REDUCED MODEL-ERROR SOURCE TERMS FOR FLUID FLOW

Wouter Edeling[1] and Daan Crommelin [1,2]

[1] Centrum Wiskunde & Informatica, Scientific Computing Group
Science Park 123, 1098 XG Amsterdam, The Netherlands
e-mail: {Wouter.Edeling, Daan.Crommelin}@CWI.nl

[2] Korteweg-de Vries Institute for Mathematics, University of Amsterdam
Science Park 105-107, 1098 XG Amsterdam, The Netherlands
e-mail: D.T.Crommelin@uva.nl

Keywords: Model error, data-driven surrogate models, ocean flow

Abstract. *It is well known that the wide range of spatial and temporal scales present in geophysical flow problems represents a (currently) insurmountable computational bottleneck, which must be circumvented by a coarse-graining procedure. The effect of the unresolved fluid motions enters the coarse-grained equations as an unclosed forcing term, denoted as the 'eddy forcing'. Traditionally, the system is closed by approximate deterministic closure models, i.e. so-called parameterizations. Instead of creating a deterministic parameterization, some recent efforts have focused on creating a stochastic, data-driven surrogate model for the eddy forcing from a (limited) set of reference data, with the goal of accurately capturing the long-term flow statistics. Since the eddy forcing is a dynamically evolving field, a surrogate should be able to mimic the complex spatial patterns displayed by the eddy forcing. Rather than creating such a (fully data-driven) surrogate, we propose to precede the surrogate construction step by a procedure that replaces the eddy forcing with a new model-error source term which: i) is tailor-made to capture spatially-integrated statistics of interest, ii) strikes a balance between physical insight and data-driven modelling , and iii) significantly reduces the amount of training data that is needed. Instead of creating a surrogate for an evolving field, we now only require a surrogate model for one scalar time series per statistical quantity-of-interest. Our current surrogate modelling approach builds on a resampling strategy, where we create a probability density function of the reduced training data that is conditional on (time-lagged) resolved-scale variables. We derive the model-error source terms, and construct the reduced surrogate using an ocean model of two-dimensional turbulence in a doubly periodic square domain.*

1 INTRODUCTION

In the numerical simulation of coarse-grained turbulent flow problems one has to cope with small-scale processes which cannot be resolved directly on the numerical grid. The effect of the unresolved eddy field enters the resolved-scale equations as an unclosed forcing term, denoted as the eddy forcing, which is highly complex, dynamic, and shows intricate spatio-temporal correlations. Traditionally, the eddy forcing is approximated by deterministic closure models, i.e. so-called parameterizations. In the context of geophysical flows, such parameterizations are based on e.g. the work of Gent-McWilliams [6], or through the inclusion of a tunable (hyper) viscosity term meant to damp the smallest resolved scales of the model [11].

It is well known that no parameterization scheme is perfect, and attempts have been made to improve their performance. For instance, the authors of [15] analysed the transfer of energy and enstrophy in spectral space for a number of parameterizations, and compared their performance to a high-fidelity reference solution of a two-dimensional turbulent flow case. They proposed a deterministic 'energy fixer' scheme, based on adding a weighted vorticity pattern to the computed vorticity field. Recently, data-driven techniques have been applied as well. For instance the recent work of [10] used artificial neural networks to learn the eddy forcing from a set of high-fidelity snapshots.

However, a general limitation of such deterministic approaches is their inability to represent the strong non-uniqueness of the unresolved scales with respect to the resolved scales [1, 16, 12]. Since the resolved scales are generally defined as the convolution of the full-scale solution with some filter, multiple unresolved states can correspond to the same resolved solution. Thus, in general there is no one-to-one correspondence between the resolved-scale state and the unresolved-scale state, and yet deterministic parameterizations do assume such correspondence. As a result, stochastic methods for representing the unresolved scales have received an increasing amount of attention. Early contributions to this topic in the context of ocean modelling includes the work of [1], where the eddy-forcing is replaced by a space-time correlated random-forcing process. Other notable examples include the work of [9, 20, 7], who construct probability density functions (pdfs) of the eddy forcing using a reference solution.

In this study, we also consider a stochastic surrogate method [17, 16], and as a performance indicator we use the degree by which it is able to capture energy and enstrophy statistics. However, we refrain from an approach that is purely data-driven, i.e. one which attempts to learn the eddy forcing directly from reference data. Instead, we replace the eddy forcing with a simpler 'model-error' source term, which we parameterize based on physical arguments. Specifically, we use the energy and enstrophy transport equations to derive a source term which tracks our chosen target statistics. The only remaining unclosed part of our model-error term is representative of the magnitude of these target statistics, i.e. scalars. As a result, the corresponding surrogate model needs to represent only one (or a few) scalar quantities rather than the full eddy forcing field. This amounts to a large dimension reduction (in this study, a reduction by four orders of magnitude), and as a consequence a large reduction in the amount of required training data, while retaining accuracy in the statistics.

The article is organised as follows. In Section 2 we describe the governing equations and multiscale decomposition. The model-error source term derivation and the surrogate method are outlined in Section 3. Initial results are shown in Section 4, and finally the conclusion and outlook are given in Section 5.

2   GOVERNING EQUATIONS

We study the same model as in [18], i.e. the forced-dissipative vorticity equations for two-dimensional incompressible flow. The governing equations read

$$\frac{\partial \omega}{\partial t} + J(\Psi, \omega) = \nu \nabla^2 \omega + \mu(F - \omega),$$
$$\nabla^2 \Psi = \omega. \tag{1}$$

Here, ω is the vertical component of the vorticity, defined from the curl of the velocity field $\mathbf{V}$ as $\omega := \mathbf{e}_3 \cdot \nabla \times \mathbf{V}$, where $\mathbf{e}_3 := (0, 0, 1)^T$. The stream function Ψ relates to the horizontal velocity components by the well-known relations $u = -\partial \Psi / \partial y$ and $v = \partial \Psi / \partial x$. As in [18], the forcing term is chosen as the single Fourier mode $F = 2^{3/2} \cos(5x) \cos(5y)$. The system is fully periodic in x and y directions over a period of $2\pi L$, where L is a user-specified length scale, chosen as the earth's radius ($L = 6.371 \times 10^6 [m]$). The inverse of the earth's angular velocity Ω^{-1} is chosen as a time scale, where $\Omega = 7.292 \times 10^{-5} [s^{-1}]$. Thus, a simulation time period of a single 'day' can now be expressed as $24 \times 60^2 \times \Omega \approx 6.3$ non-dimensional time units. Given these choices, (1) is non-dimensionalized, and solved using values of ν and μ chosen such that a Fourier mode at the smallest retained spatial scale is exponentially damped with an e-folding time scale of 5 and 90 days respectively. For more details on the numerical setup we refer to [18]. Furthermore, our Python source code for (1) can be downloaded from [4].

Finally, the key term in (1) is the Jacobian, i.e. the nonlinear advection term defined as

$$J(\Psi, \omega) := \frac{\partial \Psi}{\partial x} \frac{\partial \omega}{\partial y} - \frac{\partial \Psi}{\partial y} \frac{\partial \omega}{\partial x}. \tag{2}$$

It is this term that leads to the need for a closure model when (1) is discretized on a relatively coarse grid which lacks the resolution to capture all turbulent eddies.

2.1   Discretization

We solve (1) by means of a spectral method, where we apply a truncated Fourier expansion:

$$\omega_{\mathbf{k}}(x, y, t) = \sum_{\mathbf{k}} \hat{\omega}_{\mathbf{k}}(t) e^{i(k_1 x + k_2 y)},$$
$$\Psi_{\mathbf{k}}(x, y, t) = \sum_{\mathbf{k}} \hat{\Psi}_{\mathbf{k}}(t) e^{i(k_1 x + k_2 y)}. \tag{3}$$

The sum is taken over the components k_1 and k_2 of the wave number vector $\mathbf{k} := (k_1, k_2)^T$, and $-K' \leq k_j \leq K'$, $j = 1, 2$. These decompositions are inserted in (1), and solved for the Fourier coefficients $\hat{\omega}_{\mathbf{k}}$, $\hat{\Psi}_{\mathbf{k}}$ by means of the real Fast Fourier Transform. To avoid the aliasing problem in the nonlinear term (2), we use the pseudo spectral method, such that in practice the maximum resolved wave number is K, where $K \leq 2K'/3$ [14]. [1]

To advance the solution in time we use the second-order accurate AB/BDI2 scheme, which results in the following discrete system of equations [14]

$$\frac{3\hat{\omega}_{\mathbf{k}}^{i+1} - 4\hat{\omega}_{\mathbf{k}}^{i} + \hat{\omega}_{\mathbf{k}}^{i-1}}{2\Delta t} + 2\hat{J}_{\mathbf{k}}^{i} - \hat{J}_{\mathbf{k}}^{i-1} = -\nu k^2 \hat{\omega}_{\mathbf{k}}^{i+1} + \mu\left(\hat{F}_{\mathbf{k}} - \hat{\omega}_{\mathbf{k}}^{i+1}\right),$$
$$-k^2 \hat{\Psi}_{\mathbf{k}}^{i+1} - \hat{\omega}_{\mathbf{k}}^{i+1} = 0. \tag{4}$$

---

[1] We use $N \times N$ grids, with an even $N = 2^p$ (e.g. $p = 7$), such that $N = 2K'$ [14].

Here, $\Delta t = 0.01$ and $\hat{J}_{\mathbf{k}}^i$ is the Fourier coefficient of the Jacobian at time level i, computed with the pseudo spectral technique, and $k^2 := k_1^2 + k_2^2$.

2.2 Multiscale decomposition

As in [18], we apply a spectral filter in order to decompose the full reference solution into a resolved ($\mathcal{R}$) and an unresolved component ($\mathcal{U}$), i.e. we use

$$\hat{\omega}_{\mathbf{k}}^{\mathcal{R}} = P^{\mathcal{R}} \hat{\omega}_{\mathbf{k}}, \qquad \hat{\omega}_{\mathbf{k}}^{\mathcal{U}} = P^{\mathcal{U}} \hat{\omega}_{\mathbf{k}}, \tag{5}$$

where the projection operators $P^{\mathcal{R}}$ and $P^{\mathcal{U}}$ are depicted in Figure 1. Note that the full projection operator $P := \mathcal{P}^{\mathcal{R}} + \mathcal{P}^{\mathcal{U}}$ also removes wave numbers due to the use of the pseudo spectral method.
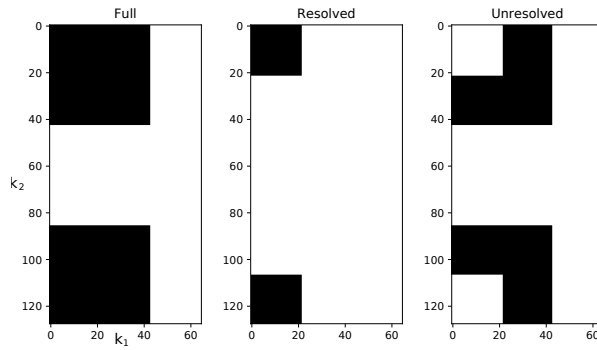

Figure 1: The spectral filter (black=1, white=0) of the full, resolved and unresolved solutions. Due to the fact that we use the real FFT algorithm, only part of the spectrum is computed, as Fourier coefficients with opposite values of $\mathbf{k}$ are complex conjugates in order to enforce real ω and Ψ fields [14].

Applying the resolved projection operator to the governing equations (1) results in the following resolved-scale transport equation

$$\frac{\partial \omega^{\mathcal{R}}}{\partial t} + \mathcal{P}^{\mathcal{R}} J \left(\Psi, \omega \right) = \nu \nabla^2 \omega^{\mathcal{R}} + \mu \left(F^{\mathcal{R}} - \omega^{\mathcal{R}} \right) \tag{6}$$

As mentioned, the key term is the Jacobian (2), since due to its non linearity, $\mathcal{P}^{\mathcal{R}} J \left(\Psi, \omega \right) \neq \mathcal{P}^{\mathcal{R}} J \left(\Psi^{\mathcal{R}}, \omega^{\mathcal{R}} \right)$. We therefore write

$$J \left(\Psi, \omega \right) - J \left(\Psi^{\mathcal{R}}, \omega^{\mathcal{R}} \right) =: r, \tag{7}$$

such that r is the exact subgrid-scale term, commonly referred to as the 'eddy forcing' [1]. The resolved-scale equation (6) can now be written as

$$\frac{\partial \omega^{\mathcal{R}}}{\partial t} + \mathcal{P}^{\mathcal{R}} J \left(\Psi^{\mathcal{R}}, \omega^{\mathcal{R}} \right) = \nu \nabla^2 \omega^{\mathcal{R}} + \mu \left(F^{\mathcal{R}} - \omega^{\mathcal{R}} \right) - \bar{r}. \tag{8}$$

We use the notation $\bar{r} := \mathcal{P}^{\mathcal{R}} r$ for the sake of brevity. A snapshot of the resolved vorticity $\omega^{\mathcal{R}}$ and corresponding resolved eddy forcing $\bar{r}$ is depicted in Figure 2. Notice the fine-grained character of the eddy forcing compared to the vorticity field.
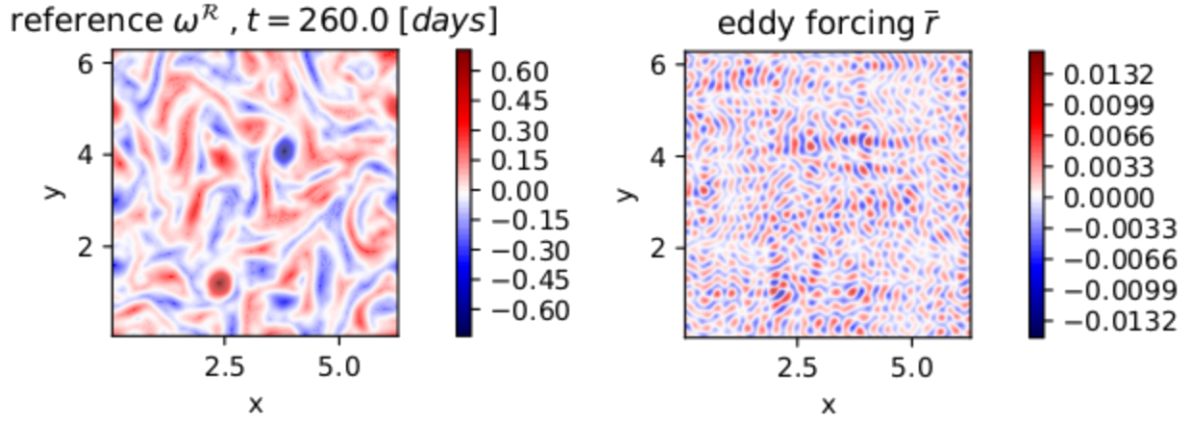
Figure 2: A snapshot of the exact, reference vorticity field $\omega^{\mathcal{R}}$ and the corresponding eddy forcing.

2.3 Prediction of climate statistics

Ultimately, our goal is to integrate (8) in time, such that we can compute the long-term climate statistics of the energy $E^{\mathcal{R}}$ and enstrophy $Z^{\mathcal{R}}$ densities, defined as

$$E^{\mathcal{R}} := \frac{1}{2}\left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi}\int_0^{2\pi} \mathbf{V}^{\mathcal{R}} \cdot \mathbf{V}^{\mathcal{R}} \mathrm{d}x\mathrm{d}y = -\frac{1}{2}\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right), \qquad (9)$$

$$Z^{\mathcal{R}} := \frac{1}{2}\left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi}\int_0^{2\pi} \left(\omega^{\mathcal{R}}\right)^2 \mathrm{d}x\mathrm{d}y = \frac{1}{2}\left(\omega^{\mathcal{R}}, \omega^{\mathcal{R}}\right). \qquad (10)$$

Here $\mathbf{V}^{\mathcal{R}}$ is the two-dimensional vector of the resolved velocity components in x and y direction. For conciseness, we use the short-hand notation

$$(\alpha, \beta) = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi}\int_0^{2\pi} \alpha\beta \,\mathrm{d}x\mathrm{d}y, \qquad (11)$$

to denote the integral of the product $\alpha\beta$ normalized by the area of the flow domain. The derivation of the last equality of (9) can be found in Appendix A.

3 EDDY-FORCING SURROGATE

We cannot integrate (8) since it is still unclosed (due to the ω and Ψ dependence of (7)), a problem which we aim to solve by creating a data-driven surrogate of $\bar{r}$, denoted by $\widetilde{r}$. For our present purpose, we define an 'ideal' surrogate $\widetilde{r}$ for the eddy forcing as one which satisfies the following set of requirements:

1. **Data-driven**: In absence of a single 'best' deterministic parameterization of r, we opt for a model inferred from a pre-computed database of high-fidelity reference data.

2. **Stochastic**: In general, the resolved scales are defined as a convolution of the full solution with some (spatial/spectral) filter. As a result there is no longer just a single unresolved-scale field that is consistent with the resolved-scale solution. This ambiguity provides us with the motivation for a stochastic model for the unresolved, small-scale fields.

3. **Correlated in space and time**: As demonstrated by Figure 2, the reference eddy forcing shows complex spatial structures. A surrogate of the full eddy forcing would ideally reflect these as well.

5

4. **Conditional on the resolved variables**: The resolved and unresolved scales are in reality two-way coupled. Hence, the eddy-forcing surrogate should not be independent from the resolved solution.

5. **Pre-computed & cheap**: While the reference database can be computationally expensive to compute, the resulting data-driven surrogate must be cheap.

6. **Extrapolates well**: To justify the cost of creating the reference database in the first place, the data-driven model must be able to predict the chosen quantity of interest well, *substantially beyond* the (time) domain of the data.

As mentioned, we will measure the performance of a surrogate model by its ability to accurately represent the statistics of (9)-(10). Thus, we do not expect from the resolved-scale model forced by the surrogate the ability to produce individual flow fields which are in absolute lockstep with the high-fidelity data, especially considering the stochastic nature of the surrogate.

One possible course of action, explored in e.g. [17, 10], is to directly create a *full-field* surrogate $\widetilde{r}(x, y; t) \in \mathbb{R}^{N \times N}$, using a database reference snapshots in time of the exact eddy forcing (7). Here, N is the number of grid points in one spatial direction, typically 2^7, 2^8 or higher. Constructing a full-field, dynamic surrogate of a quantity as complex as the eddy forcing is a challenging task, and storing a potentially large amount of reference snapshots can lead to high memory requirements [17]. We therefore propose to precede the surrogate construction step with a procedure that significantly compresses the training data.

3.1 Reduced surrogate

Note that our statistical quantities of interest (9) and (10) are scalars. Instead of creating a full-field $N \times N$ surrogate $\widetilde{r}(x, y; t)$, we will first replace the exact $\overline{r}$ in (8) with a simpler alternative, where the unclosed component is reflective of the size of the statistical quantities we aim to approximate in the first place. A simple option is to specify

$$-\overline{r}(x, y; t) = \tau(t) \omega^{\mathcal{R}}(x, y; t), \tag{12}$$

where $\tau(t)$ is an unknown, time-varying scalar. Clearly, this choice is arbitrary, and (12) will not match the eddy forcing (7). Instead, we think of (12) as an example of a 'model-error term', meant to correct the unparameterized ($\overline{r} = 0$) model in some sense. In our case, a deviation from the exact eddy forcing does not pose a problem because of the freedom that integrated quantities-of-interest give us, such that we only need our $\omega^{\mathcal{R}}$ and $\Psi^{\mathcal{R}}$ fields to approximate the truth in the weak sense of (9) and (10). We can examine the effect of (12) on the evolution equations of $E^{\mathcal{R}}$ and $Z^{\mathcal{R}}$, and subsequently combine physical insight with a data-driven approach to find the time series of τ that constrains their evolution to the reference values. A *reduced surrogate* now only needs to be constructed from this scalar time series, instead of from the full-field evolution of (7).

The evolution equation of $E^{\mathcal{R}}$ (see Appendix A) satisfies

$$\frac{\mathrm{d}E^{\mathcal{R}}}{\mathrm{d}t} = -\left(\psi^{\mathcal{R}}, \frac{\partial \omega^{\mathcal{R}}}{\partial t}\right) = -2\nu Z^{\mathcal{R}} - 2\mu U^{\mathcal{R}} - 2\mu E^{\mathcal{R}} + \left(\psi^{\mathcal{R}}, \overline{r}\right), \tag{13}$$

where we denote the integral $\left(\Psi^{R}, F\right)/2$ as $U^{\mathcal{R}}$. If we insert (12) into (13), the last term on the right-hand side becomes

$$\left(\psi^{\mathcal{R}}, \overline{r}\right) = -\tau\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right) = 2\tau E^{\mathcal{R}}. \tag{14}$$
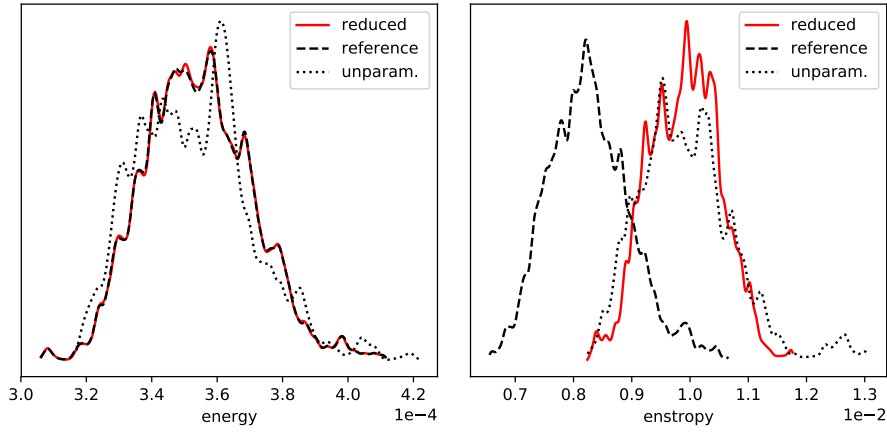
Figure 3: The pdfs of the energy (left) and enstrophy (right), of the reduced ($\overline{r} = \tau\omega^{\mathcal{R}}$), reference ($\overline{r}$ given by (7)) and unparameterised ($\overline{r} = 0$) solution.

The last equality follows from the definition (9). Thus, the physical insight is that (12) leads to the additional term $2\tau E^{\mathcal{R}}$, which either acts to produce or dissipate $E^{\mathcal{R}}$ depending on the sign of τ. Let us denote the difference between the projected reference energy and $E^{\mathcal{R}}$ as $\Delta E := E - E^{\mathcal{R}}$, where $E := -\left(\mathcal{P}^{\mathcal{R}}\Psi, \omega\right)/2$. Any quantity without superscript, e.g. E or ω, is a reference quantity computed from (1). Now, for the data-driven determination of the τ time series, we require τ to be positive when $\Delta E > 0$, i.e. to increase production when $E^{\mathcal{R}}$ is too low, and to dissipate energy when $\Delta E < 0$. We parameterize τ via an analytic relationship which reflects this property:

$$\tau := \tau_{max} \tanh\left(\frac{\Delta E}{E^{\mathcal{R}}}\right). \tag{15}$$

Here, τ_{max} is a user-specified constant, which we set to one for now. During the training period, we can compute (15) every Δt, building up a reference time series.

   To test the validity of our approach, we run the system (8) for a simulation period of 8 years. Besides τ, at every Δt we also sample the energy and enstrophy of the reference, reduced and unparameterised solution, i.e. using $\overline{r}$ given by (7), (12) and zero respectively[2]. The energy and enstrophy probability density functions (pdfs) generated from those samples can be found in Figure 3. By virtue of (15), the energy pdfs of the reference and the reduced solution practically overlap. This demonstrates that it is possible to obtain statistically-equivalent energy solutions using training data reduced by a factor of N^2 compared to the full-field surrogate case[3].

   However, we have two quantities of interest, and (12) also has an effect on the $Z^{\mathcal{R}}$ equation (a term $2\tau Z^{\mathcal{R}}$ appears). Since we train τ to track $\mathcal{P}^{\mathcal{R}}E$, we cannot expect a perfect $Z^{\mathcal{R}}$ pdf, and in fact, Figure 3 shows that the situation does not improve upon the unparameterised model, which displays a large bias in $Z^{\mathcal{R}}$ values. Rather than trying to construct a different τ which is some compromise between accuracy in $E^{\mathcal{R}}$ and $Z^{\mathcal{R}}$, we opt for two separate time series, each of which acts on either the energy or enstrophy evolution equation alone.

---

[2]Note that no surrogate is used yet, we are generating a large set of training data.
[3]In the example of Figure 3, $N^2 = 128^2 = 16384$.

3.2 Orthogonal patterns

We replace our initial simple choice (12) with

$$-\overline{r} = \tau_E \Psi' + \tau_Z \omega', \tag{16}$$

where Ψ' and ω' are patterns of the resolved vorticity and stream function. We choose Ψ' such that $\tau_E \Psi'$ only acts on the $E^{\mathcal{R}}$ equation, and produces no additional source term in the enstrophy equation. The converse must be true for the $\tau_Z \omega'$ term. This will allow us to train τ_E on ΔE alone, and τ_Z only on $\Delta Z := Z - Z^R$. Since the $E^{\mathcal{R}}$ and $Z^{\mathcal{R}}$ evolution equations are forced by $-\left(\Psi^R, \partial \omega^{\mathcal{R}}/\partial t\right)$ and $\left(\omega^{\mathcal{R}}, \partial \omega^{\mathcal{R}}/\partial t\right)$ respectively (see (13) and appendix A), this suggests a Gram-Schmidt type of approach to make Ψ' orthogonal to $\left(\omega^{\mathcal{R}}, \cdot\right)$ and likewise for ω' and $\left(\Psi^R, \cdot\right)$. Setting:

$$\Psi' = \psi^{\mathcal{R}} - \frac{\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right)}{\left(\omega^{\mathcal{R}}, \omega^{\mathcal{R}}\right)}\omega^{\mathcal{R}} \ \ \text{and} \ \ \omega' = \omega^{\mathcal{R}} - \frac{\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right)}{\left(\psi^{\mathcal{R}}, \psi^{\mathcal{R}}\right)}\psi^{\mathcal{R}}, \tag{17}$$

yields

$$\left(\omega^{\mathcal{R}}, \tau_E \Psi'\right) = 0 \ \ \text{and} \ \ \left(\psi^{\mathcal{R}}, \tau_Z \omega'\right) = 0. \tag{18}$$

The additional source term in the $E^{\mathcal{R}}$ equation now becomes

$$-\left(\psi^{\mathcal{R}}, \tau_E \Psi'\right) = -\tau_E \left(\psi^{\mathcal{R}}, \psi^{\mathcal{R}}\right) + \tau_E \frac{\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right)^2}{\left(\omega^{\mathcal{R}}, \omega^{\mathcal{R}}\right)} = 2\tau_E \left[\frac{\left(E^{\mathcal{R}}\right)^2}{Z^{\mathcal{R}}} - S^{\mathcal{R}}\right] := 2\tau_E S' \tag{19}$$

Here, we defined the integrated square stream function as $S^{\mathcal{R}} := \left(\psi^{\mathcal{R}}, \psi^{\mathcal{R}}\right)/2$. Since $\left(E^{\mathcal{R}}\right)^2/Z^R - S^{\mathcal{R}}$ has the dimension of the squared stream function, we introduce the final shorthand notation $S' := \left(E^{\mathcal{R}}\right)^2/Z^{\mathcal{R}} - S^{\mathcal{R}}$ in (19). In a similar vein, (16) produces the following source term in the $Z^{\mathcal{R}}$ equation:

$$2\tau_Z Z' \ \ \text{with} \ \ Z' := Z^{\mathcal{R}} - \frac{\left(E^{\mathcal{R}}\right)^2}{S^{\mathcal{R}}}. \tag{20}$$

We parameterise τ_E and τ_Z using the same procedure as in Section 3.1, only now we need to incorporate the sign of S' and Z' to correctly activate either the production or dissipation of $E^{\mathcal{R}}$ and $Z^{\mathcal{R}}$, i.e.

$$\tau_E := \tau_{E,max} \tanh\left(\frac{\Delta E}{E^{\mathcal{R}}}\right) \cdot \text{sgn}(S') \ \ \text{and} \ \ \tau_Z := \tau_{Z,max} \tanh\left(\frac{\Delta Z}{Z^{\mathcal{R}}}\right) \cdot \text{sgn}(Z'). \tag{21}$$

Again, we leave the proper estimation of parameters for a later study, and simply set $\tau_{E,max} = \tau_{Z,max} = 1$. Furthermore, $\text{sgn}(X) = 1$ when $X \geq 1$ and -1 otherwise. Repeating the simulation of Section 3.1, inserting (16) in (8) yields the results depicted in Figure 4. Now, both pdfs match the reference well. Only a very small discrepancy in the $E^{\mathcal{R}}$ pdf can be observed, which might fixed by tuning $\tau_{E,max}$. The corresponding τ_E, τ_Z reference time series are shown in Figure 5.
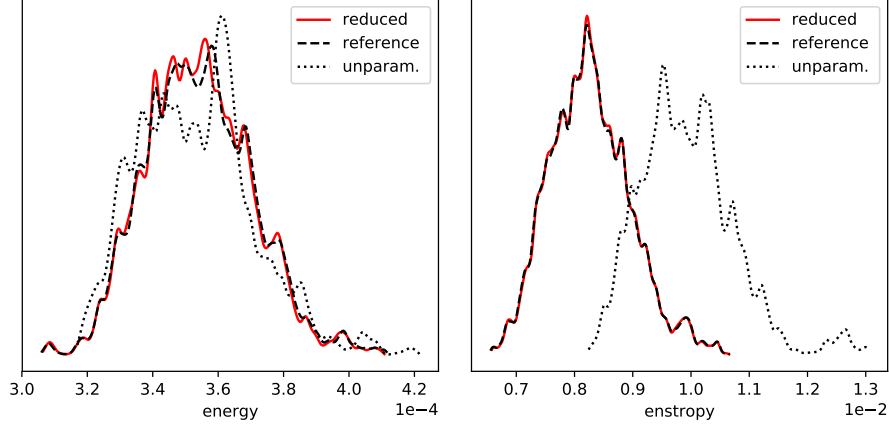
Figure 4: The pdfs of the energy (left) and enstrophy (right), of the reduced ($\bar{r} = \tau_E \Psi' + \tau_Z \omega'$), reference ($\bar{r}$ given by (7)) and unparameterised ($\bar{r} = 0$) solution.
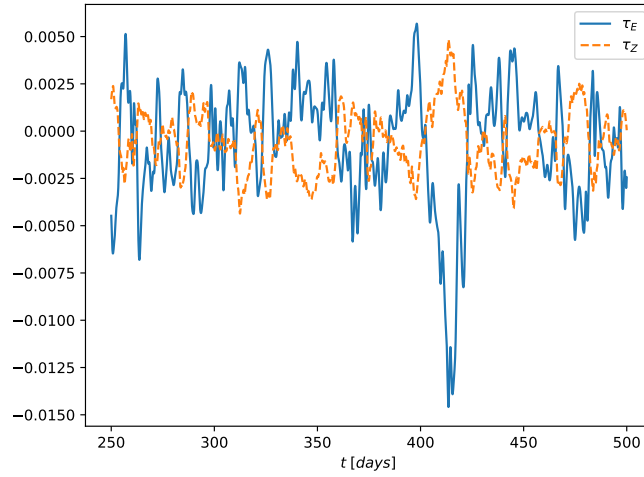

Figure 5: Training time series of τ_E and τ_Z over 500 days. Note that there seems to be a negative correlation between the two time series.

9

3.3 Surrogate construction

We will build on the resampling stategies as developed by [16, 2]. In general, these methods model the unresolved term at time t_{i+1} by sampling from the conditional probability distribution of the reference data. In our case, we keep the functional forms of (21), such that ΔE and ΔZ can be chosen as the unresolved terms in need of a surrogate model:

$$\widetilde{\Delta E}_{i+1} \sim \Delta E_{i+1} \mid \mathcal{E}_i, \ \mathcal{E}_{i-1}, \cdots$$
$$\widetilde{\Delta Z}_{i+1} \sim \Delta Z_{i+1} \mid \mathcal{Z}_i, \ \mathcal{Z}_{i-1}, \cdots \tag{22}$$

Here, $\widetilde{\Delta E}_{i+1}$ denotes the data-driven resampling surrogate at time t_{i+1}, whereas as ΔE_{i+1} represent actual reference data from the training run, and likewise for $\widetilde{\Delta Z}_{i+1}$. The set of 'conditioning variables' $\mathcal{E}_i, \mathcal{Z}_i$ etc contain variables from the resolved model. They can be (functions of) $E^{\mathcal{R}}$, S' or any other (scalar) quantity, as long as we also have access to it outside the training period. Examples of these conditional distributions are $\Delta E_{i+1} \mid E_i^{\mathcal{R}}$ and $\Delta Z_{i+1} \mid \widetilde{\Delta Z}_i, Z_i^{\mathcal{R}}$. We could assume a Markov property ($\Delta E_{i+1} \mid \mathcal{E}_i$), or build in a larger memory. Note that by design, (22) already satisfies many of the properties listed in Section 3, e.g. it is data-driven, stochastic and conditioned on resolved variables.

The main challenges with this approach are twofold. Clearly, the first challenge concerns the actual formation of the conditional distribution, i.e. how to map the observed conditioning variables to plausible subsets of ΔE_{i+1} and ΔZ_{i+1} samples from which $\widetilde{\Delta E}_{i+1}$ and $\widetilde{\Delta Z}_{i+1}$ can be randomly sampled. The second challenge concerns the proper choice of conditioning variables, which is somewhat reminiscent of the choice of 'features' in a machine-learning context.

3.4 Building the distribution

We will illustrate the approach using ΔE, the same procedure applies for ΔZ. To map $\mathcal{E}_i$ to some subset of plausible ΔE_{i+1} values we use the so-called 'binning' approach of [16]. First, consider a snapshot sequence of ΔE

$$\mathbf{\Delta E}_1^S = \{\Delta E_1, \Delta E_2, \cdots, \Delta E_i, \cdots, \Delta E_S\}, \tag{23}$$

where i is the time index. In addition, we also have snapshots of corresponding conditioning variables

$$\mathbf{E}_1^S = \{\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_S\}. \tag{24}$$

Let C be the total number of time-lagged conditioning variables used in (22). We then proceed by creating C-dimensional disjoint bins[4], each bin spanning a unique conditioning variable range, and containing a number of associated ΔE values, see Figure 6. Note that not all bins may contain samples, especially if two or more conditioning variables are used. If during prediction an empty bin is sampled, the data of the nearest bin (in Euclidean sense) is used instead. Once a bin is selected by $\mathcal{E}_i$, the resulting subset of ΔE values can be sampled randomly, or one might sample from the local bin average instead, leading to a deterministic prediction.

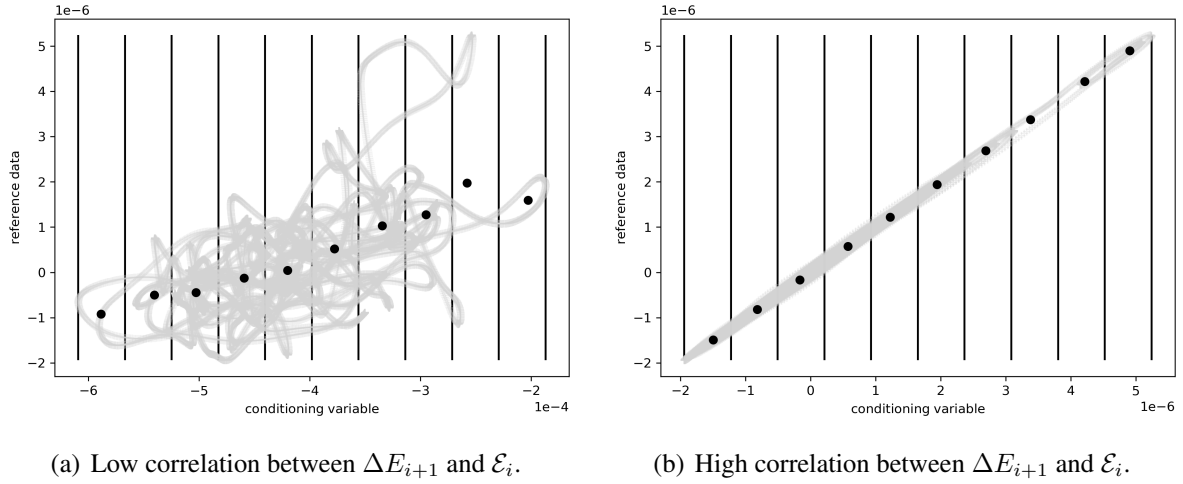[4]We used equidistant bins, but this is not a hard requirement.

(a) Low correlation between ΔE_{i+1} and $\mathcal{E}_i$.　　　　(b) High correlation between ΔE_{i+1} and $\mathcal{E}_i$.

Figure 6: Two binning objects, with the reference ΔE_{i+1} data on the vertical axis and the conditioning variable $\mathcal{E}_i$ on the horizontal axis. Vertical lines separate the different bins, and the black dots represent the local bin means.

3.5   Choice of conditioning variables

Ideally we would like the conditioning variables of (22) to display some correlation with ΔE_{i+1} and ΔZ_{i+1}. In this case, the range of plausible reference values in the selected subset is smaller. Consider the two bins depicted in Figure 6, each with 1 conditioning variable ($\Delta E_{i+1}\mid \mathcal{E}_i$). The binning object of Figure 6(a) shows considerable less correlation between $\mathcal{E}_i$ and ΔE_{i+1} than its counterpart in Figure 6(b). As a result, each bin contains a larger spread in possible ΔE values, leading to more noisy $\widetilde{\Delta E}_{i+1}$ predictions.

We continue by drawing up a list of candidate conditioning variables, and computing the temporal correlation coefficients

$$\rho\left(\Delta E_{i+1}, \mathcal{E}_i\right) = \frac{Cov\left[\Delta E_{i+1}, \mathcal{E}_i\right]}{\sigma\left(\Delta E_{i+1}\right)\sigma\left(\mathcal{E}_i\right)} \ \ \text{and} \ \ \rho\left(\Delta Z_{i+1}, \mathcal{Z}_i\right) = \frac{Cov\left[\Delta Z_{i+1}, \mathcal{Z}_i\right]}{\sigma\left(\Delta Z_{i+1}\right)\sigma\left(\mathcal{Z}_i\right)} \tag{25}$$

from a reference time series of 500 days. Here $Cov\left(\cdot,\cdot\right)$ is the covariance operator and $\sigma\left(\cdot\right)$ is the standard deviation. Specifically, we will select individual source terms from the $E^{\mathcal{R}}$ and $Z^{\mathcal{R}}$ equations as candidate $\mathcal{E}_i$ and $\mathcal{Z}_i$, the rationale being that these will also (in part) drive the evolution equations of ΔE and ΔZ. The complete list, including the correlation coefficient values, is shown in Table 1. Previously undefined conditioning variables (occurring in the $Z^{\mathcal{R}}$ equation), are $V^{\mathcal{R}} := \left(\omega^{\mathcal{R}}, F\right)/2$ and $O^{\mathcal{R}} := \left(\nabla^2\omega^{\mathcal{R}}, \omega^{\mathcal{R}}\right)/2$. This strategy for selecting candidate conditioning variables is reasonable, as many show substantial correlation with the reference data, hovering around the ± 0.5 mark. Clear exceptions are $E^{\mathcal{R}}$ (which correlates much less), and $\tau_E S'$, $\tau_Z Z'$, which show very high correlation.

4   RESULTS

This section contains the initial exploratory results of the methodology outlined in the preceding sections. For validation and training purposes we ran the reference model (1) for a simulation period of 8 years, storing reference data and conditioning variables every Δt. Here, is amounts to roughly 1.8×10^6 snapshots per variable. When predicting, the training data must be stored in memory to allow for fast resampling. If the reference snapshots are full field, this can lead to high memory requirements [17]. Subsampling the reference data reduces the memory constraints, although this leads to a surrogate with an intrinsic time step that is larger

	$\mathcal{E}_i,\ \mathcal{Z}_i$	$\rho\left(\Delta E_{i+1}, \mathcal{E}_i\right)$	$\rho\left(\Delta Z_{i+1}, \mathcal{Z}_i\right)$
$Z^{\mathcal{R}}:$	$\left(\omega^{\mathcal{R}}, \omega^{\mathcal{R}}\right)/2$	0.4017	0.336
$E^{\mathcal{R}}:$	$-\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right)/2$	0.1401	0.0951
$U^{\mathcal{R}}:$	$\left(\psi^{\mathcal{R}}, F\right)/2$	0.5497	0.598
$S^{\mathcal{R}}:$	$\left(\psi^{\mathcal{R}}, \psi^{\mathcal{R}}\right)/2$	-0.5091	-0.4857
$V^{\mathcal{R}}:$	$\left(\omega^{\mathcal{R}}, F\right)/2$	-0.5467	-0.5965
$O^{\mathcal{R}}:$	$\left(\nabla^2\omega^{\mathcal{R}}, \omega^{\mathcal{R}}\right)/2$	-0.4993	-0.4394
$\tau_E S':$	$\tau_E\left(\left(E^{\mathcal{R}}\right)^2/Z^{\mathcal{R}} - S^R\right)$	0.9484	0.8876
$\tau_Z Z':$	$\tau_Z\left(Z^R - \left(E^{\mathcal{R}}\right)^2/S^{\mathcal{R}}\right)$	0.8915	0.999

Table 1: Correlation coefficients.

than the Δt of (4), and thus can only be updated after a certain number of Δt time cycles [2]. A clear advantage of our current surrogate approach, is that we can store the full 8 year reduced training set in memory, without the need for subsampling.

We subdivide the results into tests of increasing complexity:

T1: A one-way coupled simulation where the resolved equation (8) provides the conditioning variables, without replacing $\bar{r} = \tau_E\left(\Delta E\right)\Psi' + \tau_Z\left(\Delta Z\right)\omega'$ in (8) with the surrogate $\widetilde{r} = \tau_E(\widetilde{\Delta E})\Psi' + \tau_Z(\widetilde{\Delta Z})\omega'$. The surrogates $\widetilde{\Delta E}$ and $\widetilde{\Delta Z}$ are not extrapolated, i.e. they are constructed using the full 8 year reference data set, so no simulation outside the time period of the training data takes place.

T2: A two-way coupled simulation, still without surrogate extrapolation.

T3: A two-way coupled simulation with surrogate extrapolation.

4.1 Results T1

T1 serves as a verification of our code, as in this case the exact ΔE and ΔZ are still used in (21) to compute τ_E and τ_Z. Now, if implemented correctly, surrogates such as $\widetilde{\Delta E}_{i+1} \sim \Delta E_{i+1} \mid (\tau_E S')_i$ and $\widetilde{\Delta Z}_{i+1} \sim \Delta Z_{i+1} \mid (\tau_Z Z')_i$, must follow the reference data closely, given the high correlations displayed in Table 1. This is confirmed by the results of Figure 7.

4.2 Results T2

T2 is the first real test of the surrogate method due to its two-way coupled nature. As a result, trajectories of $\widetilde{\Delta E}$ and $\widetilde{\Delta Z}$ can no longer be expected to follow the reference data. Discrepancies between the exact (reduced) eddy forcing (16) and its surrogate will cause the model forced by the surrogate to develop its own dynamics. We reiterate here that our goal is to predict the time-averaged flow statistics, which might still be feasible if we are not in absolute lockstep with ΔE and ΔZ. Even two full-scale simulations with slightly different initial conditions will diverge from each other (due to their turbulent nature), yet can converge in a statistical sense.

We tested a variety of surrogates, which differed through the set of selected conditioning variables. All were Markovian in character, conditioned on variables from the previous time step alone. Thus far, almost all considered surrogates improved upon the $Z^{\mathcal{R}}$ bias of the un-parameterized model, although they showed some varying performance amongst each other.
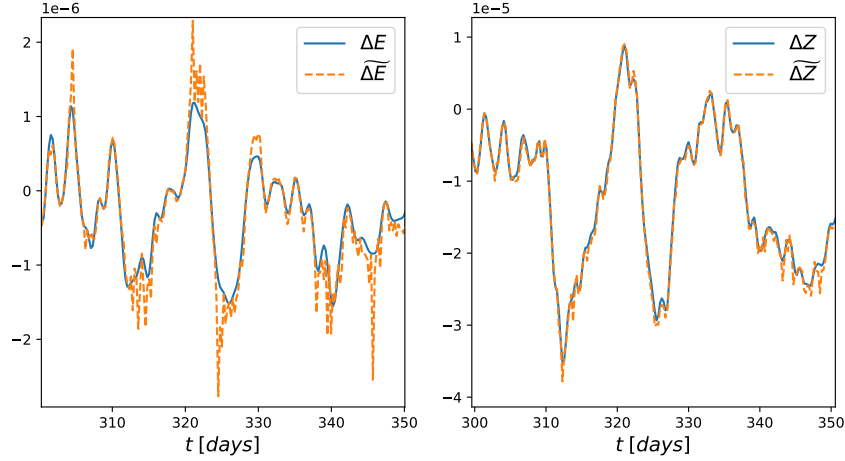
Figure 7: T1 time series for ΔE and ΔZ and their corresponding surrogates over a 50 day period. The $\widetilde{\Delta E}$ surrogate is noisier due to the lower correlation with its conditioning variable (see Table 1).
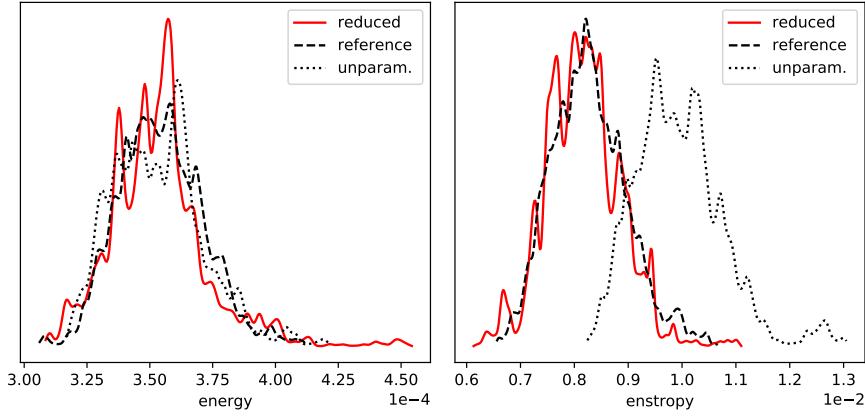

Figure 8: The pdfs of the energy (left) and enstrophy (right), of the reduced surrogate ($\widetilde{r} = \tau_E\left(\widetilde{\Delta E}\right)\Psi' + \tau_Z\left(\widetilde{\Delta Z}\right)\omega'$), reference ($\overline{r}$ given by (7)) and unparameterised ($\overline{r} = 0$) solution. The surrogates were both conditioned on $Z^{\mathcal{R}}, E^{\mathcal{R}}, U^{\mathcal{R}}, S^{\mathcal{R}}$ of the previous time step.

For brevity, we only show a representative sample of results. Consider the results of Figure 8, which shows the pdfs obtained using the surrogates $\Delta E_{i+1} \mid Z_i^{\mathcal{R}}, E_i^{\mathcal{R}}, U_i^{\mathcal{R}}, S_i^{\mathcal{R}}$ and $\Delta Z_{i+1} \mid Z_i^{\mathcal{R}}, E_i^{\mathcal{R}}, U_i^{\mathcal{R}}, S_i^{\mathcal{R}}$, with 10 bins per conditioning variable. As expected, the pdfs do not show the same (near) perfect overlap with the reference compared to the training case of Figure 4, but the match is still accurate. Surrogates conditioned on e.g. $\left(Z^{\mathcal{R}}, E^{\mathcal{R}}, U^{\mathcal{R}}\right)$ or $\left(Z^{\mathcal{R}}, U^{\mathcal{R}}, S^{\mathcal{R}}\right)$ showed fairly similar results. Somewhat degraded performance (although overall still better than $\overline{r} = 0$), is obtained when conditioning on $\left(E^R, U^{\mathcal{R}}, S^{\mathcal{R}}\right)$, see Figure 9. While the $Z^{\mathcal{R}}$ bias is still corrected for, the pdfs of the surrogate underestimate the variance. The only exception, which did not improve upon the unparameterized model, was when conditioning on $\tau_E S'$ and $\tau_Z Z'$, despite the high correlations of Table 1. A possible cause is that, when predicting, we are forced to condition on $\tau_E(\widetilde{\Delta E})S'$ instead of $\tau_E\left(\Delta E\right)S'$, as the latter is not available outside the training period. Perhaps using conditioning variables such as $\tau_E S'$ and $\tau_Z Z'$ should be viewed as some form of overfitting, leading to a surrogate which is unlikely to generalize well beyond the training set. A possible remedy might be to increase the time lag [17].
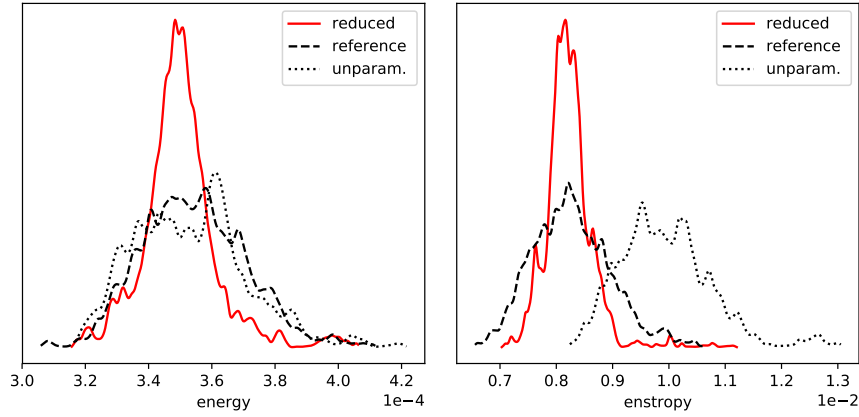
Figure 9: The pdfs of the energy (left) and enstrophy (right), of the reduced surrogate ($\widetilde{r} = \tau_E\left(\widetilde{\Delta E}\right)\Psi' + \tau_Z\left(\widetilde{\Delta Z}\right)\omega'$), reference ($\overline{r}$ given by (7)) and unparameterised ($\overline{r} = 0$) solution. The surrogates were both conditioned on $E^{\mathcal{R}}, U^{\mathcal{R}}, S^{\mathcal{R}}$ of the previous time step.
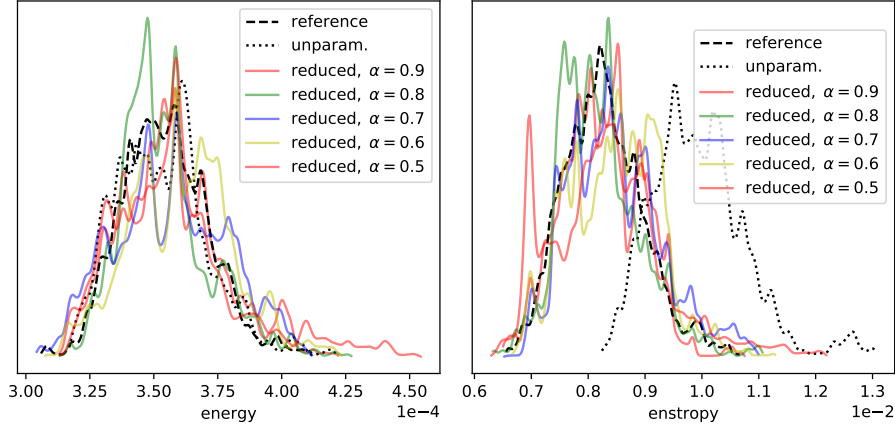

Figure 10: The pdfs of the energy (left) and enstrophy (right), of several extrapolated reduced surrogates ($\widetilde{r} = \tau_E\left(\widetilde{\Delta E}\right)\Psi' + \tau_Z\left(\widetilde{\Delta Z}\right)\omega'$), reference ($\overline{r}$ given by (7)) and unparameterised ($\overline{r} = 0$) solution. The surrogates were both conditioned on $Z^{\mathcal{R}}, E^{\mathcal{R}}, U^{\mathcal{R}}, S^{\mathcal{R}}$ of the previous time step.

4.3 Results T3

Predictive capability outside the training set should be the goal of any data-informed numerical simulation tool. In our case, this goal concerns prediction outside the time interval covered by the training set. We take tentative steps in this direction by incrementally reducing the time interval of the training set for the $\Delta E_{i+1} \mid Z_i^{\mathcal{R}}, E_i^{\mathcal{R}}, U_i^{\mathcal{R}}, S_i^{\mathcal{R}}$ and $\Delta Z_{i+1} \mid Z_i^{\mathcal{R}}, E_i^{\mathcal{R}}, U_i^{\mathcal{R}}, S_i^{\mathcal{R}}$ surrogates, while keeping the simulation time T_{sim} fixed to 8 years. Figure 10 shows the resulting pdfs, obtained using a training set spanning the first $T_{train} = \alpha T_{sim}$ years, with $\alpha \in \{0.9, 0.8, 0.7, 0.6, 0.5\}$. No significant deviation from the unextrapolated T2 test case is observed, which demonstrates the predictive capability of the surrogate method.

Finally, we note that all results can replicated via the source code and corresponding input files, available for download at [3].

5 CONCLUSION & OUTLOOK

We presented a method to create a stochastic surrogate model, conditioned on time-lagged observable variables, from a set of training data of a multiscale dynamical system. The novelty of our approach is found in the derivation of model-error source terms designed to track chosen spatially-integrated statistics of interest. We denote these as 'reduced' model error terms, as they lead to a significant reduction in the amount of training data. Although using less data might seem counter productive, we argue that this leads to an easier surrogate construction. Furthermore, our reduced framework allows us to step away from a fully-data driven, physics-blind, surrogate, and inform part of our model-error term based on the transport equations of the target statistics.

Future work includes further testing the extrapolative capability of the method. Another interesting research option would be to contrast the performance of our conditional time-lagged surrogate with machine-learning alternatives, such as random forests or neural nets. Recent relevant work also considered a combination of both approaches[13]. Finally, a further interesting avenue of future research is the *a-priori* incorporation of constraints from mathematical physics. For instance, when rewriting the eddy forcing in tensor format, certain constraints on the tensor shape can be found [19]. Such an approach opens up the possibility for efficient, physics-constrained uncertainty quantification, see e.g. [5] for examples in steady flow problems or [8] for large-eddy simulations.

REFERENCES

[1] P.S. Berloff. Random-forcing model of the mesoscale oceanic eddies. *Journal of Fluid Mechanics*, 529:71–95, 2005.

[2] D. Crommelin and E. Vanden-Eijnden. Subgrid-scale parameterization with conditional markov chains. *Journal of the Atmospheric Sciences*, 65(8):2661–2675, 2008.

[3] W.N. Edeling. Tau ez - uncecomp branch (github repository). `https://github.com/wedeling/TAU_EZ/tree/uncecomp`, 2019.

[4] W.N. Edeling. vorticity-solver (github repository). `https://github.com/wedeling/vorticity-solver`, 2019.

[5] W.N. Edeling, G. Iaccarino, and P. Cinnella. Data-free and data-driven rans predictions with quantified uncertainty. *Flow, Turbulence and Combustion*, 100(3):593–616, 2018.

[6] P.R. Gent and J.C. Mcwilliams. Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, 20(1):150–155, 1990.

[7] I. Grooms and L. Zanna. A note on 'toward a stochastic parameterization of ocean mesoscale eddies'. *Ocean Modelling*, 113:30–33, 2017.

[8] L. Jofre, S.P. Domino, and G. Iaccarino. A framework for characterizing structural uncertainty in large-eddy simulation closures. *Flow, Turbulence and Combustion*, 100(2):341–363, 2018.

[9] P. Mana and L. Zanna. Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Modelling*, 79:1–20, 2014.

[10] R. Maulik, O. San, A. Rasheed, and P. Vedula. Subgrid modelling for two-dimensional turbulence using neural networks. *Journal of Fluid Mechanics*, 858:122–144, 2019.

[11] J.C. McWilliams. The emergence of isolated coherent vortices in turbulent flow. *Journal of Fluid Mechanics*, 146:21–43, 1984.

[12] T. Palmer and P. Williams. *Stochastic physics and climate modelling*. Cambridge University Press Cambridge, UK, 2010.

[13] S. Pan and K. Duraisamy. Data-driven discovery of closure models. *SIAM Journal on Applied Dynamical Systems*, 17(4):2381–2413, 2018.

[14] R. Peyret. *Spectral methods for incompressible viscous flow*, volume 148. Springer Science & Business Media, 2013.

[15] J. Thuburn, J. Kent, and N. Wood. Cascades, backscatter and conservation in numerical models of two-dimensional turbulence. *Quarterly Journal of the Royal Meteorological Society*, 679(140):626–638, 2014.

[16] N. Verheul and D. Crommelin. Data-driven stochastic representations of unresolved features in multiscale models. *Commun. Math. Sci*, 14(5):1213–1236, 2016.

[17] N. Verheul, J. Viebahn, and D. Crommelin. Covariate-based stochastic parameterization of baroclinic ocean eddies. *Mathematics of Climate and Weather Forecasting*, 3(1):90–117, 2017.

[18] W.T.M. Verkley, P.C. Kalverla, and C.A. Severijns. A maximum entropy approach to the parametrization of subgrid processes in two-dimensional flow. *Quarterly Journal of the Royal Meteorological Society*, 142(699):2273–2283, 2016.

[19] S. Waterman and J.M. Lilly. Geometric decomposition of eddy feedbacks in barotropic systems. *Journal of Physical Oceanography*, 45(4):1009–1024, 2015.

[20] L. Zanna, P. Mana, J. Anstey, T. David, and T. Bolton. Scale-aware deterministic and stochastic parametrizations of eddy-mean flow interaction. *Ocean Modelling*, 111:66–80, 2017.

A ENERGY AND ENSTROPHY EQUATIONS

For convenience, we reproduce certain relevant derivations regarding the $E^{\mathcal{R}}$ and $Z^{\mathcal{R}}$ transport equations from [18]. The energy (density) is defined as

$$E^{\mathcal{R}} := \frac{1}{2}\left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi}\int_0^{2\pi} \mathbf{V}^{\mathcal{R}}\cdot\mathbf{V}^{\mathcal{R}}\mathrm{d}x\mathrm{d}y, \tag{26}$$

where $\mathbf{V}^{\mathcal{R}}$ is the vector containing the velocity components in x and y direction. It can be rewritten as $E^{\mathcal{R}} = -\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right)/2$ via

$$\mathbf{V}^{\mathcal{R}} \cdot \mathbf{V}^{\mathcal{R}} = \nabla\psi^{\mathcal{R}} \cdot \nabla\psi^{\mathcal{R}} = \nabla \cdot \left(\psi^{\mathcal{R}}\nabla\psi^{\mathcal{R}}\right) - \psi^{\mathcal{R}}\nabla^2\psi^{\mathcal{R}} = \nabla \cdot \left(\psi^{\mathcal{R}}\nabla\psi^{\mathcal{R}}\right) - \psi^{\mathcal{R}}\omega^{\mathcal{R}} \quad (27)$$

The first equality follows from the definition $\mathbf{V}^{\mathcal{R}} := \left(-\partial\psi^{\mathcal{R}}/\partial y, \partial\psi^{\mathcal{R}}/\partial x\right)^T$, while the second stems from the product rule of a scalar ($\psi^{\mathcal{R}}$) and a vector ($\nabla\psi^{\mathcal{R}}$):

$$\nabla \cdot \left(\psi^{\mathcal{R}}\nabla\psi^{\mathcal{R}}\right) = \nabla\psi^{\mathcal{R}} \cdot \nabla\psi^{\mathcal{R}} + \psi^{\mathcal{R}}\nabla^2\psi^{\mathcal{R}}. \quad (28)$$

Finally, the last equality of (27) simply follows from the governing equations (1). The term $\nabla \cdot \left(\psi^{\mathcal{R}}\nabla\psi^{\mathcal{R}}\right)$ disappears when integrated over the spatial domain, after application of the divergence theorem in combination with the doubly periodic boundary conditions. This leaves $E^{\mathcal{R}} = -\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right)/2$. To obtain the energy equation, start with

$$\frac{\mathrm{d}E^{\mathcal{R}}}{\mathrm{d}t} = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi}\int_0^{2\pi} \frac{\partial}{\partial t}\left[\frac{1}{2}\mathbf{V}^{\mathcal{R}} \cdot \mathbf{V}^{\mathcal{R}}\right]\mathrm{d}x\mathrm{d}y = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi}\int_0^{2\pi} \mathbf{V}^{\mathcal{R}} \cdot \frac{\partial\mathbf{V}^{\mathcal{R}}}{\partial t}\mathrm{d}x\mathrm{d}y. \quad (29)$$

Similar to the analysis above, we use the relation $\mathbf{V}^{\mathcal{R}} \cdot \mathbf{V}_t^{\mathcal{R}} = \nabla \cdot \left(\psi^{\mathcal{R}}\nabla\psi_t^{\mathcal{R}}\right) - \psi^{\mathcal{R}}\omega_t^{\mathcal{R}}$ (where the subscript t denotes $\partial/\partial t$) to obtain

$$\frac{\mathrm{d}E^{\mathcal{R}}}{\mathrm{d}t} = -\left(\Psi^R, \frac{\partial\omega^{\mathcal{R}}}{\partial t}\right) =$$
$$\left(\psi^{\mathcal{R}}, P^{\mathcal{R}}J\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right)\right) - \nu\left(\psi^{\mathcal{R}}, \nabla^2\omega^{\mathcal{R}}\right) - \mu\left(\psi^{\mathcal{R}}, F - \omega^{\mathcal{R}}\right) + \left(\psi^{\mathcal{R}}, \bar{r}\right) \quad (30)$$

Using integration by parts and the periodic boundary conditions it can be shown that the first term on the right-hand side satisfies $\left(\psi^{\mathcal{R}}, P^{\mathcal{R}}J\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right)\right) = \left(J\left(\psi^{\mathcal{R}}, \psi^{\mathcal{R}}\right), \omega^{\mathcal{R}}\right) = 0$, since the Jacobian of two equal arguments is zero [18]. Furthermore, using the self-adjoint nature of the Laplace operator, we have $\left(\psi^{\mathcal{R}}, \nabla^2\omega^{\mathcal{R}}\right) = \left(\nabla^2\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right) = \left(\omega^{\mathcal{R}}, \omega^{\mathcal{R}}\right)$. This leads to

$$\frac{\mathrm{d}E^{\mathcal{R}}}{\mathrm{d}t} = -\nu\left(\omega^{\mathcal{R}}, \omega^{\mathcal{R}}\right) - \mu\left(\psi^{\mathcal{R}}, F\right) + \mu\left(\psi^{\mathcal{R}}, \omega^{\mathcal{R}}\right) + \left(\psi^{\mathcal{R}}, \bar{r}\right), \quad (31)$$

which equals (13). Using the same procedure, the evolution equation for the enstrophy reads

$$\frac{\mathrm{d}Z^{\mathcal{R}}}{\mathrm{d}t} = \left(\omega^{\mathcal{R}}, \frac{\partial\omega^{\mathcal{R}}}{\partial t}\right) = \nu\left(\omega^{\mathcal{R}}, \nabla^2\omega^{\mathcal{R}}\right) + \mu\left(\omega^{\mathcal{R}}, F\right) - \mu\left(\omega^{\mathcal{R}}, \omega^{\mathcal{R}}\right) - \left(\omega^{\mathcal{R}}, \bar{r}\right). \quad (32)$$

17