New Perspectives on Superparameterization for Geophysical Turbulence

Andrew J. Majda^{a,b}, Ian Grooms^{a,*}

^aCenter for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, 251 Mercer St., New York, NY 10012. ^bCenter for Prototype Climate Modelling, NYU-Abu Dhabi

Center for Prototype Climate Modelling, NTO-Abu

Abstract

This is a research expository paper regarding superparameterization, a class of multi-scale numerical methods designed to cope with the intermittent multi-scale effects of inhomogeneous geophysical turbulence where energy often inverse-cascades from the unresolved scales to the large scales through the effects of waves, jets, vortices, and latent heat release from moist processes. Original as well as sparse space time superparameterization algorithms are discussed for the important case of moist atmospheric convection including the role of multi-scale asymptotic methods in providing self-consistent constraints on superparameterization algorithms and related deterministic and stochastic multi-cloud parameterizations. Test models for the statistical numerical analysis of superparameterization algorithms are discussed both to elucidate the performance of the basic algorithms and to test their potential role in efficient multi-scale data assimilation. The very recent development of grid-free seamless stochastic superparameterization methods for geophysical turbulence appropriate for "eddy-permitting" mesoscale ocean turbulence is presented here including a general formulation and illustrative applications to two-layer quasigeostrophic turbulence, and another difficult test case involving one-dimensional models of dispersive wave turbulence. This last test case has randomly generated solitons as coherent structures which collapse and radiate wave energy back to the larger scales, resulting in strong direct and inverse turbulent energy cascades.

Keywords: multi-scale algorithms, stochastic backscatter, no scale separation, subgridscale closure

1. Introduction

One of the foremost challenges of modern applied mathematics is to guide successful strategies for parameterizing unresolved scales in computational models of multi-scale turbulent systems without scale separation. Examples of such systems include oceanic and atmospheric fluid dynamics, mantle convection, turbulent generation of magnetic fields, confined plasmas, and many others. In many of these systems direct resolution of all relevant scales in numerical simulations is impossible given current computers and will remain so for the foreseeable future. The difficulty is confounded by the need to run large ensembles of simulations for state estimation, and to quantify the uncertainty in predictions.

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 generation of supercomputers. This is particularly important since the effects of rotation, stratification, and moist processes in such anisotropic turbulent flows often cause energy to flow intermittently upscale from small unresolved or marginally resolved scales to affect the largest observed scales [1]. Atmospheric weather and climate processes cover about ten 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

^{*}Corresponding author

Email address: grooms@cims.nyu.edu (Ian Grooms)

Preprint submitted to J. Comput. Phys.

the inertial range of turbulence, scales above this range and up to the planetary scales are dominated by myriad intermittent and anisotropic turbulent processes that cannot be described by traditional closures. For example, atmospheric motions on scales between 100 meters and 100 kilometers show an abundance of processes associated with dry and moist convection, clouds, waves, and boundary layer, topographic, and frontal circulations. Oceanic scales from tens of meters to hundreds of kilometers display a similar range of behaviors, albeit with less prominent phase transitions, and with a density that depends nonlinearly on temperature and salinity.

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 [2–4]. Cloud-resolving models (CRMs) 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 [5], but even these leave a vast frontier of scales, from order 50 km down to the 10 meter scale where inertial range turbulence finally takes over, almost completely unaddressed. Furthermore, the ocean mesoscale contains complex patterns of waves, jets, and vortices which inverse-cascade energy to the large scales while the submesoscale involves fronts and small regions of intense vertical mixing.

The complexity of these problems motivates the development of novel approaches to directly address the multi-scale nature of the problem. In atmospheric modeling, superparameterization (SP), called cloudresolving convection parameterization (CRCP) in its initial application [6–10], uses a horizontally periodic two-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 [6–8, 11]. However, conventional SP is still too expensive to use in large ensemble operational settings. 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 submesoscale eddies in the ocean and gravity wave drag in the atmosphere, as well as other science and engineering problems. However, the particular approach of CRCP is difficult to replicate because of the *ad hoc* nature of its development. Recently however, the authors and collaborators have shown how multi-scale models may be exploited to enable systematic, formal development of SP schemes [12–18].

Filtering or data assimilation is the process of obtaining the best statistical estimate of a natural system from partial or noisy 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 models of noisy observational signals of turbulent nonlinear dynamical systems as well as the statistical accuracy of spatio-temporal discretizations for filtering such systems. See the recent review article [19], as well as the introductory graduate textbook [20] and the many references therein. 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 parameterized 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 grid scale, and a very large dimensional phase 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 realizations through the forward dynamical operator.

The above discussion motivates the need for systematic mathematical ideas in devising algorithms for SP and multi-scale filtering/data-assimilation for large-dimensional turbulent dynamical systems, as well as new types of statistical/stochastic numerical analysis to assess the skill of various proposed algorithms. The general mathematical approach to SP for large-dimensional turbulent dynamical systems advocated here is a four-stage process [15]:

- 1. Multi-scale formulation: A multi-scale physical/mathematical formulation into large scale mean and smaller scale fluctuating components in space-time (for examples, see [12, 13, 21–26] and references therein).
- 2. 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 [10, 14]).
- 3. 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 algorithm in a larger scale model [10, 25]. This can be implemented by replacing more expensive three-dimensional models by much simpler two-dimensional [6, 27] or even cheaper stochastic models [16–18, 28–33].
- 4. A posteriori validations of the 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 algorithm to capture multi-scale interactions.

The multi-scale SP methods discussed here can be contrasted with recent complementary ideas in applied mathematics. In the work of one of the authors and collaborators [12, 13] a theoretical link has been established between SP algorithms and heterogeneous multi-scale methods (HMM) [34–36]. However, as noted in [12, 13], while HMM is proposed as a general method, there are significant differences in the regimes of nonlinear dynamics being modeled by SP algorithms as compared with the actual implemented applications of 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 (three or more orders of magnitude), and rapid local statistical equilibration in time, the skill and success of SP algorithms relies on the intermittency in space and time of physical systems with complex evolving strongly chaotic and quiescent regions, without conditional statistical equilibration of the small scales, and with only modest values of scale separation (less than an order of magnitude). Another related mathematical tool is the so-called gaptooth scheme (37) and references therein). 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 local equilibration, let alone an inertial manifold, as shown in Majda and Grote [14], and more fully in [38]. That work ([14]) introduces a class of mathematical test models for SP that are simple enough to be analyzed with large confidence, 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.

The goal of this paper is to provide a current perspective and overview of recent developments in SP algorithms and their numerical analysis. We also briefly mention connections with multi-scale filtering/data

assimilation. In section 2 we discuss current perspectives on SP for moist convection, an important and difficult contemporary application for climate science and long range weather forecasting [39], as well as the original and traditional arena for SP ideas. Section 3 presents a new general framework for seamless, efficient stochastic SP, discusses some idealized mathematical test models for statistical numerical analysis of stochastic SP and other multiscale algorithms, and reviews very recent formulations and applications of stochastic SP algorithms for geophysical turbulence [16–18] which are appropriate for mesoscale and submesoscale ocean turbulence. Section 4 is a brief concluding discussion.

2. Superparameterization and Moist Atmospheric Convection

As mentioned in the introduction, superparameterization was originally developed to improve the representation of moist atmospheric convection, one of the most significant deficiencies of operational climate and long range weather forecasting models [39]. Here we briefly discuss the original superparameterization (SP) algorithm [6, 7], as well as more recent sparse space time superparameterization (SSTSP) algorithms [10] that reduce the computational effort of the original SP by solving in smaller embedded domains of space-time and utilizing space-time periodic extension while retaining statistical accuracy. We also present an example of using systematic multi-scale modeling [13] to reveal natural algorithmic constraints in SP. Finally, we briefly discuss recent progress in alternative formulations involving deterministic and stochastic multi-cloud parameterizations that can be blended with SP.

2.1. Superparameterization and Sparse Space Time Superparameterization Algorithms

Superparameterization was originally developed as a 'cloud resolving convection parameterization' (CRCP) by Grabowski and Smolarkiewicz [6]. The idea is to use a 2D cloud resolving model (CRM) in each column of a large-scale model to explicitly represent mesoscale and smaller processes and their interactions with the large-scale model. It blends conventional parameterization on a coarse mesh with detailed CRM on a finer mesh. This approach is ideal for modern supercomputers because the CRMs embedded in the large-scale grid are independent of each other and can be run in parallel. The method has yielded promising new results regarding tropical intraseasonal behavior and has potential for many other applications in climate-atmosphere-ocean science [7, 8, 11, 27, 40–49].

There is a crucial difference between applying SP in a climate model (grid length of a few hundreds of km) and in a limited-area mesoscale (LAM) model (grid length of a few tens of km). The key difference is that in the former case the 2D CRM is representing both convective and mesoscale dynamics, while in the latter case only convective dynamics have to be treated by the SP model [50]. This has far-reaching implications. An obvious one is that the outer model time step is much shorter in the case of the LAM model (a couple minutes) than in the climate model (tens of minutes). This is important because a time step of a few minutes is significantly shorter than the lifetime of a single convective cell (typically quoted as 20 minutes [51]). This is not the case for the climate model. A similar argument can be made for the size of the domain of the SP model. In the climate model case, the fact that SP has to treat mesoscale dynamics implies that the domain can be as small as to host just a single chaotic convective cell. While the original SP was designed for coarse meshes of order a few hundred kilometers, sparse space-time superparameterization (SSTSP) is designed for operational coarse meshes of order 10 to 50 km.

2.2. The Original Superparameterization Algorithms

The strategy underlying the SP approach on mesoscales is to consider two distinct models coupled in a particular way. The first is a 3D large-scale flow model (e.g. a LAM or a general circulation model [GCM]). The large-scale model uses a horizontal grid length of a few hundreds of kilometers in a GCM case and a few tens of kilometers in a LAM case. The second model is a 2D cloud-scale model formulated on the x-z plane aligned east-west and embedded in each column of the large-scale model. Cloud-scale and large-scale models use the same vertical grid. The cloud-scale model is applied on a sufficiently fine horizontal grid (e.g. 1 km) in order to permit moist convective dynamics. The cloud-scale model is periodic in the horizontal.

The large-scale model employs moist anelastic equations, for example (see also [10]). The anelastic system of equations can be written compactly as

$$\frac{DU}{Dt} = -\nabla\Pi + kgB + iF_{CS}^{U}$$

$$\nabla \cdot (\rho_{0}U) = 0$$

$$\frac{D\Theta}{Dt} = F_{CS}^{\Theta}$$

$$\frac{DQ_{v}}{Dt} = F_{CS}^{Q_{v}}$$

$$\frac{DQ_{c}}{Dt} = F_{CS}^{Q_{c}}$$

$$\frac{DQ_{p}}{Dt} = F_{CS}^{Q_{p}}.$$
(1)

In (1), U = (U, V, W) is the large-scale fluid velocity in the east, north, and upwards directions, respectively; Θ is potential temperature, and Q_v , Q_c , and Q_p are the mixing ratios for water vapor, cloud condensate (i.e. condensed water carried by the flow), and precipitation (condensed water that falls relative to the air). Here $D/Dt = \partial_t + U \cdot \nabla$; II is the pressure perturbation with respect to a balanced ambient state, normalized by the anelastic reference density ρ_0 . The buoyancy is $B = (\Theta - \Theta_e)/\Theta_0 + \epsilon(Q_v - Q_{ve}) - Q_c - Q_p$, where Θ_e and Q_{ve} are ambient potential temperature and water vapor mixing ratios, Θ_0 is the reference potential temperature profile, $\epsilon + 1$ is the ratio of gas constants of water vapor and dry air. The F_{CS} terms in the right-hand side of (1) represent the cloud-scale model feedback which involve turbulent eddy flux-divergences and other nonlinear averages of small scale fluctuations [12, 13, 21].

The anelastic equations of the cloud-scale model are

$$\frac{d\boldsymbol{u}}{dt} = -\nabla' \pi + \boldsymbol{k}gb + \boldsymbol{i}(s_{\boldsymbol{u}} + f_{LS}^{\boldsymbol{u}})$$

$$\nabla' \cdot (\rho_{0}\boldsymbol{u}) = 0$$

$$\frac{d\theta}{dt} = \frac{\theta_{e}}{T_{e}} \left[\frac{L_{v}}{c_{p}} (CON + DEP) + r \right] + s_{\theta} + f_{LS}^{\theta}$$

$$\frac{dq_{v}}{dt} = -CON - DEP + s_{qv} + f_{LS}^{q_{v}}$$

$$\frac{dq_{c}}{dt} = CON - ACC - AUT + f_{LS}^{q_{c}}$$

$$\frac{dq_{p}}{dt} = \frac{1}{\rho_{0}} \partial_{z} (\rho_{0}v_{t}q_{p}) + ACC + AUT + DEP + f_{LS}^{q_{p}},$$
(2)

where the lowercase symbols have the same meaning as their uppercase counterparts in equation (1). We denote two-dimensionality by $\nabla' = (\partial_x, \partial_z)$ with $d/dt = \partial_t + \mathbf{u} \cdot \nabla'$. The terms f_{LS} represent the large-scale forcing for the cloud-scale model; L_v and c_p are the latent heat of condensation and the specific heat at constant pressure; and θ_e and T_e are the environmental potential temperature and temperature profiles. The sources on the right-hand side of (2) describe the formation of cloud condensate from water vapor (CON), autoconversion of cloud condensate into precipitation (AUT), accretion of cloud condensate by precipitation (ACC), and source/sink of precipitation due to deposition/evaporation of water vapor on/from precipitation particles (DEP). These sources are represented using the simple scheme of Grabowski [52].

We denote the large scale variables by Q (representing a collection of U, Θ , Q_v , etc), and small scale variables by q. The coupling formalism of the large-scale and cloud-scale models is through the terms F_{CS}^Q and f_{LS}^q . In the original superparameterization implementation, the large scale models are solved from time T to $T + \Delta T$ in the following way [8]

$$Q|^{n+1} = Q|^n + \Delta T (A_Q + S_Q)|^{n+1}_n + \Delta T F^Q_{CS}|^n,$$
(3)

where $A_Q \equiv -\boldsymbol{U} \cdot \nabla Q$ denotes the large-scale advection term, S_Q denotes the large-scale source terms (e.g. pressure gradient, Coriolis acceleration, etc.), F_{CS}^Q denotes the forcing due to small scale processes, referred to as small-scale feedback, and the notation $|_n^{n+1}$ denotes the numerical time average of these terms. The large scale forcing to the small scale models is defined as

$$f_{LS}^{q}|^{n} = \frac{Q|^{n+1} - \langle q|^{n} \rangle}{\Delta T},\tag{4}$$

where $\langle \cdot \rangle$ represents the spatial average over the small scale domain. The small scale models are solved from T to $T + N\Delta t$ ($\Delta T = N\Delta t$ is assumed) by

$$q|^{n+1} = q|^n + \Delta t \sum_{i=1}^{N} \left[(a_q + s_q)|_i^{i+1} + f_{LS}^q|^n \right],$$
(5)

where $a_q \equiv -\boldsymbol{u} \cdot \nabla' q$, s_q represents small-scale sources (e.g. the surface drag for velocity or latent heating due to phase changes, etc). Note that although the small scale models are solved for N steps within the big time step ΔT , the same large scale forcing $f_{LS}^q|^n$ is employed at all small time steps inside a single big time step. Finally, the small scale feedback is defined by

$$F_{CS}^Q|^{n+1} = \frac{\langle q|^{n+1} \rangle - Q|^{n+1}}{\Delta T}.$$
(6)

2.3. The Sparse Space Time Superparameterization Algorithms

Given the large scale time step ΔT and large scale spatial mesh ΔX , the original SP algorithm solves a small scale model on a 2D embedded, horizontally-periodic domain with width ΔX for N small time steps with $\Delta t = \Delta T/N$ where Δt is the time step of the small scale model. Savings are achieved by the dimensional reduction to 2D in the small scale models and the massively parallel nature of the algorithm. In the SSTSP algorithms, the small-scale model is solved on a spatially periodic domain of reduced size $\Delta X/p$ for a reduced number of small-scale time steps N/p where $N = \Delta T/\Delta t$; here p can be any integer divisor of N. Thus, the SSTSP-p algorithms are on the order of p^{-2} more efficient for the small-scale models than the original SP algorithm. For example, for ensemble prediction or data assimilation, p^2 ensemble members of SSTSP-p can be generated for roughly the same cost of one SP ensemble member – a significant increase in ensemble size. This assumes that the computational cost associated with the small scale models dominates the cost associated with the large scale model, which is the typical situation.

Following the formulation of the original SP in the foregoing section, we summarize and present the efficient sparse space-time algorithms for superparameterization (SSTSP) below:

• The large scale models are first updated by

$$Q|^{n+1} = Q|^n + \Delta T (A_Q + S_Q)|^{n+1}_n + \Delta T F^Q_{CS}|^n.$$
(7)

• Define the new large scale forcing, with the spatial average over a $\Delta X/p$ periodic domain

$$f_{LS}^{q}|^{n} = p \frac{Q|^{n+1} - \langle q|^{n} \rangle}{\Delta T}.$$
(8)

• The small scale models are next updated by

$$q|^{n+\frac{1}{p}} = q|^{n} + \Delta t \sum_{i=1}^{N/p} \left[(a_{q} + s_{q})|^{i+1}_{i} + f^{q}_{LS}|^{n} \right].$$
(9)



Figure 1: Contours of the surface precipitation from different simulations for the forced propagating squall lines. Clockwise from top left: CRM, SP, SSTSP-3, and SSTSP-6.

• The eddy state at $T + (N/p)\Delta t$ is used to define the eddy feedback

$$F_{CS}^{Q}|^{n+1} = \frac{\langle q|^{n+\frac{1}{p}} \rangle - Q|^{n+1}}{\Delta T}$$
(10)

and the eddies at $T + N\Delta t$ are re-initialized using $q|^{n+1} = q|^{n+\frac{1}{p}}$.

Note that the same parameter p is used here to reduce the spatial and temporal size of the embedded domains, but this is not essential. As shown in detail in Xing et al. [10], this SSTSP-p algorithm can be interpreted as solving the small scale model on a periodic spatial domain of length $\Delta X/p$ for a time $\Delta T/p$ and performing a space-time periodic extension.

In [10] a stringent test bed was used to evaluate the computational performance of the SSTSP algorithms. Four different initial large scale background shears are used to generate squall lines. Squall lines are important organized mesoscale convective systems which convert moist available potential energy to kinetic energy via interactions of moisture and shear on at least two scales ([21] and references therein). In [10] the



Figure 2: Contours of the surface precipitation from different simulations for the dying scattered convection a = 0.3. Clockwise from top left: CRM, SP, SSTSP-3, and SSTSP-6.

ambient shears are given by

$$\bar{U}(z) = \begin{cases} 10a \left(\cos\left(\frac{\pi z}{12}\right) - \cos\left(\frac{\pi z}{6}\right) \right), & \text{if } z < 12, \\ -20a, & \text{otherwise,} \end{cases}$$
(11)

for a fixed coefficient a. Xing et al. [10] ran a non-parameterized squall line experiment with full CRM resolution with grid spacing of 1 km for four different values of a = 1 (strong background shear), 0.8, 0.5, and 0.3 (weak background shear). As a decreases the Richardson number increases [ref. 53, chapter 3]. It is shown that a propagating squall line always emerges in the first three cases with the same speed [21] while the weakest shear, a = 0.3, results in dying scattered convection.

Xing et al. [10] run the SSTSP-p algorithms with a small scale grid spacing of 1 km in the stringent test suite above, where the large scale model has a grid scale $\Delta X = 32$ km and a time step 6 times longer than the high resolution reference simulations; the results are compared with the CRM and with the original SP algorithm. (Running the large scale model on this coarsened mesh but completely ignoring the eddy feedbacks gives no organized behavior in all cases [10].) It is shown in [10] that for all the regimes of propagation and dying scattered convection, the large scale variables such as horizontal velocity and specific

humidity are captured in a statistically accurate way (pattern correlations above 0.75) based on space-time reduction of the small-scale models by a factor of 1/3; thus, the efficient SSTSP algorithms result in a gain of roughly a factor of 10 in the efficiency of the small-scale component while retaining statistical accuracy on the large scale variables. Even models with 1/6 reduction in space and time with a gain of 36 in efficiency are able to distinguish between propagating squall lines and dying scattered convection with a pattern correlation above 0.6 for horizontal velocity and specific humidity. Examples comparing the performance of SSTSP-p with the CRM and original SP are presented in figure 1 for the precipitation in the propagating squall line with a = 1 and in figure 2 for dying scattered convection with a = 0.3; both figures show excellent performance of SSTSP-3 and SSTSP-6. These encouraging results suggest the possibility of using these efficient new algorithms for mesoscale ensemble forecasting and data assimilation with coarse meshes of order $\Delta X = 30$ km. Slawinska et al. [54] have developed an interesting test bed involving a high resolution CRM simulation of a Walker cell with intraseasonal multi-scale variability. An excellent topic for further research is to check the SSTSP algorithms in these and other three-dimensional prototype simulations.

2.4. Algorithmic Constraints in SP Revealed Through Multi-Scale Analysis

In the introduction we advocated the use of multi-scale asymptotic analysis to guide the design of multiscale superparameterization algorithms. Such multi-scale models on scales appropriate for the original SP or SSTSP have been developed in [12, 13] and provide a framework for solving two separate models on different scales. For applications of multi-scale models to understand phenomena in the tropical atmosphere see [22, 26]. Here we follow [13] and give a simple formal illustration of the use of multi-scale asymptotics to constrain SP algorithms.

The multi-scale models in [13] for SP utilize the two natural spatial scales, the convective scale for (x, y) with units of 10 km, and the mesoscale for $(X, Y) = \epsilon(x, y)$ with units of order 100 km so $\epsilon \approx 0.1$ is a small parameter. Consider the conservation of mass equation from equation (1),

$$\frac{1}{\rho_0}\partial_z(\rho_0 w) + \partial_x u + \partial_y v = 0 \tag{12}$$

in nondimensional units appropriate to the (x, y) convective scale where all three velocity components have the same units of 10m/s in equation (12). With the two length scales described above it is natural to expand the velocity field in two spatial scales (notation for dependence on time is suppressed)

$$u_{\epsilon}(x, y, z) = \overline{u}(\epsilon x, \epsilon y, z) + u'(x, y, \epsilon x, \epsilon y, z) + \mathcal{O}(\epsilon)$$

$$v_{\epsilon}(x, y, z) = \overline{v}(\epsilon x, \epsilon y, z) + v'(x, y, \epsilon x, \epsilon y, z) + \mathcal{O}(\epsilon)$$

$$w_{\epsilon}(x, y, z) = \overline{w}(\epsilon x, \epsilon y, z) + w'(x, y, \epsilon x, \epsilon y, z) + \mathcal{O}(\epsilon)$$
(13)

where \overline{u} varies only on the mesoscales $(X, Y) = \epsilon(x, y)$ and $\overline{\cdot}$ denotes a restriction to the large scales by averaging over small scales. The fluctuations u' are defined such that $\overline{u'} = 0$. Inserting the expansion (13) into the mass conservation equation (12) and taking the average $\overline{\cdot}$ yields

$$\frac{1}{\rho_0}\partial_z(\rho_0\overline{w}) + \epsilon \left(\partial_X\overline{u} + \partial_Y\overline{v}\right) + \overline{\left(\partial_x u' + \partial_y v'\right)} = \mathcal{O}(\epsilon).$$
(14)

In the multi-scale formalism (e.g. [13]) the fluctuations are assumed to grow sublinearly so that the average of their derivative is smaller than order one. As a result, at leading order the multi-scale expansion becomes the constraint on the vertical velocity

$$\frac{1}{\rho_0}\partial_z(\rho_0\overline{w}) = 0 \text{ or equivalently } \overline{w} \equiv 0 \text{ for all } z.$$
(15)

This means that vertical velocities w that vary only on the spatial mesoscales should be of the order 1 m/s, i.e. a factor of ϵ smaller than the horizontal velocities of 10 m/s, which is consistent with observations.

For SP algorithms, the above multi-scale derivation imposes the constraint (15) on the large scales; this condition is imposed on an *ad hoc* basis in the original SP of Grabowski [8] and multi-scale asymptotic

analysis provides a systematic justification. More generally, multi-scale asymptotic analysis also improves the coupling of large and small scales compared to the original *ad hoc* framework since it allows horizontal derivatives of large scale variables to appear in the small scale equations; terms of this sort are absent from the original formulation but are essential to certain types of eddy instability, including, for example, baroclinic instability (e.g. [23, 24]). For more systematic uses of multi-scale asymptotics expanding the formal treatment above to many other applications see [12, 13, 21-24, 26, 55] and references therein.

2.5. Deterministic and Stochastic Multi-Cloud Models

The SP and SSTSP algorithms discussed above are novel ways to improve operational coarse resolution models (GCM parameterizations) for the interaction of clouds in the tropics. The encouraging results described in section 2.3 for SSTSP-6 and shown in figures 1 and 2 where the horizontal direction of the small scale embedded domains is discretized with only six grid points and where the small scale variables are only evolved for one sixth of the coarse grid time step point toward an obvious fact that at that stage the small scale models in SSTSP are acting as coherent random models to trigger and intermittently influence the large scales. This suggests that the small scale models in SP or SSTSP might be replaced by much cheaper systematic chaotic deterministic or stochastic models to capture the missing variability in tropical convection. One of the central issues is how to restore the coherent variability of mesoscale systems on coarse meshes. One way to do this is through deterministic [26, 56–58] and stochastic [30, 31, 59] multicloud models which respect the potential organization of the three cloud types above the boundary layer, congestus, deep, and stratiform. Indeed the deterministic multi-cloud model coupled to NCAR's current operational (HOMME) dry dynamical core already yields very realistic tropical variability with a coarse mesh at the equator of 167 km in an idealized setting [58].

The stochastic multi-cloud model for tropical convection introduced by Khouider et al. [59] is a novel approach to the problem of missing tropical variability in GCMs. The stochastic parameterization is based on a Markov chain lattice model where each lattice site is either occupied by a cloud of a certain type (congestus, deep, or stratiform) or it is a clear sky site. The convective elements interact with the large scale environment and with each other through convectively available potential energy (CAPE) and middle troposphere dryness. When local interactions between the lattice sites are ignored, a coarse grained stochastic process that is intermediate between the microscopic dynamics and the mean field equations [29, 60, 61] is derived for the dynamical evolution of the cloud area fractions. Besides deep convection, the stochastic multi-cloud model includes both low-level moisture preconditioning through congestus clouds and the direct effect of stratiform clouds including downdrafts which cool and dry the boundary layer. The design principles of the multi-cloud parameterization framework are extensively explored in the deterministic version of the model developed by Khouider and Majda [56, 57].

The stochastic multi-cloud convective parameterization introduced by Khouider et al. [59] in the context of a single column model has been used recently to study flows above the equator without rotation effects [30, 31]. The stochastic model dramatically improves the variability of tropical convection compared to the conventional moderate and coarse resolution paradigm GCM parameterizations. This increase in variability comes from intermittent coherent structures such as synoptic and mesoscale convective systems, analogs of squall lines and convectively coupled waves seen in nature whose representation is improved by the stochastic parameterization. Furthermore, the statistics of extreme precipitation events as observed in nature are captured by the stochastic multi-cloud model [31]. Simulations with sea surface temperature (SST) gradient yield a realistic mean Walker-cell circulation with plausible high variability. An additional new feature of the new stochastic parameterization is a natural scaling of the model from moderate to coarse grids which preserves the variability and statistical structure of the coherent features [30, 31]. These results systematically illustrate, in a paradigm model, the benefits of using the stochastic multi-cloud framework to improve deterministic parameterizations with clear deficiencies. In figure 3 we compare qualitatively the precipitation patterns generated by the stochastic multi-cloud model and high resolution CRM simulations.

Finally, there is interesting and successful modeling of orogenic propagating precipitation generated by methods which embed explicit convection [46, 63] in a multi-scale fashion; these models embed explicit dipole heating in the vertical and this can mimic some features of organized convection in a fashion similar



Figure 3: Surface precipitation patterns from simulations using the stochastic multi-cloud model [30] (left), and using a cloudresolving model (CRM) in a similar setup on a smaller domain [62] (right; figure reproduced with permission from [62]). Time runs from top to bottom in both figures; vertical axes are labeled in units of days and horizontal axes in units of 1000 kilometers (left) and kilometers (right).

to but more crudely than the multi-cloud models discussed earlier [26, 56-58]. There are also interesting multi-scale stochastic models for incorporating convective momentum transport [64].

3. Seamless Stochastic Superparameterization Algorithms for Geophysical Turbulence

In this section we introduce a new and general mathematical framework for stochastic superparameterization, review idealized mathematical models for statistical numerical analysis of multi-scale algorithms, and review recent applications of stochastic superparameterization to two multi-scale turbulent test problems. The new framework is built on the intuition gained through the results of simulations using SP, SSTSP, and the multi-cloud models, through multi-scale asymptotic analysis of SP, and through the mathematical test model formulated by Majda and Grote [14]; the latter acts as a simplified setting in which to analyze the mathematical structure of the new formulation, and of numerical methods used for its approximation.

In both SSTSP and the multi-cloud models, the details of the small-scale nonlinear dynamics are represented using very cheap, and very loose approximations. The example of SSTSP especially demonstrates that in some cases it is not important to accurately resolve the details of the small-scale nonlinear eddy-eddy interactions; nevertheless, since the small-scale dynamics are strongly nonlinear and turbulent, the eddyeddy nonlinearity cannot be ignored. The test model of Majda and Grote [14] (discussed in more detail below) replaces the eddy-eddy nonlinearity by an additive Gaussian stochastic process conditional on the local mean, which results in a computationally cheap small-scale model. Furthermore, [14] shows that the periodic domains used in conventional SP can lead to poor representation of small scale instabilities, and therefore poor representation of the coupling between the large and small scales, and demonstrates how this can be avoided through the computationally efficient use of formally infinite embedded domains. But the test model of [14] is only a test model and leaves open the question of how the ideas could be used to model a real system.

One option for applying the ideas of [14] is to replace the nonlinear small-scale models of conventional SP by quasilinear stochastic models. Multi-scale asymptotics (e.g. [13]) provides a mathematical framework for developing coupled sets of equations for use in conventional SP, and the small scale equations derived from this framework can be replaced by quasilinear stochastic approximations. This is a reasonable application of the ideas of [14]. However, multi-scale asymptotics is not clearly applicable to turbulent physical systems without scale separation. Thus, while multi-scale asymptotics may correctly describe the interaction between very large and very small scales, it cannot describe the interaction between scales immediately above and below the coarse grid scale, since these are not well separated, by definition. It is useful therefore to develop a more flexible multi-scale mathematical framework for formally deriving equations to govern the large and small scales in an SP algorithm.

In the next section we present just such a framework, based on the so-called point approximation, and subsequently approximate the nonlinear eddy dynamics by a stochastic process using a Gaussian closure [14, 15]; together these approximations form the basis of the stochastic SP algorithms described in recent work by the authors [16-18]. We then show how the test model of [14] acts as an essential simplification of stochastic SP that allows the mathematical and computational issues involved to be addressed directly, and we close with some examples of stochastic SP applied to two different turbulent test problems [16-18].

3.1. A General Formulation of Stochastic Superparameterization

Following [15] we begin with a system of PDEs that governs the dynamics on all scales

$$\partial_t \boldsymbol{u} = \mathbf{L}\boldsymbol{u} + \boldsymbol{B}(\boldsymbol{u}, \boldsymbol{u}) + \boldsymbol{S}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{u}) + \boldsymbol{F}$$
(16)

where $\boldsymbol{u}(\boldsymbol{x},t) \in \mathbb{R}^N$, \mathbf{L} is a linear operator (e.g. the Coriolis term), $\boldsymbol{B}(\boldsymbol{u},\boldsymbol{u})$ is a quadratic nonlinearity (e.g. incompressible advection), $\boldsymbol{S}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{u})$ is a cubic nonlinearity (e.g. compressible advection or moisture terms), and \boldsymbol{F} is external forcing.

We next separate into mean and eddies by applying a Reynolds average, denoted $\overline{(\cdot)}$. Thus, the equation for the mean \overline{u} is

$$\partial_t \overline{u} = \mathbf{L}\overline{u} + B(\overline{u}, \overline{u}) + S(\overline{u}, \overline{u}, \overline{u}) + \overline{F} + \mathbf{T}.$$
(17)

The mean equation (17) is not closed because it depends on the eddies through

$$\mathbf{T} = \overline{\boldsymbol{B}(\boldsymbol{u}',\boldsymbol{u}')} + \overline{[\boldsymbol{S}_2(\overline{\boldsymbol{u}})](\boldsymbol{u}',\boldsymbol{u}')} + \overline{\boldsymbol{S}(\boldsymbol{u}',\boldsymbol{u}',\boldsymbol{u}')}.$$
(18)

The above makes use of an 'eddy' variable defined by

$$\boldsymbol{u}' = \boldsymbol{u} - \overline{\boldsymbol{u}} \tag{19}$$

and of the notation

$$[\mathbf{S}_2(\overline{\mathbf{u}})](\mathbf{u}') = \mathbf{S}(\mathbf{u}', \mathbf{u}', \overline{\mathbf{u}}) + \mathbf{S}(\mathbf{u}', \overline{\mathbf{u}}, \mathbf{u}') + \mathbf{S}(\overline{\mathbf{u}}, \mathbf{u}', \mathbf{u}').$$
(20)

The eddy equation is derived by subtracting the mean equation from the original equation

$$\partial_t \boldsymbol{u}' = \mathcal{L} \boldsymbol{u}' + \boldsymbol{B}(\boldsymbol{u}', \boldsymbol{u}') + [\boldsymbol{S}_2(\overline{\boldsymbol{u}})](\boldsymbol{u}', \boldsymbol{u}') + \boldsymbol{S}(\boldsymbol{u}', \boldsymbol{u}', \boldsymbol{u}') - \mathbf{T} + \boldsymbol{F}'$$
(21)

where

$$\mathcal{L}\boldsymbol{u}' = \mathbf{L}\boldsymbol{u}' + \boldsymbol{B}(\overline{\boldsymbol{u}}, \boldsymbol{u}') + \boldsymbol{B}(\boldsymbol{u}', \overline{\boldsymbol{u}}) + [\boldsymbol{S}_1(\overline{\boldsymbol{u}})](\boldsymbol{u}'),$$
(22)
$$[\boldsymbol{S}_1(\overline{\boldsymbol{u}})](\boldsymbol{u}') = \boldsymbol{S}(\overline{\boldsymbol{u}}, \overline{\boldsymbol{u}}, \boldsymbol{u}') + \boldsymbol{S}(\overline{\boldsymbol{u}}, \boldsymbol{u}', \overline{\boldsymbol{u}}) + \boldsymbol{S}(\boldsymbol{u}', \overline{\boldsymbol{u}}, \overline{\boldsymbol{u}}).$$

The mean and eddy equations derived so far are exact, but are no more amenable to efficient numerical solution than the original equation. To develop a multi-scale formulation similar to conventional SP and to multi-scale asymptotic methods we apply the point approximation.

The point approximation is a formalism for introducing embedded pseudo-physical domains, as is done in SP and multi-scale asymptotics. To apply the point approximation in an independent variable (e.g. x, y, or t) eddy derivatives with respect to that variable in the eddy equation (21) are re-interpreted as derivatives in a new independent coordinate $\partial_x u' \to \partial_{\bar{x}} u'$ and the overbar is re-interpreted as an average over the new coordinate. Note though, that derivatives acting on the mean variables are not re-interpreted; in this regard the point approximation is similar to multi-scale asymptotics, but not to the original SP framework [8]. As an example, in SP for atmospheric convection the physical independent coordinates \bar{x} and \tilde{y} . Thus the eddies and the mean share the vertical z and time t coordinates. However, in the point approximation derivatives of mean variables with respect to x and y can still appear in the eddy equations, whereas in the framework of original SP such derivatives are absent. In the context of tropical atmospheric moist convection there is likely little impact from the presence or absence of these terms since the primary eddy instability depends on the vertical rather than the horizontal structure of the mean. However, in extra-tropical latitudes, and in other physical settings, important small-scale instabilities like baroclinic instability are precluded by the original framework but are allowed by the point approximation.

In general the point approximation can be made in one or more spatiotemporal coordinates, and the framework does not itself specify how it should be used. Thus, use of the point approximation should be guided by physical intuition, asymptotic analysis, and computational expediency.

The point approximation acts as a systematic framework for the derivation of coupled large and small scale equations for use in SP. If the equations produced by the point approximation are solved directly it is convenient to make the new spatial coordinates periodic, as is common practice in conventional SP. This approach is computationally intensive however, and as pointed out by [14], eddy instabilities may not be well represented on periodic embedded domains. Thus, generalizing the ideas of [14, 15] we replace the eddy-eddy nonlinearities in the eddy equation (21) by an additive Gaussian stochastic forcing and linear deterministic damping

$$B(\boldsymbol{u}',\boldsymbol{u}') + [S_2(\overline{\boldsymbol{u}})](\boldsymbol{u}',\boldsymbol{u}') + S(\boldsymbol{u}',\boldsymbol{u}',\boldsymbol{u}') - \mathbf{T} \to \sigma \dot{W} - \Gamma \boldsymbol{u}'.$$
⁽²³⁾

The term $\sigma \dot{W}$ represents spatially correlated temporally white Gaussian stochastic forcing, whose properties are further discussed below. Since the eddy equations are linear and Gaussian, under the assumption of a Gaussian prior, the eddy statistics will remain Gaussian for all time. This implies that the third and higher odd moments of the eddies are zero, and that the fourth and higher even moments can be expressed in terms of the second order moments. In particular this implies that the third order moment is zero $\overline{S(u', u', u')} = 0$ in **T** (equation (18)).

Consistent with the stochastic approximation we model the eddies as random functions homogeneous in the new spatial coordinates of the point approximation with the following spectral representation

$$\boldsymbol{u}' = \int \hat{\boldsymbol{u}}'_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\tilde{\boldsymbol{x}}} \mathrm{d}W_{\boldsymbol{k}}.$$
(24)

The use of infinite embedded domains allows a continuum of scales to participate in the eddy dynamics, thereby alleviating the difficulties posed by periodic embedded domains. We note though that for a homogeneous random function the spatial average is equal to the ensemble average, and the eddy term in the mean equations \mathbf{T} is therefore not stochastic since it contains only averaged quantities. This is an important difference compared to the structure of conventional SP: in conventional SP the eddies at a given location

can be viewed as a single realization of a stochastic process (actually a realization of chaotic deterministic dynamics) generating a feedback to the mean, whereas here the expected value over an infinite ensemble of eddy realizations (all of whose dynamics are sensitive to the local mean) is used to feedback to the mean. Nevertheless, in the following we show how to develop stochastic eddy closures for \mathbf{T} whose expected value is equal to the ensemble average over the eddy realizations.

The eddy equations in stochastic superparameterization (point approximation plus Gaussian closure) are quasilinear stochastic PDEs whose linear term has coefficients that depend on the local value of the large scale mean. The term "quasilinear" is applied to this stochastic approximation because the resulting equations are nonlinear if one considers the small and large scales variables as unknown even though the small scale equations are linear given the large scales. Since the coefficients of the eddy SPDEs do not depend on the new independent coordinates, the equations are diagonalized by Fourier transformation in those directions. The eddy equation for a single Fourier mode is therefore

$$d\hat{\boldsymbol{u}}_{\boldsymbol{k}}' = (\mathcal{L}_{\boldsymbol{k}} - \gamma_{\boldsymbol{k}})\hat{\boldsymbol{u}}_{\boldsymbol{k}}'d\tau + \sigma_{\boldsymbol{k}}dW_{\boldsymbol{k}} + F_{\boldsymbol{k}}'$$
(25)

where we have made the point approximation in time so that $\partial_t u' \to \partial_\tau u'$. If the eddies and the mean share a spatial coordinate, as the vertical coordinate z is shared in conventional SP of atmospheric convection, then (25) remains a system of SPDEs. These can then be discretized by any of a variety of methods including Galerkin truncation, finite differences, etc., and we point out that the discretization of the eddy equations need not be the same as of the mean equations; however, for simplicity we will assume here that the eddies and mean do not share a coordinate so that equation (25) is a finite-dimensional system of Itō stochastic differential equations.

Since the eddies have Gaussian statistics, the components of \mathbf{T} can all be related to quadratic moments of eddy variables. The Plancherel theorem relates the spatial average of quadratic moments to an integral over quadratic products of the Fourier coefficients

$$\overline{\boldsymbol{u}'(\boldsymbol{u}')^T} = \epsilon \int_0^{\epsilon^{-1}} \int \mathbb{E}\left[\hat{\boldsymbol{u}}'_{\boldsymbol{k}} \hat{\boldsymbol{u}}^{'*}_{\boldsymbol{k}}\right] \mathrm{d}\boldsymbol{k} \mathrm{d}\tau$$
(26)

where * is the complex conjugate transpose, T denotes the transpose, \mathbb{E} denotes ensemble average, ϵ^{-1} is the length of the time average associated with the point approximation in time, and the eddy Fourier covariance is

$$C_{\boldsymbol{k}} \equiv \mathbb{E}\left[\hat{\boldsymbol{u}}_{\boldsymbol{k}}'\hat{\boldsymbol{u}}_{\boldsymbol{k}}'^{*}\right].$$
(27)

The Itō formula allows the derivation of a deterministic system of ordinary differential equations for the eddy Fourier covariance from the eddy Fourier differential equation (25)

$$\frac{\mathrm{d}}{\mathrm{d}\tau}C_{\boldsymbol{k}} = (\mathcal{L}_{\boldsymbol{k}} - \gamma_{\boldsymbol{k}})C_{\boldsymbol{k}} + C_{\boldsymbol{k}}(\mathcal{L}_{\boldsymbol{k}} - \gamma_{\boldsymbol{k}})^* + \sigma_{\boldsymbol{k}}\sigma_{\boldsymbol{k}}^*.$$
(28)

Stochastic superparameterization consists in evolving the eddy Fourier covariance according to (28) and using the results to calculate **T** using (26). The properties of the stochastic forcing and damping σ_k and γ_k are design choices for the algorithm, but can be informed by studies of the phenomenology of the eddies. In addition, stochastic closures for **T** can be developed based on random quadratures in (26); an example based on random-direction plane waves is developed in [16, 18], and reviewed below.

3.2. Mathematical Test Models

Application of the above framework for stochastic superparameterization to real multi-scale turbulent physical systems involves many details which distract from the properties of the framework itself, for example, choices in the design of the stochastic eddy approximation, and in the numerical modeling of the nonlinear mean equations, etc. The test model of Majda and Grote [14], which provides much of the inspiration for the above framework, focuses attention on the framework and its properties by refraining from modeling a specific multi-scale system, and by linearizing the mean equation except for the nonlinear eddy feedback

term. The properties of the mean-eddy coupling can be posited as desired to investigate properties of the framework.

The test model of Majda and Grote [14] has the form

$$\partial_t \overline{u} + \overline{P}(\partial_x)\overline{u} = \overline{(u')^2} + \overline{F} \tag{29}$$

$$\partial_{\tau} u' + P'(\partial_{\tilde{x}}; \overline{u})u' = -\Gamma u' + \sigma(x)\dot{W}(\tau)$$
(30)

where the mean equation is posed in a periodic domain, $\overline{P}(\partial_x)$ is a linear differential operator including linear dispersion and dissipation, and $P'(\partial_{\bar{x}}; \bar{u})$ is a quasilinear pseudo-differential operator that depends on the local mean \bar{u} . This is clearly similar to the form of the mean and eddy equations in stochastic SP, but with a linear mean equation, and the numerical methods used for this system have the same structure as for stochastic SP. Majda and Grote [14] used this system to explore the effect of using a periodic domain in SP, and to emphasize the effects of lack of conditional equilibration. To do so, they specified $P'(\partial_{\bar{x}};\bar{u})$ in such a way that for certain values of \bar{u} there is a small range of eddy wavenumbers that is linearly unstable, and does not equilibrate in the limit $\tau \to \infty$. The lack of conditional equilibration of the eddies is a feature of turbulent systems with weak or no time scale separation, and requires a short time average ϵ^{-1} in equation (26). On a periodic embedded domain the discrete spectrum of wavenumbers may miss this small unstable band and as a result severely underestimate the eddy feedback term $(u')^2$.

In Harlim and Majda [38] the potential use of superparameterization for multi-scale filtering and data assimilation is developed in the simplest setting where only the mean variable \overline{u} is observed in the context of the above test model (in reality observations at a single location alias behavior from both large and small scales). While details are not presented here, we summarize some remarkable facts learned from the test model about multi-scale data assimilation and refer the reader to [38] for the details. There is robust high filtering and statistical prediction skill when periodic embedded domains with width L are used, provided that the domain width L is not too small. The high filtering skill with periodic embedded domains is robust even when the large-scale observation network is very sparse, as long as the filter models satisfy a controllability condition with nonzero system noise. When the system is deterministically forced and the perfect model is used for filtering, the filtering algorithm fails spectacularly; in contrast, when the deterministicallyforced system is filtered by a model with additional noise, restoring controllability through judicious model errors [20], the filtering skill is restored. Two alternative closures for the eddy feedback term, based on either assuming conditional equilibration of the eddies or on ignoring the eddies altogether, have much lower filtering and prediction skill even when the true signal is very smooth, with a k^{-6} spectrum, and when controllability is satisfied. These two alternative closures cause the mean to equilibrate to a steady state without turbulence, because the eddy terms are either ignored or improperly estimated. Furthermore, in the absence of controllability these models 'collapse' and relax to the equilibrium solution ignoring the observations. This is a counter-example to naive thinking that the small-scale processes are not so important in multi-scale turbulent dynamics with steep energy spectra.

Another idealized test model for superparameterization is provided by the Lorenz-96 model [65, 66] which consists of a system of ODEs of the form

$$\dot{X}_{k} = -(X_{k-2} - X_{k+1})X_{k-1} - X_{k} + F + \frac{h_x}{J}\sum_{j=1}^{J}Y_{j,k}, \quad k = 1, \dots, K$$
(31)

$$\dot{Y}_{j,k} = \frac{1}{\epsilon} \left[-(Y_{j+2,k} - Y_{j-1,k})Y_{j+1,k} - Y_{j,k} + h_y X_k \right], \quad j = 1, \dots, J$$
(32)

where the indices k and j are periodic (the model obtained by setting $Y_{j,k} = 0$ is also known as the Lorenz-96 model). If the k index is viewed as denoting a large-scale spatial coordinate and the j index as denoting a small-scale spatial coordinate, the system of ODEs has a form similar to the original SP algorithm of [8], where the X variables are the large-scale mean, and the Y variables are the eddies. In the multi-scale asymptotic and point approximation frameworks of SP the eddies have no spatial mean, so in these frameworks the Y variables minus their local mean would correspond to the eddies. While this system offers

a valuable test-bed for analysis of multi-scale algorithms, we caution that real multiscale turbulent systems often have weak or nonexistent time scale separation so ϵ should not be taken too small. Furthermore, to guarantee a strongly mixing turbulent dynamical system the number of large-scale modes K should be roughly forty with strong forcing [20], while the number of small-scale modes J should also be relatively large. Fatkullin and Vanden-Eijnden [67] applied HMM to this system with strong scale separation, Crommelin and Vanden-Eijnden [68] used it as a test-bed for the development of stochastic parameterizations based on Markov chains, and Kang and Harlim [69] and Wilks [70] used the system to investigate aspects of multi-scale data assimilation.

3.3. Tests of Stochastic Superparameterization in the MMT Equation

In this section we review the development of stochastic SP for a difficult but idealized one-dimensional test problem [17], which is a member of a family of equations that was originally proposed as a model for testing the predictions of wave turbulence theory [71]. These equations are called the MMT models which Zakharov et al. [72] named after their inventors [71]. The system is marked by a wide range of interacting scales, a shallow $k^{-5/6}$ energy spectrum, unstable collapsing solitons, dispersive waves, and an inverse cascade of energy from small scales punctuated by strong intermittent bursts of downscale energy cascade and dissipation. The MMT equation used here which governs the dynamics at all scales is

$$i\partial_t \psi = |\partial_x|^{1/2} \psi - |\psi|^2 \psi + iF_0 \sin(4\pi x/L) + iD\psi$$
(33)

where $F_0 = 0.0163$, L = 400, and D is damping applied to Fourier modes with $|k| = 2\pi/L$ and $|k| \ge 5200\pi/L$ in a periodic domain of length L. The fractional derivative is defined by $|\partial_x|e^{ikx} = |k|^{1/2}e^{ikx}$. A highresolution reference simulation is performed on a uniform grid of $2^{13} = 8192$ points using a fourth-order exponential Runge-Kutta method [73]. Further details on the numerical method can be found in [17].

Properties of the system are described by [71, 72, 74–76]. The reference simulation exhibits complex turbulent behavior consisting of non-traveling solitons that become unstable and collapse self-similarly to a high-amplitude sharp peak, and of weakly interacting dispersive waves. The collapsing solitons carry energy to small scales in an intermittent forward cascade, while weak wave interactions cascade energy to large scales in a background inverse cascade. Solutions exhibit time-averaged energy spectra with a shallow -5/6 spectral slope $|\hat{\psi}_k|^2 \propto k^{-5/6}$ over a large inertial range of scales.

The mean and eddy equations are

$$i\partial_t \overline{\psi} = |\partial_x|^{1/2} \overline{\psi} - |\overline{\psi}|^2 \overline{\psi} + \lambda \left[2\overline{|\psi'|^2} \ \overline{\psi} + \overline{(\psi')^2} \ \overline{\psi}^* \right] + i(D\overline{\psi} + F_0 \sin(4\pi x/L)), \tag{34}$$

$$i\partial_{\tau}\psi' = |\partial_{\tilde{x}}|^{1/2}\psi' - \left[2|\overline{\psi}|^{2}\psi' + \overline{\psi}^{2}\psi'^{*}\right] + i\left[(\Gamma + D)\psi' + \sigma(\tilde{x})\dot{W}\right],\tag{35}$$

where the additional damping operator Γ is a constant set to -10^{-5} . The autocorrelation of the forcing is defined such that the eddy Fourier covariance relaxes to an equilibrium with energy spectrum equal to $A(|k|^{5/6} + \exp\{|k| - 2600(2\pi/L)\})^{-1}$ in the absence of a mean (i.e. when $\overline{\psi} = 0$) and with no correlation between the real and imaginary parts of ψ' . The initial condition for the eddy Fourier covariance equation, which always has the form (25) in stochastic SP, is set equal to the equilibrium covariance that defines the forcing.

The Plancherel integrals defining the eddy terms in the mean equation are approximated using adaptive Simpson's rule quadrature over the intervals $|k| \in [k_0, 4096]$ where k_0 is the Nyquist wavenumber of the coarse grid – there is thus no scale gap between the smallest resolvable scale on the coarse grid and the largest scale on the eddy domains. Grooms and Majda [17] saved significant computational effort by precomputing the eddy terms as functions of $\overline{\psi}$, rather than computing the required Plancherel integrals every time the eddy terms are evaluated.

We present results using the same algorithm described in [17] on a coarse grid of 256 points using $\epsilon = 2$, A = 0.0044, and damping in the mean equations (34) defined such that $Dexp\{ikx\} = -d_x exp\{ikx\}$ with $d_k = 1$ for $|k| = 2\pi/L$, $d_k = 0.5$ for $|k| \ge 84\pi/L$, and $d_k = 0$ otherwise. For comparison we plot results obtained using the same numerical method and parameters but with no eddy terms. The time-averaged



Figure 4: (a) Time averaged energy spectra $|\hat{\psi}_k|^2$ from the high resolution reference simulation (solid), the stochastic SP simulation (dashed), and the unparameterized coarse resolution simulation (dotted). Time series of the solution amplitude $|\psi(x,t)|$ from the reference simulation (b), the stochastic SP simulation (c), and the unparameterized coarse resolution simulation (d). The abscissa is normalized by $2\pi/L$ so that the smallest wavenumber is k = 1.

energy spectra of the solutions of the reference solution, the mean equation, and the unparameterized mean equation are shown in figure 4. Clearly, the stochastic SP eddy terms are able to improve the energy spectrum of the mean equations from being two orders of magnitude too weak in the case of the unparameterized simulation to being comparable to the reference solution. This indicates that the stochastic SP algorithm is able to correctly account for the inverse cascade of energy from unresolved small scales despite being a purely deterministic closure. In addition, the stochastic SP model is able to generate collapsing solitons with approximately the same spatiotemporal statistics as the true model (see figure 4b-c and [17]). A variety of values of A and ϵ , and the effect of using periodic embedded domains are discussed further in [17].

3.4. Tests of Stochastic Superparameterization in Quasigeostrophic Turbulence

In [16, 18] the authors develop stochastic superparameterization for two-layer quasigeostrophic (QG) turbulence – a paradigm model of geophysical turbulence. The governing equations are

$$\partial_t q_1 + \nabla \cdot (\boldsymbol{u}_1 q_1) + \nu \nabla^8 q_1 = -\partial_x q_1 - k_d^2 v_1, \qquad (36)$$

$$\partial_t q_2 + \nabla \cdot (\boldsymbol{u}_2 q_2) + r \nabla^2 \psi_2 + \nu \nabla^8 q_2 = + \partial_x q_2 + k_d^2 v_2 \tag{37}$$

$$q_1 = \nabla^2 \psi_1 + \frac{k_d^2}{2} (\psi_2 - \psi_1),$$
$$q_2 = \nabla^2 \psi_2 - \frac{k_d^2}{2} (\psi_2 - \psi_1),$$

where q_j is the potential vorticity in the upper (j = 1) and lower (j = 2) layers, $\nabla^2 \psi_j = \omega_j$ is the relative vorticity, the velocity-streamfunction relation is $u_j = -\partial_y \psi_j$, $v_j = \partial_x \psi_j$, $k_d = 50$ is the deformation wavenumber (k_d^{-1}) is the deformation radius), the coefficient r = 4 specifies the strength of linear bottom friction (Ekman drag) and $\nu = 1.5 \times 10^{-16}$ is the hyperviscous Reynolds number. The terms on the right hand sides of (36) and (36) result from an imposed background baroclinic shear that drives the system. The dynamics can also be described in terms of barotropic and baroclinic modes, the former being given by the vertical average $q_t = (q_1 + q_2)/2 = \nabla^2 \psi_t$ and the latter by the vertical difference $q_c = (q_1 - q_2)/2 = (\nabla^2 - k_d^2)\psi_c$. Subscripts t and c are used throughout to denote barotropic and baroclinic components, respectively.

Our reference solution uses 512 points in each direction of a square periodic domain of nondimensional width 2π , which resolves the deformation scale with ten points per deformation wavelength. The nonlinear advection terms are dealiased using the 3/2-rule, which means that they are equivalent to simulations at 768² using the 2/3-rule. Time integration is via the adaptive, fourth-order, semi-implicit Runge-Kutta time integration scheme ARK4(3)6[L]2SA of Kennedy and Carpenter [77], treating the hyperviscous terms implicitly, with PI.3.4 adaptive stepsize control based on error-per-step in the infinity norm on q_j with a tolerance of 0.1 [78]. Further details can be found in [18].

Properties of the system can be found in a variety of references, e.g. [79, 80]. The solutions are forced by an imposed, horizonally uniform zonal baroclinic shear which is associated with a meridional density gradient. The dynamics flux heat (hence density) meridionally to erode the imposed background density gradient and associated zonal shear, and feature prominent small-scale vortices that are visible in figure 5c. Potential energy, proportional to $|\psi_1 - \psi_2|^2$, cascades downscale towards the deformation radius where it is converted to kinetic energy that cascades upscale. In this setting the inverse cascade from small scales is the primary energy source for the large scales, so this constitutes a stringent test of stochastic SP.

The mean equations are

$$\partial_t \overline{q}_1 = -\nabla \cdot (\overline{u_1 q_1}) - \partial_x \overline{q}_1 - k_d^2 \overline{v}_1 - \nu \nabla^8 \overline{q}_1,$$

$$\partial_t \overline{q}_2 = -\nabla \cdot (\overline{u_2 q_2}) + \partial_x \overline{q}_2 + k_d^2 \overline{v}_2 - r \nabla^2 \overline{\psi}_2 - \nu \nabla^8 \overline{q}_2$$

$$\overline{q}_1 = \nabla^2 \overline{\psi}_1 + \frac{k_d^2}{2} (\overline{\psi}_2 - \overline{\psi}_1),$$

$$\overline{q}_2 = \nabla^2 \overline{\psi}_2 - \frac{k_d^2}{2} (\overline{\psi}_2 - \overline{\psi}_1)$$
(38)



Figure 5: Time series of the heat flux $\int v_t \psi_c dA$ from the reference simulation (a) the stochastic SP simulation (b), and an unparameterized simulation (c), and snapshots of the upper layer potential vorticity q_1 from the reference simulation (d) stochastic SP simulation (e), and an unparameterized simulation (f). The grayscale is the same in all snapshots.

and the eddy equations are

$$\partial_{\tau}q_{1}' = F_{1}' - \Gamma q_{1}' - \overline{U}_{1} \cdot \tilde{\nabla}q_{1}' - u_{1}' \cdot \nabla \overline{Q}_{1} - \nu \tilde{\nabla}^{8}q_{1}',$$

$$\partial_{\tau}q_{2}' = F_{2}' - \Gamma q_{2}' - \overline{U}_{2} \cdot \tilde{\nabla}q_{2}' - u_{2}' \cdot \nabla \overline{Q}_{2}$$

$$-r \tilde{\nabla}^{2}\psi_{2}' - \nu \tilde{\nabla}^{8}q_{2}',$$

$$q_{1}' = \nabla^{2}\psi_{1}' + \frac{k_{d}^{2}}{2}(\psi_{2}' - \psi_{1}'),$$

$$q_{2}' = \nabla^{2}\psi_{2}' - \frac{k_{d}^{2}}{2}(\psi_{2}' - \psi_{1}')$$

$$(39)$$

where $\tilde{\nabla} = (\partial_{\tilde{x}}, \partial_{\tilde{y}}), \overline{U}_j = \overline{u}_j - (-1)^j \hat{x}$ and $\overline{Q}_j = \overline{q}_j - (-1)^j k_d^2 y$. The mean and eddies are coupled through the potential vorticity flux

$$\overline{u_j q_j} = \overline{u}_j \overline{q}_j + \overline{u'_j q'_j}$$

 $(u_j q_j)' = u_j q_j - \overline{u_j q_j}$

the eddy part of which is written without approximation as

$$\nabla \cdot (\overline{u'_j q'_j}) = \frac{k_d^2 (-1)^j}{2} \nabla \cdot (\overline{u'_j (\psi'_1 - \psi'_2)}) + \left(\partial_x^2 - \partial_y^2\right) \overline{u'_j v'_j} + \partial_{xy} \left(\overline{(v'_j)^2} - \overline{(u'_j)^2}\right)$$
(40)

where the first term is a 'heat' or 'buoyancy' flux and the remaining terms are the curl of the divergence of the Reynolds stress.

The autocorrelation of the forcing is defined such that in the absence of all other influences (i.e. when $\overline{U}_c = \nabla \overline{Q}_j = r = \nu = 0$) the eddy Fourier covariance relaxes to an isotropic equilibrium with a continuous energy spectrum equal to $Ak^{-5/3}$ for $k \leq k_d$ and to $Ak_d^{4/3}k^{-3}$ for $k \geq k_d$, with no net heat flux, with barotropic energy equal to baroclinic energy at each k, and with half as much kinetic energy in the lower layer as in the upper layer. The extra damping operator Γ is chosen such that it damps modes with $k > k_d$ uniformly at rate $\gamma_0 = 30$, and modes with $k \leq k_d$ at rate $\gamma_0(k/k_d)^{2/3}$. The eddy Fourier covariance equation, which always has the form (28) in stochastic SP, requires an initial condition which is set equal to the equilibrium covariance that defines the forcing.

The Plancherel integrals defining the eddy terms in the mean equation are written in polar form (i.e. with $k = |\mathbf{k}|$ and $\theta = \tan^{-1}(k_y/k_x)$) and the radial part is approximated using trapezoid rule quadrature over the intervals $|\mathbf{k}| \in [k_0, 256]$ where k_0 is the Nyquist wavenumber of the coarse grid. Grooms and Majda [17] saved significant computational effort for the MMT model by pre-computing the eddy terms as functions of the mean, rather than computing the required Plancherel integrals every time the eddy terms are evaluated. Such an approach is not practical in this case since the eddy integrals depend on six mean parameters; however, an intermediate approach was taken in [16] where the radial part of the integrals, which depends on only three mean parameters, is precomputed while the azimuthal part of the integral is directly computed during integration of the mean equations using trapezoid rule quadrature with equispaced nodes.

However, the inverse cascade of QG turbulence is not well modeled by a deterministic process, but rather by stochastic backscatter. Thus, while the deterministic closure does succeed in generating an inverse cascade, the results are otherwise poor. As noted earlier, this can be attributed to the fact that the deterministic closure uses the expected value of an infinite number of realizations of the eddy dynamics on formally infinite domains, whereas in reality the eddies consist of a single inhomogeneous realization. To incorporate this randomness of the eddies, and following the success of similar algorithms in modelling turbulent diffusion [81–84], we approximate the azimuthal direction of the Plancherel integrals using a twopoint trapezoid rule quadrature (θ and $-\theta$) where the angle θ is chosen randomly from a uniform distribution. The value of the eddy terms in the mean equations is therefore a random variable dependent on θ , and the expected value of the eddy terms in the mean equations equals the value of the deterministic closure. For more details on the algorithm see [18].

Figure 5 compares the meridional heat flux (the domain integral of $v_t\psi_c$) and upper layer potential vorticity q_1 from the reference simulation, a simulation using the stochastic closure with $A = 2.2 \times 10^4$, $\nu = 4 \times 10^{-10}$, and $\epsilon = 200$, and a simulation where the eddy terms are ignored and the value of $\nu = 3 \times 10^{-10}$ is tuned to produce the best results. The stochastic SP and unparameterized simulations are run on a coarse grid of 64×64 points. The reference simulation results are shown in figures 5a and 5d, the stochastic SP simulation results are shown in figures 5b and 5e, and the unparameterized simulation results are shown in figures 5c and 5f. The heat flux generated by the stochastic SP simulation is nearly identical to the reference simulation despite having a factor of 8 fewer points in each horizontal dimension in the stochastic SP simulation, and despite the complete inability of the coarse grid to resolve the prominent coherent vortices that populate the high resolution reference simulation. The unparameterized simulation has less small-scale variation than the SP simulation, and the heat flux is too low by a factor of approximately 4.

Tests of the algorithm in scenarios with large-scale jets, as well as sensitivity of the algorithm to ν , ϵ and A, and alternative closures are discussed in detail in [16, 18].

4. Concluding Discussion

This research expository paper discussed superparameterization, a class of multi-scale numerical methods designed to cope with the intermittent multi-scale effects of inhomogeneous geophysical turbulence which often inverse cascades energy from the unresolved scales to the large scales through waves, jets, vortices, and for example the effect of latent heat release through moist processes. Original as well as sparse space time superparameterization algorithms were discussed for the important case of moist atmospheric convection including the role of multi-scale asymptotic methods in providing self-consistent constraints on superparameterization algorithms. The very recent development of grid-free seamless stochastic superparameterization methods for geophysical turbulence appropriate for "eddy permitting" mesoscale ocean turbulence was presented here including a general formulation and illustrative applications to two-layer quasigeostrophic turbulence and the difficult test case involving the MMT models of dispersive wave turbulence. Test models for the statistical numerical analysis of superparameterization algorithms were also discussed both to underscore their utility for analyzing the performance of multi-scale simulation and filtering/data assimilation algorithm. Future research directions include the development of multi-scale data assimilation algorithms utilizing a suitable version of superparameterization for the forecast model as well as the development of algorithms for submesoscale ocean turbulence.

Acknowledgements

This research of A.J.M. is partially supported by the Office of Naval Research grants N00014-11-1-0306 and ONR-DRI N0014-10-1-0554. I.G. is supported as a postdoctoral fellow by the NSF *Collaborations in Mathematical Geosciences* program, grant DMS-1025468.

References

- A. Majda, Real world turbulence and modern applied mathematics, in: Mathematics: Frontiers and Perspectives, American Mathematical Society, 2000, pp. 137–151.
- [2] J. Hinze, Turbulence, McGraw-Hill, New York, 1959.
- [3] A. Monin, A. Yaglom, Statistical Fluid Dynamics, MIT Press, Cambridge, MA, 1971.
- [4] A. Townsend, The Structure of Turbulent Shear Flow, Cambridge University Press, 1976.
- [5] M. Hecht, H. Hasumi (Eds.), Ocean Modeling in an Eddying Regime, American Geophysical Union Monograph Series, 2008.
- [6] W. Grabowski, P. Smolarkiewicz, CRCP: a Cloud Resolving Convection Parameterization for modeling the tropical convecting atmosphere, Physica D 133 (1999) 171–178.
- [7] W. Grabowski, Coupling cloud processes with the large-scale dynamics using the Cloud-Resolving Convection Parameterization (CRCP), J. Atmos. Sci. 58 (2001) 978–997.
- [8] W. Grabowski, An improved framework for superparameterization, J. Atmos. Sci. 61 (2004) 1940–1952.
- D. Randall, M. Khairoutdinov, A. Arakawa, W. Grabowski, Breaking the cloud parameterization deadlock, Bull. Amer. Meteor. Soc. 84 (2003) 1547–1564.
- [10] Y. Xing, A. J. Majda, W. W. Grabowski, New Efficient Sparse Space-Time Algorithms for Superparameterization on Mesoscales, Mon. Weather Rev. 137 (2009) 4307–4324.
- [11] M. Khairoutdinov, D. Randall, C. DeMott, Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes, J. Atmos. Sci. 62 (2005) 2136–2154.
- [12] A. J. Majda, New multiscale models and self-similarity in tropical convection, J. Atmos. Sci. 64 (2007) 1393–1404.
- [13] A. J. Majda, Multiscale models with moisture and systematic strategies for superparameterization, J. Atmos. Sci. 64 (2007) 2726–2734.
- [14] A. J. Majda, M. J. Grote, Mathematical test models for superparametrization in anisotropic turbulence, Proc. Natl. Acad. Sci. USA 106 (2009) 5470–5474.
- [15] A. J. Majda, Challenges in Climate Science and Contemporary Applied Mathematics, Commun. Pur. Appl. Math. 65 (2012) 920–948.
- [16] I. Grooms, A. J. Majda, Efficient stochastic superparameterization for geophysical turbulence, Proc. Natl. Acad. Sci. USA 110 (2013) 4464–4469.
- [17] I. Grooms, A. J. Majda, Stochastic superparameterization in a one-dimensional model for turbulence, Commun. Math. Sci. 12 (2014).
- [18] I. Grooms, A. J. Majda, Stochastic superparameterization in quasigeostrophic turbulence, J. Comput. Phys. (2013) Submitted.
- [19] A. J. Majda, J. Harlim, B. Gershgorin, Mathematical strategies for filtering turbulent dynamical systems, Discrete Cont. Dyn. S. 27 (2010) 441–486.

- [20] A. Majda, J. Harlim, Filtering Complex Turbulent Systems, Cambridge University Press, 2012.
- [21] A. Majda, Y. Xing, New multi-scale models on mesoscales and squall lines, Commun. Math. Sci. 8 (2010) 113–134.
- [22] R. Klein, Scale-dependent asymptotic models for atmospheric flows, Ann. Rev. Fluid Mech. 42 (2010) 249–274.
- [23] I. Grooms, K. Julien, B. Fox-Kemper, On the interactions between planetary geostrophy and mesoscale eddies, Dynam. Atmos. Oceans 51 (2011) 109–136.
- [24] I. Grooms, K. S. Smith, A. J. Majda, Multiscale models for synoptic-mesoscale interactions in the ocean, Dynam. Atmos. Oceans. 58 (2012) 95–107.
- [25] Z. Malecha, G. Chini, K. Julien, A multiscale algorithm for simulating spatially-extended langmuir circulation dynamics, J. Comput. Phys. (2013) In press.
- [26] B. Khouider, A. J. Majda, S. N. Stechmann, Climate science in the tropics: waves, vortices and PDEs, Nonlinearity 26 (2013) R1–R68.
- [27] J.-M. Campin, C. Hill, H. Jones, J. Marshall, Super-parameterization in ocean modeling: Application to deep convection, Ocean Model. 36 (2011) 90–101.
- [28] A. Kerstein, Linear-eddy modeling of turbulent transport .2. Application to shear-layer mixing, Combust. Flame 75 (1989) 397–413.
- [29] A. J. Majda, C. Franzke, B. Khouider, An applied mathematics perspective on stochastic modelling for climate, Philos. T. R. Soc. A 366 (2008) 2429–2455.
- [30] Y. Frenkel, A. J. Majda, B. Khouider, Using the Stochastic Multicloud Model to Improve Tropical Convective Parameterization: A Paradigm Example, J. Atmos. Sci. 69 (2012) 1080–1105.
- [31] Y. Frenkel, A. Majda, B. Khouider, Stochastic and deterministic multicloud parameterizations for tropical convection, Climate Dynamics (2013) 1–25.
- [32] M. Branicki, A. Majda, Dynamic stochastic superresolution of sparsely observed turbulent systems, J. Comput. Phys. 241 (2013) 333–363.
- [33] S. Keating, A. Majda, K. Smith, New methods for estimating ocena eddy heat transport using satellite altimetry, Mon. Weather Rev. 140 (2012) 1703–1722.
- [34] W. E, B. Engquist, Z. Huang, Heterogeneous multiscale method: A general methodology for multiscale modeling, Phys. Rev. B 67 (2003).
- [35] E. Vanden-Eijnden, Numerical techniques for multi-scale dynamical systems with stochastic effects, Commun. Math. Sci. 1 (2003) 385–391.
- [36] W. E, B. Engquist, X. Li, W. Ren, E. Vanden-Eijnden, Heterogeneous multiscale methods: A review, Commun. Comput. Phys. 2 (2007) 367–450.
- [37] A. Armaou, I. Kevrekidis, C. Theodoropoulos, Equation-free gaptooth-based controller design for distributed complex/multiscale processes, Comput. Chem. Eng. 29 (2005) 731–740.
- [38] J. Harlim, A. Majda, Test models for filtering with superparameterization, Multiscale Model. Simul. 11 (2013) 282–308.
 [39] M. Moncrieff, M. Shapiro, J. Slingo, F. Molteni, Collaborative research at the intersection of weather and climate, WMO
- Bulletin 56 (2007) 204–211. [40] M. Khairoutdinov, D. Randall, A cloud resolving model as a cloud parameterization in the NCAR Community Climate
- System Model: Preliminary results, Geophys. Res. Lett. 28 (2001) 3617–3620.
- [41] M. C. Wyant, M. Khairoutdinov, C. S. Bretherton, Climate sensitivity and cloud response of a gcm with a superparameterization, Geophys. Res. Lett. 33 (2006) L06714.
- [42] M. Khairoutdinov, C. DeMott, D. Randall, Evaluation of the simulated interannual and subseasonal variability in an AMIP-Style simulation using the CSU multiscale modeling framework, J. Climate 21 (2008) 413–431.
- [43] P. N. Blossey, C. S. Bretherton, M. C. Wyant, Subtropical Low Cloud Response to a Warmer Climate in a Superparameterized Climate Model. Part II: Column Modeling with a Cloud Resolving Model, J. Adv. Model. Earth Sys. 1 (2009).
- [44] W. K. Tao, D. Anderson, J. Chern, J. Entin, A. Hou, P. Houser, R. Kakar, S. Lang, W. Lau, C. Peters-Lidard, X. Li, T. Matsui, M. Rienecker, M. R. Schoeberl, B. W. Shen, J. J. Shi, X. Zeng, The Goddard multi-scale modeling system with unified physics, Ann. Geophys. 27 (2009) 3055–3064.
- [45] M. C. Wyant, C. S. Bretherton, P. N. Blossey, Subtropical Low Cloud Response to a Warmer Climate in a Superparameterized Climate Model. Part I: Regime Sorting and Physical Mechanisms, J. Adv. Model. Earth Sys. 1 (2009).
- [46] M. S. Pritchard, M. W. Moncrieff, R. C. J. Somerville, Orogenic Propagating Precipitation Systems over the United States in a Global Climate Model with Embedded Explicit Convection, J. Atmos. Sci. 68 (2011) 1821–1840.
- [47] D. Rosa, J. F. Lamarque, W. D. Collins, Global transport of passive tracers in conventional and superparameterized climate models: Evaluation of multi-scale methods, J. Adv. Model. Earth Sys. 4 (2012).
- [48] F. Li, D. Rosa, W. D. Collins, M. F. Wehner, "Super-parameterization": A better way to simulate regional extreme precipitation?, J. Adv. Model. Earth Sys. 4 (2012).
- [49] M. C. Wyant, C. S. Bretherton, P. N. Blossey, M. Khairoutdinov, Fast cloud adjustment to increasing CO2 in a superparameterized climate model, J. Adv. Model. Earth Sys. 4 (2012).
- [50] W. Grabowski, Comments on "Preliminary tests of multiscale modeling with a two-dimensional framework: Sensitivity to coupling methods", Mon. Weather Rev. 134 (2006) 2021–2026.
- [51] H. Byers, R. Braham, The thunderstorm report of the thunderstorm project, U.S. Government Printing Office, 287 pp., 1949.
- [52] W. Grabowski, Toward cloud resolving modeling of large-scale tropical circulations: A simple cloud microphysics parameterization, J. Atmos. Sci. 55 (1998) 3283–3298.
- [53] A. Majda, Introduction to P.D.E.'s and Waves for the Atmosphere and Ocean, volume 9 of Courant Lecture Notes,

American Mathematical Society & Courant Institute of Mathematical Sciences, 2002.

- [54] J. Slawinska, O. Pauluis, A. Majda, W. Grabowski, Multi-scale interactions in an idealized walker circulation: Mean circulation and intra-seasonal variability, J. Atmos. Sci. (2013) Submitted.
- [55] K. Julien, E. Knobloch, Reduced models for fluid flows with strong constraints, J. Math. Phys. 48 (2007) 065405.
- [56] B. Khouider, A. Majda, A simple multicloud parameterization for convectively coupled tropical waves. Part I: Linear analysis, J. Atmos. Sci. 63 (2006) 1308–1323.
- [57] B. Khouider, A. J. Majda, Multicloud models for organized tropical convection: Enhanced congestus heating, J. Atmos. Sci. 65 (2008) 895–914.
- [58] B. Khouider, A. St-Cyr, A. J. Majda, J. Tribbia, The MJO and Convectively Coupled Waves in a Coarse-Resolution GCM with a Simple Multicloud Parameterization, J. Atmos. Sci. 68 (2011) 240–264.
- [59] B. Khouider, J. Biello, A. J. Majda, A stochastic multicloud model for tropical convection, Commun. Math. Sci. 8 (2010) 187–216.
- [60] M. Katsoulakis, A. Majda, D. Vlachos, Coarse-grained stochastic processes and Monte Carlo simulations in lattice systems, J. Comput. Phys. 186 (2003) 250–278.
- [61] B. Khouider, A. Majda, M. Katsoulakis, Coarse-grained stochastic models for tropical convection and climate, Proc. Natl. Acad. Sci. USA 100 (2003) 11941–11946.
- [62] W. Grabowski, J.-I. Yano, M. Moncrieff, Cloud Resolving Modeling of Tropical Circulations Driven by Large-Scale SST Gradients, J. Atmos. Sci. 57 (2000) 2022–2039.
- [63] M. W. Moncrieff, C. Liu, Representing convective organization in prediction models by a hybrid strategy, J. Atmos. Sci. 63 (2006) 3404–3420.
- [64] A. J. Majda, S. N. Stechmann, Stochastic models for convective momentum transport, Proc. Natl. Acad. Sci. USA 105 (2008) 17614–17619.
- [65] E. Lorenz, Predictability: A problem partly solved, in: Proceedings of Seminar on Predicability, volume 1, ECMWF, Reading, UK, 1996, pp. 1–18.
- [66] E. Lorenz, Predictability: A problem partly solved, in: T. Palmer, R. Hagedorn (Eds.), Predictability of Weather and Climate, Cambridge University Press, 2006, pp. 40–58.
- [67] I. Fatkullin, E. Vanden-Eijnden, A computational strategy for multiscale systems with applications to Lorenz 96 model, J. Comput. Phys. 200 (2004) 605–638.
- [68] D. Crommelin, E. Vanden-Eijnden, Subgrid-scale parameterization with conditional Markov chains, J. Atmos. Sci. 65 (2008) 2661–2675.
- [69] E. L. Kang, J. Harlim, Filtering Partially Observed Multiscale Systems with Heterogeneous Multiscale Methods-Based Reduced Climate Models, Mon. Weather Rev. 140 (2012) 860–873.
- [70] D. S. Wilks, 'Superparameterization' and statistical emulation in the Lorenz '96 system, Q J Roy. Meteor. Soc. 138 (2012) 1379–1387.
- [71] A. Majda, D. McLaughlin, E. Tabak, A one-dimensional model for dispersive wave turbulence, J Nonlinear Sci. 7 (1997) 9–44.
- [72] V. Zakharov, P. Guyenne, A. Pushkarev, F. Dias, Wave turbulence in one-dimensional models, Physica D 152 (2001) 573–619.
- [73] S. Cox, P. Matthews, Exponential time differencing for stiff systems, J. Comput. Phys. 176 (2002) 430-455.
- [74] D. Cai, A. Majda, D. McLaughlin, E. Tabak, Spectral bifurcations in dispersive wave turbulence, Proc. Natl. Acad. Sci. USA 96 (1999) 14216–14221.
- [75] D. Cai, A. Majda, D. McLaughlin, E. Tabak, Dispersive wave turbulence in one dimension, Physica D 152 (2001) 551–572.
- [76] V. Zakharov, F. Dias, A. Pushkarev, One-dimensional wave turbulence, Phys. Rep. 398 (2004) 1–65.
- [77] C. Kennedy, M. Carpenter, Additive Runge-Kutta schemes for convection-diffusion-reaction equations, Appl. Numer. Math. 44 (2003) 139–181.
- [78] G. Soderlind, Automatic control and adaptive time-stepping, Numer. Algorithms 31 (2002) 281–310.
- [79] A. F. Thompson, W. R. Young, Scaling baroclinic eddy fluxes: Vortices and energy balance, J. Phys. Oceanogr. 36 (2006) 720–738.
- [80] R. Salmon, Lectures on Geophysical Fluid Dynamics, Oxford University Press, New York, 1998.
- [81] A. Majda, Random shearing direction models for isotropic turbulent-diffusion, J. Stat. Phys. 75 (1994) 1153–1165.
- [82] F. Elliott, A. Majda, A new algorithm with plane-waves and wavelets for random velocity-fields with many spatial scales, J. Comput. Phys. 117 (1995) 146–162.
- [83] F. Elliott, A. Majda, Pair dispersion over an inertial range spanning many decades, Phys. Fluids 8 (1996) 1052–1060.
- [84] A. Majda, P. Kramer, Simplified models for turbulent diffusion: Theory, numerical modelling, and physical phenomena, Phys. Rep. 314 (1999) 238–574.