (2001) stated that stepwise regression methodology to select appropriate NC atmospheric predictors might exclude predictor variables that may be crucial for climate change. The selection presented in Table 2, however, includes several predictors that are known to be sensitive to potential changes in climate conditions. They include both thermodynamic and dynamic physical processes that influence specific humidities, geopotentials and wind fields at different pressure levels within various levels in the atmosphere. Hence, both processes and their involvement in the precipitation occurrence and amounts are accounted for (e.g. Wilby & Wigley 2000; Widmann *et al*. 2002; Salathé 2003; Dibike *et al*. 2008). The MMLR parameters for daily precipitation occurrence and amount series were calibrated using predictands and predictors from the 1961 to 1990 period.

Diagnostic criteria and extreme indices

This study employed the same diagnostic criteria and extreme indices for daily precipitation series used by Hessami *et al*. (2008). They are shown in Table 3, in which PRCP1 (percentage of wet days using the threshold of $P \geq 1$ mm/day to define wet vs. dry days) and SDII (mean precipitation amount per wet days) represent occurrence and mean intensity indices, as CDD (maximum number of consecutive dry days) represent dry sequences length, and R3days (maximum 3-days' precipitation total) and PRCE90 (90th percentile of wet-day amount) are extreme indices. These indices are calculated seasonally (spring, i.e. MAM, summer, i.e. JJA, autumn, i.e. SON, and winter, i.e. DJF) to compare statistical characteristics between observed and downscaled precipitation series during the historical period from 1961 to 2000. They are also employed to analyse future daily precipitation series during the targeted period from 2060 to 2099.

Table 3 | Diagnostic and extreme indices employed to analyse climate change impacts to future daily precipitation occurrence and amount series (*Source*: Hessami *et al*. 2008)

Indices	Definition	Unit	Time scale
PRCP1	Percentage of wet days	%	Season
SDII	Mean precipitation amount at wet days	mm/day	Season
CDD	Maximum number of consecutive dry days	days	Season
R3days	Maximum 3-days' precipitation total	mm	Season
PREC90	90th percentile of daily precipitation amount	mm/day	Season

RESULTS

Historic and future precipitation time series projected by the hybrid SDM with AOGCMs predictors are presented and discussed in this section. For stability and robustness of the stochastic randomization procedure of the hybrid SDM, the multi-site hybrid SDM generates 50 realizations of precipitation occurrence and amount series over the current 1961 to 2000 and the future 2060 to 2099 periods. Therefore, results of the hybrid model presented in this section were evaluated from the 50 realizations of daily precipitation occurrence and amount series. Note that the daily MMLR and the hybrid models were calibrated separately for each month from 1961 to 1990, using the selected NC predictors.

MMLR and adjustment factors

Figure 3 reports percentages of explained variance with respect to observed precipitation occurrence and amount series by deterministic series predicted by MMLR models using NC predictors during the calibration period. The percentages varied from 20 to 47% for precipitation occurrence, and from 21 to 54% for precipitation amount, which is quite respectable considering the stochastic character of daily precipitation. Percentages of explained variance were larger in spring and in October than in the other months, with smaller values in summer for both precipitation occurrence and amount. Spatially, sites 7–10, which are located in the north-eastern part of the study area, yielded larger explained variance than the other observation sites. In winter, the spatial variability of R^2 between stations

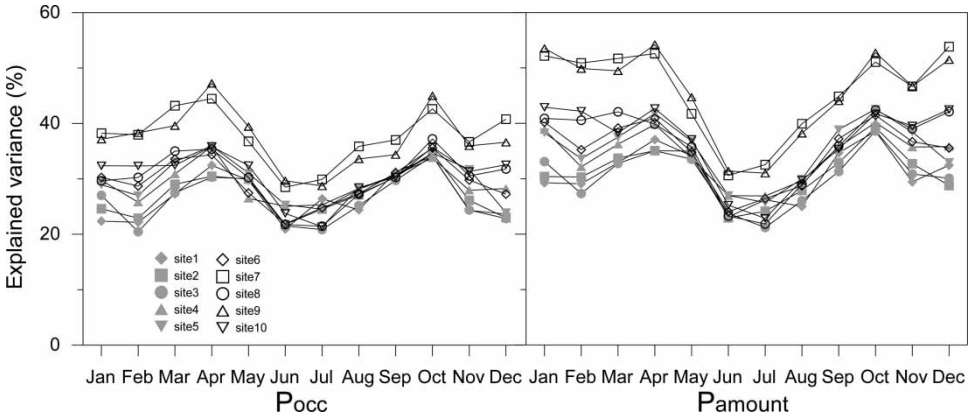

Figure 3 | Percentages of explained variances of precipitation occurrence (P_{occ}) and amount (P_{amount}) by MMLR model with NC predictors for every 10 observation sites and 12 months for the calibration period from 1961 to 1990.

is larger than in the other seasons, suggesting that the synoptic scale features of the atmospheric variables variously affect both the precipitation occurrence and total amount across the region. In spite of the fact that this last season is more affected by synoptic scale forcing than the other seasons, other regional scale effects also influence the local precipitation regime, as those are not explicitly incorporated in the large-scale variables from NCEP/NCAR.

Figure 4 shows the relative biases between the deterministic component series downscaled by MMLR with the NC predictors, and by the MMLR with the historical predictors of the three AOGCMs for the calibration period from 1961 to 1990. A negative value of relative bias implies the mean of the deterministic series of an AOGCM is larger than that of a series of NC. The relative biases varied in a range from –47 to 36% for each month and each site. The relative biases of the three AOGCMs showed quite different pattern and different seasonal variability. Especially, the relative biases of CGCM2 and HadCM3 showed stronger monthly variability than the corresponding CGCM3 predictor variables.

Figure 5 shows the VAFs between deterministic series by MMLR with the NC predictors and by MMLR with the historical predictors of the three AOGCMs for the calibration period from 1961 to 1990. A VAF smaller than one implies the standard deviation of the deterministic series of the AOGCM is larger than that of series of NC. The VAFs varied in a range from 0.75 to 1.34 for each month and each site. The VAF values of the three AOGCMs also showed quite different monthly patterns and variability. From Figures 4 and 5, it is clear that the deterministic series generated by MMLR with AOGCMs predictors could not properly reproduce the basic characteristics (i.e. mean and standard deviation) of those generated by NC predictors, in spite of the standardization procedure.

Figure 6 presents an example of the effect of the difference of the deterministic component series downscaled by the historical predictors of HadCM3 for the historical period from 1961 to 2000 versus the future downscaled precipitation series for the future period future period from 2060 to 2099. In the figure, box-whiskers plots of observed SDII, R3days and PREC90 (obs.), downscaled historical series using NC predictors (NCEP), HadCM3 predictors (H3_H) and by future HadCM3 A2 predictors (H3_A2) in winter at site 7 (Montréal airport) were compared. For the box-whisker plots, outliers were identified when a value fell below or above 1.5 IQR (inter-quartile range). For H3_H and H3_A2, box-whiskers plots of non-adjusted and adjusted series are compared. The non-adjusted series of H3_H tended to over-estimate the three indices while the adjusted series of H3_H reproduced the observed values of the three indices and those of NCEP fairly well. The systematic errors in the non-adjusted series of H3_H were propagated to the non-adjusted future series of H3_A2, thus the non adjusted series of H3_A2 yielded larger values than the adjusted series of H3_A2 for the three indices. Note that the deterministic series from HadCM3 have larger standard deviation than those with NC for winter season (see Figure 5). The adjusted deterministic series,

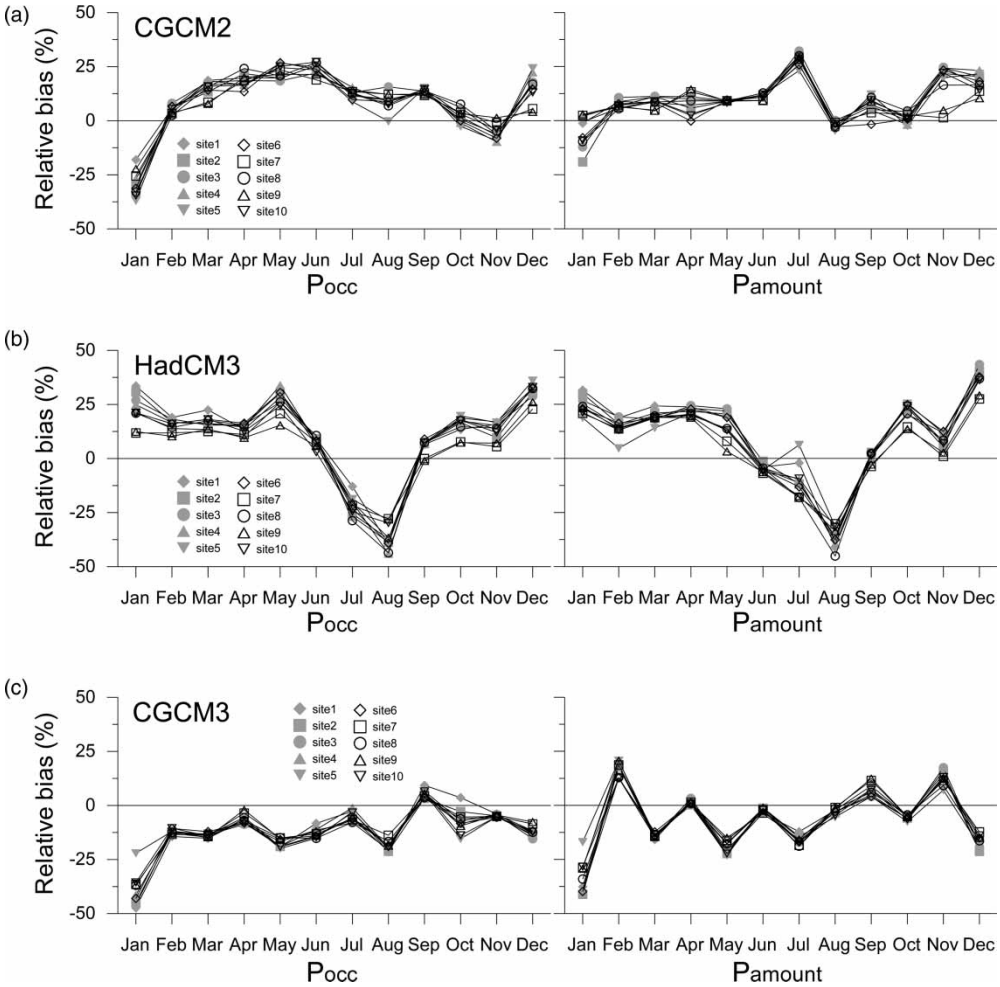

Figure 4 | Relative biases derived from deterministic component series predicted by MMLR with the NC predictors and by MMLR with the three AOGCMs' historical predictors for the 10 observation sites for all 12 months for the calibration period from 1961 to 1990.

therefore, are used to generate historical and future precipitation series from the historical and future predictors of the AOGCMs.

Spatial coherence among sites

Figure 7 reports cross-site correlation coefficients between pairs of precipitation occurrence (a) and amount (b) for observation (Obs.), deterministic series of MMLR without the stochastic component, historical series of the multi-site hybrid SDM (C2_H, H3_H and C3_H) and historical series of the single-site hybrid SDM (C2_H (S), H3_H (S) and C3_H (S)) as a function of inter-station distances for all possible combinations of station pairs for the historical period from 1961 to 2000. The single-site hybrid SDM has basically the same procedure as the multi-site hybrid SDM but does not have any consideration for spatial coherence (i.e. it adds uncorrelated random noise at the deterministic series for each site). The results of the single-site hybrid SDM connotatively show the reproduction ability of spatial coherence when precipitation series are downscaled independently at each site. Table 4 presents components included in the three models. In the figure, the deterministic precipitation (occurrence and amount) series showed large correlation coefficients for all the station pairs. Downscaled historical series by multi-site hybrid SDM with the three AOGCMs' predictors reproduced the correlation coefficients relatively well, i.e. they were quite compatible with

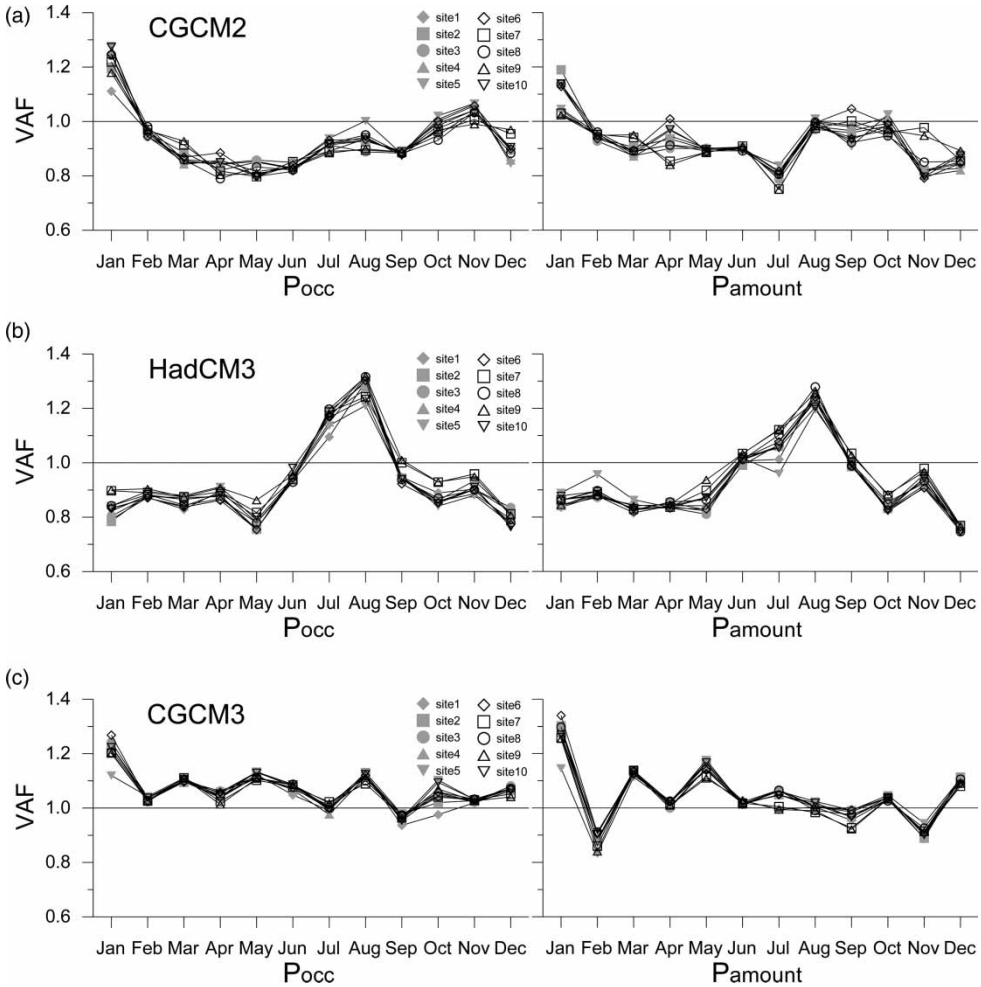

Figure 5 | VAFs (variance adjustment factors) derived from deterministic component series predicted by MMLR with the NC predictors and by MMLR with the three AOGCMs' historical predictors for the 10 observation sites for all 12 months for the calibration period from 1961 to 1990.

respect to observed series. However, the cross-site correlation coefficients among sites were systematically underestimated in the precipitation series generated by the single-site hybrid SDM.

Reproduction ability of spatial coherence of precipitation series is important, especially for hydrological analysis (e.g. flood frequency analysis on a basin). Total daily precipitation amounts in a targeted region or drainage basin can be preserved only when the spatial coherence of precipitation series on the multiple observation sites in a drainage basin is appropriately represented. Figure 8 presents box-whiskers plots for SDII, R3days and PREC90 of regionally averaged (from the results over the 10 stations) daily precipitation amount series of observation (Obs.), historical series of the multi-site hybrid SDM (C2_H, H3_H and C3_H) and historical series of the single-site hybrid SDM (C2_H (S), H3_H (S) and C3_H (S)) in summer. The historical series of the regionally averaged daily precipitation of the three AOGCMs generated by the multi-site hybrid SDM reproduced quite well the three indices, while those by single-site hybrid SDM showed underreproduction ability for these indices.

Future projection results

Table 5 provides average median values of rainfall indices from observations and downscaled historical series from the results over the 10 sites. Differences in the average

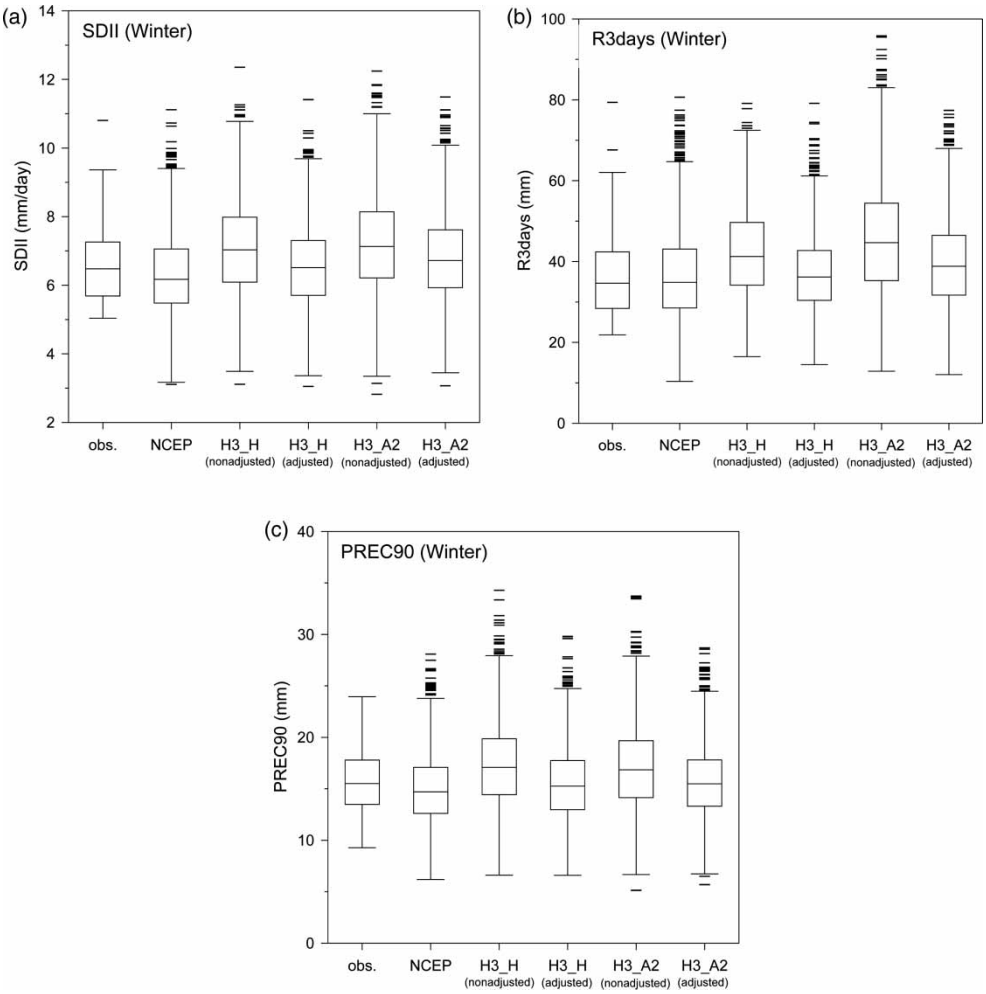

Figure 6 | Comparison of box-whiskers plots of SDII, R3days and PREC90 for observations (obs.), downscaled historical series by NC predictors (NCEP), by historical HadCM3 predictors (H3_H) and by future A2 HadCM3 predictors (H3_A2) for winter at site 7 (Montréal airport) where 'non-adjusted' and 'adjusted' refer to whether the series are adjusted by the BAF and VAF or not. The historical and future periods are 1961–2000 and 2060–2099, respectively.

median values between projected future series and their downscaled historical series are also presented in Table 5. Generally, the downscaled precipitation values with NC predictors, which are not provided in the table, produce better representation for the diagnostic and extreme indices with respect to observations than those with historical AOGCMs predictors.

Downscaled historical series at 10 observation sites tended slightly to under-represent the medians of observed PRCP1; however, differences of medians between observations and downscaled historical series were less than 2.0%, which is approximately 2 days per season. The downscaled historical series at 10 observation sites also reproduced the medians of observed SDII, CDD, R3days and PREC90 quite well and the differences between observations and downscaled historical series were less than 0.1 mm/day for SDII, 1 day for CDD, 2.8 mm for R3days and 1.6 mm/day for PREC90.

For the projected PRCP1 index, the six future series yielded a consistent increase among models in winter with a range of 2.0–7.2%, approximately 2–6 days, compared to their downscaled historical series. The future series of CGCM2 (i.e. C2_B2 and C2_A2) projected larger increases than those of HadCM3 (i.e. H3_B2 and H3_A2) and CGCM3 (i.e. C3_A1B and C3_A2). For the other seasons, AOGCMs future series showed no consistent projection results for this wet day index. CGCM2 future series also produced larger values than their historical counterparts in the

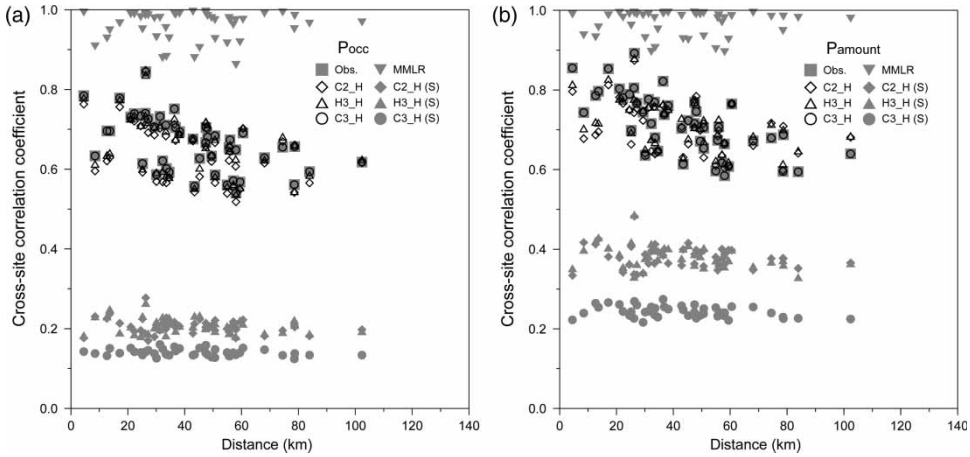

Figure 7 | Annual cross-site correlation coefficients between pairs of precipitation occurrence series (a) and precipitation amount series (b) of observation (Obs.), MMLR, historical series of the multi-site hybrid SDM (C2_H, H3_H and C3_H) and historical series of the single-site hybrid SDM (C2_H (S), H3_H (S) and C3_H (S)) versus their station distances for all possible combinations of station pairs for the historical period from 1961 to 2000.

Table 4 | Components of MMLR, single-site and multi-site hybrid models

Model	Components
MMLR	Deterministic series
Single-site hybrid	Deterministic series + uncorrelated random noise
Multi-site hybrid	Deterministic series + spatially correlated random noise

other seasons. HadCM3 future series showed decreased values in spring and summer but almost the same values as their historical series in autumn for this index. CGCM3 future series projected larger values in spring but smaller values than their historical series in summer and autumn. Generally, the future series projected by A2 scenarios yielded larger signal than those by B2 and A1B scenarios for each AOGCM.

For the projected SDII index, the six future series yielded a consistent increase among models for all seasons, except for HadCM3 and CGCM3 in summer, which suggest no change or small decrease in mean intensity per wet day. In winter, the future series of the three AOGCMs yielded an increase of 0.1–2.0 mm/day. CGCM3 future series projected the largest increase in this season, while those of HadCM3 showed the largest decrease in summer for this index. HadCM3 future series obtained larger signals in spring and autumn than in winter.

For the projected CDD index, the six future series yielded a consistent decrease of l–2 days among models in winter, i.e. consistent with the suggested increase of wet days (PRCP1). In contrast, the future series of HadCM3 and CGCM3 obtained larger values (1–2 days more) than their historical series for CDD in summer. For the other seasons, no consistent signals and/or no change are suggested.

For the projected R3days and PREC90 extreme indices, the six future series yielded a consistent increase among models for all seasons, except for HadCM3 and CGCM3 in summer. The future series of CGCM2 and CGCM3 projected a larger increase in winter and spring than in autumn, while those of HadCM3 showed a larger increase in autumn than the other seasons. Again, in most cases, the A2 runs provide higher signals than those from B2 and A1B of the same model.

Figure 9 presents box-whiskers plots of PRCP1, SDII, CDD and R3days of observation (Obs.) and downscaled historical (C2_H, H3_H and C3_H) and projected future (C2_B2, H3_B2, C3_A1B, C2_A2, H3_A2 and C3_A2) precipitation series at site 7 (Montréal airport). The results were provided only in winter, which generally showed the largest signal among the six future series and seasons. For all indices, downscaled historical series (C2_H, H3_H and C3_H) reproduced this observed index quite well. The future series projected by A2 scenarios generally yielded larger increases than those by B2 and A1B scenarios for

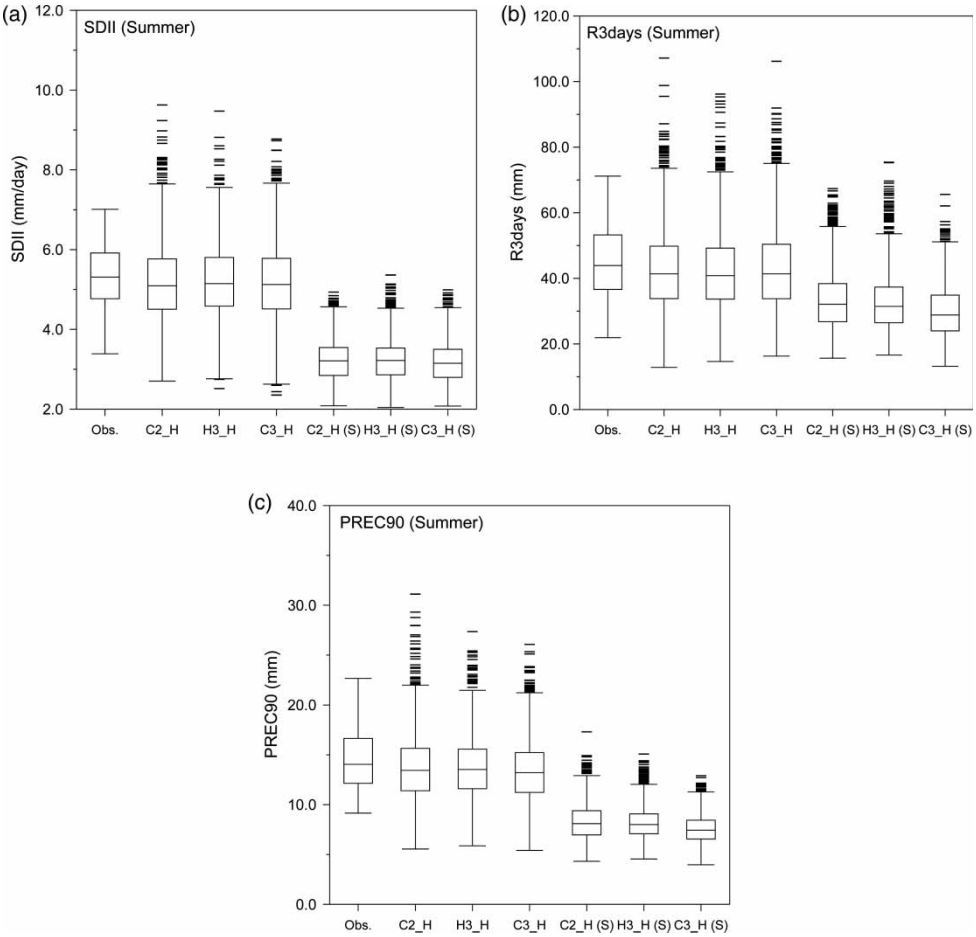

Figure 8 | Box-whiskers plots of SDII, R3days and PREC90 of regionally averaged daily precipitation amount series of observation (Obs.), historical series of the multi-site hybrid SDM (C2_H, H3_H and C3_H) and historical series of the single-site hybrid SDM (C2_H (S), H3_H (S) and C3_H (S)) in summer for the historical period from 1961 to 2000.

each AOGCM. The future series of CGCM2 and CGCM3 projected larger increases than those of HadCM3 for the PRCP1, SDII and R3days.

Table 6 provides the average values of seasonal and annual precipitation amount for observed and downscaled precipitation series from the 10 observation sites. Percentages of increase/decrease of the average values in the projected future series with respect to historical series are also provided. The three historical series tended slightly to under-estimate the annual and seasonal total amount of precipitation, by 1.4–7.5% at seasonal scale and 3.3–4.9% at annual scale. Among the three historical series, C2_H reproduced the smallest values for the annual amount. In the future, a significant and consistent increase of total precipitation is suggested in winter (and in spring and autumn) by the three AOGCMs, whereas a significant decrease is shown in summer by the HadCM3 and CGCM3, as for the SDII index. In the case of CGCM2, an increase of total precipitation is projected in the future for all seasons. Among the six future series, C2_A2 and C3_A2 showed the largest increase of total amount in winter while H3_B2 and H3_A2 showed the largest decrease in summer. At the annual scale, CGCM2 and CGCM3 future precipitation series projected a consistent increase of total precipitation for all scenarios, as the HadCM3 suggests a small consistent decrease.

SUMMARY AND CONCLUSION

Once an SDM has been appropriately calibrated to downscale a predictand using reanalysis atmospheric predictors,

Table 5 | Average median values of diagnostic indices from observations and downscaled historical series at 10 sites and differences in medians between projected future series and their downscaled historical series

						Difference between projected future series and downscaled historical series					
		Observation and downscaled historical series				B2		A1B	A2		
Statistics and indices	Season	Obs.	$C2_{\_H}$	$H3_{\_H}$	$C3_{\_H}$	$C2_{\_B2}$–$C2_{\_H}$	$H3_{\_B2}$–$H3_{\_H}$	$C3_{\_A1B}$–$C3_{\_H}$	$C2_{\_A2}$–$C2_{\_H}$	$H3_{\_A2}$–$H3_{\_H}$	$C3_{\_A2}$–$C3_{\_H}$
PRCP1	Winter	36.5	34.8	34.5	34.8	3.8	2.0	2.6	7.2	2.7	2.6
(%)	Spring	32.2	30.6	30.8	31.1	0.2	–0.3	1.4	0.8	–1.4	2.9
	Summer	33.0	32.3	32.6	32.7	1.7	–4.0	–0.7	3.2	–5.3	–1.5
	Autumn	33.9	32.8	32.9	33.7	3.3	0.1	–1.3	3.5	0.0	–0.8
SDII	Winter	6.7	6.6	6.7	6.6	0.8	0.1	1.8	1.3	0.2	2.0
(mm/day)	Spring	7.5	7.3	7.4	7.4	0.7	0.4	0.8	0.9	0.5	0.9
	Summer	8.8	8.8	8.8	8.8	0.4	–0.6	0.0	1.1	–0.7	–0.1
	Autumn	8.2	8.1	8.1	8.1	0.5	0.3	0.6	0.6	0.3	0.4
CDD	Winter	9	10	10	10	–1	–1	–1	–2	–1	–1
(day)	Spring	11	11	12	12	0	0	0	0	0	–1
	Summer	10	10	10	10	0	1	1	–1	2	1
	Autumn	10	11	11	11	–1	0	1	–1	0	0
R3days	Winter	38.0	37.0	37.2	37.7	4.6	1.6	8.9	11.3	2.5	10.4
(mm)	Spring	38.7	38.2	39.3	39.9	5.9	1.6	4.2	8.7	0.4	5.8
	Summer	50.3	48.4	47.9	49.4	1.4	–3.1	–0.4	4.7	–4.0	–1.1
	Autumn	47.8	50.4	46.7	48.6	3.2	2.2	1.4	3.8	3.7	1.2
PREC90	Winter	15.7	15.6	15.6	15.5	1.0	0.2	2.7	1.8	0.4	3.2
(mm/day)	Spring	16.8	16.7	16.9	16.9	1.3	0.6	1.4	1.8	0.7	1.7
	Summer	20.9	21.1	21.2	21.2	0.5	–0.5	0.0	1.4	–0.5	–0.2
	Autumn	18.4	19.8	19.8	19.9	0.8	0.7	0.9	0.9	0.7	0.8

it requires additional strategies to project realistic future climate scenarios from AOGCM predictors. This study has provided detailed methods to project future precipitation series at multiple observation sites in a targeted region. These methods are based on a multi-site hybrid SDM. The hybrid SDM generates precipitation occurrence and amount series separately, and uses MMLR, stochastic randomization procedure, Markov chain model and gamma distribution to reproduce spatially correlated multi-site surface precipitation series over the region. The multi-site hybrid SDM was applied to project precipitation series over 10 observation sites located in the great Montréal area, Québec, Canada. Six projections of future precipitation series were produced using predictors from a subset of combination of three AOGCMs (CGCM2, CGCM3, HadCM3) and three IPCC SRES scenarios (B2, A1B and A2).

For transformation of future continuous probability series to the binary [0 or 1] occurrence series, the Markov chain model determined two thresholds estimated by two transition probabilities p_{01} and p_{11} of observed historical precipitation occurrence series. Therefore, increased or decreased future probabilities series of precipitation occurrence can produce an increased or a decreased number of future wet and dry days, and future transition probabilities. For the projection of future precipitation amounts, future shape and scale parameters of gamma distribution of the probability mapping technique were predicted from adjusted historical and future precipitation amount series (see Equation (11)).

The MMLR models with global-scale NCEP/NCAR reanalysis predictors explained 20–47% and 21–54% of variability of observed precipitation occurrence and

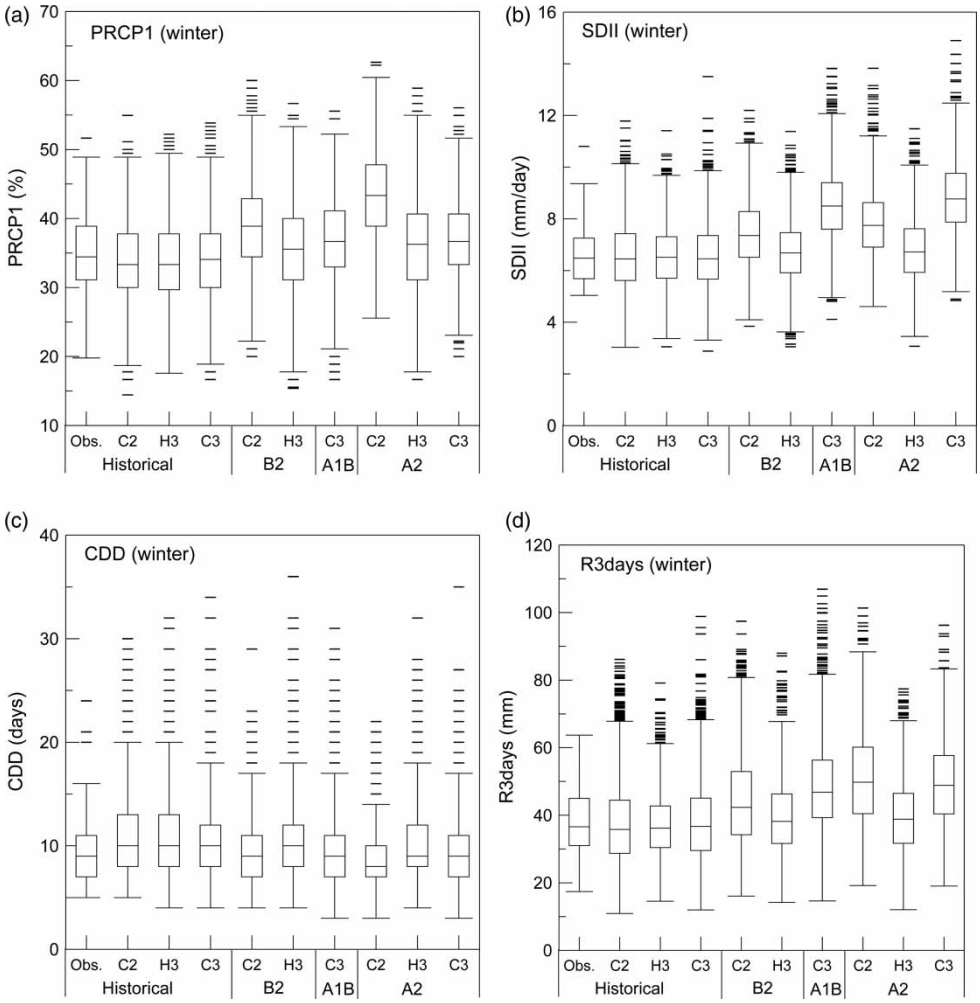

Figure 9 | Box-whiskers plots of PRCP1, SDII, CDD and R3days of observation (Obs.), downscaled historical series (C2_H, H3_H and C3_H), and projected future series (C2_B2, H3_B2, C3_A1B, C2_A2, H3_A2 and C3_A2) for winter at site 7 (Montréal airport).

Table 6 | Averaged total precipitations of observed and downscaled historical series at 10 observation sites and percentages of increase of the variables in the projected future series, where the percentages are calculated from projected future series and their downscaled historical series

					Percentages of increase in the projected future series (%)					
	Observation and downscaled historical series (mm)				B2		A1B	A2		
	Obs.	$C2_{\_H}$	$H3_{\_H}$	$C3_{\_H}$	$C2_{\_B2}$	$H3_{\_B2}$	$C3_{\_A1B}$	$C2_{\_A2}$	$H3_{\_A2}$	$C3_{\_A2}$
Winter	220.6	207.7	207.8	207.4	24.6	7.9	36.2	44.6	11.6	40.2
Spring	222.7	206.1	209.2	211.8	10.7	4.8	15.2	15.7	1.6	22.4
Summer	268.3	261.9	263.8	264.4	10.5	–17.8	–2.3	23.8	–23.0	–5.8
Autumn	251.7	240.7	242.9	248.1	17.2	4.0	2.8	18.6	4.1	3.0
Annual	963.3	916.4	923.6	931.7	15.5	–1.2	11.6	25.3	–2.5	13.2

amount series, for every 10 observation sites over twelve months. This level of explained variance has previously been reported by Hessami *et al.* (2008) over other sites in eastern Canada. Therefore, a randomization procedure is essential to incorporate the stochastic behaviour of precipitation and to analyse properly the statistical properties of its occurrence and amount. Although a large part of the natural variability of the precipitation occurrence and amount series were complemented by a randomization procedure, the downscaled precipitation series by the hybrid SDM (i.e. deterministic component plus random noise) might be more consistent with the changes of atmospheric outputs of a host AOGCM than with the weather typing and weather generator approaches, which are basically based on resampling and random generation methodologies, respectively.

Systematic differences between the deterministic series from NC predictors and from the AOGCMs predictors were found in the basic descriptive statistics (i.e. mean and standard deviation), although the standardized NC predictors were used for MMLR model calibration and for the AOGCMs predictors. Suggested bias and VAFs have the potential to remove these systematic differences on the deterministic component series of an AOGCM (vs. NCEP/NCAR), and then prevent the propagation of these systematic differences to the projected future precipitation series (see Figure 5). The adjusted deterministic component series were employed to generate historical and future precipitation series from the AOGCM predictors.

The multi-site hybrid SDM with historical AOGCMs predictors reproduced cross-site correlation coefficients of observed precipitation occurrence and amount series among the 10 observation sites. The suggested multi-site hybrid SDM therefore can be employed to analyse climate change impacts on a regional drainage basin, which usually requires precipitation scenarios that reproduce both spatial coherence among multiple observation sites and total precipitation amount over the region simultaneously.

Total amounts, statistics and extreme indices of observed and downscaled daily precipitation series by the multi-site hybrid SDM with historical (1961 to 2000) and future (2060 to 2099) periods were compared. For the total precipitation amount, the six future precipitation series suggest a consistent increase in winter precipitation of 7.9–44.6%, while the future series of HadCM3 and CGCM3 project a decrease in summer values of 2.3–23.0% with respect to historical series. CGCM2 showed increased values for all seasons. At the annual scale, CGCM2 and CGCM3 future precipitation series projected a consistent increase of total precipitation for all scenarios, as the HadCM3 suggests a small consistent decrease (i.e. ensemble range of –2.5–25.3%). These results are quite compatible with some of the projections suggested in the recent IPCC report (IPCC 2007). Based on a multi-model ensemble of 21 global model projections and A1B scenarios, Christensen *et al.* (2007) suggested a potential average increase of 11% with a range from 2 to 28% in winter, while the summer projected average increase was 1% with a range between –17 and 13% between 2080 and 2099, compared to the 1980–1999 period over Eastern North America (ENA). They also predicted a mean increase of 7% with the range from –3 to 15% of the annual total precipitation (see Table 11.1 of Christensen *et al.* 2007). Therefore, our projections for our targeted region suggest a larger increase in winter, and a larger decrease in summer for total precipitation amounts than those of Christensen *et al.* (2007) over the whole ENA region. Part of these differences could come from downscaled vs. no downscaled signal values (i.e. from raw AOGCM outputs from the work of Christensen *et al.* 2007), and from the size of the considered regions. Although our study area is in the ENA, their study area is much larger than ours and also includes a part of the Atlantic Ocean.

A consistent increase signal is suggested among all models and scenario runs in winter for all indices, as for occurrence of wet days, mean intensity per wet day and extreme values (i.e. R3days and PREC90). For summer, HadCM3 and CGCM3 downscaled runs yielded decreased values in summer for these indices, while those from CGCM2 show a small increase in mean intensity, but with a substantial increase in wet day occurrence. Some consensual signals are also obtained in spring and autumn with a small systematic increase in mean intensity per wet day and a larger increase in extreme indices, with no consistency for the other indices. Among the six future series, future precipitation series generated using CGCM2 A2 and CGCM3 A2 scenarios yielded the largest increase in winter while

those of HadCM3 B2 and A2 scenarios yielded the largest decrease in summer.

In future work, additional scenarios will be generated with the multi-site hybrid SDM and other future atmospheric predictors projected from different AOGCMs or RCMs. Our results imply that the structural differences between AOGCMs and different scenarios of GHGA can produce different projections of future precipitation series (e.g. Fowler & Ekström 2009; Semenov & Stratonovitch 2010), in spite of the fact that some consensual signals are systematically obtained in winter, and less in summer and/or in spring or fall. Therefore, multi-model ensemble approaches should be adopted to evaluate the range of potential convergence in climate change signals, and to quantify the effects of various sources of uncertainty at the regional scale. Furthermore, probabilistic and/or ensemble averaging projection methodologies of future precipitation series should be investigated based on the multi-model ensemble results (e.g. Giorgi & Mearns 2002; Murphy *et al*. 2007; Tebaldi & Knutti 2007), and in using a weighting procedure (e.g. van der Linden & Mitchell 2009; Eum *et al*. 2011), especially for the study of extreme statistics.

ACKNOWLEDGEMENTS

We acknowledge the financial support provided by the National Science and Engineering Research Council (NSERC) of Canada. The authors would like to acknowledge also the Data Access and Integration (DAI, see http://loki.qc.ec.gc.ca/DAI/) team for providing the predictors data and technical support, in particular Milka Radojevic for her help in preparing the predictor data. The DAI data download gateway is made possible through collaboration among the Global Environmental and Climate Change Centre (GEC3), the Adaptation and Impacts Research Section (AIRS) of Environment Canada and the Drought Research Initiative (DRI).

REFERENCES

Anandhi, A., Srinivas, V. V., Nanjundiah, R. S. & Nagesh Kumar, D. 2008 Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine. *Int. J. Climatol.* **28**, 401–420.

Bardossy, A. & Plate, E. J. 1992 Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resour. Res.* **28**, 1247–1259.

Bell, J. L., Sloan, L. C. & Snyder, M. A. 2004 Regional changes in extreme climatic events: a future climatic scenario. *J. Clim.* **17**, 81–87.

Bellone, E., Hughes, J. P. & Guttorp, P. 2000 A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts. *Clim. Res.* **15**, 1–12.

Buishand, T. A. & Brandsma, T. 1997 Comparison of circulation classification schemes for predicting temperature and precipitation in the Netherlands. *Int. J. Climatol.* **17**, 875–889.

Buishand, T. A. & Brandsma, T. 2001 Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling. *Water Resour. Res.* **37**, 2761–2776.

Bürger, G. & Chen, Y. 2005 Regression-based downscaling of spatial variability for hydrologic applications. *J. Hydrol.* **311**, 299–317.

Busuioc, A., Giorgi, F., Bi, X. & Ionita, M. 2006 Comparison of regional climate model and statistical downscaling simulations of different winter precipitation change scenarios over Romania. *Theor. Appl. Climatol.* **86**, 101–123.

Christensen, J. H., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, I., Jones, R., Kolli, R. K., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C. G., Räisänen, J., Rinke, A., Sarr, A. & Whetton, P. 2007 Regional climate projections. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor & H. L. Miller, eds). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

DAI CGCM3 Predictors 2008 *Sets of Predictor Variables Derived From CGCM3 T47 and NCEP/NCAR Reanalysis*. Version 1.1, April 2008, Data Access Integration team, Montreal, QC, Canada, 15 pp.

Dibike, Y. B., Gachon, P., St-Hilaire, A., Ouarda, T. B. M. J. & Nguyen, V. T. V. 2008 Uncertainty analysis of statistically downscaled temperature and precipitation regimes in Northern Canada. *Theor. Appl. Climatol.* **91**, 149–170.

Eum, H.-I., Gachon, P., Laprise, R. & Ouarda, T. B. M. J. 2012 Evaluation of regional climate model simulations versus gridded observed and regional reanalysis products using a combined weighting scheme. *Clim. Dyn.* **38** (7–8), 1433–1457.

Flato, G. M. & Boer, G. J. 2001 Warming asymmetry in climate change simulations. *Geophys. Res. Lett.* **28**, 195–198.

Flato, G. M., Boer, G. J., Lee, W. G., McFarlane, N. A., Ramsden, D., Reader, M. C. & Weaver, A. J. 2000 The Canadian centre for climate modeling and analysis global coupled model and its climate. *Clim. Dyn.* **16**, 451–467.

Fowler, H. J. & Ekström, M. 2009 Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes. *Int. J. Climatol.* **29**, 385–416.

Fowler, H. J., Kilsby, C. G., O'Connell, P. E. & Burton, A. 2005 A weather-type conditioned multi-site stochastic rainfall model for the generation of scenarios of climatic variability and change. *J. Hydrol.* **308**, 50–66.

Giorgi, F. & Mearns, L. O. 2002 Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the 'Reliability Ensemble Average' (REA) method. *J. Clim.* **15**, 1141–1158.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B. & Wood, R. A. 2000 The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Clim. Dyn.* **16**, 147–168.

Harpham, C. & Wilby, R. L. 2005 Multi-site downscaling of heavy daily precipitation occurrence and amounts. *J. Hydrol.* **312**, 235–255.

Haylock, M. R., Cawley, G. C., Harpham, C., Wilby, R. L. & Goodess, C. M. 2006 Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios. *Int. J. Climatol.* **26** (10), 1397–1415.

Hessami, M., Gachon, P., Ouarda, T. B. M. J. & St-Hilaire, A. 2008 Automated regression-based statistical downscaling tool. *Environ. Model. Softw.* **23**, 813–834.

Hewitson, B. C. & Crane, R. G. 2002 Self-organizing maps: applications to synoptic climatology. *Clim. Res.* **22**, 13–26.

Huth, R. 2002 Statistical downscaling of daily temperature in central Europe. *J. Clim.* **15** (13), 1731–1742.

Huth, R. 2004 Sensitivity of local daily temperature change estimates to the selection of downscaling models and predictors. *J. Clim.* **17** (3), 640–652.

Huth, R., Kliegrová, S. & Metelka, L. 2008 Nonlinearity in statistical downscaling: does it bring an improvement for daily temperature in Europe? *Int. J. Climatol.* **28**, 465–477.

IPCC 2000 *Special Report on Emissions Scenarios (SRES): A Special Report of Working Group III of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, UK.

IPCC 2001 Climate change 2001. In: *The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* (J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell & C. A. Johnson, eds). Cambridge University Press, Cambridge, NY, USA, UK and New York, NY, USA.

IPCC 2007 *Climate Change 2007 The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), Summary for Policymakers*. Available from: <http://www.ipcc.ch> (accessed on 10 March 2007).

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

Khalili, M., Brissette, F. & Leconte, R. 2009 Stochastic multi-site generation of daily weather data. *Stoch. Environ. Res. Risk Assess.* **23**, 837–849.

Kim, S. J., Flato, G. M., Boer, G. J. & McFarlane, N. A. 2002 A coupled climate model simulation of the Last Glacial Maximum, Part 1: transient multi-decadal response. *Clim. Dyn.* **19**, 515–537.

Kim, S. J., Flato, G. M. & Boer, G. J. 2003 A coupled climate model simulation of the Last Glacial Maximum, Part 2: approach to equilibrium. *Clim. Dyn.* **20**, 635–661.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., Dool, H., Jenne, R. & Fiorino, M. 2001 The NCEP/NCAR 50-year reanalysis. *Bull. Am. Meteorol. Soc.* **82** (2), 247–267.

Mehrotra, R. & Sharma, A. 2007 Preserving low-frequency variability in generated daily rainfall sequences. *J. Hydrol.* **345**, 102–120.

Murphy, J. M., Booth, B. B. B., Collins, M., Harris, G. R., Sexton, D. M. H. & Webb, M. J. 2007 A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philos. Trans. R. Soc. A* **365**, 1993–2028.

Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grübler, A., Jung, T. Y., Kram, T., La Rovere, E. L., Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H., Price, L., Raihi, K., Roehrl, A., Rogner, H.-H., Sankovski, A., Schlesinger, M., Shukla, P., Smith, S., Swart, R., van Rooijen, S., Victor, N. & Dadi, Z. 2000 Emissions Scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA, 599 pp.

Palutikof, J. P., Goodess, C. M., Wathkins, S. J. & Holt, J. 2002 Generating rainfall and temperature scenarios at multiple sites: example from the Mediterranean. *J. Clim.* **15** (24), 3529–3548.

Pope, V. D., Gallani, M. L., Rowntree, P. R. & Stratton, R. A. 2000 The impact of new physical parameterizations in the Hadley Centre climate model – HadCM3. *Clim. Dyn.* **16**, 123–146.

Qian, B., Corte-Real, J. & Xu, H. 2002 Multi-site stochastic weather models for impact studies. *Int. J. Climatol.* **22**, 1377–1397.

Salathé, E. P. 2003 Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. *Int. J. Climatol.* **23**, 887–901.

Schmidli, J., Goodess, C. M., Frei, C., Haylock, M. R., Hundecha, Y., Ribalaygua, J. & Schmith, T. 2007 Statistical and dynamical downscaling of precipitation: an evaluation and comparison of scenarios for the European Alps. *J. Geophys. Res.* **112**.

Semenov, M. A. & Barrow, E. M. 1997 Use of a stochastic weather generator in the development of climate change scenarios. *Clim. Change* **35**, 397–414.

Semenov, M. A. & Stratonovitch, P. 2010 Use of multi-model ensembles from global climate models for assessment of climate change impacts. *Clim. Res.* **41**, 1–14.

Tebaldi, C. & Knutti, R. 2007 The use of the multi-model ensemble in probabilistic climate projections. *Philos. Trans. R. Soc. A* **365**, 2053–2075.

Terrell, G. 2003 The Wilson-Hilferty transformation is locally saddle point. *Biometrika* **90**, 445–453.

Tolika, K., Maheras, P., Vafiadis, M., Flocas, H. A. & Arseni-Papadimitriou, A. 2007 Simulation of seasonal precipitation and raindays over Greece: a statistical downscaling technique based on artificial neural networks (ANNs). *Int. J. Climatol.* **27**, 861–881.

Trigo, R. M. & Palutikof, J. P. 2001 Precipitation scenarios over Iberia: a comparison between direct GCM output and different downscaling techniques. *J. Clim.* **14**, 4422–4446.

Tripathi, S., Srinivas, V. V. & Nanjundiah, R. S. 2006 Downscaling of precipitation for climate change scenarios: a support vector machine approach. *J. Hydrol.* **330**, 621–640.

van der Linden, P. & Mitchell, J. F. B. 2009 ENSEMBLES: climate change and its impacts: summary of research and results from the ENSEMBLES project. Met Office Hadley Centre, Exeter.

von Storch, H. 1999 On the use of 'inflation' in statistical downscaling. *J. Clim.* **12**, 3505–3506.

Widmann, M., Bretherton, C. S. & Salathé, E. P. 2002 Precipitation downscaling over the northwestern United States using numerically simulated precipitation as a predictor. *J. Clim.* **16**, 799–816.

Wilby, R. L. & Wigley, T. M. L. 2000 Precipitation predictors for downscaling: observed and general circulation model relationships. *Int. J. Climatol.* **20**, 641–661.

Wilby, R. L., Charles, S. P., Zorita, E., Timbal, B., Whetton, P. & Mearns, L. O. 2004 *Guidelines for use of climate scenarios developed from statistical downscaling methods*. Data Distribution Centre of the International Panel on Climate Change (IPCC) TGCIA, 27 pp.

Wilby, R. L., Dawson, C. W. & Barrow, E. M. 2002 SDSM – a decision support tool for the assessment of regional climate change impacts. *Environ. Model. Softw.* **17**, 147–159.

Wilby, R. L., Tomlinson, O. J. & Dawson, C. W. 2003 Multisite simulation of precipitation by conditional resampling. *Clim. Res.* **23**, 183–194.

Wilks, D. S. 1998 Multisite generation of a daily stochastic precipitation generation model. *J. Hydrol.* **210**, 178–191.

Wilks, D. S. 1999 Multisite downscaling of daily precipitation with a stochastic weather generator. *Clim. Res.* **11**, 125–136.

Yang, C., Chandler, R. E., Isham, V. S. & Wheater, H. S. 2005 Spatial-temporal rainfall simulation using generalized linear models. *Water Resour. Res.* **41**, W11415.

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

First received 24 May 2011; accepted in revised form 23 November 2011. Available online 13 August 2012