Novel parametrisation techniques in weather and climate modelling and their evaluation using simpler analogue models

Thomas D. Schanzer

Supervisor: Prof. Steven Sherwood

School of Physics

Climate Change Research Centre and ARC Centre of Excellence for Climate Extremes

University of New South Wales, Sydney, Australia

Abstract

Text

Contents

1 Introduction and background

1.1 The necessity of parametrisation

The primary task of general circulation models (GCMs) for Earth's weather and climate is to simulate the dynamics of the atmosphere and ocean, which are governed by the Navier-Stokes equations. The algorithm chosen to solve these partial differential equations is known as a model's *dynamical core* (McFarlane 2011). Since analytical solutions to the equations do not exist, the dynamical core necessarily approximates the continuous equations with finite-dimensional, exactly solvable alternatives using one of several possible discretisation schemes (e.g., the finite difference, finite element, finite volume and spectral methods) (Christensen and Zanna 2022). In practice, this usually involves representing the prognostic variables (i.e., those that affect the evolution of the flow) with sets of discrete samples in space and time, whose resolution is constrained by the available computing resources. The unavoidable consequence of discretisation is the loss of information about processes occurring on spatial and temporal scales smaller than the corresponding sampling intervals. These processes are said to be *unresolved*.

It is tempting to naïvely accept the loss of fine-scale information as a necessary sacrifice and hope the dynamical core will still make accurate predictions for the larger, resolved scales. Unfortunately, this too is impossible due to the *nonlinearity* of the governing equations. The reason for this may be seen by considering linear differential equations as a counterexample. If the governing equations were linear, they would allow arbitrary superpositions of solutions, meaning that any given solution could be partitioned into high- and low-frequency components, themselves also solutions. One would thus have the freedom to solve for the low-frequency components alone without compromise. This property may be understood more formally using the Fourier transform, defined for a function f of space and time by

$$\tilde{f}(\omega, \boldsymbol{k}) = \int \mathrm{d}t\, \mathrm{d}^3x\, e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} f(t, \boldsymbol{x}),$$

which reduces any linear differential equation to an algebraic equation relating the frequency ω to the wave vector $\boldsymbol{k}$. Each wavenumber component of the initial state propagates trivially according to its own time dependence $e^{-i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega(\boldsymbol{k})t)}$, *independently of the other components*. If the equations of fluid dynamics were linear, one could safely neglect the fine-scale dynamics because they would have no influence on the coarse scales. In reality, the equations are nonlinear cannot be solved by Fourier transform.

The consequence of the nonlinearity of the equations governing atmospheric and oceanic flows is, therefore, a coupling of the resolved coarse scales to the unresolved fine scales (McFarlane 2011). This fact may be demonstrated more explicitly by applying so-called *Reynolds averaging* to the equations (Christensen and Zanna 2022). Reynolds averaging decomposes each field q into the sum of a coarse-grained (in space or time) or statistical-ensemble-averaged field $\bar{q}$ and a perturbation q'. Note that $\overline{q'} = 0$ by definition. The coarse-graining operation is assumed to be linear, commute with differentiation and satisfy $\overline{\bar{p}q} = \bar{p}\bar{q}$ for any two fields p and q (Monin and Yaglom 2007). Following the example given by Christensen and Zanna (2022), consider the incompressible Navier-Stokes equations, which have the general form

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = \sum f_i,$$
$$\frac{\partial u_i}{\partial x_i} = 0$$

where u_i are the components of the velocity, x_i are the coordinates and f_i are various forces per unit mass. Summation over repeated indices is implied. Applying the decomposition and coarse-graining both

sides of the equations yields

$$\begin{aligned} & \frac{\partial}{\partial x_i}\left(\overline{u_i}+u_i'\right)=0 \\ \Rightarrow \qquad & \frac{\partial \overline{u_i}}{\partial x_i}+\underbrace{\frac{\partial \overline{u_i'}}{\partial x_i}}_{=0}=0 \\ \Rightarrow \qquad & \frac{\partial \overline{u_i}}{\partial x_i}=\frac{\partial u_i'}{\partial x_i}=0 \end{aligned}$$

and

$$\begin{aligned} \sum \overline{f_i} &= \overline{\frac{\partial}{\partial t}\left(\overline{u_i}+u_i'\right)}+\overline{\left(\overline{u_j}+u_j'\right)\frac{\partial}{\partial x_j}\left(\overline{u_i}+u_i'\right)} \\ &= \frac{\partial \overline{u_i}}{\partial t}+\underbrace{\frac{\partial \overline{u_i'}}{\partial t}}_{=0}+\overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j}+\overline{u_j}\underbrace{\frac{\partial \overline{u_i'}}{\partial x_j}}_{=0}+\underbrace{\overline{u_j'}}_{=0}\frac{\partial \overline{u_i}}{\partial x_j}+\overline{u_j'\frac{\partial u_i'}{\partial x_j}} \\ &= \frac{\partial \overline{u_i}}{\partial t}+\overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j}+\frac{\partial \overline{u_i'u_j'}}{\partial x_j}-\overline{u_i'\underbrace{\frac{\partial u_j'}{\partial x_j}}_{=0}} \\ \Leftrightarrow \qquad \frac{\partial \overline{u_i}}{\partial t}+\overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j} &= \sum \overline{f_i}-\frac{\partial \overline{u_i'u_j'}}{\partial x_j}. \end{aligned}$$

The physical meaning of this last equation is that the evolution of the coarse-grained velocity field depends not only on itself and the coarse-grained forces, but also on the perturbations via $\overline{u_i'u_j'}$, which does not necessarily vanish. This dependence is a consequence of nonlinearity.

The theoretical discussion in this section establishes that physical processes occurring at one place in the spectrum of temporal and spatial scales are coupled to all the other processes in the spectrum (Franzke et al. 2015). In particular, to ignore the effect of processes not explicitly resolved in numerical models would introduce unacceptable systematic biases, or errors, in the model forecasts. GCMs therefore require parametrisation schemes to estimate the effects of unresolved processes as functions of the available large-scale information (and the parameters of these functions must be chosen appropriately—hence the name "parametrisation").

1.2 Mathematical formulation of the parametrisation problem

In order to make any progress on the parametrization problem, it is necessary to formalise the notion of estimating the "effects" of unresolved processes. I will describe the general approach, which is well-established (see, e.g., Hasselmann 1976; Palmer 2001; Demaeyer and Vannitsem 2018; Brajard et al. 2021), while acknowledging possible alternative conventions where they exist.

The Earth system may be considered as a dynamical system whose exact evolution is governed by equations of the form

$$\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t}=\boldsymbol{F}(\boldsymbol{z},t), \tag{1.1}$$

where the vector $\boldsymbol{z}$ contains all the variables needed to fully specify the state of the system and $\boldsymbol{F}$ is a nonlinear differential operator. One then employs a change of variables $\boldsymbol{z}\rightarrow(\boldsymbol{x},\boldsymbol{y})$, where $\boldsymbol{x}$ contains the information or variables that are explicitly resolved by the model being studied and $\boldsymbol{y}$ contains the unresolved, 'sub-grid' information or variables. Practically, $\boldsymbol{x}$ and $\boldsymbol{y}$ can be thought of as projections of the full state $\boldsymbol{z}$ onto lower-dimensional subspaces (Brajard et al. 2021). Alternatively (and more concretely), $\boldsymbol{x}$ can be obtained by averaging the original variables over several adjacent grid points, with $\boldsymbol{y}$ being the corresponding residuals (Zacharuk et al. 2018; Alcala and Timofeyev 2021).

In principle, the change of variables splits (1.1) into two parts

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}, t),$$
$$\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \boldsymbol{H}(\boldsymbol{x}, \boldsymbol{y}, t)$$

that (still exactly) describe the coupled evolution of the resolved and unresolved variables via nonlinear differential operators $\boldsymbol{G}$ and $\boldsymbol{H}$. The goal of parametrisation is now to derive a new system of equations

$$\frac{\mathrm{d}\tilde{\boldsymbol{x}}}{\mathrm{d}t} = \tilde{\boldsymbol{G}}(\tilde{\boldsymbol{x}}, t) \tag{1.2}$$

whose solution $\tilde{\boldsymbol{x}}(t)$ approximates the true $\boldsymbol{x}(t)$ as well as possible without explicitly modelling $\boldsymbol{y}(t)$. The measure used to compare $\tilde{\boldsymbol{x}}$ and $\boldsymbol{x}$ depends on the intended function of the parametrised model; weather forecasting applications that prioritise short-term predictive skill might call for minimisation of the root-mean-square error (RMSE), while climate prediction would be more concerned with the closeness of their probability distributions (including both mean and extreme values).

Before proceeding, it is important to note two assumptions that are implicitly made in writing down (1.2). The first is that $\tilde{\boldsymbol{G}}$ is a *deterministic* function of $\tilde{\boldsymbol{x}}$. The second is that $\tilde{\boldsymbol{G}}$ depends only on the value of $\tilde{\boldsymbol{x}}$ at time t and not earlier times (that is to say, the parametrisation is *memoryless*. Recent research has shown that relaxing these assumptions has the potential to greatly improve the accuracy and reliability of the parametrised model, and this will be discussed in detail in §§ 2 and 3. However, for the purpose of this basic introduction, I will temporarily follow the traditional approach and retain these assumptions.

There is more than one possible interpretation of $\tilde{\boldsymbol{G}}(\boldsymbol{x}, t)$ and its relationship to $\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}, t)$. Hasselmann (1976) describes $\tilde{\boldsymbol{G}}(\boldsymbol{x}, t)$ as an ensemble average of $\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}, t)$, taken over the distribution of all $\boldsymbol{y}$ that are possible for a given $\boldsymbol{x}$, i.e.

$$\tilde{\boldsymbol{G}}(\boldsymbol{x}, t) = \langle \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}, t) \rangle_{\boldsymbol{y}}.$$

Demaeyer and Vannitsem (2018) instead assert that the ultimate goal of parametrisation is to literally approximate $\boldsymbol{y}$ as a function $\boldsymbol{\xi}(\boldsymbol{x})$, in which case

$$\tilde{\boldsymbol{G}}(\boldsymbol{x}, t) = \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{\xi}(\boldsymbol{x}), t).$$

In practice, $\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}, t)$ usually separates into a known resolved part $\boldsymbol{D}(\boldsymbol{x}, t)$ independent of $\boldsymbol{y}$ and an unresolved coupling term $\boldsymbol{C}(\boldsymbol{x}, \boldsymbol{y}, t)$. It then suffices to approximate only the unresolved part by a parametrisation $\boldsymbol{P}(\boldsymbol{x}, t)$ using one of the above methods, so that the parametrised model reads

$$\frac{\mathrm{d}\tilde{\boldsymbol{x}}}{\mathrm{d}t} = \boldsymbol{D}(\tilde{\boldsymbol{x}}, t) + \boldsymbol{P}(\tilde{\boldsymbol{x}}, t). \tag{1.3}$$

1.3 Physical processes requiring parametrisation

Following the broad theoretical argument in the previous sections, I now turn to concrete examples of processes that are often represented by parametrisations, restricting the discussion to atmospheric processes for brevity. This section will give a basic introduction to each process, its effect on the larger-scale behaviour of the climate system and the role of its corresponding parametrisation scheme. The purpose of these examples is to provide real-world context and motivation for current parametrisation research in more idealised settings, which will form one of the main topics of this review.

Cloud microphysics parametrisations model the composition of clouds in terms of the amount of water in the solid (e.g., hail), liquid (cloud droplets and rain) and gaseous phases, and the rate of transitions between these phases. Accurate representation of cloud formation and evolution is crucial for several reasons. First, the interaction of clouds with solar and terrestrial radiation has a major influence on the overall energy balance of the atmosphere (McFarlane 2011). Second, cloud water phase transitions lead to

the formation of precipitation, and are a source and sink of latent heat that drives convection (McFarlane 2011). Furthermore, the statistics of cloud formation are linked to global temperatures in a feedback loop; uncertainty in the sign and magnitude of this feedback effect is a major contributor to uncertainty in the sensitivity of global temperatures to increases in atmospheric CO_2 concentration (Andrews et al. 2012; Christensen and Zanna 2022; Stevens and Bony 2013). Microphysical processes occur on the spatial scales of single water droplets and ice particles (of order 10^{-6} m to 10^{-2} m; Lamb (2003)), among the smallest scales in the atmosphere. All atmospheric models must parametrise the amount of cloud water in each phase, the size distribution and concentration of liquid and ice particles, and the resulting rate and type (rain, hail, snow, etc.) of precipitation (Christensen and Zanna 2022).

Moist convection encompasses vertical motions in the atmosphere that are accompanied (and driven) by phase changes of water, and must be parametrised when it occurs on sub-grid scales. Moist convection is generally triggered by warming and moistening at low levels, which create convective instability, and is manifested by narrow updrafts and downdrafts that can rarely be resolved explicitly (McFarlane 2011). Convection transports heat and moisture vertically, removing the instability and generating storms where the condensed water falls out as precipitation; more broadly, it is a key component of the global atmospheric circulation in spite of its small spatial scale (Christensen and Zanna 2022).

It should be noted that parametrisation also encompasses estimation techniques for processes that are too complex to model exactly for reasons unrelated to spatial and temporal resolution (McFarlane 2011). A good example of such a process is *radiative transfer*. Air and its constituent gases, as well as clouds, absorb, emit, reflect and/or scatter solar and terrestrial radiation. This leads to differential heating and cooling that drives the atmospheric circulation. While the theory of radiative transfer is understood well enough to allow precise calculations in principle, the prohibitive computational cost of such calculations necessitates a parametrisation based on simplifying assumptions (the details of which are beyond the scope of this review) (Christensen and Zanna 2022).

1.4 Traditional solutions to the problem and their limitations

With the examples of cloud microphysics, moist convection and radiative transfer in mind, I now broadly review the traditional approaches used to construct parametrisation schemes in practice. In particular, this section will identify the key assumptions upon which many traditional schemes are founded, and the circumstances under which the assumptions may be violated. This will motivate research into novel approaches.

A parametrisation scheme is traditionally constructed by formulating a simplfied, easily solvable and deterministic conceptual model of the physical process in question. The solution of this model is then used to estimate the effect of the process on the coarse-scale state of the parent model, known as the *unresolved tendency* (McFarlane 2011). For example, the earliest convective parametrisation was developed by Manabe, Smagorinsky, and Strickler (1965) and simply assumed that the net effect of convection is to relax the vertical structure of the atmosphere towards a neutrally stable state whenever it becomes convectively unstable. One obvious deficiency of simple conceptual models is that they cannot possibly capture the full range of variability in the processes they simulate. One major branch of modern parametrisation research therefore studies *data-driven* schemes that instead use observational or high-resolution simulated data to fit empirical models for the unresolved processes, naturally capturing a wider range of variability (Christensen and Zanna 2022). Data-driven parametrisation will be discussed in much further detail in the next section.

In general, a traditional parametrisation scheme aims to capture the net unresolved tendency due to all occurrences of the unresolved process (e.g., all convective updrafts and downdrafts) within each grid cell of the parent model. A deterministic prediction of this type is valid when each grid cell contains many independent realisations whose varying contributions may be expected to yield a reliable average tendency. This requires a *scale separation* between the unresolved process and the resolution of the parent model. Scale separation breaks down when model development and increases in available computing resources allow simulations at resolutions approaching the previously unresolved scales. In this case, knowledge

of the coarse-scale state cannot be expected to uniquely determine the unresolved tendency because the process is only realised a few times in each grid cell. The resulting (seemingly) random nature of the true unresolved tendency motivates stochastic treatments (McFarlane 2011; Christensen and Zanna 2022; Berner et al. 2017). These will be discussed in the next section.

The conceptual models core to traditional parametrisations usually contain free parameters that are determined from the coarse-scale model state using additional assumptions called *closures*. Closures often postulate a state of quasi-equilibrium between the unresolved processes and their large-scale environment, such as a balance between the accumulation of convective available potential energy (CAPE) and its removal by convection, or between the horizontal convergence of moisture at low levels and its convective transport to higher levels (McFarlane 2011; Christensen and Zanna 2022; Palmer 2019). However, there is no guarantee that such an equilibrium exists; in fact, it has been demonstrated that the CAPE balance is violated by fluctuations on sub-diurnal time scales (Donner and Phillips 2003) and by midlatitude continental convection (Zhang 2002). Newer parametrisation schemes have allowed departures from equilibrium by representing the unresolved processes in a prognostic rather than diagnostic manner (i.e., allowing the processes to have their own self-governing time dependence rather than calculating them from the large-scale state at each time step independently of their values at the previous step) (Rio, Del Genio, and Hourdin 2019; Berner et al. 2017). This creates *memory* or *latency* in the parametrised tendencies, meaning that the tendencies have some nonzero response time when subjected to sudden changes in the large-scale state. Memory will be discussed further in the next section.

Parametrisation schemes commonly suffer from several other issues that I will briefly address here. Firstly, while the division of the general atmospheric dynamics into a set of separately parametrised processes (microphysics, convection, etc.) is physically motivated, it remains somewhat arbitrary because these processes, strictly speaking, form a continuum without well-defined boundaries (Christensen and Zanna 2022; McFarlane 2011). It is a goal of contemporary research to unify the parametrisations as much as possible. Secondly, given the importance of future climate projections, it is natural to ask whether the parametrisation schemes that have been developed and tuned on today's climate remain valid as the climate changes over decade- to century-long simulations. This is a matter of particular concern for data-driven parametrisations, since there is little reason to trust empirical models once they are extrapolated beyond the range of the data originally used to fit them (Christensen and Zanna 2022). Finally, unless very special care is taken, parametrisation schemes can cause the parent model to violate known physical conservation laws (e.g., mass, energy and momentum) (Christensen and Zanna 2022). Efforts to resolve this issue are ongoing.

2 Novel approaches to the parametrisation problem

Since the 1990s, the limitations identified in § 1.4 have prompted many to reconsider the principles upon which parametrisation schemes are founded. The main advance has been the development of stochastic parametrisations that incorporate random noise in the predicted tendencies. More recently, others have approached the problem from an entirely new direction, developing data-driven parametrisations that learn to predict the tendencies empirically. This section will introduce the stochastic and data-driven approaches in general terms while omitting the technical details of individual implementations, the aim being to contextualise and motivate the study of these approaches in more idealised frameworks in § 3. Other methods (such as superparametrisation) exist but are beyond the scope of this review.

2.1 Stochastic parametrisation and memory

The potential value of stochasticity for climate modelling was first established by Hasselmann (1976), whose seminal paper sought to explain the characteristics of long-term climate variability. Knowing that climate depends on interactions between all components of the Earth system (atmosphere, ocean, cryosphere, biosphere, etc.), Hasselmann argued that the effect of the more rapidly-evolving atmosphere on the other, more slowly-varying components is that of a stochastic forcing. Owing to their long response time, the slowly-varying components effectively integrate this stochastic forcing, allowing long-term climate variability to be characterised as a type of random walk process akin to the Brownian motion of a massive particle in a fluid.

Further motivation for stochastic parametrisation in particular stems from the need to reliably estimate uncertainties in weather forecasts and climate projections. Weather forecasting centres typically propagate initial condition uncertatinties (due to imperfect observations) through to the final forecast in a Monte Carlo fashion, by initialising an ensemble of model runs with perturbed initial conditions and measuring the spread of the resulting forecasts. It has been observed that deterministic models produce systematically under-dispersed ensembles that fail to span the range of actual weather outcomes, indicating that these models are failing to capture additional sources of variability (Palmer et al. 2005; Berner et al. 2017; Palmer 2019). Knowing that poor scale separation and departures from quasi-equilibrium should preclude deterministic relationships between unresolved tendencies and the large-scale state (see § 1.4), it should seem highly likely that deterministic parametrisation contributes to this deficiency.

The principle of stochastic parametrisation is, therefore, that unresolved tendencies should be randomly sampled from an appropriate distribution at each point in space and time, not simply set to the mean of the distribution (Franzke et al. 2015). This choice has now been theoretically justified using statistical mechanical arguments; most notably, Wouters and Lucarini (2012, 2013) showed that, assuming a weak coupling to the resolved variables, the effect of unresolved dynamics should be parametrised by a combination of deterministic and stochastic terms, as well as a *non-Markovian* memory term depending on the past states of the resolved variables.

The simplest and earliest approach to stochastic parametrisation is the method of *stochastically perturbed parametrisation tendencies* (SPPT), which takes an existing deterministic parametrisation and randomly scales its output with a multiplicative noise field (Palmer 2019; Christensen 2020). Re-using the notation of § 1.2, an SPPT model for the resolved variables $\boldsymbol{x}$ takes the form

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{D}(\boldsymbol{x}, t) + [I + \mathrm{diag}(\boldsymbol{e})]\boldsymbol{P}(\boldsymbol{x}, t),$$

where $\boldsymbol{P}$ is the existing deterministic parametrisation, $\boldsymbol{e}$ is a mean-zero random vector with the same length as $\boldsymbol{x}$ and I is the identity matrix. The choice of multiplicative (and thus inherently state-dependent) rather than additive noise is intuitively motivated by the expectation that variability in the effect of unresolved processes should be greatest when those processes are most active (Franzke et al. 2015). Christensen (2020) performed a comparison of high-resolution simulations to parametrised single-column model output which justified the use of multiplicative perturbations.

In GCMs, where the variables being modelled have spatial dependence, the multiplicative perturbation takes the form of a random *field* $e(x, y, t)$. It has been argued that this random field should be spatially and temporally correlated (in contrast to uncorrelated "white" noise) in order to emulate the organisation of unresolved processes on larger scales and their persistence in time (Christensen and Zanna 2022; Franzke et al. 2015). In particular, perturbations that explictly depend on their own past values constitute a type of memory, albeit distinct from the memory term advocated by Wouters and Lucarini (2012, 2013), which would instead couple the perturbations to the past values of the large-scale variables.

Stochastic parametrisations have several known advantages over their deterministic counterparts, and are now implemented in operational weather forecast models. They have been found to remedy the aforementioned issues of ensemble underdispersion and prediction unreliability (Palmer et al. 2005; Berner et al. 2017), and (as of 2009) even make the skill of five-day forecasts comparable to that of deterministically parametrised two-day forecasts (Palmer 2019). Despite the zero-mean nature of the noise, it has been shown that stochastic parametrisations can reduce systematic model biases ("noise-induced drift"; Palmer et al. (2005)) and stabilise the simulation of regime-based behaviour (such as the El Niño-Southern Oscillation) in the climate system (Berner et al. 2017).

Reviewing the field, Palmer (2019) identifies outstanding issues for further research. The main concern is the lack of rigour in most stochastic schemes; SPPT, for example, modifies existing deterministic schemes *ad hoc* rather than incorporating stochasticity *ab initio* in the development process. More objective approaches are yet to gain widespread acceptance. This motivates both data-driven methods and further testing using simpler dynamical systems where objectivity is more feasible.

2.2 Data-driven parametrisation and machine learning

The fitting of predictive statistical models to data is, of course, ubiquitous in the sciences, but attempts to use such models as parametrisation schemes and couple them into fully fledged GCMs are a relatively new phenomenon—certainly more so than stochastic parametrisations. The prerequisite for all data-driven parametrisations is obviously training data: mathematically, an approximate solution of (1.1) of sufficient accuracy and resolution to be considered "truth" for the application at hand. Training data are typically derived from high-resolution simulations, such as regional weather model runs or large eddy simulations (LES). The other ingredient is the imperfect low-resolution model that will later be augmented by parametrisation.

To generate an appreciation of how these data might be used to construct a parametrisation in practice, I roughly follow the argument and notation of Brajard et al. (2021). Define a map $\mathcal{M}$ that takes each state $\boldsymbol{z}(t)$ in the training dataset to the state $\boldsymbol{z}(t + \delta t)$ at the next time step. Similarly, denote by $\mathcal{M}^{\mathrm{r}}$ the low-resolution model (the superscript r meaning "reduced"), which maps a low-resolution state $\boldsymbol{x}(t)$ to a prediction for $\boldsymbol{x}(t+\Delta t)$, where Δt is not necessarily equal to δt (usually larger). The spaces of high- and low-resolution states, having different dimensions, may be linked by an operator $\langle\cdot\rangle$ that projects high-resolution states onto the low-resolution state space. Practically, the projection operation is simply a coarse-graining of $\boldsymbol{z}$ by averaging values at adjacent grid points to match the resolution of $\boldsymbol{x}$. Now, for each state $\boldsymbol{z}$ in the training dataset, one may compute the difference

$$\boldsymbol{\epsilon}(\boldsymbol{z}) = \frac{\langle \mathcal{M}(\boldsymbol{z})\rangle - \langle \boldsymbol{z}\rangle}{\delta t} - \frac{\mathcal{M}^{\mathrm{r}}(\langle \boldsymbol{z}\rangle) - \langle \boldsymbol{z}\rangle}{\Delta t}$$

between the coarse-grained true tendency (on the left) and the tendency predicted by the coarse model when it sees the same state (on the right). This is the unresolved tendency. Knowledge of $\langle \boldsymbol{z}\rangle$ alone does not uniquely determine $\boldsymbol{\epsilon}$ because $\langle \mathcal{M}(\boldsymbol{z})\rangle$ depends on the original $\boldsymbol{z}$. However, if one can fit a statistical model $\boldsymbol{P}$ to the dataset of $(\langle \boldsymbol{z}\rangle, \boldsymbol{\epsilon})$, then $-\boldsymbol{P}(\boldsymbol{x}(t))\Delta t$ will be able to serve as an estimate of the error incurred by the coarse model in predicting $\boldsymbol{x}(t + \Delta t)$ given $\boldsymbol{x}(t)$. This motivates the construction of a parametrised model $\mathcal{M}^{\mathrm{p}}$, defined by

$$\mathcal{M}^{\mathrm{p}}(\boldsymbol{x}) = \mathcal{M}^{\mathrm{r}}(\boldsymbol{x}) + \boldsymbol{P}(\boldsymbol{x})\Delta t$$

that simply subtracts the estimated error at each time step of the coarse model. In theory, the solution $\boldsymbol{x}(t)$ obtained by iteration of $\mathcal{M}^{\mathrm{p}}$ will be more accurate than that obtained by iteration of the unparametrised $\mathcal{M}^{\mathrm{r}}$.

The current method of choice for constructing $\boldsymbol{P}$ in weather and climate modelling contexts is machine learning (ML). I will briefly introduce the key concepts, strengths and weaknesses of ML, following Beucler et al. (2022). ML encompasses a broad class of algorithms that programmatically and autonomously develop and apply rules for performing tasks, such as image classification, removing the need for a human to explicitly program the rules. They do this using supplied *training data* that exemplify the task to be performed. When it comes to data-driven parametrisation, the task is function approximation or regression, and the relevant algorithms usually require a training dataset of example inputs (i.e., coarse-scale variables) *labelled* with the desired outputs (i.e., unresolved tendencies). One commonly-used class of algorithms called *neural networks* consist of layers of inter-communicating calculation nodes called neurons and learn to approximate functions by optimising sets of weights associated with the neurons. Another type of algorithm is the *random forest*, which approximates functions by passing the inputs through a series of nested `if/else` decision trees.

The distinguishing advantage offered by ML data-driven parametrisations is their ability to systematically capture complex relationships without relying on artificially simplified physical models derived by humans (Irrgang et al. 2021; Beucler et al. 2022). In addition to constructing new parametrisations from scratch, ML has been used to emulate the action of existing parametrisation schemes in GCMs (e.g., Gentine et al. 2018). These ML emulators have the potential to reduce the computational burden that parametrisation usually imposes, leaving more resources for other tasks such as calling the parametrisation scheme more frequently or increasing the model resolution (Beucler et al. 2022). Furthermore, the ability to emulate multiple existing schemes at once is a step towards unifying the parametrisation of unresolved processes (a long-standing issue; see § 1.4).

ML parametrisations are not without their issues; they are known to be prone to overfitting and often cause their host models to become numerically unstable. They are also, by their very nature, far more difficult to interpret than traditional schemes. It is therefore very difficult to determine *a priori* whether or not they will continue to be valid in climate conditions outside the bounds of their training datasets (Irrgang et al. 2021; Beucler et al. 2022). "Interpretable" or "explainable" ML is an active area of research that aims to address these issues. Another active research topic is "physics-guided ML", the aim of which is to augment ML with known physical constraints and conservation laws that would otherwise be violated (e.g., Yuval, O'Gorman, and Hill 2021).

3 Simple dynamical systems as testbeds for novel approaches

Historically, it has been common practice to develop and test new parametrisation techniques using simpler dynamical systems that share the key nonlinear, multi-scale and chaotic properties of the climate system. These analogue systems provide testbeds where the parametrisation problem is more tractable, avoiding the expense, effort and technicalities of fully-fledged climate models. The simplicity of these systems also helps to ensure the reproducibility of the results obtained. This section will review the progress that has been made using simple analogue systems and argue that, considering the outstanding issues identified in §§ 1 and 2, further research using these systems may yet be warranted. The relevant techniques for constructing and testing parametrisation schemes in simple model frameworks, and the remaining open questions, will be discussed in order to inform future research.

3.1 The Lorenz '96 system

The Lorenz '96 model (henceforth L96) was proposed by eminent meteorologist Edward N. Lorenz in a 1995 paper (Lorenz 1995) on the growth of errors in solutions of dynamical systems and the resulting limits of predictability for those systems. Lorenz proposed to mimic the multi-scale nature of the climate system by coupling two dynamical systems with different characteristic time scales. Following his notation, there is a discrete set of "slow" variables $\{X_k\}_{k=1}^{K}$, each of which has an associated discrete set of "fast" variables $\{Y_{j,k}\}_{j=1}^{J}$. The indices of the slow variables are periodic ($X_{K+1} = X_1$ and $X_0 = X_K$) and the sets of fast variables are arranged end-to-end so that $Y_{J+1,k} = Y_{1,k+1}$ and $Y_{0,k} = Y_{J,k-1}$. The intuition is that the variables represent samples of a field along a latitude circle, as shown in Figure 1 (a reproduction of Figure 1 by Russell et al. (2017)). The evolution of the system is governed by the ordinary differential equations

$$\frac{\mathrm{d}X_k}{\mathrm{d}t} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F - \frac{hc}{b}\sum_{j=1}^{J} Y_{j,k}, \tag{3.1a}$$

$$\frac{\mathrm{d}Y_{j,k}}{\mathrm{d}t} = -cbY_{j+1,k}(Y_{j+2,k} - Y_{j-1,k}) - cY_{j,k} + \frac{hc}{b}X_k. \tag{3.1b}$$

The constant $c \geq 1$ dictates the ratio of the time scale of the Y variables to the time scale of the X variables, b the typical magnitude of X relative to Y, h the strength of the coupling between the two subsystems and F the constant forcing applied to each X variable.

As discussed in § 1.2, the objective is to find $P_k(X_1, \ldots, X_K) \approx -(hc/b)\sum_{j=1}^{J} Y_{j,k}$ such that the solution of the parametrised system

$$\frac{\mathrm{d}X_k}{\mathrm{d}t} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F + P_k(X_1, \ldots, X_K)$$

approximates as accurately as possible the true X_k obtained by solving (3.1). The simple form of L96 allows one to generate a training dataset by numerically solving (3.1) and directly diagnosing the unresolved tendency of X_k as

$$U_k(t) = \frac{X_k(t + \Delta t) - X_k(t)}{\Delta t} - \left[-X_{k-1}(X_{k-2} - X_{k+1}) - X_k\right]. \tag{3.2}$$

As we shall see, it is usually assumed that P_k depends only on X_k (i.e., the parametrisation is *local*) but may also depend on the value of X_k at earlier times. The symmetry of (3.1) under shifts of the k index implies that all the X_k have the same long-term statistics, so it suffices to aggregate all the U_k into a single training dataset and use the resulting parametrisation for all the X_k rather than constructing a separate parametrisation for each variable.

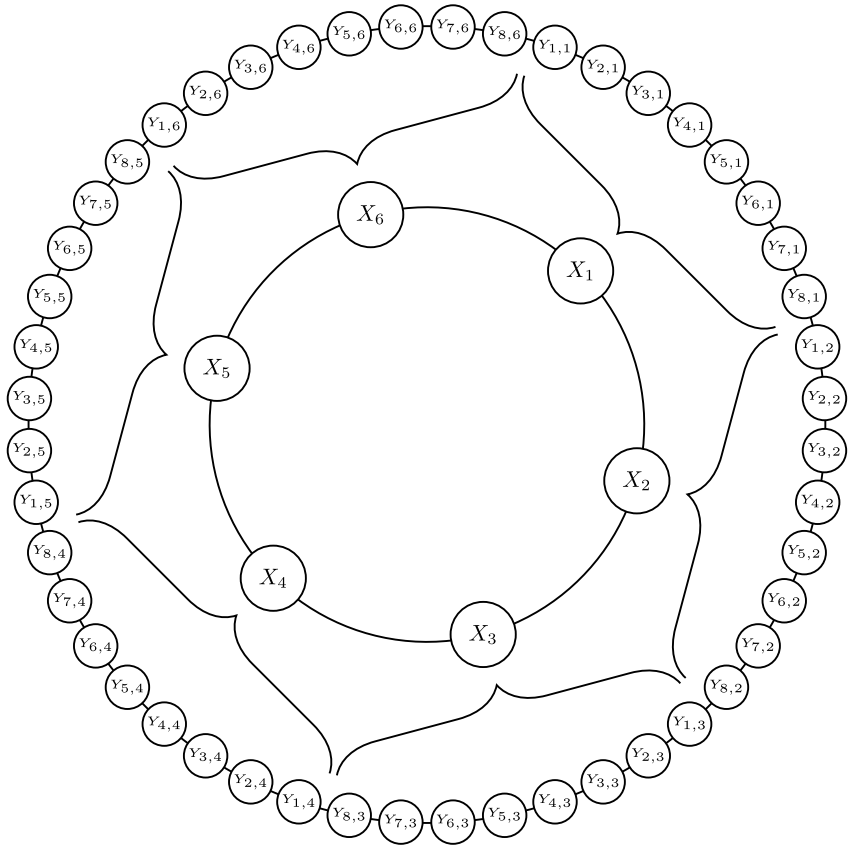

Figure 1: Illustration of the periodic arrangement of the L96 variables, with the slow variables X_k arranged in a circle. Each slow variable is coupled to a neighbouring subset of the similarly arranged fast variables $Y_{j,k}$. Reproduced from Russell et al. (2017), Figure 1.

3.2 Statistical models

Constructing a data-driven parametrisation scheme for L96 requires first choosing the structure of the scheme and then using the training dataset to find the optimal values of the associated free parameters. The simplest structure, popularised in an influential paper by Wilks (2005), consists of a polynomial regression of U against X as a deterministic base, modified by stochastic noise. Wilks' scatterplot of U and X and quartic polynomial least-squares regression, shown in Figure 2, confirm that such a structure is reasonable.

Denote the deterministic polynomial part of the parametrisation by $P_{\text{det}}(X)$. Wilks' proposal was to then model the residuals as an additive, mean-zero noise term $e(t)$, independent of X and updated at each time step, so that

$$P(X) = P_{\text{det}}(X) + e(t). \tag{3.3}$$

Arnold, Moroz, and Palmer (2013) and later Christensen, Moroz, and Palmer (2015) considered the SPPT method (see § 2.1), where instead

$$P(X) = [1 + e(t)]P_{\text{det}}(X). \tag{3.4}$$

The most common form for $e(t)$ in both (3.3) and (3.4) is that of a *first-order autoregressive* or AR(1) model

$$e(t) = \phi e(t - \Delta t) + \sigma z \tag{3.5}$$

where $\phi \in [0, 1]$ and $\sigma \geq 0$ are constants and z is drawn independently from the standard normal distribution at each time step. The AR(1) model introduces memory by having $e(t)$ relax from its value

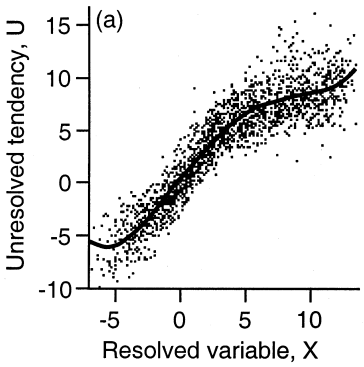

Figure 2: Scatterplot of unresolved tendencies U against large-scale variables X in L96 (black dots), and corresponding quartic polynomial least-squares regression (black line). Data are for forcing $F = 18$. Reproduced from Figure 2a of Wilks (2005).

at the previous time step towards zero, while also adding independent random jumps with standard deviation σ. AR(1) noise is commonly also referred to as *red* noise. The two important special cases of an AR(1) model are $\phi = 0$, which reduces $e(t)$ to white noise without memory, and $\phi = \sigma = 0$, which results in a deterministic parametrisation $P(X) = P_{\text{det}}(X)$. The best estimates of ϕ and σ are straightforwardly obtained by examining the autocorrelation and standard deviation of the residual time series $U(t) - P_{\text{det}}(X(t))$ in the training dataset (see Arnold, Moroz, and Palmer (2013) and Wilks (2011, Chapter 9) for details).

It must be noted that the additive and SPPT schemes (3.3) and (3.4) are inherently limited by their simple form. The additive scheme assumes that the variance of the unresolved tendency is independent of the value of X, and SPPT implies that the variance vanishes when $P_{\text{det}}(X) = 0$. Inspection of Figure 2 indicates that neither of these conditions is strictly satisfied. Arnold, Moroz, and Palmer (2013) therefore proposed two possible modifications of (3.3), with the standard deviation of $e(t)$ being a linear function of either $|X(t)|$ or $|P_{\text{det}}(X(t))|$.

Other studies have experimented with more complex statistical models for the unresolved tendencies. Chorin and Lu (2015) tested NARMAX (nonlinear autoregression moving average with exogenous inputs), a model which represents the tendency as a function of (i) the tendencies estimated at previous time steps (autoregression), (ii) the current and previous values of X (nonlinear; exogenous data), (iii) independent Gaussian noise and (iv) the previous values of the Gaussian noise (moving average). The motivation for NARMAX is that it may be able to capture more complex relationships and memory effects by using more predictors and free parameters. Crommelin and Vanden-Eijnden (2008) and Kwasniok (2012) used Markov chain models, which approximate the continuous range of possible unresolved tendencies U for a given X by a discrete set of allowed states. The model transitions from state to state according to a set of probabilities that depend on X and are estimated from the training dataset. The latest studies have tested machine learning algorithms for L96, reflecting the increasing research interest in ML-based parametrisation schemes for GCMs (see § 2.2). Gagne II et al. (2020) used generative adversarial networks (GANs), which involve pairs of competing neural networks and have the advantage of being stochastic with no need for *ad hoc* perturbations. Bhouri and Gentine (2023) used deterministic but memory-based neural networks trained to directly optimise short-term forecast accuracy without requiring the calculation of the unresolved tendencies U in the training dataset.

3.3 Evaluating parametrisation performance

After choosing and fitting a parametrisation scheme, the next step is to assess its performance. It is important to recognise that there are several contexts in which one might want a parametrisation to perform well. The first important distinction is between *offline* and *online* testing. Offline testing

involves feeding large-scale states into the parametrisation scheme from a pre-computed test dataset that was not used for training (such as a new high-resolution simulation) and measuring the level of agreement between the tendencies predicted by the scheme and the true tendencies. Online testing, on the other hand, involves coupling the parametrisation scheme into the host model and comparing the output of the parametrised model to the corresponding "truth" solution (which might be the output of a high-resolution unparametrised model). Online performance can further be assessed for short-term ("weather") forecasts or long-term ("climate") predictions. The reliability of ensemble forecasts (see § 2.1) must also be determined. As we shall see, the extent to which good performance in more than one category is achievable is a topic of ongoing research.

3.3.1 Offline testing

Offline testing is rarely documented in the L96 literature because it is somewhat more straightforward than online testing. The aim is to ensure that the parametrisation scheme accurately captures the distribution of unresolved tendencies that exist in the true solution. Gagne II et al. (2020) achieved this by integrating the full system (3.1) to obtain a test dataset and estimating the time-aggregated probability density functions (PDFs) of the true unresolved tendencies (3.2) and the tendencies predicted by the parametrisation schemes. They then expressed the difference between these distributions as a scalar quantity using the Hellinger distance, defined for two PDFs p and q by

$$H(p,q) = \frac{1}{2}\int \mathrm{d}x \left(\sqrt{p(x)} - \sqrt{q(x)}\right)^2 . \tag{3.6}$$

3.3.2 Forecast performance

When evaluating online forecast accuracy for L96, studies are primarily concerned with the agreement between a "truth" (or verification) solution of (3.1) and the mean of an ensemble of independently realised, stochastically parametrised forecasts starting from the same or nearby initial conditions. This reflects standard practice in operational numerical weather prediction. A standard approach is to calculate the root mean square error (RMSE)

$$\mathrm{RMSE}(t) = \left(\left\langle \sum_k (X_k^{\mathrm{ens}}(t) - X_k^{\mathrm{ver}}(t))^2 \right\rangle\right)^{1/2}$$

between the ensemble mean forecast $X_k^{\mathrm{ens}}(t)$ and the verification $X_k^{\mathrm{ver}}(t)$, where the mean $\langle\cdot\rangle$ is taken over many forecast-verification pairs with a range of initial conditions (Crommelin and Vanden-Eijnden 2008; Gagne II et al. 2020).

An alternative for single-integration (non-ensemble) forecasts is to take the mean over time, from the initial time up to time t (Bhouri and Gentine 2023). Another accuracy metric is the anomaly correlation (ANCR), which is the Pearson correlation coefficient

$$\mathrm{ANCR}(t) = \left\langle \frac{\sum_k A_k^{\mathrm{ens}}(t) A_k^{\mathrm{ver}}(t)}{\left[\left(\sum_k A_k^{\mathrm{ens}}(t)^2\right)\left(\sum_k A_k^{\mathrm{ver}}(t)^2\right)\right]^{1/2}} \right\rangle$$

between the forecast anomaly $A_k^{\mathrm{ens}}(t) = X_k^{\mathrm{ens}}(t) - \langle X_k^{\mathrm{ver}}(t)\rangle_t$ and the verification anomaly $A_k^{\mathrm{ver}}(t) = X_k^{\mathrm{ver}}(t) - \langle X_k^{\mathrm{ver}}(t)\rangle_t$, again averaged over forecast-verification pairs (Crommelin and Vanden-Eijnden 2008). Other studies (Kwasniok 2012; Arnold, Moroz, and Palmer 2013) have used more complex skill scores to assess forecast accuracy.

Recall from § 2.1 that an ensemble forecast is said to be reliable if the spread of the ensemble members is an accurate estimate of the uncertainty in the ensemble mean. In the context of L96, there are two ways to measure reliabilty. The first (Wilks 2005; Crommelin and Vanden-Eijnden 2008; Kwasniok 2012) is to sort the states X_k of the ensemble members at the chosen lead time from largest to smallest (or

vice versa), and find the rank that the truth state would occupy within the list. The ranks from many ensemble forecasts are aggregated and their distribution plotted on a *rank histogram*, which should ideally be uniform for a well-dispersed ensemble. A U-shaped rank histogram indicates underdispersion; that is, the truth too frequently lies at the extreme ends of or outside the ensemble. An inverted U shape indicates the opposite (overdispersion). Another method is to simply compare the ensemble standard deviation and the RMSE, which should be roughly equal for a reliable forecast (Arnold, Moroz, and Palmer 2013; Gagne II et al. 2020).

3.3.3 Climate prediction performance

There are several classes of metrics and methods available for assessing the accuracy of long-term climate predictions made by parametrised L96 models. The simplest is to compare the low-order moments (mean, variance) of the long-term distributions of X between the parametrised and truth solutions. However, it is often the case for the relatively simple L96 model that parametrised models reproduce these moments quite accurately (Wilks 2005), motivating comparison of the entire PDFs—a more stringent test. In addition to visual comparison of the PDFs, one can use scalar measures such as the Hellinger distance (3.6) (Arnold, Moroz, and Palmer 2013; Gagne II et al. 2020) and the Kolmogorov-Smirnov statistic

$$D_n = \max_X |\Psi_{\text{true}}(X) - \Psi_{\text{model}}(X)|,$$

which is the maximum absolute difference between the cumulative distribution functions Ψ of X in the truth and parametrised models (Wilks 2005; Chorin and Lu 2015; Kwasniok 2012). As a diagnostic tool, it is also possible to compare the joint distributions of X and the unresolved tendency U between the truth and parametrised models (Gagne II et al. 2020).

Another approach is to study the correlation and spectral properties of X. In the time domain, one can compute the autocorrelation function $\langle X_k(t)X_k(t+\tau)\rangle_{t,k}$ or the cross-correlation $\langle X_k(t)X_{k+1}(t+\tau)\rangle_{t,k}$ between neighbouring variables and compare these to their truth counterparts (Crommelin and Vanden-Eijnden 2008; Kwasniok 2012; Chorin and Lu 2015; Gagne II et al. 2020). Gagne II et al. (2020) applied the continuous wavelet transform (CWT) to the $X(t)$ time series to measure the amount of power in the signal as a function of oscillation period. Similar analyses may be performed in "space" by computing the autocorrelation $\langle X_k(t)X_{k+n}(t)\rangle_{t,k}$ with respect to shifts in the k index (Gagne II et al. 2020) or by taking a discrete Fourier transform of X_k with respect to k and deriving wavenumber statistics (Crommelin and Vanden-Eijnden 2008; Kwasniok 2012).

3.4 Lessons learnt from Lorenz '96

3.5 Beyond Lorenz '96

4 Dynamical system case study: Rayleigh-Bénard convection

4.1 Problem statement

Rayleigh-Bénard convection is the motion of a fluid confined between two horizontal isothermal plates, the temperature of the bottom plate being higher than that of the top plate. The governing equations for the flow follow from the Navier-Stokes equations of mass, energy and momentum conservation. The reader is referred to Chandrasekhar (1961) for a detailed derivation; I summarise the assumptions and approximations involved below.

The density, ρ, of the fluid is related to its temperature T by the linear equation of state

$$\rho = \rho_0[1 - \alpha(T - T_0)],$$

where α is the (constant) volume coefficient of thermal expansion and ρ_0 and T_0 are the base-state density and temperature such that $\rho = \rho_0$ when $T = T_0$. The key assumption is that density variations are small ($\alpha(T - T_0) \ll 1$), which allows the governing equations to be simplified under the *Boussinesq approximation*. The Boussinesq approximation involves first writing the pressure, p, of the fluid as

$$p = p_0 - \rho_0 g z + p',$$

where p_0 is an arbitrary constant, g is the acceleration due to gravity and z is the vertical coordinate. p' is the (time-varying) deviation from a hydrostatically balanced background profile $p_0 - \rho_0 g z$ in which the upward pressure gradient force per unit volume $\rho_0 g$ cancels the downward weight force per unit volume $-\rho_0 g$. Since $\alpha(T - T_0) \ll 1$, density variations are neglected everywhere except in their contribution to the weight force, leading to a net buoyant (background pressure gradient plus weight) force per unit mass

$$\frac{\rho_0 - \rho}{\rho_0} g = \alpha(T - T_0) g.$$

With these assumptions in mind, I adopt the governing equations as they are derived by Chandrasekhar (1961):

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} = -\frac{1}{\rho_0} \boldsymbol{\nabla} p' + \alpha(T - T_0) g \hat{\boldsymbol{z}} + \nu \nabla^2 \boldsymbol{u} \quad \text{(momentum conservation)}, \tag{4.1}$$

$$\frac{\partial T}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} T = \kappa \nabla^2 T \quad \text{(energy conservation), and} \tag{4.2}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 \quad \text{(incompressibility)}. \tag{4.3}$$

$\boldsymbol{u}$ is the fluid velocity, t is time, $\hat{\boldsymbol{z}}$ is the upward unit vector, ν is the (constant) kinematic viscosity and κ is the thermal diffusivity (also constant). Notice that the aforementioned buoyancy term $\alpha(T - T_0)g$ appears on the right-hand side of (4.1).

The parametrisation test-bed developed in this work solves the governing equations in a two-dimensional domain $(x, z) \in [0, d] \times [0, L]$, subject to no-slip, isothermal boundary conditions on the top and bottom plates,

$$\boldsymbol{u} = \boldsymbol{0}, \quad T = T_0 + \frac{\delta T}{2} \qquad \text{at } z = 0 \text{ and} \tag{4.4}$$

$$\boldsymbol{u} = \boldsymbol{0}, \quad T = T_0 - \frac{\delta T}{2} \qquad \text{at } z = d, \tag{4.5}$$

and periodic boundary conditions in the horizontal,

$$\boldsymbol{u}(x = 0) = \boldsymbol{u}(x = L) \quad \text{and} \quad T(x = 0) = T(x = L). \tag{4.6}$$

δT is the constant temperature difference between the plates.

4.2 Nondimensionalisation and scale analysis

It is helpful to nondimensionalise the governing equations (4.1)–(4.6); this is not only useful for numerical work but also gives insight into the different flow regimes that are possible. A range of nondimensionalisations are used in fluid dynamics literature; I adopt a common one (see, e.g., Grötzbach 1983; Ouertatani et al. 2008; Stevens, Verzicco, and Lohse 2010) which is suitable for the turbulent convective regime.

For low-viscosity, turbulent flow, a suitable time scale is the *free-fall time* t_0, which is the time for a fluid element with constant temperature $T = T_0 - \delta T$ to fall from the top plate to the bottom plate under the influence of buoyancy ($-g\alpha\delta T$) alone. It is simple to show that

$$t_0 \sim \left(\frac{d}{g\alpha\delta T}\right)^{1/2},$$

ignoring a factor of $\sqrt{2}$. The obvious length and temperature scales are the plate separation d and temperature difference δT, respectively.

Making the substitutions $p'/\rho_0 \to \pi$ and $T - T_0 \to \theta$ in (4.1)–(4.6) and expressing all variables in units of t_0, d and δT leads to the dimensionless equations

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} = -\boldsymbol{\nabla}\pi + \left(\frac{\mathrm{Pr}}{\mathrm{Ra}}\right)^{1/2}\nabla^2\boldsymbol{u} + \theta\hat{\boldsymbol{z}}, \tag{4.7}$$

$$\frac{\partial \theta}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\theta = (\mathrm{Ra}\,\mathrm{Pr})^{-1/2}\nabla^2\theta, \quad \text{and} \tag{4.8}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \tag{4.9}$$

with boundary conditions

$$\boldsymbol{u} = \boldsymbol{0}, \quad \theta = +\frac{1}{2} \qquad \text{at } z = 0, \tag{4.10}$$

$$\boldsymbol{u} = \boldsymbol{0}, \quad \theta = -\frac{1}{2} \qquad \text{at } z = 1, \tag{4.11}$$

$$\boldsymbol{u}(x=0) = \boldsymbol{u}(x=\Gamma) \quad \text{and} \quad \theta(x=0) = \theta(x=\Gamma). \tag{4.12}$$

There are three dimensionless parameters: the aspect ratio of the domain

$$\Gamma \equiv \frac{L}{d},$$

the *Prandtl number*

$$\mathrm{Pr} \equiv \frac{\nu}{\kappa}$$

which measures the relative importance of viscosity (momentum diffusivity) and thermal diffusivity, and the *Rayleigh number*

$$\mathrm{Ra} \equiv \frac{g\alpha d^3\delta T}{\kappa\nu}.$$

The Rayleigh number can be interpreted as the ratio of the time scale for thermal transport by conduction to the time scale for thermal transport by convection. It determines the importance of diffusion for the evolution of $\boldsymbol{u}$ and θ; inspection of (4.7) and (4.8) indicates that low Ra implies strong diffusion and high Ra weak diffusion. Detailed theoretical analysis of the governing equations (see, e.g., Chandrasekhar (1961) and the seminal work by Lord Rayleigh (1916)) reveals that there exists a critical Rayleigh number (dependent on boundary conditions but of order 10^3), below which the equations have a stable equilibrium with the fluid at rest and a linear conductive temperature profile. Above the critical value, the equilibrium is unstable and small perturbations lead to the formation of a regular series of steady, rotating convection cells. If the Rayleigh number is increased much further (Le Quéré (1991) cites $\mathrm{Ra} \approx 2\times10^8$), the solution becomes unsteady and increasingly turbulent. This work is concerned with the turbulent regime, since Rayleigh numbers for atmospheric deep moist convection can be as large as 10^{22} (Chillà and Schumacher 2012).

4.3 Thermal properties

The rate of heat transport across the fluid layer has physical significance for natural realisations of thermally driven convection. It is widely analysed in terms of the dimensionless *Nusselt number*, whose accurate determination is a common theme in the Rayleigh-Bénard convection literature; I therefore introduce this quantity before proceeding. The Nusselt number measures the rate of (vertical) heat transport across a horizontal plane at height z, normalised by the purely conductive rate that would exist if the fluid were at rest (Verzicco and Camussi 1999). Following Chillà and Schumacher (2012), I use the definition (before nondimensionalisation)

$$\mathrm{Nu}(z,t) \equiv \frac{\langle wT\rangle_{A,t} - \kappa\partial\langle T\rangle_{A,t}/\partial z}{\kappa\delta T/d} \tag{4.13}$$

where $\langle\cdot\rangle_{A,t}$ denotes averaging over time and the horizontal plane at height z, and $w = \boldsymbol{u}\cdot\hat{\boldsymbol{z}}$ is the vertical velocity. The rate of heat transport in the numerator has two terms: advection $\langle wT\rangle_{A,t}$ and conduction $-\kappa\partial\langle T\rangle_{A,t}/\partial z$. The denominator $\kappa\delta T/d$ is the rate of heat transport for a linear conductive temperature profile with the fluid at rest.

Another important quantity is the thickness δ_T of the *thermal boundary layer* at each plate where large temperature gradients exist. Chillà and Schumacher (2012) define δ_T as follows: if, on average, the fluid temperature changes with height from $+\delta T/2$ at the lower plate to 0 (the mean value in the well-mixed interior) over a distance δ_T, then

$$\left.\frac{\partial\langle T\rangle_{A,t}}{\partial z}\right|_{z=0} \approx -\frac{\delta T}{2\delta_T}.$$

But if one considers the definition of the Nusselt number (4.13) at $z = 0$, the advection term $\langle wT\rangle_{A,t}$ vanishes due to the $\boldsymbol{u} = \boldsymbol{0}$ boundary condition and

$$\mathrm{Nu}(z=0) = -\frac{d}{\delta T}\left.\frac{\partial\langle T\rangle_{A,t}}{\partial z}\right|_{z=0}.$$

Thus,

$$\delta_T = \frac{d}{2\,\mathrm{Nu}(z=0)}. \tag{4.14}$$

4.4 Resolution dependence of numerical solutions

In constructing a parametrisation test-bed, I will firstly seek reasonably accurate, high-resolution "truth" solutions of (4.7)–(4.12). I will then deliberately reduce the resolution, aiming to produce solutions that are sufficiently "different" that they might reasonably be "improved" by a parametrisation scheme. In this section, I review relevant literature with the aim of establishing how, exactly, under-resolved simulations of Rayleigh-Bénard convection might "differ" from well-resolved ones, and thus what one would hope to "improve". Specifically, I will address the following questions:

- What resolution is necessary for a converged solution?
- Which quantities are most sensitive to insufficient resolution?

4.4.1 Theoretical resolution requirements for accurate simulations

Grötzbach (1983) is recognised as the first to formulate resolution requirements for accurate simulations of Rayleigh-Bénard convection (Chillà and Schumacher 2012; Scheel, Emran, and Schumacher 2013). He identified separate constraints on the mean (i.e., averaged in each spatial direction) grid spacing and the

vertical spacing near the plates; I first discuss the former. Grötzbach reasoned that a numerical model that neglects subgrid-scale effects must have a geometric mean grid spacing $h = (\Delta x \Delta y \Delta z)^{1/3}$ such that

$$h \leq \pi\eta = \pi \left(\frac{\nu^3}{\langle \epsilon \rangle} \right)^{1/4} \tag{4.15}$$

where $\eta \equiv (\nu^3/\langle\epsilon\rangle)^{1/4}$ is the *Kolmogorov length*, the universal smallest relevant length scale for general turbulent flow, and $\langle\epsilon\rangle$ is the spatial and temporal average of the kinetic energy dissipation rate defined by

$$\epsilon(\boldsymbol{x}, t) \equiv \frac{\nu}{2} \sum_{ij} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)^2 \tag{4.16}$$

(Chillà and Schumacher 2012). The inequality (4.15) between h and η can be understood using the Nyquist-Shannon theorem, which states that a sampling frequency $f \geq k/\pi$ is needed to unambiguously reconstruct a signal with maximum wavenumber k; substituting $f = 1/h$, $k = 1/\eta$ leads to the claimed relation.

Grötzbach recognised that the above reasoning was only valid for the mean grid spacing; large gradients in temperature and velocity near the top and bottom plates require finer resolution in those regions. The notion of nearness can be formalised using the thermal boundary layer thickness (4.14), and one asks how many grid points are necessary in this layer. Grötzbach did not give a theoretical argument to derive this number but claimed that 3 points are sufficient for turbulent flows. Shishkina et al. (2010) presented a theoretical argument based on the (experimentally and numerically justified) assumption of laminar *Prandtl-Blasius* flow conditions in the boundary layer and were able to calculate the minimum number of grid points (e.g., 9 for $\mathrm{Ra} = 2 \times 10^9$ and $\mathrm{Pr} = 0.7$). The results agreed with criteria derived in previous numerical experiments. Importantly, the results of Shishkina et al. allow *a priori* determination of vertical resolution requirements, potentially bypassing the time-consuming and expensive process of iteratively running simulations, checking their convergence and updating the resolution.

4.4.2 Resolution-dependence tests and consequences of under-resolution

Performing numerical experiments for a 3D fluid layer, Grötzbach found that RMS velocity and Nusselt number were the most sensitive quantities to insufficient mean grid spacing, but even they increased "only slightly" above the values obtained from well-resolved simulations. He concluded that condition (4.15) was overly restrictive and recommended (for $\mathrm{Pr} > 0.59$) the simplified, approximate version

$$h \lesssim 5.23\,\mathrm{Pr}^{-1/4}\mathrm{Ra}^{-0.3205}.$$

Later work also supports the finding that the Nusselt number is sensitive to under-resolution. Even studying only steady-state convective solutions at moderate Rayleigh number, Le Quéré (1991) found that the maximum and minimum Nusselt numbers were most sensitive to changes in resolution and had the largest uncertainty among existing benchmark solutions. Other studies have used the convergence of the Nusselt number as an indicator that the spatial resolution is sufficient to produce an accurate solution (Ouertatani et al. 2008).

Stevens, Verzicco, and Lohse (2010) performed 3D simulations in a finite cylindrical cavity with the aim of reconciling the apparent disagreement between the Nusselt numbers in previous numerical studies and experimental observations. They found that agreement with experiment could be achieved, but only by using a much higher resolution than the previous studies. They offered the physical explanation that horizontally under-resolved simulations produce insufficient thermal diffusion, leading to systematic overestimation of the buoyancy of convective plumes near the side-walls of the cylinder; this results in Nusselt numbers that exceed experimentally observed values. This led them to conclude that the two criteria of Grötzbach (1983)—for mean grid spacing and for the vertical spacing near the upper and lower plates—are not independent; the definition $h = (\Delta x \Delta y \Delta z)^{1/3}$ in (4.15) allows the horizontal spacing to remain relatively coarse near the plates, provided the vertical spacing is small. Since fine horizontal

resolution is also necessary to accurately capture the dynamics of the thin plumes, they proposed that (4.15) be applied with $h = \max(\Delta x, \Delta y, \Delta z)$ instead.

Some more recent work, however, casts doubt on the notion that the Nusselt number is sensitive to under-resolution and that its convergence is a good indicator that the flow is well-resolved. In assessing the performance of several published computational fluid dynamics codes on the Rayleigh-Bénard problem in a cylindrical cavity, Kooij et al. (2018) identified one higher-order code that reproduced the theoretically predicted scaling of Nu as a function of Ra even when the flow was deliberately under-resolved. On the other hand, the presence of numerical artefacts in the instantaneous temperature field near the bottom plate was a clear indicator of insufficient resolution. Figure 3, a reproduction of part of their Figure 5, shows these artefacts.

Scheel, Emran, and Schumacher (2013) performed similar high-resolution simulations for a cylindrical cavity and also found that the Nusselt number, among other global transport properties, were "fairly insensitive to insufficient resolution, as long as the mean Kolmogorov length [was] resolved" (i.e., (4.15) was satisfied). However, they found that the horizontally averaged or local kinetic energy dissipation rate (4.16) and the corresponding thermal dissipation rate

$$\epsilon_T(\boldsymbol{x}, t) \equiv \kappa \sum_i \left(\frac{\partial T}{\partial x_i} \right)^2 \tag{4.17}$$

were much more sensitive, with their convergence requiring even stricter conditions than (4.15).

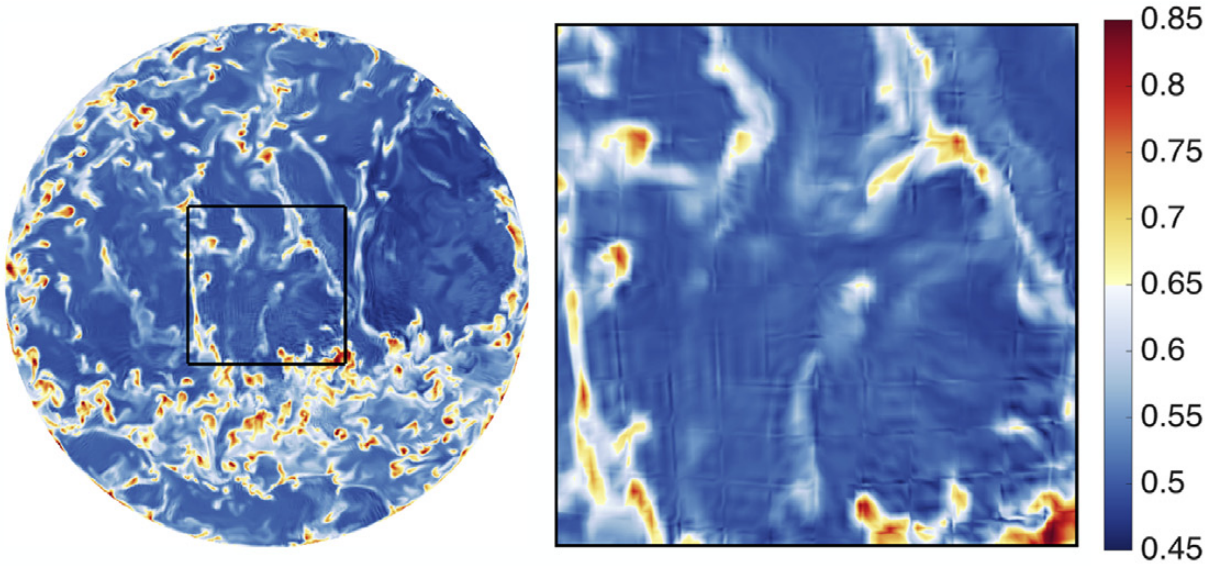

Figure 3: Reproduction of part of Figure 5 by Kooij et al. (2018), showing an instantaneous horizontal temperature profile simulated by the Nek5000 code for Rayleigh-Bénard convection at $\mathrm{Ra} = 10^{10}$. The profile is taken near the bottom plate of a cylindrical cavity and shows a grid-like pattern of numerical artefacts. The right panel is an enlargement of the region inside the black square in the left panel.

5 Summary

References

Alcala, J. and I. Timofeyev (2021). "Subgrid-scale parametrization of unresolved scales in forced Burgers equation using generative adversarial networks (GAN)". *Theor. Comput. Fluid Dyn.* **35**(6). DOI: 10.1007/s00162-021-00581-z.

Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor (2012). "Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models". *Geophys. Res. Lett.* **39**(9). DOI: 10.1029/2012GL051607.

Arnold, H. M., I. M. Moroz, and T. N. Palmer (2013). "Stochastic parametrizations and model uncertainty in the Lorenz '96 system". *Phil. Trans. R. Soc. A* **371**(1991). DOI: 10.1098/rsta.2011.0479.

Berner, J., U. Achatz, L. Batté, L. Bengtsson, A. de la Cámara, H. M. Christensen, M. Colangeli, D. R. B. Coleman, D. Crommelin, S. I. Dolaptchiev, C. L. E. Franzke, P. Friederichs, P. Imkeller, H. Järvinen, S. Juricke, V. Kitsios, F. Lott, V. Lucarini, S. Mahajan, T. N. Palmer, C. Penland, M. Sakradzija, J.-S. von Storch, A. Weisheimer, M. Weniger, P. D. Williams, and J.-I. Yano (2017). "Stochastic parameterization: toward a new view of weather and climate models". *Bull. Am. Meteorol. Soc.* **98**(3). DOI: 10.1175/BAMS-D-15-00268.1.

Beucler, T., I. Ebert-Uphoff, S. Rasp, M. Pritchard, and P. Gentine (2022). "Machine learning for clouds and climate". *ESS Open Archive*. Preprint. DOI: 10.1002/essoar.10506925.1.

Bhouri, M. A. and P. Gentine (2023). "Memory-based parameterization with differentiable solver: application to Lorenz '96". *Chaos* **33**(7). DOI: 10.1063/5.0131929.

Brajard, J., A. Carrassi, M. Bocquet, and L. Bertino (2021). "Combining data assimilation and machine learning to infer unresolved scale parametrization". *Phil. Trans. R. Soc. A* **379**(2194). DOI: 10.1098/rsta.2020.0086.

Chandrasekhar, S. (1961). *Hydrodynamic and hydromagnetic stability*. Oxford: Clarendon Press. ISBN: 9780486319209.

Chillà, F. and J. Schumacher (2012). "New perspectives in turbulent Rayleigh-Bénard convection". *Eur. Phys. J. E* **35**(7). DOI: 10.1140/epje/i2012-12058-1.

Chorin, A. J. and F. Lu (2015). "Discrete approach to stochastic parametrization and dimension reduction in nonlinear dynamics". *PNAS* **112**(32). DOI: 10.1073/pnas.1512080112.

Christensen, H. M., I. M. Moroz, and T. N. Palmer (2015). "Simulating weather regimes: impact of stochastic and perturbed parameter schemes in a simple atmospheric model". *Clim. Dyn.* **44**(7). DOI: 10.1007/s00382-014-2239-9.

Christensen, H. and L. Zanna (2022). "Parametrization in weather and climate models". *Oxford Research Encyclopedia of Climate Science*. Oxford University Press. DOI: 10.1093/acrefore/9780190228620.013.826.

Christensen, H. M. (2020). "Constraining stochastic parametrisation schemes using high-resolution simulations". *Q. J. R. Meteorolog. Soc.* **146**(727). DOI: 10.1002/qj.3717.

Crommelin, D. and E. Vanden-Eijnden (2008). "Subgrid-scale parameterization with conditional Markov chains". *J. Atmos. Sci.* **65**(8). DOI: 10.1175/2008JAS2566.1.

Demaeyer, J. and S. Vannitsem (2018). "Stochastic parameterization of subgrid-scale processes: a review of recent physically based approaches". *Advances in Nonlinear Geosciences*. Ed. by A. A. Tsonis. Cham: Springer International Publishing, pp. 55–85. ISBN: 978-3-319-58895-7. DOI: 10.1007/978-3-319-58895-7_3.

Donner, L. J. and V. T. Phillips (2003). "Boundary layer control on convective available potential energy: implications for cumulus parameterization". *J. Geophys. Res.: Atmos.* **108**(D22). DOI: 10.1029/2003JD003773.

Franzke, C. L. E., T. J. O'Kane, J. Berner, P. D. Williams, and V. Lucarini (2015). "Stochastic climate theory and modeling". *WIREs Clim. Change* **6**(1). DOI: 10.1002/wcc.318.

Gagne II, D. J., H. M. Christensen, A. C. Subramanian, and A. H. Monahan (2020). "Machine learning for stochastic parameterization: generative adversarial networks in the Lorenz '96 model". *J. Adv. Model. Earth Syst.* **12**(3). DOI: 10.1029/2019MS001896.

Gentine, P., M. Pritchard, S. Rasp, G. Reinaudi, and G. Yacalis (2018). "Could machine learning break the convection parameterization deadlock?" *Geophys. Res. Lett.* **45**(11). DOI: 10.1029/2018GL078202.

Grötzbach, G. (1983). "Spatial resolution requirements for direct numerical simulation of the Rayleigh-Bénard convection". *J. Comput. Phys.* **49**(2). DOI: 10.1016/0021-9991(83)90125-0.

Hasselmann, K. (1976). "Stochastic climate models: Part I. Theory". *Tellus* **28**(6). DOI: 10.3402/tellusa.v28i6.11316.

Irrgang, C., N. Boers, M. Sonnewald, E. A. Barnes, C. Kadow, J. Staneva, and J. Saynisch-Wagner (2021). "Towards neural Earth system modelling by integrating artificial intelligence in Earth system science". *Nat. Mach. Intell.* **3**(8). DOI: 10.1038/s42256-021-00374-3.

Kooij, G. L., M. A. Botchev, E. M. A. Frederix, B. J. Geurts, S. Horn, D. Lohse, E. P. van der Poel, O. Shishkina, R. J. A. M. Stevens, and R. Verzicco (2018). "Comparison of computational codes for direct numerical simulations of turbulent Rayleigh-Bénard convection". *Comput. Fluids* **166**. DOI: 10.1016/j.compfluid.2018.01.010.

Kwasniok, F. (2012). "Data-based stochastic subgrid-scale parametrization: an approach using cluster-weighted modelling". *Phil. Trans. R. Soc. A* **370**(1962). DOI: doi:10.1098/rsta.2011.0384.

Lamb, D. (2003). "Cloud microphysics". *Encyclopedia of Atmospheric Sciences*. Ed. by J. R. Holton. Oxford: Academic Press. ISBN: 978-0-12-227090-1. DOI: 10.1016/B0-12-227090-8/00111-1.

Le Quéré, P. (1991). "Accurate solutions to the square thermally driven cavity at high Rayleigh number". *Comput. Fluids* **20**(1). DOI: 10.1016/0045-7930(91)90025-D.

Lord Rayleigh (1916). "On convection currents in a horizontal layer of fluid, when the higher temperature is on the under side". *Philos. Mag.* **32**(192). DOI: 10.1080/14786441608635602.

Lorenz, E. N. (1995). "Predictability: a problem partly solved". Seminar on Predictability (Shinfield Park, Reading). ECMWF. URL: https://www.ecmwf.int/en/elibrary/75462-predictability-problem-partly-solved.

Manabe, S., J. Smagorinsky, and R. F. Strickler (1965). "Simulated climatology of a general circulation model with a hydrologic cycle". *Mon. Weather Rev.* **93**(12). DOI: 10.1175/1520-0493(1965)093<0769:SCOAGC>2.3.CO;2.

McFarlane, N. (2011). "Parameterizations: representing key processes in climate models without resolving them". *WIREs Clim. Change* **2**(4). DOI: 10.1002/wcc.122.

Monin, A. S. and A. M. Yaglom (2007). *Statistical fluid mechanics: mechanics of turbulence*. Ed. by J. L. Lumley. Vol. 1. Mineola, New York: Dover Publications. ISBN: 0-486-45883-0.

Ouertatani, N., N. Ben Cheikh, B. Ben Beya, and T. Lili (2008). "Numerical simulation of two-dimensional Rayleigh-Bénard convection in an enclosure". *C.R. Mec.* **336**(5). DOI: 10.1016/j.crme.2008.02.004.

Palmer, T. N. (2001). "A nonlinear dynamical perspective on model error: a proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models". *Q. J. R. Meteorolog. Soc.* **127**(572). DOI: 10.1002/qj.49712757202.

— (2019). "Stochastic weather and climate models". *Nat. Rev. Phys.* **1**(7). DOI: 10.1038/s42254-019-0062-2.

Palmer, T., G. Shutts, R. Hagedorn, F. Doblas-Reyes, T. Jung, and M. Leutbecher (2005). "Representing model uncertainty in weather and climate prediction". *Annu. Rev. Earth Planet. Sci.* **33**(1). DOI: 10.1146/annurev.earth.33.092203.122552.

Rio, C., A. D. Del Genio, and F. Hourdin (2019). "Ongoing breakthroughs in convective parameterization". *Curr. Clim. Change Rep.* **5**(2). DOI: 10.1007/s40641-019-00127-w.

Russell, F. P., P. D. Düben, X. Niu, W. Luk, and T. N. Palmer (2017). "Exploiting the chaotic behaviour of atmospheric models with reconfigurable architectures". *Comput. Phys. Commun.* **221**. DOI: 10.1016/j.cpc.2017.08.011.

Scheel, J. D., M. S. Emran, and J. Schumacher (2013). "Resolving the fine-scale structure in turbulent Rayleigh-Bénard convection". *New J. Phys.* **15**(11). DOI: 10.1088/1367-2630/15/11/113063.

Shishkina, O., R. J. A. M. Stevens, S. Grossmann, and D. Lohse (2010). "Boundary layer structure in turbulent thermal convection and its consequences for the required numerical resolution". *New J. Phys.* **12**(7). DOI: 10.1088/1367-2630/12/7/075022.

Stevens, B. and S. Bony (2013). "What are climate models missing?" *Science* **340**(6136). DOI: doi:10.1126/science.1237554.

Stevens, R. J. A. M., R. Verzicco, and D. Lohse (2010). "Radial boundary layer structure and Nusselt number in Rayleigh-Bénard convection". *J. Fluid Mech.* **643**. DOI: 10.1017/S0022112009992461.

Verzicco, R. and R. Camussi (1999). "Prandtl number effects in convective turbulence". *J. Fluid Mech.* **383**. DOI: 10.1017/S0022112098003619.

Wilks, D. S. (2005). "Effects of stochastic parametrizations in the Lorenz '96 system". *Q. J. R. Meteorolog. Soc.* **131**(606). DOI: 10.1256/qj.04.03.

— (2011). *Statistical methods in the atmospheric sciences*. 3rd ed. International Geophysics. Amsterdam, Boston: Elsevier/Academic Press. ISBN: 0123850231.

Wouters, J. and V. Lucarini (2012). "Disentangling multi-level systems: averaging, correlations and memory". *J. Stat. Mech: Theory Exp.* **2012**(03). DOI: 10.1088/1742-5468/2012/03/P03003.

— (2013). "Multi-level dynamical systems: connecting the Ruelle response theory and the Mori-Zwanzig approach". *J. Stat. Phys.* **151**(5). DOI: 10.1007/s10955-013-0726-8.

Yuval, J., P. A. O'Gorman, and C. N. Hill (2021). "Use of neural networks for stable, accurate and physically consistent parameterization of subgrid atmospheric processes with good performance at reduced precision". *Geophys. Res. Lett.* **48**(6). DOI: 10.1029/2020GL091363.

Zacharuk, M., S. I. Dolaptchiev, U. Achatz, and I. Timofeyev (2018). "Stochastic subgrid-scale parametrization for one-dimensional shallow-water dynamics using stochastic mode reduction". *Q. J. R. Meteorolog. Soc.* **144**(715). DOI: 10.1002/qj.3396.

Zhang, G. J. (2002). "Convective quasi-equilibrium in midlatitude continental environment and its effect on convective parameterization". *J. Geophys. Res.: Atmos.* **107**(D14). DOI: 10.1029/2001JD001005.