

# Chapter 3

## Response and Alternative Theories in Climate Change



### Introduction by Guido Visconti

The response theory applied to study climate change has its origins in a couple of papers published by Cecil Leith in the 1970s. It is based on the Fluctuations Dissipation Theorem (FDT) whose classical application is to the Brownian motion. In this case, the random motion of a particle, in the fluid, is forced by the thermal movement of the molecules. During the motion the energy of the particle is dissipated by the viscosity of the fluid and converted to heat that contribute to maintain the temperature of the fluid. Following the words of Leith, “In this analogy the detailed motions of gas molecules correspond to the weather and the statistical properties of a gas such as temperature and pressure correspond to the climate” with some additional problems. For practical purposes climate is defined as the average of weather in some location and this must include the mean but also the standard deviation from the mean. Beside the mean must be carried out over some time interval and this is a crucial point. As a matter of fact, climate changes with time so that the interval must be long enough but not too long that would eliminate long term variations. In statistical mechanics such a problem is circumvented by recurring to the concept of ensemble mean. The ensemble is constituted by a large number (infinity?) of identical systems. This concept is not very practical to define climate because it would require studying a large number of Earths, subject to the same conditions including forcing. There is an alternative definition of climate as the distribution of the different variables (temperature, precipitation, etc.) with the weather being a sample from that distribution.

The problem was solved by Leith assuming that the system recovers from a small natural anomaly in the same way that it does from one induced by external forcing and this leads to expressing the sensitivity of the model (that not necessarily coincides with that of climate) as a matrix response function. As observed in a lucid paper by Thomas Bell some years later this matrix is of fundamental importance and in theory can be calculated but not “easily accessible to observation”. Consequently, the only

way to apply FDT theorem to climate was to do everything in the model, that is, calculate the sensitivity of the model to different perturbations.

We have to consider however that the climate system, just because it is under perturbation, is not in a steady state and it is known that fluctuations in a perturbed system have different spectral character with respect to an unperturbed one. For these reasons FDT is not strictly applicable to such a system and the Linear Response Theory (LRT) was developed to solve this problem. In practice LRT boils down to calculate the Green function for the system and then it is applied to predict the signal from an assigned perturbation. The Green function can be calculated either from a single General Circulation Model (GCM) or from an ensemble of such models like those of CMIP (Coupled Model Intercomparison Project) and it is calculated as the time derivative of the mean response to an assigned perturbation. Usually reference is to an increase of 1% per year of CO<sub>2</sub> or to an abrupt doubling of the same greenhouse gas. Then the same function can be used to evaluate the climate signal for any perturbation.

It is to note the fundamentally LRT gives average value of the Global Mean Surface Temperature (GMST) so it is fine for evaluating the signal from an assigned scenario of greenhouse gas emission but cannot reproduce the details of a GCM run that predicts the behavior of several climate variables and their geographical distribution. Beside it assumes the GCM runs as reference neglecting all the philosophical and technical problems that plague GCMs. As a matter of fact, the results as far as temperature is concerned are quite reasonable even examining the latitudinal distribution. However, the results for precipitation are quite poor making clear that the method has not the necessary physical details and mechanisms that produce precipitation. However, the method remains very much suitable for all the exercises of the IPCC (Intergovernmental Panel on Climate Change) and actually LRT has been used also to develop a so called Stochastic Space State Model that combines emission scenario with the climate forcing introducing some stochastic forcing in both carbon dioxide and temperature. This produces not just the average value of temperature changes but the temperature distribution for the different scenarios.

This application of LRT is quite interesting but at the least now cannot predict the geographical distribution of the climate changes and so does not constitute an input for example to evaluate regional climate change. Sometimes you got the impression that theoretical (mathematical) work on climate change is taking revenge on several years of honest and neglected work of the GCM practitioners. There are some people that insist on putting climate studies in some sophisticated mathematical framework. This follows an old habit of the physicists that believe in the old reductionist practice.

## Machine-Generated Summaries

Keywords: *Ensemble, parameter, error, space, projection, theory, ppe, member, approach, small, future, parameter space, surface, run, scale.*

## ***Finding Plausible and Diverse Variants of a Climate Model. Part 1: Establishing the Relationship Between Errors at Weather and Climate Time Scales***

<https://doi.org/10.1007/s00382-019-04625-3>

### **Abstract-Summary**

In this first part, the extent to which climate biases develop at weather forecast timescales is assessed with two PPEs, which are based on 5-day forecasts and 10-year simulations with a relatively coarse resolution (N96) atmosphere-only model.

The study confirms more robustly than in previous studies that investigating the errors on weather timescales provides an affordable way to identify and filter out model variants that perform poorly at short timescales and are likely to perform poorly at longer timescales too.

The use of PPEs also provides additional information for model development, by identifying parameters and processes responsible for model errors at the two different timescales, and systematic errors that cannot be removed by any combination of parameter values.

### **Extended**

In this first part, we build on ideas from previous studies to use model performance at short timescales (here 5 days) to filter the parameter space (Rodwell and Palmer [1]), and to use 5-day forecast errors to infer something about model errors at longer timescales (e.g. Ma and others [2]).

In part II, we show how this result can be exploited, in an application which is to select a number of model variants capable of providing plausible simulations of historical climate and diverse projections of future climate change.

### **Introduction**

This paper is the first of two aimed at designing a “small” perturbed parameter ensemble (PPE) of plausible simulations based on a relatively expensive global climate model that can be used in producing climate projections for adaptation planning.

The second paper (Karmalkar and others [3]; hereafter Part II) focuses on a methodology that uses this relationship in model performance across weather and climate timescales to identify a small PPE of plausible simulations by screening out parameter combinations.

Some national projections like those from the Netherlands (van den Hurk and others [4]) or Australia (CSIRO and Bureau of Meteorology [5]) use the CMIP5 multimodel ensemble to represent modelling uncertainty, but there are advantages to also providing a PPE derived from a single climate model.

We want to achieve this with coupled ocean–atmosphere simulations but for this two part study we limit ourselves to atmosphere-only models to test the basic concept of screening out poorly performing parameter combinations whilst maintaining a diversity of credible process behaviour and future climate response.

### **Experimental Design and Elicitation**

Design of a PPE first requires selection of a model configuration to perturb, then eliciting prior probability distributions for the parameters to perturb as chosen by the parameterisation experts, and finally deciding how to sample the parameters.

SHELF also allows for the elicitation to be completed after the meeting as long as the experts understand what is required of them and have experience with one example of a parameter already.

To elicit the plausible range for each parameter, the most important aspect is to explain to the experts that the simulations will be evaluated against a wide range of observational metrics so that (1) PPE members can be ruled out as implausible and (2) the final uncertainty quantification can be based on constrained parameter ranges.

Many parameter ranges were based on the experts' own analyses of very high resolution process models such as Large Eddy Simulations or Cloud Resolving Models.

### **Data and Methods**

For TAMIP, in this study, we focus on evaluating the mean forecast error across the 16 start dates for day 2 and day 5 of the simulations.

Using the average errors across all 16 initial conditions makes the results more robust, though it limits us to relating the TAMIP MSEs to annual mean MSEs from the longer term ATMOS simulations.

The high correlations for surface air temperature and precipitation suggest that the error growth of each variable is largely due to the same parameters at days 2 and 5.

The day 5 errors for outgoing shortwave radiation in the clear sky are slightly less than the day 2 errors, but this comes from the Arctic and Antarctic where sea ice is fixed over the 5-day forecast to its initial value, and so should be considered a feature of the design of the TAMIP experiment.

### **Results**

The similarity between the error patterns at the two timescales also applies for most variables and single model variants, with only a few variables (in particular 250 hPa eastward wind and specific humidity) showing correlations close to zero for an appreciable fraction of the ensemble.

The fraction might be reduced with smoother TAMIP patterns as described above, but until this is tested, the distributions for these variables show that the link between the two timescales is not robust across all parameter space, and so it cannot be assumed that the forecast errors are indicative of the climate biases.

The negative uncentred correlation in this region reflects the development of process errors on timescales longer than 5 days that lead to the change the sign of the ensemble mean bias noted above.

For model tuning where the search is only across a more focussed parameter space, there is still information in the patterns of 5-day forecast errors about the mean climate biases.

## Discussion

The results show that this size of PPE, which is several times the number of parameters, offers new insights into the relationship between errors at different timescales and the underlying processes, and is potentially valuable for model tuning and prioritising which model errors need to be reduced by model development.

The impact of our results is discussed below in terms of the emergent relationships between errors on the two timescales and how the influences of the parameters affects these, followed by the implications for experimental design and then model development.

Strong emergent relationships exist between the model errors at 5-day and 5-year timescales, they can be exploited to inform the efficient design of a PPE suitable for predictions across multiple timescales.

It is very likely that good performance at the 5-day timescale would also be an important indicator of credibility of climate model projections of climate variability and extremes, although our simulations were not long enough to support investigation of such links in the present paper.

## Acknowledgement

*A machine generated summary based on the work of Sexton, D. M. H.; Karmalkar, A. V.; Murphy, J. M.; Williams, K. D.; Boutle, I. A.; Morcrette, C. J.; Stirling, A. J.; Vosper, S. B. (2019 in Climate Dynamics).*

## ***Finding Plausible and Diverse Variants of a Climate Model. Part II: Development and Validation of Methodology***

<https://doi.org/10.1007/s00382-019-04617-3>

## Abstract-Summary

Exploratory work towards developing a strategy to select variants of a state-of-the-art but expensive climate model suitable for climate projection studies.

The strategy combines information from a set of relatively cheap, idealized perturbed parameter ensemble (PPE) and CMIP5 multi-model ensemble (MME) experiments, and uses two criteria as the basis to select model variants for a PPE suitable for future projections: (a) acceptable model performance at two different timescales, and (b) maintaining diversity in model response to climate change.

This relationship is used to filter out parts of parameter space that do not give credible simulations of present day climate, while minimizing the impact on ranges in forcings and feedbacks that drive model responses to climate change.

We use statistical emulation to explore the parameter space thoroughly, and demonstrate that about 90% can be filtered out without affecting diversity in global-scale climate change responses.

This leads to the identification of plausible parts of parameter space from which model variants can be selected for projection studies.

Comparisons with the CMIP5 MME demonstrate that our approach can produce a set of plausible model variants that span a relatively wide range in model response to climate change.

### **Extended**

This work in progress will be documented in future papers, underpinned by the proof of concept developments described here.

### **Introduction**

Multi-model ensembles (MMEs) comprising climate model simulations carried out by various institutions all over the world (e.g. CMIP3, CMIP5 archive) have been used widely to provide a range in climate change projections (Meehl et al. [6]; Taylor et al. [7]).

As models become more sophisticated, for example, due to increases in their horizontal and vertical resolutions to improve representation of various aspects of climate variability and extremes (Scaife et al. [8]), creating a large ensemble for probabilistic projections becomes increasingly expensive.

The PPE is based on the atmospheric component of the Hadley Centre Global Environmental Model version 3 (HadGEM3-A; Hewitt et al. [9]) and is described in detail in Part I. This paper is heuristic in the sense that we describe a set of atmosphere-only PPE simulations and evaluation techniques capable of informing the subsequent definition of a climate projection system, but without progressing to the final step of evaluating our identified atmosphere model variants in coupled (AOGCM) simulations with a dynamic ocean component.

### **Principles of Methodology**

The selection is based on the following criteria: (1) Assessment of model performance: The plausible variants must have satisfactory performance at Numerical Weather Prediction (NWP; 5 days) and climate (5 or more years) timescales. (2) Diversity in model response to climate change: The selected variants should explore the range of forcings and feedbacks from the entire plausible sub-region of parameter space identified in (1), as far as possible.

The seamless assessment approach, based on the idea that one can diagnose and characterize model errors by assessing performance at different timescales ranging from weather to climate, is very useful in this regard.

The spread in model responses to increasing GHGs is mainly determined by uncertainties in radiative forcings and climate feedback processes (Bony et al. [10]; Webb et al. [11]).

The success of a PPE in either matching or augmenting MME ranges in relevant aspects of climate response depends on the underlying model (Yokohata et al. [12]) and the experimental design.

### **Experimental Details**

We used the Latin hypercube sampling technique (McKay et al. [13]), which ensures that the prior probability of each parameter is sampled evenly over the 21-dimensional

parameter space, to create a 250-member ensemble that allows perturbing all 21 independent parameters simultaneously.

This was determined from the model development exercise for GA4.0 (Walters et al. [14]), and brought the total ensemble size to 251.

These AMIP-style experiments (Gates et al. [15]), which are forced at the ocean–atmosphere interface using observed estimates of SSTs and sea ice, are suitable for studying the impact of poorly constrained atmospheric and land surface parameters on uncertainties in the performance and response of the model.

In order to diagnose forcings and feedbacks, the second phase of the Cloud Feedback Model Intercomparison Project (CFMIP) proposed a set of idealized experiments (CFMIP-2 Experimental Design; Bony et al. [16]).

### **PPE Results**

While a majority of variables show positive relationships over land and oceans, surface variables such as surface air temperature (*tas*) and downwelling longwave radiation at the surface (*rlds*) have much stronger correlations over land because cross-ensemble variations over the oceans are heavily constrained by the use of prescribed SSTs in both experiments.

We must include: (1) variables such as temperature and precipitation that are commonly used and important for understanding impacts of future climate change. (2) Variables that show strong relationships between TAMIP and ATMOS errors, which will allow us to find model variants that perform well at both time scales. (3) Variables that have a ratio of parametric uncertainty to structural uncertainty greater than 1, where the latter denotes the component of error that cannot be reduced by changing parameter values (Rougier [17]).

One of the most important variables, surface air temperature (*tas*) was not chosen, in spite of its TAMIP-ATMOS errors being correlated, because it shows relatively small spread across the ensemble (see Part I).

### **Filtering Out Parts of Parameter Space**

Although our numbers of completed ensemble members (194 for TAMIP and 80 for ATMOS) are relatively small samples of a 21-dimensional parameter space, the Latin Hypercube design fills the space efficiently enough to allow us to build an emulator for the six assessment metrics at each of the two timescales to predict selected model output variables at untried combinations of parameter values.

We determine MSE-based tolerances to rule out parts of parameter space based on performance benchmarks relative to CMIP5, emulator predictions for model crashes and on maintaining ranges of the diversity metrics.

To quantify uncertainty associated with internal variability, we calculated the variance in MSE in a 16-member ATMOS ensemble of the standard version of the model produced by varying only the initial conditions. (4) Emulator error In addition to best-estimate predictions for the ATMOS and TAMIP MSE of untried parameter combinations, our emulators provide estimates of the associated uncertainties.

### **Plausible Model Variants: Selection and Validation**

Once the parameter space is reduced to contain only acceptable and diverse variants of the atmosphere model, the challenge is to pick a small subset of variants—called ‘plausible’, suitable for (notional) use to provide climate change projections using the AOGCM configuration of the model.

The algorithm did have difficulty in picking model variants at the extremes of diversity metrics, due to the presence of relatively few acceptable variants to pick from.

The second criterion may appear counter-intuitive, since a more obvious aim might be to pick the best performing 50 model variants (subject to spanning a range of forcing and feedbacks), rather than variants that sample a range of performance across each of the individual assessment metrics.

Tests showed that without the second criterion, the 50 chosen variants would not include enough of the better performing models across each of the 12 assessment metrics.

### **Emergent Properties**

AMIP-style experiments, where SSTs are not allowed to adjust to changes in the atmosphere, can potentially result in wide variations in the net TOA radiation flux, and provide a suitable design to expose the full consequences of atmospheric modelling errors on this metric.

Also a risk that excessively restricting the range of acceptable values, through comparison with a set of highly tuned multi-model results, could artificially restrict the range of outcomes consistent with uncertainties in the large set of processes that contribute to global energy balance (e.g., Collins et al. [18]).

Our discovery of a range of TOA net fluxes outside the CMIP5 range occurs because the overall ranges in net TOA fluxes and albedo for the ‘acceptable’ model variants are not reduced significantly compared to the emulated range across the full model parameter space.

### **Conclusions**

The methodology—that includes (1) an assessment of model performance at weather and climate timescales for a variety of metrics, (2) the use of benchmarking information from structurally different models [specifically the CMIP5 multi-model ensemble (MME)] and (3) the maintenance of diversity in forcings and feedbacks—allows us to reduce significantly the prior parameter space specified by modelling experts to a sub-region suitable for the selection of ensemble projection system members.

The seamless assessment approach, in particular, shows that the large parameter space can be efficiently explored by running the climate model in weather forecast mode using 5 day “Transpose AMIP” (TAMIP) experiments, in conjunction with the statistical technique of emulation.

In simulations of present day climate, the PPE explores the ranges of skill spanned by the CMIP5 MME for most key climate variables, often with a few model variants



better than the best performer from CMIP5 and a few variants worse than the worst CMIP5 performer.

### **Discussion**

The presence of strong relationships between weather and climate errors for many variables will enable us to use inexpensive NWP hindcasts as an efficient way of pre-screening the parameter space to exclude parts giving rise to physically unrealistic model behavior, before investing in longer climate simulations either in atmospheric or coupled mode.

While an initial assessment of seasonal mean errors showed strong relationships with their annual mean counterparts, this does not necessarily imply that seasonal errors could not play a useful role in refining future assessments of model performance.

We also build emulators separately for TAMIP and ATMOS runs, but given that there is a strong relationship between model performance across weather and climate timescales, it may be more optimal to consider building emulators in future work that linked the two timescales.

### **Acknowledgement**

*A machine generated summary based on the work of Karmalkar, Ambarish V.; Sexton, David M. H.; Murphy, James M.; Booth, Ben B. B.; Rostron, John W.; McNeall, Doug J. (2019 in Climate Dynamics).*

## ***Multivariate Probabilistic Projections Using Imperfect Climate Models Part I: Outline of Methodology***

<https://doi.org/10.1007/s00382-011-1208-9>

### **Abstract-Summary**

This method combines information from a perturbed physics ensemble, a set of international climate models, and observations.

This is important if different sets of impacts scientists are to use these probabilistic projections to make coherent forecasts for the impacts of climate change, by inputting several uncertain climate variables into their impacts models.

Unlike a single metric, multiple metrics reduce the risk of rewarding a model variant which scores well due to a fortuitous compensation of errors rather than because it is providing a realistic simulation of the observed quantity.

The method also has a quantity, called discrepancy, which represents the degree of imperfection in the climate model i.e. it measures the extent to which missing processes, choices of parameterisation schemes and approximations in the climate model affect our ability to use outputs from climate models to make inferences about the real system.

Discrepancy also provides a transparent way of incorporating improvements in subsequent generations of climate models into probabilistic assessments.

The set of international climate models is used to derive some numbers for the discrepancy term for the perturbed physics ensemble, and associated caveats with doing this are discussed.

### **Introduction**

Perturbed physics ensembles (PPEs) provide an alternative strategy for exploring uncertainty in climate modelling (Murphy et al. [19]; Stainforth et al. [20]; Webb et al. [21]; Yokohata et al. [12]), by generating ensembles where each member differs from the standard version of a climate model by having a different set of values for the model parameters.

Murphy et al. [19] used an interpolation technique so that the model variants sampled by the PPE were used to predict the climate sensitivity and the relative skill in simulating some observable aspects of the climate system (in that case, fields of multiannual means of multiple climate variables) for untried points in parameter space.

The method used by Murphy et al. [19] demonstrates several features: use of probability to represent uncertainty; emulation, a technique in which a statistical model is trained on a PPE and then used to predict the output of untried model variants; using observations to constrain the probabilistic projection to higher quality parts of parameter space.

### **Data**

These were the model parameters, observations, model output that corresponds to the observations and prediction variables, the true climate (of which the observations and the model output are uncertain estimates), and the discrepancy which is a link between the model output and the true climate.

Webb et al. [21] used the data from the first stage to choose 128 members that the method of Murphy et al. [19] predicted to be relatively credible model variants that spanned a wide range of parameter space and climate sensitivity.

As the spread of the PPE at larger spatial scales is generally much greater than internal variability, the leading eigenvectors are mainly driven by the changes to model parameters and are therefore representative of the major changes in the physics of the climate model across the PPE.

This is estimated from a 600-year long control integration of HadSM3, the model variant with standard values for the parameters (Barnett et al. [22]).

### **Outline of the Calculations**

The first stage of any Bayesian analysis is the specification of the uncertain objects which make up the joint probability distribution e.g. the model data, the observations.

The second term in the integrand is called the likelihood function of  $x$  given some observed values and is equal to the probability of obtaining the observed values given those values of  $x$ . The third term is the prior distribution of the values for the model parameters.

The general problem of predicting several model outputs constrained by several observations requires an emulator for each element in  $\mathbf{o}$ , and for each prediction variable in  $\mathbf{y}_f$ .

The close relationship between the emulated and modelled values is not guaranteed as some PPEs sample parameter space in a way that is very different to a prior distribution of where the best input is.

### **Specification of Discrepancy Distribution**

In searching for points in the HadSM3 parameter space which best match the physics of a multimodel ensemble member, it is important to base the search for analogues on a wide range of climate variables, in order to reduce the risk that a fortuitous match could be found through a compensation of errors.

For each multimodel ensemble, we used four different points in parameter space from the initial 100,000, chosen randomly from the leading good fits of the initial sample, and estimate four different best analogues.

Variations within the set of four analogues for each multimodel ensemble member are small compared to variations between members, though there are examples (e.g. when attempting to find analogues for the UIUC model), which confirm the importance of sampling initial conditions in the optimisation of the best analogues.

### **Results and Discussion**

For  $N_h = 6$ , 93% of the sampled values had lower probability density than the actual observed values, indicating that the joint prior distribution, which combines climate model data, emulators, parametric uncertainty, the discrepancy, and the choices we make like the number of dimensions used to represent the observed quantities, compares adequately with the actual observed values and we do not expect strong sensitivity in the results or “surprises” as O’Hagan and Forster [23] call them.

By removing the historical discrepancy, fewer sampled points receive a relatively high weight so that the effective sample size becomes smaller by a factor of 4; this leads to a less smooth PDF which underestimates the range of climate sensitivity in comparison with the full posterior distribution, and would lead to an increased risk of poor decisions based on this PDF.

### **Conclusions**

We have simply allowed the multimodel data to determine the relationship between historical and future discrepancy, with the unavoidable caveat that structural errors common to all current climate models are not included.

We believe that including a defensible estimate of discrepancy leads to a more realistic quantification of prediction uncertainties, and allows us to obtain an improved estimate of the spread of possible future climate outcomes consistent with current modelling technology and understanding of climate feedback processes, because we have combined information from a PPE and a multimodel ensemble.

The final advantage is that the framework, especially the emulator, allows us to assess the robustness of our results to a number of key methodological choices,

including the prior distribution of model parameters, discrepancy, the set of multi-model ensemble members, and the choice of the observational metrics used to constrain the prediction.

### **Acknowledgement**

*A machine generated summary based on the work of Sexton, David M. H.; Murphy, James M.; Collins, Mat; Webb, Mark J. (2011 in Climate Dynamics).*

## ***Multivariate Probabilistic Projections Using Imperfect Climate Models. Part II: Robustness of Methodological Choices and Consequences for Climate Sensitivity***

<https://doi.org/10.1007/s00382-011-1209-8>

### **Abstract-Summary**

A method for providing probabilistic climate projections, which applies a Bayesian framework to information from a perturbed physics ensemble, a multimodel ensemble and observations, was demonstrated in an accompanying paper.

This information allows us to account for the combined effects of more sources of uncertainty than in any previous study of the equilibrium response to doubled CO<sub>2</sub> concentrations, namely parametric and structural modelling uncertainty, internal variability, and observational uncertainty.

Such probabilistic projections are dependent on the climate models and observations used but also contain an element of expert judgement.

Two expert choices in the methodology involve the amount of information used to (1) specify the effects of structural modelling uncertainty and (2) represent the observational metrics that constrain the probabilistic climate projections.

We are therefore confident that, despite sampling sources of uncertainty more comprehensively than previously, the improved multivariate treatment of observational metrics has narrowed the probability distribution of climate sensitivity consistent with evidence currently available.

The main caveat is that the handling of structural uncertainty does not account for systematic errors common to the current set of climate models and finding methods to assess the impact of this provides a major challenge.

### **Introduction**

The method in Part 1 uses a Bayesian framework based on Rougier [17] where a joint probability distribution is constructed to contain probabilistic information about the uncertain objects in the climate projection problem: model parameters; observations; the true climate, consisting of the future that we want to predict and the past which we can compare with the observations; and model output, corresponding to our choice of observed climate variables and also variables we want to predict.

By assuming that this set of structural differences are exchangeable with the structural differences of our climate model with the real system i.e. they are effectively sampled from the same distribution, we can pool these prediction errors over the multimodel ensemble and use them to inform the mean and covariance of the discrepancy term.

### **Effect of Dimensionality Used to Represent Historical Climate**

ESS is a measure of how effectively the observational information restricts the prior parameter space to regions of parameter space that are consistent with the observations used to constrain the PDF; so if all weight was assigned to one sample, ESS would be 1 (though this would be a strong indication that the posterior PDF would not be robust if the full sample was repeated); if all samples were assigned equal weight, ESS would be equal to the sample size, indicating no constraint at all from the observations.

For the observational constraint, an increase in  $N_h$  makes it harder to randomly select a point in parameter space that is a reasonable match to the observed values of all  $N_h$  historical eigenvectors, making the weights less evenly distributed, and so ESS decreases.

This indicates that interactions between estimated model errors (obtained by projecting emulated and observed values of relevant climate variables onto our  $N_h$  eigenvectors) and the off-diagonal terms in the discrepancy covariance matrix (representing relationships between the structural component of model errors in different variables) can play a significant role in determining variations in the weights across parameter space, and hence the ESS.

### **Sensitivity Tests**

We check the sensitivity of our probabilistic climate projections to a number of subjective choices that affect the expert prior probability on the model parameters, the discrepancy term, and number of eigenvectors of historical climate variables used to estimate the relative likelihood of points in parameter space.

Based on these sensitivity tests, the median of the posterior PDF of climate sensitivity is between 3.2 and 3.3 K, with the 5th percentile between 2.2 and 2.4 K, and the 95th percentile is between 4.1 and 4.5 K. We can make a direct comparison between our results and those of our previous study Murphy et al. [19].

Considering first the prior PDFs, Murphy et al. [19] used a uniform sampling of parameter space, and found a 95th percentile of the prior PDF of climate sensitivity of 5.3 K. This compares to 5.0 K for the prior PDF in the present study, when an equivalent uniform sampling is tried as one of our sensitivity tests.

### **Conclusions**

To make an expert assessment about the likely range of equilibrium climate sensitivity, Meehl et al. [24] used PDFs of climate sensitivities from two main categories of study (see their Box 10.2 for details).

This is because the AR4 assessment included evidence based on observational constraints offered by past climate change (the first category identified above), which is not considered in our study.

Extensions of our approach to include constraints based on historical climate change are feasible, and offer the prospect of a transparent, quantitative and testable synthesis of much of the evidence from both major categories assessed in AR4.

In our method, the prior knowledge can be based on expert elicitation (e.g. the prior distribution for the model parameters), or on our judgement (e.g. that the other climate models sample a distribution of climate processes not explored by the variants of HadSM3, and so provides a meaningful way to inform the discrepancy term).

### **Acknowledgement**

*A machine generated summary based on the work of Sexton, David M. H.; Murphy, James M. (2011 in Climate Dynamics).*

## ***Climate Model Errors, Feedbacks and Forcings: A Comparison of Perturbed Physics and Multi-Model Ensembles***

<https://doi.org/10.1007/s00382-010-0808-0>

### **Abstract-Summary**

Ensembles of climate model simulations are required for input into probabilistic assessments of the risk of future climate change in which uncertainties are quantified.

Model-error characteristics derived from time-averaged two-dimensional fields of observed climate variables indicate that the perturbed physics approach is capable of sampling a relatively wide range of different mean climate states, consistent with simple estimates of observational uncertainty and comparable to the range of mean states sampled by the multi-model ensemble.

The perturbed physics approach is also capable of sampling a relatively wide range of climate forcings and climate feedbacks under enhanced levels of greenhouse gases, again comparable with the multi-model ensemble.

### **Extended**

The perturbed physics ensembles described here, together with others documented elsewhere, are combined with a statistical emulator of the model parameter space (see e.g., Rougier and others 25 for an example) and a “time-scaling” technique (Harris and others [26] which maps equilibrium to transient responses taking into account any errors that may arise because of a mismatch between the patterns of transient and equilibrium).

The perturbed physics approach can sample a wide range of different model “errors” in two-dimensional time-averaged climate fields for a number of different variables that for many variables are comparable with uncertainties in the observations and comparable with the errors in the members of the multi-model archive.

The perturbed physics approach can sample a wide range of global-mean feedbacks under climate change.

## Introduction

The main motivation for this paper is to document the design and characteristics of a number of perturbed physics ensembles that have been produced as part of an extensive programme of research at the Met Office Hadley Centre to produce regional climate projections (e.g. Murphy and others [27, 28]) and to contrast aspects of those perturbed physics ensembles with corresponding multi-model ensembles.

We might naively assume that the multi-model ensemble contains members with a wide range of different error characteristics, whereas the perturbed-physics approach produces members with very similar baseline climates and thus very similar errors.

We know that the perturbed physics approach is capable of producing model variants with a wide range of different feedback strengths under climate change (e.g. Webb and others [21]; Sanderson and others [29]).

Question 4 is highly relevant when we use ensembles of climate model projections to generate predictions of climate change expressed in terms of PDFs which provide a measure the uncertainty (or credibility) in that prediction.

## Climate Model Ensembles and Variables

For more complex versions of the model (e.g. using a dynamical ocean component rather than a mixed-layer, q-flux or slab component) fewer ensemble members are possible because of the extra resources required to spin-up model versions and run scenario experiments.

In the design of the ensemble, an attempt was made to minimise the average of the root mean squared error of a number of time-averaged model fields while sampling a wide range of surface and atmospheric feedbacks under climate change.

The model versions are therefore suitable for quantifying uncertainty and examining feedbacks, etc. This ensemble uses the fully coupled version of HadCM3 but with perturbations only to parameters in the atmosphere component (an updated version of the ensemble described in Collins and others [18]).

For historical reasons, the sea-ice scheme in HadCM3 is contained in the atmosphere component of the model and parameters in the scheme are perturbed in line with the equivalent S-PPE-M ensemble.

## Model “Errors”

In the slab-ocean multi model ensemble, S-MME, we see a similar range of land SAT biases as in the case of the perturbed physics ensembles, but a somewhat wider range of RMS errors.

Both SST bias and RMS errors are of a similar magnitude in slab-ocean perturbed physics and multi model ensembles and are in many cases smaller than those errors seen in the non-flux-adjusted CMIP3 coupled models (AO-MME).

Global mean biases in precipitation in the slab-model ensembles follow a similar pattern to those in global land surface air temperature and SST in the different ensembles, except that the S-MME has a relatively wider range of biases than any of the other slab-ocean perturbed physics ensembles.

For the surface sensible heat flux, the range of both biases and RMS errors is generally smaller in the perturbed physics ensembles in comparison with the multi-model ensembles.

### Feedbacks and Forcings

In the case of the AO-PPE-O ensemble, with identical HadCM3 atmosphere components but perturbations to parameters in the ocean model, there is a similarly small spread.

Despite the fact that the volcanic forcing time series of stratospheric optical depth is precisely the same in each member of the perturbed physics ensemble, the spread in total negative volcanic radiative forcing is comparable with the spread in the multi-model case in which different input forcing data are used.

LW forcing in 1995–2004 is centred around  $2.4 \text{ W m}^{-2}$  in both the multi-model and perturbed-physics ensembles, with a range of  $1.5\text{--}3.1 \text{ W m}^{-2}$  in the AO-MME case and a smaller range of  $2.1\text{--}2.7 \text{ W m}^{-2}$  in the AO-PPE-A case (in both cases the range is greater than would be expected from natural variability).

### Relating Model Errors to Feedbacks

Having examined model errors and climate change feedbacks in the multi-model and perturbed physics ensembles, we now examine the relationships between them.

To improve models we need to know how to target research to do this, i.e., by quantifying the relationship between error and climate feedback, we may learn which improvements to different aspects of the model simulations will lead to the most progress in reducing uncertainty in predictions.

The only variable in which there is a reasonably high correlation between errors and feedbacks in both perturbed physics and multi model ensembles are the biases in the global mean cloud amount (coefficients around 0.6–0.7, see also Yokohata and others [30]).

For the perturbed physics ensembles there are weak to moderately strong correlations for a number of variables suggesting that the combination of those (and other) variables into a single metric would be a way of constraining the climate feedback parameter.

### Discussion and Conclusions

The perturbed physics approach can sample a wide range of different model “errors” in two-dimensional time-averaged climate fields for a number of different variables that for many variables are comparable with uncertainties in the observations and comparable with the errors in the members of the multi-model archive.

It is possible to produce quite different baseline climates with the perturbed physics approach such that the ensemble-mean appears as the “best” model in comparison with any individual ensemble member.

For regional measures, and for variables not examined here such as variability or extremes, there may be differences between perturbed physics and multi model ensembles which do not fit with these general conclusions.



In our companion work on producing probabilistic climate change projections, we combine perturbed physics and multi-model ensemble information together with observations and estimates of uncertainty in observations to produce projects based on as much information about the climate system as is possible (Murphy and others [27, 28]).

### **Acknowledgement**

*A machine generated summary based on the work of Collins, Matthew; Booth, Ben B. B.; Bhaskaran, B.; Harris, Glen R.; Murphy, James M.; Sexton, David M. H.; Webb, Mark J. (2010 in Climate Dynamics).*

## ***Predicting Climate Change Using Response Theory: Global Averages and Spatial Patterns***

<https://doi.org/10.1007/s10955-016-1506-z>

### **Abstract-Summary**

The provision of accurate methods for predicting the climate response to anthropogenic and natural forcings is a key contemporary scientific challenge.

Response theory allows one to practically compute the time-dependent measure supported on the pullback attractor of the climate system, whose dynamics is non-autonomous as a result of time-dependent forcings.

We assess strengths and limitations of the response theory in predicting the changes in the globally averaged values of surface temperature and of the yearly total precipitation, as well as in their spatial patterns.

We also show how it is possible to define accurately concepts like the inertia of the climate system or to predict when climate change is detectable given a scenario of forcing.

### **Extended**

Response theory allows to practically compute such a time-dependent measure starting from the invariant measure of a suitably chosen reference autonomous dynamics.

### **Introduction**

One needs to consider that the study of climate faces, on top of all the difficulties that are intrinsic to any nonequilibrium system, the following additional aspects that make it especially hard to advance its understanding: the presence of well-defined subdomains—the atmosphere, the ocean, etc.—featuring extremely different physical and chemical properties, dominating dynamical processes, and characteristic time-scales; the complex processes coupling such subdomains; the presence of a continuously varying set of forcings resulting from, e.g., the fluctuations in the incoming solar radiation and the processes—natural and anthropogenic—altering the atmospheric

composition; the lack of scale separation between different processes, which requires a profound revision of the standard methods for model reduction/projection to the slow manifold, and calls for the unavoidable need of complex parametrization of subgrid scale processes in numerical models; the impossibility to have detailed and homogeneous observations of the climatic fields with extremely high-resolution in time and in space, and the need to integrate direct and indirect measurements when trying to reconstruct the past climate state beyond the industrial era; the fact that we can observe only one realization of the process.

### **Pullback Attractor and Climate Response**

After a sufficiently long time, related to the slowest time scale of the system, at each instant the statistical properties of the ensemble of simulations do not depend anymore on the choice of the initial conditions.

A prominent example of this procedure is given by how simulations of past and historical climate conditions are performed in the modeling exercises such as those demanded by the IPCC [31, 32], where time-dependent climate forcings due to changes in greenhouse gases, volcanic eruptions, changes in the solar irradiance, and other astronomical effects are taken into account for defining the radiative forcing to the system.

In order to construct the time dependent measure following directly the definition of the pullback attractor, we need to construct a different ensemble of simulations for each choice of  $F(x, t)$ .

In other terms, from the knowledge of the time dependent measure of one specific pullback attractor, we can derive the time dependent measures of a family of pullback attractors.

### **A Climate Model of Intermediate Complexity: The Planet Simulator—PLASIM**

A detailed study of the impact of changing oceanic heat transports on the dynamics and thermodynamics of the atmosphere can be found in [33].

We remark that previous analyses have shown that using a spatial resolution approximately equivalent to T21 allows for obtaining an accurate representation of the major large scale features of the climate system.

While the lack of a dynamical ocean hinders the possibility of having a good representation of the climate variability on multidecadal or longer timescales, the climate simulated by PLASIM is definitely Earth-like, featuring qualitatively correct large scale features and turbulent atmospheric dynamics.

We are confident of the thermodynamic consistency of our model, which is crucial for evaluating correctly the climate response to radiative forcing resulting from changes in the opacity of the atmosphere.

### **Results**

One expects that coarse grained (in space) quantities will have a better signal-to-noise ratio and will allow for performing higher precision climate projection using response theory.

We will begin by looking into globally averaged quantities, and then address the problem of predicting the spatial patterns of climate change.

We dedicate some additional care in studying the climate response in terms of changes in the globally averaged surface temperature.

We would like to be able to assess when not only the projected change in the ensemble average is distinguishable from the statistics of the control run, but, rather, when an actual individual simulation is incompatible with the statistics of the unperturbed climate, because we live in one of such realizations, and not on any averaged quantity.

The methods of response theory allow us to treat seamlessly also the problem of predicting the climate response for (spatially) local observables.

### **A Critical Summary of the Results**

The performance of response theory in predicting the change in the globally averaged surface temperature and precipitation is rather good at all time horizons, with the predicted response falling within the ensemble variability of the direct simulations for all time horizons except for a minor discrepancy in the time window 40–60 y. Additionally, our results confirm the presence of a strong linear link in the form of modified Clausius–Clapeyron relation between changes in such quantities, as already discussed in the literature.

Response theory provides an excellent tool also for predicting the change in the zonal mean of the surface temperature, except for an underestimation of the warming in the very high latitude regions in the time horizons of 40–60 y. This is, in fact, the reason for the small bias found already when looking at the prediction of the globally averaged surface temperature.

### **Challenges and Future Perspectives**

The ab-initio construction of the linear response operator has proved elusive because of the difficulties associated with dealing effectively with both the unstable and stable directions in the tangent space.

What is extremely interesting about BVs is that (1) their growth factors are strongly dependent on the region of the phase space where the system is; and (2) the choice of the reference norm of the perturbation and of the time interval between two successive renormalization procedures (breeding period) effects strongly the properties of the dominant instabilities specifically active on the chosen time scales.

Of meteo-climatic relevance it has been shown that a relatively low number of BVs is extremely effective for reconstructing the properties of the unstable space, and that BVs contain useful information on spatially localized features, so that it may be worth trying to construct an approximation to the Ruelle response operator using the BVs.

Using such results in a reduced state space might provide a novel and effective method for approaching the problem of climate response.

### **Acknowledgement**

*A machine generated summary based on the work of Lucarini, Valerio; Ragone, Francesco; Lunkeit, Frank (2016 in Journal of Statistical Physics).*

## ***Beyond Forcing Scenarios: Predicting Climate Change Through Response Operators in a Coupled General Circulation Model***

<https://doi.org/10.1038/s41598-020-65297-2>

### **Abstract-Summary**

Global Climate Models are key tools for predicting the future response of the climate system to a variety of natural and anthropogenic forcings.

We show how to use statistical mechanics to construct operators able to flexibly predict climate change.

We perform our study using a fully coupled model—MPI-ESM v.1.2—and for the first time we prove the effectiveness of response theory in predicting future climate response to CO<sub>2</sub> increase on a vast range of temporal scales, from inter-annual to centennial, and for very diverse climatic variables.

The change in the Atlantic Meridional Overturning Circulation (AMOC) and of the Antarctic Circumpolar Current (ACC) is accurately predicted.

We are able to predict accurately the temperature change in the North Atlantic.

### **Introduction**

Global climate models (GCMs) are currently the most advanced tools for studying future climate change; their future projections are key ingredients of the reports of the Intergovernmental Panel on Climate Change (IPCC) and are key for climate negotiations [34].

For IPCC-class GCMs, future climate projections are usually constructed by defining a few climate forcing scenarios, given by changes in the composition of the atmosphere and in the land use, each corresponding to a different intensity and time modulation of the equivalent anthropogenic forcing.

No rigorous prescription exists for translating the climate change projections if one wants to consider different time modulations of a given forcing, e.g. a faster or slower CO<sub>2</sub> increase.

### **Response Theory and Climate Change**

The FDT has recently been key to inspiring the theory of emergent constraints, which are tools for reducing the uncertainties on climate change by looking at empirical relations between climate response and variability of some given observables [35, 36].

Response theory is a generalisation of the FDT that allows one to predict how the statistical properties of general—near or far from equilibrium, deterministic or stochastic—systems change as a result of forcings.

Encouragingly, response theory has recently been shown to have a great potential for predicting climate change in multi-model ensembles of CMIP5 atmosphere–ocean coupled GCMs outputs [37].

The response of a slow (oceanic) climatic observable of interest has been investigated so far in relation to the change in the dynamical properties of some other climatic observable, by constructing a linear regression between the predictand and predictor using the properties of the natural variability of the system [38–41].

### **Predicting Climate Change Using the Ruelle Response Theory**

We then focus on two key aspects of the large-scale ocean circulation, namely the Atlantic Meridional Overturning Circulation (AMOC) [42, 43] and the Antarctic Circumpolar Current (ACC) [44], and show that we can achieve excellent skill in predicting the the slow modes of the climate response.

In current conditions, the ocean is well-known to absorb a large fraction of the Earth’s energy imbalance due to global warming and to store it through its large thermal inertia, up to time scales defined by the deep ocean circulation [45].

### **Results**

Of the presence of slow oceanic time scales, the Green function significantly departs from a simple exponential relaxation behavior, which is sometimes adopted to describe the relaxation of the climate system to forcings [46, 47].

On short time scales, we have a reduction of AMOC, as a result of the negative value of the Green function.

On longer time scales ( $>100$  y), a negative feedback acts as a restoring mechanism, associated with a positive sign in the Green function.

On decadal scales, we have a loss in the correlation between wind stress and ACC, corresponding to the Green function turning negative after about 30 y. Beyond these time scales, we have time-wise coherent response of the AMOC and ACC, underlying the response of the global ocean circulation.

This has profound implications for setting the time scales of the ACC and AMOC response.

### **Discussion and Conclusions**

In all considered cases response theory successfully predicts the time-dependent change.

Ruelle’s response theory provides a relatively simple yet robust and powerful set of diagnostic and prognostic tools to study the response of climatic observables to external forcings.

The availability of a large number of ensemble members allows for constructing more accurate Green functions and for studying effectively the response of a broader class of climatic observables.

A promising application is the definition of functional relations between the response of different observables of a system to forcings, in the spirit of some recent investigations (see, e.g. Zappa and others [48]).

Being based on a perturbative approach, response theory (linear and nonlinear) has, by definition, only a limited range of applicability (e.g. one cannot use it to treat arbitrarily strong forcings).

## **Appendix A: Methods**

The susceptibility gives a spectroscopic description of the properties of the response of the observable, and its analysis can give interesting information on the most relevant time scales and related processes that determine the response of the observable.

From the Green function, one could in principle compute the susceptibility and perform a spectral analysis of the properties of the response.

This translates into the fact that, despite the Green function and the susceptibility being strictly connected, to obtain a satisfactory estimate for the latter require a statistics orders of magnitude larger than for the former, and possibly different and dedicated numerical estimation approaches [49–51].

An analysis of the susceptibility in experiments similar to what done in this work was attempted in Ragone and others [52], but using ten times more ensemble members.

While the analysis of the detailed frequency response of a climate model remains a very interesting and promising topic, it has to likely wait until experiments with at least several hundreds ensemble members will be available.

## **Acknowledgement**

*A machine generated summary based on the work of Lembo, Valerio; Lucarini, Valerio; Ragone, Francesco (2020 in Scientific Reports).*

## ***Improving Prediction Skill of Imperfect Turbulent Models Through Statistical Response and Information Theory***

<https://doi.org/10.1007/s00332-015-9274-5>

## **Abstract-Summary**

Statistical uncertainty quantification (UQ) to the response to the change in forcing or uncertain initial data in such complex turbulent systems requires the use of imperfect models due to the lack of both physical understanding and the overwhelming computational demands of Monte Carlo simulation with a large-dimensional phase space.

The systematic development of reduced low-order imperfect statistical models for UQ in turbulent dynamical systems is a grand challenge.

The forty mode Lorenz 96 (L-96) model which mimics forced baroclinic turbulence is utilized as a test bed for the calibration and predicting phases for the hierarchy of computationally cheap imperfect closure models both in the full phase space and in a reduced three-dimensional subspace containing the most energetic modes.

For reduced-order model for UQ in the three-dimensional subspace for L-96, the systematic low-order imperfect closure models coupled with the training strategy provide the highest predictive skill over other existing methods for general forced

response yet have simple design principles based on a statistical global energy equation.

The systematic imperfect closure models and the calibration strategies for UQ for the L-96 model serve as a new template for similar strategies for UQ with model error in vastly more complex realistic turbulent dynamical systems.

### **Introduction**

A conceptual framework intermediate between detailed dynamical physical modeling and purely statistical analysis based on empirical information theory has been proposed (Majda and Gershgorin [53, 54]; Gershgorin and Majda [55]) to address imperfect model fidelity and sensitivity problems.

In Majda and Gershgorin [56], a direct link by utilizing fluctuation–dissipation theorem (FDT) for complex systems together with the framework of empirical information theory for improving imperfect models is developed.

We investigate and develop systematic strategies for improving the imperfect model prediction skill for complex turbulent dynamical systems by employing ideas in both the information-theoretic framework and linear response theory mentioned above.

Following the direct link between the linear response and empirical information theory demonstrated in Majda and Gershgorin [56] for models with equilibrium fidelity, it is shown that they can be seamlessly combined into a precise systematic framework to improve imperfect model sensitivity through measuring the information error of the linear response operator in the training phase with unperturbed statistics.

### **Theories for Improving Imperfect Model Prediction Skill**

The information theory offers a least biased measure for quantifying the error between the imperfect model prediction and the truth; and the linear response theory gives an important tool relating the model responses to stationary state statistics of the dynamical system.

With the help of these theories (Majda and Gershgorin [53, 56]), one systematic process to tune model parameters in a training phase to possibly achieve the optimal model with sensitivity to all kinds of perturbations is discussed.

It is reasonable to claim that an imperfect model with precise prediction of this linear response operator should possess uniformly good sensitivity to different kinds of perturbations.

Considering all these good features of the linear response operator, information barrier due to model sensitivity to perturbations can be overcome by minimizing the information error in the imperfect model kicked response distribution relative to the true response from observation data (Majda and Gershgorin [56]).

### **L-96 System as a Test Bed and Its Statistical Dynamics**

To quantify these uncertainties, we are interested in resolving the statistical features of this dynamical system, especially the first two-order moments.

The equations will never be closed under this process by calculating the dynamical equations for each order moments.

Even though we derive the moment equations above from the L-96 system under homogeneous assumption for the sake of analysis and will focus on them in the following discussions, the moment equations are actually quite representative and are easy to be extended to general nonlinear systems with conservative quadratic forms.

We can focus on the simplified moment equations and investigate the statistical properties inside this system.

Pointwise statistics by only considering the variance at each grid point, and ignoring the correlations between different grids may not be sufficient for accurate model predictions.

This will end up with large information barrier when only one-point statistics are considered in the imperfect model.

### **Statistical Closure Methods in Full Phase Space**

The imperfect model prediction skill as well as the improvement through the information–response framework will be compared through checking the models' ability to capture the responses to several different types of perturbed external forcing terms.

The dynamical imperfect model using the closure method offers more precise prediction for the nonlinear responses for both the mean and variance.

Two rows, we show the model outputs for the mean and total variance with closure methods GC1, GC2, and MQG compared with the truth from Monte Carlo simulation.

The model prediction skill increases as more and more detailed calibration about the nonlinear flux are proposed from GC1 to GC2, MQG.

In the signal part, the error in the mean can be minimized to small amount for all three models under optimal parameter; while for dispersion part, MQG and GC2 have much better prediction for the prediction in variance compared with GC1.

### **Low-Order Models in a Reduced Subspace**

Keeping all these shortcomings in mind, we propose the following further corrections to the reduced-order methods and refer the resulting model as the corrected model.

We are interested in checking whether these correction strategies for the reduced methods can actually improve the model prediction skill.

In the full space case, we need to first tune the reduced model parameters in a training phase for optimal responses.

Observing the errors in signal and dispersion part separately for the original model in the second row, it can be found that large inherent information barrier (especially for the mean prediction) exists for improving the model prediction skill no matter how well we tune the parameters in the training phase.

GC2 can even offer better prediction in the reduced-order case than ROMQG model considering that it is also cheaper in computation.



### **Conclusion and Future Work**

Several important points can be concluded from the theoretical analysis and numerical tests using the L-96 test bed: The second-order statistical closure models outperform the linear FDT predictions for capturing responses to external perturbations, especially in regimes with larger perturbations and stronger nonlinearities.

This is an important result showing that higher-order moments can be determined by the lower-order approximations and offers important guideline for designing imperfect closure schemes; Still accurate single-point statistics prediction is not sufficient for the imperfect models to break information barriers (Majda and Gershgorin [53, 56]; Majda and Branicki [57]).

Imperfect model prediction skill can be improved uniformly regardless of the specific perturbation form applied; It is important for practical applications that the information–response framework can also be applied systematically to reduced-order models which focus on capturing the uncertainties in the dominant modes.

### **Acknowledgement**

*A machine generated summary based on the work of Majda, Andrew J.; Qi, Di (2015 in Journal of Nonlinear Science).*

## **Book Reading List**

### **An Introduction to the Theory of Climate**

By Monin, A.S. (1986).

During the last 20 years the study of, and the prediction of, changes in the climate of our planet have become an urgent social imperative, addressed to scientists the world over. The first principles on which to base such a study were formulated in 1974 in Stockholm, at the international GARP conference on the physical fundamentals of climate theory and climate modeling. In 1979 the World Meteorological Organization and the International Council of Scientific Unions decided to conduct a global program of climate research.

Please see <https://www.springer.com/gp/book/9789027719355> for original source.

### **Non-equilibrium Thermodynamics and the Production of Entropy**

By Kleidon, A. (Ed), Lorenz, R. D. (Ed) (2005).

The present volume studies the application of concepts from non-equilibrium thermodynamics to a variety of research topics. Emphasis is on the Maximum Entropy Production (MEP) principle and applications to Geosphere-Biosphere couplings. Written by leading researchers from a wide range of background, the book proposed to give a first coherent account of an emerging field at the interface of thermodynamics, geophysics and life sciences.

Please see <https://www.springer.com/gp/book/9783540224952> for original source.

### **Theory and Practice of Climate Adaptation**

By Alves, F. (Ed), Leal Filho, W. (Ed), Azeiteiro, U. (Ed) (2018).

Climate change is one of the greatest challenges of our time. As such, both the Fifth Assessment Report (AR5) released by the Intergovernmental Panel on Climate Change (IPCC) and the 25th Conference of the Parties (COP 25) recommendations call for action not only from government, but also from various stakeholders. Apart from the knowledge offered by modeling and forecasts, which allows the readers to understand the problem and how it is likely to develop in the future, the book highlights approaches, methods and tools that can help readers cope with the social, economic and political problems posed by climate change.

Please see <https://www.springer.com/gp/book/9783319728735> for original source.

### **Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics**

By Ghil, M., Childress, S. (1987).

The vigorous stirring of a cup of tea gives rise, as we all know, to interesting fluid dynamical phenomena, some of which are very hard to explain. In this book our “cup of tea” contains the currents of the Earth’s atmosphere, oceans, mantle, and fluid core. Our goal is to understand the basic physical processes which are most important in describing what we observe, directly or indirectly, in these complex systems. While in many respects our understanding is measured by the ability to predict, the focus here will be on relatively simple models which can aid our physical intuition by suggesting useful mathematical methods of investigation.

Please see <https://www.springer.com/gp/book/9780387964751> for original source.

## **References**

1. Rodwell MJ, Palmer TN (2007) Using numerical weather prediction to assess climate models. *Q J R Meteorol Soc* 133:129–146. <https://doi.org/10.1002/qj.23>
2. Ma HY, Xie S, Klein S, Williams K, Boyle J, Bony S, Douville H, Fermepein S, Medeiros B, Tyteca S (2014) On the correspondence between mean forecast errors and climate errors in CMIP5 models. *J Clim* 27(4):1781–1798
3. Karmalkar AV, Sexton David MH, James M, Ben Booth BB, Rostron John MD (2019) Finding plausible and diverse variants of a climate model: part 2 development and validation of methodology. <https://doi.org/10.1007/s00382-019-04617-3>
4. van den Hurk B, Siegmund P, Klein Tank A et al (2014) KNMI’ 14: climate change scenarios for the 21st century—A Netherlands perspective. Scientific report WR2014-01, KNMI, Bilt, Netherlands. <https://www.climate-scenarios.nl/15>. Accessed 1 Jan 2016
5. CSIRO and Bureau of Meteorology (2015) Climate change in Australia. <https://www.climatechangeinaustralia.gov.au/>. Accessed 1 Jan 2016
6. Meehl GA et al (2007) The WCRP CMIP3 multimodel dataset: a new era in climate change research. *Bull Am Meteorol Soc* 88:1383–1394
7. Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bull Am Meteorol Soc* 93(4):485–498

8. Scaife AA, Copsey D, Gordon C, Harris C, Hinton T, Keeley S, O'Neill A, Roberts M, Williams K (2011) Improved atlantic winter blocking in a climate model. *Geophys Res Lett* 38(23):L23703. <https://doi.org/10.1029/2011GL049573>
9. Hewitt H, Copsey D, Culverwell I, Harris C, Hill R, Keen A, McLaren A, Hunke E (2011) Design and implementation of the infrastructure of HadGEM3: The next-generation Met Office climate modelling system. *Geosci Model Dev* 4(2):223–253
10. Bony S, Colman R, Kattsov VM, Allan RP, Bretherton CS, Dufresne JL, Hall A, Hallegatte S, Holland MM, Ingram W (2006) How well do we understand and evaluate climate change feedback processes? *J Clim* 19(15):3445–3482
11. Webb MJ, Lambert FH, Gregory JM (2013) Origins of differences in climate sensitivity, forcing and feedback in climate models. *Clim Dyn* 40(3–4):677–707
12. Yokohata T, Webb MJ, Collins M, Williams KD, Yoshimori M, Hargreaves JD, Annan JD (2010) Structural similarities and differences in climate responses to CO<sub>2</sub> increase between two perturbed physics ensembles. *J Clim* 23:1392–1410. <https://doi.org/10.1175/2009JCLI2917.1>
13. McKay MD, Beckman RJ, Conover WJ (1979) Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21(2):239–245
14. Walters D, Williams K, Boutle I, Bushell A, Edwards J, Field P, Lock A, Morcrette C, Stratton R, Wilkinson J et al (2014) The Met Office Unified Model global atmosphere 4.0 and JULES global land 4.0 configurations. *Geosci Model Dev* 7(1):361–386
15. Gates WL, Boyle JS, Covey C, Dease CG, Doutriaux CM, Drach RS, Fiorino M, Gleckler PJ, Hnilo JJ, Marlais SM (1999) An overview of the results of the atmospheric model intercomparison project (AMIP I). *Bull Am Meteorol Soc* 80(1):29–55
16. Bony S, Webb M, Stevens B, Bretherton C, Klein S, Tselioudis G (2009) The cloud feedback model intercomparison project: summary of activities and recommendations for advancing assessments of cloud-climate feedbacks. CFMIP Doc
17. Rougier J (2007) Probabilistic inference for future climate using an ensemble of climate model evaluations. *Clim Change* 81:247–264
18. Collins M, Booth BBB, Harris GR, Murphy JM, Sexton DMH, Webb MJ (2006) Towards quantifying uncertainty in transient climate change. *Clim Dyn* 27:127–147
19. Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* 430:768–772
20. Stainforth DA, Aina T, Christensen C, Collins M, Frame DJ, Kettleborough JA, Knight S, Martin A, Murphy J, Piani C, Sexton D, Smith LA, Spicer RA, Thorpe AJ, Allen MR (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* 433:403–406
21. Webb MJ et al (2006) On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Clim Dyn* 27:17–38
22. Barnett DN, Brown SJ, Murphy JM, Sexton DMH, Webb MJ (2006) Quantifying uncertainty in changes in extreme event frequency in response to doubled CO<sub>2</sub> using a large ensemble of GCM simulations. *Clim Dyn* 26:489–511
23. O'Hagan A, Forster J (2004) Bayesian inference, vol 2b of Kendall's advanced theory of statistics, 2nd edn. Edward Arnold, London
24. Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z (2007) Global climate projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) *Climate change 2007: the physical science basis. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge
25. Rougier J, Sexton DMH, Murphy JM, Stainforth DA (2009) Analyzing the climate sensitivity of the HadSM3 climate model using ensembles from different but related experiments. *J Clim* 22:1327–1353

26. Harris GR, Sexton DMH, Booth BBB, Collins M, Murphy JM, Webb MJ (2006) Frequency distributions of transient regional climate change from perturbed physics ensembles of general circulation model simulations. *Clim Dyn* 27:357–375
27. Murphy JM, Booth BBB, Collins M, Harris GR, Sexton D, Webb MJ (2007) A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philos Trans R Soc Lond a* 365:1993–2028
28. Murphy JM, Sexton DMH, Jenkins G, Boorman P, Booth BBB, Brown K, Clark R, Collins M, Harris GR, Kendon E (2009) *Climate change projections*. ISBN 978-1-906360-02-3
29. Sanderson BM et al (2008) Constraints on model response to greenhouse gas forcing and the role of subgrid-scale processes. *J Clim* 21:2384–2400
30. Yokohata T, Webb MJ, Collins M, Williams KD, Yoshimori M, Hargreaves JC, Annan JD (2010) Structural similarities and differences in climate responses to CO<sub>2</sub> increase between two perturbed physics ensembles. *J Clim* 23(6):1392–1410
31. Houghton J (ed) (2001) *IPCC third assessment report: Working Group I report, “The Physical Science Basis.”* Cambridge University Press, Cambridge
32. Intergovernmental Panel on Climate Change: In: Stocker T et al (eds) *The physical science basis IPCC working group I contribution to AR5*. Cambridge University Press, Cambridge
33. Knietzsch M-A, Schröder A, Lucarini V, Lunkeit F (2015) The impact of oceanic heat transport on the atmospheric circulation. *Earth Syst Dyn* 6(2):591–615
34. Intergovernmental Panel on Climate Change (2013) *Climate change 2013: the physical science basis*. Press, Cambridge University, Cambridge Mass
35. Cox PM, Huntingford C, Williamson MS (2018) Emergent constraint on equilibrium climate sensitivity from global temperature variability. *Nature* 553(7688):319–322
36. Cox PM (2019) Emergent constraints on climate-carbon cycle feedbacks. *Curr Clim Change Rep* 5(4):275–281
37. Aengenheyster M, Feng QY, van der Ploeg F, Dijkstra HA (2018) The point of no return for climate action: effects of climate uncertainty and risk tolerance. *Earth Syst Dyn* 9(3):1085–1095
38. Pillar HR, Heimbach P, Johnson HL, Marshall DP (2016) Dynamical attribution of recent variability in Atlantic overturning. *J Clim* 29(9):3339–3352
39. Kostov Y et al (2017) Fast and slow responses of Southern Ocean sea surface temperature to SAM in coupled climate models. *Clim Dyn* 48(5):1595–1609
40. Johnson HL, Cornish SB, Kostov Y, Beer E, Lique C (2018) Arctic Ocean freshwater content and its decadal memory of sea-level pressure. *Geophys Res Lett* 45(10):4991–5001
41. Cornish SB, Kostov Y, Johnson HL, Lique C (2020) Response of Arctic freshwater to the arctic oscillation in coupled climate models. *J Clim* 33(7):2533–2555
42. Kuhlbrodt T et al (2007) On the driving processes of the Atlantic meridional overturning circulation. *Revi Geophys* 45(2):RG2001
43. Hirschi J et al (2003) A monitoring design for the Atlantic meridional overturning circulation. *Geophys Res Lett* 30:1413
44. Orsi AH, Whitworth T, Nowlin WD (1995) On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Res Part I: Oceanogr Res Pap* 42(5):641–673
45. von Schuckmann K et al (2016) An imperative to monitor Earth’s energy imbalance. *Nat Clim Chang* 6:138–144
46. Hasselmann K, Sausen R, Maier-Reimer E, Voss R (1993) On the cold start problem in transient simulations with coupled atmosphere-ocean models. *Clim Dyn* 9(2):53–61
47. Held IM et al (2010) Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J Clim* 23(9):2418–2427
48. Zappa G, Ceppi P, Shepherd TG (2020) Time-evolving sea-surface warming patterns modulate the climate change response of subtropical precipitation over land. *Proc Natl Acad Sci* 117(9):4539–4545
49. Lucarini V (2009) Evidence of dispersion relations for the nonlinear response of the Lorenz 63 system. *J Stat Phys* 134(2):381–400
50. Gritsun A, Lucarini V (2017) Fluctuations, response, and resonances in a simple atmospheric model. *Physica D* 349:62–76

51. Lucarini V, Sarno S (2011) A statistical mechanical approach for the computation of the climatic response to general forcings. *Nonlinear Process Geophys* 18(1):7–28
52. Ragone F, Lucarini V, Lunkeit F (2016) A new framework for climate sensitivity and prediction: a modelling perspective. *Clim Dyn* 46(5–6):1459–1471
53. Majda AJ, Gershgorin B (2011) Improving model fidelity and sensitivity for complex systems through empirical information theory. *Proc Natl Acad Sci USA* 108(25):10044–10049
54. Majda AJ, Gershgorin B (2010) Quantifying uncertainty in climate change science through empirical information theory. *Proc Natl Acad Sci USA* 107(34):14958–14963
55. Gershgorin B, Majda AJ (2012) Quantifying uncertainty for climate change and long-range forecasting scenarios with model errors. Part I: Gaussian models. *J Clim* 25(13):4523–4548
56. Majda AJ, Gershgorin B (2011) Link between statistical equilibrium fidelity and forecasting skill for complex systems with model error. *Proc Natl Acad Sci USA* 108(31):12599–12604
57. Majda AJ, Branicki M (2012) Lessons in uncertainty quantification for turbulent dynamical systems. *Discrete Cont Dyn Syst* 32(9):3133–3221