

Effects of variance adjustment techniques and time-invariant transfer functions on heat wave duration indices and other metrics derived from downscaled time-series. Study case: Montreal, Canada

Carlos F. Gaitán^{1,2,3}

Received: 25 November 2015 / Accepted: 19 May 2016 / Published online: 27 May 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Statistical downscaling techniques are often used to generate finer-scale projections of climate variables affected by local-scale processes not resolved by coarse resolution numerical models like global climate models (GCMs). Statistical downscaling models rely on several assumptions in order to produce finer-/local-scale projections of the variable of interest; one of these assumptions is the time-invariance of the relationships between predictors (e.g. coarse resolution GCM output) and the local-scale predictands (e.g. gridded observation-based time-series or weather station observations). However, in the absence of future observations, statistical downscaling studies use historical data to evaluate their models and assume that these historical simulation skills will be retained in the future. In addition, regression-based downscaling models fail to reproduce the observed variance, and hence their projections need to be adjusted accordingly. Two approaches are usually employed to perform this adjustment: randomization and variance inflation. Here, we study the effect of the stationarity assumption when downscaling daily maximum temperatures and using the downscaled information to estimate historical and future metrics like return periods and heat waves durations over Montreal, Canada; and the effect of the two variance adjustment techniques on the historical and future time-series. To do so, we used regional climate model (RCM) output from the Canadian RCM 4.2, as proxies of historical and future local climates, and daily maximum temperatures obtained from the Canadian GCM 3.1. The results show that the root-mean-squared errors between the pseudo-observations and the statistically downscaled time-series (historical and future)

Electronic supplementary material The online version of this article (doi:[10.1007/s11069-016-2381-2](https://doi.org/10.1007/s11069-016-2381-2)) contains supplementary material, which is available to authorized users.

✉ Carlos F. Gaitán
carlos.gaitan@ou.edu

¹ South Central Climate Science Center, Norman, OK, USA

² University of Oklahoma, Norman, OK, USA

³ NOAA – GFDL, 201 Forrestal Road, Princeton, NJ 08540, USA

varied over time, with higher errors in the future period; and the effects of randomization and variance inflation on the tails of the statistically downscaled time-series.

Keywords Heat wave duration index · Daily maximum temperature · Transfer functions · Statistical downscaling · Stationarity

1 Introduction

Global climate models (GCMs) are numerical representations of the main chemical, physical and biological components of the global climate system. These can be used to simulate historical climates and to project future climates under different emission scenarios—now representative concentration pathways (RCPs)—resulting from different assumptions about socio-economical trends (IPCC 2000). However, currently there is a necessity to generate finer-scale projections of climate variables at a higher spatio-temporal resolution, as users from diverse communities and disciplines require information at a spatial scale that is not provided by the current generation of GCMs (Franklin et al. 2013; Semenov and Stratonovitch 2010; Stock et al. 2011).

Among the typical users of high-resolution data, we can find hydrologists (e.g. Anandhi et al. 2008; Ahrens 2003; Chen et al. 2011), biologists (e.g. Cossarini et al. 2008; Bucklin et al. 2012; Saba et al. 2012), agronomists, economists (e.g. van Vuuren et al. 2010; Matsumura et al. 2014), planning agencies (e.g. Mirhosseini et al. 2011), insurance companies (e.g. Cheng 2012), policy-makers and government agencies interested in providing local-scale, relevant information to their stakeholders (e.g. Barrow et al. 2004; Fuhrer et al. 2006; Eum et al. 2011; Dibike et al. 2007; Khalili et al. 2013).

Finer-scale climate information is usually obtained in three different ways: (1) by running a higher-resolution GCM, (2) by using a high-resolution nested climate model using the boundary conditions provided by the surrounding GCM, or (3) by using statistical downscaling methods. The first approach involves running the already computationally intensive GCMs at a higher resolution (e.g. Salathé et al. 2008); therefore, it is unlikely to be followed by practitioners without access to the computational resources where the GCMs are generally run, as doubling the spatial resolution of a GCM generally implies 16 times the volume of computations to be made (Coiffier 2011). The second approach, also known as dynamical downscaling, extracts regional scale information using regional climate models (RCMs) (e.g. Lim et al. 2011; Amengual et al. 2007; Chan et al. 2012). This approach is less computationally intensive than the previous one, as the RCMs use as lateral boundary conditions information from a coarser resolution GCM model (Mearns et al. 2003; Laprise 2008), but it also has high-computational requirements that makes it impractical to practitioners with access to single workstations; on the other hand, the third technique, known as statistical downscaling, is based on finding statistical relationships between the atmospheric variables from coarse resolution model outputs and the finer-scale variables (e.g. Jarosch et al. 2012; Nicholas and Battisti 2012; Gaitan et al. 2013; Gaitan and Cannon 2013; Cannon and Whitfield 2002). The statistical technique is less computationally demanding than the aforementioned methods and has the advantage that the statistical relationships can be obtained using desktop computers. Additionally, downscaling techniques can be used to understand the underlying relationships between the

coarse resolution predictors and surface observations and to challenge the GCMs with observations in ways that can feed back to inform GCM development (Gaitan et al. 2013).

Statistical downscaling techniques can be divided in three general categories: (1) regression models, (2) weather typing/classification schemes and (3) weather generators. This document will deal with regression techniques. Regression techniques represent linear or nonlinear relationships between predictands—i.e. local-scale variables—and large-scale predictors—i.e. coarse resolution output from the numerical models (Fowler et al. 2007). Wigley et al. (1990) provided an earlier reference to regression techniques and documented significant spatiotemporal variations in model performance near mountain and coastal regions, areas where sub-grid local process can be found. Detailed descriptions of the other statistical downscaling approaches can be found in Benestad et al. (2008). In general, weather generators replicate the statistical attributes of a local climate variable (mean and variance) but not observed sequences of events (Wilby et al. 2004), while the weather typing/classification techniques select the date in the historical (training) period when the situation most closely resembled the day for which the projection is made (Benestad et al. 2008), but because the projected values are taken from the historical observations, the finer-scale projections are limited to values inside the training interval (i.e. they do not extrapolate).

Similarly, statistical downscaling models rely on several assumptions in order to produce finer-/local-scale projections of the climate variable of interest; one of these assumptions is the time-invariance of the relationships between predictors (e.g. coarse resolution GCM output) and the local-scale predictands (e.g. gridded observation-based time-series, like PRISM (Daly et al. 2008), or weather station or buoy observations). However, in the absence of future observations, statistical downscaling studies rely on historical data to evaluate their models and assume that these historical simulation skills will be retained in the future (Wilby et al. 1998); hence, the results should be used with cautious confidence (Bardossy and Pegram 2011).

Recent studies on non-stationarity in statistical downscaling include: (1) the sensitivity analysis proposed by Wilby (1994) where the historical record could be fragmented into warm/cold or wet/dry years, with the statistical models trained in one fragment (usually the coldest one, if interested in future projections) and validated in the other (usually the warmest, assuming that the local mean temperature will increase in the future), (2) the validation methodology from Vrac et al. (2007), and its recent application to daily precipitation downscaling and daily wind speeds implemented by Gaitan et al. (2014b) and Gaitan and Cannon (2013), respectively; (3) the evaluation of bias correction methods using different conditions from those used for model calibration (Teutschbein and Seibert 2013), and (4) the “Perfect Model” evaluation (Dixon et al. 2016; Gaitan et al. 2014a) where the observations used for training the downscaling model are substituted by high-resolution GCM output and the predictors are derived from a coarsened version of the same high-resolution GCM. These pseudo-reality experiments allow the comparison of historical and future downscaling skills, while the information gathered from these experiments is crucial for a better-informed decision-making process and for the determination of the uncertainties associated with the statistical downscaling models (given their time-invariance assumption).

On the other hand, because a changing climate is expected to increase average summer temperatures and the frequency and intensity of extremes—including the number of summer hot days—(Walsh et al. 2012). We analyzed the annual maximum duration of summer hot days by calculating the heat wave duration index (HWDI). The HWDI is defined as the maximum number of days for a given summer (i.e. the months of June, July

and August) with maximum temperatures at least 5° warmer than the daily climatology from the historical (1971–2000) period, for at least 5 consecutive days. Heat waves usually occur in synoptic situations with pronounced slow air mass development and movement, leading to intensive and prolonged heat stress (Gad-el-Hak 2008), and are associated with significant morbidity and mortality (Huth et al. 2000). For example, it is estimated that the 2003 European heat wave caused more than 14,802 excess deaths in France, and 2045 and 2009 excess deaths in UK and Portugal, respectively (Gad-el-Hak 2008). In 2013, extreme heat wave lengths between five and sixteen days over Argentina and Uruguay produced the warmest December on record at various locations over central southern South America (Blunden and Arndt 2014). Heat waves, and heat stress in general, can also affect plant growth during key development stages (Kirschbaum 1995a), soil organic matter decomposition (Kirschbaum 1995b), and milk production (West 2003). In general, a warming climate will likely increase environmental stresses and may result in less resilient ecosystems that are unable to combat invasive species (Hellmann et al. 2008); as a warming climate will likely increase the rate of new species invasions and may promote the spread of already established species (Ontario Ministry of Natural Resources 2012).

In particular, our study case shows whether the statistically downscaled relationships between the coarse resolution global climate model and the local-scale predictands—daily maximum temperature for Montreal, Canada—are time-invariant, when using Vrac et al. (2007) evaluation methodology. The study also shows whether the skill simulating heat wave durations (from downscaled data) varied over time. Additionally, we will discuss the effect of two post-processing variance adjustment methodologies [randomization and variance inflation, following the terminology used by von Storch (1999) and Bürger et al. (2012)] on the statistically downscaled time-series. In particular, the evaluation methodology is shown for one GCM/RCM combination, as done by Vrac et al. (2007); however, the comparison can be extended in different dimensions, by evaluating other climate models, statistical downscaling methods (e.g. weather typing approaches, weather generators), climate variables, future projections, and different GCM/RCM runs.

2 Datasets

The North American Regional Climate Change Assessment Program-NARCCAP- (Mearns et al. 2007) includes twelve combinations of RCMs driven by GCMs (see supplementary material for details). Our application used daily values of maximum temperatures from the Canadian global climate model 3.1 (CGCM3.1) and the Canadian regional climate model 4.2 (CRCM4.2) as a proof of concept. The CGCM3.1 was considered one of the five top performing models simulating the mean annual cycle and inter-annual variability by Radić and Clarke (2011). Similar experiments can be performed using the other eleven available GCM/RCM combinations from NARCCAP.

Our setup used information from the historical (1971–2000) and future (2041–2070) time periods. Both periods correspond to the time windows used by the NARCCAP. The historical period uses the atmospheric component of CGCM 3.1 20C3M transient run, while the future period uses the SRES A2 scenario (IPCC 2000) forced with the CGCM 3.1 T47 run number 4. This scenario assumes a very heterogeneous world, with a high population growth and rapid economic development. CGCM3.1 outputs have been used as potential predictors for downscaling temperatures and precipitation in Quebec, Canada (e.g. Khan et al. 2006; Dibike and Coulibaly 2006; Jeong et al. 2012). The datasets were

obtained through the Data Access Integration (DAI) portal (<http://climat-quebec.qc.ca/CC-DEV/trunk/index.php/pages/dai>) from Environment Canada (DAI 2008). To help our partners from the Adaptation and Climate Monitoring Unit of the Meteorological Service of Canada (responsible for providing climate information for Quebec, Canada), we decided to focus our study on Montreal, Quebec.

In particular, we extracted historical and future daily maximum temperature (tasmax) outputs from the CRCM4.2 grid point over Montreal and used them as pseudo-observations. In addition, as predictors should carry the climate change signal, have a strong relationship with the predictand, and be realistically represented by the GCMs (Benestad et al. 2008), we used daily maximum temperatures from nine CGCM3.1 grid points near Montreal, Quebec. We realize that the use of other predictors might improve the models' skill, as shown by Huth (2003) and Gaitan et al. (2013). Similarly, we acknowledge the relationship between heat waves and upper-air high pressure shown by Alexander and Arblaster (2009). However, as the main focus of this manuscript is time-invariance and not predictor selection, our methodology uses simple(r) statistical models with the same meteorological variable being used as predictor and predictand (i.e. CGCM3.1 tasmax and CRCM4.2 tasmax, respectively). This approach of matching predictors-predictands is used in direct downscaling studies (e.g. Maurer and Hidalgo 2008; Stoner et al. 2013; Gaitan and Cannon 2013; Gaitan et al. 2014a).

3 Methods

In general, when using statistical downscaling techniques for climate applications, the most widely used regression and classification techniques rely on the time synchronous behavior between predictors (from a coarse resolution model) and the predictand/target (local-scale observations). However, as the outputs from the GCMs historical runs are not synched with observations (e.g. weather station data or gridded observation-based products), it is common to train the downscaling models with reanalysis data (e.g. NCEP/NCAR) instead of GCM outputs because the reanalysis products are synched with observations, as they are produced using information from land surface, ship, rawinsonde, pibal, aircraft and satellite data (Kalnay et al. 1996). Once the transfer functions (e.g. equations) are obtained using the reanalysis output and the observations, these functions are then used in conjunction with GCM data (as predictors) to produce downscaled (historical and future) estimates of the variable of interest. This process is done after determining that the predictors from the historical reanalysis are well simulated by the Global Climate Model of interest.

3.1 Using RCM outputs as pseudo-observations

As the statistical downscaling models used in this study rely on the time synchronous behavior between predictors and predictand, and because we were interested in testing the stationarity assumption, common to all downscaling methods, we decided to use Vrac et al. (2007) validation methodology.

This methodology uses historical and future output from a regional climate model as pseudo-observations, thus allowing the future comparison between the downscaled time-series and the target. However, as the RCM historical outputs are not synched in time with the reanalysis products, we need to use a predictor set that is synched with the pseudo-observations/RCM output. Fortunately, the RCMs use as boundary conditions information

from coarse resolution GCMs “driving” the RCM. Therefore, it is of utmost importance to guarantee that the RCM output (used as pseudo-observations) has been produced by the same GCM run used to extract the predictors.

As a counterexample, if one tries to replicate this analysis using RCM outputs obtained from different boundary conditions (i.e. a GCM run not used to obtain the predictors) the correlations will drop significantly, as the time synchronous behavior between the local predictand and the coarse scale predictors is not present. As mentioned earlier, one can use any GCM/RCM combination if the GCM and the RCM are synched in time; therefore, the methodology is not limited to the NARCCAP models. Here we evaluate one of the NARCCAP combinations (CGCM3/CRCM4.2) as a proof of concept.

3.2 Statistical downscaling

The statistical downscaling procedure we used consists of three different steps:

1. Train the statistical downscaling model using the historical CGCM3.1 coarse resolution daily maximum temperature output as predictors and CRCM4.2 pseudo-observed daily maximum temperatures as predictands;
2. Cross-validate the downscaled output (i.e. compute the model’s error using independent data);
3. Once the downscaling models have been trained and cross-validated using historical data, we use future CGCM3.1 model output as predictors of the cross-validated models, and obtain downscaled future possible climates.

For this particular application, we removed the climatological seasonal mean from the predictors and the predictands (yielding anomalies from the mean datasets) and trained the regression models with these anomalies. As we are interested in evaluating different quantiles and return periods, we added back the climatological seasonal mean to the statistically downscaled time-series (anomalies) to perform the evaluations. The regression methods used twofold cross-validation to prevent overfitting (Bishop 2006).

Nevertheless, the reader must be aware of several limitations when downscaling to gridded data (e.g. RCM output and gridded observation-based datasets), including: (1) the gridded predictand represents an area average not point measurements; (2) the variance of a variable averaged over a large area is expected to be smaller than the variance of the same variable at a particular weather station/point, and (3) in the case of precipitation analyses, the wet spells calculated from the gridded data likely last longer than the observed ones (Gaitan Ospina 2013). Furthermore, Chen and Knutson (2008) cautioned the practitioners about using gridded observations as point estimates; similarly when using RCM output as pseudo-observations, one should be aware that since RCMs simulate climate over a specified area of interest, they require nesting information which describes the evolution of the atmospheric circulation at their lateral boundaries (Music and Sykes 2011), and thus are affected by the driving GCM uncertainties. Additionally, when using the CRCM4.2 output as pseudo-observations, it is worth noting that according to Bourdages and Huard (2010), the simulated temperatures generated by the CRCM 4.2 (driven by the CGCM3) are lower than the observed values.

3.2.1 Regression-based statistical downscaling

Regression analysis seeks to find the relationship between one or more independent variables and a dependent variable. Regression models used to downscale can represent

linear or nonlinear relationships between predictands and large-scale predictors (Fowler et al. 2007). In particular, when downscaling temperatures over Europe, Huth et al. (2008) found out that nonlinear methods did not necessarily improve the solutions, when compared to simpler linear methods. Similarly, when downscaling temperatures over southern Ontario and Quebec, Gaitan et al. (2013) found that the improvements of the nonlinear methods over the linear ones were primarily noticed when comparing climate indices in terms of indices of agreement (Willmott et al. 2012), and only marginally in terms of mean absolute errors. Additionally, linear methods using only 2 m. temperature output [from the NCEP/NCAR reanalysis (Kalnay et al. 1996)] as predictors performed similarly than complex non-linear methods, when comparing the HWDI index of agreement (Gaitan et al. 2013).

Here, we compared two popular regression methods: (1) multiple linear regression with stepwise selection and (2) robust multiple linear regression.

3.2.1.1 Multiple linear regression with stepwise selection Stepwise techniques for regression analysis are described in Darlington (1990). In general, multiple linear regression with stepwise selection—commonly known as stepwise regression (hereafter SWLR)—is a systematic method for adding and removing predictors from a multiple linear regression model. In particular, an initial model is created at the first iteration and then the p value is computed to test models with and without a potential predictor. The null hypothesis is that the predictor to be added or removed has a zero regression coefficient (Hill and Lewicki 2006). This multiple linear regression approach uses forward selection and/or backward elimination to select a predictors subset from an original pool containing all the available predictors (Wilks 2011; Darlington 1990). For a recent application of SWLR in statistical downscaling, refer Hessami et al. (2008).

Here, we used the MATLABTM statistics toolbox to implement the stepwise multiple linear regression (SWLR) and kept its default p -values for adding and removing a variable at 0.05 and 0.1, respectively.

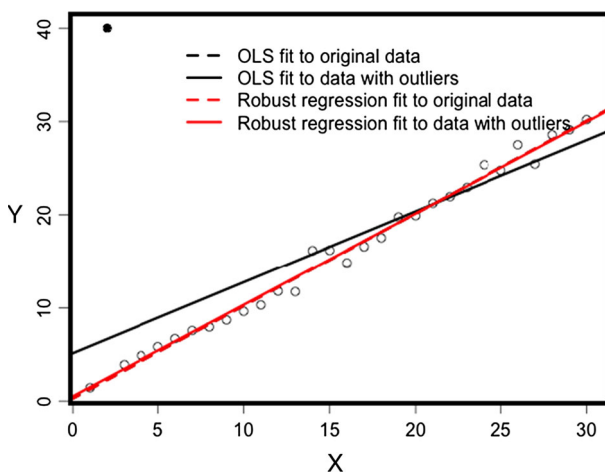
3.2.1.2 Robust regression Robust regression is an alternative to ordinary least squares regression (OLS) that is less sensitive than OLS to outliers. As with OLS, each point is assigned equal weight in the first iteration; however, during the next iterations, the weights are readjusted so less weight is given to the points farther from the predictions. Then the model coefficients are adjusted using weighted least squares, until a stopping criterion is met. Hereafter, the statistically downscaled time-series produced using this type of regression will be identified with the word “Robust”.

Here, we used MATLAB’s robust regression (*robustfit*). This implementation uses iteratively reweighted least squares with the bi-square weighting function.

Figure 1 illustrates the main differences between the OLS regression and the robust regression. The figure compares regression estimates of two datasets that only differ in one point [$y = 2$, in the original data, and $y = 40$ in the modified dataset—see black dot]. The regressions from the first dataset are identified with dashed lines, and the regressions from the dataset containing the outlier are plotted with solid lines. The figure shows that the OLS estimates from the two datasets differ in terms of intercept and slope, while the estimates from the robust regression are closer to the ones from the first dataset even in the presence of the outlier.

Fig. 1 Schematic representation about the differences between OLS and Robust regression.

Dashed red and black lines (Robust and OLS, respectively) show the model fits to a dataset containing no outliers; *solid lines* show the corresponding fit when substituting one point from the original dataset for one outlier (solid black dot ($X = 2, Y = 4$)). Notice the behavior (change in slope and intercept) of the *solid black line* in the presence of the outlier



3.3 Variance adjustment techniques: randomization (RN) versus variance inflation (VI)

As regression methods underestimate the variance (i.e. the simulated variance is less than the observed one), the downscaled estimates will have lower variances than the target dataset (i.e. high-resolution observations or predictand). To address this drawback, two solutions have been proposed: (1) variance inflation (Karl et al. 1990) and (2) to add noise to the downscaled time-series in order to match the observed variance (von Storch 1999), this technique is also known as randomization (Bürger et al. 2012; von Storch 1999).

In general, to obtain the variance-adjusted time-series using variance inflation (Y_{VI}), one needs to multiply the downscaled output (\tilde{y}) by the square root of the ratio of the variances between the observations (y) and downscaled time-series.

$$Y_{vi} = \tilde{y}(\text{var}(y)/\text{var}(\tilde{y}))^{0.5}, \quad (1)$$

However, this method is criticized by von Storch (1999) and others because the predictors do not completely specify the small-scale feature of interest, and because the VI process affects the MSE between the target and the downscaled time-series.

On the other hand to obtain variance-adjusted downscaled estimates using randomization (Y_{RN}), one needs to add noise (white or red) to account for the unexplained variance (von Storch 1999):

$$Y_{RN} = \tilde{y} + \text{noise}, \quad (2)$$

However, as pointed out by Huth et al. (2001), the addition of noise breaks the temporal correlation of the downscaled data.

3.4 Model evaluation

We evaluated the downscaled results in terms of daily variability and their ability to reproduce heat waves, using the HWDI as a proxy. The independent evaluation error of the entire dataset was obtained using cross-validation (e.g. James et al. 2013; Bishop 2006).

Specifically, we divided the data into two adjacent sections (of equal length), used one section to train the models, and the remaining section was used to test the model on independent data. We repeated the procedure so both sections were used to test predictions.

The daily variability validation involved calculating the RMSEs between the pseudo-observations and the statistically downscaled series, while the ability to reproduce heat waves was determined by RMSEs between the pseudo-observations and the downscaled time-series.

4 Results

In this section we analyze the downscaled datasets in three different categories; first, we evaluate the statistically downscaled time-series' daily variability statistics; then, we assess the performance of the downscaled datasets in terms of different quantiles and return

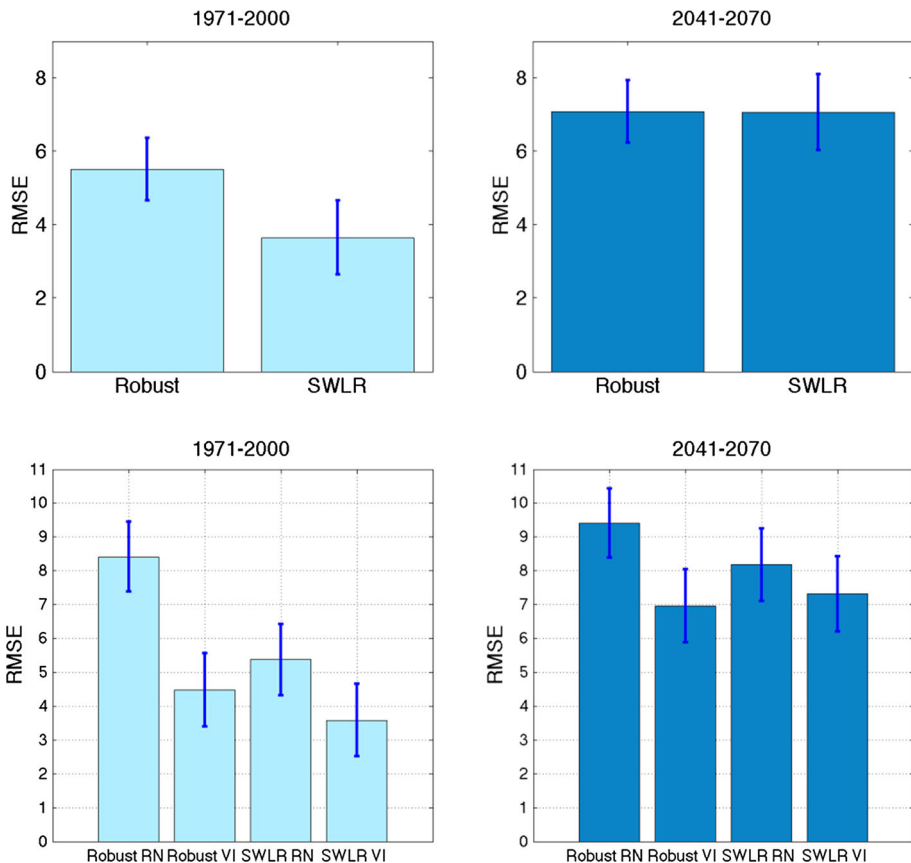


Fig. 2 RMSE between the statistically downscaled time-series and the CRCM4.2 pseudo-observations. Error bars indicate standard errors. Time-series identified with “VI” were variance-inflated, while the ones identified with “RN” include random noise to adjust the modeled and observed variances. Top figures show historical (left) and future (right) RMSEs without using variance adjustment techniques. Bottom left and right figures show historical and future RMSEs after using variance adjustment techniques

periods; and then we finish the section by comparing the results from the heat wave duration indices calculated from the downscaled and pseudo-observed time-series.

4.1 Daily maximum temperature time-series

After adding back the seasonal cycle to the statistically downscaled anomalies, we compared the resulting time-series with those from the CRCM4.2 in terms of Pearson correlation to evaluate whether the downscaled values were synchronized in time with the pseudo-observations; and in terms of simulated variance, to test whether the regression methods produced values from a wide temperature range.

The analysis shows that both methods adequately simulated the observed temperatures in terms of daily variability (correlation coefficients >0.94); however, when examining the variances one of the known drawbacks of the regression methods was evident, the simulated variance was less than the observed one. In particular, the SWLR estimates explained more variance than the robust regression estimates (89 and 69 %, respectively). This behavior was partially caused because the robust regression is less sensitive to data points marked as outliers than the traditional SWLR using OLS.

As we are interested in evaluating the performance of the downscaled time-series in terms of quantiles, return periods and heat wave durations, we created two different post-processed versions of each statistically downscaled time-series, one using randomization (RN) and another one using variance inflation (VI). Therefore, we compared four time-series with the observations. Hereafter these time-series will be denoted as Robust RN, Robust VI, SWLR RN and SWLR VI.

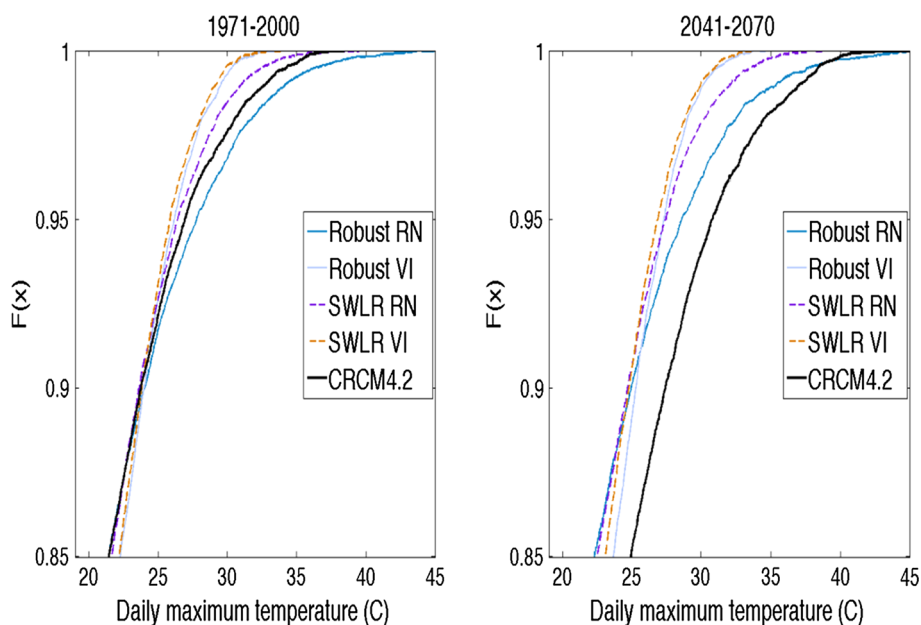


Fig. 3 Statistically downscaled time-series empirical cumulative density functions for the historical and future periods (*left and right panels*, respectively)—upper tails ($F(x) > 0.85$) of the distributions

Table 1 Quantiles (degrees celsius) from the statistically downscaled time-series

Dataset	Quantiles							
	0.05		0.25		0.50		0.75	
	20C3 M	A2	20C3 M	A2	20C3 M	A2	20C3 M	A2
Robust VI	-11.16	-9.35	-1.56	-0.52	8.26	9.69	18.97	20.31
Robust RN	-11.81	-10.54	-0.31	0.71	8.49	9.55	17.05	18.21
SWLR VI	-11.65	-10.89	-1.28	-0.98	8.60	9.52	18.88	19.73
SWLR RN	-11.86	-11.08	-0.81	-0.51	8.77	9.35	18.04	18.97
CRCM4.2	-12.20	-8.80	-0.18	1.64	8.00	10.89	17.83	20.93

Values in bold typeface indicate the closest value to the CRCM 4.2 pseudo-observations (last row)

Figure 2 shows the RMSE between the aforementioned variance-adjusted statistically downscaled time-series, and the CRCM4.2 output used as pseudo-observations. The figure shows the non-stationarity of the errors between the historical (1971–2000) and future (2041–2070) time periods, with higher errors in the future, and considerable differences (in RMSE) between periods for Robust VI, SWLR RN and SWLR VI. Similarly, it is clear that SWLR VI, SWLR RN and Robust VI outscored the downscaled output of Robust RN in the historical period; however for the future period, the performance of SWLR RN is comparable (in terms of RMSE) with the performance of Robust RN, and the performance of Robust VI is comparable to the one from SWLR VI.

The higher future errors are also caused by differences in the pseudo-observed variances between periods. In particular the variance of the future CRCM4.2 maximum temperature output, used as pseudo-observations, is 5 % higher than the historical pseudo-observed variance.

4.2 Quantiles and return periods

Figure 3 shows the historical and future empirical cumulative distribution functions (ecdf) from the variance-adjusted statistically downscaled time-series and from the pseudo-observations. The left panel, displaying the historical simulations, shows that in general there is a good agreement between the statistically downscaled time-series and the pseudo-observations; with higher differences found around the 90th percentile, where Robust VI, SWLR VI and SWLR RN underpredicted the percentiles of the upper tail, while Robust RN overpredicted the same quantiles. On the right panel, showing the future projections, we noticed that in general the downscaled datasets had a cold bias versus the pseudo-observations.

Please refer to Table 1 for a brief comparison of the 5th, 25th, 50th, 75th and 95th quantiles from the statistically downscaled time-series and from the pseudo-observations

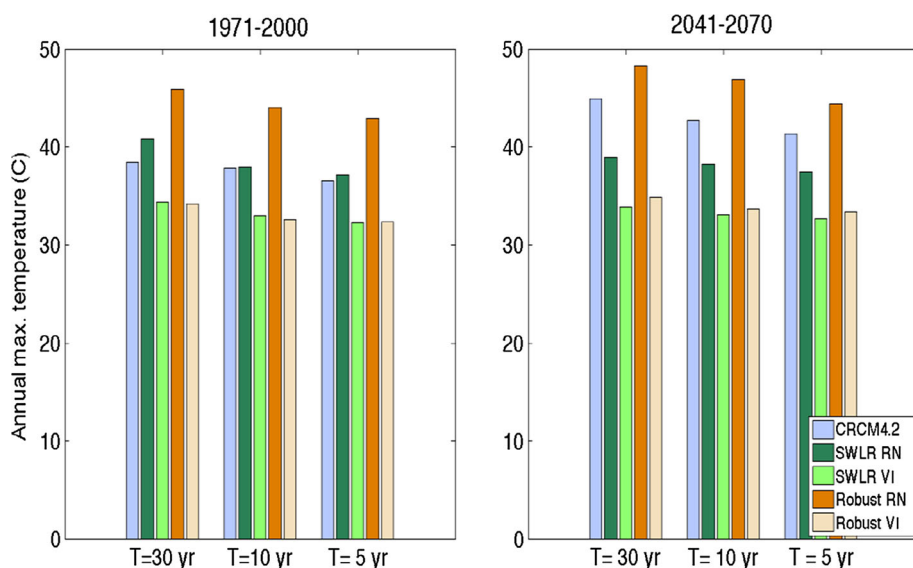


Fig. 4 Empirical return periods. Daily maximum temperature annual maxima

for the historical and future periods. Values in bold typeface indicate the closest value to the CRCM 4.2 pseudo-observations (last row).

When looking at the last 15 quantiles from the empirical cumulative density functions (Fig. 3), we can see that SWLR VI and Robust VI exhibited a similar tail behavior, while Robust RN and SWLR RN had bigger differences between them. In particular, when analyzing the future period simulations, all the downscaled time-series underpredicted the daily maximum temperatures, with Robust RN showing lower differences to the pseudo-observations than the other three time-series.

Regarding the study of the return periods (Fig. 4), we calculated the 30-, 10- and 5-year return periods (denoted by the letter T) from the annual maximum temperatures of the variance-adjusted statistically downscaled time-series using the California Method (Monsalve Saenz 2002) to determine the frequencies. The historical results (Fig. 4, left) show that more than 10° separate the return periods from Robust VI, SWLR VI and Robust RN. A similar pattern was found in the future period (Fig. 4, right). In general, Robust RN overpredicted the historical and future return periods, while SWLR VI and Robust VI

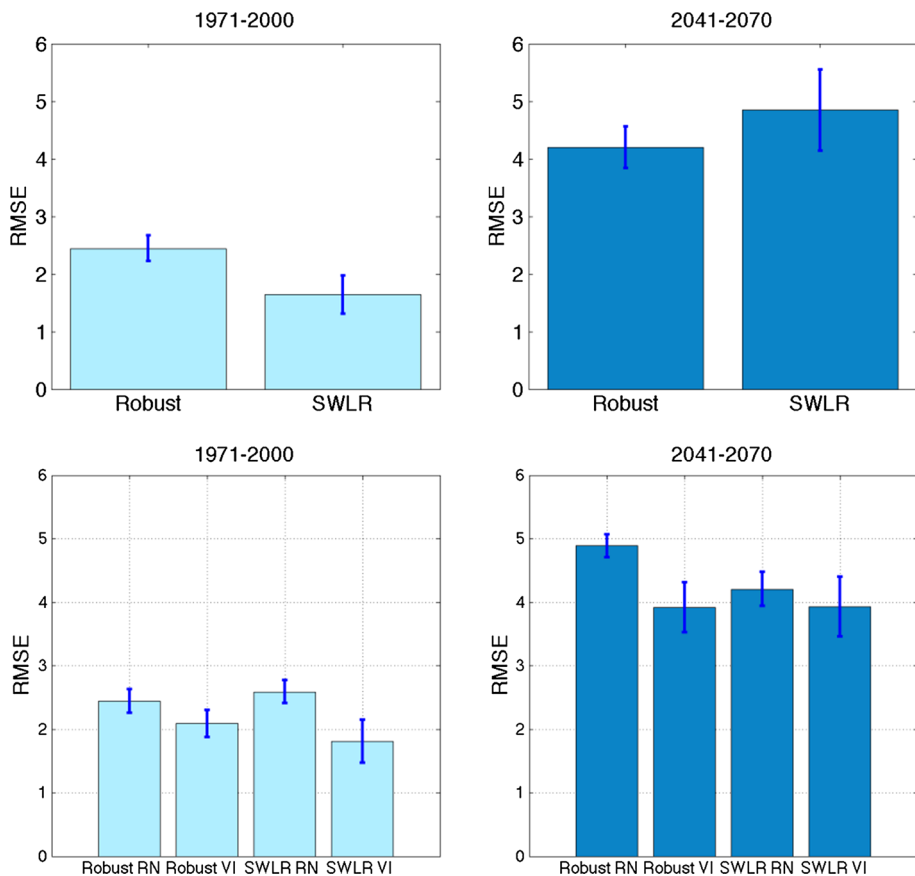


Fig. 5 Heat wave duration index RMSE versus the CRCM4.2 pseudo-observations. *Error bars* indicate standard errors. *Top left and right* figures, historical and future comparison, respectively, without variance adjustment. *Bottom left and right*, historical and future comparisons with variance adjustment techniques

underpredicted the return periods for both periods. On the other hand, the historical return periods from SWLR RN show a better agreement with the pseudo-observed ones; however, when looking at the future period, SWLR RN underpredicted the pseudo-observed return periods.

The return periods obtained from the statistically downscaled time-series indicate that having the best agreement (vs the return periods from the pseudo-observations) in the historical period does not guarantee having the same bias in the future (e.g. SWLR RN), especially as the return periods from the pseudo-observations increased over time. For example, the 30-year (pseudo-observed) return period changed from $\sim 38^{\circ}\text{C}$ (historical period) to $\sim 45^{\circ}\text{C}$ in the future ($+7^{\circ}\text{C}$), while the return periods calculated from SWLR RN changed from ~ 39 to $\sim 40^{\circ}\text{C}$ in the same period. On the other hand, the underprediction of the historical return periods shown by SWLR VI and Robust VI persisted in the future period, while Robust RN overpredicted the return periods on both historical and future periods. This performance agrees with the upper tail behavior shown in Fig. 3 by Robust RN, where higher than “observed” values were shown in the upper quantiles of the daily ecdf.

4.3 Heat wave duration index (HWDI)

As mentioned earlier, for the purpose of this study, the heat wave duration index (HWDI) is defined as the maximum number of days for a given year with maximum temperatures at least 5°C warmer than the daily climatology from the historical (1971–2000) period, for at least 5 consecutive days, during the summer months (June, July and August).

After calculating the annual HWDI for the variance-adjusted statistically downscaled time-series and for the pseudo-observations (historical and future), we compared their indices in terms of RMSEs (Fig. 5). The results show that the HWDI errors also vary greatly over time, with markedly higher errors in the future (even taking into consideration the standard errors). Additionally, the historical results show that SWLR VI outscored the other models in terms of historical RMSE and that Robust RN and SWLR RN exhibited similar errors, while the future results show that the indices from Robust VI, SWLR RN and SWLR VI were comparable in terms of RMSEs.

Other statistics (maximum, mean and median) calculated from the historical and future heat wave duration indices are shown in Table 2. In general, SWLR VI outscored the other models in terms of maximum and mean heat wave durations (historical and future), and median HWDI for the historical period. All the models underrepresented the median and minimum HWDI, with Robust VI and SWLR VI producing the closest values to the median HWDI from the future “observations”. In particular, the biggest differences appear when calculating the maximum heat wave duration for the future period; as Robust VI and

Table 2 HWDI statistics—1971–2000 (20C3 M) and 2041–2070 (A2)

Dataset	Maximum		Mean		Median	
	20C3 M	A2	20C3 M	A2	20C3M	A2
Robust RN	5	6	3.03	3.10	3	3
Robust VI	6	12	3.43	5.07	3	5
SWLR RN	5	7	2.90	4.13	3	4
SWLR VI	9	13	4.07	5.37	4	5
CRCM4.2	8	13	4.96	7.80	5	7

Values in bold typeface indicate the closest value to the CRCM 4.2 pseudo-observations (last row)

SWLR VI produced heat waves of 12 and 13 days, respectively, (vs 13-day heat waves from the pseudo-observations), while SWLR RN and Robust RN only produced heat waves of maximum 7 and 6 days, respectively.

Overall the comparisons made in this manuscript show that the future errors vary in time, when compared to the historical ones. This indicates that the statistical relationships defined during the historical (training) period were not time-invariant, and thus the results derived from the statistical models should be taken with caution. In particular, when analyzing daily variability, the future RMSEs were considerably higher than the historical ones. Similarly, when looking at the HWDI calculated from the downscaled data, we noticed that the errors changed over time and that model-derived results were less reliable in the future; but in general the variance-inflated models outscored their counterparts using randomization.

5 Summary and recommendations

The present work evaluated whether the statistical relationship(s) between the coarse resolution predictors and the finer scale predictand, used in statistical downscaling, remained constant over time when downscaling daily maximum temperatures over Montreal, Canada. In particular, we evaluated whether the errors between two statistically downscaled datasets and pseudo-observed time-series obtained from the CRCM4.2 were time-invariant. In addition we evaluated the downscaled output in terms of quantiles, means, medians, maximum and minimum values, and 30-, 10- and 5-year return periods. We also evaluated whether the heat wave duration indices derived from the statistically downscaled time-series differed from the ones calculated from the pseudo-observations. We encourage future studies to extend our analyses to other GCM/RCM combinations, other scenarios (or RCPs) and to other downscaling methods. However, it is worth mentioning that currently NARCCAP only has 12 GCM/RCM combinations and a very limited number of scenarios available.

Here, we used two popular regression approaches: linear regression with stepwise selection and robust regression, to statistically downscale daily maximum temperatures from the Canadian GCM 3.1. As regression methods underestimate the variance, we variance-adjusted the downscaled time-series using two different methodologies: variance inflation and randomization. The results suggest that both regression approaches showed non-stationarity of their errors when comparing the historical and future variance-adjusted time-series (in terms of RMSE), with SWLR RN, SWLR VI, and Robust VI marginally outscoring Robust RN.

Overall, the use of the randomization post-processing method might not be recommended when the variance to be added represents ~ 50 % of the explained variance like in the case of the robust regression; however, when the variance to be added is ~ 10 % of the explained variance (e.g. SWLR) the method could be used with more confidence, as shown in the bottom-left (historical) panel of Fig. 5. Nevertheless, when looking at the results for the future period, one might consider the use of a second variance adjustment step in order to better reproduce the future variance of the pseudo-observations, as their future variance is greater than the historical one. In practice we cannot adjust the future downscaled series to match the variance of the future pseudo-observations as we do not have future local-scale information; alternatively one might further adjust the downscaled future values by

also taking into account the variance difference between the future and historical GCM predictors and then using randomization or variance inflation to account for this variance.

In terms of HWDI, after calculating the indices from the observed and downscaled datasets and then comparing the indices' RMSEs, we observed that SWLR VI and Robust VI outscored both SWLR RN and Robust RN. This suggests that the use of variance-inflated time-series might be preferable than using the ones with randomization (when interested in the analysis of heat wave durations). This is likely because the variance inflation approach preserves temporal behavior of the downscaled time-series; consequently preserving the simulated heat spells. However, as with the analysis of the downscaled time-series, the future RMSEs were notably higher than the historical ones. Hence, it is not recommended to assume that the downscaling models' historical performance (simulating heat wave durations) would be kept in the future. Additionally, the results suggest that although the linear regression with stepwise selection method seemed to outscore the robust regression when downscaling daily maximum temperature, it might not necessarily produce better results than the robust regression when calculating the median duration of heat waves.

On the other hand, the return periods obtained from the statistically downscaled time-series indicate that having the best agreement (vs the return periods from the pseudo-observations) in the historical period does not guarantee having the same bias in the future (e.g. SWLR RN), especially as the pseudo-observations' return periods increased over time. Overall, here we show that the future errors vary in comparison with the historical ones, indicating that the statistical relationships defined during the historical (training) period are not time-invariant.

Our findings have significant repercussions given that one of the statistical downscaling paradigms is to assume that present simulation skills will be kept in the future, it is likely that a stakeholder using statistically downscaled data to make decisions, or a practitioner needing to use high-resolution local-scale data as input to other ecological, biological, hydrological or economical models, could end up selecting a model with poor future performance, by assuming stationary relationships. As mentioned earlier, there are many causes of uncertainty in climate change regional projections in addition to the uncertainty caused by the choice of downscaling method; here, we found that after obtaining the downscaled results, the choice of post-processing technique (for variance adjustment) can also affect the final local-scale projections, and the differences among the post-processed projections were as important as the differences between the downscaling methods used (Robust and SWLR). An ongoing study is being conducted to analyze the effect of post-processing techniques on statistical downscaling methods, and their impact on the downscaled projections uncertainty.

We speculate that downscaling uncertainty and scenario uncertainty are still the predominant sources of uncertainty, with non-stationarity uncertainty being a sub component of the statistical downscaling uncertainty process. For example, when comparing 28 climate change projections from seven GCMs and three scenarios, over Central Quebec, Chen et al. (2011) concluded that the uncertainty envelopes from the downscaling methods were similar to the envelopes from the emission scenarios. Therefore, it is possible that if we add the (unaccounted) non-stationarity uncertainty to the downscaling uncertainty, this envelope could become the predominant one, as the regression-based statistical downscaling methods employed by Chen et al. (2011) contributed significantly to the uncertainty envelope.

On the other hand, once the practitioners establish the non-stationarity of the statistical relationships used, we envision a series of steps that might improve their downscaled

estimates. First, the practitioners should try to use downscaled output from downscaling models that had been trained and evaluated in different climate regimes (e.g. positive and negative ENSO) and include best practices like cross-validation of the downscaled model output. Second, when making decisions, the practitioners will need to expand the downscaling model uncertainty to include non-stationarity. Therefore, their decision-making process might need to include deep uncertainty analysis (Hallegatte et al. 2012), or adopt practices from other disciplines, like cost–benefit analysis under uncertainty (Arrow and Fisher 1974), cost–benefit analysis with regret minimization (Hahn et al. 1996), climate informed decision analysis (CIDA, Brown et al. 2011), and/or robust decision making (RDM, Lempert and Collins 2007; Hallegatte 2009).

Recently Hall (2014) argued that the climate science community must identify downscaling’s strengths and limitations and develop best practices to prevent bad decisions. In the end, we aspire that by knowing that there are differences between the models’ historical and future performances the practitioners will get valuable information regarding the level of confidence one should attribute to the downscaled climate projections. Even though our evaluation illustrated two simple regression-type statistical downscaling models, the main conclusions may also be valid for more complicated models, like the nonlinear classification and regression models used by Gaitan et al. (2014b) to downscale daily precipitation. Our results also corroborate the cautionary notes from Chen et al. (2011) and Ouyang et al. (2014) regarding the confidence that should be attributed to climate change impact studies based on only one downscaling method.

Acknowledgments The author would like to acknowledge the Data Access Integration (DAI) Team for providing the data and technical support. The DAI Portal (<http://loki.qc.ec.gc.ca/DAI/>) is made possible through collaboration among the Global Environmental and Climate Change Centre (GEC3), the Adaptation and Impacts Research Division (AIRD) of Environment Canada, and the Drought Research Initiative (DRI). The College of Atmospheric and Geographic Sciences at the University of Oklahoma provided the funds to support the author. Support for the lead author’s workspace and computational environment were provided by NOAA’s Geophysical Fluid Dynamics Laboratory (GFDL).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ahrens B (2003) Rainfall downscaling in an alpine watershed applying a multiresolution approach. *Journal of Geophysical Research* 108(D8). doi:[10.1029/2001jd001485](https://doi.org/10.1029/2001jd001485)
- Alexander LV, Arblaster JM (2009) Assessing trends in observed and modelled climate extremes over Australia in relation to future projections. *Int J Climatol* 29(3):417–435. doi:[10.1002/Joc.1730](https://doi.org/10.1002/Joc.1730)
- Amengual A, Romero R, Homar V, Ramis C, Alonso S (2007) Impact of the lateral boundary conditions resolution on dynamical downscaling of precipitation in mediterranean Spain. *Clim Dyn* 29(5):487–499. doi:[10.1007/S00382-007-0242-0](https://doi.org/10.1007/S00382-007-0242-0)
- Anandhi A, Srinivas VV, Nanjundiah RS, Nagesh Kumar D (2008) Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine. *Int J Climatol* 28(3):401–420. doi:[10.1002/joc.1529](https://doi.org/10.1002/joc.1529)
- Arrow K, Fisher A (1974) Environmental preservation, uncertainty, and irreversibility. *Q J Econ* 88:312–319
- Bardossy A, Pegram G (2011) Downscaling precipitation using regional climate models and circulation patterns toward hydrology. *Water Resour Res* 47(4):1–8. doi:[10.1029/2010wr009689](https://doi.org/10.1029/2010wr009689)

- Barrow EM, Maxwell B, Gachon P (2004) Climate variability and change in Canada. Past, present and future. ACSD Science Assessment Series. Environment Canada, Toronto, Ontario
- Benestad RE, Chen D, Hanssen-Bauer I (2008) Empirical-statistical downscaling. World Scientific, Singapore
- Bishop CM (2006) Pattern recognition and machine learning. Springer, Cambridge
- Blunden J, Arndt DS (2014) State of the climate in 2013. *Bull Am Meteorol Soc* 95:1–279
- Bourdages L, Huard D (2010) Climate change scenario over Ontario based on the Canadian regional climate model (CRCM4.2). Ouranos, Montreal
- Brown C, Werick W, Leger W, Fay D (2011) A decision-analytic approach to managing climate risks: application to the upper great lakes. *J Am Water Resour As* 47(3):524–534
- Bucklin DN, Watling JI, Speroterra C, Brandt LA, Mazzotti FJ, Romañach SS (2012) Climate downscaling effects on predictive ecological models: a case study for threatened and endangered vertebrates in the southeastern United States. *Reg Environ Change* 13(S1):57–68. doi:[10.1007/s10113-012-0389-z](https://doi.org/10.1007/s10113-012-0389-z)
- Bürger G, Murdock TQ, Werner AT, Sobie SR, Cannon AJ (2012) Downscaling extremes—an intercomparison of multiple statistical methods for present climate. *J Clim* 25(12):4366–4388. doi:[10.1175/jcli-d-11-00408.1](https://doi.org/10.1175/jcli-d-11-00408.1)
- Cannon AJ, Whitfield PH (2002) Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. *J Hydrol* 259:136–151
- Chan SC, Kendon EJ, Fowler HJ, Blenkinsop S, Ferro CAT, Stephenson DB (2012) Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? *Clim Dyn* 41(5–6):1475–1495. doi:[10.1007/s00382-012-1568-9](https://doi.org/10.1007/s00382-012-1568-9)
- Chen C-T, Knutson TR (2008) On the verification and comparison of extreme rainfall indices from climate models. *J Clim* 21(7):1605–1621. doi:[10.1175/2007jcli1494.1](https://doi.org/10.1175/2007jcli1494.1)
- Chen J, Brissette FP, Leconte R (2011) Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *J Hydrol* 401(3–4):190–202. doi:[10.1016/j.jhydrol.2011.02.020](https://doi.org/10.1016/j.jhydrol.2011.02.020)
- Cheng CS (2012) Climate change and heavy rainfall-related water damage insurance claims and losses in Ontario, Canada. *J Water Resour Prot* 04(02):49–62. doi:[10.4236/jwarp.2012.42007](https://doi.org/10.4236/jwarp.2012.42007)
- Coiffier J (2011) Fundamentals of numerical weather prediction. Cambridge University Press, New York
- Cossarini G, Libralato S, Salon S, Gao X, Giorgi F, Solidoro C (2008) Downscaling experiment for the Venice lagoon. II. Effects of changes in precipitation on biogeochemical properties. *Clim Res* 38(1):43–59. doi:[10.3354/Cr00758](https://doi.org/10.3354/Cr00758)
- DAI (2008) Predictor datasets derived from the CGCM3.1 T47 and NCEP/NCAR reanalysis
- Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int J Climatol* 28(15):2031–2064. doi:[10.1002/joc.1688](https://doi.org/10.1002/joc.1688)
- Darlington RB (1990) Regression and linear models. Chapter 18, McGraw-Hill, New York
- Dibike YB, Coulibaly P (2006) Temporal neural networks for downscaling climate variability and extremes. *Neural Netw* 19(2):135–144. doi:[10.1016/j.neunet.2006.01.003](https://doi.org/10.1016/j.neunet.2006.01.003)
- Dibike YB, Gachon P, St-Hilaire A, Ouarda TBMJ, Nguyen VTV (2007) Uncertainty analysis of statistically downscaled temperature and precipitation regimes in Northern Canada. *Theor Appl Climatol* 91(1–4):149–170. doi:[10.1007/s00704-007-0299-z](https://doi.org/10.1007/s00704-007-0299-z)
- Dixon KW, Lanzante JR, Nath MJ, Hayhoe K, Stoner A, Radhakrishnan A, Balaji V, Gaitán CF (2016) Evaluating the stationarity assumption in statistically downscaled climate projections: is past performance an indicator of future results? *Climatic Change*. doi:[10.1007/s10584-016-1598-0](https://doi.org/10.1007/s10584-016-1598-0)
- Eum H-I, Gachon P, Laprise R, Ouarda T (2011) 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. doi:[10.1007/s00382-011-1149-3](https://doi.org/10.1007/s00382-011-1149-3)
- Fowler HJ, Blenkinsop S, Tebaldi C (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int J Climatol* 27(12):1547–1578. doi:[10.1002/joc.1556](https://doi.org/10.1002/joc.1556)
- Franklin J, Davis FW, Ikegami M, Syphard AD, Flint LE, Flint AL, Hannah L (2013) Modeling plant species distributions under future climates: how fine scale do climate projections need to be? *Glob Change Biol* 19(2):473–483. doi:[10.1111/gcb.12051](https://doi.org/10.1111/gcb.12051)
- Fuhrer J, Beniston M, Fischlin A, Frei C, Goyette S, Jasper K, Pfister C (2006) Climate risks and their impact on agriculture and forests in Switzerland. *Clim Change* 79(1–2):79–102. doi:[10.1007/s10584-006-9106-6](https://doi.org/10.1007/s10584-006-9106-6)
- Gad-el-Hak M (2008) Large-scale disasters: prediction, control and mitigation. Cambridge University Press, Hong Kong

- Gaitan CF, Cannon AJ (2013) Validation of historical and future statistically downscaled pseudo-observed surface wind speeds in terms of annual climate indices and daily variability. *Renew Energy* 51:489–496. doi:[10.1016/j.renene.2012.10.001](https://doi.org/10.1016/j.renene.2012.10.001)
- Gaitan Ospina CF (2013) Comparison of linearly and nonlinearly statistically downscaled atmospheric variables in terms of future climate indices and daily variability. University of British Columbia, Vancouver
- Gaitan CF, Hsieh WW, Cannon AJ, Gachon P (2013) Evaluation of linear and non-linear downscaling methods in terms of daily variability and climate indices: surface Temperature in Southern Ontario and Quebec, Canada. *Atmos Ocean* 52(3):211–221. doi:[10.1080/07055900.2013.857639](https://doi.org/10.1080/07055900.2013.857639)
- Gaitan CF, Dixon KW, McPherson R, Balaji V (2014) Statistically downscaled north american precipitation using support vector regression and the perfect model evaluation framework. In: 11th International conference on hydroinformatics, New York
- Gaitan CF, Hsieh WW, Cannon AJ (2014b) Comparison of statistically downscaled precipitation in terms of future climate indices and daily variability for southern Ontario and Quebec, Canada. *Clim Dyn* 43(12):3201–3217. doi:[10.1007/s00382-014-2098-4](https://doi.org/10.1007/s00382-014-2098-4)
- Hahn RW, Lave LB, Noll RG, Portney PR, Russel M, Schmalensee RL (1996) Benefit-cost analysis in environmental, health, and safety regulation. American Enterprise Institute Books and Monographs, Washington
- Hall A (2014) Projecting regional change. *Science* 346(6216):1461–1462. doi:[10.1126/science.aaa0629](https://doi.org/10.1126/science.aaa0629)
- Hallegatte S (2009) Strategies to adapt to an uncertain climate change. *Glob Environ Change* 19:240–247
- Hallegatte S, Shah A, Lempert R, Brown C, Gill S (2012) Investment decision making under deep uncertainty-application to climate change (trans: Economist OotC). The World Bank,
- Hellmann JJ, Byers JE, Bierwagen BG, Dukes JS (2008) Five potential consequences of climate change for invasive species. *Conserv Biol J Soc Conserv Biol* 22(3):534–543. doi:[10.1111/j.1523-1739.2008.00951.x](https://doi.org/10.1111/j.1523-1739.2008.00951.x)
- Hessami M, Gachon P, Ouarda TBMJ, St-Hilaire A (2008) Automated regression-based statistical downscaling tool. *Environ Model Softw* 23(6):813–834. doi:[10.1016/j.envsoft.2007.10.004](https://doi.org/10.1016/j.envsoft.2007.10.004)
- Hill T, Lewicki P (2006) Statistics: methods and applications: a comprehensive reference for science, industry and data mining. StatSoft, Tulsa
- Huth R (2003) Sensitivity of local daily temperature change estimates to the selection of downscaling models and predictors. *J Clim* 17:640–652
- Huth R, Kysely J, Pokorna L (2000) A GCM simulation of heat waves, dry spells, and their relationships to circulation. *Clim Change* 46:32
- Huth R, Kysely J, Dubrovsky M (2001) Time structure of observed, GCM-simulated, downscaled, and stochastically generated daily temperature series. *J Clim* 14(20):4047–4061
- Huth R, Kliegrová S, Metelka L (2008) Non-linearity in statistical downscaling: does it bring an improvement for daily temperature in Europe? *Int J Climatol* 28(4):465–477. doi:[10.1002/joc.1545](https://doi.org/10.1002/joc.1545)
- IPCC (2000) Emissions scenarios. A special report of working group iii of the intergovernmental panel on climate change. Cambridge University Press, Cambridge
- James G, Witten D, Hastie T, Tibshirani R (2013) An Introduction to statistical learning: with applications in R. Springer texts in statistics. Springer, New York
- Jarosch AH, Anslow FS, Clarke GKC (2012) High-resolution precipitation and temperature downscaling for glacier models. *Clim Dyn* 38(1–2):391–409. doi:[10.1007/S00382-010-0949-1](https://doi.org/10.1007/S00382-010-0949-1)
- Jeong DI, St-Hilaire A, Ouarda TBMJ, Gachon P (2012) CGCM3 predictors used for daily temperature and precipitation downscaling in Southern Quebec, Canada. *Theo Appl Climatol* 107(3–4):389–406. doi:[10.1007/S00704-011-0490-0](https://doi.org/10.1007/S00704-011-0490-0)
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77(3):437–471
- Karl TR, Wang WC, Schlesinger ME, Knight RW, Portman D (1990) A method of relating general circulation model simulated climate to the observed local climate. Part I: seasonal statistics. *J Clim* 3(10):1053–1079
- Khalili M, Van Nguyen VT, Gachon P (2013) A statistical approach to multi-site multivariate downscaling of daily extreme temperature series. *Int J Climatol* 33(1):15–32. doi:[10.1002/joc.3402](https://doi.org/10.1002/joc.3402)
- Khan M, Coulibaly P, Dibike Y (2006) Uncertainty analysis of statistical downscaling methods. *J Hydrol* 319(1–4):357–382. doi:[10.1016/j.jhydrol.2005.06.035](https://doi.org/10.1016/j.jhydrol.2005.06.035)
- Kirschbaum MUF (1995a) Ecophysiological, ecological, and soil processes in terrestrial ecosystems: a primer on general concepts and relationships. Impacts, adaptations and mitigation of climate change: scientific-technical analyses. IPCC WGII, Geneva

- Kirschbaum MUF (1995b) The temperature-dependence of soil organic-matter decomposition, and the effect of global warming on soil organic-C storage. *Soil Biol Biochem* 27(6):753–760. doi:[10.1016/0038-0717\(94\)00242-S](#)
- Laprise R (2008) Regional climate modelling. *J Comput Phys* 227(7):3641–3666. doi:[10.1016/j.jcp.2006.10.024](#)
- Lempert RJ, Collins MT (2007) Managing the risk of uncertain thresholds responses: comparison of robust, optimum, and precautionary approaches. *Risk Anal* 27:1009–1026
- Lim YK, Stefanova LB, Chan SC, Schubert SD, O'Brien JJ (2011) High-resolution subtropical summer precipitation derived from dynamical downscaling of the NCEP/DOE reanalysis: how much small-scale information is added by a regional model? *Clim Dyn* 37(5–6):1061–1080. doi:[10.1007/S00382-010-0891-2](#)
- Matsumura K, Gaitan CF, Sugimoto K, Cannon AJ, Hsieh WW (2014) Maize yield forecasting by linear regression and artificial neural networks in Jilin, China. *J Agric Sci* 153(03):399–410. doi:[10.1017/S0021859614000392](#)
- Maurer EP, Hidalgo HG (2008) Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrol Earth Syst Sc* 12(2):551–563
- Mearns LO, Giorgi F, Whetton P, Pabon D, Hulme M, Lal M (2003) Guidelines for use of climate scenarios developed from regional climate model experiments (trans: IPCC DDCot). IPCC Technical Report. IPCC
- Mearns LO, Gutowski WJ, Jones R, Leung R, McGinnis S, Nunes A, Qian Y (2007) The North American regional climate change assessment program dataset. Boulder, CO doi:[10.5065/D6RN35ST](#)
- Mirhosseini M, Sharifi F, Sedaghat A (2011) Assessing the wind energy potential locations in province of Semnan in Iran. *Renew Sustain Energy Rev* 15(1):449–459. doi:[10.1016/j.rser.2010.09.029](#)
- Monsalve Saenz G (2002) Hidrologia en la Ingenieria, 2nd edn. Escuela Colombiana de Ingenieria, Bogota
- Music B, Sykes C (2011) CRCM diagnostics for future water resources in opg priority watersheds
- Nicholas RE, Battisti DS (2012) Empirical downscaling of high-resolution regional precipitation from large-scale reanalysis fields. *J Appl Meteorol Clim* 51(1):100–114. doi:[10.1175/Jamc-D-11-04.1](#)
- Ontario_Ministry_of_Natural_Resources (2012) Ontario invasive species strategic plan 2012. Ontario_Ministry_of_Natural_Resources, Toronto
- Ouyang F, Lu H, Zhu Y, Zhang J, Yu Z, Chen X, Li M (2014) Uncertainty analysis of downscaling methods in assessing the influence of climate change on hydrology. *Stoch Env Res Risk Assess* 28:991–1010. doi:[10.1007/s00477-013-0796-9](#)
- Radić V, Clarke GKC (2011) Evaluation of IPCC models' performance in simulating late-twentieth-century climatologies and weather patterns over North America. *J Clim* 24(20):5257–5274. doi:[10.1175/jcli-d-11-00011.1](#)
- Saba VS, Stock CA, Spotila JR, Paladino FV, Santidrian Tomillo P (2012) Projected response of an endangered marine turtle population to climate change. *Nat Clim Change* 2(11):814–820. doi:[10.1038/nclimate1582](#)
- Salathé EP, Steed R, Mass CF, Zahn PH (2008) A high-resolution climate model for the U.S. Pacific Northwest: mesoscale feedbacks and local responses to climate change*. *J Clim* 21(21):5708–5726. doi:[10.1175/2008jcli2090.1](#)
- Semenov MA, Stratonovitch P (2010) Use of multi-model ensembles from global climate models for assessment of climate change impacts. *Clim Res* 41:1–14. doi:[10.3354/cr00836](#)
- Stock CA, Alexander MA, Bond NA, Brander KM, Cheung WWL, Curchitser EN, Delworth TL, Dunne JP, Griffies SM, Haltuch MA, Hare JA, Hollowed AB, Lehodey P, Levin SA, Link JS, Rose KA, Rykaczewski RR, Sarmiento JL, Stouffer RJ, Schwing FB, Vecchi GA, Werner FE (2011) On the use of IPCC-class models to assess the impact of climate on living marine resources. *Prog Oceanogr* 88(1–4):1–27. doi:[10.1016/j.pocan.2010.09.001](#)
- Stoner AMK, Hayhoe K, Yang X, Wuebbles DJ (2013) An asynchronous regional regression model for statistical downscaling of daily climate variables. *Int J Climatol* 33(11):2473–2494. doi:[10.1002/joc.3603](#)
- Teutschbein C, Seibert J (2013) Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? *Hydrol Earth Syst Sci* 17(12):5061–5077. doi:[10.5194/hess-17-5061-2013](#)
- van Vuuren DP, Smith SJ, Riahi K (2010) Downscaling socioeconomic and emissions scenarios for global environmental change research: a review. *Wiley Interdiscip Rev Clim Change* 1(3):393–404. doi:[10.1002/wcc.50](#)
- von Storch H (1999) On the use of “inflation” in statistical downscaling. *J Clim* 12:3505–3506
- Vrac M, Stein ML, Hayhoe K, Liang XZ (2007) A general method for validating statistical downscaling methods under future climate change. *Geophys Res Lett*. doi:[10.1029/2007gl030295](#)

- Walthall CL, Hatfield J, Backlund P (2012) Climate change and agriculture in the United States: effects and adaptation. USDA Technical Bulletin 1935. USDA, Washington
- West JW (2003) Effects of heat-stress on production in dairy cattle. *J Dairy Sci* 86(6):2131–2144
- Wilby R (1994) Stochastic weather type simulation for regional climate change impact assessment. *Water Resour Res* 30(12):3395–3403
- Wigley TML, Jones PD, Briffa KR, Smith G (1990) Obtaining subgrid scale information from coarse resolution general circulation model output. *J Geophys Res* 95(D2):1943–1953
- Wilby RL, Wigley TML, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS (1998) Statistical downscaling of general circulation model output: a comparison of methods. *Water Resour Res* 34(11):2995–3008
- Wilby R, Charles S, Zorita E, Timbal B, Whetton P, Mearns LO (2004) Guidelines for use of climate scenarios developed from statistical downscaling methods IPCC 2004. Intergovernmental Panel on Climate Change
- Wilks DS (2011) Statistical methods in the atmospheric sciences. International geophysics series, vol 100, 3rd edn. Academic Press
- Willmott CJ, Robeson SM, Matsuura K (2012) A refined index of model performance. *Int J Climatol* 32(13):2088–2094. doi:[10.1002/joc.2419](https://doi.org/10.1002/joc.2419)