

# The new national climate change documents of Mexico: what do the regional climate change scenarios represent?

Francisco Estrada · Benjamín Martínez-López ·  
Cecilia Conde · Carlos Gay-García

Received: 21 June 2010 / Accepted: 22 April 2011 / Published online: 8 June 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** This paper presents a review of the methodology applied for generating the regional climate change scenarios utilized in important National Documents of Mexico, such as the Fourth National Communication to the United Nations Framework Convention on Climate Change, the Fourth National Report to the Convention on Biological Diversity and The Economics of Climate Change in Mexico. It is shown that these regional climate change scenarios, which are one of the main inputs to support the assessments presented in these documents, are an example of the erroneous use of statistical downscaling techniques. The arguments presented here imply that the work based on such scenarios should be revised and therefore, these documents are inadequate for supporting national decision-making.

## 1 Introduction

In 2009 the Mexican government published several documents to support decision-making processes regarding climate change. Some of the most important technical documents to assess the potential effects of climate change and the risk this phenomenon could possess for the country are the Fourth National Communication to the United Nations Framework Convention on Climate Change (INE-SEMARNAT 2009), the Fourth National Report to the Convention on Biological Diversity

---

F. Estrada (✉) · B. Martínez-López · C. Conde · C. Gay-García  
Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México,  
Mexico, D.F. 04510, Mexico  
e-mail: feporrua@atmosfera.unam.mx

(CONABIO and SEMARNAT 2009)<sup>1</sup> and *The Economics of Climate Change in Mexico* (SEMARNAT and SHCP 2009).

These documents address different aspects of climate change that the government considers particularly important for the development of climate change policies regarding the potential impacts of climate change, the convenience of adopting adaptation and mitigation strategies and for building a national position towards the imminent post-Kyoto international agreement.

One of the main inputs for conducting national climate change assessments are undoubtedly regional climate change scenarios. The assessment results are therefore highly sensitive to factors such as the spatial distribution and the magnitude of the changes in climate (Wilby and Harris 2006). This is in fact the main motivation for the development and application of downscaling techniques for climate change studies (Christensen et al. 2007; Fowler et al. 2007; Giorgi et al. 2001; Wilby et al. 2004).

It has been widely recognized that the spatial resolution of current state-of-the-art climate models is too coarse for most of the regional climate change assessments needed for assisting decision-making (Benestad et al. 2008; Christensen et al. 2007). Downscaling techniques aim to satisfy this need by complementing the information obtained by physical climate models with the effect of the local scale physiographic features. Statistical downscaling methods have shown to be a competitive and cheaper alternative for achieving this task in comparison to dynamical downscaling (Wilby et al. 2004).

The least that should be required of a statistical downscaling method is that the spatial patterns it produces are not a statistical artifact but that they really reflect some of the most important local scale features. If the method cannot fulfill this requirement in a reliable manner, the result amounts to trading physically consistent patterns produced by General Circulation Models (GCMs) with possibly irrelevant spatial patterns generated by statistically inadequate models. In this case, the scaling coefficients are spurious and the resulting patterns are arbitrary (Estrada et al., A cautionary note on statistical downscaling methods for climate change, submitted). That is, even though the original climate change scenarios could have been produced by state-of-the-art climate models, which provide the best representation available of the climate system, after performing a poorly applied or incorrect downscaling method there are no longer reasons to believe that the new spatial patterns are physically consistent or even physically possible.

The resulting downscaled scenarios are not only clearly worse than not performing any downscaling at all, but entail a cascade of errors that propagate to the assessment studies and ultimately to the decision-making processes. As it is shown in this paper, the regional climate change scenarios in all of the above mentioned national documents are an example of the erroneous use of statistical downscaling techniques. Our results indicate that incorrect application of statistical downscaling techniques can affect relevant aspects of the work based on such scenarios. Therefore, these

---

<sup>1</sup>Climate change and its potential impacts are not the main issue discussed in this report. Nevertheless, all of the information that is presented in this publication regarding these topics is based on Magaña and Gómez (2008). This latter publication is based on the same regional scenarios described in Magaña and Caetano (2007), Zermeño (2008) and Magaña (2010) that are reviewed in this paper and are subject to the critique that is presented here.

documents should be revised and corrected in order to make them adequate for supporting decision-making.

There is a large literature on how the statistical methods for downscaling can be most properly applied for constructing regional and local scale climate change scenarios, as well as on some of the potential pitfalls that can occur when using these techniques and how to prevent them. Accordingly, this paper does not offer a review nor examples of the available methodologies, but instead refers the interested reader to Benestad et al. (2008), Maraun et al. (2010), Frias et al. (2006), Vrac et al. (2007), von Storch et al. (2000), Giorgi et al. (2001), Wilby and Wigley (1997, 2000), Fowler et al. (2007), Estrada et al. (submitted), among others.

## 2 Description and discussion of the proposed statistical downscaling methodology

### 2.1 Data and methods

The methodology used for producing the regional climate change scenarios in the Fourth National Communication to the United Nations Framework Convention on Climate Change (INE-SEMARNAT 2009), the Fourth National Report to the Convention on Biological Diversity (CONABIO and SEMARNAT 2009) and The Economics of Climate Change in Mexico (SEMARNAT and SHCP 2009) is documented in Magaña and Caetano (2007), Zermeño (2008), INE-SEMARNAT (2009) and Magaña (2010) and is briefly described in the following paragraphs.

The downscaling approach adopted for generating the regional climate change scenarios is the Model Output Statistics (MOS) proposed by Glahn and Lowry (1972), and was implemented by means of the Climate Predictability Tool (<http://iri.columbia.edu>), which is an automated statistical downscaling toolbox based on canonical correlations and principal component linear regression. It is important to note that the CPT was not used in its original version, nor for its original purpose (seasonal prediction). Magaña (2010, page 30) states that he modified the CPT's original programming code in order to make it "adequate for generating climate change scenarios". The modification consisted basically of making the CPT consider only the first principal component as the predictor variable.

According to Glahn and Lowry (1972) the MOS downscaling approach consists in estimating a statistical relationship between an observed local predictand and one or more large scale predictors that are the output of a dynamical model at some projection time. Then this relationship is applied to model output to estimate the projected values at local scales.

For the regional climate change scenarios for Mexico, the predictands were the Climate Research Unit (CRU) TS3.0  $0.5^\circ \times 0.5^\circ$  gridded database of observed climate variables produced by the University of East Anglia (Mitchell and Jones 2005).

The predictors were chosen to be the first principal components obtained from the 20th Century Climate Experiment (20c3m) runs produced by a variety of GCMs for the IPCC's Fourth Assessment Report. The principal components were estimated for a region encompassing Mexico and the southern part of the United States of America. According to Magaña and Caetano (2007), Zermeño (2008) and Magaña (2010), the first principal component is supposed to represent the warming trend

over the region during the twentieth century. Magaña (2010) argues that the first principal component can be used to evaluate the climate sensitivity to changes in radiative forcing and therefore he considers it as the most important explanatory variable for downscaling purposes. These arguments of course are true only if (1) the warming trend is the dominant mode of interannual variability (i.e. explains the most predictand variability), and (2) the first principal component is well-separated statistically from the second (North et al. 1982). In fact, neither condition is satisfied for the applications of modified CPT examined here.

The statistical model chosen was simple linear regression. The climate variables considered for the regional climate change scenarios were monthly temperature and precipitation, while the calibration period for the downscaling models was from 1901 to 1969.<sup>2</sup>

## 2.2 Discussion

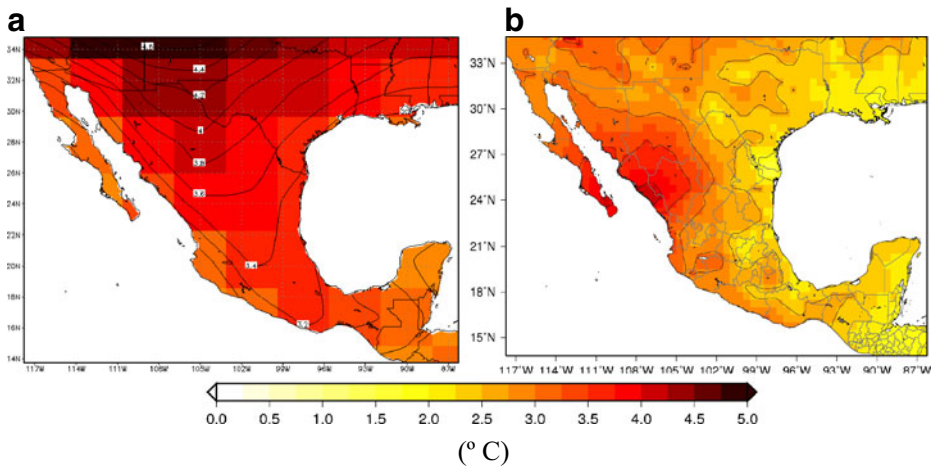
Figure 1 was taken from Zermeño (2008) and Magaña (2010),<sup>3</sup> and provides an example of the proposed methodology. As can be seen from this figure, after applying this statistical downscaling methodology, the original GCM—in this case the ECHO-G model (Legutke and Voss 1999), developed at the Max-Planck-Institute for Meteorology—and the downscaled spatial patterns (panel a and b, respectively) are markedly different. Notice that, after the downscaling has been applied, the magnitude of the temperature increment has decreased notably for most of the country. Moreover, in the model simulation (Fig. 1a) the largest increments in temperature are shown in the north of the country and the temperature increment gradually decreases as the latitude decreases. In contrast, the downscaled pattern (Fig. 1b) shows that the largest temperature increment would occur in a well defined region in the northwest part of Mexico, while the temperature increments are reduced importantly for the rest of the country and for the south of the United States of America. On average, the increments in temperatures are about 1/2 what the climate model projects.

If these changes are *in fact* due to physical processes that are a consequence of climate change and not captured by the GCM, this is one of the most important corrections that downscaling techniques can offer, and is one of the most important reasons for using these methodologies. Nevertheless, especially when such large differences arise, it should be mandatory to verify that they are not produced by some statistical artifact.

---

<sup>2</sup>It is important to note that the objective of this paper is to review the downscaling methodology that was proposed and used for conducting some of the most important climate change documents of Mexico. The authors do not intend to present an evaluation of the performance or capabilities of the different downscaling toolboxes that are available, nor to compare the different approaches and techniques that are used for downscaling climate change scenarios. The reader is referred to Benestad et al. (2008) and the references therein for a comprehensive review of downscaling methods.

<sup>3</sup>Note that this same downscaled scenario (Fig. 1, panel b) can also be found in the Fourth National Communication to the United Nations Framework Convention on Climate Change (INE-SEMARNAT 2009) and in The Economics of Climate Change in México (SEMARNAT and SHCP 2009).



**Fig. 1** Climate change scenario for 2080–2099 using the ECHO G model and the A1B emissions scenario. **a** Shows the original GCM change patterns with a spatial resolution of  $3.7^\circ \times 3.7^\circ$  and **b** shows the downscaled scenario with a spatial resolution of  $0.5^\circ \times 0.5^\circ$

As will be shown in the following section, all the physics embedded in the original GCM scenarios have been replaced by arbitrary and random patterns and magnitudes produced by a flawed downscaling methodology. As such, the results of impact, vulnerability, adaptation and risk assessments produced using the scenario in Fig. 1b) will be based on an arbitrary climate scenario for which there is no reason to believe is physically relevant or consistent with climate change.

Due to the fact that the underlying methodology is flawed, the results of the assessments that were conducted using these scenarios should be revised and they should not be used for assisting decision-making, much less to support the development of Mexico's climate change policies.

### 3 Testing the proposed downscaling methodology

In this section, we test the validity of the proposed downscaling methodology by means of simple statistical analysis and by analyzing the behavior and statistical significance of the slope coefficient in the linear regression model.

For these purposes, the same Climate Research Unit monthly temperature database in Magaña (2010), Zermeño (2008) and Magaña and Caetano (2007), is used as predictands. We use the five ECHOG 20c3m runs as predictors (Min 2006a, b, c; Lee 2006a, b). Although the analyses presented here were performed for all the months in a year, only the results for January temperature are shown because the results are very similar for all of them.

As will be shown in this section our findings are conclusive: the downscaling methodology used for the new National Documents of Mexico is flawed.

### 3.1 Are the first principal components of the different ECHOG runs correlated?

Do they share a warming trend? Can it be assumed that the first principal component necessarily represents the warming trend?

A necessary condition for the proposed methodology to be consistent would be that the first principal components of different simulations of the same model and under the same forcing scenario would be well correlated, because they are assumed to represent the warming trend. In this case, no matter which model simulation is used for downscaling, a significant relation of roughly similar magnitude and sign would be obtained.

Table 1 shows the correlation matrix of the five first principal components for January temperatures of the different ECHOG 20c3m runs. It is important to keep in mind that, given the assumed time-series properties of the variables (trending and autocorrelated), it should be easy to find significant (and possibly spurious) correlations for any significance level (see, for example, Yule 1926). Nevertheless, as can be seen from Table 1, although the correlation coefficients range from  $-0.16$  to  $0.14$ , these variables can be considered as linearly independent as revealed by their corresponding *p*-values.

All of these time series are independent realizations of the same data generating process: which, if any, should be used for statistical downscaling? In order for the proposed methodology to be useful, all of these simulations should reflect climate change to some extent, and would have similar relationships with the observed variable. If not, results would depend on which model realization is used and the value of the estimated coefficients would be meaningless for downscaling purposes. Clearly, if the simulated series are independent, this condition cannot be satisfied.

The independence between these time series is a direct consequence of the concept of climate simulation. A climate simulation cannot, and is not intended to, predict the actual realized values of a climate variable, but only to describe some of the climate's characteristics under a particular experiment. As such, different monthly model runs under the same emission scenario need not to be correlated, nor they have to be correlated with the observed climate series, unless there was a prominent trend on them. Even in such a case, and as stated before, the statistical relationship could then be spurious.

According to Magaña (2010), Magaña and Caetano (2007) the reason for using the first principal component of the simulated fields is that this variable is supposed to represent the warming trend over the region during the twentieth century. This hypothesis can be tested: do the first principal components of the different 20c3m runs show a warming trend? and if so, do they share the same warming trend? Based

**Table 1** Correlation matrix of the five ECHOG 20c3m runs for January

	Run 1	Run 2	Run 3	Run 4	Run 5
Run 1	1.000				
Run 2	$-0.023$ (0.822)	1.000			
Run 3	$-0.117$ (0.246)	$-0.055$ (0.587)	1.000		
Run 4	$0.137$ (0.176)	$-0.157$ (0.119)	$0.009$ (0.928)	1.000	
Run 5	$0.111$ (0.274)	$0.046$ (0.651)	$0.112$ (0.266)	$-0.054$ (0.591)	1.000

*P* values in parentheses. None of the correlations are statistically significant at even the 10% level

**Table 2** Estimations of the slope coefficient in regression (1) for the five first principal components of January temperature of the ECHOG 20c3m runs

Model run	$\beta$	<i>P</i> value
Run 1	0.003	0.505
Run 2	0.000	0.925
Run 3	0.009	0.016
Run 4	0.006	0.072
Run 5	0.005	0.148

on Table 1 correlations, it is unlikely more than one of the five has any warming trend at all.

To examine these questions more closely, consider the following regression model

$$T_{r,t} = \alpha + \beta t + u_t, t = 1, \dots, n. \quad (1)$$

where  $T_{r,t}$  is the first principal component of the January temperature series simulated by the ECHOG 20c3m run  $r$ , *time* is a linear time trend<sup>4</sup> and  $u_1, \dots, u_n$  a sequence of independent and identically distributed random errors such that  $u_t \sim N(0, \sigma^2)$ , while  $\alpha$  and  $\beta$  are unknown but fixed, constant parameters.

Table 2 shows the slope coefficient  $\beta$  and the corresponding p-value obtained by estimating the linear regression (1) by Ordinary Least Squares (OLS). The coefficient  $\beta$  is not significant at conventional levels (5%) for four of the five first principal components.

That is, only in one of the five time series a warming trend *could* be present but not in the other four. Clearly, it can not be assumed that these time series share a common warming trend. Therefore, as in the example of the correlation coefficients above, the downscaling results would depend on which model run is used.

Once again, considering that there is no common trend in the first principal component of the five available runs, it is clear that the proposed methodology cannot produce consistent results.

Furthermore, is it correct to assume that the first principal component necessarily represents the warming trend? As is well known, the first principal component represents the principal mode of variability in any particular multivariate data set. As such, unless the warming trend in the region encompassing Mexico and the south of the United States of America is large in comparison with other sources of interannual variability, the first principal component will not represent the warming trend of the twentieth century. In addition, even if a warming trend could be present in the first (unrotated) principal component it may be poorly separated statistically from other modes of variation and consequently be contaminated with the influence of other signals that are known to have a large impact on the Northern Hemisphere such as El Niño/Southern Oscillation (ENSO), the North Atlantic Oscillation (NAO) and the Pacific/North American Pattern (PNA), for example.

In order to test if the first principal component represents the warming trend, unrotated (PCA) and rotated principal component (RPCA) analyses were

<sup>4</sup>The CUSUM, CUSUMQ and Quandt-Andrews tests were performed to evaluate if a segmented trend (non-linear trend) was a more adequate representation than a linear trend. All these tests provided evidence of parameter stability, supporting the use of a linear trend.

**Table 3** Correlation between the first principal component and the PNA, NAO and ENSO

	PNA	NAO	ENSO
PC1 (unrotated)	<b>0.57</b>	<b>-0.50</b>	0.21
PC1 (rotated)	<b>-0.41</b>	<b>0.38</b>	-0.22

PC1 represents the first principal component. Statistically significant correlations at 5% levels or lower are shown in bold

performed on the observed temperature field over the twentieth century in the region of interest.<sup>5</sup>

Similar to the results presented in Table 2, no significant trend could be found in the (rotated and unrotated) first principal components estimated from the observations. Conducting simple correlation analyses, it can be shown that, in both the PCA and RPCA, the first principal components represent natural internal interannual variability associated with the PNA and NAO patterns in opposite polarities (Table 3). This is in agreement with the climate variability literature which has shown that the PNA and NAO teleconnection patterns are two of the most prominent modes of low-frequency climate variability in the Northern Hemisphere (see, for example, Wallace and Gutzler 1981; Livezey and Smith 1999; Barnston and Livezey 1987). Their projection onto the same rotated mode with opposite polarities is consistent with a result obtained by Livezey and Smith (1999).

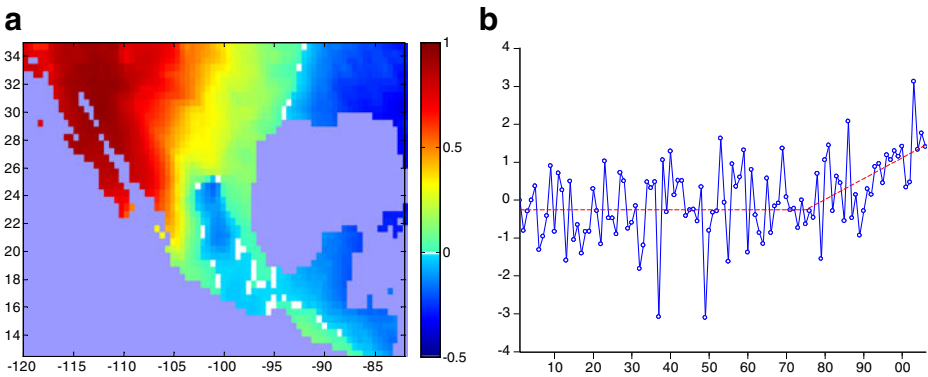
If climate models do a fair job reproducing the observed temperatures during the twentieth century, the first principal components that are obtained from them over the same region should represent the natural internal variability associated mainly with the PNA and the NAO, not to global warming. In this manner, the methodology of Magaña and Caetano (2007), Zermelo (2008) and Magaña (2010) does not use the climate change signal to produce the regional climate change scenarios as they intended to do, but uses instead a combination of natural internal variability signals, i.e. climate noise. Consequently, the predictor variable chosen for downscaling can convey little, if any, of the climate change information produced by the GCM.

Evidently, it was erroneous to assume a priori that the first principal component must represent the warming trend over the last century, as was done by Magaña and Caetano (2007), Zermelo (2008) and Magaña (2010). Instead, an analysis should have been carried out first to determine which low-frequency signals the first modes represent.

Analyzing the other rotated and unrotated principal components, obtained from the observations, it becomes apparent that there is a significant warming trend in the second principal component. This warming signal is even clearer in the second rotated principal component and it shows a structural break or “hinge” around 1975. A structural break in the slope of the trend function around this year has also been found in the global, regional and even in local mean temperatures and is commonly interpreted as a possible manifestation of global warming (Gay et al. 2009; Livezey et al. 2007; Gil-Alana 2008). A simple areal average of all of the input temperature series indicates a similar trend with a hinge in the same year (not shown here). In the

<sup>5</sup>RPCA requires the determination of how many eigenmodes should be rotated. This paper uses the suggestions for truncation found in O’Lenic and Livezey (1988).





**Fig. 2** Rotated second principal component of observed temperature series during the period 1901–2006. **a** Shows the loadings and **b** the scores

five ECHO-G 20c3m runs this warming trend can be found in either the second or third rotated principal components.

It is important to note that even if the trend mode is adequately extracted by means of RPCA, the signal to noise ratio may be too low (the trend might be too faint)<sup>6</sup> and consequently it may not be possible to establish a consistent (and statistically significant) relationship between observed and different realizations of the simulated variables. In this manner, results would still depend on which model simulation is used and would not be consistent. Performing the same type of analysis presented in Table 1 it can be shown that this is the case of the ECHO-G second and third principal components. The model signal over this small domain is extremely weak to produce consistent results, even if the sample is restricted to the period 1975–2000 where the climate change signal is stronger. Furthermore, even the correlation between the observed trend mode and the ensemble of the model trend modes over this period is very low (0.2 with a p-value of 0.33), confirming that the climate change signal in the model runs is extremely weak.

Figure 2 shows the loadings and the scores (panel a and b, respectively) of the rotated second principal component. The spatial pattern in panel a) shows that most of the warming in the region of study, during the twentieth century, occurred in the northwest of Mexico (Southwest of the United States of America) while most of the country has experienced a much modest warming and even cooling in the Yucatan peninsula. This is in accordance with previous studies on warming trends in North America and at the global scale (Livezey et al. 2007; Hansen et al. 2006). Panel b) of Fig. 2 shows the second principal component with a hinge fit and, obvious from the figure, the warming has taken place dominantly over the last three and a half decades. Note that, in addition, since the calibration period chosen by Magaña and Caetano (2007), Zermeño (2008) and Magaña (2010) is 1901–1969, most of the warming signal in the observed data would have been anyway excluded.

<sup>6</sup>The use of long non-stationary records may also dilute the global warming signal which is much more clear over the last three and a half decades, not over the entire 20<sup>th</sup> century.

### 3.2 A “Perfect Climate Model”

This and the next section consist of graphic demonstrations of the impact of downscaling climate noise rather than climate signal (the warming trend), particularly the inconsistencies that result.

Even if a climate model could perfectly represent the climate of the twentieth century, can the proposed downscaling methodology produce consistent downscaled spatial patterns?

Assume that one of the five ECHOG 20c3m runs is the *observed predictand* and choose one of the other four runs to represent the *predictive variable* produced by a climate model. This would be equivalent to assuming that the climate model is indeed perfect: the “observed” variable and the model simulations are realizations of the same data generating process. All these realizations represent the same variable, share the same scale, model physics and forcings.

Consider the linear regression

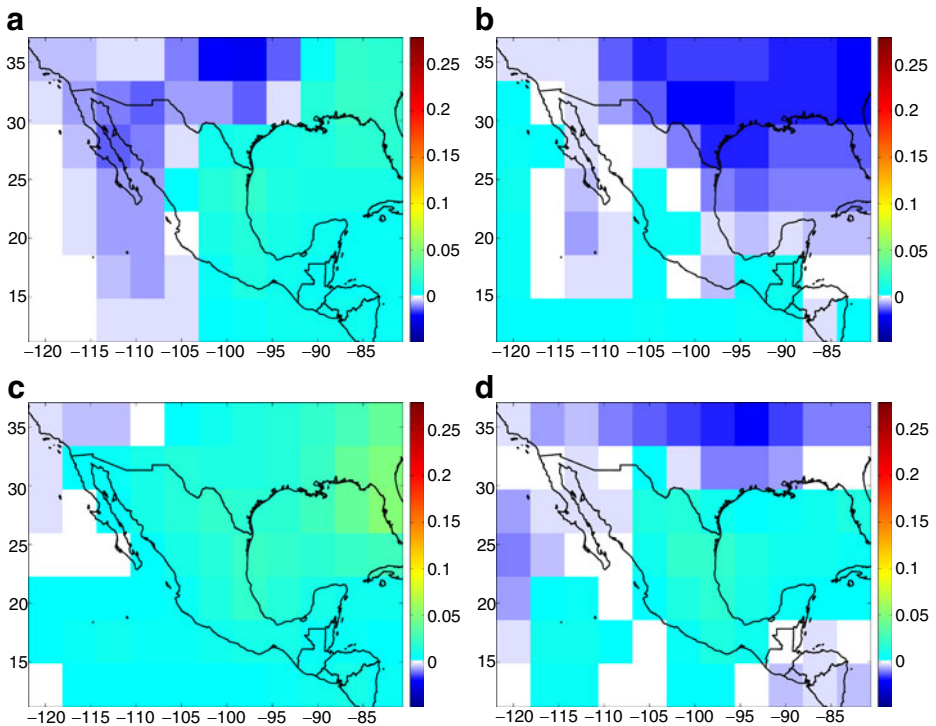
$$Tobs_{i,j,t} = \alpha + \beta T_{r,t} + u_t \quad (2)$$

where  $Tobs_{i,j,t}$  represents the “observed” January temperature series for the coordinate  $i, j$ ;  $T_{r,t}$  represents the first principal component of the January temperature series simulated by the ECHOG 20c3m run  $r$ ;  $u_1, \dots, u_n$  a sequence of independent and identically distributed random errors such that  $u_t \sim N(0, \sigma^2)$ . As in (1),  $\alpha$  and  $\beta$  are unknown but fixed, constant parameters.

Figures 3 and 4 show the values of the slope parameters  $\beta$  and their statistical significance at approximately 5% levels (a t-statistic larger than 1.96 in absolute value) for each grid point in the analyzed region, respectively. From Fig. 3, it becomes apparent that the spatial patterns in each of them are quite different and would lead to very different conclusions in terms of impact, vulnerability and risk assessments. Furthermore, as can be seen from Fig. 4, the value of most of the slopes shown in Fig. 3 is not statistically different from zero: that is, the coefficients  $\beta$  are random variables with a zero mean, therefore the signs and magnitudes in Fig. 3 are meaningless and depend on the particular sample. It is worth noticing that in these examples, none of the slope coefficients are statistically different from zero over Mexico’s territory. Using these coefficients for downscaling would produce random and nonsensical patterns and magnitudes.

The differences between runs of the same model and under the same forcing scenario, are basically due to the model’s internal variability and to the initial conditions chosen for each of the runs, not climate change. For practical matters, these differences can be regarded as random and so can the downscaling patterns obtained using the methodology proposed by Magaña and Caetano (2007), Zermeno (2008) and Magaña (2010). Instead of adding relevant local-scale features for producing regional climate change scenarios that are more suitable for conducting impact, vulnerability and risk assessments, this downscaling methodology completely destroys the patterns produced by the climate model associated with climate change, and replaces them with random patterns and magnitudes. This would invalidate the results of the assessment obtained using such scenarios.

Even under the extreme assumption of having “perfect” models, the results of the proposed downscaling methodology are irrelevant and inappropriate to the



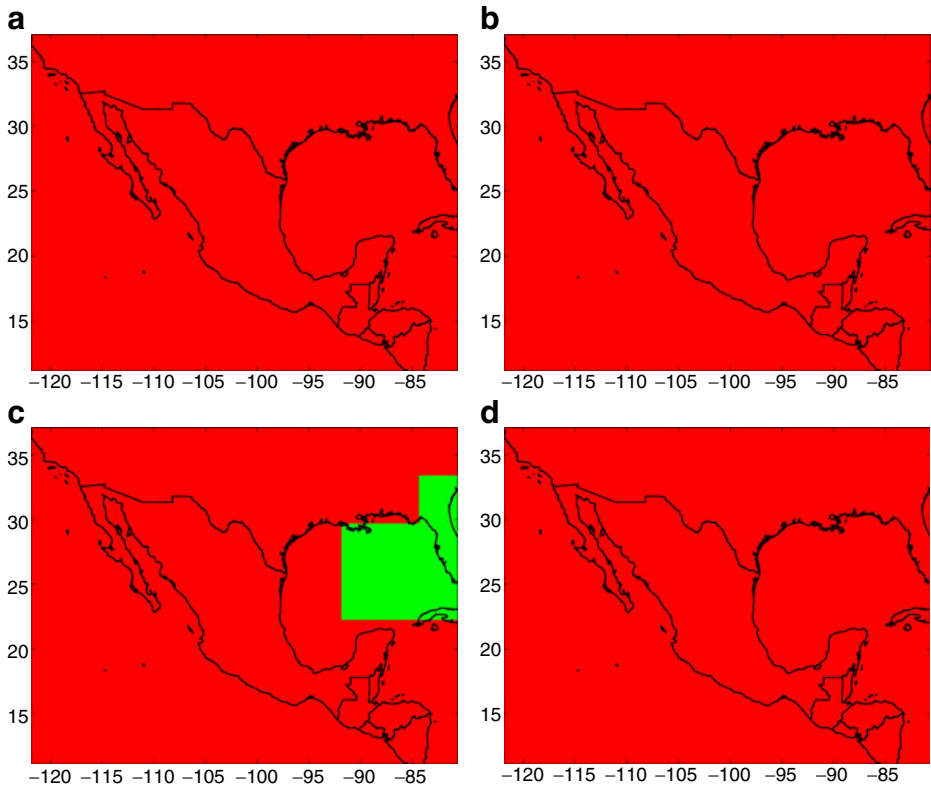
**Fig. 3** Spatial patterns for January's temperature produced by the proposed downscaling methodology. The map shows the slope coefficients in regression (2), using the ECHOG model run 1 under the 20c3m scenario as the predictand and using as predictor: (1) the first principal component of the ECHOG 20c3m run 2 (**a**); (2) the first principal component of the ECHOG 20c3m run 3 (**b**); the first principal component of the ECHOG 20c3m run 4 (**c**); the first principal component of the ECHOG 20c3m run 5 (**d**)

downscaling objectives. The spatial patterns that can be constructed using this methodology are arbitrary and misleading. As is shown in the following section, using observed series as predictands and different model runs as predictors leads to a different set of spatial patterns that are also arbitrary and nonsensical.

Note that even when the trend mode is adequately extracted using RPCA, repeating this experiment leads to the same type of inconsistent spatial patterns and not statistically significant slope coefficients because the model signal is too weak over this small domain to produce a usable response.

### 3.3 The behavior of the slope coefficients of the transfer function when using the observed temperature series and different model runs as explanatory variables

This section follows closely from Section 3.2. It presents the same type of analysis but using as predictand the observed climate. As is shown, the results confirm that the statistical downscaling methodology is flawed.



**Fig. 4** Statistical significance of the slope coefficients in Fig. 2. Green areas denote statistical significance at approximately 5% significance levels (t-statistic larger than 1.96 in absolute value)

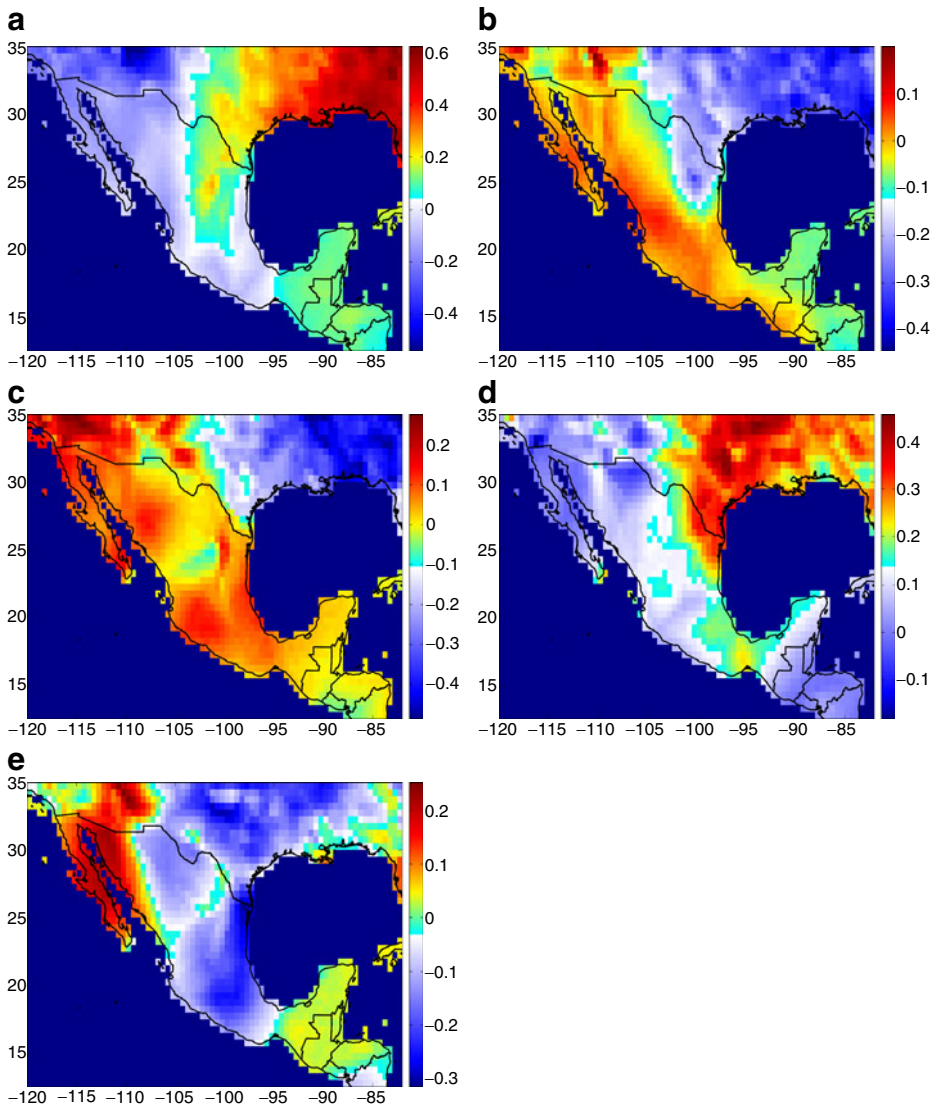
Consider the linear regression

$$Tcr_{i,j,t} = \alpha + \beta T_{r,t} + u_t \quad (3)$$

where  $Tcr_{i,j,t}$  represents the observed January temperature series obtained from the Climate Research Unit database for the coordinate  $i, j$ ;  $T_{r,t}$  represents the first principal component of the January temperature series simulated by the ECHOG 20c3m run  $r$ ;  $u_1, \dots, u_n$  a sequence of independent and identically distributed random errors such that  $u_t \sim N(0, \sigma^2)$ . As in (1) and (2),  $\alpha$  and  $\beta$  are unknown but fixed, constant parameters.

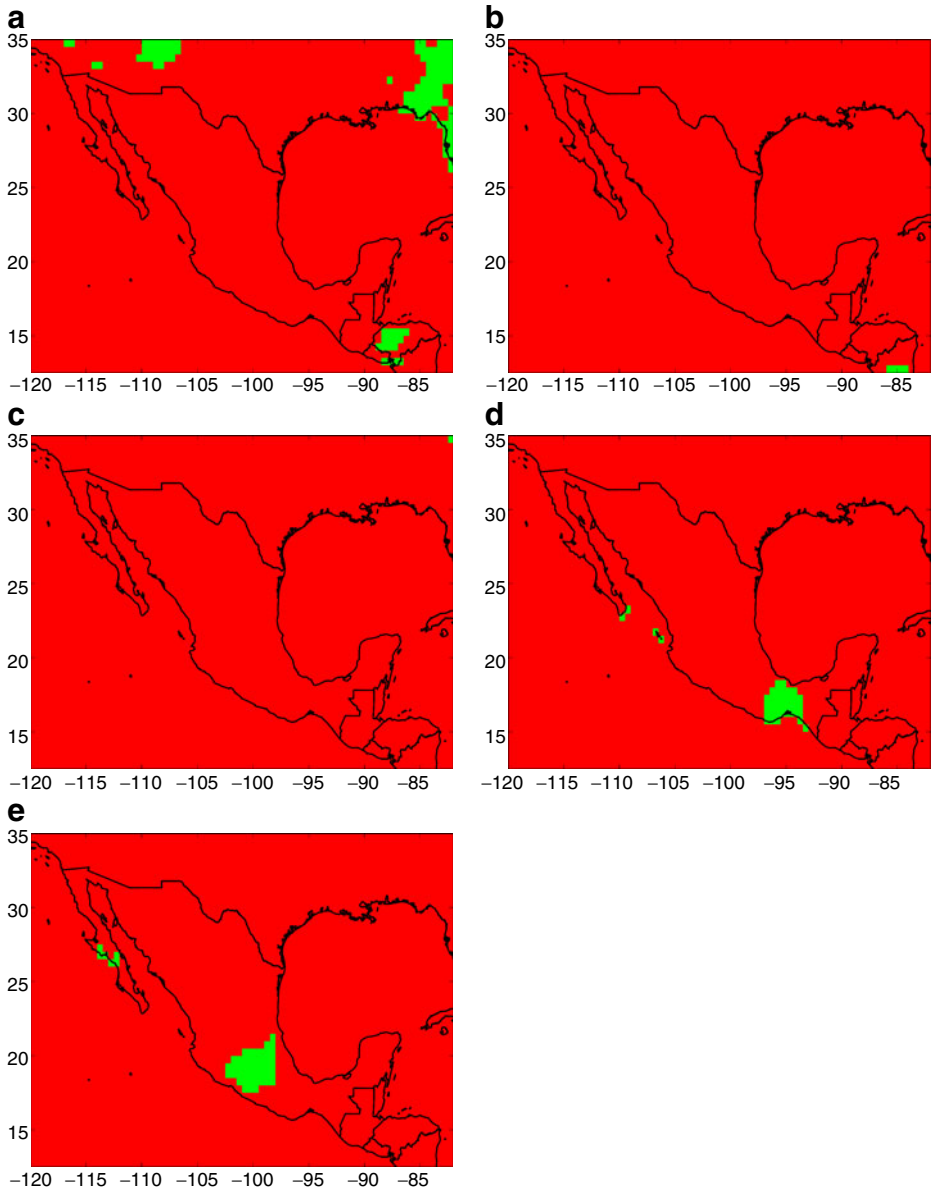
As could be expected from results in the previous sections, when the downscaling methodology is applied to observed data, conclusions are very similar to those inferred from Figs. 3 and 4.

Figures 5 and 6 confirm that: the proposed methodology cannot produce consistent results; the spatial patterns and the magnitudes of the change in such scenarios are arbitrary and misleading; the model physics has been replaced by random patterns and magnitudes: climate noise is represented, not climate change.



**Fig. 5** Spatial patterns for January's temperature produced by the proposed downscaling methodology. The map shows the slope coefficients in regression (3), using the observed January temperature series as the predictand and using as predictor: (1) the first principal component of the ECHOG 20c3m run 1 (**a**); (2) the first principal component of the ECHOG 20c3m run 2 (**b**); the first principal component of the ECHOG 20c3m run 3 (**c**); the first principal component of the ECHOG 20c3m run 4 (**d**); the first principal component of the ECHOG 20c3m run 5 (**e**)

Consequently, assessments constructed using these scenarios are invalidated. For that reason we believe that it is urgent to revise and correct the above mentioned national documents, and that they should not be used for decision and policy-making until then.



**Fig. 6** Statistical significance of the slope coefficients in Fig. 4. *Green areas* denote statistical significance at approximately 5% significance levels (t-statistic larger than 1.96 in absolute value)

#### 4 Conclusions

The downscaling methodology used for scenarios that form bases for Mexican assessment documents failed to represent climate change, replacing it with climate noise. One fundamental error of the methodology was the presumption that the first

unrotated principal component cleanly and separately resolved the climate change signal, when in fact only a higher-order rotated mode could do this. This error could have been avoided easily, if the “black box” predictors had been separately examined at some point, simple tests like those here undertaken, or peer-review conducted.

Another fundamental error was to assume that all of the climate simulations under the 20c3m experiment produced by the climate models included in the IPCC’s Fourth Assessment Report could capture observed warming signal over such a small domain well enough or at all. In addition, using the 1901–1969 sample to calibrate the downscaling models further diluted the warming trend which has taken place in the region mainly over the last 35 years. Consequently, even if the trend mode is extracted adequately by means of rotated principal components, the signal to noise ratio may be too low (trend might be too faint) to establish a consistent, statistically significant and usable relationship between observed and different realizations of the simulated variables. The ECHOG 20c3m runs presented here provide a clear example of this: even if the rotated trend modes are used, the results of the reviewed downscaling methodology would still have been inconsistent and the spatial patterns and magnitudes random. The reason behind this is that, over this small domain, many of the model runs are either missing the warming signal or their signal to noise ratio precludes resolving it. As has been shown in the climate change downscaling literature, the MOS methodology is not adequate for producing climate change scenarios unless the GCM is forced to closely follow observational data during the calibration period (see, for example, Maraun et al. 2010).

A great deal of the social relevance of climate change science depends on its usefulness for decision-making. As such, the methodologies that are proposed for conducting climate change studies should be evaluated before their application in order to minimize the possibility of misinforming decision-making. This paper stresses this need and illustrates the consequences of not having a rigorous review process for producing climate change assessments.

Magaña and Caetano (2007) stress a well known fact in downscaling literature: it is crucial to produce sound climate change scenarios for conducting climate change risk assessments, because the results of these assessments critically depend on them. Among other factors, they underline the importance of using methodologies that ensure that the spatial patterns are physically consistent.

Besides the results in Estrada et al. (submitted) which show that, when using an automated “black-box” downscaling toolboxes for constructing climate change scenarios, it is highly probable to find spurious relations<sup>7</sup> which would lead to arbitrary spatial patterns, there is a deeper problem with the methodology examined here.

In this paper it is shown that the methodology applied for generating the regional climate change scenarios for important National Documents of Mexico, such as the Fourth National Communication to the United Nations Framework Convention on Climate Change, the Fourth National Report to the Convention on Biological Diversity and The Economics of Climate Change in Mexico is flawed, to the extent

---

<sup>7</sup>A spurious relation occurs when a badly specified statistical model provides false evidence of a “significant” relation when in fact there is none: the true value of the parameter relating the dependent and independent variables is zero (Pindyck and Rubinfeld 1998; Estrada et al. submitted).

that the usefulness of the assessments that were conducted using these scenarios is put in to question.

The arguments presented in this paper are not only important in that they alert potential users of these national documents (national and international academy, government, civil society, among others), but also to help prevent other climate change studies from repeating the errors.

In the context of the current debate regarding the use of “grey literature” for supporting decision making processes, we believe that it is important to revise how such documents are being evaluated. The example presented in this paper stresses the need for government agencies to conduct a peer-review process to assess the quality of the technical reports before publishing national documents and making them available for assisting decision-making processes. Given the influence that national documents can have on a wide variety of decision-makers and social agents, as well as for developing national and international public policy, we believe that the review process for these documents should be, at least, as rigorous as it is for scientific publications.

**Acknowledgements** The authors are grateful for the very valuable comments of Robert Livezey and two anonymous reviewers.

## References

- Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon Weather Rev* 115:1083–1126
- Benestad RE, Hanssen-Bauer I, Chen D (2008) Empirical-statistical downscaling. World Scientific Publishing Company, Chennai, p 228
- Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon W-T, Laprise R, Magaña Rueda V, Mearns L, Menéndez CG, Räisänen J, Rinke A, Sarr A, Whetton P (2007) Regional climate projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) *Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, pp 849–940
- CONABIO and SEMARNAT (2009) Fourth National Report to the Convention on Biological Diversity (Cuarto Informe Nacional de México al Convenio sobre Diversidad Biológica (CDB)). Comisión Nacional para el Conocimiento y Uso de la Biodiversidad y Secretaría de Medio Ambiente y Recursos Naturales, México D.F.
- 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* 27:1547–1578
- Frias M, Zorita E, Fernandez J, Rodriguez-Puebla C (2006) Testing statistical downscaling methods in simulated climates. *Geophys Res Lett* 33(L19):807
- Gay C, Estrada F, Sanchez A (2009) Global and hemispheric temperature revisited. *Climatic Change* 94:333–349. doi:10.1007/s10584-008-9524-8
- Gil-Alana LA (2008) Time trend estimation with breaks in temperature time series. *Climatic Change* 89(3–4):325–337. doi:10.1007/s10584-008-9407-z
- Giorgi F, Hewitson B, Christensen J, Hulme M, von Storch H, Whetton P, Jones R, Mearns L, Fu C (2001) Regional climate information—evaluation and projections In: Houghton JT (ed) *Climate change 2001: the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, 583–638
- Glahn HR, Lowry DA (1972) The use of model output statistics (MOS) in objective weather forecasting. *J Appl Meteorol* 11:1203–1211
- Hansen J, Sato Mki, Ruedy R, Lo K, Lea DW, Medina-Elizade M (2006) Global temperature change. *Proc Natl Acad Sci* 103:14288–14293. doi:10.1073/pnas.0606291103



- INE-SEMARNAT (2009) Fourth National Communication to the United Nations Framework Convention on Climate Change (Cuarta Comunicación Nacional ante la Convención Marco de las Naciones Unidas sobre el Cambio Climático). Instituto Nacional de Ecología y Secretaría de Medio Ambiente y Recursos Naturales, México D.F
- Lee (2006a) ECHOG\_A45\_20C3M\_RUN4: 4th member of 20C3M simulations for IPCC AR4. World Data Center for Climate. CERA-DB “ECHOG\_A45\_20C3M\_RUN4” [http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG\\_A45\\_20C3M\\_RUN4](http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG_A45_20C3M_RUN4)
- Lee (2006b) ECHOG\_A46\_20C3M\_RUN5: 5th member of 20C3M simulations for IPCC AR4. World Data Center for Climate. CERA-DB “ECHOG\_A46\_20C3M\_RUN5” [http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG\\_A46\\_20C3M\\_RUN5](http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG_A46_20C3M_RUN5)
- Legutke S, Voss R (1999) The hamburg atmosphere-ocean coupled circulation model ECHO-G. Technical report, No. 18, German Climate Computer Centre (DKRZ), Hamburg, p 62
- Livezey RE, Smith TM (1999) Covariability of aspects of North American climate with global sea surface temperatures on interannual to interdecadal timescales. *J Clim* 12:289
- Livezey RE, Vinnikov KY, Timofeyeva MM, Tinker R, van den Dool HM (2007) Estimation and extrapolation of climate normals and climate trends. *J Appl Meteor Climatol* 46:1759–1776
- Magaña V (2010) Guía para Generar y Aplicar Escenarios Probabilísticos Regionales de Cambio Climático en la Toma de Decisiones. Instituto Nacional de Ecología, Secretaría de Medio Ambiente y Recursos Naturales, México D.F. Available at <http://zimbra.ine.gob.mx/escenarios/>
- Magaña V, Caetano E (2007) Pronóstico climático estacional regionalizado para la República Mexicana como elemento para la reducción de riesgo, para la identificación de opciones de adaptación al cambio climático y para la alimentación del sistema: cambio climático por estado y por sector. Technical report. Instituto Nacional de Ecología, Secretaría de Medio Ambiente y Recursos Naturales, México D.F. Available at <http://www.ine.gob.mx/descargas/climatico/e2007o.pdf>
- Magaña V, Gómez L (2008) Cambio climático y biodiversidad: avances y retos en México, En: Manson, R. y Jardel, Perturbaciones y desastres naturales: impactos sobre las ecorregiones, la biodiversidad y el bienestar socioeconómico en México. En R. Dirzo, R. Gonzalez y I. March (eds), Capital natural de México, Vol. II: Estado de conservación y tendencias de cambio. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México
- Maraun D, et al (2010) Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Rev Geophys* 48:RG3003. doi:10.1029/2009RG000314
- Min (2006a) ECHOG\_A42\_20C3M\_RUN1: 1st member of 20C3M simulations for IPCC AR4. World Data Center for Climate. CERA-DB “ECHOG\_A42\_20C3M\_RUN1” [http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG\\_A42\\_20C3M\\_RUN1](http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG_A42_20C3M_RUN1)
- Min (2006b) ECHOG\_A43\_20C3M\_RUN2: 2nd member of 20C3M simulations for IPCC AR4. World Data Center for Climate. CERA-DB “ECHOG\_A43\_20C3M\_RUN2” [http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG\\_A43\\_20C3M\\_RUN2](http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG_A43_20C3M_RUN2)
- Min (2006c) ECHOG\_A44\_20C3M\_RUN3: 3rd member of 20C3M simulations for IPCC AR4. World Data Center for Climate. CERA-DB “ECHOG\_A44\_20C3M\_RUN3” [http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG\\_A44\\_20C3M\\_RUN3](http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ECHOG_A44_20C3M_RUN3)
- Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int J Climatol* 25:693–712. doi:10.1002/joc.1181
- North GR, Bell TL, Cahalan RF, Moeng FJ (1982) Sampling errors in the estimation of empirical orthogonal functions. *Mon Weather Rev* 110:699–706
- O’Lenic E, Livezey RE (1988) Practical considerations in the use of rotated principal components analysis (RPCA) in diagnostic studies of upper-air height fields. *Mon Wea Rev* 116:1682–1689
- Pindyck RS, Rubinfeld DL (1998) *Econometric models & economic forecasts*, 4th edn. McGraw-Hill, Singapore
- SEMARNAT and SHCP (2009) The economics of climate change in Mexico (La Economía del Cambio Climático en México). Secretaría de Medio Ambiente y Recursos Naturales y Secretaría de Hacienda y Crédito Público, México D.F
- von Storch H, Hewitson B, Mearns L (2000) Review of empirical downscaling techniques. In: Iversen T, Høiskar BAK (eds) Regional climate development under global warming. General Technical Report 4. [http://regclim.met.no/rapport 4/presentation02/presentation02.htm](http://regclim.met.no/rapport%204/presentation02/presentation02.htm)
- Vrac M, Stein ML, Hayhoe K, Liang X-Z (2007) A general method for validating statistical downscaling methods under future climate change. *Geophys Res Lett* 34:L18701. doi:10.1029/2007GL030295
- Wallace JM, Gutzler DS (1981) Teleconnections in the geopotential height field during the Northern Hemisphere Winter. *Mon Weather Rev* 109:784–812

- Wilby RL, Harris I (2006) A framework for assessing uncertainties in climate change impacts: low-flow scenarios for the River Thames, UK. *Water Resour Res* 42:W02419. doi:[10.1029/2005WR004065](https://doi.org/10.1029/2005WR004065)
- Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. *Prog Phys Geogr* 21:530–548
- Wilby RL, Wigley TML (2000) Downscaling general circulation model output: a reappraisal of methods and limitations. In: Sivakumar MVK (ed) *Climate prediction and agriculture*. Proceedings of the START/WMO International Workshop, 27–29 September 1999, Geneva. International START Secretariat, Washington, DC, pp 39–68
- Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO (2004) Guidelines for use of climate scenarios development from statistical downscaling methods. Supporting material of the Intergovernmental Panel on Climate Change. Available from the DDC of IPCC TGCIA, p 27
- Yule GU (1926) Why do we sometimes get nonsense-correlations between time-series?—a study in sampling and the nature of time-series. *J Roy Stat Soc* 89:1–63
- Zermeño D (2008) Análisis probabilístico de escenarios escalados de precipitación y temperatura bajo cambio climático en México. Master thesis. Posgrado en Ciencias de la Tierra, UNAM