**Nonlinear Processes
in Geophysics**

Bias correction and post-processing under climate change

S. Vannitsem

Institut Royal Météorologique de Belgique, Avenue Circulaire, 3, 1180 Brussels, Belgium

Abstract. The statistical and dynamical properties of bias correction and linear post-processing are investigated when the system under interest is affected by model errors and is experiencing parameter modifications, mimicking the potential impact of climate change. The analysis is first performed for simple typical scalar systems, an Ornstein-Uhlenbeck process (O-U) and a limit point bifurcation. It reveals system's specific (linear or non-linear) dependences of biases and post-processing corrections as a function of parameter modifications. A more realistic system is then investigated, a low-order model of moist general circulation, incorporating several processes of high relevance in the climate dynamics (radiative effects, cloud feedbacks...), but still sufficiently simple to allow for an extensive exploration of its dynamics. In this context, bias or post-processing corrections also display complicate variations when the system experiences temperature climate changes up to a few degrees. This precludes a straightforward application of these corrections from one system's state to another (as usually adopted for climate projections), and increases further the uncertainty in evaluating the amplitudes of climate changes.

1 Introduction

Long term measurement environmental records indicate that climate is experiencing modifications on a wide variety of time and space scales. Under this changing environment, one can wonder whether how predictable is the system. In this context, two types of questions arise: (i) what are the modifications of the statistical and dynamical properties of the system under slow variations of its forcings, usually referred

to as system's sensitivity, and (ii) can we predict the slow modifications of the system.

If the model at hand representing the dynamics of reality is good enough and if we have a good knowledge about the uncertainty in the forcings and in the initial conditions, one could imagine to perform projections to get a reliable estimate of the possible outcomes. However models are always affected by *model errors*, increasing the uncertainty on the possible outcome (see e.g. Stainforth et al., 2007; Knutti et al., 2008), in particular in case of (non-linear) systems that could experience catastrophic changes (i.e. bifurcations). These errors may indeed considerably affect the specific structure of the system's attractor and of its bifurcation diagram. Their presence is amply demonstrated by the difficulty of atmospheric models in reproducing correctly the climate, as illustrated by the robust systematic errors that are still present in the climate of state-of-the-art models (e.g. Berner et al., 2008).

In order to partly correct the impact of these model errors, post-processing techniques are used. One of the most simple approach often used in the climate community is to remove the systematic errors, and to work with anomalies instead of absolute values (e.g. Solomon et al., 2007). More sophisticated approaches consist in correcting higher order moments, usually the second one, in order to reproduce the variability of the observation field as often used for the correction of weather forecasts (Johnson and Swinbank, 2009; Vannitsem, 2009; Glahn et al., 2009; Van Schaeybroeck and Vannitsem, 2011), or to apply downscaling schemes producing an additional variability at smaller scales (e.g. Benestad et al., 2008). But one of the main assumptions behind these different approaches are the stationarity of the correction parameters. The reduction of performances associated with non-stationarity of regression relationships has been discussed notably in Schmith (2008) in the context of downscaling of European winter temperatures and precipitation.

One can therefore wonder whether this type of post-processing can still be used in the context of a transient climate, in particular in the context of decadal forecasts. The obvious answer would be no in a strict sense since modifications of external parameters generically imply modifications of the variability of the system. But in view of the strong necessity and pressure in getting answers on the potential impact of climate changes on our society, these post-processing are used assuming that the impact of model errors is the same before and after the climate change, or is varying linearly in time (Buser et al., 2009). An interesting attempt in understanding the dependence of the bias on the temperature and precipitation amplitudes has been performed by Christensen et al. (2008) in the context of current climate simulations. This experiment, although not realistic since it does not compare biases before and after climate change, has enlightened the possible nonlinear character of the model biases as a function of increasing temperatures or precipitation amounts.

In the present work, the statistical and dynamical properties of bias correction and linear post-processing, in the presence of model errors, are investigated when the system under interest is experiencing parameter modifications, mimicking the potential impact of climate change. The sources of model errors are restricted here to parametric errors. Section 2 is devoted to the analysis of simple typical scalar systems, an Ornstein-Uhlenbeck process (O-U) and a limit point bifurcation (both paradigms of important processes acting in the atmosphere and climate). The analysis reveals (potentially) complicate dependences of biases and post-processing corrections under parameter evolution. The investigation of a more realistic system, the low-order model of moist general circulation developed by Lorenz (1984) briefly described in the Appendix, also supports this statement (Sect. 3). This system is incorporating several processes of high relevance in the climate dynamics (radiative effects, cloud feedbacks...), but is still sufficiently simple to allow for an extensive exploration of its dynamics. The conclusions are then drawn in Sect. 4.

2 Post-processing in scalar systems

2.1 Bias correction

Bias correction is certainly the simplest approach to post-process weather forecasts or climate runs. It simply consists in removing the mean of an ensemble of previous forecasts from current forecasts or of long climate runs. How this bias is affected by climate modifications is the central question of this section in the context of scalar systems.

2.1.1 Ornstein-Uhlenbeck process

Let us consider first the case where the real dynamics and its model are described by Ornstein-Uhlenbeck processes,

$$\frac{dx}{dt} = -\lambda x + K + \xi \tag{1}$$

$$\frac{dy}{dt} = -\lambda' y + K' + \xi' \tag{2}$$

where ξ and ξ' are white noise processes such that

$$< \xi(t) > \; = \; < \xi'(t) > = 0$$
$$< \xi(t)\xi(t') > \; = \; Q^2 \delta(t - t')$$
$$< \xi'(t)\xi'(t') > \; = \; Q'^2 \delta(t - t')$$

where $< . >$ refers to an ensemble average. The evolution of the probability densities of the solutions of these equations is described by a Fokker-Planck equation and the moments of the distribution can then be evaluated (e.g. Gardiner, 1985). Let us focus first on the properties of the first moments,

$$\frac{d < x >}{dt} = -\lambda < x > + K \tag{3}$$

$$\frac{d < y >}{dt} = -\lambda' < y > + K' \tag{4}$$

Assume furthermore that the parameter K is constant and identical to K'. The asymptotic stationary solutions are then staigthforwardly obtained, $x_s = K/\lambda$ and $y_s = K/\lambda'$, which gives a *climate bias* of

$$(x_s - y_s)_K = K \frac{\lambda - \lambda'}{\lambda \lambda'} \tag{5}$$

Let us now assume that the dynamical change is expressed as a jump in the value of K, $K'' = K + \delta K$. Under these circumstances, the bias is given by

$$x_s' - y_s' = (K + \delta K)\frac{\lambda - \lambda'}{\lambda \lambda'}$$
$$= (x_s - y_s)_K + \delta K \frac{\lambda - \lambda'}{\lambda \lambda'} \tag{6}$$

indicating that the bias depends linearly on the "climate" perturbation δK. This straightforward analysis already suggests that the bias is not a stationary quantity, as often assumed in the climate community when working with anomalies. Let us now consider that the model error is present in the parameter K, $\delta K = K' - K$ and the climate perturbation will affect the parameter λ. Assuming for simplicity that $\lambda = \lambda'$ and that the new parameter value after the change is given by $\lambda'' = \lambda + \delta \lambda$, one gets

$$(x_s - y_s)_\lambda = \frac{\delta K}{\lambda} \tag{7}$$

$$(x_s' - y_s')_\lambda = \frac{\delta K}{(\lambda + \delta \lambda)}$$
$$= (x_s - y_s)_\lambda - \frac{\delta K \delta \lambda}{\lambda} + \frac{\delta K \delta \lambda^2}{\lambda} + O(\delta \lambda^3) \tag{8}$$

indicating a nonlinear dependence of the bias on the climate perturbation, imposed through $\delta\lambda$.

Up to now, climate change was expressed in the form of an abrupt change of a parameter, but usually parameter changes are time dependent, such as the CO_2 increase supposed to affect the global temperature as a linear time dependent forcing (see Nicolis, 1988, for a discussion on this point). Therefore let us now consider that one has a climate change expressed through a progressive increase of the additive parameter in the form of *a ramp*, namely that $K = \gamma + \epsilon t$ and $K' = \gamma + \epsilon't$ and the model error is incorporated in the parameter value λ, $\lambda' = \lambda + \delta\lambda$. Note that for completeness the climate change scenario is also affected by an uncertainty incorporated through a modification of the parameter ϵ. In this case, the solutions of Eqs. (3)–(4) are given by

$$
\begin{aligned}
<x(t)> =\ & <x(0)> e^{-\lambda t} + \frac{\gamma}{\lambda}(1 - e^{-\lambda t}) \\
& + \frac{\epsilon}{\lambda}\left[\frac{e^{-\lambda t}}{\lambda} + (t - \frac{1}{\lambda})\right] \\
<y(t)> =\ & <y(0)> e^{-\lambda' t} + \frac{\gamma}{\lambda'}(1 - e^{-\lambda' t}) \\
& + \frac{\epsilon'}{\lambda'}\left[\frac{e^{-\lambda' t}}{\lambda'} + (t - \frac{1}{\lambda'})\right]
\end{aligned}
\tag{9}
$$

and the time dependent bias is the difference between the two equations. For times larger than $1/\lambda$ and $1/\lambda'$, one gets

$$
\begin{aligned}
<y(t)> - <x(t)> \approx\ & \gamma\frac{\lambda - \lambda'}{\lambda\lambda'} + (\frac{\epsilon'}{\lambda'} - \frac{\epsilon}{\lambda})t \\
& - (\frac{\epsilon'}{\lambda'^2} - \frac{\epsilon}{\lambda^2})
\end{aligned}
\tag{10}
$$

Clearly a (linear) time dependence of the bias is present that will depend on the amplitude of the model error $\delta\lambda$, the ramp velocity ϵ and the uncertainty of the ramp velocity $\delta\epsilon = \epsilon' - \epsilon$.

These results indicate that even in a linear system, a non trivial dynamics of the bias associated with the modelling error is present under climate change. This feature contrasts with the classical view that the bias associated with model deficiencies are similar under climate change.

For comparison let us also consider the case of a multiplicative ramp. The evolution equation for the first moment of the O-U process is

$$
\frac{d<x>}{dt} = -(\lambda + \epsilon t)<x> + K
\tag{11}
$$

The formal solution involves error functions that are not very transparent. One therefore resorts to an adiabatic formulation – considering that the parameter ϵ is small – in order to get the solution for long times (Nicolis and Nicolis, 2000). It consists in transforming the time as $\tau = \lambda + \epsilon t$ and seeking for a solution of the form

$$
<x> = <x>_0 + \epsilon <x>_1 + ...
\tag{12}
$$

Equating the terms corresponding to the same order in ϵ, one gets for the two first orders,

$$
<x>_0 = \frac{K}{\tau}
\tag{13}
$$

$$
<x>_1 = \frac{K}{\tau^3}
\tag{14}
$$

Note that $<x>_0$ is simply the stationary solution of Eq. (11), and is referred as the *adiabatic approximation*. Similar expansion can be performed for the model version of Eq. (11) in which the parameters are modified (becoming K' and ϵ'). The modification of the bias in time up to the first order in ϵ and ϵ' is now given by

$$
\begin{aligned}
<y>(t) - <x>(t) \approx\ & \frac{K'}{\lambda + \epsilon't} - \frac{K}{\lambda + \epsilon t} \\
& + (\frac{\epsilon'K'}{(\lambda + \epsilon't)^3} - \frac{\epsilon K}{(\lambda + \epsilon t)^3})
\end{aligned}
\tag{15}
$$

revealing a complicate non-linear variation in time, even in the case of this simple linear system.

2.1.2 A prototypical non-linear system

Assume now that the system under consideration displays a limit point bifurcation. This is one typical bifurcation, known to play an important role in many components of the climate systems and in particular in the Ocean dynamics (e.g. Dijkstra, 2005; Dijkstra and Ghil, 2005). In the absence of fluctuations, it obeys a normal form (see Nicolis, 1988),

$$
\frac{dx}{dt} = x(x - m) + K
\tag{16}
$$

whose steady state solutions are given by

$$
x_\pm = \frac{1}{2}[m \pm \sqrt{m^2 - 4K}]
\tag{17}
$$

Let us now assume that a model error is present in m, the model steady states are then given by

$$
y_\pm = \frac{1}{2}[m' \pm \sqrt{m'^2 - 4K'}]
\tag{18}
$$

Let us focus on the stable branch of solution and one assumes that parameters K and K' are progressively increasing in time, $K = \gamma + \epsilon t$ and $K' = \gamma + \epsilon't$. In order to clarify analytically the properties of the bias, one resorts to the adiabatic approximation – considering that the parameters ϵ and ϵ' are small – in order to get the solution for long times. Up to the first order in ϵ, one gets

$$
x_0 = \frac{1}{2}[m - \sqrt{m^2 - 4\tau}]
\tag{19}
$$

$$
x_1 = \frac{-1}{m^2 - 4\tau}
\tag{20}
$$

$$
y_0 = \frac{1}{2}[m' - \sqrt{m'^2 - 4\tau'}]
\tag{21}
$$

$$
y_1 = \frac{-1}{m'^2 - 4\tau'}
\tag{22}
$$

where $\tau = \gamma + \epsilon t$ and $\tau' = \gamma + \epsilon' t$ are the rescaled times. The bias is therefore approximately equal to

$$y - x \approx (\frac{1}{2}[m' - \sqrt{m'^2 - 4\tau'}] - \frac{1}{2}[m - \sqrt{m^2 - 4\tau}])$$
$$+ (\frac{-\epsilon'}{m'^2 - 4\tau'} - \frac{-\epsilon}{m^2 - 4\tau}) \qquad (23)$$

This relation is represented in Fig. 1 as a function of time, together with the direct numerical integration of the system defined by Eq. (16) and its model version for different values of ϵ and ϵ'. First it clearly indicates the nonlinear character of the bias under climate change expressed in the form of a ramp. Second, relation (23) is a good approximation of the variation of this bias. This relation displays a complicate structure far from the simple approximation made in the literature, namely constant bias or linear variation of the bias (Buser et al., 2009). Note that a similar analysis can be performed when parameter m is affected by the climate modification with similar conclusions. It also clearly suggests that the temporal variation of the bias highly depends on the velocity increase of the parameter affected by the climate change (as for the O-U process with multiplicative ramp presented in the previous section). This further complicates the evaluation of the bias and of the future climate trends.

2.2 Linear post-processing

Usually the mean of the process is not the only moment affected by the presence of model errors. One could therefore have a need to correct higher order moments too, and in particular the variance of the field of interest. *Under stationary conditions*, linear post-processing can allow for providing the correction of both the mean and the variance as discussed in Vannitsem (2009) for weather forecasts. These approaches can be extended to seasonal or decadal forecasts, but in the later case changes in system forcings (i.e. increase of anthropogenic gases) could affect the potential of correction.

Let us assume that before any climate change, one can develop a linear post-processing algorithm for a forecast, allowing for the correction of both the mean and the variance of the field of interest,

$$X'_C(t) = \alpha'(t) + \beta'(t) X_M(t) \qquad (24)$$

where t refers to the lead time of the forecast, X_M the model predictor, and the parameters $\alpha'(t)$ and $\beta'(t)$ are obtained by minimizing a cost function based on past forecasts (see Vannitsem, 2009). To get a correction of both the mean and the variance, the parameters should be equal to,

$$\alpha'(t) = <X_R>_{\text{ref}} - \beta'(t) <X_M>_{t,\text{ref}} \qquad (25)$$

$$\beta'(t) = \sqrt{\frac{\sigma^2_{R,\text{ref}}}{\sigma^2_{M,\text{ref}}(t)}} \qquad (26)$$

where $<X_R>_{\text{ref}}$ and $\sigma^2_{R,\text{ref}}$ are the mean and the variance of the reference system (in other word, the reality), and
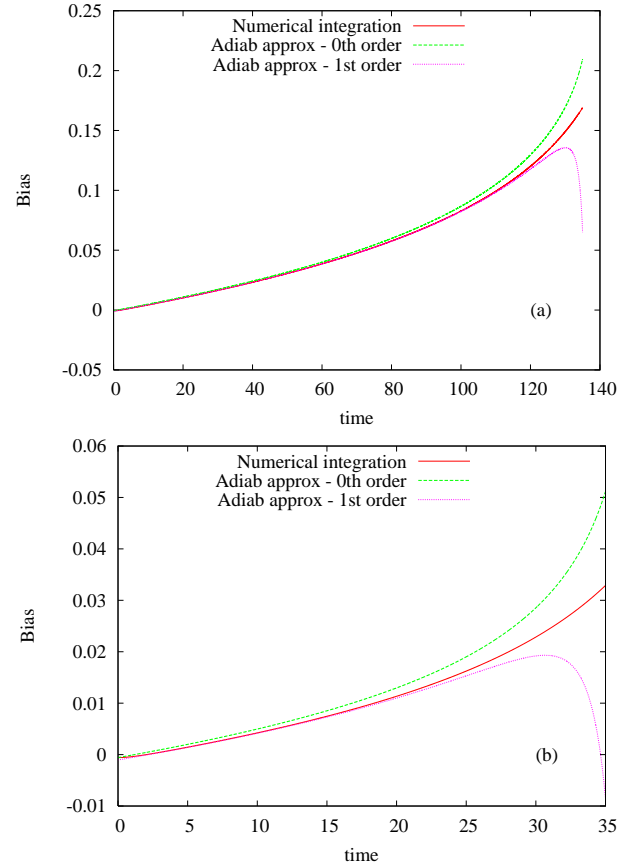

Fig. 1. (a) Temporal evolution of the bias obtained by the direct numerical integration of Eq. (16) with $m = 1$, $\epsilon = 0.001$ and $\gamma = 0.05$ and its model version with $m' = 1.01$, $\epsilon' = 0.0015$ and $\gamma = 0.05$ (continuous line). The 0th and first order approximations (Eq. 23) are given by the short-dashed and dotted curves. **(b)** as in **(a)** but with $\epsilon = 0.005$ and $\epsilon' = 0.0055$.

$<X_M>_{t,\text{ref}}$ and $\sigma^2_{M,\text{ref}}(t)$ are the corresponding quantity of the model starting from the initial conditions defined by the reality, before system's modification. This post-processing approach is referred as EVMOS (Error-in-Variable Model Output Statistics, see Vannitsem, 2009).

One can also define optimal parameters that would have been relevant during or after a system's modification. Let us assume that at a certain lead time t, one can write

$$X_C(t) = \alpha(t) + \beta(t) X_M(t) \qquad (27)$$

where

$$\alpha(t) = <X_R>_t - \beta(t) <X_M>_t \qquad (28)$$

$$\beta(t) = \sqrt{\frac{\sigma^2_R(t)}{\sigma^2_M(t)}} \qquad (29)$$

where now $<X_R>_t$ and $\sigma^2_R(t)$ are the mean and the variance of the reality during or after the system's change. The indice t reflects the time dependence of these statistical properties due to the transient effects affecting the real system.

Implicitely one assumes here that we have access to a set of realisations of the reality under this transient forcing for defining the optimal parameters. $< X_M >_t$ and $\sigma_M^2(t)$ are the corresponding quantities for the model starting from the initial conditions of the reality. The question now is to know to what extent the Eqs. (24)–(26) can replace the optimal scheme defined by Eqs. (27)–(29) in post-processing future forecasts.

In order to evaluate this impact, one can decompose the mean square error of the corrected forecasts as,

$$
\begin{aligned}
< (X_C' - X_R)^2 >_t = {} & < (X_C - X_R)^2 >_t + (\alpha'(t) - \alpha(t))^2 \\
& + (\beta'^2(t) - \beta^2(t))\sigma_M^2(t) \\
& - 2(\beta'(t) - \beta(t))C(X_R(t), X_M(t)) \quad (30)
\end{aligned}
$$

where $C(X_R(t), X_M(t))$ is the correlation between the actual reality and the model forecasts after system's change. The second term is clearly positive, increasing the error of the suboptimal scheme. The other terms can be either positive or negative depending on the properties of β' and β and therefore of the variability before and after the change.

To be more concrete, Let us now illustrate this evaluation to the O-U process, as described by Eqs. (2) and with the parameters $K = \gamma + \epsilon t$ and $K' = \gamma + \epsilon' t$. In this case, the variances of the reference system and of the future climate are the same (no modification due to the temporal variation of K and K', see Nicolis, 1988). This does simplify the problem and one gets for times $t > \max(1/\lambda, 1/\lambda')$,

$$
\begin{aligned}
\beta^2(t) &= \frac{\sigma_R^2}{\sigma_M^2(t)} = \frac{Q^2}{Q'^2}\frac{\lambda'}{\lambda} \\
&= \beta'^2(t) \\
\alpha(t) &= \frac{\gamma}{\lambda} + \frac{\epsilon}{\lambda}(t - \frac{1}{\lambda}) - \frac{Q}{Q'}\sqrt{\frac{\lambda'}{\lambda}}(\frac{\gamma}{\lambda'} + \frac{\epsilon'}{\lambda'}(t - \frac{1}{\lambda'})) \\
\alpha'(t) &= \frac{\gamma}{\lambda} - \frac{Q}{Q'}\sqrt{\frac{\lambda'}{\lambda}}\frac{\gamma}{\lambda'} \quad (31)
\end{aligned}
$$

and

$$
\begin{aligned}
< (X_C' - X_R)^2 >_t \approx {} & < (X_C - X_R)^2 >_t \\
& + (\frac{\epsilon}{\lambda}t - \frac{Q}{Q'}\sqrt{\frac{\lambda'}{\lambda}}\frac{\epsilon'}{\lambda'}t)^2 \quad (32)
\end{aligned}
$$

indicating that the quality of the correction as measured with this quadratic norm is degrading as a function of the square of the lead time. This also suggests that as it could be anticipated, the post-processing could be useful (provided it is useful before the change) when the system's modifications are kept small.

The important message of this section is that the quality of post-processing (bias correction or linear regression) is system's specific and strongly depends on the model error source and on the way parameters are affected by the change (together with the uncertainty of the scenario). This precludes the possibility to deduce universal evolution relations

for these corrections and therefore to make some general assumptions on the structure of the variations.

3 Analysis of a low order moist general circulation model

3.1 Predictability experiments and post-processing strategies

In the previous section, (potentially) complicate variations of the bias and regression parameters in the presence of model error and system's modification have been illustrated in the context of scalar systems. In particular, it has been shown that it depends critically (linearly or non-linearly) on the amplitude of the model error, the uncertainty of the climate change scenario and the velocity of this change. In the present section, post-processing is explored in the context of a low-order moist general circulation system (an atmosphere coupled to a slab ocean).

In order to simplify matters, we will consider only the case for which the system has reached its own attractor after climate change, or in other words, there is no transient evolution of the external parameters but just a small instantaneous jump leading to a modification of the attractor's properties. Although transient forcing variation certainly plays an important role in the actual dynamics, this will add a further complication, not of high relevance in the context of this idealized model. Furthermore one assumes for simplicity that the external climate forcing jump is the same for both the reality and the model. This will emphasize the role of model errors and not of the forcing uncertainty. Last point will be addressed in the future in the context of a more realistic climate model.

In the following, the experiments based on the model described in the Appendix will therefore be performed by fixing the value of a' (without an explicit time dependence). Once this value is fixed, long runs are performed in order to get asymptotic results. The runs presented below are typically performed for periods of the order of 50 000 days, after a transient period of convergence toward the respective attractors.

3.2 Results

3.2.1 Natural variability

Let us start briefly with the analysis of the solutions of this system. Figure 2 shows the averaged air temperature, dew point, total atmospheric heat forcing, relative humidity and albedo as a function of $T0$ (the spatially averaged solar forcing temperature, see the Appendix) for $T3 = 10\,\mathrm{K}$ (the amplitude of the latitudinal variation of the solar forcing temperature) and $D = 1.0$ (parameter related to the Ocean volume). The different curves correspond to different values of a'. A quite complex structure is found with a sensitivity to changes
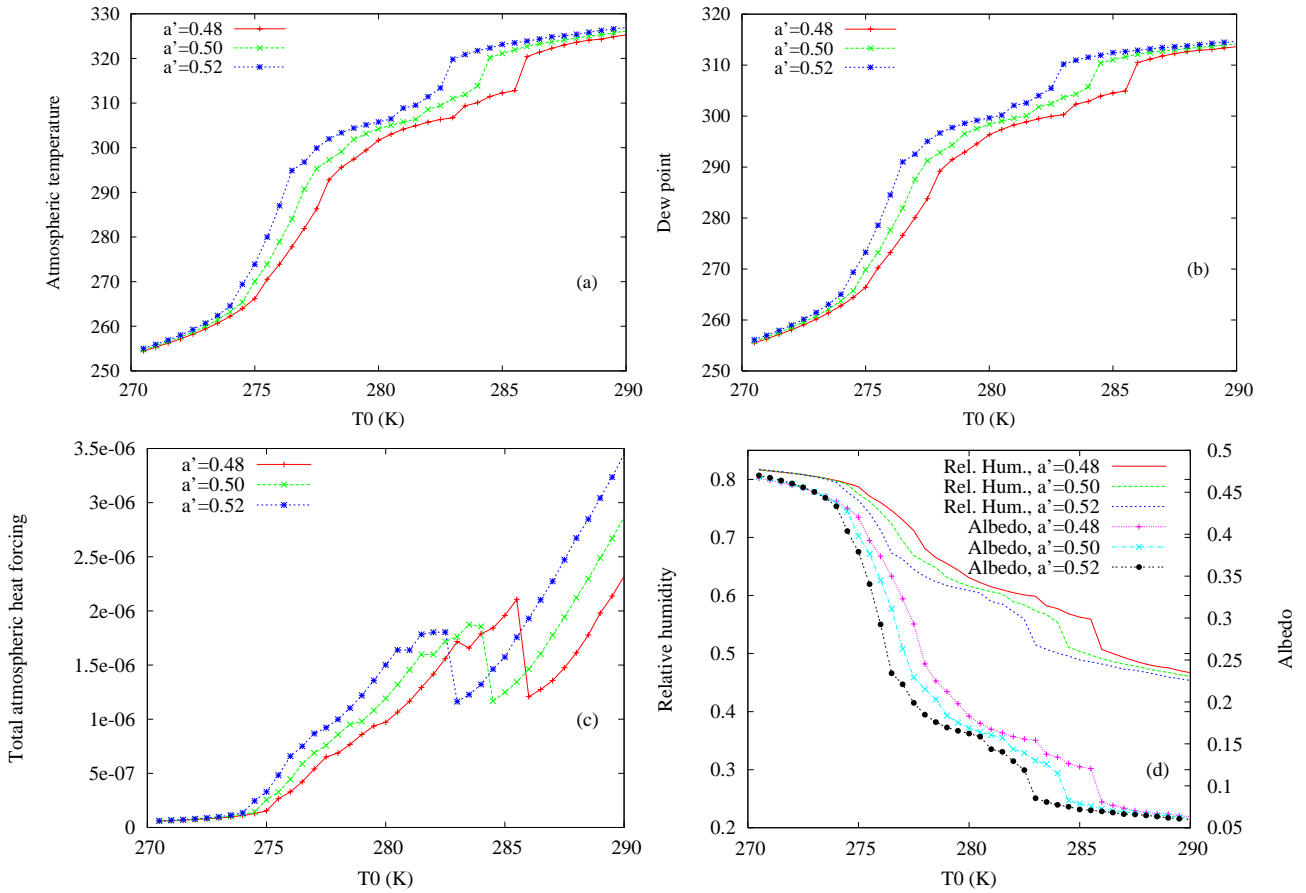
Fig. 2. Mean state as a function of $T0$ for different values of a' and with $D = 1.0$ and $T3 = 10$ K, for **(a)** the atmospheric temperature, **(b)** the dew point, **(c)** the heat forcing (Eq. A8), and **(d)** the relative humidity and the albedo.

of $T0$ which is highly variable depending on the range of $T0$ values considered. In particular for large values of $T0$, a large jump is found in the diagram reflecting a drastic change of dynamics, as indicated also by the drastic decrease of the total atmospheric heat forcing shown in Fig. 2c. It is also worth noting that by increasing the parameter a', a global shift of the curves toward smaller values of $T0$ is found. In other word, for a specific value of the external forcing $T0$, the increase of a' implies an increase of atmospheric temperature and dew point, and the albedo and relative humidity are decreasing. Note that this feature, specific to the present idealized model, could be different in other models for which cloud feedbacks and atmospheric absorption are more complicated.

Let us now look at the dynamical instability of this system (e.g. Kalnay, 2003). The Lyapunov exponents have been computed through a simple algorithm based on the (nonlinear) evolution of a set of small amplitude orthogonal perturbations. Different amplitudes have been tested without substantial differences.

Figure 3 displays the dominant Lyapunov exponent as a function of $T0$ for (a) $T3 = 5$ K and $D = 1$, (b) $T3 = 10$ K and $D = 1$, (c) $T3 = 15$ K and $D = 1$, (d) $T3 = 10$ K and

$D = 0.1$, (e) $T3 = 15$ K and $D = 0.1$. Different curves are plotted for different values of a'. In panel (a), Stable fixed points, (quasi-)periodic or chaotic solutions are found. In the other panels, there is no stable fixed points solutions for the range of values of $T0$ explored.

Let us now be more specific by investigating the structure of the bifurcation diagram for $T3 = 10$ K, $a' = 0.5$ and $D = 1$ (panel b) which corresponds to the solution presented in Fig. 2. For small values of $T0$ and for values ranging from 282 K to 284 K, the solution is (quasi-)periodic. For intermediate values (between 275.5 K and 281.5 K) and the largest values explored (284.5 K and 285 K), the solution is chaotic. For these large values of $T0$, mean temperature reaches very high values 315 to 320 K. These values are unrealistic for our midlatitude climate and in order to keep some resemblance with our climate atmosphere, we focus on smaller values of $T0$. For values between 276 and 280, the mean temperature field, dew point, relative humidity and albedo are closer to the values effectively observed for the Earth climate. In this range the dominant Lyapunov exponent is of the order of 0.15 to 0.20 day^{-1}, a value not far from estimates obtained in more complex models (e.g. Vannitsem and Nicolis, 1997). For $T3 = 15$ K, one can reach values up to 0.30 day^{-1}
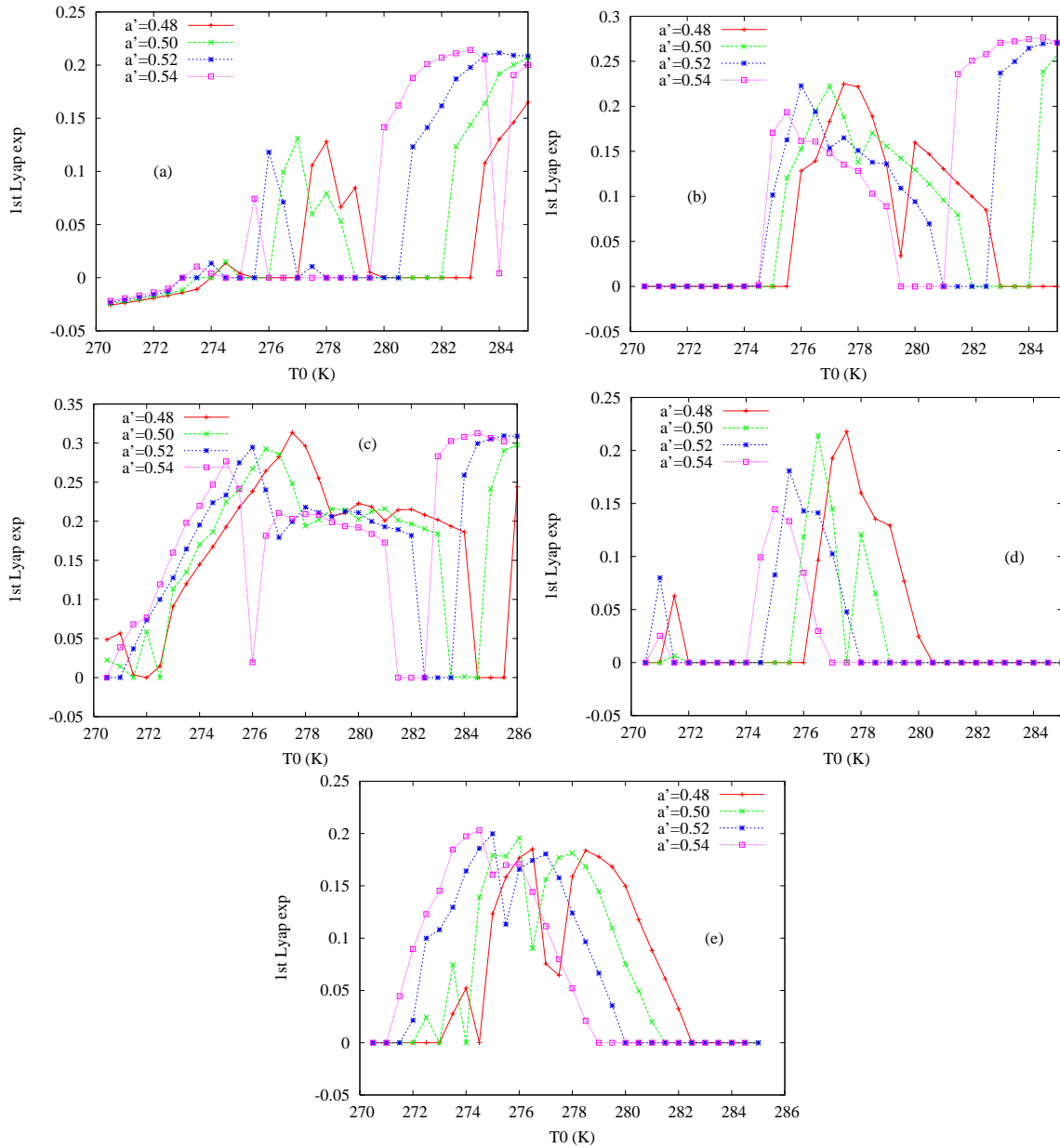
Fig. 3. Dominant Lyapunov exponent as a function of $T0$ for different values of a', for **(a)** $T3 = 5\,\mathrm{K}$ and $D = 1.0$, **(b)** $T3 = 10\,\mathrm{K}$ and $D = 1.0$, **(c)** $T3 = 15\,\mathrm{K}$ and $D = 1.0$, **(d)** $T3 = 10\,\mathrm{K}$ and $D = 0.1$ and **(e)** $T3 = 15\,\mathrm{K}$ and $D = 0.1$.

(see panel c). When the volume of the Ocean is increased ($D = 0.1$, panels d and e), a reduction of these maximum values is experienced. The effect of increasing a' is mainly to displace the whole Lyapunov diagram toward smaller values of $T0$. This suggests that depending on the actual value of $T0$, the system can experience either an increase or decrease of short term predictability as a function of the increase of a'.

3.2.2 Post-processing

The central question of the present work is to know whether post-processing techniques can be used, after a climate change. In order to investigate this aspect, we first introduce

a model error in the system affecting its climate. Model errors are present in the various parameterizations used in atmospheric or climate models, but in order to keep things simple, we just introduce an error in the estimate of the coupling coefficient k between the Ocean and the Atmosphere. The actual value of k is 0.015 and we fix for most of the experiments presented below a value for the model to $k = 0.014$, so just a few percent errors.

For the forecasting experiments, we also introduce a random initial condition error sampled from a uniform distribution between $[-5 \times 10^{-5}, 5 \times 10^{-5}]$. This experimental design assumes that the initial condition is better known that
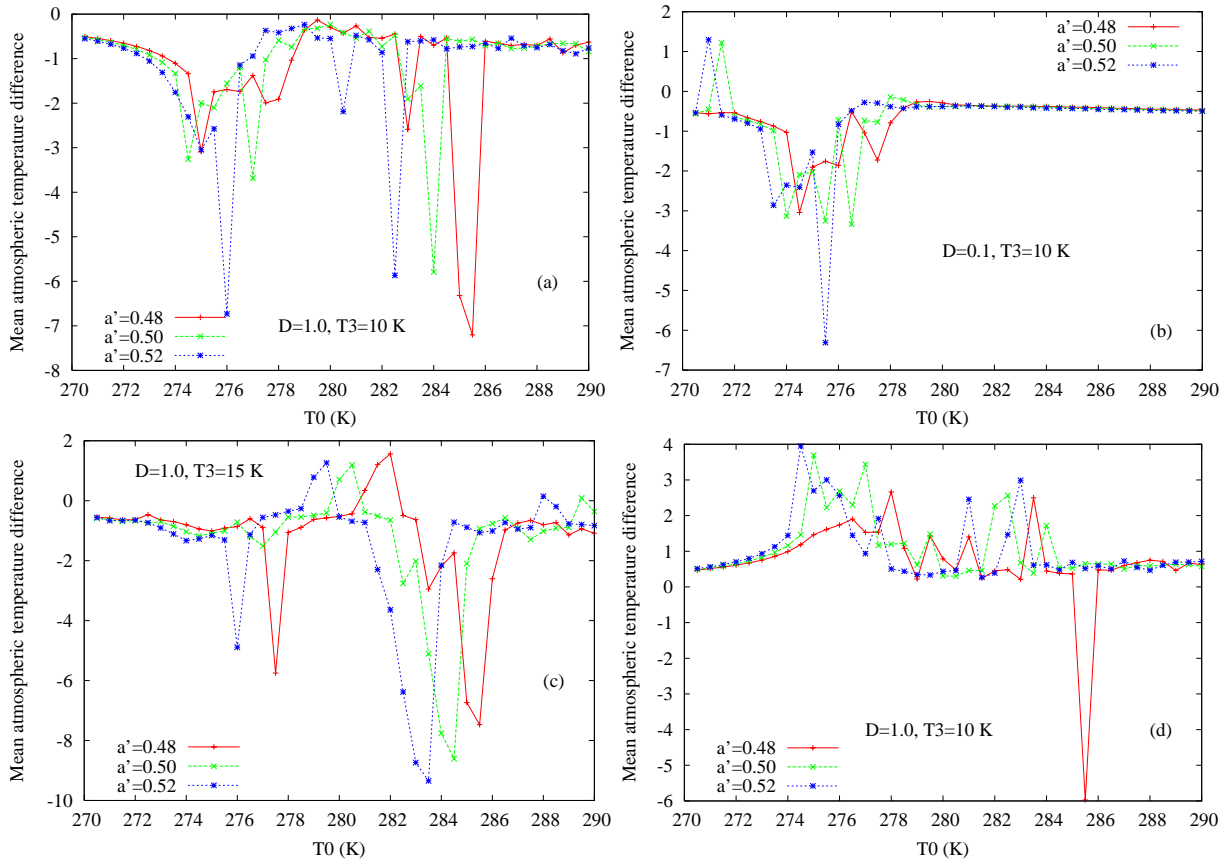
Fig. 4. **(a)** Temperature bias between the reference system with $k = 0.015$ and the model version with $k = 0.014$ for different values of a', and with $D = 1.0$ and $T3 = 10$ K; **(b)** as in **(a)** but with $D = 0.1$ and $T3 = 10$ K; **(c)** as in **(a)** but with $D = 1.0$ and $T3 = 15$ K; as in **(a)** but with a model version with $k = 0.016$, and with $D = 1.0$ and $T3 = 10$ K.

the model structure. This specific choice has an impact on the post-processing correction for short times because the larger is the initial condition error, the smaller is the potential correction of the method (see Vannitsem and Nicolis, 2008). However it does not qualitatively modify the conclusions drawn below.

The climate mean

Let us first consider the case where only the bias between reality and model is corrected as it is usually the case for climate projections. Figure 4 displays the difference between the climatological averages of the air temperature for different parameter settings and different values of a', reality-model. The difference is strongly dependent on the value of $T0$, from just a few tenth of degrees up to 7 degrees. Once this quantity is known, one can correct the forecast for the mean temperature. Let us now assume that the system is experiencing a climate change through an increase of the parameter a'. In this case the mean difference (the systematic bias of the model) can be very different as illustrated in Fig. 4. For instance if one focuses on the case where $T0 = 277$ K, $T3 = 10$ K and $D = 1.0$, the difference can pass

from about -1.5 degrees for $a' = 0.48$ to -3.5 for $a' = 0.50$ and back to -1 for $a' = 0.52$. This of course has strong influences when one wants to infer the absolute values of the change in mean temperature expected in the future. In the present example, one would overestimate the mean temperature after climate change by about 2 degrees (for a' going from 0.48 to 0.50), even if we know exactly how a' has changed. For other parameter values, the effect can be much milder but sometimes it can be very dramatic like when the system is close to a qualitative change of dynamics as for $T0 = 283.5$ K or 276 K (up to 6 degrees) in panel (a), or even more as in panel (c) and can sometimes even change sign (panels b and c). Note that there is almost no change of the biases when the dynamics becomes (quasi-)periodic (see Fig. 3). For larger value of the model parameter k (0.016 in panel d) a similar picture holds, but the bias is most of the time of the opposite sign to the ones corresponding to $k = 0.014$.

The main message is therefore that the systematic correction associated with the presence of model errors cannot be straithgforwardly transposed from one climate condition to another, as usually done in long term climate projections

(Solomon et al., 2007). In other words, the impact of model errors cannot be simply considered as a purely constant systematic additive term. This introduces a further (probably rather large) uncertainty on the possible future outcomes of the climate system, associated with the presence of model errors.

Finally, one can wonder whether developing an ensemble scheme based on a set of forecasts using different parameters (as discussed for instance in Stainforth et al., 2007) is worth doing in order to evaluate the range of possible outcomes. 100 runs are performed with different values of k ranging from [0.0145, 0.0155]. The reference value is chosen as one model coming from this ensemble ($k = 0.01511$). The mean temperature is computed for this set of models and for three different values of a'. The mean temperature biases are plotted as a function of k in Fig. 5. First the range of possible biases for this ensemble of model runs differs considerably from one value of a' to another, and there is no clear dependence of this quantity as a function of a'. This indicates that the actual uncertainty of the model biases cannot be straightforwardly transposed to another value of a' (or in other words for different absorption forcings). But what is interesting to mention is that there is an almost linear dependence of the bias as a function of k, suggesting that the best models are the ones having a value of k close to the actual one, whatever is the value of a'. This supports the usefulness of classifying the quality of the models under the current climate conditions as in Giorgi and Mearns (2003), as a confidence in their climate projections. But still one should be cautious because these results have been obtained for a small range of uncertainty of k and assuming that the model error is only associated with the error of a parameter. Other sources of model errors can play an even more important role, like processes not represented in the model (referred to as *model inadequacy* in Stainforth et al., 2007). These could change the conclusions but the analysis of this aspect is beyond the scope of the present work and is worth performing in a more complicate model.

Linear post-processing of forecasts

Long term projections on centennial or millennial time scales are one aspect of climate prospective. Another domain of very active research is to try to perform some climate forecasts on shorter time scales, from seasons to decadal time scales. These forecasts are also affected by model errors and furthermore are subject to slow changes of the climate forcings. Let us now focus on a more "sophisticated" (linear) post-processing technique, already discussed in Sect. 2, aiming at correcting forecasts. Again as in the previous section, one assumes that the climate change is simply represented by an increase of a' without any additional transient dynamics. The training of the post-processing technique is first performed on forecasts obtained with one value of a' and then applied and verified on a second set of forecasts obtained
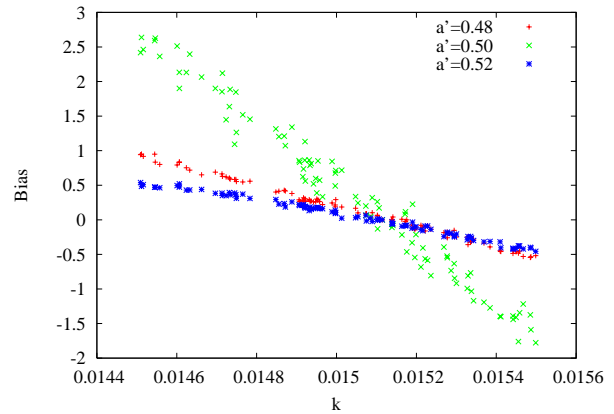

Fig. 5. Mean temperature bias for different values of k and a'. The reference value of k is 0.01511, for each a'. $D = 1.0$, $T3 = 10$ K and $T0 = 277$ K.

with another value of a'. The number of forecasts for the training and verification phases is fixed to 5000. The regression is performed for each lead time of the forecast and at each grid point of the system grid.

Figure 6 displays the time variation of the parameter $\beta(t)$ and $\alpha(t)$ for the air temperature and dew point for parameter values $D = 1.0$, $T0 = 277$ K and $T3 = 10$ K and for one specific location of the central row of the grid of this coupled system. Three different parameter values a' are displayed. The lines without symbols represent a running average of the parameters starting after about 15 days. Two time scales should be distinguished, before and after about 10 days. For the atmospheric temperature at short time scales, the parameters are close for $a' = 0.48$ and 0.50, while for $a' = 0.52$ it varies in an opposite direction. For dew point, the picture is slightly different where the variations of the parameters are closer to each other for short time scales. For longer time scales, the variability of the parameters is pronounced and the running mean values (lines without symbols) are substantially different for the different parameter values a'. This variability is associated with the rather "small" number of forecasts used for the training.

In order to clarify whether a correction can be obtained when applying the parameters defined for a specific a' to forecasts generated with another parameter a', we have computed the ratio of the Mean Square Error (MSE) after correction and the MSE of the raw forecasts (Fig. 7). If its value is below one, a correction is obtained. In each panel the actual value of a' is 0.52 and the results obtained with post-processing based on parameters $a' = 0.48$, 0.50 and 0.52, are displayed. For $D = 1.0$, the use of parameters corresponding to $a' = 0.50$ is still improving the atmospheric temperature and dew point for short times, but not with the smallest value of a'. For $D = 0.1$, even the use of parameters corresponding to $a' = 0.50$ does not improve the atmospheric temperature.
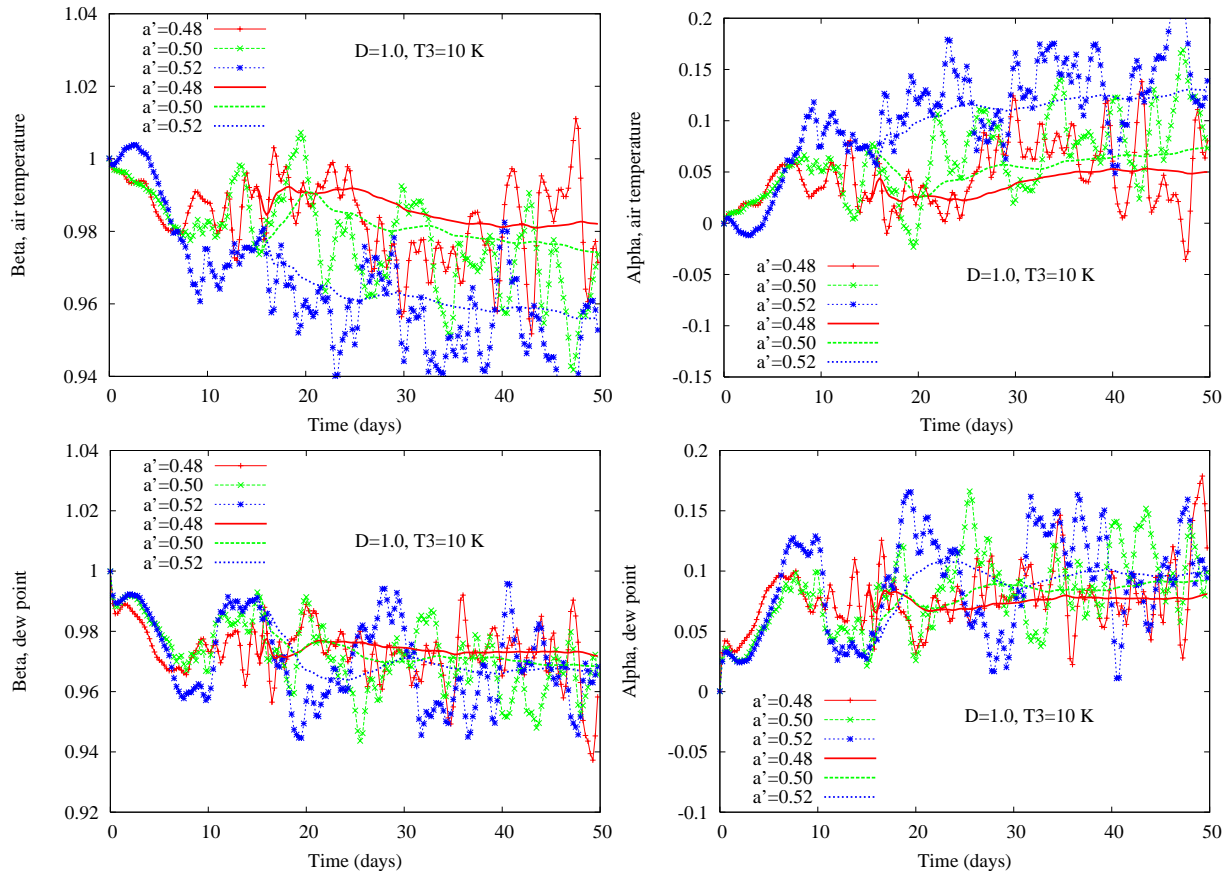
Fig. 6. Variations of the parameter α and β at a function of lead time for different values of a' at one specific location of the grid (one grid point in the middle row). Panels **(a)** and **(b)** for air temperature and **(c)** and **(d)** for dew point. $D = 1.0$ and $T3 = 10$ K. The continuous lines without symbols are running averages starting after 15 days in order to evaluate the average level values of the parameters at long lead times.

Figure 8 displays the variations of the mean square error ratio for post-processing trained with different values of a' and applied on a set of forecasts produced at $a' = 0.52$. The lead time is fixed to 5 days and two different values of D are considered, namely 1.0 and 0.1. This results indicate that the traditional post-processing approach can be useful but in some restricted situations, for which the dynamics does not vary much between the training and the verification phases. This corroborates the results of Sect. 2. Moreover, the variations of the quality of post-processing display a complicate nonlinear dependence, as a function of the difference between the values of a' used for verification and for training.

4   Conclusions

Statistical post-processing is a common approach to correct weather and climate forecasts, as well as climate projections. The main assumption behind these approaches is the stationarity of the correction parameters allowing for an application of the techniques for new runs or forecasts. Under climate change, this hypothesis is not realistic, as it has been amply

demonstrated in the present paper. In particular biases can display linear or non-linear variations depending on many aspects, the nature of the underlying dynamics, the source of model errors, the intensity (velocity of change) of the forcing scenario and its uncertainty. These features are system specific, increasing considerably the difficulty in assessing the quality of future forecasts or projections, and their potential corrections.

This analysis has been performed here in the context of very simple systems. A natural extension will be the investigation of an intermediate order climate model in the same spirit in the presence of realistic model errors. This analysis could provide a more quantitative information on the variations of biases and post-processing parameters, that could guide future estimates of the uncertainties of climate forecasts or projections and the development of correction schemes based on non-stationary coefficients.

The present work has also been limited to the most simple post-processing approaches for single model forecasts or runs. Recently there were a lot of efforts in developing and using multi-model ensembles to evaluate the uncertainty of climate projections. This question has been just slightly
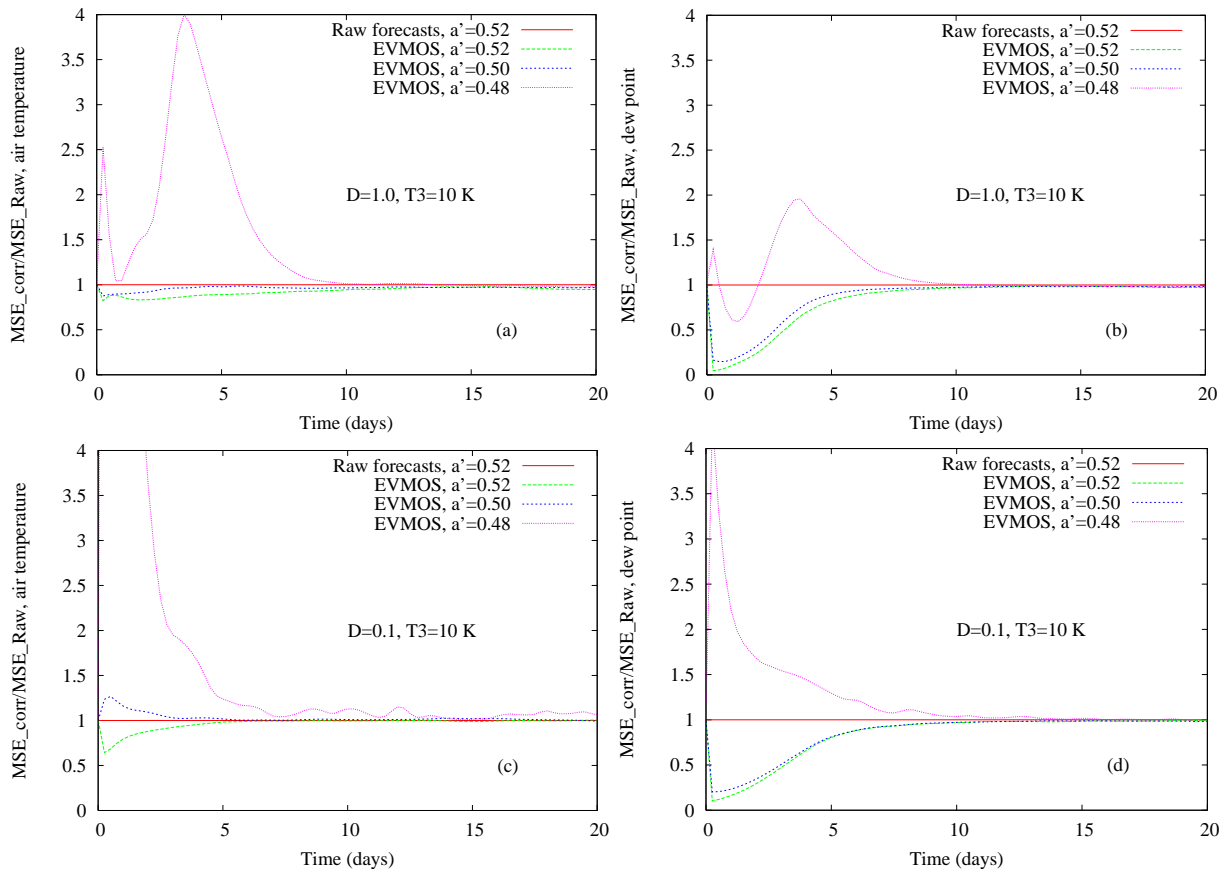
Fig. 7. Ratio of the Mean Square Error (MSE) of the corrected forecasts and the MSE of the raw forecasts for a value of $a' = 0.52$. The MSE is evaluated over the whole grid. The different curves correspond to the use of post-processing schemes trained using different values of a'. $D = 1.0$ and $T3 = 10$ K.
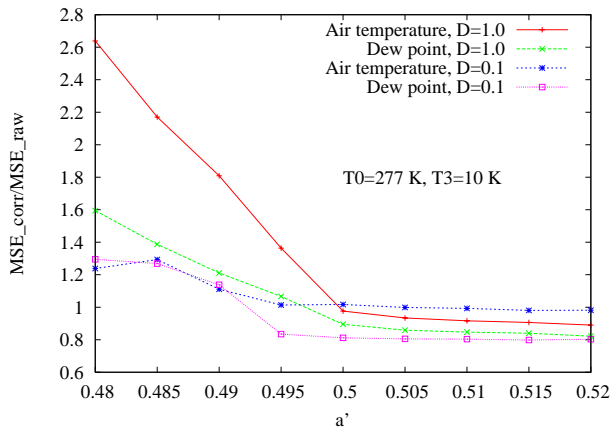

Fig. 8. Ratio of the Mean Square Error (MSE) of the corrected forecasts and the MSE of the raw forecasts for a value of $a' = 0.52$, as a function of the value of a' used for the training of the post-processing scheme. The MSE is evaluated over the whole grid and the lead time is fixed to 5 days.

touched upon by performing a few runs with different parameter values reflecting the model uncertainty. In this case the result suggests that the closest is the model parameter value from the true value, the smaller is the bias. In other words, the ranking of the quality of the models seems to be more robust than the amplitude of the bias. This could justify the usefulness of some statistical approaches (e.g. Giorgi and Mearns, 2003; Duan and Phillips, 2010 and reference therein) providing a larger weight to projections generated by the models matching best the current observations. Weigel et al. (2010) however pointed out the difficulty in finding correct weights for each model of the ensemble, that could lead to worse results than equal weighting. This clearly points out the necessecity for a thorough analysis of these approaches in simple dynamical systems, in a similar vein as in the present work.

Finally, it has been clearly demonstrated that working with anomalies (as usually presented in international reports) masks some deficiencies of the system and introduces overconfidence which is not appropriate especially when mitigation strategies of the impact of climate changes are envisaged.

Appendix A

The low-order moist general circulation model

The model developed by Lorenz (1984) is a coupled Ocean-Atmosphere model, whose atmosphere is composed of a mixture of dry air and water into two possible phases, vapor or liquid. The Ocean is just a heat bath at which the atmosphere is coupled and with which it exchanges water and heat through simple parameterizations. The domain is periodic in the west-east direction and bounded on the North and South by frictionless vertical walls. The atmosphere is described as a baroclinic dynamics defined on a β plane. The external driving force is the incoming solar radiation. The basic prognostic equations are: the vorticity equation,

$$\frac{\partial \nabla^2 \psi}{\partial t} + J(\psi, \nabla^2 \psi) + \beta \frac{\partial \psi}{\partial x} = -f \nabla^2 \xi + F \tag{A1}$$

where ψ, ξ and F are the streamfunction, the velocity potential and the curl of the viscous drag, respectively. The operator J is the Jacobian, f and β are the coriolis parameter defined for the midlatitudes and its first derivative, respectively; The thermodynamic equation

$$\frac{d(c_p T + Lv)}{dt} = \frac{RT\omega}{p} + H \tag{A2}$$

where T, v, ω and H are the air temperature, the water vapor mixing ratio, the vertical velocity in pressure coordinates and the atmospheric diabatic heating per unit mass, respectively, and L and c_p are the latent heat of condensation and the specific heat of air at constant pressure; the water content equation,

$$\frac{dw}{dt} = G \tag{A3}$$

where w and G are the total water mixing ratio and the gain or loss of water by evaporation or precipitation, respectively; and the oceanic thermodynamic equation,

$$\frac{dcS}{dt} = E \tag{A4}$$

where S and E are the oceanic temperature and the oceanic heating per unit mass, respectively. The coefficient c is the specific heat of water. These prognostic equations are complemented by the diagnostic continuity and thermal wind equations. Note that the total water mixing ratio is itself transformed into a dew point temperature through the diagnostic relation,

$$w = c' W^\mu / p \tag{A5}$$

where μ is a constant fixed to $L/R_w T_s \approx 20$, derived through a simplification of the Clausius-Clapeyron equation. Equation (A3) is replaced by an equation for the dew point temperature.

This system of equations is then simplified through a two-layer vertical discretization for which the streamfunction is defined at two levels and temperature (air or dew point) at an intermediate level, and in the horizontal direction by focusing only on the largest modes. This approximation reads

$$X = \sum_{n=0}^{6} X_n \Phi_n$$

where

$$\Phi_0 = 1$$

$$\Phi_{3j-2} = 2\sin(jy)\cos(2x)$$

$$\Phi_{3j-1} = 2\sin(jy)\sin(2x)$$

$$\Phi_{3j} = \sqrt{2}\cos(jy)$$

for $j = 1, 2$. All the variables are adimensionalized, with the charactristic scales fixed to $f^{-1} = 3$ h, $L = 1830$ km, $p_0 = 1000$ mb, and the unit of temperature is equal to 100 K.

Lorenz (1984) also has proposed a pseudo-specral method to evaluate the sources and sinks terms and the nonlinear terms appearing in the equations. The grid developed by Lorenz is composed of 9 grid points. All the parameters are fixed as in Lorenz (1984), except the lapse rate fixed to 0.215 as in Nese et al. (1996). The integration time step is fixed to 0.02 time units.

The vertically averaged forcings are given by the relations

$$\overline{\overline{F}} = -k\nabla^2 \psi \tag{A6}$$

$$\overline{F} = -k'\nabla^2 T \tag{A7}$$

$$\overline{\overline{H}} = -k\{c_p(T-S) + L(v-s)\} + \overline{\overline{H_R}} \tag{A8}$$

$$\overline{\overline{G}} = -k(v-s) - (lv/\lambda)(w-v) \tag{A9}$$

$$\overline{\overline{E}} = -\frac{p_0}{p_1 - p_0}k\{c_p(S-T) + L(s-v)\} + \overline{\overline{E_R}} \tag{A10}$$

where s is the saturation mixing ratio at temperature S and mean sea level pressure p_0. p_1 is the pressure at bottom of oceanic layer. k is the coupling coefficient between the ocean and the atmosphere. Another damping term (A7) with coefficient k' is reducing the shear between the two levels. $1/l$ is a damping time scale corresponding to the half life of clouds (water delivered as rain). A new constant containing information on the volume of the oceanic layer is defined as $D = p_0 c_p / ((p_1 - p_0)c)$. The radiative terms are defined as

$$\overline{\overline{H_{\mathrm{R}}}} = (\sigma g/p_0)v'a''(S^4 - T''^4 - T'^4) \tag{A11}$$

$$\overline{\overline{E_{\mathrm{R}}}} = (\sigma g/(p_1 - p_0))(-S^4 + v'a''T''^4 - (1 - v'a)T*^4) \tag{A12}$$

where

$$v'a'' = v'(a + (1 - a)a') \tag{A13}$$

accounts for the fraction of heat effectively absorbed by the atmosphere where $1 - a'$ is the fraction of longwave radiation that will go out the atmosphere without absorption. Note that in the original discussion of Lorenz (1984), a' was considered as the part transmitted toward space, but it cannot coincide with the effective fraction of heat absorbed by the atmosphere, Eq. (A13), except in the case of $a' = 0.5$ which is precisely the value chosen by E. Lorenz. σ is the Stefan constant, T' and T'' are temperatures at which the cloud-free portion of the atmosphere is radiated upward and downward, respectively, and $v' = v/(v + v_{\mathrm{s}})$ where v_{s} is a reference water vapor mixing ratio.

Finally the solar forcing entering in the term E_{R} depends on 2 Fourier coefficients,

$$T* = T0\Phi_0 + T3\Phi_3$$

In this idealized coupled ocean-atmosphere system, one can wonder how one can simulate a climate change associated with the variations or atmospheric absorption. The answer can be simply by modifying the net absorption of longwave radiations coming from the ocean by the atmosphere. So we will simply focus on variations of the parameter a', appearing in Eq. (A13). This parameter will be modified in such a way that the mean air temperature difference reaches a few degrees (temperature differences expected in a climate change).

Edited by: W. Hsieh
Reviewed by: two anonymous referees

References

Benestad, R. E., Hanssen-Bauer, I., and Chen, D.: Empirical-statistical downscaling, World Scientific, Singapore, 215 pp., 2008.

Berner, J., Doblas-Reyes, F. J., Palmer, T. N., Schutts, G., and Weisheimer, A.: Impact of a quasi-stochastic cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model, Phil. Trans. R. Soc. A, 366, 2561–2579, 2008.

Buser, Ch. M., Künsch, H. R., Lüthi, D., Wild, M., and Schär, C.: Bayesian multi-model projection of climate: bias assumptions and interannual variability, Clim. Dynam., 33, 849–868, 2009.

Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, Ph.: On the need for bias correction of regional climate change of temperature and precipitation, Geophys. Res. Lett., 35, L20709, doi:10.1029/2008GL035694, 2008.

Dijkstra, H.: Nonlinear Physical Oceanography, Springer Dordrecht, New York, 532 pp., 2005.

Dijkstra, H. and Ghil, M.: Low-frequency variability of the large scale ocean circulation: A dynamical system approach, Rev. Geophys., 43, RG3002, doi:10.1029/2002RG000122, 2005.

Duan, Q. and Phillips, Th. J.: Bayesian estimation of local signal and noise in multimodel simulations of climate change, J. Geophys. Res., 115, D18123, doi:10.1029/2009JD013654, 2010.

Gardiner, C. W.: Handbook of stochastic methods, Springer-Verlag, New-York, 442 pp., 1985.

Giorgi, F. and Mearns, L. O.: Probability of regional climate change based on the reliability ensemble averaging (REA) method, Geophys. Res. Lett., 30, 1629–1632, 2003.

Glahn, B., Peroutka, M., Wiedenfeld, J., Wagner, J., Zylstra, G., Schuknecht, B., and Jackson, B.: MOS uncertainty estimates in an ensemble framework, Mon. Weather Rev., 137, 246–268, 2009.

Johnson, Ch. and Swinbank, R.: Medium-range multimodel ensemble combination and calibration, Q. J. R. Meteorol. Soc., 135, 777–794, 2009.

Kalnay, E.: Atmospheric modeling, data assimilation, and predictability, Cambridge University Press, 341 pp., 2003.

Knutti, R., Allen, M. R., Friedlingstein, P., Gregory, J. M., Hegerl, G. C., Meehl, G. A., Meinshausen, M., Murphy, J. M., Plattner, G.-K., Raper, S. C. B., Stocker, T. F., Stott, P. A., Teng, H., and Wigley, T. M. L.: A review of uncertainties in global temperature projections over the twenty-first century, J. Climate, 21, 2651–2663, 2008.

Lorenz, E. N.: Formulation of a low-order model of a moist general circulation, J. Atmos. Sci., 41, 1933–1945, 1984.

Nese, J. M., Miller, A. J., and Dutton, J. A.: The nature of predictability enhancement in a low order Ocean-Atmosphere model, J. Climate, 9, 2167–2172, 1996.

Nicolis, C.: Transient climatic response to increasing CO_2 concentration: some dynamical scenarios, Tellus, 40A, 50–60, 1988.

Nicolis, C. and Nicolis, G.: Passage through a barrier with a slowly increasing control parameter, Phys. Rev. E, 62, 197–203, 2000.

Schmith, T.: Stationarity of regression relationships: Application to empirical downscaling, J. Climate, 21, 4529–4537, 2008.

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

Stainforth, D. A., Allen, M. R., Tredger, E. R., and Smith, L. A.: Confidence, uncertainty and decision-support relevance in climate predictions, Phil. Trans., R. Soc., 365, 2145–2161, 2007.

Vannitsem, S.: A unified linear model output statistics scheme for both deterministic and ensemble forecasts, Q. J. R. Meteorol. Soc., 135, 1801–1815, 2009.

Vannitsem, S. and Nicolis, C.: Lyapunov vectors and error growth

patterns in a T21L3 quasigeostrophic model, J. Atmos. Sci., 54, 347–361, 1997.

Vannitsem, S. and Nicolis, C.: Dynamical properties of model output statistics forecasts, Mon. Weather Rev., 136, 405–419, 2008.

Van Schaeybroeck, B. and Vannitsem, S.: Post-processing through linear regression, Nonlin. Processes Geophys., 18, 147–160, doi:10.5194/npg-18-147-2011, 2011.

Weigel, A., Knutti, R., Liniger, M. A., and Appenzeller, Ch.: Risks of model weighting in multimodel climate projections, J. Climate, 23, 4175–4191, 2010.