

Challenges in Climate Science and Contemporary Applied Mathematics

ANDREW J. MAJDA

Morse Professor of Arts and Sciences

Department of Mathematics and Center for Atmosphere and Ocean Science

Courant Institute of Mathematical Sciences

New York University, 251 Mercer St., New York, N.Y. 10012

Abstract

This article discusses the challenges in climate science from the emerging viewpoint of stochastic-statistical properties of turbulent dynamical systems. The mathematical topics discussed here include empirical information theory, fluctuation-dissipation theorems, reduced order stochastic modelling, and the development of mathematically unambiguous exactly solvable test models for climate science which capture crucial features of vastly more complex scientific problems. The applied mathematics topics include the emerging development of multi-scale algorithms for filtering/data assimilation and superparameterization for climate science and other problems in science and engineering, as well as suitable unambiguous mathematical test problems for their behavior. Interesting contemporary research directions and specific open problems are mentioned throughout the article. The perspective here also should be useful for applications to other complex dynamical systems arising in neural science, material science, and environmental/mechanical engineering. © 2000 Wiley Periodicals, Inc.

1 Introduction

The climate is an extremely complex coupled system involving significant physical processes for the atmosphere, ocean, and land over a wide range of spatial scales from millimeters to thousands of kilometers and time scales from minutes to decades or centuries [21, 111]. Climate change science focuses on predicting the coarse-grained planetary scale long time changes in the climate system due to either changes in external forcing or internal variability such as the impact of increased carbon dioxide [113]. For several decades the predictions of climate change science have been carried out with some skill through comprehensive computational atmospheric and oceanic simulation (AOS) models [21, 111, 113, 117] which are designed to mimic the complex physical spatio-temporal patterns in nature. Such AOS models either through lack of resolution due to current computing power or

through inadequate observation of nature necessarily parameterize the impact of many features of the climate system such as clouds, mesoscale and submesoscale ocean eddies, sea ice cover, etc. Thus, there are intrinsic model errors in the AOS models for the climate system and a central scientific issue is the effect of such model errors on predicting the coarse-grained large scale long time quantities of interest in climate change science. The central difficulty in climate change science is that the dynamical equations for the actual climate are unknown. All that is available from the true climate in nature are some coarse-grained observations of functions such as mean or variance of temperature, tracer greenhouse gases such as carbon dioxide, or the large scale horizontal winds. Thus, climate change science must cope with predicting the coarse-grained dynamic changes of an extremely complex system only partially observed from a suite of imperfect models for the climate. Basic questions arise such as the following:

- (A) How to measure the skill (i.e., the statistical accuracy) of a given model in reproducing the present climate and predicting the future climate in an unbiased fashion?
- (B) How to make the best possible estimate of climate sensitivity to changes in external or internal parameters by utilizing the imperfect knowledge available of the present climate? What are the most sensitive parameters for climate change given uncertain knowledge of the present climate?
- (C) How do coarse-grained measurements of different functionals of the present climate affect the assessments in (A, B)? What are the weights which should be assigned to different functionals of the present climate as targets to improve the performance of the imperfect AOS models? Which new functionals of the present climate should be observed in order to improve the assessments in (A), (B)?

Predicting how climate will change is one of the great societal and intellectual challenges of our time. Furthermore, predicting and understanding the seasonal, yearly, decadal and centennial impacts of climate change for issues ranging from extreme weather events to sea level rise to the evolving extent of deserts involves assessing the impacts of climate change on a variety of significant temporal and spatial scales. This is especially difficult because energy often flows intermittently from the smaller unresolved or marginally resolved scales in contemporary AOS models to impact much larger and longer spatio-temporal scales of motion of interest in the above problems [79]. While contemporary AOS models have high skill for the midlatitude upper troposphere, notable deficiencies in contemporary climate models involve assessing the multi-scale impacts of clouds in the tropics [108], sea ice and land ice in the polar regions [50], as well as the role of the observed mesoscale and submesoscale turbulence in the ocean [119]. During the past fifteen years,

the faculty members at the Courant Institute in the Center for Atmosphere and Ocean Science (CAOS) have been at the cutting edge in contributing to all these issues through the modus operandi of modern applied mathematics involving rigorous mathematical theory, quantitative and qualitative models for crucial issues in climate science, novel multi-scale asymptotic and numerical models, and data-driven methods to constrain complex models by observations and theories. Examples include the first theories and models for large-scale behavior of clouds in the tropics consistent with observations [91, 107, 63, 92, 84, 13, 61, 62], the development and use of ideas from empirical information theory to quantify and improve long range forecasting skill and climate sensitivity [67, 81, 68, 1, 47, 95, 36, 37, 104, 105], the first detailed observations and theory for the Greenland and West Antarctic ice sheet collapse [51, 58], and the first accurate assessment of the central role of moist entropy and low level moisture transports in determining subtropical and midlatitude climate circulations from observed data [114, 71, 115].

The goal of the present article is not to review all of the above interdisciplinary developments but instead to emphasize the author's personal view on the statistical-stochastic, multi-scale framework for large dimensional turbulent dynamical systems which is emerging at the present time and the central role it is likely to play for uncertainty quantification and sensitivity in climate change science in the near future. The mathematical toolkit utilized below includes empirical information theory, fluctuation-dissipation theorems and systematic physics-constrained, statistical-stochastic modelling for large-dimensional turbulent dynamical systems; the use of these ideas in climate change science is only beginning. The author hopes that this article inspires other mathematicians to contribute to these important emerging topics. While the bibliography of this paper is not comprehensive, many of the cited papers contain substantial additional references to the climate science literature.

2 Climate Change Science and the Statistical Dynamics of Complex Systems

While the actual equations governing climate dynamics on earth are unknown, it is natural to assume that these dynamics are Markovian, i.e., the future state depends only on the present state, on a suitably large space of (hidden) variables $\mathbf{v} \in \mathbb{R}^P$, $P \gg 1$. Thus, it is reasonable to assume that the perfect dynamical system for the climate is given by

$$(2.1) \quad \mathbf{v}_t = F(\mathbf{v}) + \sigma(\mathbf{v})\dot{W},$$

for $\mathbf{v} \in \mathbb{R}^P$ where σ is a $P \times K$ noise matrix and $\dot{W} \in \mathbb{R}^K$ is K -dimensional white noise. Already in (2.1), for simplicity in exposition, the important time varying effects of the seasonal cycle and diurnal cycle [97, 31] have been

ignored; furthermore, again for simplicity in exposition, no jump process contributions in (2.1) have been included.

The associated Fokker-Planck equation for the probability density $p(\mathbf{v}, t)$ is

$$(2.2) \quad p_t = -\operatorname{div}_{\mathbf{v}}(F(\mathbf{v})p) + \frac{1}{2}\operatorname{div}_{\mathbf{v}}\nabla_{\mathbf{v}}(Qp) \equiv L_{FP}p,$$

where $Q = \sigma\sigma^T$. To illustrate the reasonable nature of the Markovian assumption on a sufficiently large phase space, consider the micro-scale process in the atmosphere where cloud water forms on condensation nuclei from dust particles in the atmosphere in the presence of small scale turbulence; there is no doubt here that the dynamics is Markovian and microphysics models have been made but a detailed precise description of the dynamics is not available. This example motivates the crucial difficulty in climate change science; the detailed dynamics of the climate system are unknown and even the dimension of the phase space \mathbb{R}^P , $P \gg 1$ is unknown. In fact, all that is actually known about the present climate are certain coarse-grained statistical measurements of functionals

$$(2.3) \quad \mathbf{E}(\mathbf{u}), \quad \text{for } \mathbf{u} \in \mathbb{R}^N, \quad N \ll P,$$

for a training interval of time. These measurements encompass satellite data, weather station data, ocean buoys, ice cores, coral data, etc., where the extensive earth observing or training period has only occurred in the last fifty to one hundred years. The coarse-grained statistical measurements are quantities such as the mean and variance of temperature in the northern and southern hemisphere or over the continents, tracer gases in the atmosphere, like carbon dioxide, or geochemical tracers in the ocean. Successful predictions in climate change science are hampered by the fact that the actual dynamics in (2.1) is a turbulent large-dimensional system with positive Lyapunov exponents on essentially all spatio-temporal scales, as verified in our common experience with weather, storms, and gazing at the turbulent surface of the ocean. The use of statistical descriptions like (2.1), (2.2) for the climate system is not new and goes back to early predictability studies for simplified atmosphere models [73, 74, 75, 24].

2.1 The imperfect models

The imperfect models are naturally assumed to be given by a known dynamical system

$$(2.4) \quad (\mathbf{v}_M)_t = F_M(\mathbf{v}_M) + \sigma_M(\mathbf{v}_M)\dot{W}, \quad \mathbf{v}_M \in \mathbb{R}^M,$$

which has a similar structure as in (2.1) but the phase space for the imperfect model, \mathbb{R}^M , is often completely different from the natural system with usually $M \ll P$; the natural system in (2.1) and the imperfect model in (2.4) share in common the coarse-grained variables $\mathbf{u} \in \mathbb{R}^N$. The imperfect models in (2.4) range from comprehensive AOS climate models with billions

of variables [21, 111, 113, 117], to lower dimensional statistical-stochastic models for suitable low-frequency teleconnections ([90, 93] and references therein) to purely data driven regression models with varying degrees of statistical sophistication ([98, 99] and references therein). A simple example illustrating the fundamental difficulties in climate science in trying to use imperfect models like those in (2.4) to predict the sensitivity issues in (2.1) for the perfect model is presented next [104].

A typical situation with model error for complex systems arises when the true system has additional degrees of freedom that are hidden from the family of imperfect models utilized to study this system either through lack of scientific understanding or the practical lack of computational resolution. The simplest example with these features is to consider the true system as given by the two linear stochastic equations

$$(2.5) \quad \begin{aligned} \frac{du}{dt} &= au + v + F, \\ \frac{dv}{dt} &= qu + Av + \sigma \dot{W}, \end{aligned}$$

where \dot{W} is white noise; the system of equation in (2.5) has a smooth Gaussian statistical steady state provided that

$$(2.6) \quad a + A < 0, \quad aA - q > 0.$$

Assume that the variable v in (2.5) is hidden from the modeling process where all imperfect models are given by the scalar stochastic equation

$$(2.7) \quad \frac{du_M}{dt} = -\gamma_M u_M + F_M + \sigma_M \dot{W}_M.$$

The natural requirement $\gamma_M > 0$ is needed for (2.7) to have a Gaussian statistical steady state. Now consider the situation where the model in (2.7) has been tuned to match the single time statistics for u in (2.5) with perfect fidelity by matching the mean and variance of u_M with u ; elementary calculations show this is true for a one parameter family of models parameterized by $\gamma_M > 0$ provided that F_M, σ_M^2 satisfy the equilibrium mean and variance equations

$$(2.8) \quad \frac{F_M}{\gamma_M} = -\frac{AF}{aA - q}, \quad \frac{\sigma_M^2}{2\gamma_M} = -\frac{\sigma^2}{2(a + A)(aA - q)} \equiv E.$$

Thus, the conditions in (2.8) for F_M and σ_M guarantee perfect model fidelity for any $\gamma_M > 0$. In many practical situations such as actual experiments or climate science, it is important to understand the response of the natural system to external forcing, δF , and to hope that the response of the imperfect model captures the features of this response. The natural system response for (2.5) occurs by replacing F in (2.5) by $F + \delta F$ while the same experiment in the model for (2.7) involves replacing F_M by $F_M + \delta F$. For

both the natural system in (2.5) and the model system in (2.7), the only change in the equilibrium response is through the change in mean

$$(2.9) \quad \delta u = -\frac{A}{aA - q} \delta F, \quad \delta u_M = \frac{1}{\gamma_M} \delta F,$$

while the variance of u for the perfect model and u_M for the imperfect model stays constant at the same value E determined through the second equality in (2.8). Now assume that the natural system satisfies the stability conditions in (2.6) with $A > 0$. We claim that no model from (2.7), even with perfect fidelity in (2.8) for any $\gamma_M > 0$, can match the sensitivity of the natural system correctly; this is easy to see from (2.9) since for $A > 0$, $\text{sign}(\delta u) = -\text{sign}(\delta F)$ but for all models from (2.7), $\text{sign}(\delta u_M) = \text{sign}(\delta F)$ and the perfect and model sensitivity are always anti-correlated! Thus, even though the climate models satisfying (2.7), (2.8) are tuned to exactly match the true climate, these imperfect models are intrinsically deficient in calculating the crucial climate sensitivity for $A > 0$.

2.2 Systematically improving Climate Models Through Empirical Information Theory

With a subset of variables $\mathbf{u} \in \mathbb{R}^N$ and a family of measurement functionals $\mathbf{E}_L(\mathbf{u}) = (E_j(\mathbf{u}))$, $1 \leq j \leq L$, for the perfect system, empirical information theory [57, 86] builds the least biased probability measure $\pi_L(\mathbf{u})$ consistent with the L measurements of the present climate, $\bar{\mathbf{E}}_L$. There is a unique functional on probability densities [57, 86] to measure this given by the entropy

$$(2.10) \quad \mathcal{S} = - \int \pi \log \pi,$$

and $\pi_L(\mathbf{u})$ is the unique probability so that $\mathcal{S}(\pi_L(\mathbf{u}))$ has the largest value among those probability densities consistent with the measured information, $\bar{\mathbf{E}}_L$. All integrals as in (2.10) are over the phase space \mathbb{R}^N unless otherwise noted. For example, measurements of the mean and second moments of the perfect system necessarily lead to a Gaussian approximation [81, 86] to the perfect system from measurements, $\pi_L(\mathbf{u}) = \pi_G(\mathbf{u})$. Any model of the perfect system produces a probability density, $\pi^M(\mathbf{u})$. The natural way [70, 86] to measure the lack of information in one probability density, $q(\mathbf{u})$, compared with the true probability density, $p(\mathbf{u})$, is through the relative entropy, $\mathcal{P}(p, q)$, given by

$$(2.11) \quad \mathcal{P}(p, q) = \int p \log \left(\frac{p}{q} \right).$$

This asymmetric functional on probability densities, $\mathcal{P}(p, q)$, has two attractive features [70, 81, 86] as a metric for model fidelity : (i) $\mathcal{P}(p, q) \geq 0$ with equality if and only if $p = q$; (ii) $\mathcal{P}(p, q)$ is invariant under general nonlinear

changes of variables. The first issue to contend with is the fact that $\pi_L(\mathbf{u})$ is not the actual perfect model density but only reflects the best unbiased estimate of the perfect model given the L measurements, $\bar{\mathbf{E}}_L$. Let $\pi(\mathbf{u})$ denote the probability density of the perfect model, which is not actually known. Nevertheless, $\mathcal{P}(\pi, \pi_L)$ precisely quantifies the intrinsic error in using the L measurements of the perfect model, $\bar{\mathbf{E}}_L$. Consider an imperfect model with its associated probability density, $\pi^M(\mathbf{u})$; then the intrinsic model error in the climate statistics is given by $\mathcal{P}(\pi, \pi^M)$. In practice, $\pi^M(\mathbf{u})$ is determined by no more information than that available in the perfect model.

Consider a class of imperfect models, \mathcal{M} . The best imperfect model for the coarse-grained variable \mathbf{u} is the $M_* \in \mathcal{M}$ so that the perfect model has the smallest additional information beyond the imperfect model distribution $\pi^{M_*}(\mathbf{u})$, i.e.,

$$(2.12) \quad \mathcal{P}(\pi, \pi^{M_*}) = \min_{M \in \mathcal{M}} \mathcal{P}(\pi, \pi^M).$$

Also, actual improvements in a given imperfect model with distribution $\pi^M(\mathbf{u})$ resulting in a new $\pi_{post}^M(\mathbf{u})$ should result in improved information for the perfect model, so that $\mathcal{P}(\pi, \pi_{post}^M) \leq \mathcal{P}(\pi, \pi^M)$. Otherwise, objectively, the model has not been improved compared with the original perfect model. The following general principle [95, 85] facilitates the practical calculation of (2.12)

$$(2.13) \quad \begin{aligned} \mathcal{P}(\pi, \pi_{L'}^M) &= \mathcal{P}(\pi, \pi_L) + \mathcal{P}(\pi_L, \pi_{L'}^M) \\ &= (\mathcal{S}(\pi_L) - \mathcal{S}(\pi)) + \mathcal{P}(\pi_L, \pi_{L'}^M) \quad \text{for } L' \leq L. \end{aligned}$$

The entropy difference, $\mathcal{S}(\pi_L) - \mathcal{S}(\pi)$ in (2.13) precisely measures an intrinsic error from the L measurements of the perfect system. With (2.13) and a fixed family of L measurements of the actual climate, the optimization principle in (2.12) can be computed explicitly by replacing the unknown density π by the hypothetically known π_L in these formulas so that, for example, π^{M_*} is calculated by

$$(2.14) \quad \mathcal{P}(\pi_L, \pi_{L'}^{M_*}) = \min_{M \in \mathcal{M}} \mathcal{P}(\pi_L, \pi_{L'}^M).$$

The most practical setup for applying the framework of empirical information theory developed above arises when both the perfect system measurements and the model measurements involve only the mean and covariance of the variables \mathbf{u} so that π_L is Gaussian with climate mean $\bar{\mathbf{u}}$ and covariance R while π^M is Gaussian with model mean $\bar{\mathbf{u}}_M$ and covariance R_M . In this

case, $\mathcal{P}(\pi_L, \pi^M)$ has the explicit formula [86, 67]

$$(2.15) \quad \mathcal{P}(\pi_L, \pi^M) = \left[\frac{1}{2}(\bar{\mathbf{u}} - \bar{\mathbf{u}}_M)^*(R_M)^{-1}(\bar{\mathbf{u}} - \bar{\mathbf{u}}_M) \right] \\ + \left[-\frac{1}{2} \log \det(RR_M^{-1}) + \frac{1}{2}(\text{tr}(RR_M^{-1}) - N) \right].$$

Note that the first term in brackets in (2.15) is the signal, reflecting the model error in the mean but weighted by the inverse of the model covariance, R_M^{-1} while the second term in brackets, the dispersion, involves only the model error covariance ratio, RR_M^{-1} . The intrinsic metric in (2.15) is invariant under any (linear) change of variables which maps Gaussian distributions to Gaussians and the signal and dispersion terms are individually invariant under these transformations; this property is very important.

As a simple illustration of these concepts, let's assume the elementary perfect and imperfect climate models discussed in (2.5) and (2.7) above, where as shown below, empirical information theory reveals an intrinsic barrier for the imperfect models to prediction of the sensitivity for $A > 0$. The formula in (2.15) applies exactly to these models with perfect fidelity with

$$(2.16) \quad \mathcal{P}(\pi_\delta, \pi_\delta^M) = \frac{1}{2}E^{-1} \left| -\frac{A}{aA - q} - \frac{1}{\gamma_M} \right|^2 |\delta F|^2.$$

In this situation with $A > 0$, the attempt to minimize the information theoretic model error in the sensitivity through the general principle in (2.12) is futile because no finite minimum over γ_M of (2.16) is achieved and necessarily $\gamma_M \rightarrow \infty$ in the approach to this minimum value; in other words, there is an intrinsic barrier to skill in sensitivity which cannot be overcome with the imperfect models in (2.7) even though they satisfy perfect model fidelity in (2.8). In this situation, information theory predicts that one needs to enlarge the class of models beyond (2.7) by introducing more degrees of freedom in the model. On the other hand, if the natural system satisfies (2.6) with $A < 0$, then using (2.16) to minimize the lack of information in the sensitivity in the models which satisfy perfect fidelity in (2.8) results in the unique model with

$$(2.17) \quad \gamma_M^* = -A^{-1}(aA - q), \quad A < 0,$$

and this model captures both the model fidelity and model sensitivity to this forcing parameter exactly.

The relative entropy in (2.11) occupies a central role in statistics [10, 11, 128] and large deviation theory in the limit of large sample sizes [122, 123]. The empirical point of view presented here is useful for developing unbiased empirical statistical/physics-based models and has been utilized to predict the location and structure of Jupiter's Red Spot from observations of the

Galileo mission, as well as the behavior of large-scale quantities in statistical fluid dynamics [86]. Kleeman [67] first applied these ideas to the prediction skill for long range forecasting in perfect models and these concepts have been developed extensively by scientists in CAOS at CIMS with many applications and associated theory [69, 81, 68, 1, 47] in the context of perfect models. Recent research utilizing empirical information theory has focussed on important coarse-grained descriptions of perfect and imperfect models and improving the long range forecasting and sensitivity of imperfect models [95, 36, 37, 104, 105]. This is an active and important area blending concepts from mathematics, statistics and physics and an exciting area for future research developments. One of the current directions involves utilizing the fluctuation-dissipation theorem (FDT) for (2.1) and (2.4) which is briefly discussed next.

2.3 Fluctuation Dissipation Theorems for Turbulent Dynamical Systems and Climate Change Science

The fluctuation-dissipation theorem is one of the cornerstones of the statistical physics of identical molecules of gases and liquids [109]. In a very brief seminal article from 1975, Leith [72] suggested that if FDT can be established for suitable coarse-grained functionals in climate science, then climate change assessments can be performed simply by gathering suitable statistics in the present climate. Here is a brief summary of FDT for the stochastic dynamical system in (2.1) [85, 97].

The ideal equilibrium state associated with (2.1) is the probability density $p_{eq}(\mathbf{v})$ that satisfies $L_{FP} p_{eq} = 0$ and the equilibrium statistics of some functional $A(\mathbf{v})$ are determined by

$$(2.18) \quad \langle A(\mathbf{v}) \rangle = \int A(\mathbf{v}) p_{eq}(\mathbf{v}) d\mathbf{v}.$$

Next, perturb the system in (2.11) by the change $\delta\mathbf{w}(\mathbf{v})f(t)$, that is, consider the perturbed equation

$$(2.19) \quad \mathbf{v}_t^\delta = F(\mathbf{v}^\delta) + \delta\mathbf{w}(\mathbf{v})f(t) + \sigma(\mathbf{v}^\delta)\dot{W}.$$

Calculate perturbed statistics by utilizing the Fokker-Planck equation associated with (2.19) with initial data given by the unperturbed statistical equilibrium. Then, FDT [85] states that if δ is small enough, the leading order correction to the statistics in (2.18) becomes

$$(2.20) \quad \delta\langle A(\mathbf{v}) \rangle(t) = \int_0^t \mathcal{R}(t-s)\delta f(s)ds,$$

where $\mathcal{R}(t)$ is the linear response operator that is calculated through correlation functions in the unperturbed climate

$$(2.21) \quad \mathcal{R}(t) = \langle A(\mathbf{v}(t))B(\mathbf{v}(0)) \rangle, \quad B(\mathbf{v}) = -\frac{\operatorname{div}_{\mathbf{v}}(\mathbf{w} p_{eq})}{p_{eq}}.$$

The noise in (2.1) is not needed for FDT to be valid but, in this form, the equilibrium measure needs to be smooth. Such a rigorous FDT response is known to be valid for a wide range of dynamical systems under minimal hypothesis [45].

There are important practical and computational advantages for climate change science when a skillful FDT algorithm is established. The FDT response operator can be utilized directly for multiple climate change scenarios, multiple changes in forcing, and other parameters, such as damping and inverse modelling directly [42, 43], without the need of running the complex climate model in each individual case. Note that FDT is a type of dynamic statistical linearization and does not involve linearizing the underlying nonlinear dynamics. The direct application of FDT to the natural perfect model in (2.1) is hampered by the fact that the dynamics in (2.1), the equilibrium measure in (2.18), and even the dimension of the phase space in (2.1) and (2.18) are unknown. Recently, an important link [105] was established through empirical information theory and FDT between the skill of specific prediction experiments in the training phase for the imperfect model when the climate is observed and the skill of the model for long range perturbed climate sensitivity.

There is a growing literature in developing theory [85, 97, 98, 102, 31] and algorithms for FDT [72, 8, 16, 44, 42, 43, 2, 3, 4, 5, 6] for forced dissipative turbulent systems far from equilibrium. In fact, the earliest algorithms which tested the original suggestion of Leith [72] utilized kicked perturbations without model error to evaluate the response operator [8, 16] and these algorithms have been improved recently [3, 5]; their main limitation is that they can diverge at finite times when there are positive Lyapunov exponents [16, 3, 5]. Alternative algorithms utilize the quasi-Gaussian approximation [85] in the formulas in (2.21); these algorithms have been demonstrated to have high skill in both mean and variance response in the midlatitude upper troposphere to tropical forcing [42, 43] as well as for a variety of other large dimensional turbulent dynamical systems which are strongly mixing [85, 2, 4]. There are recent blended response algorithms which combine the attractive features of both approaches and give very high skill for both the mean and variance response for the L-96 model [76, 2] as well as suitable large dimensional models of the atmosphere [4] and ocean [6] in a variety of weakly and strongly chaotic regimes. Finally, there are linear regression models [116] which try to calculate the mean and variance response directly from data; these linear regression models can have very good skill in the

mean response but necessarily have no skill [98] in the variance response; they necessarily have an intrinsic barrier [104, 105, 95] for skill in model response when the perfect model has a large variance response. In fact, one can regard all of the above approximations as defining various systems with model error in calculating the ideal response of a perfect model [85]; this is a useful exercise for understanding the information theoretic framework for model error and response proposed recently [105] and examples are presented there.

2.4 Statistically Exactly Solvable Test Models Capturing Crucial Features of Climate Change Science

An important role of mathematics in applied sciences is to develop simpler exactly or easily solvable test models with unambiguous mathematical features which nevertheless capture crucial features of vastly more complex systems in science and engineering. Such models provide firm underpinning for advancing scientific understanding and developing new numerical or statistical algorithms. With all of the difficult issues in climate science mentioned in the present article, such unambiguous test models assume a crucial role. Here, two such models are briefly described.

First, introduce a family of test models for climate change science which have direct qualitative relevance for actual observed features for tracers in the atmosphere [112, 12] with the additional attractive feature of exactly solvable statistics for the mean and covariance [31, 35, 106] with many degrees of freedom despite the inherent statistical nonlinearity. Thus, they are physically relevant unambiguous test models for uncertainty in climate change science [95, 104, 105]. The models have a zonal (east-west) mean jet, $U(t)$, a family of planetary and synoptic scale waves with north-south velocity $v(x, t)$ with x , a spatially periodic variable representing a fixed mid-latitude circle in the east-west direction, and tracer gas $T(x, t)$ with a north-south environmental mean gradient α and molecular diffusivity κ [106, 12]. The dynamical equations for these variables are

$$\begin{aligned}
 (2.22) \quad A) \quad & \frac{dU}{dt} = -\gamma U + f(t) + \sigma \dot{W}, \\
 B) \quad & \frac{dv}{dt} = P \left(\frac{\partial}{\partial x} \right) v + \sigma_v(x) \dot{W}_v + f_v(x, t), \\
 C) \quad & \frac{\partial T}{\partial t} + U(t) \frac{\partial T}{\partial x} = -\alpha v(x, t) + \kappa \frac{\partial^2 T}{\partial x^2} - d_T T.
 \end{aligned}$$

The functions $f(t)$, $f_v(x, t)$ are known time-periodic functions with period of 1 yr reflecting the changing external forcing of the seasonal cycle, while \dot{W} , \dot{W}_v , represent random white noise fluctuations in forcing arising from hidden nonlinear interactions and other processes [96, 88]. The equation in

(2.22B) for the turbulent planetary waves is solved by Fourier series with independent scalar complex variable versions of the equation in (2.22A) for each different wave number k [96, 88]; in Fourier space the operator \hat{P}_k has the form $\hat{P}_k = -\gamma_k + i\omega_k$ with frequency $\omega_k = \beta k / (k^2 + F_s)$ corresponding to the dispersion relation of baroclinic Rossby waves and dissipation $\gamma_k = \nu(k^2 + F_s)$ where β is the north-south gradient of rotation, F_s is the stratification, and ν is a damping coefficient; the white noise forcing for (2.22)B is chosen to vary with each spatial wave number k to generate an equipartition energy spectrum for planetary scale wave numbers $1 \leq |k| \leq 10$ and a $|k|^{-5/3}$ turbulent cascade spectrum for $11 \leq |k| \leq 52$ (see [96, 88]). Any other turbulent energy spectrum can be imposed on v . The zonal jet $U(t) = \bar{U}(t) + U'(t)$, where $\bar{U}(t)$ is the climatological periodic mean with γ , and σ chosen so that this jet is strongly eastward while the random fluctuations, $U'(t)$, have a standard deviation consistent with such eastward dynamical behavior. While $U(t), v(x, t)$ have exactly solvable Gaussian statistics mimicking features of the atmosphere, the tracer $T(x, t)$ has non-Gaussian behavior due to the nonlinear tracer flux term $U'(t) \frac{\partial T}{\partial x}$ in (2.22C) with intermittent fat tails like realistic tracers in the atmosphere [106, 112]; nevertheless, $T(x, t)$ has exactly solvable mean and covariance climate statistics following [31, 30, 32] with explicit formulas. These procedures define the exactly solvable statistics for the perfect climate. Actual AOS models utilized in climate change science typically have too much additional damping and one can mimic this here in the representative AOS models by increasing the two parameters γ, v for (2.22A,B) to γ_M, v_M to define the AOS model velocity fields $U(t)_M = \bar{U}_M(t) + U'_M(t), v_M(x, t)$, with model error. The turbulent tracer in an AOS model is usually calculated roughly by an eddy diffusivity [21, 111, 113, 117], $\overline{U'_M(t) \frac{\partial T}{\partial x}} = -\kappa_M^* T_{xx}$, and in the present models there is an exact explicit formula for κ_M^* . Thus, the AOS model tracer satisfies

$$(2.23) \quad \frac{\partial T_M}{\partial t} + \bar{U}_M(t) \frac{\partial T_M}{\partial x} = -\alpha v_M(x, t) + (\kappa + \kappa_M^*) \frac{\partial^2 T_M}{\partial x^2} - d_T T_M + \sigma_T \dot{W}(x, t),$$

where $\dot{W}(x, t)$ denotes space-time white noise forcing with variance σ_T to overcome deterministic model error. With (2.23) the AOS model with (U_M, v_M, T_M) has Gaussian statistics.

Note that the above perfect and imperfect climate models do not have positive Lyapunov exponents but nevertheless exhibit non-normal transient growth through the non-zero mean gradient, $\alpha > 0$, for the tracer. These models have been utilized as unambiguous test models for all the issues of climate change science, information theory, prediction, and FDT described earlier in this section [95, 104, 105]. These are also important test models for the real-time recovery of turbulent tracer fields from partial observations, an

important topic with much practical interest in climate science, as well as other disciplines [35]. A complete development of the turbulent statistics of such test models is presented in [106]. Similar exactly solvable test models with intermittent positive Lyapunov exponents are developed elsewhere [33, 34, 15] and mentioned briefly in section 3 in the context of filtering.

There is recent interest in deriving reduced stochastic models for climate and extended-range weather prediction. An attractive property of atmospheric low-frequency variability is that it can be efficiently described by just a few large-scale teleconnection patterns (see [26, 9, 14], and references therein). These patterns exert a huge impact on surface climate and seasonal predictability. Reduced stochastic models are an attractive alternative for climate sensitivity studies via FDT [98] because they are computationally much more efficient than state-of-the-art climate models and often have been shown to have comparable long-range prediction skill [28, 87, 37]. Systematic mathematical stochastic-mode reduction strategies [90, 82, 83] have been utilized recently to develop normal forms for reduced, stochastic climate models [93]. The one-dimensional, normal form was applied in a regression strategy in [93] for data from a prototype AOS model [26] to build one-dimensional stochastic models for low-frequency patterns such as the North Atlantic Oscillation (NAO) and the leading principal component (PC-1) that has features of the Arctic Oscillation. These one-dimensional, normal form stochastic models have been utilized to show the high skill of FDT algorithms despite deterministic, structural instability to both changes in external forcing and dissipation parameters as well as test models for climate sensitivity and model error via information theory [102, 95]. The canonical, one-dimensional stochastic models for low-frequency variability [93] are given by the scalar stochastic equation

$$(2.24) \quad dx = [F + ax + bx^2 - cx^3]dt + (A - Bx)dW + \sigma dW_A,$$

with corresponding Fokker-Planck equation

$$(2.25) \quad \frac{\partial p}{\partial t} = -\frac{\partial}{\partial x} [(F + ax + bx^2 - cx^3)p] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [((Bx - A)^2 + \sigma^2)p].$$

As calculated elsewhere [93], the Fokker-Planck equation in (2.25) has an explicit, smooth equilibrium distribution $p_{eq}(x)$ with a Gaussian tail provided that the physically imposed restriction, $c > 0$, is satisfied. The explicit form of the PDF, $p_{eq}(x)$, allows one to calculate explicit forms of the ideal response operator to perturbations in forcing, F , or the dissipation parameter, a , as well as explicit analytic expressions for the FDT linear response operator and various approximations [102]. One of the striking features of atmospheric general circulation models is that there are different regimes of low-frequency behavior despite uni-modal, nearly-Gaussian PDFs for the

low-frequency variables [9, 14, 27, 28, 87]. The stochastic models are studied in parameter regimes where this behavior occurs and in the vicinity of where there is deterministic, structural instability. Figure 1 from [102] illustrates how intermittent regimes can happen in a stochastic dynamical system with nearly unimodal PDF's and (2.24), (2.25) is the simplest model which illustrates this phenomenon, which is very different from the usual (often incorrect!) association of regimes necessarily with multiple peaks in the PDF's. The information theoretic perspective on model error and long range prediction has been applied to these stochastic models recently [95, 38] as an elementary unambiguous test problem.

3 Multi-scale Algorithms for Turbulent Dynamical Systems: Superparameterization, Filtering or Data Assimilation, and Judicious Model Errors

The complexity of anisotropic turbulent processes over a wide range of spatial and temporal scales in atmospheric and oceanic flows requires novel computational strategies, even with the current and next generations of supercomputers. This is especially important since energy often flows intermittently from the smaller unresolved or marginally resolved scales to affect the largest observed scales in such anisotropic turbulent flows due to the effects of rotation, stratification and moist processes [79]. Atmospheric weather and climate processes cover about 10 decades of spatial scales, from a fraction of a millimeter to planetary scales. A similarly staggering range of interconnected scales characterizes the oceanic circulation. While the smaller scales (of order millimeters to tens of meters) are comparatively less complex, as they fall within the inertial range of turbulence, scales above this range and up to the planetary scales, are dominated by an array of intermittent and anisotropic turbulent processes that cannot be described by traditional closures. For example, atmospheric motions on scales between 100 m and 100 km show an abundance of processes associated with dry and moist convection, clouds, waves, boundary layer, topographic, and frontal circulations. Oceanic scales from 10 meters to 100 km display a similar range of behaviors, albeit without phase transitions, but with a two-component density reflecting temperature and salt variations. On the atmospheric side, a major stumbling block in the accurate prediction of weather and short term climate on the planetary and synoptic scales is the accurate parameterization of moist convection. Moist convective processes involve intermittency in space and time due to complex evolving chaotic and quiescent regions, without statistical equilibration and with only moderate scale separation, so that traditional turbulence closure modeling fails [49, 110, 120]. Cloud-resolving models realistically represent convective-scale and mesoscale processes with fine computational grids. However, due to their extremely high

computational cost, they cannot be applied to large ensemble-size weather prediction or climate simulations. This state of affairs, unfortunately, will remain for the foreseeable future. In ocean models used for coupled climate simulations, the situation is arguably even worse. Here, the computational grid is typically on order 100 km, near the spectral peak of the oceans kinetic energy (which is dominated by baroclinic eddies somewhat larger than the deformation scale). Eddy-permitting simulations for ocean-only process studies are now becoming common [48], but even these leave a vast frontier of scales, from order 50 km down to the 10 m scale where inertial range turbulence finally takes over, almost completely unaddressed.

The complexity of these problems motivates the development of novel approaches that would directly address the multi-scale nature of the problem. In atmospheric modeling, superparameterization (SP), or more specifically to its initial application, cloud-resolving convection parameterization, [39, 40, 118, 41, 129] uses a spatially periodic 2-dimensional cloud-system-resolving model in each column of a large-scale model to explicitly represent small-scale and mesoscale processes, and interactions among them. In this context, SP blends conventional parameterization on a coarse mesh with detailed cloud-resolving modeling on a finer mesh. This approach has been shown to be ideal for parallel computations on supercomputers and has yielded promising new results regarding tropical intraseasonal behavior [39, 40, 41, 65]. The SP approach to convective parameterization in the atmosphere is powerful and invites application of SP to a broader array of problems in climate-atmosphere-ocean science such as mesoscale and sub-mesoscale eddies in the ocean and gravity wave drag in the stratosphere, as well as other science and engineering problems. However, that particular approach is difficult to replicate because of the ad-hoc nature of its development. Recently, however, the author and collaborators have shown how multi-scale models may be exploited to enable rigorous, systematic development of SP schemes [89, 129]. Moreover, a general statistical numerical analysis framework has been introduced recently [94] which illustrates why such methods can successfully model systems with only modest scale separation and without statistical equilibration of the small-scale dynamics [94].

Filtering or data assimilation is the process of obtaining the best statistical estimate of a natural system from partial observations of the true signal from nature. In many contemporary applications in science and engineering, real time filtering of a turbulent signal from nature involving many degrees of freedom is needed to make accurate predictions of the future state. This is obviously a problem with significant practical impact. Important contemporary examples involve the real time filtering and prediction of weather and climate as well as the spread of hazardous plumes and pollutants or the prediction of storm surges in environmental science and engineering. Thus, an

important emerging scientific issue is the real time filtering through observations of noisy signals for turbulent nonlinear dynamical systems as well as the statistical accuracy of spatio-temporal discretizations for filtering such systems. See the recent review article [96] and the many references therein, as well as the introductory graduate text book [103]. From the practical standpoint, the demand for operationally practical filtering methods escalates as the model resolution is significantly increased. In the coupled atmosphere-ocean system, the current practical models for prediction of both weather and climate involve general circulation models where the physical equations for these extremely complex flows are discretized in space and time and the effects of unresolved processes are parametrized according to various recipes; the result of this process involves a model for the prediction of weather and climate from partial observations of an extremely unstable, chaotic dynamical system with several billion degrees of freedom. These problems typically have many spatiotemporal scales, rough turbulent energy spectra in the solutions near the mesh scale, and a very large dimensional state space yet real time predictions are needed. There is an inherently difficult practical issue of small ensemble size in filtering statistical solutions of these complex problems due to the large computational overload in generating individual ensemble members through the forward dynamical operator.

The above discussion motivates the need for systematic mathematical ideas in devising algorithms for superparameterization (SP) and filtering/data assimilation (FDA) for large-dimensional turbulent dynamical systems, as well as new types of statistical/stochastic numerical analysis to assess the performance skill of various proposed algorithms. Thus, there is a natural link between the viewpoint developed here and the earlier discussion in section 2. The recent review paper [96] and graduate text [103] contain much more detailed material for the interested reader on this emerging viewpoint for FDA for turbulent dynamical systems.

The general mathematical approach to SP or FDA for large-dimensional turbulent dynamical systems advocated here is a four-stage process:

- (3.1)
- (i) Multi-scale formulation: A multi-scale physical/mathematical formulation into larger scale mean and smaller scale fluctuating components in space-time (for examples, see [129, 89, 101, 66] and references therein).
 - (ii) Small-scale model: A mathematical model to represent the behavior of the smaller scales, typically involving a spatial periodic approximation and an imposed scale-gap (see [129, 94]).
 - (iii) Computational strategies to reduce the cost of the small scale models by making judicious model errors: Mathematical algorithms that allow for computationally efficient but statistically accurate implementation of the small-scale model as a SP or FDA algorithm in a larger scale model while retaining statistical accuracy [129, 89]. This can be implemented by replacing more expensive three-dimensional models by much simpler two-dimensional [39, 40] and even cheaper stochastic models [60, 94, 90, 46, 96, 103].
 - (iv) A-posteriori validations of the FDA and/or SP approximations: The accuracy of the approximations made during steps (ii) and (iii) must be evaluated, with particular attention on the ability of the SP representation or FDA algorithm to capture multi-scale interactions.

The multi-scale SP and FDA methods discussed here can be contrasted to recent complementary ideas in applied mathematics. In the work of the author and collaborators [89, 129] a theoretical link has been established between SP algorithms and heterogeneous multi-scale methods (HMM) [19, 20, 121, 25]. HMM algorithms are a mathematical synthesis of earlier work (see [85, 78] and references therein) as well as an abstract formulation that leads to new multi-scale algorithms for complex systems with widely disparate time scales [19, 20, 121, 25]. However, as noted in [89, 129], there are significant differences in the regimes of nonlinear dynamics being modeled by SP algorithms as compared with HMM. A key difference between SP and HMM lies in that while reduced HMM time-steppers have been analyzed and applied for various physical systems with wide scale separation, $\epsilon = 10^{-3}, 10^{-4}$, with ϵ the scale separation ratio between large and small scales, and rapid local statistical equilibration in time [19, 20, 121, 25], the skill and success of superparameterization algorithms relies on intermittency in space and time due to complex evolving strongly chaotic and quiescent regions without statistical equilibration despite only modest values of scale

separation, $\epsilon = 1/6$ to $1/10$ [129, 89, 124, 127]. Another related mathematical tool is the so-called gaptooth scheme [7]. The gaptooth method has formal similarity with SP but only works well on problems with an inertial manifold and for systems in which most modes are strongly decaying. The SP methods discussed here, by contrast, work in the strongly wave-like unstable regimes where there is intermittency and without even local equilibration, let alone an inertial manifold, as shown in a recent paper [94]. This work introduces a class of mathematical test models for SP that are simple enough to be analyzed with large confidence in a given physical context, yet reveal essential mechanisms and features of both SP and HMM numerical algorithms. This non-classical numerical analysis of model test problems provides firm mathematical underpinnings for the proposed new algorithms. Such test models can be designed in any physical context following the recipe developed there. The emphasis is on models with intermittent strongly unstable fluctuations and only moderate scale separation without statistical equilibration, so that more traditional numerical closure methods such as HMM cannot be applied. In the remaining parts of this section we illustrate a general idealized framework for mathematical issues related to steps i) and ii) in (3.1) and then discuss examples and important issues regarding judicious model error for FDA and SP.

3.1 Simple Gaussian Closure Models for Turbulent Dynamical Systems

In climate atmosphere ocean science, it is often useful [86, 80, 85] to consider the turbulent dynamical system from (2.1) with a special structure (note that the notation here is different from that in (2.1)),

$$(3.2) \quad \mathbf{u}_t = L\mathbf{u} + B(\mathbf{u}, \mathbf{u}) + \mathcal{S}(\mathbf{u}) + \bar{F} + F', \quad \mathbf{u} \in \mathbb{R}^N,$$

where \bar{F} is time-dependent deterministic forcing and F' is zero-mean random forcing. In applications the linear operator L involves rotation, stratification, topography, drag dissipation, etc.; $B(\mathbf{u}, \mathbf{u})$ denotes the quadratic effect of nonlinear advection, while $\mathcal{S}(\mathbf{u})$ denotes nonlinear source terms such as heating from clouds. Here and below, the decomposition of a variable such as \mathbf{u} into

$$(3.3) \quad \mathbf{u} = \bar{\mathbf{u}} + \mathbf{u}',$$

denotes the formal decomposition of the random field \mathbf{u} into its mean, $\bar{\mathbf{u}}$, and fluctuations \mathbf{u}' with $\overline{\mathbf{u}'} \equiv 0$. For simplicity in exposition it is assumed here that the source term $\mathcal{S}(\mathbf{u})$ is cubic so that

$$(3.4) \quad \mathcal{S}(\mathbf{u}) = \mathcal{S}(\bar{\mathbf{u}} + \mathbf{u}') \equiv \mathcal{S}(\bar{\mathbf{u}}) + \mathcal{S}_1(\bar{\mathbf{u}})\mathbf{u}' + \mathcal{S}_2(\bar{\mathbf{u}})(\mathbf{u}', \mathbf{u}') + \mathcal{S}_3(\bar{\mathbf{u}})(\mathbf{u}', \mathbf{u}', \mathbf{u}'),$$

where the last two terms are bilinear and trilinear forms in the fluctuations, \mathbf{u}' . Closure approximations for statistical solutions of (3.2) are developed by using the (Reynolds) decomposition $\mathbf{u} = \bar{\mathbf{u}} + \mathbf{u}'$ in (3.2) and computing

separate dynamic equations for the mean and fluctuations. First, the exact average equation for the mean is given by

$$(3.5) \quad \bar{\mathbf{u}}_t = L\bar{\mathbf{u}} + B(\bar{\mathbf{u}}, \bar{\mathbf{u}}) + \mathcal{S}(\bar{\mathbf{u}}) + \bar{F} + \overline{B(\mathbf{u}', \mathbf{u}')} + \overline{\mathcal{S}_2(\bar{\mathbf{u}})(\mathbf{u}', \mathbf{u}')} + \overline{\mathcal{S}_3(\bar{\mathbf{u}})(\mathbf{u}', \mathbf{u}', \mathbf{u}')}.$$

To derive an exact equation for the fluctuations it is convenient to introduce the linear operator depending on $\bar{\mathbf{u}}$ defined by

$$(3.6) \quad \mathcal{L}(\bar{\mathbf{u}})\mathbf{v}' \equiv L\mathbf{v}' + B(\bar{\mathbf{u}}, \mathbf{v}') + B(\mathbf{v}', \bar{\mathbf{u}}) + \mathcal{S}_1(\bar{\mathbf{u}})\mathbf{v}'.$$

The exact equations for the fluctuations are given by

$$(3.7) \quad \mathbf{u}'_t = \mathcal{L}(\bar{\mathbf{u}})\mathbf{u}' + F' + [B(\mathbf{u}', \mathbf{u}') + \mathcal{S}_2(\bar{\mathbf{u}})(\mathbf{u}', \mathbf{u}') + \mathcal{S}_3(\bar{\mathbf{u}})(\mathbf{u}', \mathbf{u}', \mathbf{u}')],$$

where in (3.7) $[f] = f' = f - \bar{f}$. So far both (3.5) and (3.7) are exact formulas. The Gaussian closure for (3.2) consists of replacing the last two terms in (3.7), the forcing and the term in brackets, which are turbulent fluctuations by a model with additional damping, $-\Gamma_M \mathbf{u}'$, with $\Gamma_M > 0$ and Gaussian random forcing, assumed here for simplicity to be white noise forcing, $\sigma_M \dot{W}$. Thus, the Gaussian Closure Model for the Fluctuations is given by

$$(3.8) \quad (\mathbf{u}'_M)_t = \mathcal{L}(\bar{\mathbf{u}}_M)\mathbf{u}'_M - \Gamma_M \mathbf{u}'_M + \sigma_M \dot{W}.$$

Note that \mathbf{u}'_M is a Gaussian random field with zero mean at each instant of time so that

$$(3.9) \quad A) \quad \overline{(\mathbf{u}'_M, \mathbf{u}'_M, \mathbf{u}'_M)} = 0,$$

and

$$(3.10) \quad B) \quad \text{the statistics of } \mathbf{u}'_M \text{ are completely determined by the symmetric covariance matrix, } \mathcal{C}_M \equiv \overline{\mathbf{u}'_M \mathbf{u}'_M^T}.$$

In particular, it is easy to calculate from (3.8) that \mathcal{C}_M satisfies the Lyapunov equation for the covariance

$$(3.11) \quad (\mathcal{C}_M)_t = (\mathcal{L}(\bar{\mathbf{u}}_M) - \Gamma_M)\mathcal{C}_M + \mathcal{C}_M(\mathcal{L}(\bar{\mathbf{u}}_M) - \Gamma_M)^T + Q_M,$$

with $Q_M > 0$ given by $Q_M = \sigma_M \sigma_M^T$. With (3.7) and (3.9) the equation for the mean of the Gaussian closure model is given by

$$(3.12) \quad (\bar{\mathbf{u}}_M)_t = L\bar{\mathbf{u}}_M + B(\bar{\mathbf{u}}_M, \bar{\mathbf{u}}_M) + \mathcal{S}(\bar{\mathbf{u}}) + \bar{F} + \overline{B(\mathbf{u}'_M, \mathbf{u}'_M)} + \overline{\mathcal{S}_2(\bar{\mathbf{u}}_M)(\mathbf{u}'_M, \mathbf{u}'_M)}.$$

The equations for the covariance in (3.11) and the mean (3.12) completely specify the entire dynamics for the Gaussian closure model; furthermore this is a realizable closure since $\mathcal{C}_M(0) \geq 0$ and (3.11) guarantee $\mathcal{C}_M(t) \geq 0$ for all times. While the Gaussian closure models provide an important

theoretical framework for illustrating the general use of the decomposition in (3.1 i), practical implementation is hampered by the fact that the covariance equation in (3.11) cannot be solved directly for large-dimensional turbulent dynamical systems with $N \gg 1$. Nevertheless, with judicious model error, such closure models can have very high skill for filtering turbulent dynamical systems [15, 96, 103]; non-Gaussian variants involving nonlinear stochastic parameters, finite state Markov chains and Gaussian mixtures can have even more skill for filtering and superparameterization (see §3.3 below). Such approximations become much more relevant for both SP and FDA when they are implemented on the small scales in a multi-scale environment as illustrated next [94].

3.2 Multi-scale Test Models for Superparameterization and Filtering

Following [94], the goal here is to show briefly how to develop simple multi-scale Gaussian mathematical test models for studying the issues in (3.1) for SP and FDA, as well as the accuracy of HMM algorithms with much wider scale separation. In the derivation here the multi-scale test models mimic (3.8), (3.11), (3.12) in a formal multi-scale environment. This is made explicit by introducing two spatial scales, X and $x \in \mathbb{R}^N$ with $X = \epsilon x$ and two time scales t, τ with $\tau = t/\epsilon$ with $\epsilon < 1$, a scale separation parameter. Assume that the physical field has the multi-scale decomposition

$$(3.13) \quad \mathbf{u} = \bar{\mathbf{u}}(X, t) + \mathbf{u}'(X, x, t, \tau).$$

For a function $f(t, \tau)$,

$$(3.14) \quad \langle f \rangle(t) = \epsilon \int_0^{\epsilon^{-1}} f(t, \tau) dt,$$

denotes the empirical time average over the fluctuations for a fixed value of ϵ . Repeating the derivation of the Gaussian closure model in this multi-scale context yields the mean model for $\bar{\mathbf{u}}_M(X, t)$ involving only the large scale variables (X, t)

$$(3.15) \quad (\bar{\mathbf{u}}_M)_t = L\bar{\mathbf{u}}_M + B(\bar{\mathbf{u}}_M, \bar{\mathbf{u}}_M) + \mathcal{S}(\bar{\mathbf{u}}) + \bar{F} + \langle \overline{B(\mathbf{u}'_M, \mathbf{u}'_M)} \rangle + S_2(\bar{\mathbf{u}}_M) \langle \overline{(\mathbf{u}'_M, \mathbf{u}'_M)} \rangle,$$

and the leading order multi-scale equation for the covariance $\mathcal{C}_M(X, t, \tau)$ with (X, t) regarded as frozen variables,

$$(3.16) \quad \frac{\partial \mathcal{C}_M}{\partial \tau} = \epsilon (\mathcal{L}(\bar{\mathbf{u}}_M) - \Gamma_M) \mathcal{C}_M + \epsilon \mathcal{C}_M (\mathcal{L}(\bar{\mathbf{u}}_M) - \Gamma_M)^T + \epsilon Q_M.$$

Note that the two averaging terms in the large scale equation from (3.15) are determined from the small scale time averaged covariance by

$$(3.17) \quad \langle \overline{B(\mathbf{u}'_M, \mathbf{u}'_M)} \rangle + S_u(\bar{\mathbf{u}}_M) \langle \overline{(\mathbf{u}'_M, \mathbf{u}'_M)} \rangle \equiv \mathcal{L}(\bar{\mathbf{u}}_M(X, t)) \langle \mathcal{C}_M \rangle(X, t),$$

where \mathcal{L} is just a pointwise linear operator on the large scales. Thus, the dynamics of the mean in (3.15) is completely determined by the empirical time average of the covariance matrix in (3.16). In turn, this small scale covariance depends nonlinearly (and nonlocally !) on this mean state. The equations in (3.15), (3.16), (3.17) are a more systematic version of the test models for SP proposed in [94] and the above viewpoint should be relevant for future applications and statistical numerical analysis.

In [94], a test model is developed for a single, real-valued scalar field u , in a single space dimension. A scalar differential operator is chosen to include advection, dispersion and dissipation, as in typical anisotropic systems. In the Fourier representation of the small-scale dynamics, a uniform damping and variance are chosen to yield a $-5/3$ turbulent spectrum. Thus, if the interaction with the large scale field is ignored, the statistical equilibrium state for the small-scale dynamics is an energetic turbulent field without scale separation. Intermittency is built in by making the large time behavior of the small-scale dynamics dependent on u . The resulting small-scale dynamics can then be solved exactly, and its effects on the large-scale dynamics explored precisely. In one parameter regime limit, the small-scale dynamics equilibrates on the short time-scale (the HMM limit), leading to a solvable equilibrium statistical closure on the large scales. Even here, however, non-trivial pattern formation in the large-scale dynamics can be generated solely by interaction with the small scales. The more interesting and relevant parameter regime leads to no small-scale equilibration on the short time scale, and hence no closed statistical equilibrium model for the large scale dynamics. The regimes of success and failure of the large-scale dynamics in this limit are then delineated systematically. This is the great advantage of the test-model approach: the error entailed in a specific SP scheme can be determined. Much more systematic mathematical work understanding step iii) from (3.1) needs to be developed in the context of superparameterization; namely, how can cheaper models capture the statistical dynamics of more complex systems. Examples already exist in the context of turbulent diffusion where time alternating superpositions of one-dimensional plane wave random fields can be used to simulate accurate large scale statistics of a turbulent tracer in a field with many spatio-temporal scales [78, 77, 22, 23].

3.3 Judicious Model Errors in Filtering Turbulent Dynamical Systems: Stochastic Parameterization Extended Kalman Filters (SPEKF)

All of the above theoretical developments utilize Gaussian closures as test models for highly anisotropic inhomogeneous turbulent systems. Can simple models incorporate non-Gaussian features of turbulent dynamical systems

yet have the advantage of cheap computational overhead for filtering turbulent dynamical systems from sparse observations? A key feature of turbulence is bursts of energy across multiple scales with intermittent instability and random forcing. Stochastic Parameterization Extended Kalman Filters (SPEKF) have been introduced and analyzed recently [33, 34, 96, 103] as computationally cheap algorithms which make judicious model errors which retain high filtering skill for complex turbulent signals [46, 59, 15, 96]. For example, aliasing is usually viewed as a bad feature of numerical algorithms; in the present context, judicious use of aliasing yields stochastic superresolution [96, 59, 103].

The basis for the SPEKF algorithms is the following system for the complex scalar partially observed turbulent signal u (the reader can think of a Fourier amplitude of turbulence at a given spatial wavenumber) coupled with stochastic additive forcing and multiplicative damping/instability coefficients, b, γ , which are learned “on the fly” from the observed turbulent signal

$$\begin{aligned}
 (a) \quad & du(t) = [(-\gamma(t) + i\omega)u(t) + b(t) + f(t)]dt + \sigma_u dW_u(t), \\
 (3.18) \quad (b) \quad & db(t) = [(-\gamma_b + i\omega_b)(b(t) - \hat{b})]dt + \sigma_b dW_b(t), \\
 (c) \quad & d\gamma(t) = -d_\gamma(\gamma(t) - \hat{\gamma})dt + \sigma_\gamma dW_\gamma(t),
 \end{aligned}$$

where W_u, W_b are independent complex Wiener processes with independent components and W_γ is a real Wiener process. There are nine parameters in the system (3.18): two damping parameters γ_b, d_γ , two oscillation frequencies ω and ω_b , two stationary mean terms \hat{b} and $\hat{\gamma}$ and noise amplitudes $\sigma_u, \sigma_b, \sigma_\gamma$; f is a deterministic forcing. The advantage of the equations in (3.18) is that they have non-Gaussian dynamics but nevertheless exactly solvable first and second-order statistics, so they are readily implemented practically in a filtering algorithm. The equations in (3.18) have rich statistical behavior in a variety of regimes and this complex behavior can be utilized to test the filter performance of a wide variety of Gaussian filter approximations [15]. Such models are also useful as an unambiguous test bed for all of the issues of prediction and model error discussed in section 2.

4 Concluding Discussion and Future Directions

Here, we briefly mention several topics for mathematical research directly connected with this expository article and not discussed in detail in sections 2 and 3 above.

4.1 Mathematically Rigorous FDT and Stochastic-Statistical Numerical Analysis

The recent paper [45] only begins the mathematically rigorous analysis of fluctuation-dissipation theorems for turbulent dynamical systems. Much more rigorous work should be done for time-periodic systems, general multiplicative noise, and rigorous FDT representation formulas. Further developments of the important role of information theory for model error and sensitivity are needed beyond references [95, 98, 102, 104, 105]. The recent papers [97, 31] contain much of the formal research program and demonstrate it on an exactly solvable test model. Besides the statistical/stochastic numerical analysis research program described in section 3, there is a great need for mathematical theory and the assessment of numerical algorithms which capture the long-time statistical dynamics of turbulent dynamical systems with high accuracy. Wang has carried out this important research program for the example of a turbulent dynamical system arising in the large Prandtl number limit of classical Rayleigh-Benard convection [125, 126] and this work serves as a model for further research.

4.2 Physics Constrained Data-Driven Statistical-Stochastic Models

It is extremely important to develop data-driven reduced stochastic-statistical models of turbulent dynamical systems for long range forecasting and uncertainty quantification. Standard linear regression models can have some skill but suffer from inherent mathematical limitations and intrinsic barriers in skill [98]. Ad-hoc nonlinear regression models can exhibit improved skill in a training time series (see references in [99]) but can suffer unphysical finite-time blow-up of statistical solutions, as well as unphysical pathology in their invariant measure [99]. There are rigorous proofs [130] that physics-constrained stochastic mode reduction models which are Markovian have the physically correct asymptotic behavior for their invariant measure for low-frequency variability but they require further generalizations to include non-Markovian memory effects for many applications. There is a wide array of data-driven clustering algorithms [87, 52, 53, 54, 29, 28, 55, 56] to develop multiple regime Markov models for use in prediction. Giannakis and the author [36, 37, 38] apply empirical information theory to assess the skill of coarse-grained partitions of phase space and reduced Markov models for long-range prediction. The methods of Horenko [52, 53, 54, 55, 56, 38] are especially promising in this context but need further physical constraints to be more useful for long range forecasting. This is an exciting area for future interdisciplinary research.

4.3 Multi-scale Models, Waves, and PDE's for the Tropics

This is a very important topic for climate science research to explain observations, develop theories, and improve numerical models [92]. It is also a very interesting topic with many new phenomena for rigorous PDE analysis [100, 18, 17, 108] with many open problems. Due to the lack of space, it is not discussed here despite the author's enthusiasm for these topics. Nevertheless, the interested reader can consult the above references, as well as the current research/expository article [64] for these developments.

Bibliography

- [1] Abramov R.; Majda A.J.; Kleeman R. Information theory and predictability for low frequency variability. *J. Atmos. Sci.* **62** (2005), 65-87.
- [2] Abramov R.; Majda A.J. Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems. *Nonlinearity* **20** (2007), 2793–2821.
- [3] Abramov R.; Majda A.J. New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems. *J. Nonlinear Sci.* **18** (2008), 303–341.
- [4] Abramov R.; Majda A.J. A new algorithm for low-frequency climate response. *J. Atmos. Sci.* **66** (2009), 286-309.
- [5] Abramov R. Short-time linear response with reduced-rank tangent map. *Chinese Ann. Math.* **30** (2009), no. 5, 447–462.
- [6] Abramov R.; Majda A.J. Low frequency climate response of quasigeostrophic wind-driven ocean circulation. Submitted to *J. Phys. Oceanogr.* (2011).
- [7] Armaou A.; Kevrekidis I.; and Theodoropoulos C. Equation-free gaptooth-based controller design for distributed complex/multiscale processes. *Computers Chem. Eng.* **29** (2004), 731–740.
- [8] Bell T. Climate sensitivity from fluctuation dissipation: Some simple model tests. *J. Atmos. Sci.* **37** (1980), 1700–1707.
- [9] Berner J.; Branstator G. Linear and nonlinear signatures in the planetary wave dynamics of an AGCM: Probability density functions. *J. Atmos. Sci.* **64** (2007), 117–136.
- [10] Bernardo J.M.; Smith A.F.M. *Bayesian theory*, Wiley series in probability and mathematical statistics, John Wiley & Sons, Hoboken NJ, 2000.
- [11] Berger J.O. *Statistical decision theory and Bayesian analysis*, Springer, New York, 1985.
- [12] Bourlioux A.; Majda A.J. Elementary models with probability distribution function intermittency for passive scalars with a mean gradient. *Phys Fluids* **14** (2002), 881–897
- [13] Biello J.A.; Majda A.J. A new multiscale model for the Madden-Julian oscillation. *J. Atmos. Sci.* **62** (2005), 1694–1721.
- [14] Branstator G.; Berner J. Linear and nonlinear signatures in the planetary wave dynamics of an AGCM: Phase space tendencies. *J. Atmos. Sci.* **62** (2005), 1792–1811.

- [15] Branicki M.; Gershgorin B.; Majda A.J.; Filtering skill for turbulent signals for a suite of nonlinear and linear Kalman filters. Submitted to *J. Comp. Phys.* (2011).
- [16] Carnevale G.; Falcioni M.; Isola S.; Purini R.; Vulpiani A. Fluctuation-response relations in systems with chaotic behavior. *Phys Fluids* **2** (1991), 2247–2254.
- [17] Dutrifoy A.; Majda A.J. The Dynamics of Equatorial Long Waves: A Singular Limit with Fast Variable Coefficients. *Comm. Math. Sci.* **4** (2006), no. 2, 375–397.
- [18] Dutrifoy A.; Majda A.J.; Schochet S. A Simple Justification of the Singular Limit for Equatorial Shallow-Water Dynamics, *Commun. Pur. Appl. Math.* **62** (2009), 322–333.
- [19] E W.; Engquist B. The heterogeneous multiscale methods. *Comm. Math. Sci.* **1** (2003), 87–132.
- [20] Engquist B.; Tsai Y-H. Heterogeneous multiscale methods for stiff ordinary differential equations. *Math. Comp.* **74** (2005), 1707–1742.
- [21] Emanuel K.A.; Wyngaard J.C.; McWilliams J.C.; Randall D.A.; Yung Y.L. *Improving the scientific foundation for atmosphere-land ocean simulations*, Natl. Acad. Press, Washington DC, 2005.
- [22] Elliott Jr F.W.; Horntrop D.J.; Majda A.J. Monte Carlo methods for turbulent tracers with long range and fractal random velocity fields. *Chaos* **7** (1997), no. 1, 39–48.
- [23] Elliott Jr F.W.; Majda A.J. A wavelet Monte Carlo method for turbulent diffusion with many spatial scales. *J. Comput. Phys.* **113** (1994), no. 1, 82–111.
- [24] Epstein E.S. Stochastic dynamic predictions, *Tellus* **21** (1969), 739–759.
- [25] Fatkulin I.; Vanden-Eijnden E. A computational strategy for multiscale systems with application to Lorenz 96 model. *J. Comp. Phys.* **200** (2004), 605–638.
- [26] Franzke C.; Majda A.J. Low-order stochastic mode reduction for a prototype atmospheric GCM. *J. Atmos. Sci.* **63** (2006), 1620–1635.
- [27] Franzke C.; Majda A.J.; Branstator G. The origin of nonlinear signatures of planetary wave dynamics: Mean phase space tendencies and contributions from non-gaussianity. *J. Atmos. Sci.* **64** (2007), 3987–4003.
- [28] Franzke C.; Horenko I.; Majda A.J.; Klein R. Systematic metastable regime identification in an AGCM, *J. Atmos. Sci.* **66** (2009), 1997–2012.
- [29] Franzke C.; Crommelin D.; Fischer A.; Majda A.J. A hidden Markov model perspective on regimes and metastability in atmospheric flows. *J. Climate* **21** (2008), 1740–1757.
- [30] Gershgorin B.; Majda A.J. A nonlinear test model for filtering slow-fast systems, *Commun. Math. Sci.* **6** (2008), 611–649.
- [31] Gershgorin B.; Majda A.J. A Test Model for Fluctuation-Dissipation Theorems with Time Periodic Statistics. *Physica D* **239** (2010), no. 17, 1741–1757.
- [32] Gershgorin B.; Majda A.J. Filtering a nonlinear slow-fast system with strong fast forcing. *Commun. Math. Sci.* **8** (2010), 67–92.

- [33] Gershgorin B.; Harlim J.; Majda A.J. Test models for improving filtering with model errors through stochastic parameter estimation. *J. Comp. Phys* **229** (2010), no.1, 1–31.
- [34] Gershgorin B.; Harlim J.; Majda A.J. Improving filtering and prediction of spatially extended turbulent systems with model errors through stochastic parameter estimation. *J. Comp. Phys* **229** (2010), no. 1, 32–57.
- [35] Gershgorin B.; Majda A.J. Filtering a statistically exactly solvable test model for turbulent tracers from partial observations. *J. Comp. Phys.* **230** (2011), 1602–1638.
- [36] Giannakis D.; Majda A.J. Quantifying the Predictive Skill in Long-Range Forecasting. Part I: Coarse-grained Predictions in a Simple Ocean Model. Submitted to *J. Climate* (2011).
- [37] Giannakis, D.; Majda, A.J.; Quantifying the Predictive Skill in Long-Range Forecasting. Part II: Model Error in Coarse-Grained Markov Models with Application to Ocean-Circulation Regimes. Submitted to *J. Climate* (2011).
- [38] Giannakis D.; Majda A.J.; Horenko I. Information theory, model error, and predictive skill of stochastic models for complex nonlinear systems. Submitted to *Physica D* (2011).
- [39] Grabowski W.W.; Smolarkiewicz P.K. CRCP: A cloud resolving convection parameterization for modeling the tropical convecting atmosphere. *Physica D* **133** (1999), 171–178.
- [40] Grabowski W.W. Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP). *J. Atmos. Sci.* **58** (2001), 978–997.
- [41] Grabowski W.W. An improved framework for superparameterization. *J. Atmos. Sci.* **61** (2004), 1940–1952.
- [42] Gritsun A.; Branstator G. Climate response using a three-dimensional operator based on the fluctuation-dissipation theorem. *J. Atmos. Sci.* **64** (2007), 2558–2575.
- [43] Gritsun A.; Branstator G.; Majda A.J. Climate response of linear and quadratic functionals using the fluctuation-dissipation theorem. *J. Atmos. Sci.* **65** (2008), 2824–2841.
- [44] Gritsun A.; Dymnikov V. Barotropic atmosphere response to small external actions. Theory and numerical experiments. *Izv. Atmos. Oceanic Phys.* **35** (1999), 511–525.
- [45] Hairer M.; Majda A.J. A Simple Framework to Justify Linear Response Theory. *Nonlinearity* **23** (2010), no. 4, 909–922.
- [46] Harlim J.; Majda A.J. Filtering Turbulent Sparsely Observed Geophysical Flows. *Mon. Wea. Rev.* **138** (2010), no. 4, 1050–1083.
- [47] Haven K.; Majda A.J.; Abramov R. Quantifying predictability through information theory: small sample estimation in non-Gaussian framework. *J. Comp. Phys.* **206** (2005), 334–362.
- [48] *Ocean Modeling in an Eddying Regime*. Edited by M.W. Hecht and H. Hasumi. American Geophysical Union Monograph Series, 2008.
- [49] Hinze J. *Turbulence*. McGraw-Hill, New York, 1959.

- [50] Holland D.M. An impact of sub-grid-scale ice-ocean dynamics on sea-ice cover. *J. Climate* **14** (2001), no. 7, 1585–1601.
- [51] Holland D.M.; Thomas R.H.; deYoung B.; Ribergaard M.H.; Lyberth B. Acceleration of Jakobshavn Isbrae Triggered by Warm Subsurface Ocean Waters. *Nature Geosci.* **1** (2008), 659–664.
- [52] Horenko I. Finite element approach to clustering of multidimensional time series. *SIAM J. Sci. Comput.* **32** (2010), 62–83.
- [53] Horenko I. On the identification of nonstationary factor models and their application to atmospheric data analysis. *J. Atmos. Sci.* **67** (2010), 1559–1574.
- [54] Horenko I. On clustering of non-stationary meteorological time series. *Dyn. Atmos. Oceans* **49** (2010), 164–187.
- [55] Horenko I. On robust estimation of low-frequency variability trends in discrete Markovian sequences of atmospheric circulation patterns. *J. Atmos. Sci.* **66** (2009), 2059–2072.
- [56] Horenko I. Nonstationarity in multifactor models of discrete jump processes, memory and application to cloud modeling. *J. Atmos. Sci.* (2011), Early online release, doi: 10.1175/2011JAS3692.1.
- [57] Jaynes E.T. Information Theory and Statistical Mechanics. *Phys. Rev.* **106** (1957), 620–630.
- [58] Joughin I.; Smith B.; Holland D.M. Sensitivity of 21st Century Sea Level to Ocean-Induced Thinning of Pine Island Glacier, Antarctica. *Geophys. Res. Lett.* **37** (2010), L20502.
- [59] Keating S.; Majda A.J.; Smith K.S. New methods for estimating poleward eddy heat transport using satellite altimetry. Submitted to *Mon. Wea. Rev.* (2011).
- [60] Kerstein A.R. Linear eddy modeling of turbulent transport II: Application to shear layer mixing. *Combust. Flame* **75** (1989), 397–413.
- [61] Khouider B.; Majda A.J. A simple multcloud parameterization for convectively coupled tropical waves. Part I: Linear analysis. *J. Atmos. Sci.* **63** (2006), 1308–1323.
- [62] Khouider B.; Majda A.J. Equatorial convectively coupled waves in a simple multcloud model. *J. Atmos. Sci.* **65** (2008), 3376–3397.
- [63] Khouider B.; St-Cyr A.; Majda A.J.; Tribbia J. MJO and convectively coupled waves in a coarse resolution GCM with a simple multcloud parameterization. *J. Atmos. Sci.* **68** (2011), 240–264.
- [64] Khouider B.; Majda A.J.; Stechmann S. Climate Science, Waves, and PDE’s for the tropics. Submitted to *Nonlinearity* (2011).
- [65] Khairoutdinov M.; Randall D.; DeMott C. Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *J. Atmos. Sci.* **62** (2005), 2136–2154.
- [66] Klein R. Scale-Dependent Asymptotic Models for Atmospheric Flows. *Ann. Rev. Fluid Mech.* **42** (2010), 249–274.
- [67] Kleeman R. Measuring dynamical prediction utility using relative entropy, *J. Atmos. Sci.* **59** (2002), 2057–2072.

- [68] Kleeman R.; Majda A.J.; Timofeyev I. Quantifying predictability in a model with statistical features of the atmosphere. *Proc. Natl. Acad. Sci. USA* **99** (2002), 15291–15296.
- [69] Kleeman R. Information theory and dynamical system predictability. *Entropy* **13** (2011), 612–649.
- [70] Kullback S.; Leibler R. On information and sufficiency. *Ann. Math. Stat.* **22** (1951), 79–86.
- [71] Laliberte F.; Pauluis O. Winter intensification of the moist branch of the circulation in simulations of 21st century climate. *Geophys. Res. Lett.* **37** (2010), L20707.
- [72] Leith C. Climate response and fluctuation dissipation. *J. Atmos. Sci.* **32** (1975), 2022–2025.
- [73] Lorenz E.N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20** (1963), 130–141.
- [74] Lorenz E.N. A study of predictability of a 28-variable atmospheric model. *Tellus* **17** (1968), 321–333.
- [75] Lorenz E.N. The predictability of a flow which possesses many scales of motion. *Tellus* **21** (1969), 289–307.
- [76] Lorenz E.N. Predictability: A problem partly solved. *Proc. Seminar on Predictability*, Reading, UK, ECMWF, **1** (1995), 1–18.
- [77] Majda A.J. Random shearing direction models for isotropic turbulent diffusion. *J. Statist. Phys.* **25** (1994), no. 5/6, 1153–1165.
- [78] Majda A.J.; Kramer P. Simplified models for turbulent diffusion: theory, numerical modeling, and physical phenomena. *Phys. Reports* **314** (1999), no. 4-5, 238–574.
- [79] Majda A.J.; Real world turbulence and modern applied mathematics. In *Mathematics: Frontiers and Perspectives 2000*, International Mathematical Union, 2000, 137–151, American Math. Society.
- [80] Majda A.J. *Introduction to P.D.E.s and Waves for the Atmosphere and Ocean*. Courant Lecture Notes Vol. 9, American Mathematical Society & Courant Institute of Mathematical Sciences, 2002
- [81] Majda A.J.; Kleeman R.; Cai D. A framework for predictability through relative entropy. *Meth. Appl. Anal.* **9** (2002), 425–444.
- [82] Majda A.J.; Timofeyev I.; Vanden-Eijnden E. Systematic strategies for stochastic mode reduction in climate. *J. Atmos. Sci.* **60** (2003), 1705–1722.
- [83] Majda A.J.; Timofeyev I. Low-dimensional chaotic dynamics versus intrinsic stochastic noise: A paradigm model. *Physica D* **199** (2004), 339–368.
- [84] Majda A.J.; Biello J.A. A multiscale model for the intraseasonal oscillation. *Proc. Natl. Acad. Sci. USA* **101** (2004), 4736–4741.
- [85] Majda A.J.; Abramov R.; Grote M. *Information Theory and Stochastics for Multiscale Nonlinear Systems*, CRM Monogr. Series, American Mathematical Society, 2005.
- [86] Majda A.J.; Wang X. *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows*. Cambridge University Press, 2006.

- [87] Majda A.J.; Franzke C.; Fischer A.; Crommelin D.T. Distinct metastable atmospheric regimes despite nearly Gaussian statistics: A paradigm model. *Proc. Natl. Acad. Sci. USA* **103** (2006), 8309–8314.
- [88] Majda A.J.; Grote M. Explicit off-line criteria for stable accurate time filtering of strongly unstable spatially extended systems. *Proc Natl Acad Sci USA* **104** (2007), 1124–1129.
- [89] Majda A.J. Multiscale models with moisture and systematic strategies for superparameterization. *J. Atmos. Sci.* **64** (2007), 2726–2734.
- [90] Majda A.J.; Franzke C.; Khouider B. An Applied Mathematics Perspective on Stochastic Modelling for Climate. *Phil. Trans. Roy. Soc.* **366** (2008), no. 1875, 2427–2453.
- [91] Majda A.J.; Stechmann S. The skeleton of tropical Intraseasonal Oscillations. *Proc. Natl. Acad. Sci. USA* **16** (2009), no. 21, 8417–8422.
- [92] Majda A.J.; Stechmann S. The Madden-Julian oscillation and the multiscale hierarchy of organized convection. *UCLA Tropical Meteorology and Climate Newsletter* (2009), available at <http://math.nyu.edu/faculty/majda/publicationrevised.html>.
- [93] Majda A.J.; Franzke C.; Crommelin D. Normal forms for reduced stochastic climate models, *Proc. Natl. Acad. Sci. USA* **106** (2009), 3649–3653.
- [94] Majda A.J.; Grote M. Mathematical test models for superparameterization in anisotropic turbulence. *Proc. Nat. Acad. Sci. USA* **106** (2009), no. 14, 5470–5474.
- [95] Majda A.J.; Gershgorin B. Quantifying Uncertainty in Climate Change Science Through Empirical Information Theory. *Proc. Natl. Acad. Sci. USA* **107** (2010), no. 34, 14958–14963.
- [96] Majda A.J.; Harlim J.; Gershgorin B. Mathematical strategies for filtering turbulent dynamical systems. *Discret Contin. Dyn. Sys.* **27** (2010), 441–486.
- [97] Majda A.J.; Wang X. Linear Response Theory for Statistical Ensembles in Complex Systems with Time-Periodic Forcing. *Comm. Math. Sci.* **8** (2010), no. 1, 145–172.
- [98] Majda A.J.; Gershgorin B.; Yuan Y. Low Frequency Climate Response and Fluctuation-Dissipation Theorems: Theory and Practice. *J. Atmos. Sci.* **67** (2010), no. 4, 1186–1201.
- [99] Majda A.J.; Yuan Y. Fundamental Limitations of Ad Hoc Linear and Quadratic Multi-Level Regression Models for Physical Systems. Submitted to *Discrete Cont. Dyn. Systems* (2010).
- [100] Majda A.J.; Souganidis P. The Existence and Uniqueness of Weak Solutions for Precipitation Fronts: A Novel Hyperbolic Free Boundary Problem in Several Space Variables. *Comm. Pure Appl. Math.* **63** (2010), no. 10, 1351–1361.
- [101] Majda A.J.; Xing Y. New Multi-Scale Models on Mesoscales and Squall Lines. *Comm. Math. Sci.* **8** (2010), no. 1, 113–134.
- [102] Majda A.J.; Abramov R.; Gershgorin B. High skill in low frequency climate response through fluctuation dissipation theorems despite structural instability. *Proc. Natl. Acad. Sci. USA* **107** 2010, no. 2, 581–586.

- [103] Majda A.J.; Harlim J. *Filtering Turbulent Signals in Complex Systems*. Cambridge University Press, UK, 2011.
- [104] Majda A.J.; Gershgorin B. Improving model fidelity and sensitivity for complex systems through empirical information theory. *Proc. Natl. Acad. Sci. USA* (2011), in press.
- [105] Majda A.J.; Gershgorin B. The link between statistical equilibrium fidelity and forecasting skill for complex systems with model error. *Proc. Natl. Acad. Sci. USA* (2011), in press.
- [106] Majda A.J.; Gershgorin B. Elementary models for turbulent diffusion with complex physical features: eddy diffusivity, spectrum, and intermittency. Submitted to *Phil. Trans. Roy. Soc* (2011).
- [107] Majda A.J.; Stechmann S. Nonlinear Dynamics and Regional Variations in the MJO Skeleton. Submitted to *J. Atmos. Sci.* (2011).
- [108] Moncrieff M.W.; Shapiro M.; Slinge J.; Molteni F. Collaborative research at the intersection of weather and climate, *WMO Bulletin* **56** (2007), 204–211.
- [109] Marconi U.; Puglisi A.; Rondoni L.; Vulpani A. Fluctuation-dissipation: Response theory in statistical physics. *Physics Reports* **461** (2008), 111–195.
- [110] Monin A.S.; Yaglom A.M. *Statistical Fluid Dynamics*. MIT Press, Cambridge, MA, 1971.
- [111] Neelin J.D.; Munnich M.; Su H.; Meyerson J.E.; Holloway C.E. Tropical drying trends in global warming models and observations. *Proc. Natl. Acad. Sci. USA* **103** (2003), 6110–6115.
- [112] Neelin, J.D.; Lintner B.R.; Tian B.; Li Q.; Zhang L.; Patra P.K.; Chahine M.T.; Stechmann S.N. Long tails in deep columns of natural and anthropogenic tropospheric tracers. *Geophys. Res. Lett.* **37** (2009), L05804.
- [113] National Research Council/Ad Hoc Study Group on Carbon Dioxide and Climate (1979) Carbon dioxide and climate: A scientific assessment: Report of an ad hoc study group on carbon dioxide and climate, Woods Hole, Massachusetts, July 23–27, 1979, to the Climate Research Board, Assembly of Mathematical and Physical Sciences, Natl. Acad. Press, Washington DC, pp 22.
- [114] Pauluis O.; Czaja A.; Korty R. The global atmospheric circulation on moist isentropes. *Science* **321** (2008), 1075–1078.
- [115] Pauluis O.; Czaja A.; Korty R. The Global Atmospheric Circulation in Moist Isentropic Coordinates. *J. Climate* **23** (2010), 3077–3093.
- [116] Penland C.; Sardeshmukh P. The optimal growth of tropical sea surface temperature anomalies. *J. Climate* **8** (1995), 1999–2024.
- [117] Randall D.A.; et al. Climate models and their evaluation. *Climate change 2007: The physical science basis, contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change*, eds S. Solomon et al., Cambridge Univ. Press, Cambridge, UK, 2007, 589–662.
- [118] Randall D.A.; Khairoutdinov M.; Arakawa A.; Grabowski W.W. Breaking the cloud parameterization deadlock. *Bull. Amer. Meteor. Soc.* **84** (2003), 547–1564.

- [119] Smith K.S. The geography of Linear Baroclinic Instability in Earth's Oceans. *J. Marine Res.* **65** (2007), 655–683.
- [120] Townsend A.A. *The Structure of Turbulent Shear Flow*. Cambridge University Press, UK, 1976.
- [121] Vanden-Eijnden E. Numerical techniques for multiscale dynamical systems with stochastic effects. *Comm. Math. Sci.* **1** (2003), 385–391.
- [122] Varadhan S.R.S. *Large Deviations and Applications*. CBMS-NSF Regional Conference Series in Applied Mathematics 46, SIAM, 1984.
- [123] Varadhan S.R.S. Large Deviations and Applications. *Expo. Math.* **3** (1985), 251–272.
- [124] Visbeck M.; Marshall J.; Haine T.; Spall M. Specification of eddy transfer coefficients in coarse-resolution ocean circulation models. *J. Phys. Oceanogr.* **27** (1997), 381–402.
- [125] Wang X. Stationary statistical properties of Rayleigh-Bénard convection at large Prandtl number. *Comm. Pur. Appl. Math.* **61** (2008), no. 6, 789–815.
- [126] Wang X. Approximation of stationary statistical properties of dissipative dynamical systems: time discretization. *Math. Comp.* **79** (2010), 259–280.
- [127] Wardle R.; Marshall J. Representation of eddies in primitive equation models by a PV flux. *J. Phys. Oceanogr.* **30** (2000), 2481–2503.
- [128] Williams D. *Weighing the odds: A course in probability and statistics*. Cambridge University Press, UK, 2001.
- [129] Xing Y.; Majda A.J.; Grabowski W.W. New efficient sparse space-time algorithms for superparametrization on mesoscales. *Mon. Wea. Rev.* **137** (2009), 4307–4324.
- [130] Yuan Y.; Majda A.J. Invariant Measures and Asymptotic Gaussian Bounds for Normal Forms of Stochastic Climate Model. *Chinese Annals of Mathematics, Series B* (2010), in press, doi: 10.1007/s11401-011-0647-2.

Received Month 200X.