Geoscientific
Model Development

High dimensional decision dilemmas in climate models

A. Bracco[1], J. D. Neelin[2], H. Luo[1], J. C. McWilliams[2], and J. E. Meyerson[2]

[1]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA
[2]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA, USA

Correspondence to: A. Bracco (abracco@gatech.edu)

Abstract. An important source of uncertainty in climate models is linked to the calibration of model parameters. Interest in systematic and automated parameter optimization procedures stems from the desire to improve the model climatology and to quantify the average sensitivity associated with potential changes in the climate system. Building upon on the smoothness of the response of an atmospheric circulation model (AGCM) to changes of four adjustable parameters, Neelin et al. (2010) used a quadratic metamodel to objectively calibrate the AGCM. The metamodel accurately estimates global spatial averages of common fields of climatic interest, from precipitation, to low and high level winds, from temperature at various levels to sea level pressure and geopotential height, while providing a computationally cheap strategy to explore the influence of parameter settings. Here, guided by the metamodel, the ambiguities or dilemmas related to the decision making process in relation to model sensitivity and optimization are examined. Simulations of current climate are subject to considerable regional-scale biases. Those biases may vary substantially depending on the climate variable considered, and/or on the performance metric adopted. Common dilemmas are associated with model revisions yielding improvement in one field or regional pattern or season, but degradation in another, or improvement in the model climatology but degradation in the interannual variability representation. Challenges are posed to the modeler by the high dimensionality of the model output fields and by the large number of adjustable parameters. The use of the metamodel in the optimization strategy helps visualize trade-offs at a regional level, e.g., how mismatches between sensitivity and error spatial fields yield regional errors under minimization of global objective functions.

1 Introduction

General circulation models (GCMs) are an invaluable tool to understand and predict climate variability and change. Climate simulations, however, involve complex interactions among many processes, including turbulent mixing in both ocean and atmosphere, cloud physics, small-scale moist convection, and aerosol dynamics, among others. These processes are too small scale or too complex to be explicitly resolved, and are commonly replaced by parameterizations of various sophistication. The modeling of those processes and their interactions remains imperfect, and GCM predictions have large uncertainties that depend on the climate variable of interest. For example, over tropical regions the representation of mean precipitation or wind patterns is prone to biases comparable in magnitude to the observed signal in certain areas (Covey et al., 2003; Dai, 2006; Meel et al., 2007; Kidston and Gerber, 2010; Stephens et al., 2010; Swart and Fyfe, 2012; van Oldenborgh et al., 2012; Barimalala et al., 2012).

Model biases depend on both parameterizations and on the choice of parameters used, and there is no general agreement on the set of key model parameters that are uncertain yet critical in GCM performances. It is also not clear if all relevant uncertainties are linked to parameterization and parameter choices, or else if GCMs are prone to structural instabilities that cannot be expressed only as parameter variations (McWilliams, 2007). Systematic investigations of parameter space in climate models have been performed for single models using the so called "perturbed physics" strategy (Murphy et al., 2004; Knight et al., 2007; Rougier et al., 2009; Collins et al., 2010; Rowlands et al., 2012). The outcome of those simulations has to then be compared with observations to narrow the range of acceptable parameters.

Alternatively to the perturbed physics approach, for a given GCM it is common practice to settle on an optimized set of parameters – optimized according to the modeler needs – and retest and *tune* multiple aspects of the model every time a given parameter value or parameterization scheme are modified. This exercise is carried out primarily by trial and error, and in most, if not all, cases a model revision in the GCM parameter setting yields improvement in one field or geographic region, but degradation in another. In Neelin et al. (2010) we proposed an algorithm that would partially automate this process in a computationally efficient manner, helping to condense information for the modeler. This approach stems from the engineering and theoretical optimization literature and uses a multi-objective optimization technique, while relaying on standard packages available for constrained optimization problems. It is based on approximating the model's parameter dependence to a low-order polynomial by using a limited number of model integrations, and it assumes that the error metric varies smoothly whenever parameters are changed. The smoothness assumption relies on the GCM experiments shown here, and on other unpublished explorations performed using the Community Atmospheric Model (CAM) version 4 (Neal et al., 2010), and it is supported by the non-hydrostatic regional simulations presented in Bellprat et al. (2012). It is not an a priori fundamental property of GCMs, and it needs to be verified for each variable and parameter of interest. While Hairer and Majda (2010) present a theoretical argument that could justify the framework of linear response theory in climate science, simple idealized models have been found to provide non-smooth responses (Hakkarainen et al., 2012). A significant advantage of the procedure that we outlined in Neelin et al. (2010) is that optimization can be performed repeatedly for as many objective functions as the user desires at low computational cost.

Here we further explore basic issues in model parameter dependence using an atmospheric circulation model (AGCM) to assess the potential for a wide implementation of this procedure, while focusing on the ambiguities related to the parameter decision process. When trying to optimize a model setup, it is necessary to define an objective or cost function that measures the distance between selected metrics of the model output for a given set of parameters, and those of the observed climate. We focus on exploring optimization differences for varying metrics of interest, for example using square- or root-mean-square (RMS) error of key climate variables versus using regressions on a widely used El Niño Southern Oscillation index. We also investigate regional spatial patterns of signal, sensitivity, and error, as the choice of the objective function is expected to be user dependent and possibly focused on a given spatial region.

2 Model and experimental design

The atmospheric model used in this study is the International Centre for Theoretical Physics (ICTP) AGCM (Molteni, 2003). It is based on a hydrostatic spectral dynamical core (Held and Suarez, 1994), and adopts the vorticity-divergence form described in Bourke (1974). The ICTP AGCM includes, in the parameterized processes, short- and long-wave radiation, large-scale condensation, convection, surface fluxes of momentum, heat and moisture, and lateral and vertical diffusion. Convection is represented by a mass-flux scheme that is activated where conditional instability is present, and boundary layer fluxes are obtained by stability-dependent bulk formulae. Land and ice temperature anomalies are determined by a simple one-layer thermodynamic model. In this study, the AGCM is configured with eight vertical (sigma) levels and with a spectral truncation at total wavenumber 30. Applications of the ICTP AGCM can be found in e.g., Bracco et al. (2004) and Kucharski et al. (2006, 2007, 2009).

As in Neelin et al. (2010), we analyze a suite of AGCM integrations forced by observed sea surface temperatures (SST) from the HadISST data set (Rayner et al., 2003). Four parameters are varied, specifically the subgrid scale wind gustiness (WGust), a horizontal viscosity parameter corresponding to a damping time (Damp), the cloud albedo parameter (Cl_{Alb}), and the relative humidity parameter from the deep convective parameterization (RH_{Conv}). To each of those we assign an admissible range based on the properties of the parameterization and the model numerics, and we choose nine values centered at the standard setting recommended and validated by ICTP (http://users.ictp.it/~kucharsk/speedy8_clim_v41.html). We verified that most of those parameters, in the selected ranges, do not influence substantially the globally averaged top-of-atmosphere net energy flux (variations are less than 10 %). A noticeable exception is Cl_{Alb}, which causes unrealistic variations of the energy flux from -10 to $26\,W\,m^{-2}$. Ensembles of ten members, each member starting from slightly different initial conditions, are carried out over the period 1977–2002 for each parameter value considered. The last 25 yr are used in the following analysis and the first year is discarded as spin-up.

3 Metamodel formulation and accuracy

In Neelin et al. (2010) we explored the parameter dependence of the ICTP AGCM in a suite of global measures and found it generally smooth. Such characteristic has been confirmed by Bellprat et al. (2012) for a state-of-the-art regional climate model (RCM) when varying a suite of five parameters, and by Archibald et al. (2012) for CAM in relation to changes to the critical relative humidity thresholds for the formation of low level or middle to high level clouds. The smoothness leads us to adopt a low-order polynomial strategy of

fitting parameter dependencies. More specifically, we constructed a computationally cheap metamodel to estimate the root-mean-square error of any model quantity for a given combination of the input parameters, assuming that changes due to a perturbation of one parameter are approximated by a second-order polynomial regression. If two parameters are varied at once, a nonlinear term for each parameter pair is introduced in the metamodel computation, and perturbations of more than two parameters are approximated by the sum of the nonlinear terms of all possible pairs.

Following Neelin et al. (2010), the quadratic metamodel fit to a climate field $\phi(\boldsymbol{x},t)$ that depends on space and time can be expressed as

$$\phi_{mm} = \phi_{std} + \sum_{i=1}^{N} a_i \mu_i + \sum_{i=1}^{N} \sum_{j=1}^{N} b_{i,j} \mu_i \mu_j, \tag{1}$$

where $\mu_i = \mu_{i_{pert}} - \mu_{i_{std}}$ is the parameter i taken relative to its standard value, N the number of parameters considered, and $\phi(\boldsymbol{x},t)$ is any statistic from the model output chosen by the modeler. Here $a_i(\boldsymbol{x},t)$ is a high-dimensional vector containing the linear coefficients for each parameter at each grid point in time, and $b_{i,j}(\boldsymbol{x},t)$ represents the quadratic (diagonal) and interaction (off-diagonal) terms, with the assumption $b_{i,j}(\boldsymbol{x},t) = b_{j,i}(\boldsymbol{x},t)$. Thus a fit procedure of order N allows to estimate the linear sensitivity and the quadratic nonlinearity, while the off-diagonal coefficients, obtained with a number of simulations of order N^2, can be calculated from the corners of pairwise planes.

Figure 1 exemplifies the procedure by showing the RMS error of the ensemble mean climatologies compared to the National Centers for Environmental Predictions (NCEP) reanalysis (Kalnay et al., 1996) (panels a–c) or the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP, Xie and Arkin, 1997) (panel d) for slices along the four parameter directions for different climate variables. The variables vary from zonal wind at 200 hPa (WGust), to meridional wind at 925 hPa (Cl$_{Alb}$), geopotential height anomalies at 500 hPa (Damp), and rainfall (RH$_{Conv}$). We consider climatologies in boreal summer (June–August, JJA), winter (December–February, DJF), and annual averages. In Fig. 1, together with the RMS errors evaluated from the model output, we present the error reconstructions by the quadratic metamodel (solid lines) and its linear counterpart (dashed lines) using only the endpoints and the standard case, and an estimation of the ensemble spread of the RMS error. As already mentioned, in all cases the parameter dependence is smooth in large-scale measures, and the quadratic fit reproduces it extremely well when the ten-member ensemble mean is used, independently of the parameter or climate variable selected. If only one 25 yr long ensemble member is considered, on the other hand, it is possible to achieve an estimate of sensitivity but the variation in ϕ is not very smooth for many parameters and variables.
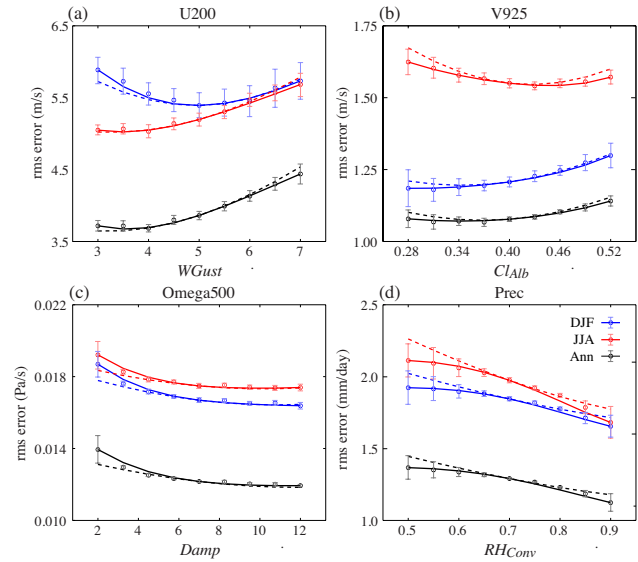

Fig. 1. RMS error of the ensemble mean AGCM climatology of (a) zonal wind at 200 hPa for varying WGust, (b) meridional wind at 925 hPa for varying Cl$_{Alb}$, (c) vertical velocity at 500 hPa for varying Damp, and (d) precipitation for varying RH$_{Conv}$ in December–February (DJF) in blue, in June–August (JJA) in red, and for annual averages in black relative to the National Centers for Environmental Prediction reanalysis for wind fields and vertical velocity, and to CPC Merged Analysis of Precipitation for precipitation. The AGCM values are compared to the quadratic metamodel reconstruction based on the endpoints for each parameter, and to its linear counterpart. Note that the linear metamodel gives quadratic terms (with positive curvature) in the RMS error. Units on abscissa are $\mathrm{m\,s^{-1}}$ for WGust, days for Damp and nondimensional for the remaining parameters. Error bars provide an estimate of the spread of the RMS error in the ten ensemble members (± 1 standard error of the ensemble).

Often, the linear metamodel provides a reasonable fit over the entire parameter range, with the quadratic term contributing only a small correction. In few cases, however, the quadratic term is essential to capture the negative curvature of the dependence, as for the RMS error of precipitation on RH$_{Conv}$. In Fig. 1 it is also evident that the parameter dependence varies with seasons, and the parameter optima for summer, winter or annual averages may not coincide, as in the case of the wind fields analyzed. This is a known, common problem, especially in coupled general circulation models, and identifies a first dilemma associated with parameter optimization and sensitivity: its temporal dependence. For global quantities, the RMS errors in the annual averages tend to be smaller than the corresponding seasonal ones, again suggesting that the model error may change sign within the year, and cancelations are common whenever all months are considered.

We have shown so far that the metamodel is helpful in reconstructing the globally averaged model error for varying parameters. In doing so, we synthesized the parameter dependence in one number. In the spatial structure, however, discrepancies between the AGCM output and the metamodel reconstruction could be large but homogeneously distributed around zero, and cancel each other when averaged globally. Before proceeding in our optimization exercise and in the analysis of regional parameter dependencies, we need to quantify those discrepancies. By construction, the model and metamodel error patterns coincide at the standard μ_i values and at the two extremes (μ_{i_Min} and μ_{i_Max}). The four ensembles at intermediate μ values, on the other hand, allows us to compare the model output with the reconstructed one. As an example, the maps of the DJF modeled precipitation error for $RH_\text{Conv} = 0.8$, of the reconstructed error from the metamodel, and of their difference, are shown in Fig. 2. In the difference, both positive and negative contributions are indeed present, but overall their amplitude is small (at most 10 % of the AGCM error). We verified for a large number of variables of climate interest that the metamodel provides indeed a very good approximation of the globally averaged RMS error and of its pattern. Videos displaying the evolution across the parameter space of the modeled precipitation error, of the metamodel error, and their difference for all μ_i can be found in the Supplement. Obviously, the metamodel allows for evaluating the error patterns for continuous variations of μ_i, and not just for a subset of discrete values. Videos of the metamodel representation of precipitation error changes (evaluated against CMAP) for all four parameters continuously varied in their acceptable range are also available in the Supplement.

4 Global-scale optimization ambiguities

Having established the ability of the metamodel in identifying the optima parameter space over the feasible range, we summarize such space in Table 1 for a number of variables of common interest to climate scientists, from precipitation, low level winds, land surface temperature and mean sea level pressure, to vertical velocity at 500 hPa, and temperature and winds at the top of the atmosphere, for boreal winter and summer. The parameter optima are easily obtained from the metamodel (here including the quadratic terms) using a standard Matlab function for constrained optimization (fmincon), which is part of the Optimization Toolbox. The location of the optima in parameter space differs significantly for different climate variables and in the two seasons, highlighting the contradictions faced by the modeler whenever optimizing one variable versus another, or summer versus winter climatologies. The metamodel, on the other hand, proves to be a practical and computationally cheap tool to estimate the model sensitivity and the trade-offs between objective functions.
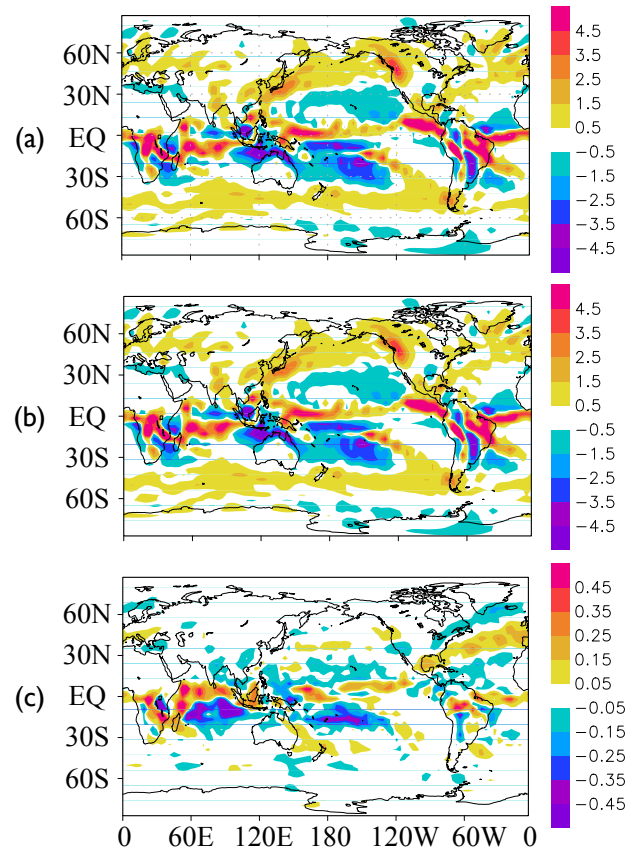
Fig. 2. (a) Spatial distribution the AGCM precipitation error relative to CPC Merged Analysis of Precipitation for $RH_\text{Conv} = 0.8$ and all other parameters kept to standard values in December–February. **(b)** Same as above but reconstructed using the quadratic metamodel. **(c)** Difference between AGCM and metamodel reconstructed error. Note the different color scale in **(c)** compared to **(a)** and **(b)**. Unit: mm day^{-1}.

Once the optima have been identified, it is important to evaluate them together with trade-offs associated with parameter changes around the optima. In other words, it is important to quantify the error introduced adopting a parameter setting different from the optimized one, for any given variable. This information can be easily obtained with the metamodel and is provided in Tables 2 and 3, where we summarize the trade-offs for all other climate variables at the optimum parameter setting of each of them, one at a time, for boreal winter and summer, respectively. The examination of the steepness of the fits provides also hints to this end, with flat curves being indicative of small dependence, and vice versa. Two-dimensional plots as the one shown in Fig. 3 for DJF precipitation help in visualizing the relative dependence of the parameters, two at a time. In the figure, together with the evolution of the RMS error of the modeled climatology in the parameter space, we also indicate the areas in which the increase in RMS error from its minimum at the optima is

Table 1. Optimized parameter setting for nine commonly analyzed climate variables in boreal winter, December–January, and summer, June–August.

	WGust DJF	Cl_{Alb} DJF	RH_{Conv} DJF	Damp DJF	WGust JJA	Cl_{Alb} JJA	RH_{Conv} JJA	Damp JJA
Precip	4.42	0.43	0.90	10.29	5.00	0.44	0.90	7.89
U 925 hPa	3.00	0.28	0.88	6.98	4.85	0.37	0.90	4.90
V 925 hPa	5.40	0.28	0.90	5.70	7.00	0.41	0.90	7.04
MSLP	3.00	0.28	0.72	3.91	4.43	0.52	0.77	9.79
LSTA	3.00	0.52	0.9	2.00	4.19	0.52	0.90	2.85
Ω 500 hpa	4.97	0.4	0.9	7.01	5.00	0.42	0.90	7.00
U 200 hPa	4.53	0.51	0.75	7.02	3.00	0.49	0.89	2.86
V 200 hPa	5.74	0.31	0.82	8.32	6.50	0.45	0.90	8.31
T 200 hPa	3.00	0.52	0.77	2.00	3.00	0.52	0.87	2.00

less than 1, 5, and 10 %. The examination of Fig. 3 reveals that the choice of Damp plays a lesser role in the representation of precipitation (and, as it appears, of most variables), while the optimum in the RH_{Conv} direction occurs at or near the upper boundary of the feasible range, with rapid degradation away from this, suggesting the need for careful scrutiny of the convective parameterization scheme.

It is important to stress that the approximate smoothness found for the RMS error of the model climatologies holds, even if not as precisely, for all other quantities we experimented with. The modeler may indeed be more interested in optimizing the representation of specific climate modes of variability and their teleconnections, instead of the model climatology, and the metamodel still proves useful. Here we consider, as example, the difference between observed and modeled El Niño Southern Oscillation (ENSO) regressed patterns, constructed using the Niño 3.4 index (Trenberth and Stepaniak, 2001). In Fig. 4 we show results for precipitation varying all four parameters, for all seasons. The RMS error in the regression patterns is far smaller than in the case of the climatological error, indicating that the ICTP AGCM correctly simulates the response to changes in SST forcing in the equatorial Pacific, and it is almost insensitive to the parameter choices in winter and fall, where the SST anomalies associated with ENSO are the strongest. The spread of the RMS error within the ensemble considered, on the other hand, is large, due to the internal variability in the atmospheric response to the (few) ENSO events recorded between 1977 and 2002. The error grows from boreal winter and fall, to summer and then spring, following inversely the strength of the SST signal, and in summer and spring a negative curvature characterizes most fits, pointing to a nonlinear dependence in those seasons. The seasonally dependent evolution and negative curvatures persist for all other variables, from surface winds to geopotential height, if their level is 500 hPa or closer to the surface. The model representation of the atmospheric circulation at higher levels regressed onto the ENSO index, on the other hand, does not display any seasonal dependence, and the ENSO regressions of 200 hPa fields are almost unaffected by the parameter changes.
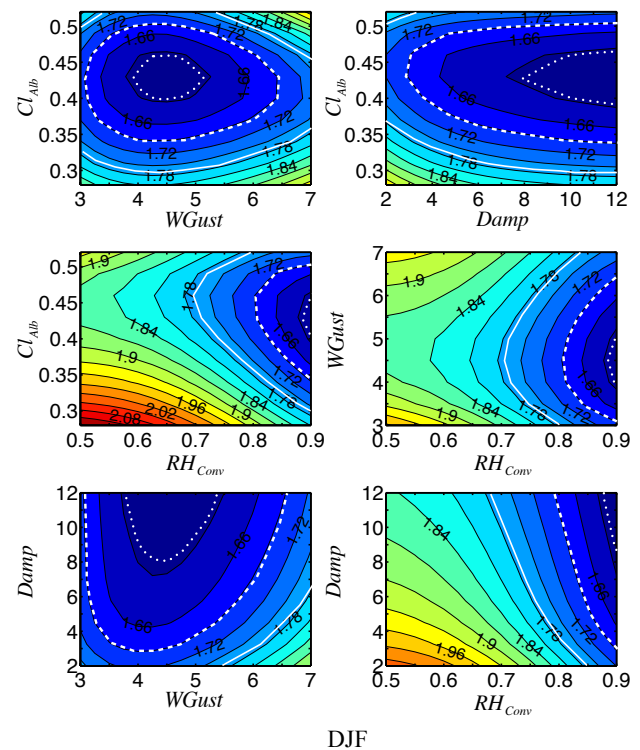
Fig. 3. Pairwise planes of the four-dimensional parameter space showing the root-mean-square globally averaged precipitation error in mm day^{-1} estimated by the quadratic metamodel in December–February. The white dotted (dashed; solid) contours indicate regions where the model error increases by 1 % (5 %; 10 %) compared to its minimum value at the parameter optima. Units on axes are m s^{-1} for WGust, days for Damp and nondimensional for Cl_{Alb} and RH_{Conv}.

A different picture emerges if the North Atlantic Oscillation (NAO) index is used instead of the Niño 3.4. The error in the ICTP AGCM representation of the atmospheric variability associated with changes in the NAO index is quite large, and so is the internal model variability (Bracco et al., 2004). The spread within the ten-member ensembles considered

Table 2. Trade-offs associated with optimizing one variable at a time according to the setting in Table 1 for nine climatic variables in boreal winter, December–February. The last column provides the sum of the trade-offs given the optimization of one given variable. Trade-offs are quantified as global root-mean-square error at the given parameter setting divided by the error at the variable optima.

	Precip	U 925 hPa	V 925 hPa	MSLP	LSTA	Ω 500 hPa	U 200 hPa	V 200 hPa	T 200 hPa	Total
Precip	1	1.08	1.05	1.13	1.17	1.01	1.04	1.09	1.16	9.73
U 925 hPa	1.18	1	1.03	1.02	1.35	1.07	1.22	1.10	1.34	10.31
V 925 hPa	1.17	1.07	1	1.12	1.38	1.09	1.10	1.01	1.40	10.34
MSLP	1.32	1.03	1.04	1	1.34	1.15	1.23	1.11	1.32	10.54
LSTA	1.23	1.10	1.14	1.06	1	1.19	1.09	1.26	1.01	10.08
Ω 500 hPa	1.02	1.08	1.03	1.13	1.22	1	1.04	1.05	1.22	9.79
U 200 hPa	1.10	1.09	1.10	1.11	1.06	1.07	1	1.15	1.07	9.75
V 200 hPa	1.17	1.11	1.01	1.17	1.36	1.08	1.09	1	1.17	10.34
T 200 hPa	1.26	1.11	1.13	1.05	1.01	1.19	1.08	1.24	1	10.07

Table 3. As in Table 2 but for boreal summer, June–August.

	Precip	U 925 hPa	V 925 hPa	MSLP	LSTA	Ω 500 hPa	U 200 hPa	V 200 hPa	T 200 hPa	Total
Precip	1	1.02	1.03	1.01	1.20	1.01	1.10	1.02	1.19	9.58
U 925 hPa	1.08	1	1.05	1.04	1.37	1.04	1.11	1.06	1.27	10.02
V 925 hPa	1.07	1.05	1	1.03	1.27	1.03	1.20	1.01	1.32	9.98
MSLP	1.13	1.08	1.09	1	1.02	1.10	1.14	1.04	1.09	9.69
LSTA	1.16	1.08	1.09	1.03	1	1.14	1.08	1.03	1.06	9.65
Ω 500 hPa	1.01	1.01	1.03	1.02	1.24	1	1.09	1.02	1.21	9.63
U 200 hPa	1.21	1.07	1.12	1.05	1.08	1.14	1	1.06	1.03	9.76
V 200 hPa	1.05	1.06	1.01	1.02	1.18	1.03	1.20	1	1.25	9.80
T 200 hPa	1.29	1.09	1.15	1.06	1.01	1.20	1.00	1.06	1	9.86

here is too large for the metamodel to be well determined (not shown), due to the large internal variability relative to the parameter sensitivity.

Once the metric of interest has been defined and the metamodel has been fitted to the chosen points (here the extremes in the available parameter ranges), the AGCM output of any parameter combination can be estimated, and the associated temporal and regional changes can be visualized. As an example, in Fig. 5 we present the spatial pattern of the linear (a_i) and quadratic (b_{ii}) contributions of the metamodel (Eq. 1), constructed using the climatological RMS error as metric for boreal winter and summer, in the Cl_{Alb} direction for meridional wind at 925 hPa, and in the RH_{Conv} direction for precipitation. All fits are from the parameter endpoints only, i.e for for $\mu_{i_{Max}}$ relative to the standard value ($Cl_{Alb} = 0.52$, and $RH_{Conv} = 0.90$). The sum of the two fields quantifies the change at $\mu_{i_{Max}}$ from the standard case. For the wind fields, the linear contribution is about three times larger than the nonlinear one, is stronger in summer than in winter, and off-phase between the two seasons considered. Both linear and nonlinear terms show an alternation of positive and negative areas, so that the global RMS error does not change significantly between $Cl_{Alb} = 0.52$ and the standard value (see Fig. 1), but the overall modeled wind field does. In the case of rainfall, the linear sensitivity and the quadratic nonlinearity

tend to reinforce each other in most regions, as for example in the Indian basin and Pacific Warm Pool in JJA, or over South America in DJF, and are both sufficiently strong to modify the overall patterns, especially in the tropics, where convective precipitation dominates. The model response to an increase in RH_{Conv} corresponds to stronger dependence of the convection on the humidity content above the boundary layer. This yields sensitivity to factors such as ventilation by dry air from adjacent regions tending to reduce convection on certain margins of the convection zones. When the metamodel is constructed using ENSO regression patters as metric, the linear and quadratic contributions for all near surface variables are of comparable intensity. In Fig. 6 we show both contributions to the near surface zonal wind for WGust = 7.0 compared to the standard. An increase of WGust has the overall impact of strengthening the near surface wind anomalies associated with ENSO in the central equatorial Pacific, while relative changes at higher latitudes are small (less than 7 % of climatological values). It is also evident that in the same season linear and quadratic contributions in several regions have opposite signs and tend to cancel each other. The analysis of those spatial patterns help in visualizing where changes are occurring at the regional scale, and the geographical importance of nonlinearity as function of the desired metric.
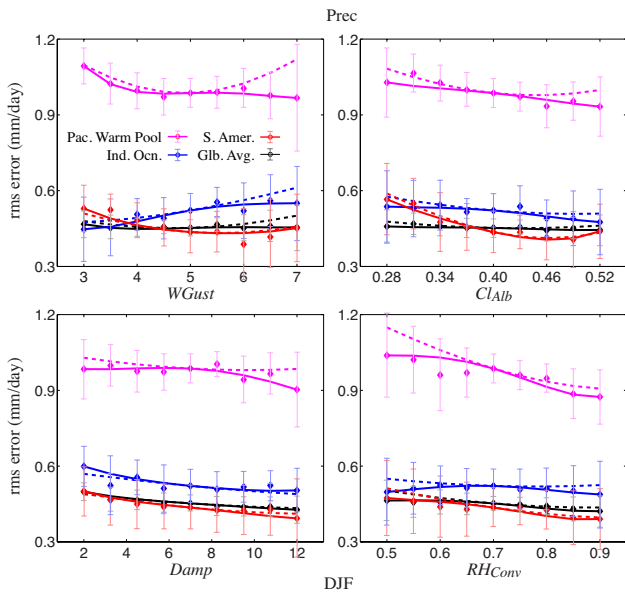
Fig. 4. Root-mean-square precipitation error of the modeled regression coefficients for the El Niño 3.4 index in December–February in blue, March–May in magenta, June–August in red, and September–November in black, relative to the CPC Merged Analysis of Precipitation data set. The AGCM values are compared to the quadratic metamodel reconstruction based on the endpoints for each parameter, and to its linear counterpart. Units on abscissa are $\mathrm{m\,s^{-1}}$ for WGust, days for Damp and nondimensional for Cl_{Alb} and RH_{Conv}. Error bars quantify the spread of the RMS error in the ten ensemble members (± 1 standard error of the ensemble).

5   Regional ambiguities

As already remarked, some of the more pressing ambiguities in selecting climate model parameters are linked to considerable pattern biases in different regions. When tuning global general circulation models there are multiple potential weightings that depend on assessments at regional scales. It is not uncommon, for example, to verify separately the model representation of the Asian monsoon, or of precipitation patterns over land, or of the variability associated with the North Atlantic Oscillation over North America and Europe. Those biases may vary substantially between climate variables, and/or on the performance metrics, and a model revision that improves one field or geographical area, may cause degradation in another. This is exemplified in Fig. 7, where we show the JJA RMS error in the precipitation climatology zooming over the Indian Ocean and western Pacific at RH_{Conv} standard value (top), at its maximum ($RH_{Conv} = 0.90$) – where the global error is at its minimum – (middle), and the absolute value of the error difference (bottom). By increasing RH_{Conv}, the effect of shallow convection on the moisture flux is increased. This causes greater convective activity where convection is limited by the humidity above the boundary layer, as for example over land, but it also carries

away moisture from the boundary layer so that, in some locations, the threshold amount of humidity is not fulfilled anymore. As a result, the model representation of the Indian summer monsoon and of precipitation over parts of Vietnam and China improves substantially, and the RMS error is reduced to half of its standard counterpart. On the other hand, the model rainfall signal over the Indian Ocean south of the Equator, part of Indonesia and Papua New Guinea degrades, approximately by the same amount.

The metamodel allows for visualizing those trade-offs very effectively, at little computational cost. The evolution of the absolute value of the error difference between the standard case and any other value can be constructed for each parameter and any values within a realistic range simply using the ensembles at $\mu_{i_{Std}}$, $\mu_{i_{Max}}$ and $\mu_{i_{Min}}$. Notwithstanding the approximation error associated with the quadratic metamodel, usually limited to less than 10 % of the amplitude of the reconstructed error field, and not significant in terms of overall pattern identification, the analysis of the regional evolution of the model biases for varying parameter values offers a powerful tool to select parameter combinations.

The quadratic metamodel can also be used to construct objective functions for targeted areas. Given the interest of the climate community for certain key regions, where model biases are large, we further explore the regional ambiguities associated with model tuning focusing on the Indian Ocean basin and Indian subcontinent (40–110° E, 30° S–22° N), South America (75–36° W, 30° S–10° N), and the Pacific Warm Pool (120–170° E, 20° S–20° N). For each region, we construct the metamodel as before, using only information from the ensemble mean at the standard setup, and the integrations at $\mu_{i_{Max}}$ and $\mu_{i_{Min}}$. In Fig. 8 we portray the regional parameter dependence of the RMS error of the precipitation climatology in boreal winter. The quadratic fit provides an accurate approximation of the parameter dependence in all regions and for all four parameters. The model error is greater over South America, independently of the parameter investigated, identifying a deficiency of the model in representing convective precipitation over land. Over the Indian Ocean, improvements in the WGust and Cl_{alb} space are achieved in the opposite direction than in the other regions or in the global average.

A different picture emerges if the model representation of ENSO teleconnections is alternatively used as metric (Fig. 9). The quadratic metamodels in this case provide valuable information on the general tendencies, but they do not always capture the functional form of the dependency. It should be noted, however, that the absolute value of the difference between the metamodel fit and the actual model response is small in all cases. This is verified for most variables, not only for precipitation. A good example is provided by the Pacific's Warm Pool region dependency on RH_{Conv}. The Pacific Warm Pool undergoes the strongest parameter dependence in association to ENSO, followed by the Indian Ocean, and then by South America.

V925

Prec


DJF   JJA
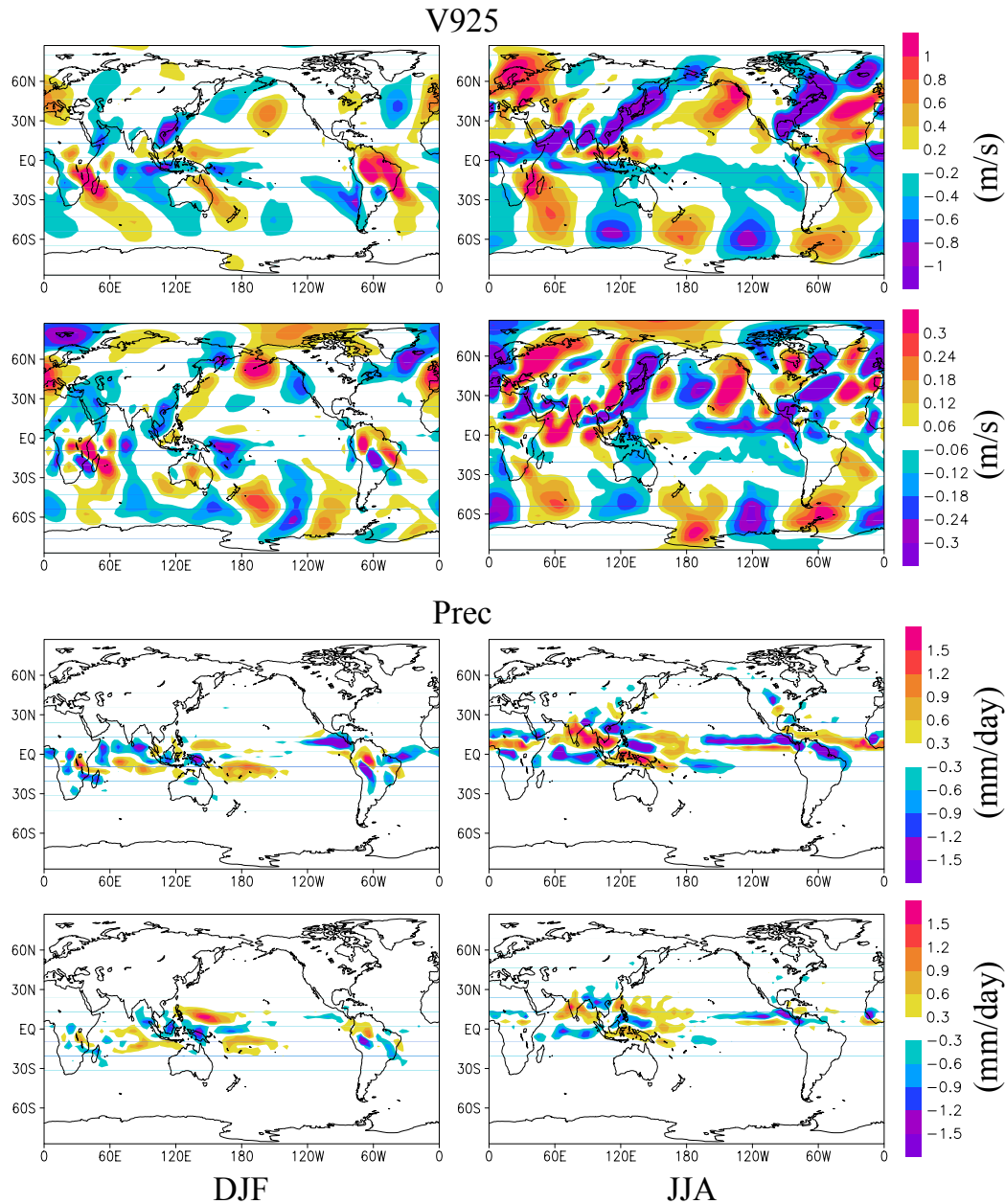
Fig. 5. Ensemble-mean measures of the sensitivity of the modeled error for meridional wind at 925 hPa in Cl_{Alb} (top two panels) and precipitation in RH_{Conv} (bottom two panels), in boreal winter (December–February) and summer (June–August). The linear contributions a_i are shown in the first and third pair of panels multiplied by $\mu_{i_{Max}}$ to give units of $m\,s^{-1}$ and $mm\,day^{-1}$ respectively, for the values that would occur at the positive end of the feasible range of Cl_{Alb} and RH_{Conv}. The quadratic contributions b_{ii}, similarly given as $b_{ii}\mu_{i_{Max}}^2$ in $m\,s^{-1}$ and $mm\,day^{-1}$, are shown in the second and bottom pairs of panels. The two contributions added together represent the total difference between $\mu_{i_{Max}}$ and the standard case. Here the subscript i denotes Cl_{Alb} for the wind maps and RH_{Conv} for precipitation.

This is consistent with the decreasing relative importance of ENSO-associated anomalies in the precipitation field from the Warm Pool to South America. Globally, the model reproduces well the ENSO response for all parameter choices, the error in the regressions being very small (at least three times smaller than the error in the representation of the precipitation climatology). Independent of the parameter choices, the model is capable of simulating the globally averaged ENSO response in rainfall (and in all low level atmospheric variables we tested). The error is larger where the climate signal is the strongest, i.e., over the Pacific Warm Pool, where it follows the global tendencies. An opposite curvature
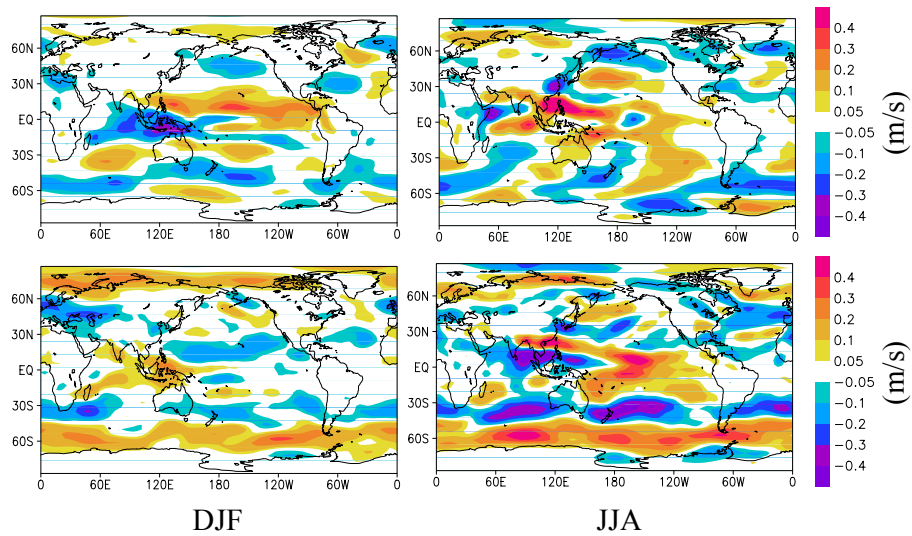
Fig. 6. Ensemble-mean measures of the sensitivity of the modeled error of the Niño 3.4 regression of zonal wind at 925 hPa at WGust = 7.0 in boreal winter (December–February) on the left, and summer (June–August) on the right. The linear contribution a_i is shown on top multiplied by $\mu_{i_{Max}}$ to give units of $m\,s^{-1}$. The quadratic contribution b_{ii}, similarly given as $b_{ii}\mu_{i_{Max}}^2$ in $m\,s^{-1}$, is shown in the bottom panels. The two contributions added together represent the total difference between $\mu_{i_{Max}}$ and the standard case. Here the subscript i denotes WGust.
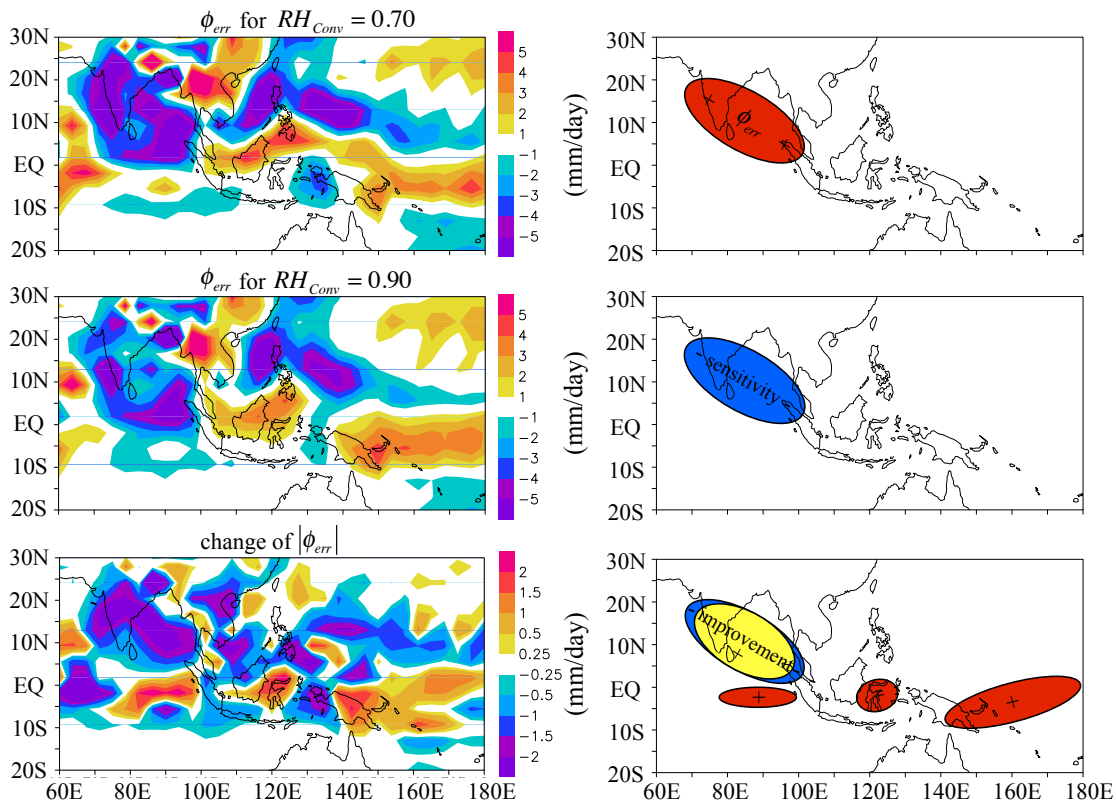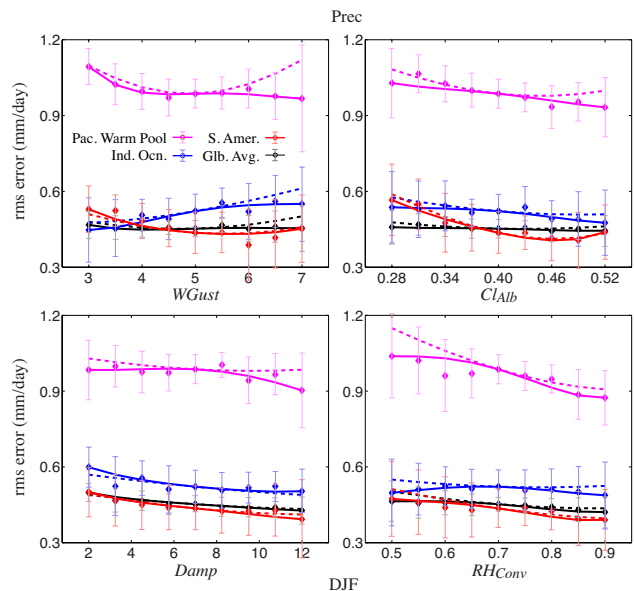

Fig. 7. On left, top: error of the AGCM precipitation climatology with all μ_i at standard value, middle: as above but with RH_{Conv} at its optimum ($RH_{Conv} = 0.90$), bottom: absolute value of the model error difference between top and middle panels. On right: schematic showing the region where error is largest (top), where sensitivity to RH_{Conv} is greatest (middle), and where parameter optimization improves (in yellow) or downgrades (in red) model performance (bottom). All plots focus on the Indian Ocean and western Pacific areas.
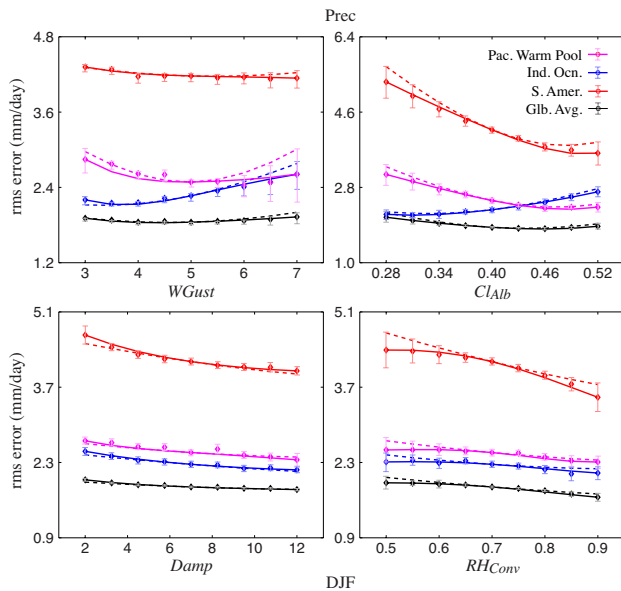
Fig. 8. Root-mean-square error of the AGCM precipitation climatology in December–February for target regions relative to the CPC Merged Analysis of Precipitation data set. The AGCM values are compared to the quadratic metamodel reconstruction based on the endpoints for each parameter, and to its linear counterpart. Global average in black, Pacific Warm Pool in magenta, Indian Ocean basin in blue and South America in red. Units on abscissa are $\mathrm{m\,s^{-1}}$ for WGust, days for Damp and nondimensional for Cl_{Alb} and RH_{Conv}. Error bars quantify the spread of the RMS error within the ten ensemble members (± 1 standard error of the ensemble).

Fig. 9. Root-mean-square precipitation error of the AGCM regression coefficients for the El Niño 3.4 index in December–February for target regions relative to the CPC Merged Analysis of Precipitation data set. The AGCM values are compared to the quadratic metamodel reconstruction based on the endpoints for each parameter, and to its linear counterpart. Global average in black, Pacific Warm Pool in magenta, Indian Ocean basin in blue and South America in red. Units on abscissa are $\mathrm{m\,s^{-1}}$ for WGust, days for Damp and nondimensional for Cl_{Alb} and RH_{Conv}. Error bars estimate the spread of the RMS error within the ten ensemble members (± 1 standard error of the ensemble).

to that for South America is found over the Indian Ocean for varying WGust and Cl_{alb}, while error and signal averaged globally are small.

Figure 10 compares the regional objective functions with the global ones for the low level meridional winds (left) and for zonal 200 hPa zonal winds (right) when varying WGust and RH_{Conv} in DJF. The modeled climatologies for both wind fields are not very sensitive to RH_{Conv} changes, and a larger than standard value of RH_{Conv} will decrease the precipitation error (see Fig. 8) in most of the globe without disrupting the near surface wind, while causing a limited downgrading in the upper level wind over the Indian Ocean and the Pacific Warm Pool. Variations in WGust, on the other hand, show a far more complex response. At 200 hPa the globally averaged climatology displays an optimum around the standard value, but the error over the Indian Ocean and the warm pool is minimized only at the maximum WGust explored, with a significant improvement in the Indian basin, while a smaller than standard WGust is beneficial over South America. Close to the Earth's surface, the meridional wind dependence in the Indian and Pacific regions is in the same direction of the upper level zonal wind field, and over South America the minimum WGust value minimizes the RMS error. For both wind fields considered, the tendency for varying WGust is

opposite to that of precipitation over the Indian Ocean and South America.

We summarize the dilemmas associated with choosing model parameters while optimizing specific regional patterns (a problem of interest also to the regional modeling community) by visualizing the different relative dependence of the parameters considered, two at a time, for the Indian Ocean basin (Fig. 11) and the Pacific Warm Pool (Fig. 12). We present results for DJF precipitation, to allow for a comparison with the correspective global analysis shown in Fig. 3, and again we indicate the areas in which the increase in regional RMS error from its minimum is less than 1, 5, and 10 %. The optimization of WGust and Cl_{Alb} in any one basin, for example, causes a large degradation in the other, while the optimization based on the global RMS error provides intermediate values between the ones found here. It is evident that the parameter dependence of regional patterns is complicated by the fact that improvement in one variable over one region may cause strong deterioration in another area.
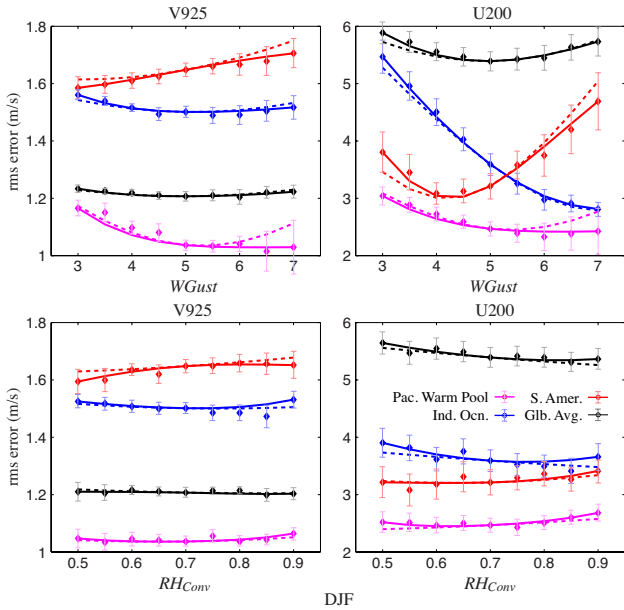
Fig. 10. Root-mean-square error of the AGCM meridional wind at 925 hPa (left) and zonal wind at 200 hPa (right) climatology in December–February for target regions relative to the National Centers for Environmental Prediction reanalysis. The AGCM values are compared to the quadratic metamodel reconstruction based on the endpoints for each parameter, and to its linear counterpart. Global average in black, Pacific Warm Pool in magenta, Indian Ocean basin in blue and South America in red. Units on abscissa are $\mathrm{m\,s^{-1}}$ for WGust (top panels) and nondimensional for $\mathrm{RH_{Conv}}$ (bottom panels). Error bars quantify the spread of the error within the ten ensemble members (± 1 standard error of the ensemble).

5.1 Analytical solutions

We finally recall, as discussed in Neelin et al. (2010), that analytic approximations based on the metamodel provide an alternative, computationally cheap, tool to highlight the regional dilemma typical of the optimization problem whenever spatial averages, instead of global ones, are of interest. The analysis of the spatial patterns in the analytic solution allows the identification of regions where the parameter dependence does not follow the global average one, without the need of calculating a large number of regionally targeted objective functions. Given that the quadratic fit to the climate field ϕ provides RMS error objective functions f (here omitting subscript k) with fourth-order terms in μ, we can obtain an analytical approximation using the square error (which has the same extrema as RMS), expand in μ about a reference value (for simplicity here we use the standard case), and retain the second-order terms in μ for expediency. After differentiating, we obtain

DJF

Indian Ocean Region
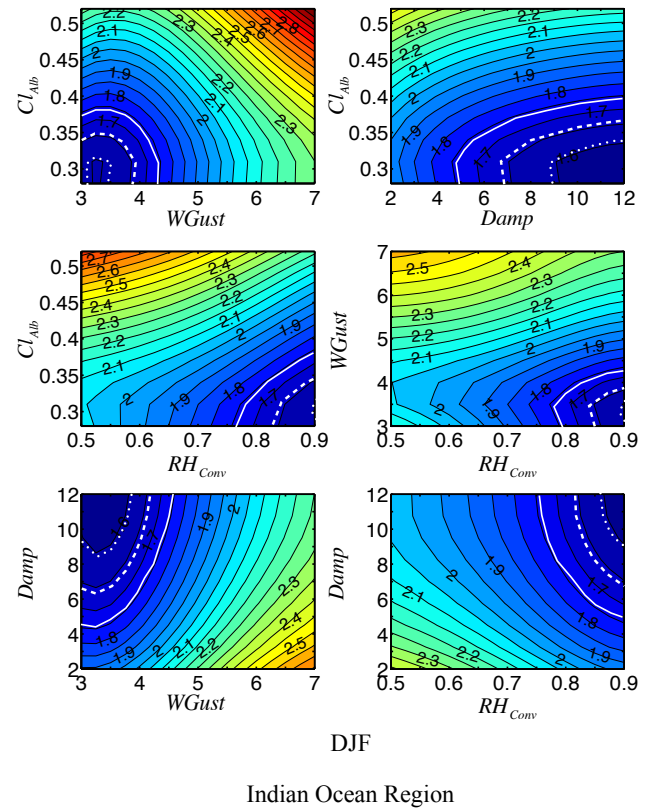
Fig. 11. Pairwise planes of the four-dimensional parameter space showing the root-mean-square precipitation error in $\mathrm{mm\,day^{-1}}$ estimated by the quadratic metamodel over the Indian Ocean basin in December–February. The white dotted (dashed; solid) contours indicate regions where the model error increases by 1 % (5 %; 10 %) compared to its minimum value at the parameter optima. Units on axes are $\mathrm{m\,s^{-1}}$ for WGust, days for Damp and nondimensional for $\mathrm{Cl_{Alb}}$ and $\mathrm{RH_{Conv}}$.

$$\nabla_\mu f = g + \mathbf{A}\mu \tag{2}$$

$$g_i = 2\langle a_i \phi_{\mathrm{err}} \rangle; \tag{3}$$

$$\mathbf{A}_{ij} = \mathbf{A}_{ji} = 2\left(\langle a_i a_j \rangle + 2\langle b_{ij} \phi_{\mathrm{err}} \rangle\right). \tag{4}$$

Here g is the gradient in μ, $\mathbf{A}$ is the Hessian matrix derived from the curvature of the metamodel fit, both evaluated at the standard case, and ϕ_{err} is the error of the standard case with respect to observations.

The off-diagonal terms in $\mathbf{A}$ originate from the linear contributions $\langle a_i a_j \rangle$ and from $\langle b_{ij} \phi_{\mathrm{err}} \rangle$. They are usually small compared to the diagonal given that parameter pairs do not interact much (Neelin et al., 2010; Bellprat et al., 2012), and they can be neglected when trying to get a general idea of the basic model behavior. Practically, g_i provides a spatial projection of the sensitivity a_i with ϕ_{err}, the error of the standard case, and a reduction in error at regional scales is achieved only if the spatial pattern of a_i matches the error pattern. If,

for example, a_i has large amplitude in a localized area that does not project on the error, then the global RMS optimization will introduce a significant error in such area despite yielding a solution that reduces the model error globally. Therefore, maps of the spatial patterns of the different contributions in Eq. (2) allow for quantifying at a regional level the properties of objective functions defined for global quantities. Examples are provided in Fig. 13 for 200 hPa zonal wind dependence on WGust in JJA, and for rainfall dependence on RH_{Conv}, again in DJF. On top we show $a_i\phi_{err}$ normalized by the diagonal contribution $\langle a_i^2 \rangle$, which is proportional to the objective function gradient g_i. If the majority of points is of the same sign, as in the case for the precipitation dependence on RH_{Conv} in boreal summer, presented in Neelin et al. (2010, see their Fig. 5), then it is to be expected that regionally focused optimizations will yield optima in the same direction as found for the global averages. Specifically, optima should be found close to the WGust and Cl_{alb} minima for the wind fields, for which, however, the global RMS is only slightly reduced compared to the standard case, given that the model error is large at high latitudes, where changes in the surface wind gustiness cannot influence significantly the model climatology, and at the RH_{Conv} maximum for precipitation (see Fig. 1). Here, however, the same sign requirement is not satisfied everywhere; for precipitation it does not hold over the equatorial Indian Ocean, for the physical reason noticed in the discussion of Fig. 7, or in part of Brazil's Nordeste, and in the case of the wind fields, positive and negative patterns alternate throughout the domain.

The middle panels in Fig. 13 display the quantity $2b_{ii}\phi_{err}\langle a_i^2 \rangle$, and provides a measure of the importance of nonlinearity in the diagonal term. If the average of this term over the area of interest is negative, and its magnitude is greater than one, then the curvature of the metamodel fit is reversed, as in the case of global precipitation (left panels). Finally, at the bottom we present the model error difference between the case when μ_i is at its minimum (maximum) value for winds (precipitation), and when μ_i is at the standard. Negative values are indicative of regions where the model error decreases compared to the error of the standard case, and vice versa for positive values. It is clear that the analysis of the $a_i\phi_{err}$, together with $2b_{ii}\phi_{err}\langle a_i^2 \rangle$ whenever the metamodel fit points to a significant role for nonlinearities, provides an alternative low-cost way to gather information on the model parameter dependence at regional scales that assists the modeler's ability to digest high dimensional spatial information on the impact of a parameter change.

6   Conclusions

In this work, we evaluate the sensitivity of an atmospheric circulation model to changes in four of its tunable parameters using a metamodeling approach in which the coefficients of the quadratic expansion in parameter space include spatial
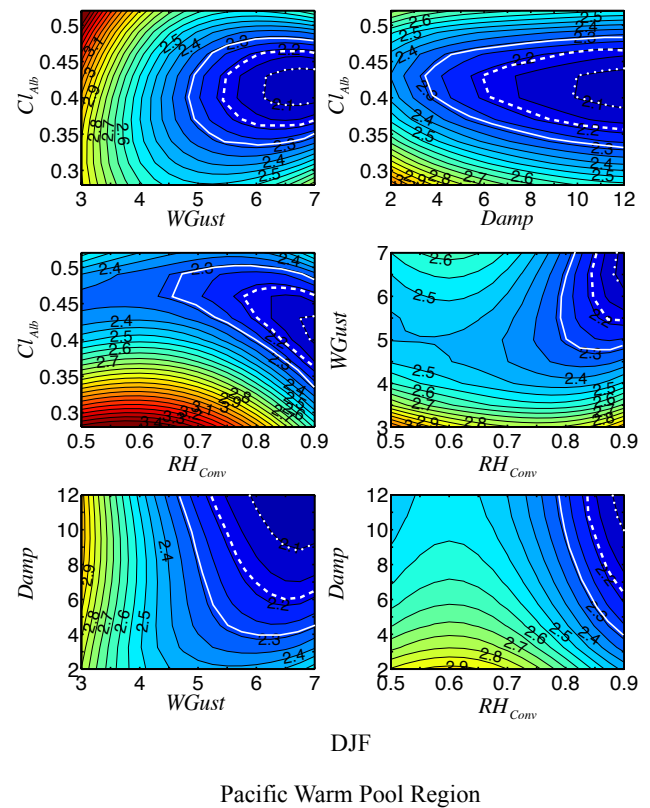

DJF

Pacific Warm Pool Region

Fig. 12. As in Fig. 11 but for the Pacific Warm Pool region.

and seasonal dependence as in Neelin et al. (2010). The metamodel offers a strategy to systematize the calibration and analysis of the parameter dependence in climate models whenever the smoothness of the error metrics for varying parameter values is verified. Such strategy consists in identifying the parameters and metrics, performing few GCM runs, fitting the metamodel, exploring the decision dilemmas encountered by varying the parameters, and finally choosing the best, or least unsatisfactory, parameter set. Rather than employing these tools in "blind" automated optimization based on global objective functions, we emphasize their use for identification of sensitive parameter ranges and of trade-offs among multiple objective functions involving different physical climate variables and spatial regions of interest. From the metamodel, we estimate objective functions based on both global spatial averages and the distribution of regional patterns of several common fields of climatic interest, from precipitation to winds, in different seasons and in annual averages. We also present the parameter dependence analysis of the ICTP AGCM response to ENSO sea surface temperatures, finding substantial nonlinearity in most low level variables. In all cases, we find that the error objective functions vary sufficiently smoothly through the explored parameter ranges for an analytic metamodel approach

U 200 hPa, JJA, *WGust*   Prec, DJF, *RH*$_{Conv}$
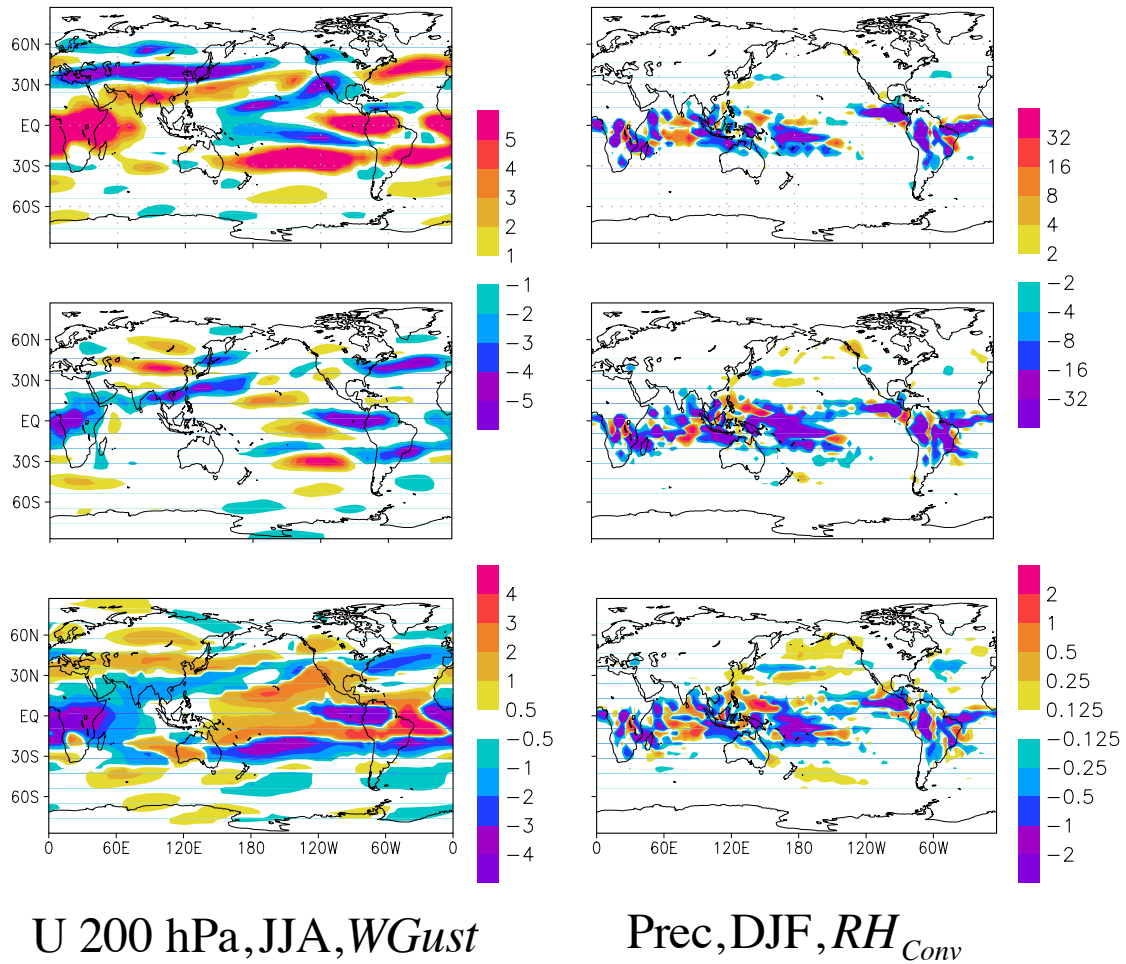
Fig. 13. Contributions to the analytic solution for the ensemble mean zonal wind at 200 hPa in June–August in the WGust direction (left panels), and for precipitation in December–February in the RH$_{Conv}$ direction (right panels). Top panels: $a_i \phi_{err}$ normalized by the diagonal contribution $\langle a_i^2 \rangle$. Middle panels: contribution $2 b_{ii} \phi_{err} / \langle a_i^2 \rangle$ (see Eq. 4). Bottom panels: the model error difference between μ_i at its optimum ($\mu_{i_{Min}}$ for the wind fields and $\mu_{i_{Max}}$ for precipitation), and at standard value.

to be useful, and find the quadratic truncation adequate to estimate leading nonlinear effects.

Guided by the metamodel, we present a strategy to visualize the dilemmas associated with the parameter selection, and to quantify the trade-offs of given parameter choices. Common ambiguities result from not having direct control over how climate patterns change with parameter changes. Most dilemmas faced by modelers are associated with parameter (or parameterization) changes that improve one field or regional pattern over a given season, but cause degradation in another, or improve the model climatology but degrade its representation of interannual variability. Because different applications would place different weight on the accuracy of the simulation in particular regions or variables, this results in the situation typical of multi-objective optimization: while some sets of parameter choice yield improvements in all objective functions of interest, within these sets trade-offs arise

among different objective functions. In climate applications, the fidelity of the simulation evaluated for multiple variables over many small regions in space and season can be viewed as a very high dimensional set of objective functions. A key aim is thus to present information on the trade-offs among these in a way that is digestible to the modeler. We have shown that using the metamodel it is possible (a) to visualize global- and regional-scale biases for any climate variable of interest, any season, and different performance metrics at little computational cost (using climate model simulations similar to those typically carried out to evaluate more basic elements of parameter sensitivity), and (b) to quantify objectively those biases without performing an excessively large number of simulations. For example, information such as the location of the optimum and the minimum global-average RMS error for individual climate variables over the parameter space provides useful information on just how severe

the trade-offs are among different variables. Spatial maps of fields that directly enter the optimization problem, such as the product of the linear sensitivity spatial pattern with the error field at the standard parameter settings, can rapidly provide intuition regarding spatial regions that will be improved or degraded by a particular parameter change.

Many of the dilemmas are unresolvable as different modelers have different priorities regarding which pattern should be improved at the expense of others. The metamodel, however, provides a constructive tool to identify those spatial patterns and to characterize what the most acute dilemmas are for each given metric.

Supplementary material related to this article is available online at http://www.geosci-model-dev.net/6/1673/2013/gmd-6-1673-2013-supplement.zip.

References

Archibald, R., Chakoumakos, M., and Zhuang, T.: Characterizing the elements of Earth's radiative budget: applying uncertainty quantification to the CESM, Elsevier Science Journal, Procedia Comput. Sci., 9, 1014–1020, doi:10.1016/j.procs.2012.04.109, 2012.

Barimalala, R., Bracco, A. and Kucharski, F.: The representation of the South Tropical Atlantic teleconnection to the Indian Ocean in the AR4 coupled models, Clim. Dynam., 38, 1147–1166, doi:10.1007/s00382-011-1082-5, 2012.

Bellprat, O., Kotlarski, S., Lüthi, D., and Schär, C.: Objective calibration of regional climate models, J. Geophys. Res.-Atmos., 117, D23115, doi:10.1029/2012JD018262, 2012.

Bourke, W.: A multilevel spectral model, I. Formulation and hemispheric integrations, Mon. Weather Rev., 102, 687–701, 1974.

Bracco, A., Kucharski, F., Rameshan, K., and Molteni, F.: Internal variability, external forcing and climate trends in multi-decadal AGCM ensembles, Clim. Dynam., 23, 659–678, 2004.

Collins, M., Booth, B. B. B., Bhaskaran, B., Harris, G. R., Murphy, J. M., Sexton, D. M. H., and Webb, M. J.: Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles, Clim. Dynam., 36, 1737–1766, doi:10.1007/s00382-010-0808-0, 2010.

Covey, C., AchutaRao, K. M., Cubasch, U., Jones, P., Lambert, S. J., Mann, M. E., Phillips, T. J., and Taylor, K. E.: An overview of results from the Coupled Model Intercomparison Project, Global Planet. Change, 37, 103–133, 2003.

Dai, A.: Precipitation characteristics in eighteen coupled climate models, J. Climate, 19, 4605–4630, 2006.

Hakkarainen, J., Ilin, A., Solonen, A., Laine, M., Haario, H., Tamminen, J., Oja, E., and Järvinen, H.: On closure parameter estimation in chaotic systems, Nonlin. Processes Geophys., 19, 127–143, doi:10.5194/npg-19-127-2012, 2012.

Hairer, M. and Majda, A. J.: A simple framework to justify linear response theory, Nonlinearity, 23, 909–922, doi:10.1088/0951-7715/23/4/008, 2010.

Held, I. M. and Suarez, M. J. A.: proposal for the intercomparison of the dynamical cores of atmospheric general circulation models, B. Am. Meteorol. Soc., 75, 1825–1830, 1994.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, B. Am. Meteorol. Soc., 77, 437–471, 1996.

Kidston, J. and Gerber, E. P.: Intermodel variability of the poleward shift of the austral jet stream in the CMIP3 integrations linked to biases in 20th century climatology, Geophys. Res. Lett., 37, L09708, doi:10.1029/2010GL042873, 2010.

Knight, C. G., Knight, S. H. E., Massey, N., Aina, T., Christensen, C., Frame, D. J., Kettleborough, J. A., Martin, A. S. P., Sanderson, B., Stainforth, D. A., and Allen, M. R.: Association of parameter, software, and hardware variation with large-scale behavior across 57,000 climate models, P. Natl. Acad. Sci. USA, 104, 12259–12264, 2007.

Kucharski, F., Molteni, F., and Bracco, A.: Decadal interactions between the western tropical Pacific and the North Atlantic oscillation, Clim. Dynam., 26, 79–91, 2006.

Kucharski, F., Bracco, A., Yoo, J. H., and Molteni, F.: Low-frequency variability of the Indian monsoon – ENSO relation and the Tropical Atlantic: the "weakening" of the '80s and '90s, J. Climate, 20, 4255–4266, 2007.

Kucharski, F., Bracco, A., Yoo, J. H., Tompkins, A., Feudale, L., Ruti, P., and dell'Aquila, A.: A Gill-Mastuno-type mechanism explains the Tropical Atlantic influence on African and Indian Monsoon rainfall, Q. J. Roy. Meteor. Soc., 135, 569–579, doi:10.1002/qj.406, 2009.

McWilliams, J. C.: Irreducible imprecision in atmospheric and oceanic simulations, P. Natl. Acad. Sci. USA, 104, 8709–8713, 2007.

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J., and Taylor, K. E.: The WCRP CMIP3 multi-model dataset: a new era in climate change research, B. Am. Meteorol. Soc., 88, 1383–1394, 2007.

Molteni, F.: Atmospheric simulations using a GCM with simplified physical parameterizations. I. Model climatology and variability in multi-decadal experiments, Clim. Dynam., 20, 175–191, 2003.

Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, doi:10.1038/nature02771, 2004.

Neale, R. B., Richter, J. H., Conley, A. J., Park, S., Lauritzen, P. H., Gettelman, A., and Williamson, D. L.: Description of the NCAR Community Atmospheric Model (CAM 4.0), NCAR Technical Note, NCAR TN-485+STR, available at: http://www.cesm.ucar.edu/models/ccsm4.0/cam/docs/description/cam4_desc.pdf, 2010.

Neelin, J. D., Bracco, A., Luo, H., McWilliams, J. C., and Meyerson, J. E.: Considerations for parameter optimization and sensitivity in climate models, P. Natl. Acad. Sci. USA, 107, 21349–21354, doi:10.1073/pnas.1015473107, 2010.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C. and Kaplan, A.: Global analyses of SST, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res., 108, 4407, doi:10.1029/2002JD002670, 2003.

Rougier, J., Sexton, D. M. H., Murphy, J. M., and Stainforth, D.: Analyzing the climate sensitivity of the HadSM3 climate model using ensembles from different but related experiments, J. Climate, 22, 3540–3557, 2009.

Rowlands, D. J., Frame, D. J., Ackerley, D., Aina, T., Booth, B. B. B., Christensen, C., Collins, M., Faull, N., Forest, C. E., Grandey, B. S., Gryspeerdt, E., Highwood, E. J., Ingram, W. J., Knight, S., Lopez, A., Massey, N., McNamara, F., Meinshausen, N., Piani, C., Rosier, S. M., Sanderson, B. M., Smith, L. A., Stone, D. A., Thurston, M., Yamazaki, K., Yamazaki, Y. H., and Allen, M. R.: Broad range of 2050 warming from an observationally constrained large climate model ensemble, Nat. Geosci., 5, 256–260, doi:10.1038/NGEO1430, 2012.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettlemen, A., Golaz, J. C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P., and Haynes, J.: Dreary state of precipitation in global models, J. Geophys. Res., 115, D24211, doi:10.1029/2010JD014532, 2010.

Swart, N. C. and Fyfe, J. C.: Observed and simulated changes in the Southern Hemisphere surface westerly wind-stress, Geophys. Res. Lett., 39, L16711, doi:10.1029/2012GL052810, 2012.

Trenberth, K. E. and Stepaniak, D. P.: Indices of El Niño evolution, J. Climate, 14, 1697–1701, 2001.

van Oldenborgh, G. J., Doblas-Reyes, F. J., Wouters, B., and Hazeleger, W.: Decadal prediction skill in a multi-model ensemble, Clim. Dynam., 38, 1263–1280, doi:10.1007/s00382-012-1313-4, 2012.

Xie, P. and Arkin, P. A.: Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates and numericalmodel outputs, B. Am. Meteorol. Soc., 78, 2539–2558, 1997.