

Stochastic climate theory and modeling

Article

Accepted Version

Franzke, C. L. E., O'Kane, T. J., Berner, J., Williams, P. D. and Lucarini, V. (2015) Stochastic climate theory and modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 6 (1). pp. 63-78. ISSN 1757-7799 doi: <https://doi.org/10.1002/wcc.318> Available at <http://centaur.reading.ac.uk/38700/>

It is advisable to refer to the publisher's version if you intend to cite from the work.

To link to this article DOI: <http://dx.doi.org/10.1002/wcc.318>

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online



Stochastic Climate Theory and Modelling

Christian L. E. Franzke^{*1}, Terence J. O’Kane², Judith Berner³, Paul D. Williams⁴, and Valerio Lucarini^{1,5}

¹*Meteorological Institute and Centre for Earth System Research and Sustainability (CEN), University of Hamburg, Hamburg, Germany*

²*Centre for Australian Weather and Climate Research, CSIRO Marine and Atmospheric Research, Hobart, Australia*

³*National Center for Atmospheric Research, Boulder, USA*

⁴*Department of Meteorology, University of Reading, Reading, UK*

⁵*Department of Mathematics and Statistics, University of Reading, Reading, UK*

August 27, 2014

Abstract

Stochastic methods are a crucial area in contemporary climate research and are increasingly being used in comprehensive weather and climate prediction models as well as reduced order climate models. Stochastic methods are used as subgrid-scale parameterizations as well as for model error representation, uncertainty quantification, data assimilation and ensemble prediction. The need to use stochastic approaches in weather and climate models arises because we still cannot resolve all necessary processes and scales in comprehensive numerical weather and climate prediction models. In many practical applications one is mainly interested in the largest and potentially predictable scales and not necessarily in the small and fast scales. For instance, reduced order models can simulate and predict large scale modes. Statistical mechanics and dynamical systems theory suggest that in reduced order models the impact of unresolved degrees of freedom can be represented by suitable combinations of deterministic and stochastic components and non-Markovian (memory) terms. Stochastic approaches in numerical weather and climate prediction models also lead to the reduction of model biases. Hence, there is a clear need for systematic stochastic approaches in weather and climate modelling. In this review we present evidence for stochastic effects in laboratory experiments. Then we provide an overview of stochastic climate theory from an applied mathematics perspectives. We also survey the current use of stochastic methods in comprehensive weather and climate prediction models and show that stochastic parameterizations have the potential to remedy many of the current biases in these comprehensive models.

^{*}Corresponding Author; Meteorological Institute and Centre for Earth System Research and Sustainability, University of Hamburg, Grindelberg 7, D-20144 Hamburg, Germany; Email: christian.franzke@uni-hamburg.de

1 Introduction

The last few decades have seen a considerable increase in computing power which allows the simulation of numerical weather and climate prediction models with ever higher resolution and the inclusion of ever more physical processes and climate components (e.g. cryosphere, chemistry). Despite this increase in computer power many important physical processes (e.g. tropical convection, gravity wave drag, clouds) are still not or only partially resolved in these numerical models. Despite the projected exponential increase in computer power these processes will not be explicitly resolved in numerical weather and climate models in the foreseeable future^{120,171}. For instance, Dawson et al.²⁴ have demonstrated using the ECMWF integrated forecast system that extremely high resolutions (T1279, which corresponds to a grid spacing of about 16km) are required to accurately simulate the observed Northern hemispheric circulation regime structure. This resolution, typical for limited area weather and climate models used for short term prediction, remains unfeasible for the current generation of high resolution global climate models due to computational and data storage requirements. Hence, these missing processes need to be parameterized, i.e. they need to be represented in terms of resolved processes and scales¹⁵³. This representation is important because small-scale (unresolved) features can impact the larger (resolved) scales^{84,162} and lead to error growth, uncertainty and biases.

At present, these parameterizations are typically deterministic, relating the resolved state of the model to a unique tendency representing the integrated effect of the unresolved processes. These “bulk parameterizations” are based on the notion that the properties of the unresolved subgrid-scales are determined by the large-scales. However, studies have shown that resolved states are associated with many possible unresolved states^{22,144,167}. This calls for stochastic methods for numerical weather and climate prediction which potentially allow a proper representation of the uncertainties, a reduction of systematic biases and improved representation of long-term climate variability. Furthermore, while current deterministic parameterization schemes are inconsistent with the observed power-law scaling of the energy spectrum^{5,142} new statistical dynamical approaches that are underpinned by exact stochastic model representations have emerged that overcome this limitation. The observed power spectrum structure is caused by cascade processes. Recent theoretical studies suggest that these cascade processes can be best represented by a stochastic non-Markovian Ansatz. Non-Markovian terms are necessary to model memory effects due to model reduction¹⁹. It means that in order to make skillful predictions the model has to take into account also past states and not only the current state (as for a Markov process).

We first review observational evidence of stochasticity in laboratory geophysical fluid experiments (section 2), then discuss stochastic climate theory in fast-slow systems (system 3). In section 4 we present statistical physics approaches and in section 5 we review the current state of stochastic-dynamic weather and climate modelling. We close with an outlook and challenges for the future of weather and climate modelling (section 6).

2 Laboratory Evidence of Stochasticity

Research on the climate system is somewhat hindered by the obvious difficulties of performing reproducible experiments on the atmosphere and ocean in different parameter regimes. For example, an optical physicist studying the nonlinear response of isolated atoms to intense electromagnetic waves can easily change the incident wavelength¹¹⁰. In contrast, climate scientists cannot (and arguably should not!) change the rotation rate of the planet or the intensity of the incoming solar radiation. To some extent, numerical simulations come to the rescue, by allowing us to perform virtual experiments. However, the grid spacing in climate models is orders of magnitude larger than the smallest energized scales in the atmosphere and ocean, introducing biases.

Fortunately, there is another option available to us. It is possible to exploit dynamical similarity³⁰ to study analogues of planetary fluid flow in bespoke laboratory experiments. The traditional set-up is the classic rotating annulus, which has been used for decades to study baroclinic instability and other large-scale phenomena⁶¹. Recent observations of small-scale inertia-gravity waves embedded within a large-scale baroclinic wave^{85,172,173} have allowed the scale interactions between these two modes to be studied in a laboratory fluid for the first

74 time. The experimental apparatus consists of a two-layer isothermal annulus forced by a differentially rotating
75 lid, which drives a shear across the internal interface and represents the mid-latitude tropospheric wind shear.

76 The large-scale baroclinic wave in these laboratory experiments exhibits regime behavior, equilibrating at
77 finite amplitude with a zonal wavenumber of typically 1, 2, or 3. These simple wave modes are regarded as
78 prototypes of the more complicated regime behavior in the atmosphere, such as mid-latitude blocking^{160,164}.
79 A notable finding from repeated experiments using this apparatus is that small-scale inertia-gravity waves can
80 induce large-scale regime transitions, despite the separation of wavelengths by an order of magnitude¹⁶⁸. An
81 example of this process is illustrated in Figure 1. A wavenumber 2 mode *without* co-existing inertia-gravity
82 waves (upper row) remains a wavenumber 2 mode indefinitely, drifting around the annulus with the zonal-mean
83 flow. In contrast, with the same parameter values, a wavenumber 2 mode *with* co-existing inertia-gravity waves
84 (lower row) is found to have a finite probability of transitioning to a wavenumber 1 mode. The amplitude of
85 the inertia-gravity waves is controlled here without directly affecting the large-scale mode, by slightly varying
86 the interfacial surface tension between the two immiscible fluid layers.

87 The laboratory transitions discussed above are reminiscent of noise-induced transitions between different
88 equilibrium states in a meta-stable dynamical system¹⁵⁸. To test this interpretation, a quasi-geostrophic nu-
89 merical model that captures the meta-stability of the large-scale flow in the rotating annulus¹⁷⁴ was run with
90 and without weak stochastic forcing added to the potential vorticity evolution equation for each fluid layer.
91 The stochastic forcing was an approximate representation of the inertia-gravity waves, which are inherently
92 ageostrophic and are therefore forbidden from the quasi-geostrophic model. Consistent with the laboratory
93 experiments, only when the noise term was activated did the numerical simulations exhibit large-scale wave
94 transitions in the equilibrated flow¹⁶⁹. In further numerical experiments, the noise was found to be able to
95 influence wavenumber selection during the developing baroclinic instability.

96 In summary, the above laboratory experiments constitute the first evidence in a real fluid that small-scale
97 waves may trigger large-scale regime transitions. In a numerical model in which the small-scale waves were
98 absent, the transitions were captured through the addition of stochastic noise. Note that the small-scale waves
99 satisfy the dispersion relation for inertia-gravity waves and are therefore coherent in space and time, and yet
100 apparently they are 'sensed' by the large-scale flow as if they were random fluctuations. These results have led
101 to a possible interpretation of sudden stratospheric warmings as noise-induced transitions⁹. Furthermore, these
102 laboratory results help to motivate the development of stochastic parameterizations in climate models and a
103 more general development of stochastic climate theory.

104 3 Stochastic Climate Theory

105 Climate is a multi-scale system in which different physical processes act on different temporal and spatial
106 scales⁶⁹. For instance, on the micro-scale are turbulent eddies with time scales of seconds to minutes, on
107 the meso-scale is convection with time scales of hours to days, on the synoptic scale are mid-latitude weather
108 systems and blocking with time scales from days to weeks, on the large-scale are Rossby waves and teleconnection
109 patterns with time scales of weeks to seasons. And there is the coupled atmosphere-ocean system with time
110 scales of seasons to decades. The crucial point here is that all these processes acting on widely different temporal
111 and spatial scales, interact with each other due to the inherent nonlinearity of the climate system. We have
112 shown an illustrative laboratory example for this in the previous section.

113 For many practical applications we are only interested in the processes on a particular scale and not in
114 the detailed evolution of the processes at smaller scales. Often the scales of interest are linked to inherently
115 predictable processes, while the smaller scales processes are unpredictable. For instance, in the above laboratory
116 experiment we are interested in the regime behavior and not in the detailed evolution of the inertia-gravity
117 waves. In numerical simulations the fastest scales, which are typically also the smallest scales, use up the bulk
118 of computing time, slowing down the computation of the processes of actual interest. In numerical weather and
119 climate prediction many of the small scale processes are currently not explicitly resolved and won't be in the

120 foreseeable future. This neglect of these processes can lead to biases in the simulations. Because of that the
 121 unresolved processes need to be parameterized as demonstrated in the previous section.

122 Stochastic climate theory is based on the concept of scale separation in space or time. Hasselmann⁵⁶ was
 123 the first to propose to split the state vector \vec{z} into slow climate components \vec{x} and fast weather fluctuations \vec{y}
 124 and then to derive an effective equation for the slow climate variables only. In this equation the effect of the
 125 now unresolved variables is partially represented as a noise term. The physical intuition behind this idea is,
 126 for example, that the aggregated effect of 'fast' weather fluctuations drives fluctuations in the 'slower' ocean
 127 circulation. To first order such a model can explain the 'red' spectrum of oceanic variables^{36,64}. It has to be
 128 noted that there is no scale separation in the climate system. This lack of time-scale separation introduces
 129 non-Markovian (memory) effects and complicates the derivation of systematic parameterizations.

130 Rigorous mathematical derivations for this approach have been provided by Gottwald and Melbourne⁵²,
 131 Khasminsky⁶⁶, Kurtz⁷⁶, Melbourne and Stuart¹⁰¹, Papanicolaou¹²⁴, Pavliotis and Stuart¹²⁵. For accessible
 132 reviews see Givon et al.⁵⁰ and the text book by Pavliotis and Stuart¹²⁵. This approach has been applied to
 133 climate models by Majda and coworkers^{29,37,38,89-94,96}. Climate models have the following general functional
 134 form

$$d\vec{z} = \left(\tilde{F} + \tilde{L}\vec{z} + \tilde{B}(\vec{z}, \vec{z}) \right) dt \quad (1)$$

135 where \tilde{F} denotes an external forcing, \tilde{L} a linear operator and \tilde{B} a quadratic nonlinear operator. Eq. (1)
 136 constitutes the form of the dynamical cores of weather and climate prediction models.

137 Now splitting the state vector \vec{z} into slow \vec{x} and fast \vec{y} components (which amounts to assuming a time
 138 scale separation) and assuming that the nonlinear self-interaction of the fast modes $\tilde{B}(\vec{y}, \vec{y})$ can be represented
 139 by a stochastic process^{37,38,89,90} leads to a stochastic differential equation. The stochastic mode reduction
 140 approach^{37,38,89,90} then predicts the functional form of reduced climate models for the slow variable \vec{x} alone:

$$d\vec{x} = (F + L\vec{x} + B(\vec{x}, \vec{x}) + M(\vec{x}, \vec{x}, \vec{x})) dt + \sigma_A d\vec{W}_A + \sigma_A(\vec{x}) d\vec{W}_M \quad (2)$$

141 Structurally new terms are a deterministic cubic term which acts predominantly as nonlinear damping and
 142 both additive and multiplicative (state-dependent) noise terms. The fundamentals of stochastic processes and
 143 calculus are explained in Box 1. The multiplicative noise and the cubic term stem from the nonlinear interaction
 144 between the resolved and unresolved modes³⁷.

145 The above systematic procedure allows also a physical interpretation of the new deterministic and stochastic
 146 terms³⁷. The additive noise stems both from the nonlinear interaction amongst the unresolved modes and the
 147 linear interaction between resolved and unresolved modes³⁷.

BOX 1**STOCHASTIC PROCESSES**

In contrast to deterministic processes stochastic processes have a random component. See the books by Lemons⁸⁶ and Gardiner⁴⁹ for intuitive introductions to stochastic processes. Typically stochastic processes are driven by white noise. White noise is a serially uncorrelated time series with zero mean and finite variance⁴⁹.

A stochastic differential equation (SDE) is a combination of a deterministic differential equation and a stochastic process. In contrast to regular calculus, stochastic calculus is not unique; i.e. different discretizations of its integral representation lead to different results even for the same noise realization. The two most important calculi are Ito and Stratonovich. See details in Gardiner⁴⁹. The physical difference is that Ito calculus has uncorrelated noise forcing while Stratonovich allows for finite correlations between noise increments. Hence, physical systems have to be typically approximated by Stratonovich SDEs. On the other hand, it is mathematically straightforward to switch between the two calculi. So one only needs to make a decision at the beginning which calculus is more appropriate for modeling the system under consideration and can then switch to the mathematically more convenient form.

SDEs describe systems in a path wise fashion. The Fokker-Planck equation (FPE) describes how the probability distribution evolves over time⁴⁹. Thus, SDEs and the FPE offer two different ways at looking at the same system. The parameters of SDEs and their corresponding FPE are linked; thus, one can use the FPE to estimate the parameters of the corresponding SDE^{3,146}.

Multiplicative (or state-dependent) noise is important for deviations from Gaussianity and thus extremes. The intuition behind multiplicative noise is as follows: On a windless day the fluctuations are very small, whereas on a windy day not only is the mean wind strong but also the fluctuations around this mean are large; thus, the magnitude of the fluctuations dependent on the state of the system.

The first practical attempts at stochastic climate modelling were made using Linear Inverse Models (LIM)^{127,128,166,176} and dynamically based linear models with additive white noise forcing^{26,27,34,35,179}. These approaches linearise the dynamics and then add white noise and damping¹⁶⁶ in order to make the models numerically stable (i.e. the resulting linear operator should only have negative eigenvalues to ensure stability and reliability of the solutions). While these models have encouraging predictive skill, especially for ENSO, they can only produce Gaussian statistics and, thus, are less useful for predictions of high impact weather.

Recently there are encouraging attempts in fitting nonlinear stochastic models to data. These include multi-level regression^{70,74}, fitting the parameters via the Fokker-Planck equation^{3,146}, stochastic averaging^{23,103}, optimal prediction^{18,154} or Markov Chains²¹. Most of the previous approaches fitted the parameters of the stochastic models without taking account of physical constraints, e.g. global stability. Many studies linearized the dynamics and then added additional damping to obtain numerically stable models^{1,2,166,179}. Majda et al.⁹⁶ developed the nonlinear normal form of stochastic climate models and also physical constraints for parameter estimation. Recent studies use these physical constraints to successfully derive physically consistent stochastic climate models^{57,97,126}.

Most of the above approaches are based on an implicit assumption of time scale separation. However, the climate system has a spectrum with no clear gaps which would provide the basis of scale separation and the derivation of reduced order models. Such a lack of time scale separation introduces memory effects into the truncated description. Memory effects mean that the equations become non-Markovian and that also past states need to be used in order to predict the next state. This can be explained by considering the interaction between a large-scale Rossby wave with a smaller scale synoptic wave. At some location the Rossby wave will favor the development of the synoptic wave. Initially this synoptic wave grows over some days without affecting the Rossby wave. Once the synoptic wave has reached a sufficient large amplitude it will start affecting the Rossby wave. Now in a reduced order model only resolving the Rossby wave but not the synoptic wave this interaction cannot be explicitly represented. However, because the Rossby wave initially triggered the synoptic wave which then in turn affects it some days later, this can be modeled with memory terms which takes into account that

178 the Rossby wave has triggered at time t_0 an anomaly which will affect it at some later time t_n .

179 Recently, Wouters and Lucarini¹⁷⁷ have proposed to treat comprehensively the problem of model reduction in
180 multi-scale systems by adapting the Ruelle response theory^{136,137} for studying the effect of the coupling between
181 the fast and slow degrees of freedom of the system. This theory has previously been used in a geophysical context
182 to study the linear and nonlinear response to perturbations^{87,88}, which also allows climate change predictions.
183 This approach is based on the *chaotic hypothesis*⁴⁸ and allows the general derivation of the reduced dynamics of
184 the slow variables able to mimic the effect of the fast variables in terms of matching the changes in the expectation
185 values of the observables of the slow variables. The ensuing parametrization includes a deterministic correction,
186 which is a mean field result and corresponds to linear response, a general correlated noise and a non-Markovian
187 (memory) term. These results generalize Eq. (2). In the limit of infinite time-scale separation, the classical
188 results of the averaging method is recovered. Quite reassuringly, the same parametrizations can be found using
189 a classical Mori-Zwanzig approach¹⁹, which is based on projecting the full dynamics on the slow variables and
190 general mathematical results provide evidence that deterministic, stochastic and non-Markovian components
191 should constitute the backbone of parameterizations^{17,178}. Recent studies show improvements over approaches
192 based on time-scale separation^{17,177,178}.

193 Recent studies have shown that stochastic approaches are also important for the prediction of extreme events
194 and tipping points^{40,41,155,156}. Sura¹⁵⁶ discusses a stochastic theory of extreme events. He especially focuses
195 on deviations from a Gaussian distribution; i.e. skewness and kurtosis, as first measures of extremes. He shows
196 that multiplicative noise plays a significant role in causing non-Gaussian distributions. Franzke⁴⁰ shows that
197 both deterministic nonlinearity and multiplicative noise are important in predicting of extreme events.

198 4 Statistical Physics Approaches to Stochastic Climate Theory

199 Significant progress has been achieved in the development of tractable and accurate statistical dynamical clo-
200 sures for general inhomogeneous turbulent flows that are underpinned by exact stochastic models (see Box 2).
201 For an accessible review see the text book by Heinz⁵⁸. The statistical dynamical closure theory, pioneered by
202 Kraichnan⁷¹, has been recognized as a natural framework for a systematic approach to modelling turbulent
203 geophysical flows. Closure theories like the Direct Interaction Approximation (DIA),⁷¹ for homogeneous tur-
204 bulance and the Quasi-Diagonal Direct Interaction Approximation (QDIA),⁴² for the interaction of mean flows
205 with inhomogeneous turbulence have exact generalized Langevin model representations⁶⁰. This means that
206 such closures are realizable; i.e. they have non-negative energy.

207 The first major application of turbulence closures has been the examination of the predictability of geo-
208 physical flows. Early approaches applied homogeneous turbulence models to predicting error growth^{77,79,83}
209 whereas more recent advances by Frederiksen and O’Kane⁴⁶, O’Kane and Frederiksen¹¹³, building on the pi-
210 oneering studies of Epstein³² and Pitcher¹³⁰, have enabled predictability studies of inhomogeneous strongly
211 non-Gaussian flows typical of the mid-latitude atmosphere. Turbulence closures have also been used for Subgrid-
212 Scale Parameterisation (SSP) of the unresolved scales, for example eddies in atmospheric and ocean general
213 circulation models. Since it is generally only possible to represent the statistical effects of unresolved eddies
214 while their phase relationships with the resolved scales are lost¹⁰⁰, statistical dynamical turbulence closures are
215 sufficient to allow SSPs to be formulated in a completely transparent way^{42,43,73,77,80,112,134}. Insights gained
216 through the development of inhomogeneous turbulence closure theory have motivated the recent development
217 of general stochastic forms for subgrid-scale parameterisations for geophysical flows⁶⁸.

4.1 Statistical dynamical closure theory

BOX 2 CLOSURE PROBLEM

In order to describe the statistical behavior of a turbulent flow the underlying nonlinear dynamical equations must be averaged. For simplicity we consider a generic equation of motion with quadratic nonlinearity for homogeneous turbulence, in which the mean field is zero, and the fluctuating part of the vorticity in Fourier space, $\hat{\zeta}_{\mathbf{k}}$, satisfies the equation:

$$\frac{\partial}{\partial t} \hat{\zeta}_{\mathbf{k}}(t) = K_{\mathbf{k}\mathbf{p}\mathbf{q}} \hat{\zeta}_{-\mathbf{p}}(t) \hat{\zeta}_{-\mathbf{q}}(t). \quad (3)$$

219

where \mathbf{p} and \mathbf{q} are the other wave numbers describing triad interactions i.e. $\mathbf{k} = (k_x, k_y)$ where $\delta(\mathbf{k} + \mathbf{p} + \mathbf{q}) = 1$ if $\mathbf{k} + \mathbf{p} + \mathbf{q} = 0$ and 0 otherwise. Here $K_{\mathbf{k}\mathbf{p}\mathbf{q}}$ are the interaction or mode coupling coefficients. The correlation between the eddies can now be represented by an equation for the covariance (cumulant in terms of wavenumbers \mathbf{k} and \mathbf{l}) which is found to depend on the third order cumulant in Fourier space:

$$\frac{\partial}{\partial t} \langle \hat{\zeta}_{\mathbf{k}}(t) \hat{\zeta}_{-\mathbf{l}}(t') \rangle = K_{\mathbf{k}\mathbf{p}\mathbf{q}} \langle \hat{\zeta}_{-\mathbf{p}}(t) \hat{\zeta}_{-\mathbf{q}}(t) \hat{\zeta}_{-\mathbf{l}}(t') \rangle. \quad (4)$$

Similarly the third order cumulant depends on the fourth order and so on such that we see that an infinite hierarchy of moment or cumulant equations is produced. Statistical turbulence theory is principally concerned with the methods by which this moment hierarchy is closed and the subsequent dynamics of the closure equations. The fact that for homogeneous turbulence the covariance matrix is diagonal greatly simplifies the problem. The majority of closure schemes are derived using perturbation expansions of the nonlinear terms in the primitive dynamical equations. The most successful theories use formal renormalization techniques^{19,58}.

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

The development of renormalized turbulence closures was pioneered by Kraichnan's Eulerian DIA⁷¹ for homogeneous turbulence. The DIA, so named because it only takes into account directly interacting modes, can be readily regularised to include approximations to the indirect interactions^{45,111} which are required to obtain the correct inertial range scaling laws. Other homogeneous closures such as Herring's self consistent field theory (SCFT⁵⁹) and McComb's local energy transfer theory (LET),⁹⁹ were independently developed soon after. The DIA, SCFT and LET theories have since been shown to form a class of homogeneous closures that differ only in whether and how a fluctuation dissipation theorem (FDT i.e. the linear response of a system to an infinitesimal perturbation as it relaxes toward equilibrium)^{15,25,44,71} is applied. As noted earlier, a consequence of the DIA having an exact stochastic model representation is that it is physically realizable, ensuring positive energy spectra. This is in contrast with closures based on the quasi-normal hypothesis which require further modifications in order to ensure realizability; an example of such a closure is the eddy damped quasi-normal Markovian (EDQNM) closure^{77,109,116} developed as a better Markovian fit to the DIA. The EDQNM is dependent on a choice of an eddy-damping parameter which can be tuned to match the phenomenology of the inertial range. This Markovian assumption assumes that the rate at which the memory integral decays is significantly faster than the time scale on which the covariances evolve. The relative success of these turbulence closures has enabled the further study of the statistics of the predictability of homogeneous turbulent flows^{77-79,102}.

Frederiksen⁴² formulated a computationally tractable non-Markovian (memory effects) closure, the quasi-diagonal direct interaction approximation (QDIA), for the interaction of general mean and fluctuating flow components with inhomogeneous turbulence and topography. The QDIA assumes that, prior to renormalisation, a perturbative expansion of the covariances are diagonal at zeroth order. In general, very good agreement has been found between the QDIA closure results and the statistics of DNS^{45,46,111}.

The non-Markovian closures discussed above are systems of integro-differential equations with potentially long time-history integrals posing considerable computational challenges, however various ways to overcome

243 these challenges exist^{44,46,111–115,135} and have been generalised to allow computationally tractable closure mod-
 244 els for inhomogeneous turbulent flow over topography to be developed^{42,46,111}. An alternative derivation of a
 245 stochastic model of the Navier-Stokes equations has been put forward by Memin¹⁰⁷. It is based on a decompo-
 246 sition of the velocity fields into a differentiable drift and a stochastic component.

247 4.2 Statistical dynamical and stochastic subgrid modelling

248 Many subgrid-scale stress models assume the small scales to be close to isotropic and in equilibrium such that
 249 energy production and dissipation are in balance, similar to the eddy viscosity assumption of the Smagorinsky
 250 model¹⁴⁸. Using the DIA, Kraichnan⁷¹ showed that for isotropic turbulence the inertial transfer of energy could
 251 be represented as a combination of both an eddy viscous (on average energy drain from retained to subgrid
 252 scales) and stochastic back-scatter (positive semi-definite energy input from subgrid to retained scales) terms.
 253 The nonlinear transfer terms represented by eddy viscosity and stochastic back-scatter are the subgrid processes
 254 associated with the respective eddy-damping and nonlinear noise terms that constitute the right hand side of
 255 the DIA tendency equation for the two-point cumulant $\frac{\partial}{\partial t} \langle \hat{\zeta}_{\mathbf{k}}(t) \hat{\zeta}_{-\mathbf{k}}(t') \rangle$. Leith⁷⁷ used the EDQNM closure to
 256 calculate an eddy dissipation function that would preserve a stationary k^{-3} kinetic energy spectrum for two-
 257 dimensional turbulence. Kraichnan⁷³ developed the theory of eddy viscosity in two and three dimensions and
 258 was the first to identify the existence of a strong cusp in the spectral eddy viscosity near the cutoff wavenumber
 259 representing local interactions between modes below and near the cusp. Rose¹³⁴ argued for the importance of
 260 eddy noise in subgrid modelling.

261 O’Kane and Frederiksen¹¹² calculated QDIA based SSPs considering observed atmospheric flows over global
 262 topography and quantifying the relative importance of the subgrid-scale eddy-topographic, eddy-mean field,
 263 quadratic mean and mean field-topographic terms. They also compared the QDIA based SSPs to heuristic
 264 approaches based on maximum entropy, used to improve systematic deficiencies in ocean climate models⁶².
 265 While closure models may be the natural starting place for developing subgrid-scale parameterisations, their
 266 complexity makes them difficult to formulate and apply to multi-field models like GCMs, even though successful
 267 studies exist^{68,180}.

268 5 Stochastic Parametrisation Schemes in Comprehensive Models

269 Climate and weather predictions are only feasible because the governing equations of motion and thermody-
 270 namics are known. To solve these equations we need to resort to numerical simulations that discretize the
 271 continuous equations onto a finite grid and parameterize all processes that cannot be explicitly resolved. Such
 272 models can be characterized in terms of their dynamical core, which describe the resolved scales of motion, and
 273 the physical parameterizations, which provide estimates of the grid-scale effect of processes which cannot be
 274 resolved by the dynamical core. This general approach has been hugely successful in that nowadays predictions
 275 of weather and climate are made routinely. On the other hand, exactly through these predictions it has become
 276 apparent that uncertainty estimates produced by current state-of-the-art models still have shortcomings.

277 One shortcoming is that many physical parameterizations are based on bulk formula which are based on the
 278 assumption that the subgrid-scale state is in equilibrium with the resolved state¹¹⁸. Model errors might arise
 279 from a misrepresentation of unresolved subgrid-scale processes which can affect not only the variability, but also
 280 the mean error of a model^{129,141}. An example in a comprehensive climate model is e.g., the bias in the 500hPa
 281 geopotential height pattern, which is reduced when the representation of the subgrid-state is refined⁷ (Fig. 2).

282 In recent years, methods for the estimation of flow-dependent uncertainty in predictions have become an
 283 important topic. Ideally, uncertainties should be estimated within the physical parameterizations and uncer-
 284 tainty representations should be developed alongside the model. Many of these methods are “ad hoc” and
 285 added *a posteriori* to an already tuned model. Only first steps to develop uncertainty estimates from within
 286 the parameterizations have been attempted^{20,131}.

287 The representation of model-error in weather and climate models falls in one of two major categories:
288 Multi-model approaches and stochastic parameterizations. In the multi-model approach each ensemble mem-
289 ber consists of an altogether different model. The models can differ in the dynamical core and the physical
290 parameterizations^{55,63,75} or use the same dynamical core but utilize either different static parameters in their
291 physical parameterizations¹⁵¹ or altogether different physics packages^{6,31,106,152}. Both approaches have been
292 successful in improving predictions of weather and climate over a range of scales, as well as their uncertainty.
293 Multi-model ensembles provide more reliable seasonal forecasts¹²² and are commonly used for the uncertainty
294 assessment of climate change predictions e.g., as in the Assessment Report 5 of the Intergovernmental Panel
295 on Climate Change (IPCC)¹⁵⁷. Stochastic parameterizations are routinely used to improve the reliability of
296 weather forecasts in the short-⁶ and medium-range^{5,10,123} as well as for seasonal predictions^{4,28,165}.

297 In the stochastic approach, the effect of uncertainties due to the finite truncation of the model are treated
298 as independent realizations of stochastic processes that describe these truncation uncertainties. This treatment
299 goes back to the idea of stochastic-dynamic prediction^{33,118,130}. While the verdict is still open if subgrid-scale
300 fluctuations must be included explicitly via a stochastic term, or if it is sufficient to include their mean influence
301 by improved deterministic physics parameterizations, one advantage of stochastically perturbed models is that all
302 ensemble members have the same climatology and model bias; while for multi-parameter, multi-parameterization
303 and multi-model ensembles each ensemble member is *de facto* a different model with its own dynamical attractor.
304 For operational centers the maintenance of different parameterizations requires additional resources and due to
305 the different biases makes post-processing very difficult.

306 5.1 Stochastic Parameterizations in Numerical Weather Prediction

307 Due to the chaotic nature of the dynamical equations governing the evolution of weather, forecasts are sensitive
308 to the initial condition limiting the intrinsic predictability of the weather system^{82,84}. Probabilistic forecasts
309 are performed by running ensemble systems, where each member is initialized from a different initial state and
310 much effort has gone into the optimal initialization of such ensemble systems^{63,108,161}. Nevertheless state-of-
311 the-art numerical weather predictions systems continue to produce unreliable and over-confident forecasts¹⁴.
312 Consequently, the other source of forecast uncertainty – model-error – has received increasing attention^{117,118}.
313 Since for chaotic systems model-error and initial condition error will both result in trajectories that will diverge
314 from the truth, it is very difficult to disentangle them¹⁴⁹.

315 The first stochastic parameterization used in an operational numerical weather prediction model was the
316 stochastically perturbed physics tendency scheme (SPPT), sometimes also referred to as stochastic diabatic
317 tendency or Buizza-Miller-Palmer (BMP) scheme¹³. SPPT is based on the notion that – especially as the
318 horizontal resolution increases – the equilibrium assumption no longer holds and the subgrid-scale state should
319 be sampled rather than represented by the equilibrium mean. Consequently, SPPT multiplies the accumulated
320 physical tendencies of temperature, wind and moisture at each grid-point and time step with a random pattern
321 that has spatial and temporal correlations. In other words, SPPT assumes that parameterization uncertainty can
322 be expressed as a multiplicative noise term. Ensemble systems perturbed with the SPPT scheme show increased
323 probabilistic skill mostly due to increased spread in short and medium-range ensemble forecasts^{8,13,132,159}.

324 A second successful stochastic parameterization scheme, is the so-called stochastic kinetic energy-backscatter
325 scheme (SKEBS) whose origin lies in Large-Eddy Simulation modeling⁹⁸ and has recently been extended to
326 weather and climate scales^{142,143}. The key idea is that energy associated with subgrid processes is injected
327 back onto the grid using a stochastic pattern generator. This method has been successfully used in a number
328 of operational and research forecasts across a range of scales^{5,6,8,10,11,16,138}. Similar to the SPPT scheme,
329 ensemble systems with SKEBS increase probabilistic skill by increasing spread and decreasing the root-mean-
330 square (RMS) error of the ensemble mean forecast. First results of these schemes at a convection-permitting
331 resolution of around 4km report also a positive impact on forecast skill, in particular more reliable precipitation
332 forecasts^{12,133}.

5.2 Stochastic Parameterizations in Climate Models

The use of stochastic parameterization in climate models is still in its infancy. Climate prediction uncertainty assessments, e.g., IPCC¹⁵⁰, are almost exclusively based on multi-models, mostly from different research centers. Part of the problem is that on climate timescales, limited data for verification exists. A second reason is that on longer timescales, bias is a major source of uncertainty and traditional multi-models are very efficient at sampling biases, although such an experiment is poorly designed for an objective and reliable uncertainty assessment.

However, in recent years first studies have emerged which demonstrate the ability of stochastic parameterizations to reduce longstanding biases and improve climate variability (see Fig. 2 for an example). Jin and Neelin⁸¹ developed a stochastic convective parameterization that includes a random contribution to the convective available potential energy (CAPE) in the deep convective scheme. They find that adding convective noise results in enhanced eastward propagating, low-wavenumber low-frequency variability. Berner et al.⁷ investigate the impact of SKEBS on systematic model-error and report an improvement in the representation of convectively-coupled waves leading to a reduction in the tropical precipitation bias. Furthermore, Majda and colleagues developed systematic stochastic multi-cloud parameterizations for organized convection^{47,67,94,95}. The multi-cloud approach is based on the assumption that organized convection involves three types of clouds and the evolution from one cloud type to another can be described by a transition matrix.

A longstanding systematic error of climate models is the underestimation of the occurrence of Northern Hemispheric blocking. Stochastic parameterizations have been demonstrated to be one way to increase their frequency^{4,28,53,165}, although, e.g. increasing horizontal resolution, leads to similar improvements^{7,24}. This suggests that while it might be necessary to include subgrid-scale variability in some form, the details of this representation might not matter. On the other side, Berner et al.⁷ argue that this degeneracy of response to different subgrid-scale forcings warrants a cautionary note: namely that a decrease in systematic error might not necessarily occur for the right dynamical reasons. The opposite holds true, as well: Due to the necessary tuning of parameters in the parameterizations of comprehensive climate models, an improvement in the formulation of a physical process might not immediately lead to an improved model performance. A striking example of compensating model errors is described in Palmer and Weisheimer¹¹⁹, who report how an inadequate representation of horizontal baroclinic fluxes resulted in a model error that was equal and opposite to the systematic error caused by insufficiently represented vertical orographic gravity wave fluxes. Improvements to wave drag parameterization without increasing resolution unbalanced the compensating model errors, leading to an increase in systematic model bias.

Williams¹⁷⁵ studied the effect of including a stochastic term in the fluxes between the atmospheric and oceanic components in a coupled ocean-atmosphere model. He reports changes to the time-mean climate and increased variability of the El Nino Southern Oscillation, suggesting that the lack of representing of sub-grid variability in air-sea fluxes may contribute to some of the biases exhibited by contemporary climate models.

On seasonal timescales where sufficient observational data for a probabilistic verification exist, stochastic parameterizations have been reported to increase predictive skill. For example, ensemble forecasts of the sea surface temperatures over the Nino3.4 region showed increased anomaly correlation, decreased bias and decreased root mean square error in coupled ocean-atmosphere models^{4,28,165}.

6 Conclusion

We postulate the use of stochastic-dynamical models for uncertainty assessment and model-error representation in comprehensive Earth-System models. This need arises since even state-of-the-art weather and climate models cannot resolve all necessary processes and scales. Here we reviewed mathematical methods for stochastic climate modeling as well as stochastic subgrid-scale parameterizations and postulate their use for a more systematic strategy of parameterizing unresolved and unrepresented processes.

In the last decade, a number of studies emerged that demonstrate the potential of this approach, albeit

378 applied in an ad hoc manner and tuned to specific applications. Stochastic parameterizations have been shown
379 to provide more skillful weather forecasts than traditional ensemble prediction methods, at least on timescales
380 where verification data exists. In addition, they have been shown to reduce longstanding climate biases, which
381 play an important role especially for climate and climate change predictions.

382 Here we argue, that rather than pushing out the limit of skillful ensemble predictions by a few days, more
383 attention should be given on the assessment of uncertainty (as already proposed, e.g., Smith¹⁴⁹). Ideally, it
384 should be carried out alongside the physical parameterization and dynamical core development and not added
385 *a posteriori*. The uncertainty should be directly estimated from within the parameterization schemes and not
386 tuned to yield a particular model performance, as is current practice. For example, Sapsis and Majda^{139,140}
387 propose a statistical framework which systematically quantifies uncertainties in a stochastic fashion.

388 The fact that according to the last two assessment reports (AR) of the IPCC (AR4¹⁴⁷ and AR5¹⁵⁰) the
389 uncertainty in climate predictions and projections has not decreased may be a sign that we might be reaching
390 the limit of climate predictability, which is the result of the intrinsically nonlinear character of the climate
391 system (as first suggested by Lorenz⁸²).

392 Recently Palmer¹²¹ argued that due to limited computational and energy power resources, predictable scales
393 should be solved accurately, while the unpredictable scales could be represented inaccurately. This strategy is at
394 the core of the systematic mode reduction reviewed here, but has only recently been considered for comprehensive
395 Earth-System Models. Stochastic models focus on the accurate simulation of the large, predictable, scales, while
396 only the statistical properties of the small, unpredictable, scales are captured. This has been demonstrated,
397 e.g. by Franzke and Majda³⁸, Kravtsov et al.⁷⁴, who successfully applied mode reduction strategies to global
398 atmospheric circulation models. They showed that these reduced models consisting of only 10-15 degrees of
399 freedom reproduced many of the important statistics of the numerical circulation models which contained a few
400 hundreds degrees of freedom. Vanden-Eijnden¹⁶³ proposed numerical approaches for multi-scale systems where
401 only the largest scales are explicitly computed and the smaller scales are approximated on the fly.

402 The recent result of Wouters and Lucarini^{177,178} provide a promising path towards a general theory of
403 parametrizations for weather and climate models, and give theoretical support that parameterization schemes
404 should include deterministic, stochastic and non-Markovian (memory) components. Moreover, Wouters and
405 Lucarini's results suggest that there is common ground in developing parameterizations for weather and climate
406 prediction models. Optimal representations of the reduced dynamics based on Ruelle's response theory and
407 the Mori-Zwanzig formalism coincide, thus, providing equal optimal representations of the long-term statistical
408 properties and the finite-time evolution of the slow variables.

409 One exciting future research area is the use of stochastic methods for use in data assimilation, which is already
410 an active field of research^{51,54,65,104,114,115,133}. Stochastic methods have been shown to increase the ensemble
411 spread in data assimilation, leading to a better match between observations and model forecasts^{54,105,133}.
412 A cutting-edge frontier is the use of order moments and memory effects in Kalman filter data assimilation
413 schemes¹¹⁵. Another emerging field is the use of stochastic parameterizations in large climate ensembles, which
414 would allow the comparison of uncertainty estimates based in multi-models to that of stochastically perturbed
415 ones.

416 Our hope is that basing stochastic-dynamic prediction on sound mathematical and statistical physics con-
417 cepts will lead to substantial improvements, not only in our ability to accurately simulate weather and climate,
418 but even more importantly to give proper estimates on the uncertainty of these predictions.

419 **Acknowledgments:**

420 CLEF is supported by the German Research Foundation (DFG) through the cluster of Excellence CliSAP, TJO
421 is an Australian Research Council Future Fellow, PDW acknowledges a University Research Fellowship from
422 the Royal Society (UF080256) and VL funding from the European Research Council (NAMASTE).

References

- [1] Achatz, U. and G. Branstator, 1999: A two-layer model with empirical linear corrections and reduced order for studies of internal climate variability. *J. Atmos. Sci.*, 56, 3140-3160.
- [2] Achatz, U. and J. D. Opsteegh, 2003: Primitive-equation-based low-order models with seasonal cycle. Part I: Model construction. *J. Atmos. Sci.*, 60, 465-477.
- [3] Berner, J., 2005: Linking nonlinearity and non-Gaussianity of planetary wave behavior by the Fokker-Planck equation. *J. Atmos. Sci.*, 62, 2098-2117.
- [4] Berner, J., F. J. Doblas-Reyes, T. N. Palmer, G. Shutts and A. Weisheimer, 2008: Impact of a quasi-stochastic cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model. *Phil. Trans. R. Soc. A*, **366**, 2559–2577.
- [5] Berner, J., G. Shutts, M. Leutbecher, and T. Palmer, 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.*, **66**, 603–626.
- [6] Berner, J., S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Snyder, 2011: Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multi-physics representations. *Mon. Wea. Rev.*, **139**, 1972–1995.
- [7] Berner, J., T. Jung and T. N. Palmer, 2012: Systematic model error: The impact of increased horizontal resolution versus improved stochastic and deterministic parameterizations. *J. Climate*, **25**, 4946–4962.
- [8] Berner, J., K. R. Smith, S.-Y. Ha, J. P. Hacker, and C. Snyder, 2014: Increasing the skill of probabilistic forecasts: Model-error representations versus calibration and debiasing. *Mon. Wea. Rev.*, *submitted*.
- [9] Birner, T. and P. D. Williams, 2008: Sudden Stratospheric Warmings as Noise-Induced Transitions. *J. Atmos. Sci.*, 65, 3337-3343.
- [10] Bowler, N. E., A. Arribas, S. E. Beare, K. R. Mylne, and G. J. Shutts, 2009: The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **135**, 767–776.
- [11] Bowler, N. E., A. Arribas, K. R. Mylne, K. B. Robertson, and S. E. Beare, 2008: The MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **134**, 703–722.
- [12] Bouttier, F., B. Vié, O. Nuissier, and L. Raynaud, 2012: Impact of stochastic physics in a convection-permitting ensemble. *Monthly Weather Review*, **140**, 3706–3721.
- [13] Buizza, R., M. Miller and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.*, 125, 2887-2908.
- [14] Buizza, R., P. L. Houtekamer, Z. Toth, G. Pellerini, M. Wei and Y. Zhu, 2005: A comparison of the ECMWF, MSC and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, 133, 1076-1097.
- [15] Carnevale G. F. and J. S. Frederiksen, 1983: Viscosity renormalization based on direct-interaction closure. *J. Fluid Mech.*, 131, 289.
- [16] Charron, M., G. Pellerin, L. Spacek, P. L. Houtekamer, N. Gagnon, H. L. Mitchell, and L. Michelin, 2010: Toward random sampling of model error in the canadian ensemble prediction system. *Mon. Wea. Rev.*, **138**, 1877–1901.
- [17] Chekroun, M., H. Liu and S. Wang, 2013: On stochastic parameterizing manifolds: Pullback characterization and non-Markovian reduced equations. arxiv:1310.3896
- [18] Chorin, A. J., O. H. Hald and R. Kupferman, 2002: Optimal prediction with memory. *Physica D*, **166**, 239-257.
- [19] Chorin, A. J. and O. H. Hald, 2013: Stochastic tools in mathematics and science. Springer, New York, 208pp.

- 462 [20] Cohen, B. G. and G. C. Craig, 2005: Fluctuations in a equilibrium convective ensemble. Part II: Numerical
463 experiments. *J. Atmos. Sci.*, **63**, 2005-2015.
- 464 [21] Crommelin, D. and E. Vanden-Eijnden, 2006: Reconstruction of diffusions using spectral data from timeseries.
465 *Comm. Math. Sci.*, 4, 651-668.
- 466 [22] Crommelin, D. and E. Vanden-Eijnden, 2008: Subgrid-scale parameterization with conditional Markov chains. *J.*
467 *Atmos. Sci.*, 65, 2661-2675.
- 468 [23] Culina, J., S. Kravtsov and A. H. Monahan, 2011: Stochastic parameterization schemes for use in realistic climate
469 models. *J. Atmos. Sci.*, 68, 284-299.
- 470 [24] Dawson, A., T. N. Palmer and S. Corti, 2012: Simulating regimes structures in weather and climate prediction
471 models. *Geophys. Res. Lett.*, 39, L21805.
- 472 [25] Deker U. and F. Haake, 1957: Fluctuation-dissipation theorems for classical processes. *Phys. Rev. A*, 6, 2043.
- 473 [26] DelSole, T. and A. Y. Hou, 1999: Empirical stochastic models for the dominant climate statistics of a General
474 Circulation Model. *J. Atmos. Sci.*, 56, 3436-3456.
- 475 [27] DelSole, T., 2004: Stochastic models of quasigeostrophic turbulence. *Surveys Geophys.*, 25, 107-149.
- 476 [28] Doblas-Reyes, F., A. Weisheimer, M. Déqué, N. Keenlyside, M. McVean, J.M. Murphy, P. Rogel, D. Smith, T.N.
477 Palmer, 2009: Addressing model uncertainty in seasonal and annual dynamical seasonal forecasts. *Quart. J. Roy.*
478 *Meteor. Soc.*, **135**, 1538–1559.
- 479 [29] Dolaptchiev, S. U. Achatz and I. Timofeyev, 2013: Stochastic closure for local averages in the finite-difference
480 discretization of the forced Burgers equation. *Theor. Comput. Fluid Dyn.*, **27**, 297-317.
- 481 [30] Douglas, J. F. and J. M. K. Gasiorek, 2000: Fluid Mechanics. Pearsons Higher Education.
- 482 [31] Eckel, F. A. and C. F. Mass, 2005: Aspects of effective mesoscale, short-range ensemble forecasting. *Wea. For.*,
483 20, 328-350.
- 484 [32] Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus*, 21, 739-759.
- 485 [33] Epstein, E. S. and Pitcher E. J., 1972: Stochastic analysis of meteorological fields. *J. Atmos. Sci.*, 29, 244-257.
- 486 [34] Farrell, B. F. and P. J. Ioannou, 1993: Stochastic Dynamics of baroclinic waves. *J. Atmos. Sci.*, 50, 4044-4057.
- 487 [35] Farrell, B. F. and P. J. Ioannou, 1995: Stochastic Dynamics of the midlatitude atmospheric jet. *J. Atmos. Sci.*,
488 52, 1642-1656.
- 489 [36] Frankignoul, C. and K. Hasselmann, 1977: Stochastic climate models, Part II: Application to sea-surface temper-
490 ature anomalies and thermocline variability. *Tellus*, 29, 289-305.
- 491 [37] Franzke, C., A. J. Majda and E. Vanden-Eijnden, 2005: Low-Order Stochastic Mode Reduction for a Realistic
492 Barotropic Model Climate. *J. Atmos. Sci.*, 62, 1722-1745.
- 493 [38] Franzke C. and A. J. Majda, 2006: Low-Order Stochastic Mode Reduction for a Prototype Atmospheric GCM. *J.*
494 *Atmos. Sci.*, 63, 457-479.
- 495 [39] Franzke, C., A. J. Majda and G. Branstator, 2007: The origin of nonlinear signatures of Planetary Wave Dynam-
496 ics: Mean Phase Space Tendencies and Contributions from Non-Gaussianity. *J. Atmos. Sci.*, 64, 3987-4003. doi:
497 10.1175/2006JAS2221.1
- 498 [40] Franzke, C., 2012: Predictability of extreme events in a nonlinear stochastic-dynamical model. *Phys. Rev. E*, 85,
499 doi: 10.1103/PhysRevE.85.031134

- 500 [41] Franzke, C., 2013: Predictions of critical transitions with non-stationary reduced order models. *Physica D*, 262,
501 35-47.
- 502 [42] Frederiksen J. S., 1999: Subgrid-scale parameterizations of eddy-topographic force, eddy viscosity, and stochastic
503 backscatter for flow over topography. *J. Atmos. Sci.*, 56, 1481.
- 504 [43] Frederiksen J. S. and A. G. Davies, 1997: Eddy viscosity and stochastic backscatter parameterizations on the
505 sphere for atmospheric circulation models. *J. Atmos. Sci.*, 54, 2475.
- 506 [44] Frederiksen J.S., A. G. Davies and R. C. Bell, 1994: Closure equations with non-Gaussian restarts for truncated
507 two-dimensional turbulence. *Phys. Fluids*, 6, 3153.
- 508 [45] Frederiksen J. S. and A. G. Davies, 2000: Dynamics and spectra of cumulant update closures for two-dimensional
509 turbulence. *Geophys. Astrophys. Fluid Dynam.*, 92, 197.
- 510 [46] Frederiksen J. S. and T. J. O’Kane, 2005: Inhomogeneous closure and statistical mechanics for Rossby wave
511 turbulence over topography. *J. Fluid Mech.*, 539, 137.
- 512 [47] Frenkel, Y., A. J. Majda and B. Khouider, 2011: Using the stochastic multicloud model to improve tropical
513 convective parameterizations: A paradigm example. *J. Atmos. Sci.*, 69, 1080-1105.
- 514 [48] Gallavotti, G. and E. G. D. Cohen, 1995: Dynamical ensembles in stationary states. *J. Stat. Phys.*, **80**, 931-970.
- 515 [49] Gardiner, C. W., Handbook of stochastic methods, 1985. Springer-Verlag, 442pp.
- 516 [50] Givon, D., P. Kupferman and A. Stuart, 2004: Extracting macroscopic dynamics: model problems and algorithms.
517 *Nonlinearity*, **17**, R55.
- 518 [51] Gottwald, G. and J. Harlim, 2013: The role of additive and multiplicative noise in filtering complex dynamical
519 systems. *Proc. Roy. Soc. A*, **469**, 20130096.
- 520 [52] Gottwald, G. and I. Melbourne, 2013: Homogenization for deterministic maps and multiplicative noise. *Proc. Roy.*
521 *Soc. A*, **469**, 20130201.
- 522 [53] Jung, T., T. N. Palmer, and G. J. Shutts, 2005: Influence of a stochastic parameterization on the frequency of
523 occurrence of north pacific weather regimes in the ECMWF model. *Geophys. Res. Lett.*, **32**, L23811.
- 524 [54] Ha, S.-Y., J. Berner, and C. Snyder, 2014: Model error representation in mesoscale wrf-dart cycling. *to be submitted*
525 *to Mon. Wea. Rev.*.
- 526 [55] Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer, 2005: The rationale behind the success of multi-model
527 ensembles in seasonal forecasting – i. basic concept. *Tellus*, **57A**, 219-233.
- 528 [56] Hasselmann, K., Stochastic Climate Models: Part I. Theory. *Tellus*, 28 (1976), 473-484.
- 529 [57] Harlim, J., A. Mahdi and A. J. Majda, 2013: An ensemble Kalman filter for statistical estimation of physics
530 constrained nonlinear regression models. *J. Compt. Phys.*, 257, 782-812.
- 531 [58] Heinz, S., 2003: Statistical mechanics of turbulent flows. Springer, Berlin, 215pp.
- 532 [59] Herring J. R., 1965: Self-consistent-field approach to turbulence theory. *Phys. Fluids*, 8, 2219.
- 533 [60] Herring J.R. and Kraichnan R.H., 1972: *Statistical Models of Turbulence*. Springer, 1972.
- 534 [61] Hide, R., P. J. Mason and R. A. Plumb, 1977: Thermal convection in a rotating fluid subject to a horizontal
535 temperature gradient: spatial and temporal characteristics of fully developed baroclinic waves. *J. Atmos. Sci.*, 34,
536 930-950.
- 537 [62] Holloway G., 1992: Representing topographic stress for large-scale ocean models. *J. Phys. Oceanogr.*, 22, 1033.

- 538 [63] Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach
539 to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225-1242.
- 540 [64] Imkeller, P. and J.-S. von Storch, 2001: *Stochastic Climate Models*. Birkhäuser, 398 pages.
- 541 [65] Isaksen, L., M. Fisher, and J. Berner, 2007: *Use of analysis ensembles in estimating flow-dependent background*
542 *error variances*. Proceedings of ECMWF Workshop on Flow Dependent Background Errors ECMWF, June 11-13,
543 2007, Reading, United Kingdom, 65–86 pp.
- 544 [66] Khasminsky, R. Z., 1963: Principle of averaging for parabolic and elliptic differential equations and for Markov
545 processes with small diffusions. *Theory Probab. Appl.*, **8**, 1-21.
- 546 [67] Khouider, B., J. Biello and A. J. Majda, 2010: A stochastic multicloud model for tropical convection. *Commun.*
547 *Math. Sci.*, **8**, 187-216.
- 548 [68] Kitsios V., J. S. Frederiksen and M. J. Zidikheri, 2012: Subgrid model with scaling laws for atmospheric flows. *J.*
549 *Atmos. Sci.*, **69**, 1427.
- 550 [69] Klein, R., 2010: Scale-dependent models for atmospheric flows. *Ann. Rev. Fluid Mech.*, **42**, 249-274.
- 551 [70] Kondrashov, D., S. Kravtsov and M. Ghil, 2006: Empirical mode reduction in a model of extratropical low-
552 frequency variability. *J. Atmos. Sci.*, **63**, 1859-1877.
- 553 [71] Kraichnan R. H., 1959: The structure of isotropic turbulence at very high reynolds number. *J. Fluid. Mech.*, **5**,
554 497.
- 555 [72] Kraichnan R. H., 1972: Test-field model for inhomogeneous turbulence. *J. Fluid Mech.*, **56**, 287.
- 556 [73] Kraichnan R. H., 1976: Eddy viscosity in two and three dimensions. *J. Atmos. Sci.*, **33**, 1521.
- 557 [74] Kravtsov, S., D. Kondrashov and M. Ghil, 2005: Multilevel regression modeling of nonlinear processes: derivation
558 and applications to climatic variability. *J. Climate*, **18**, 4404-4424.
- 559 [75] Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, , and S. Suren-
560 dran, 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196-4216.
- 561 [76] Kurtz, T. G., 1973: A limit theorem for perturbed operator semi-groups with applications to random evolutions.
562 *J. Funct. Anal.*, **12**, 55-67.
- 563 [77] Leith C. E., 1971: Atmospheric predictability and two-dimensional turbulence. *J. Atmos. Sci.*, **28**, 145.
- 564 [78] Leith C. E., Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409.
- 565 [79] Leith, C. E., and R. H. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.*, **29**, 1041-1058.
- 566 [80] Leslie D.C. and G. L. Quarini, 1979: The application of turbulence theory to the formulation of subgrid modelling
567 procedures. *J. Fluid Mech.*, **91**, 65.
- 568 [81] Lin, J. W. B. and J. D. Neelin, 2000: Influence of a stochastic moist convective parameterization on tropical
569 climate variability. *Geophys. Res. Lett.*, **27**, 3691–3694.
- 570 [82] Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130-141.
- 571 [83] Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321-333.
- 572 [84] Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**,
573 636-646.
- 574 [85] Lovegrove, A. F., P. L. Read and C. J. Richards, 2000: Generation of inertia-gravity waves in a baroclinically
575 unstable fluid. *Q. J. Roy. Meteorol. Soc.*, **126**, 3233-3254.

- 576 [86] Lemons, D. S., 2002: An introduction to stochastic processes in physics. Johns Hopkins University Press, 115pp.
- 577 [87] Lucarini, V., 2009: Evidence for dispersion relations for the nonlinear response of the Lorenz 1963 system. *J. Stat.*
578 *Phys.*, 134, 381-400.
- 579 [88] Lucarini, V. and S. Sarno 2011: A statistical mechanic approach for the computation of the climatic response to
580 general forcings. *Nonlin. Proc. Geophys.*, 18, 7-28.
- 581 [89] Majda, A. J., I. Timofeyev and E. Vanden-Eijnden, 1999: Models for stochastic climate prediction. *Proc. Nat.*
582 *Acad. Sci. USA*, 96, 14687-14691.
- 583 [90] Majda, A. J., I. Timofeyev and E. Vanden-Eijnden, 2001: A mathematical framework for stochastic climate models.
584 *Commun. Pure Appl. Math.*, 54, 891-974.
- 585 [91] Majda, A. J., I. Timofeyev and E. Vanden-Eijnden, 2002: A priori tests of a stochastic mode reduction strategy.
586 *Physica D*, 170, 206-252.
- 587 [92] Majda, A. J., I. Timofeyev and E. Vanden-Eijnden, 2003: Systematic strategies for stochastic mode reduction in
588 climate. *J. Atmos. Sci.*, 60, 1705-1722.
- 589 [93] Majda, A. J., R. Abramov and M. Grote, 2005: Information Theory and Stochastics for Multiscale Nonlinear
590 Systems. CRM Monograph Series, American Mathematical Society.
- 591 [94] Majda, A. J., C. Franzke and B. Khouider, 2008: An applied mathematics perspective on stochastic modelling for
592 climate. *Phil. Trans. R. Soc. A*, 366, 2429-2455.
- 593 [95] Majda, A. J. and S. Stechmann, 2008: Stochastic models for convective momentum transport. *Proc. Nat. Acad.*
594 *USA*, 105, 17614-17619.
- 595 [96] Majda, A. J., C. Franzke and D. Crommelin, 2009: Normal forms for reduced stochastic climate models. *Proc.*
596 *Nat. Acad. Sci. USA*, 106, 3649-3653. doi: 10.1073/pnas.0900173106
- 597 [97] Majda, A. J. and J. Harlim, 2013: Physics constrained nonlinear regression models for time series. *Nonlinearity*,
598 26, 201-217.
- 599 [98] Mason, P. and D. Thomson, 1992: Stochastic backscatter in large-eddy simulations of boundary layers. *J. Fluid*
600 *Mech.*, 242, 51-78.
- 601 [99] McComb W. D., 1974: A local energy-transfer theory of isotropic turbulence. *J. Phys. A*, 7, 632.
- 602 [100] McComb W. D. A. Hunte and C. Johnson, 2001a: Conditional mode elimination and the subgrid-modelling
603 problem for isotropic turbulence. *Phys. Fluids*, 13, 2030.
- 604 [101] Melbourne, I. and A. M. Stuart, 2011: A note on diffusion limits of chaotic skew-product flows. *Nonlinearity*, 24,
605 1361-1367.
- 606 [102] Métais O. and M. Lesieur, 1986: Statistical predictability of decaying turbulence. *J. Atmos. Sci.*, 43, 857.
- 607 [103] Monahan, A. H. and J. Culina, 2011: Stochastic averaging of idealized climate models. *J. Climate*, 24, 3068-3088.
- 608 [104] Miller, R. N., E. F. Carter and S. T. Blue, 1999: Data assimilation into nonlinear stochastic models. *Tellus*, 51A,
609 167-194.
- 610 [105] Mitchell, L. and G. A. Gottwald, 2012: Data assimilation in slow-fast systems using homogenized climate models.
611 *J. Atmos. Sci.*, 69, 1359-1377.
- 612 [106] Murphy, J. M., D. M. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M Collins and D. A. Stainforth, 2004:
613 Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, 430, 768-772.

- 614 [107] Memin, E., 2014: Fluid flow dynamics under location uncertainty. *Geophys. Astrophys. Fluid Dyn.*, **108**, 119-146.
- 615 [108] Molteni, F. and T. N. Palmer, 1993: Predictability and finite-time instability of the Northern winter circulation.
616 *Quart. J. Roy. Meteor. Soc.*, **119**, 269-298.
- 617 [109] Ogura Y., 1963: A consequence of the zero fourth cumulant approximation in the decay of isotropic turbulence.
618 *J. Fluid Mech.*, 16, 33.
- 619 [110] O’Kane T. J., R. E. Scholten, M. R. Walkiewicz and P. M. Farrell, 1999: Nonlinear interactions of multi-level
620 atoms with a near resonant standing wave. *Phys. Rev. A*, 59, 4485-4493.
- 621 [111] O’Kane T. J. and J. S. Frederiksen, 2004: The QDIA and regularized QDIA closures for inhomogeneous turbulence
622 over topography. *J. Fluid Mech.*, 504, 133.
- 623 [112] O’Kane T. J. and J. S. Frederiksen, 2008a: Statistical dynamical subgrid-scale parameterizations for geophysical
624 flows. *Phys. Scr.*, T132, 014033.
- 625 [113] O’Kane T. J. and J. S. Frederiksen, 2008b: A comparison of statistical dynamical and ensemble prediction methods
626 during blocking. *J. Atmos. Sci.*, 65, 426.
- 627 [114] O’Kane T. J. and J. S. Frederiksen, 2008c: Comparison of statistical dynamical, square root and ensemble Kalman
628 filters. *Entropy*, 10, 684-721. DOI: 10.3390/e10040684.
- 629 [115] O’Kane T. J. and J. S. Frederiksen, 2010: The application of statistical dynamical turbulence closures to data
630 assimilation. *Phys. Scr.*, T142, 014042.
- 631 [116] Orszag S. A., 1970: Analytical theories of turbulence. *J. Fluid Mech.* 41, 363.
- 632 [117] Palmer, T. N., 1999: A nonlinear dynamical perspective on climate prediction. *J. Climate*, **12**, 575-591.
- 633 [118] Palmer, T. N., 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-
634 dynamic parameterization in weather and climate prediction. *Quart. J. Roy. Meteor. Soc.*, **127**, 279-304.
- 635 [119] Palmer, T. N., and A. Weisheimer, 2011: Diagnosing the causes of bias in climate models - why is it so hard?
636 *Geophysical & Astrophysical Fluid Dynamics*, **105**, 351-365.
- 637 [120] Palmer, T. N. and P. D. Williams, 2008: Introduction. Stochastic physics and climate modelling. *Phil. Trans. Roy.*
638 *Soc. A*, **366**(1875), 2421-2427.
- 639 [121] Palmer, T., 2014: More reliable forecasts with less precise computations: a fast-track route to cloud-resolved
640 weather and climate simulators? *Phil. Trans. A*, **372**, 20130391.
- 641 [122] Palmer, T., U. Andersen, P. Cantelaube, M. Davey, M. Deque, F. Doblas-Reyes, H. Feddersen, R. Graham,
642 S. Gualdi, J.-F. Gueremy, et al., 2004: Development of a European multi-model ensemble system for seasonal to
643 inter-annual prediction (DEMETER). *Bulletin of the American Meteorological Society*, **85**, 853-872.
- 644 [123] Palmer, T. N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. Shutts, M. Steinheimer, and A. Weisheimer,
645 2009: Stochastic Parametrization and Model Uncertainty. *ECMWF Technical Memorandum*, **598**, available at
646 <http://www.ecmwf.int/publications/>.
- 647 [124] Papanicolaou, G. C., 1976: Some probabilistic problems and methods in singular perturbations. *Rocky Mountain*
648 *J. Math.*, 6, 653-674.
- 649 [125] Pavliotis, G. A. and A. M. Stuart, 2008: *Multiscale Methods: Averaging and Homogenization*. Springer, 307pp.
- 650 [126] Peavoy, D., C. Franzke and G. Roberts, 2014: Systematic Physics Constrained Parameter Estimation of Stochastic
651 Differential Equations. *Comp. Stat. Data Anal.*, accepted pending minor revisions.

- 652 [127] Penland, C. and L. Matrosova, 1994: A balance condition for stochastic numerical models with application to the
653 El Nino-Southern Oscillation. *J. Climate.*, 7, 1352-1372.
- 654 [128] Penland, C. and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J.*
655 *Climate.*, 8, 1999-2014.
- 656 [129] Penland, C., 2003: Noise out of chaos and why it won't go away. *Bull. Amer. Meteorol. Soc.*, 84, 921-925.
- 657 [130] Pitcher, E. J., 1977: Application of stochastic dynamic prediction to real data. *J. Atmos. Sci.*, 34, 3-21. Imer
- 658 [131] Plant, R. S. and G. C. Craig, 2008: A stochastic parameterization for deep convection based on equilibrium
659 statistics. *J. Atmos. Sci.*, 65, 87-105.
- 660 [132] Reynolds, C. A., J. G. McLay, J. S. Goerss, E. A. Serra, D. Hodyss and C. R. Sampson, 2011: Impact of resolution
661 and design on the U.S. Navy global ensemble performance in the tropics. *Mon. Wea. Rev.*, 139, 2145-2155.
- 662 [133] Romine, G. S., C. S. Schwartz, J. Berner, K. R. Smith, C. Snyder, J. L. Anderson, and M. L. Weisman, 2014:
663 Representing forecast error in a convection-permitting ensemble system. *Mon. Wea. Rev.*, *submitted*.
- 664 [134] Rose H. A., 1977: Eddy diffusivity, eddy noise and subgrid-scale modelling. *J. Fluid Mech.*, 81, 719.
- 665 [135] Rose H. A., 1985: An efficient non-Markovian theory of non-equilibrium dynamics. *Physica D*, 14, 216.
- 666 [136] Ruelle, D., 1999: Smooth dynamics and new theoretical ideas from nonequilibrium statistical mechanics. *J. Stat.*
667 *Phys.*, 95, 393-468.
- 668 [137] Ruelle, D., 2009: A review of linear response theory for general differentiable dynamical systems. *Nonlinearity*, 22,
669 855.
- 670 [138] Sanchez, C., K. D. Williams, G. Shutts, and M. Collins, 2014: Impact of a stochastic kinetic energy backscatter
671 scheme across time-scales and resolutions. *Q.J.R. Meteorol. Soc.*, doi: 10.1002/qj.2328, 577-599.
- 672 [139] Sapsis, T. and A. J. Majda, 2013a: A statistically accurate modified quasilinear Gaussian closure for uncertainty
673 quantification in turbulent dynamical systems. *Physica D*, 252, 34-45.
- 674 [140] Sapsis, T. and A. J. Majda, 2013b: Statistically accurate low-order models for uncertainty quantification in
675 turbulent dynamical systems. *Proc. Nat. Acad. USA*, 110, 13705-13710.
- 676 [141] Sardeshmukh, P., C. Penland and M. Newman, 2001: Rossby waves in a stochastically fluctuating medium.
677 *Stochastic Climate Models*, Birkhäuser Basel, 396-384.
- 678 [142] Shutts, G. J., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy.*
679 *Meteor. Soc.*, **612**, 3079-3102.
- 680 [143] Shutts, G. J. and T. N. Palmer, 2003: The use of high resolution numerical simulations of tropical circulation to
681 calibrate stochastic physics schemes. *ECMWF/CLIVAR workshop on simulation and prediction of intra-seasonal*
682 *variability with the emphasis on the MJO*, ECMWF, available from <http://www.ecmwf.int/publications>, 83-102.
- 683 [144] Shutts, G. J., and T. N. Palmer, 2007: Convective Forcing Fluctuations in a Cloud-Resolving Model: Relevance
684 to the Stochastic Parameterization Problem. *J. Climate*, **20**, 187-202.
- 685 [145] Shutts, G., T. Allen, and J. Berner, 2008: Stochastic parametrization of multiscale processes using a dual-grid
686 approach. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*,
687 **366**, 2623-2639.
- 688 [146] Siegert, S., R. Friedrich and J. Peinke, 1998: Analysis of data sets of stochastic systems. *Phys. Lett. A*, 243,
689 275-280.

- 690 [147] Solomon, S., E., D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, 2007: *Climate*
691 *Change 2007: The Physical Science Basis*. Cambridge University Press, New York, 996pp pp.
- 692 [148] Smagorinsky J., 1963: General circulation experiments with the primitive equations. I. The basic experiment.
693 *Mon. Wea. Rev.*, 91, 99.
- 694 [149] Smith, L. A., 2001: *Disentangling uncertainty and error: On the predictability of nonlinear systems*. Springer.
- 695 [150] Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M.
696 Midgley, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, Cambridge, 1535
697 pp.
- 698 [151] Stainforth, D., T. Aina , C. Christensen , M. Collins , N. Faull , D. J. Frame , J. A. Kettleborough , S. Knight , A.
699 Martin , J. M. Murphy , C. Piani , D. Sexton , L. A. Smith , R. A. Spicer , A. J. Thorpe and M. R. Allen, 2005:
700 Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406.
- 701 [152] Stensrud, D. J., J.-W. Bao and T. T. Warner, 2000: Using initial condition and model physics perturbations in
702 short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, 128, 2077-2107.
- 703 [153] Stensrud, D. J., 2007: Parameterization Schemes: Keys to understanding Numerical Weather prediction models.
704 Cambridge University Press, Cambridge, 459pp.
- 705 [154] Stinis, P., 2006: A comparative study of two stochastic mode reduction methods. *Physica D*, 213, 197-213.
- 706 [155] Sura, P., 2011: A general perspective of extreme events in weather and climate. *Atmos. Res.*, 101, 1-21.
- 707 [156] Sura, P., 2013: Stochastic models of climate extremes: Theory and observations. In *Extremes in a changing world:*
708 *detection, analysis and uncertainty*. Eds.: AghaKouchak et al., Springer, 432pp.
- 709 [157] Stocker, E. T. F., D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and
710 P. Midgley, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, Cambridge,
711 United Kingdom and New York.
- 712 [158] De Swart, H. E. and J. Grasman, 1987: Effect of stochastic perturbations on a low-order spectral model of the
713 atmospheric circulation. *Tellus A*, 39, 10-24.
- 714 [159] Teixeira, J. and C. A. Reynolds, 2008: Stochastic nature of physical parameterization in ensemble prediction: a
715 stochastic convection approach. *Mon. Wea. Rev.*, 136, 483-496.
- 716 [160] Tian, Y., E. R. Weeks, K. Ide, J. S. Urbach, C. N. Baroud, M. Ghil and H. L. Swinney, 2001: Experimental and
717 numerical studies of an eastward jet over topography. *J. Fluid Mech.*, 438, 129-157.
- 718 [161] Toth, Z. and E. Kalnay, 1993: Ensemble forecasting at nmc: The generation of perturbations. *Bull. Amer. Meteor.*
719 *Soc.*, **74**, 2317–2330.
- 720 [162] Tribbia, J. J. and D. P. Baumhefner, 2004: Scale interactions and atmospheric predictability: an updated per-
721 spective. *Mon. Wea. Rev.*, 132, 703-713.
- 722 [163] Vanden-Eijnden, E., 2003: Numerical techniques for multi-scale dynamical systems with stochastic effects. *Comm.*
723 *Math. Sci.*, 1, 385-391.
- 724 [164] Weeks, E. R., Y. Tian, J. S. Urbach, K. Ide, H. L. Swinney and M. Ghil, 1997: Transitions Between Blocked and
725 Zonal Flows in a Rotating Annulus with Topography. *Science*, 278, 1598-1601.
- 726 [165] Weisheimer, A., T. Palmer, and F. Doblas-Reyes, 2011: Assessment of representations of model uncertainty in
727 monthly and seasonal forecast ensembles. *Geophysical Research Letters*, **38**.
- 728 [166] Whitaker, J. S. and P. D. Sardeshmukh, 1998: A linear theory of extratropical synoptic eddy statistics. *J. Atmos.*
729 *Sci.*, 55, 237-258.

- 730 [167] Wilks, D. S., 2005: Effects of stochastic parameterizations in the Lorenz'96 system. *Q. J. R. Meteorol. Soc.*, 131,
731 389-407.
- 732 [168] Williams, P. D., P. L. Read and T. W. N. Haine, 2003: Spontaneous generation and impact of inertia-gravity
733 waves in a stratified, two-layer shear flow. *Geophys. Res. Lett.*, 30, 2255.
- 734 [169] Williams, P. D., T. W. N. Haine and P. L. Read, 2004a: Stochastic resonance in a nonlinear model of a rotating,
735 stratified shear flow, with a simple stochastic inertia-gravity wave parameterization. *Nonlinear. Proc. Geophys.*,
736 11, 127-135.
- 737 [170] Williams, P. D., P. L. Read and T. W. N. Haine, 2004b: A calibrated, non-invasive method for measuring the
738 internal interface height field at high resolution in the rotating, two-layer annulus. *Geophys. Astrophys. Fluid
739 Dyn.*, 6, 453-471.
- 740 [171] Williams, P. D., 2005: Modelling climate change: the role of unresolved processes. *Phil. Trans. Roy. Soc. A*,
741 363(1837), 2931-2946.
- 742 [172] Williams, P. D., T. W. N. Haine and P. L. Read, 2005: On the generation mechanisms of short-scale unbalanced
743 modes in rotating two-layer flows with vertical shear. *J. Fluid Mech.*, 528, 1-22.
- 744 [173] Williams, P. D., T. W. N. Haine and P. L. Read, 2008: Inertia-gravity waves emitted from balanced flow: Obser-
745 vations, properties, and consequences. *J. Atmos. Sci.*, 65, 3543-3556.
- 746 [174] Williams, P. D., T. W. N. Haine, P. L. Read, S. R. Lewis and Y. H. Yamazaki, 2009: QUAGMIRE v1.3: a
747 quasi-geostrophic model for investigating rotating fluids experiments. *Geoscientific Model Development*, 2, 13-32.
- 748 [175] Williams, P. D., 2012: Climatic impacts of stochastic fluctuations in air-sea fluxes. *Geophysical Research Letters*,
749 **39**.
- 750 [176] Winkler, C. R., M. Newman and P. D. Sardeshmukh, 2001: A linear model of wintertime low-frequency variability.
751 Part I: Formulation and forecast skill. *J. Climate*, 14, 4474-4494.
- 752 [177] Wouters, J. and V. Lucarini, 2012: Disentangling multi-level systems: averaging, correlations and memory. *J. Stat.*
753 *Mech.*, PO3003.
- 754 [178] Wouters, J. and V. Lucarini, 2013: Multi-level dynamical systems: Connecting the Ruelle response theory and
755 the Mori-Zwanzig approach. *J. Stat. Phys.*, 151, 850-860.
- 756 [179] Zhang, Y. and I. M. Held, 1999: A linear stochastic model of a GCM's midlatitude storm tracks. *J. Atmos. Sci.*,
757 56, 3416-3435.
- 758 [180] Zidikheri M. J. and J. S. Frederiksen, 2009: Stochastic subgrid parameterizations for simulations of atmospheric
759 baroclinic flows. *J. Atmos. Sci.*, 66, 2844.

760 **List of Figures**

761 1 Regime transitions in a rotating two-layer annulus laboratory experiment, viewed from above.
762 Different colours correspond to different internal interface heights, through the use of a sophisti-
763 cated visualisation technique¹⁷⁰. In the upper row, small-scale inertia-gravity waves are absent,
764 and large-scale regime transitions do not occur. In the lower row, small-scale inertia-gravity waves
765 are present locally in the troughs of the large-scale wave, and a large-scale regime transition does
766 occur. From the laboratory experiments of Williams et al.^{168,169,172,173}. 22

767 2 Mean systematic error of 500 hPa geopotential height fields (shading) for extended boreal winters
768 (December–March) of the period 1962-2005. Errors are defined with regard to the observed mean
769 field (contours), consisting of a combination of ERA-40 (1962-2001) and operational ECMWF
770 analyses (2002-2005). (a) Systematic error in a numerical simulation with the ECMWF model
771 IFS, version CY32R1, run at a horizontal resolution of T_L95 (about 210km) and 91 vertical
772 levels. (b) Systematic error in a simulation with a stochastic kinetic-energy backscatter scheme
773 (SKEBS). Significant differences at the 95% confidence level based on a Student’s t-test are
774 hatched. After Berner et al.⁷. 23

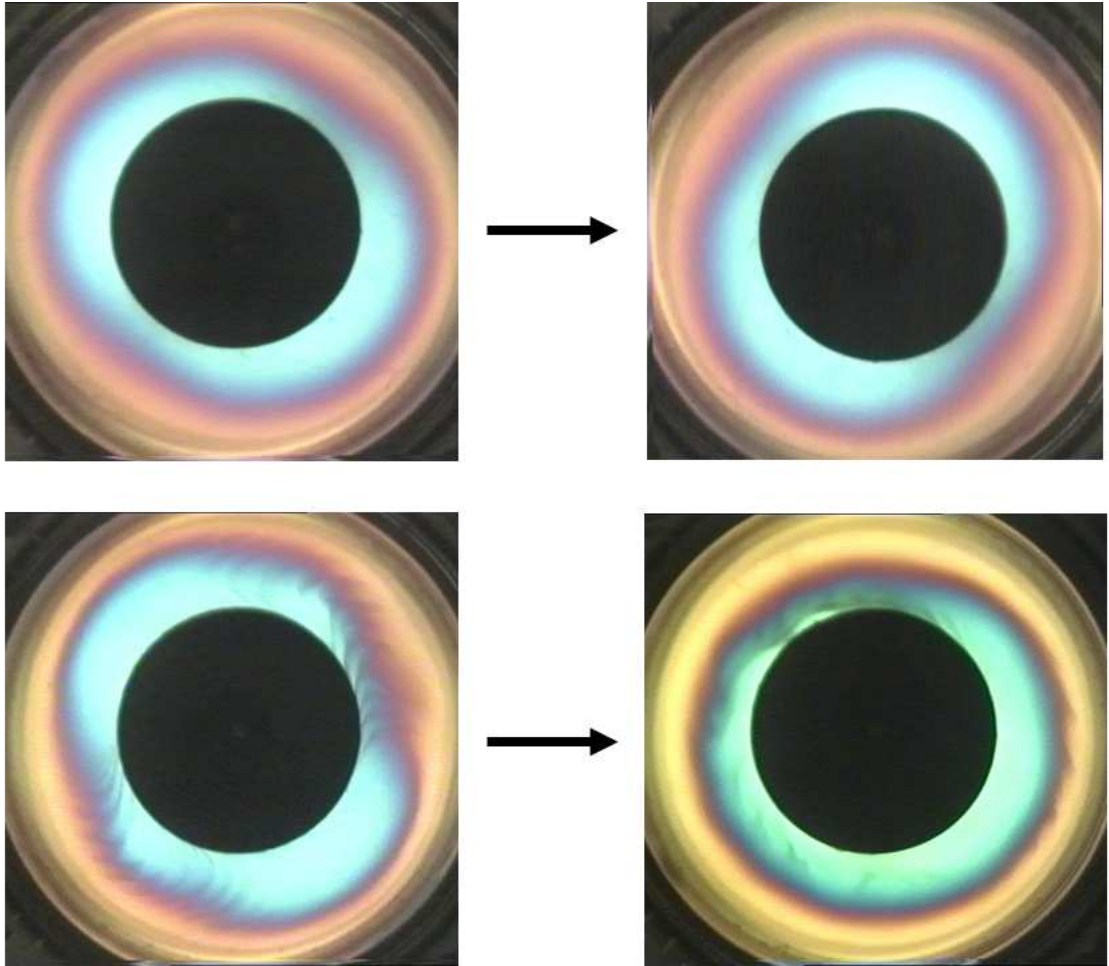


Figure 1: Regime transitions in a rotating two-layer annulus laboratory experiment, viewed from above. Different colours correspond to different internal interface heights, through the use of a sophisticated visualisation technique¹⁷⁰. In the upper row, small-scale inertia-gravity waves are absent, and large-scale regime transitions do not occur. In the lower row, small-scale inertia-gravity waves are present locally in the troughs of the large-scale wave, and a large-scale regime transition does occur. From the laboratory experiments of Williams et al.^{168,169,172,173}.

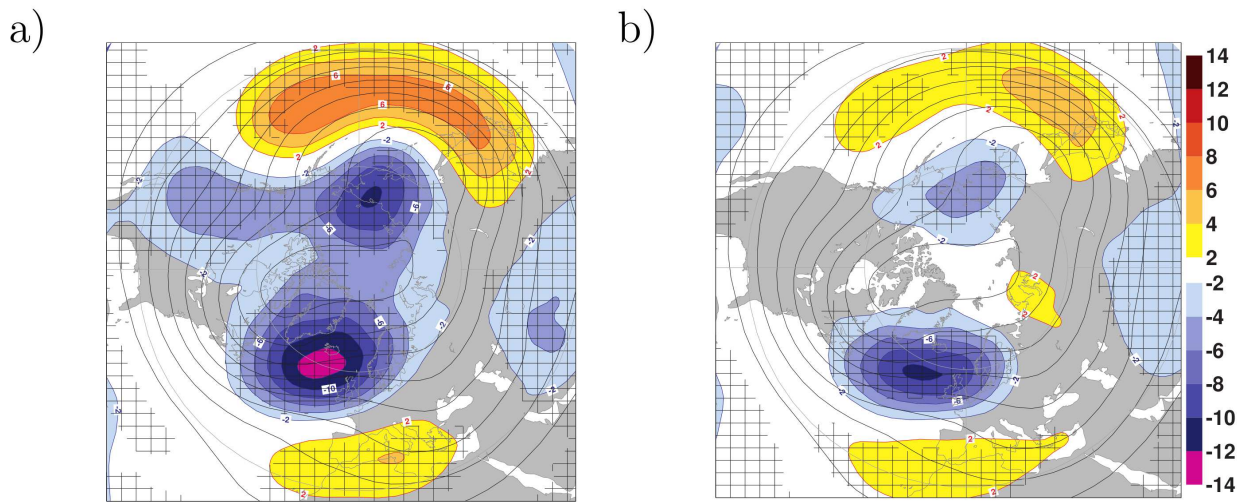


Figure 2: Mean systematic error of 500 hPa geopotential height fields (shading) for extended boreal winters (December–March) of the period 1962–2005. Errors are defined with regard to the observed mean field (contours), consisting of a combination of ERA-40 (1962–2001) and operational ECMWF analyses (2002–2005). (a) Systematic error in a numerical simulation with the ECMWF model IFS, version CY32R1, run at a horizontal resolution of T_L95 (about 210km) and 91 vertical levels. (b) Systematic error in a simulation with a stochastic kinetic-energy backscatter scheme (SKEBS). Significant differences at the 95% confidence level based on a Student's t-test are hatched. After Berner et al.⁷.