Available online at www.sciencedirect.com

ScienceDirect

Ocean Modelling 00 (2021) 1–13

Ocean Modelling

Short communication

A method for preserving nominally-resolved flow patterns in low-resolution ocean simulations: Dynamical system reconstruction

I. Shevchenko[a,*], P. Berloff[a,b]

[a] *Department of Mathematics, Imperial College London, Huxley Building, 180 Queen's Gate, London, SW7 2AZ, UK*

[b] *Institute of Numerical Mathematics of the Russian Academy of Sciences, Moscow, Russia*

Abstract

Accurate representation of large-scale flow patterns in low-resolution ocean simulations is one of the most challenging problems in ocean modelling. The main difficulty is to correctly reproduce effects of unresolved small scales on the resolved large scales. For this purpose, most of current research is focused on development of parameterisations directly accounting for the small scales. In this work we propose an alternative to the mainstream ideas by showing how to reconstruct a dynamical system from the available reference solution data (our proxy for observations) and, then, how to use this system for modelling not only large-scale but also nominally-resolved flow patterns at low resolutions. This approach is advocated as a part of the novel framework for data-driven hyper-parameterisation of mesoscale oceanic eddies in non-eddy-resolving models. The main characteristic of this framework is that it does not require to know the physics behind large-small scale interactions to reproduce both large and small scales in low-resolution ocean simulations. We tested it in the context of a three-layer, statistically equilibrated, steadily forced quasigeostrophic model for the beta-plane configuration and showed that non-eddy-resolving solution can be substantially improved towards the reference eddy-resolving benchmark. The proposed methodology robustly allows to retrieve a system of equations governing reduced dynamics of the observed data, while the additional adaptive nudging counteracts numerical instabilities by keeping solutions in the region of phase space occupied by the reference fields. Remarkably, its solutions simulate not only large-scale but also small-scale flow features, which can be nominally resolved by the low-resolution grid. In addition, the proposed method reproduces realistic vortex trajectories. One of the important and general conclusions that can be drawn from our results is that not only mesoscale eddy parameterization is possible in principle but also it can be highly accurate (up to reproducing individual vortices) for significantly reduced dynamics (down to 30 degrees of freedom). This conclusion provides great optimism for the ongoing parameterization studies, which are still far away from being completed.

Keywords: Ocean general circulation and dynamics, Multi-layer quasi-geostrophic model, Mesoscale eddies and parameterizations, Dynamical system reconstruction, Adaptive nudging

*Corresponding author at:
Email address: i.shevchenko@imperial.ac.uk (I. Shevchenko)

1. Introduction

It is typical of low-resolution ocean simulations to have significantly distorted or even absent large-scale flow structures that are otherwise present in the high-resolution simulations. This failure is due to missing effects of the small scales, which are not adequately resolved in low-resolution simulations. To mitigate the problem, many parameterisations for both comprehensive and idealized ocean models have been proposed (e.g., Gent and Mcwilliams (1990); Duan and Nadiga (2007); Frederiksen et al. (2012); Jansen and Held (2014); Mana and Zanna (2014); Cooper and Zanna (2015); Grooms et al. (2015); Berloff (2015, 2016, 2018); Danilov et al. (2019); Ryzhov et al. (2019); Juricke et al. (2020a,b); Cotter et al. (2019); Ryzhov et al. (2020); Cotter et al. (2020a,b,c)), but overall the problem remains largely unresolved for several reasons. First, defining the small and large scales is ambiguous, because they are not separated by a clear spectral gap or otherwise. Second, definition of the small and large scales should be consistent with the specific resolving capabilities of a low-resolution model in which their interactions are to be parameterized. Third, dynamical interactions across the scales are remarkably complex, as well as spatially inhomogeneous and non-stationary.

In this work we have further developed the hyper-parameterisation approach to reproduce the effects of mesoscale oceanic eddies on the large-scale ocean circulation. The main characteristic of this approach is that it does not require to know the physics behind large-small scale interactions to reproduce them. Complimentary to the mainstream physics-based perspective, we propose to deal with the eddy effects from the dynamical systems point of view, and interpret the lack of them as the persistent tendency of phase space trajectories representing the low-resolution solution to escape the right region of the corresponding phase space, which is occupied by the reference eddy-resolving solution. Therefore, we approach the problem from a different direction: instead of parameterizing small-scale effects, we retrieve an underlying dynamical system and use it to model evolution of the nominally-resolved flow patterns at low resolutions. The nominally-resolved flow patterns are those that are resolved by the grid. In our case, we need at least 7 grid points per the flow pattern to ensure it is resolved, since we use the CABARET scheme (Karabasov et al., 2009), while conventional methods would require at least 10 points. For example, a vortex of 70km in diameter requires the grid step to be at least 10km. Although the basic idea has long research history, our application of it is novel, and the proposed methodology has many novel features. Let us first discuss below the relevant background.

Retrieving reduced equations underlying the observed flow evolution is one of the most challenging problems in dynamical systems (e.g.,(Aguirre and Letellier, 2009; Brunton et al., 2016)). Although, this field has been researched for decades, most of the efforts used low-dimensional dynamical systems with 3-5 degrees of freedom (e.g., Brunton et al. (2017); Mangiarotti and Huc (2019)), and even this turned out difficult. The other problem is about frequent numerical instabilities of the retrieved dynamical systems (e.g., Sceller et al. (1999)). This implies that applying known methodologies for thousands of degrees of freedom, typical for describing low-resolution oceanic flows, is unfeasible.

For developing and testing the approach, we considered an intermediate-complexity, quasigeostrophic, eddy-resolving model of the wind-driven midlatitude ocean circulation — this is a respected and widely used (e.g., Siegel et al. (2001); Karabasov et al. (2009); Shevchenko and Berloff (2016); Shevchenko et al. (2016) and references there in) paradigm for process studies involving large-/small-scale turbulent interactions and their parameterizations. To mitigate the model size problem, we applied the Empirical Orthogonal Function (EOF) analysis (Preisendorfer, 1988; Hannachi et al., 2007) to the reference flow, defined here as the high-resolution solution subsampled on a coarse grid, and reconstructed a dynamical system for the evolving Principal Components (PCs) corresponding to the leading EOFs. Successful examples of reduced-order modelling with EOF-PC description can be found in (Kondrashov and Berloff, 2015; Kondrashov et al., 2018). Other types of space reduction are possible and can improve the outcome even further, but they are not considered in this study. To resolve the problem with numerical instabilities, we used adaptive nudging

methodology, which is an upgraded extension of the nudging method proposed in (Shevchenko and Berloff, 2021).

2. The method

The main objective of this study is to reconstruct a dynamical system from the reference solution (say $\mathbf{x}(t)$, $\mathbf{x} \in \mathbb{R}^n$); the reference solution is the high-resolution solution projected onto the coarser grid. Note that $\mathbf{x}(t)$ cannot be evolved in terms of $\mathbf{x}(t)$ only (without knowledge of the high-resolution solution), i.e. the evolution equation for $\mathbf{x}(t)$ is not closed. This is why we reconstruct a dynamical system for the evolution of $\mathbf{x}(t)$.

This dynamical system is meant to correctly simulate the reference large-scale flow patterns on a low-resolution grid. The full dimensionality of the problem is the total number of the grid nodes, and for the reconstructed dynamical system we aim to reduce it by orders of magnitude via EOFs/PCs decomposition of the reference data. Next, we formulate general dynamical system in terms of the leading PCs:

$$\mathbf{y}'(t) = \mathbf{F}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}^m, \quad t \in [0, \widetilde{T}], \quad m << n\,, \tag{1}$$

where the PCs are combined in the vector and denoted by $\mathbf{y}(t)$, and the dash means time differentiation. In our case the dimensionality has been eventually reduced by three orders of magnitude (from $n = 16441$ to $m = 30$).

In (1) we used 30 leading PCs that captured 98% of the reference flow variance. The right hand side of (1) is approximated with polynomial of order two in all the variables, $\mathbf{P}(\mathbf{y})$, and with the Fourier series, $\mathcal{F}(\mathbf{y})$, containing 50 leading harmonics:

$$\mathbf{F}(\mathbf{y}) \approx \mathbf{P}(\mathbf{y}) + \mathcal{F}(\mathbf{y})\,, \tag{2}$$

where

$$\mathbf{P}(\mathbf{y}) := a_0 + \sum_{i=1}^{30} a_i y_i + b_i y_i^2 + c_i y_i y_j, \quad j = 1, \ldots, m, \quad i \neq j, \tag{3}$$

and

$$\mathcal{F}(\mathbf{y}) := \sum_{k=1}^{25} d_k \cos\left(\frac{2\pi k t}{\widetilde{T}}\right) + e_k \sin\left(\frac{2\pi k t}{\widetilde{T}}\right), \tag{4}$$

with unknown coefficients $\mathbf{c} = \{a_0, a_i, b_i, c_i, d_k, e_k\}$, $i = 1, \ldots, 30$, $k = 1, \ldots, 25$ to be defined with the least squares method from the system of equations:

$$\mathbf{A}\mathbf{c} = \mathbf{y}'. \tag{5}$$

The derivative on the right hand side of (5) is approximated with the forward finite difference, and it is computed over the time interval $[0, \widetilde{T}]$ for which the leading PCs (computed from the reference solution) are available.

Note that $\mathbf{F}(\mathbf{y})$ can be approximated differently, and its optimal choice (beyond the scope of this work) is a challenge for the dynamical system reconstruction. Without proper information for tailoring the right hand side more specifically, a polynomial expansion is justified by the Weierstrass approximation theorem, while the use of the Fourier series allows one to approximate the mean flow more accurately. We will get back to this choice when discussing the results. We note that the choice of the polynomial basis is justified by the Weierstrass theorem as one can approximate the right hand side with a desired order of accuracy using polynomial functions. It does not mean though that the Weierstrass theorem justifies the use of quadratic polynomials for the dynamical system reconstruction as the latter is not guaranteed to be accurate even if the right hand side is accurately approximated.

Having approximated $\mathbf{F}(\mathbf{y})$ up to a given order of accuracy, one can solve the reconstructed dynamical system

$$\mathbf{z}'(t) = \mathbf{P}(\mathbf{z}) + \mathcal{F}(\mathbf{z}), \quad \mathbf{z} \in \mathbb{R}^m, \quad t \in [0, T], \quad T > \widetilde{T}\,. \tag{6}$$

Note that this system is integrated over a time interval which is longer (here, 2 times longer) than that of the original system (1). In all further simulations we will have $\widetilde{T}$ =2 years and T =4 years. However, an accurate approximation of $\mathbf{F}(\mathbf{y})$ does not guarantee that system (6) can be easily solved, because the integration errors can quickly contaminate the solution and result in severe numerical instability — this is what actually happened in our case. In order to stabilize the numerical integration, we used the nudging methodology (Shevchenko and Berloff, 2021):

$$\mathbf{z}'(t) = \mathbf{P}(\mathbf{z}) + \mathcal{F}(\mathbf{z}) + \eta\left(\frac{1}{N}\sum_{k\in\mathcal{U}(\mathbf{z}(t))}\mathbf{y}(t_k) - \mathbf{z}(t)\right), \quad t \in [0, T]\,, \tag{7}$$

where $\mathcal{U}(\mathbf{z}(t))$ is a neighbourhood of $\mathbf{z}(t)$, and index k is the timestep of the corresponding PC $\mathbf{y}(t_k)$; the timestep of the PC is the timestep with which the actual data for the EOF analysis was generated. The notation $k \in \mathcal{U}(\mathbf{z}(t))$ means that k's are taken for those $\mathbf{y}(t_k)$ that are in the neighbourhood of $\mathbf{z}(t)$. The neighbourhood is computed in l_2 norm as the average of $N = 5$ points nearest to the solution $\mathbf{z}(t)$. Note that the number of neighbourhood points is a parameter, and its sensitivity should be explored and taken into account for each application of the proposed methodology. We have defined its value from a series of experiments with values 2, 5, 10, 20; $N = 5$ gives the most accurate solution.

Having solved equation (7), we approximated the reference solution by using the leading EOF-PC pairs as follows:

$$\mathbf{x}(t) \approx \sum_{i=1}^{m} z_i(t)\mathbf{E}_i\,, \tag{8}$$

with $\mathbf{E}_i$ and z_i being the i-th EOF and PC, respectively. The choice of m results from the variability of the reference solution to be reproduced.

Note that N in equation (7) can be made time-dependent and adaptive, like the nudging coefficient η, which is the other important parameter. In order to make the numerical integration stable with the Euler method, we used the following adaptive nudging:

$$\eta(t_i) = \begin{cases} \eta(t_{i-1}) + \eta_h & \text{if } \sigma(\mathbf{z}(t_i)) > \max_{t\in[0,\widetilde{T}]}\sigma(\mathbf{y}(t)), \\ \eta(t_{i-1}) - \eta_h & \text{if } \sigma(\mathbf{z}(t_i)) \le \max_{t\in[0,\widetilde{T}]}\sigma(\mathbf{y}(t)), \quad i = 1, 2, \ldots \\ 0 & \text{if } \eta(t_{i-1}) - \eta_h < 0. \end{cases} \tag{9}$$

with σ being the standard deviation, $\eta_h = 0.001$, and $\eta(t_0) = 0$.

We opted for an adaptive nudging, as it keeps the system within a neighbourhood of the phase space region occupied by the reference solution. As an alternative, a constant η can be also used with some tuning and caution, keeping in mind that its small value may not be enough for keeping the solution within the right region and its large value may result in an over-stabilized solution with suppressed flow variability (slow flow dynamics).

3. Multilayer quasi-geostrophic model

We consider a 3-layer quasi-geostrophic (QG) model with forcing and dissipation for the evolution of the potential vorticity (PV) anomaly $\mathbf{q} = (q_1, q_2, q_3)$ in domain Ω (Pedlosky, 1987):

$$\partial_t q_j + \mathrm{J}(\psi_j, q_j + \beta y) = \delta_{1j} F_\mathrm{w} - \delta_{j3}\,\mu\nabla^2\psi_j + \nu\nabla^4\psi_j, \quad j = 1, 2, 3\,, \tag{10}$$

where $\mathrm{J}(f,g) = f_x g_y - f_y g_x$, δ_{ij} is the Kronecker symbol, and $\boldsymbol{\psi} = (\psi_1, \psi_2, \psi_3)$ is the velocity streamfunction in three layers. The planetary vorticity gradient is $\beta = 2 \times 10^{-11}\,\mathrm{m}^{-1}\,\mathrm{s}^{-1}$, the bottom friction parameter is $\mu = 4 \times 10^{-8}\,\mathrm{s}^{-1}$, and the lateral eddy viscosity is $\nu = 50\,\mathrm{m}^2\,\mathrm{s}^{-1}$. The asymmetric wind curl forcing, driving the double-gyre ocean circulation, is given by

$$F_{\mathrm{w}} = \begin{cases} -1.80\,\pi\,\tau_0 \sin\left(\pi y / y_0\right), & y \in [0, y_0), \\ 2.22\,\pi\,\tau_0 \sin\left(\pi (y - y_0)/(L - y_0)\right), & y \in [y_0, L], \end{cases}$$

with the wind stress amplitude $\tau_0 = 0.03\,\mathrm{N\,m}^{-2}$ and the tilted zero forcing line $y_0 = 0.4L + 0.2x$, $x \in [0, L]$. The computational domain $\Omega = [0, L] \times [0, L] \times [0, H]$ is a closed, flat-bottom basin with $L = 3840\,\mathrm{km}$, and the total depth $H = H_1 + H_2 + H_3$ given by the isopycnal fluid layers of depths (top to bottom): $H_1 = 0.25\,\mathrm{km}$, $H_2 = 0.75\,\mathrm{km}$, $H_3 = 3.0\,\mathrm{km}$.

The PV anomaly $\boldsymbol{q}$ and the velocity streamfunction $\boldsymbol{\psi}$ are coupled through the system of elliptic equations:

$$\boldsymbol{q} = \nabla^2 \boldsymbol{\psi} - \mathbf{S}\boldsymbol{\psi}\,, \tag{11}$$

with the stratification matrix

$$\mathbf{S} = \begin{pmatrix} 1.19 \cdot 10^{-3} & -1.19 \cdot 10^{-3} & 0.0 \\ -3.95 \cdot 10^{-4} & 1.14 \cdot 10^{-3} & -7.47 \cdot 10^{-4} \\ 0.0 & -1.87 \cdot 10^{-4} & 1.87 \cdot 10^{-4} \end{pmatrix}.$$

The stratification parameters are given in units of km^{-2} and chosen so, that the first and second Rossby deformation radii are $Rd_1 = 40\,\mathrm{km}$ and $Rd_2 = 23\,\mathrm{km}$, respectively; the choice of these parameters is typical for the North Atlantic, as it allows to simulate a more realistic than in different QG setups but yet idealized eastward jet extension of the western boundary currents

System (10)-(11) is augmented with the integral mass conservation constraint (McWilliams, 1977):

$$\partial_t \iint\limits_{\Omega} \left(\psi_j - \psi_{j+1}\right)\, dydx = 0, \quad j = 1, 2 \tag{12}$$

with the zero initial condition, and with the partial-slip lateral boundary condition (Haidvogel et al., 1992):

$$\left(\partial_{\mathbf{nn}}\boldsymbol{\psi} - \alpha^{-1}\partial_{\mathbf{n}}\boldsymbol{\psi}\right)\Big|_{\partial\Omega} = 0\,, \tag{13}$$

where $\alpha = 120\,\mathrm{km}$ is the partial-slip parameter, and $\mathbf{n}$ is the normal-to-wall unit vector; no-flow-through boundary condition is also implemented (as part of the elliptic solver). The value of the parameter α is chosen based on the study by Shevchenko and Berloff (2015), where it has been shown that smaller values of α inhibit the eastward jet extension penetration length and volume transport, while larger values have much less pronounced influence on the jet. As with other governing parameters used in this study, our choice of α is justified by a more realistic eastward jet. The QG system (10)-(13) is solved using the high-resolution CABARET method, which is based on a second-order, non-dissipative and low-dispersive, conservative advection scheme.

For this study we need both high- and low-resolution solutions. In order to compute them, we first spin up the model (10)-(13) for 100 years and then solve it for the other 4 years on 2 uniform horizontal grids: 513×513 (high resolution) and 129×129 (low resolution). Note that all the parameters in the QG model are held fixed for both high- and low-resolution setups. In order to obtain the reference solution (denoted as q_1), we project the high-resolution solution on the coarse grid 129×129 by using point-to-point projection, i.e. the high-resolution solution is subsampled at the low-resolution grid points, (Figure 1a). The low-resolution solution (denoted as $\widehat{q}_1$) is the solution of the QG model on grid 129×129 (Figure 1b). Our goal is to find a dynamical system that can model the leading PCs (which are then used to approximate the reference

solution given by (8)), so that the approximate solution (denoted as $\widetilde{q}_1$) simulates the reference large-scale flow patterns in qualitatively correct way. We would like to draw the reader's attention to the fact that in the general description of the method the reference solution is denoted as $\mathbf{x}(t)$, while in the context of the QG model $\mathbf{x}(t) \coloneqq q_1$.

For the purpose of this work, it is enough to consider only the first layer, as it consists of both large- and small-scale features (Figure 1a) which we aim to reproduce. Moreover, the upper layer is more difficult to model than the deep ones. As seen in Figure 1a, the solution is characterized by the well-pronounced eastward jet extension of the western boundary currents and surrounding small-scale coherent vortices. Both of these features are missed in the low-resolution solution (Figure 1b) due to the under-resolved eddy effects. In order to restore nominally-resolved flow patterns (the eastward jet and surrounding vortices), we first reconstruct a reduced dynamical system (for the leading PCs) which is based on the second-order polynomials and then the one based on the second-order polynomials and Fourier series. The solution corresponding to the former is presented in Figure 1c. Although the snapshots show that both the eastward jet and vortices are successfully reproduced, the time-mean flow significantly differs from the reference solution: the eastward jet separation point is shifted north and the jet itself manifests fluctuations unseen in the reference solution.

For a better approximation we combined the second-order polynomial basis with the Fourier series. The corresponding solution (8) computed from the leading EOF-PC pairs is significantly improved (Figure 1d), mostly due to the better approximation of the PCs (Figure 2). The low-resolution solutions q_1^* and $\widetilde{q}_1$ (Figures 1c,d) have excessive spatial variance in the 4-year averages compared to the reference solution q_1 (Figure 1a). This is improved by adding the Fourier basis (Figure 1d). Also, the low resolution solutions have excessive time variance (Figure 1, standard deviation panels). This is made worse by adding the Fourier basis (for example in the jet detachment region).

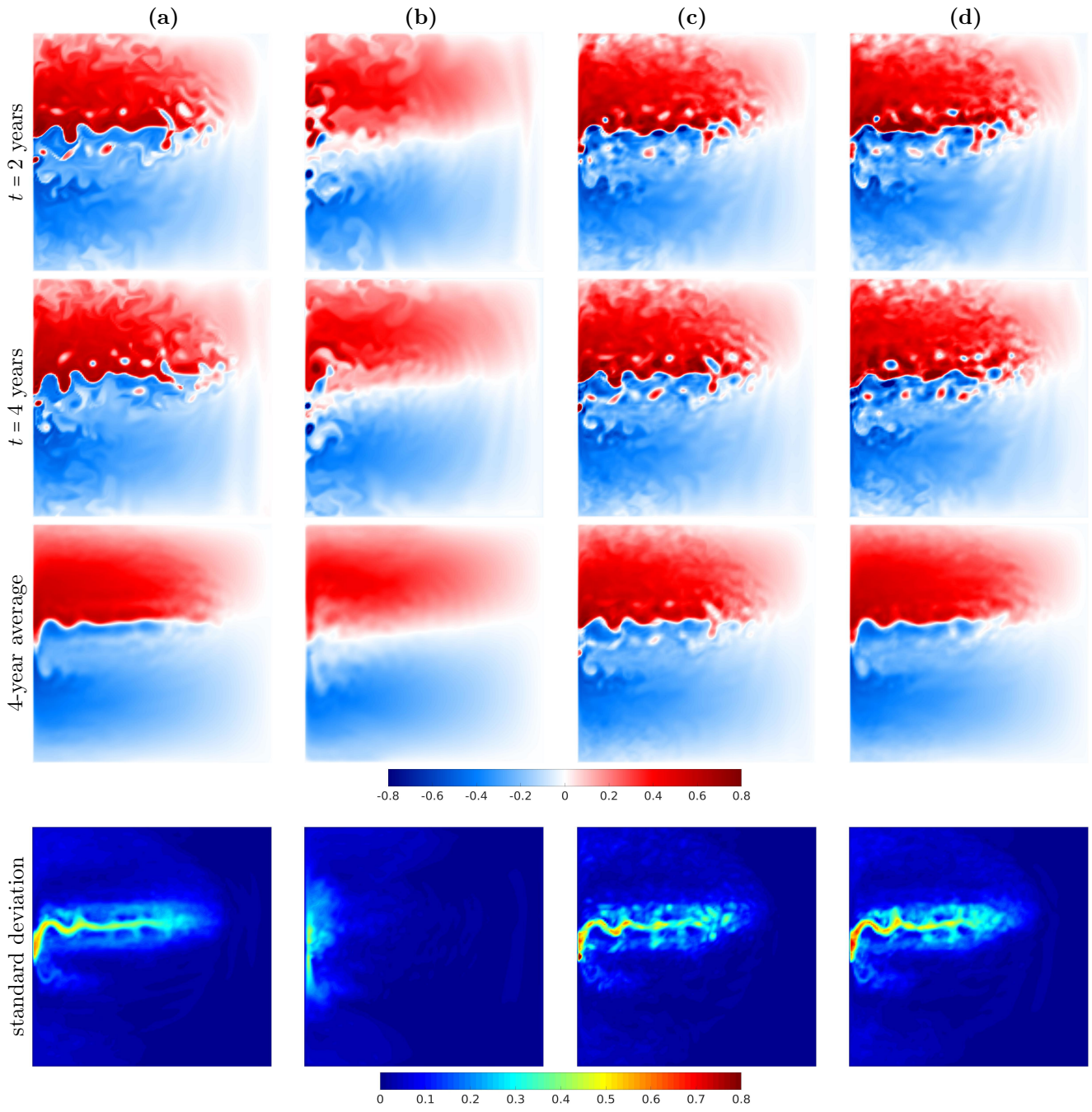

Figure 1: Shown is a series of snapshots, 4-year average, and standard deviation of the top layer PV anomaly of **(a)** the reference solution q_1 (computed on grid 513×513 and projected on grid 129×129), **(b)** low-resolution solution $\widehat{q}_1$ computed on grid 129×129, **(c)** low-resolution solution q_1^* on grid 129×129 (with the second-order polynomial basis used for the reconstruction), **(d)** low-resolution solution $\widetilde{q}_1$ on grid 129×129 (with the second-order polynomials and Fourier basis used for the reconstruction). The solution is given in units of $[s^{-1} f_0^{-1}]$, where $f_0 = 0.83 \times 10^{-4}\,\mathrm{s}^{-1}$ is the Coriolis parameter. The results in panels **(c)**, **(d)** demonstrate that the proposed method preserves not only large-, but also small-scale features (nominally resolved on the coarse-grid) like those seen in the reference solution **(a)** but absent in the low-resolution solution **(b)**.

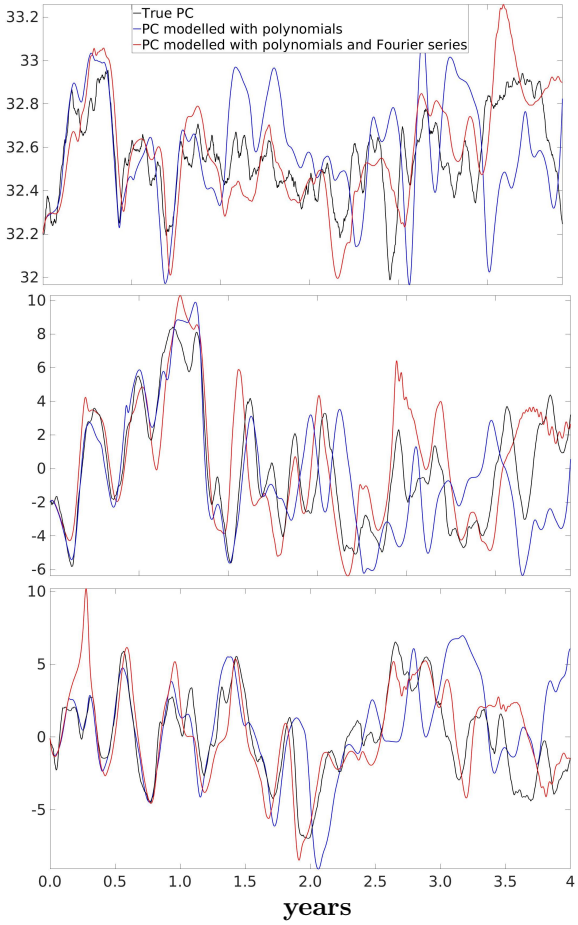

Figure 2: Shown are the first three leading PCs and their dependence on the basis functions used for the reconstruction of the dynamical system: true PC (black), PC modelled with the second-order polynomial-only basis (blue), and PC modelled with both the second-order polynomials and Fourier series (red). The results demonstrate that using the basis consisting of both the second-order polynomials and Fourier series yields significantly more accurate approximation of the PCs.

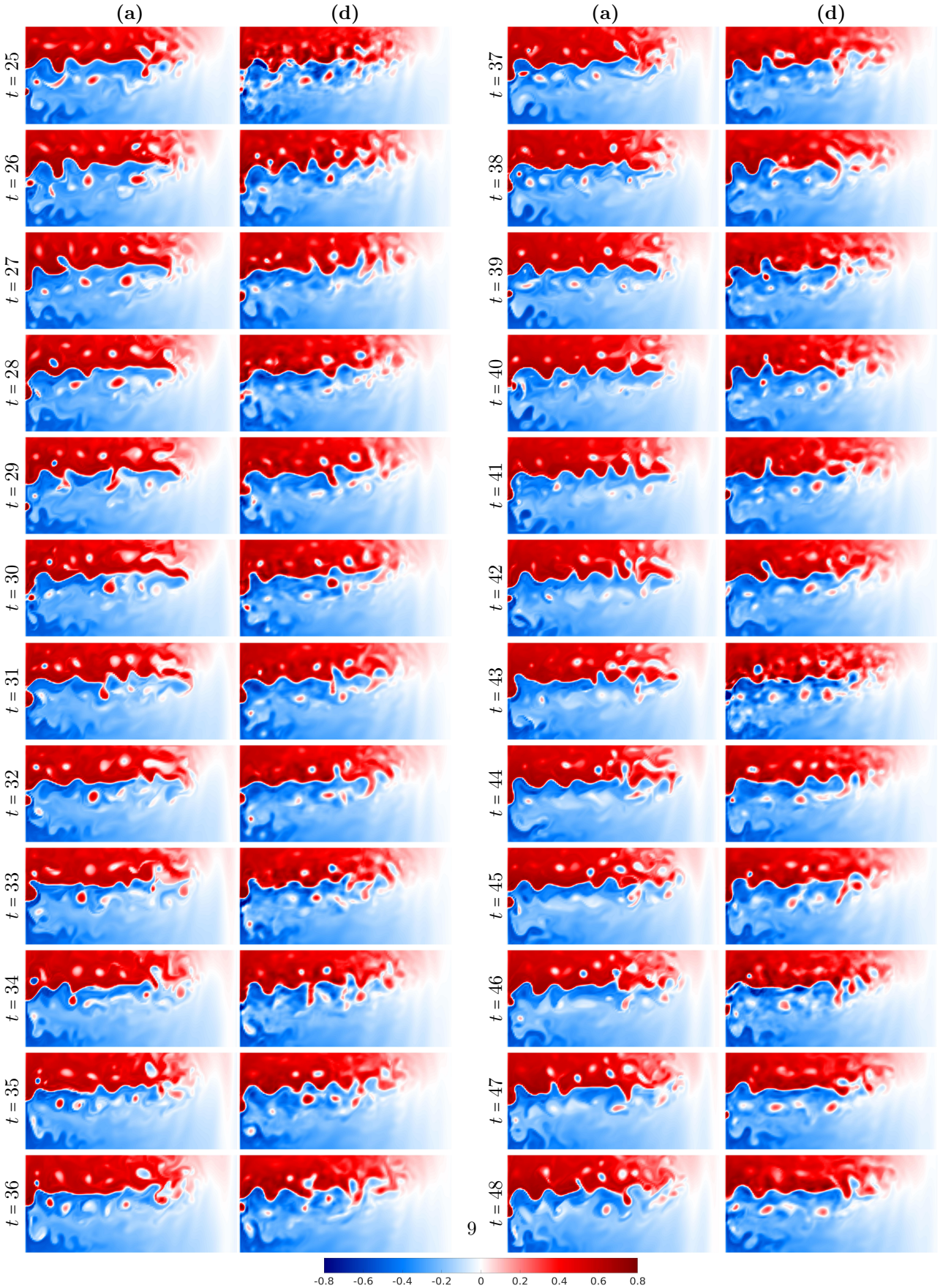

Figure 3: The same panels as in Figure 1 but with the time step 1 month. Shown is a neighbourhood of the eastward jet for the last two years of the 4-year long simulation. The results clearly demonstrate that not only the nominally-resolved flow structures (the eastward jet and coherent vortices) are present in the modelled solution but also realistic vortex trajectories are correctly reproduced.

Recall that the solution in Figure 1 is over 4 years, and only the first 2 years were used to reconstruct the dynamical system. This shows that the proposed method preserves not only the large-scale flow structure but also the small-scale flow features, all of them over a long time interval. It means that the proposed method maintains the large-scale flow structure (the eastward jet extension of the western boundary currents) and small-scale features (nominally resolved on the coarse grid eddies) akin to those of the reference solution. Figure 3 shows that not only the nominally-resolved flow structures are present in the modelled solution but also realistic vortex trajectories (vortices drift westward) are correctly reproduced. Moreover, vortices are formed at the tip of the jet and through meanders along the jet. However, it does not mean that the exact same individual eddies (vortices) in the sub-sampled high resolution solution are preserved. For example, the individual eddies in Figures 1a,d and Figures 3a,d are different. Besides, the dispersion characteristics of the reference and the reconstructed flow have not been checked and may be different. The ability of the method to reproduce small-scale features may look surprising, but since these features were present in the reference data, their reconstruction is a matter of the high-quality reconstruction of the dynamical system.

A key ingredient that makes the method work is the adaptive nudging which keeps the solution in the right region of the phase space that is occupied by the reference solution. As an approximation of the reference region, we used a sphere, $S(\mathbf{q}_1)$, centered at the time-mean of the solution, $\langle \mathbf{q}_1 \rangle$, and the sphere radius, r, is the mean distance of the solution from the centre:

$$S(\mathbf{q}_1) \coloneqq (\mathbf{q}_1 - \langle \mathbf{q}_1 \rangle)^2 - r^2 \le 0, \quad \langle \mathbf{q}_1 \rangle \coloneqq \frac{1}{T}\int_0^T \mathbf{q}_1(t)\,dt, \quad r \coloneqq \frac{1}{T}\int_0^T \|\mathbf{q}_1(t) - \langle \mathbf{q}_1 \rangle\|_2\,dt\,. \tag{14}$$

The mean distance between two points $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ in the phase space of the QG model is given by

$$\langle \mathcal{D}(\mathbf{x}_1, \mathbf{x}_2) \rangle \coloneqq \frac{1}{T}\int_0^T \|\mathbf{x}_1(t) - \mathbf{x}_2(t)\|_2\,dt\,. \tag{15}$$

The mean distance computed below is given in non-dimensional units.

The mean distances for the reference and low-resolution solutions are $\langle \mathcal{D}(q_1, \overline{q}_1) \rangle = 11.9$ and $\langle \mathcal{D}(\widehat{q}_1, \overline{\widehat{q}}_1) \rangle = 7.2$, respectively, showing that the latter is confined in a smaller region. The l_2-norm distance between the time means of these solutions (denoted as barred quantities) is $\mathcal{D}(\overline{q}_1, \overline{\widetilde{q}}_1) = 12.92$. The application of the adaptive nudging decreases the distance between the time means to $\mathcal{D}(\overline{q}_1, \overline{\widetilde{q}}_1) = 2.65$, thus shifting the whole solution $\widetilde{q}_1$ much closer to the phase space region occupied by the reference solution. It also yields a lot more accurate mean distance $\langle \mathcal{D}(\widetilde{q}_1, \overline{\widetilde{q}}_1) \rangle = 12.6$, thus suggesting that the solution has correct amplitude. On the other hand, $\langle \mathcal{D}(q_1, \overline{\widetilde{q}}_1) \rangle = 12.2$ (which is quite close to $\langle \mathcal{D}(\widetilde{q}_1, \overline{\widetilde{q}}_1) \rangle = 12.6$) thus reassuring once again that the reconstructed model gives an adequate approximation of the reference flow dynamics. Note that the perfect reconstruction over a period $[0, T]$ would mean that $\mathcal{D}(q_1(t), \widetilde{q}_1(t)) = 0$, $t \in [0, T]$.

4. Conclusions and discussion

In this study we proposed a method for preserving nominally-resolved flow patterns in low-resolution ocean model simulations. The method utilizes the well-known idea of reconstructing the dynamical system that underlies the observed flow evolution. However, direct application of this idea to the quasi-geostrophic model studied in this work is numerically unfeasible task because of the high dimensionality of the observed flow. Moreover, a numerical integration of the reconstructed dynamical system can be unstable, but our methodology can cope with this and ensure stability. We solved the problem of large dimensionality by applying the Empirical Orthogonal Function decomposition of the reference solution (the high-resolution solution subsampled on the coarse grid) that allowed to reduce the dimension by three orders of magnitude. In order to solve the unstable integration problem, we developed the adaptive nudging method following

(Shevchenko and Berloff, 2021). This method keeps the solution in the neighbourhood of the phase space region occupied by the reference solution. This is sufficient for accurate reproduction of both the large- and small-scale flow features at low resolutions, despite the fact that these features are not present in the dynamical solutions of the low-resolution model. The proposed method aims to operate with hundreds of degrees of freedom thus offering orders-of-magnitude acceleration compared to low-resolution ocean models which have at least 3-4 orders of magnitude more.

The proposed method can formally be classified as a reduced order modelling method (e.g., Brunton et al. (2016)), since it reduces the dimensionality of the original system. The principal difference of the proposed method (compared to the reduced order modelling technique) is that it does not use the original equation to substitute the reference solution projected onto the EOF basis functions. Instead, it reconstructs a dynamical system describing PCs directly and then finds the solution by using leading EOF-PC pairs. The dynamical system reconstruction used in this study is not a new idea, while using the adaptive nudging method is a novel approach in itself not to mention its application in the context of dynamical systems reconstruction. Moreover, a combination of the polynomial basis with the Fourier series is another new idea which can find its use beyond the scope of this work.

The proposed method was tested on a 3-layer quasi-geostrophic ocean circulation model at low non-eddy-resolving resolution, such that it cannot simulate the correct large-scale flow structure. Our results show that if the reconstructed dynamical model is based only on the second-order polynomials, then it is not sufficiently accurate (compared with the reference solution), because its time-mean eastward jet separation point is shifted north, and the jet itself has unrealistic fluctuations which are not observed in the reference solution. We tried to use higher order polynomials, but the reconstructed system became very sensitive to errors leading to sever numerical instabilities which we failed to stabilize. We resolved this problem by augmenting the polynomial basis with the additional Fourier series. With all this in place, not only the large-scale flow structure becomes correct but also the small-scale coherent vortices, which are unresolved in the low-resolution full-dynamics model, appear in the solution. All in all, this shows that the method has potential for modelling even more complicated oceanic flows. Being small-scales-unaware (not relying on reproducing the effect of small scales onto large ones like parameterisations), the proposed method can be thought of as an alternative to the modern (small-scales-aware) parameterisations, which try to reproduce effects of small dynamically unresolved scales on the large scales, in the hope that the solution will stay in the right region of the phase space. The proposed approach is quite the opposite: it forces the solution to stay in the right phase space region and predicts the flow evolution via the reconstructed reduced dynamical system. Note that the method does not require the original quasi-geostrophic model to be solved at low-resolution.

It is worth reiterating again that the proposed method maintains the large-scale flow structure (the eastward jet extension of the western boundary currents) and small-scale features (nominally resolved on the coarse grid eddies) akin to those of the reference solution. Moreover, not only the nominally-resolved flow structures are present in the modelled solution but also realistic vortex trajectories are correctly reproduced. However, it does not mean that the exact same individual eddies in the sub-sampled high resolution solution are preserved.

One of the important and general conclusions that can be drawn from our results is that not only mesoscale eddy parameterization is possible in principle but also it can be highly accurate (up to reproducing individual vortices) for significantly reduced dynamics (down to 30 degrees of freedom). This conclusion provides great optimism for the ongoing parameterization studies, which are still far away from being completed.

The reference data is used twice: first, for reconstructing the dynamical system; second, for augmenting the solution of this system by nudging (to compute the PCs, $\mathbf{y(t)}$, in the nudging term in equation (7)). The method can be further improved by using a more sophisticated nudging methodologies and different dynamical systems which can better represent the underlying flow dynamics. The proposed method can

be applied to primitive equations, but in this case reconstruction of the dynamical system will be more subtle, as it will include more PCs and can require changes of the basis functions. Moreover, EOF analysis might cease to work for time-dependent forcing and more sophisticated flow decompositions will therefore be needed (see, e.g., Xie et al. (2018); Shady et al. (2021)). It is also important to note that the proposed data-driven method requires the system to be statistically equilibrated, which is not the case in applications that undergo regime transitions (e.g., forced climate projections or systems with multiple (statistically) steady states). In other words, if the system has multiple attractors and only some of them are presented in data then the proposed method can only reproduce the flow dynamics which is presented in data.

Another future extensions of this study can be exploring the possibility of generating a forcing for the low-resolution ocean model, based on the EOFs and the adaptive nudging, perhaps with some ingredients such as stochastic forcing.

5. Acknowledgments

The authors thank The Leverhulme Trust for the support of this work through the grant RPG-2019-024. Pavel Berloff was supported by the NERC grants NE/R011567/1 and NE/T002220/1, and by the Moscow Center of Fundamental and Applied Mathematics (supported by the Agreement 075-15-2019-1624 with the Ministry of Education and Science of the Russian Federation).

References

Aguirre, L. and Letellier, C. (2009). Modeling nonlinear dynamics and chaos: A review. *Mathematical Problems in Engineering*, 2009:1–35.

Arakawa, A. (1966). Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow. Part I. *J. Comput. Phys.*, 1:119–143.

Berloff, P. (2015). Dynamically consistent parameterization of mesoscale eddies. Part I: simple model. *Ocean Model.*, 87:1–19.

Berloff, P. (2016). Dynamically consistent parameterization of mesoscale eddies. Part II: eddy fluxes and diffusivity from transient impulses. *Fluids*, 1:1–19.

Berloff, P. (2018). Dynamically consistent parameterization of mesoscale eddies. Part III: Deterministic approach. *Ocean Model.*, 127:1–15.

Brunton, S., Brunton, W., Proctor, J., Kaiser, E., and Kutz, N. (2017). Chaos as an intermittently forced linear system. *Nat. Commun.*, 8:1–9.

Brunton, S., Proctor, J., and Kutz, N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *PNAS*, 113:3932–3937.

Cooper, F. and Zanna, L. (2015). Optimization of an idealised ocean model, stochastic parameterisation of sub-grid eddies. *Ocean Model.*, 88:38–53.

Cotter, C., Crisan, D., Holm, D., Pan, W., and Shevchenko, I. (2019). Numerically modelling stochastic Lie transport in fluid dynamics. *Multiscale Model. Simul.*, 17:192–232.

Cotter, C., Crisan, D., Holm, D., Pan, W., and Shevchenko, I. (2020a). A Particle Filter for Stochastic Advection by Lie Transport (SALT): A case study for the damped and forced incompressible 2D Euler equation. *SIAM/ASA Journal on Uncertainty Quantification*, 8:1446–1492.

Cotter, C., Crisan, D., Holm, D., Pan, W., and Shevchenko, I. (2020b). Data assimilation for a quasi-geostrophic model with circulation-preserving stochastic transport noise. *Journal of Statistical Physics*, 179:1186–1221.

Cotter, C., Crisan, D., Holm, D., Pan, W., and Shevchenko, I. (2020c). Modelling uncertainty using stochastic transport noise in a 2-layer quasi-geostrophic model. *Foundations of Data Science*, 2:173–205.

Danilov, S., Juricke, S., Kutsenko, A., and Oliver, M. (2019). Toward consistent subgrid momentum closures in ocean models. In Eden, C. and A.Iske, editors, *Energy Transfers in Atmosphere and Ocean*, chapter 5, pages 145–192. Springer-Verlag.

Duan, J. and Nadiga, B. (2007). Stochastic parameterization for large eddy simulation of geophysical flows. *Proc. Am. Math. Soc.*, 135:1187–1196.

Frederiksen, J., OKane, T., and Zidikheri, M. (2012). Stochastic subgrid parameterizations for atmospheric and oceanic flows. *Phys. Scr.*, 85:068202.

Gent, P. and Mcwilliams, J. (1990). Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, 20:150–155.

Grooms, I., Majda, A., and Smith, K. (2015). Stochastic superparametrization in a quasigeostrophic model of the Antarctic Circumpolar Current. *Ocean Model.*, 85:1–15.

Haidvogel, D., McWilliams, J., and Gent, P. (1992). Boundary current separation in a quasigeostrophic, eddy-resolving ocean circulation model. *J. Phys. Oceanogr.*, 22:882 – 902.

Hannachi, A., Jolliffe, I., and Stephenson, D. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.*, 27:1119–1152.

Hundsdorfer, W., Koren, B., van Loon, M., and Verwer, J. (1995). A positive finite-difference advection scheme. *J. Comput. Phys.*, 117:35–46.

Jansen, M. and Held, I. (2014). Parameterizing subgrid-scale eddy effects using energetically consistent backscatter. *Ocean Model.*, 80:36–48.

Juricke, S., Danilov, S., Koldunov, N., Oliver, M., Sein, D., Sidorenko, D., and Wang, Q. (2020a). A kinematic kinetic energy backscatter parametrization: From implementation to global ocean simulations. *J. Adv. Model. Earth Syst.*, 12:2020MS002175.

Juricke, S., Danilov, S., Koldunov, N., Oliver, M., and Sidorenko, D. (2020b). Ocean kinetic energy backscatter parametrization on unstructured grids: Impact on global eddypermitting simulations. *J. Adv. Model. Earth Syst.*, 12:2019MS001855.

Karabasov, S., Berloff, P., and Goloviznin, V. (2009). CABARET in the ocean gyres. *Ocean Model.*, 2–3:155–168.

Kondrashov, D. and Berloff, P. (2015). Stochastic modeling of decadal variability in ocean gyres. *Geophys. Res. Lett.*, 42:1543–1553.

Kondrashov, D., Chekroun, M., and Berloff, P. (2018). Multiscale Stuart-Landau emulators: Application to wind-driven ocean gyres. *Fluids*, 3:1–32.

Mana, P. P. and Zanna, L. (2014). Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Model.*, 79:1–20.

Mangiarotti, S. and Huc, M. (2019). Can the original equations of a dynamical system be retrieved from observational time series? *Chaos*, 29:023133.

McWilliams, J. (1977). A note on a consistent quasigeostrophic model in a multiply connected domain. *Dynam. Atmos. Ocean*, 5:427–441.

Pedlosky, J. (1987). *Geophysical fluid dynamics*. Springer-Verlag, New York.

Preisendorfer, R. W. (1988). *Principal Component Analysis in Meteorology and Oceanography*. Elsevier, Amsterdam.

Ryzhov, E., Kondrashov, D., Agarwal, N., and Berloff, P. (2019). On data-driven augmentation of low-resolution ocean model dynamics. *Ocean Model.*, 142:101464.

Ryzhov, E., Kondrashov, D., Agarwal, N., McWilliams, J., and Berloff, P. (2020). On data-driven induction of the low-frequency variability in a coarse-resolution ocean model. *Ocean Model.*, 153:101664.

Sceller, L. L., Letellier, C., and G., G. (1999). Structure selection for global vector field reconstruction by using the identification of fixed points. *Phys. Rev. E*, 60:1600–1606.

Shady, A., Suraj, P., Omer, S., Adil, R., and Mandar, T. (2021). A nudged hybrid analysis and modeling approach for realtime wake-vortex transport and decay prediction. *Computers and Fluids*, 221:104895.

Shevchenko, I. and Berloff, P. (2015). Multi-layer quasi-geostrophic ocean dynamics in eddy-resolving regimes. *Ocean Modell.*, 94:1–14.

Shevchenko, I. and Berloff, P. (2016). Eddy backscatter and counter-rotating gyre anomalies of midlatitude ocean dynamics. *Fluids*, 1(3):1–16.

Shevchenko, I. and Berloff, P. (2021). A method for preserving large-scale flow patterns in low-resolution ocean simulations. *Ocean Model.*

Shevchenko, I., Berloff, P., Guerrero-Lóopez, D., and Roman, J. (2016). On low-frequency variability of the midlatitude ocean gyres. *J. Fluid Mech.*, 795:423–442.

Shu, C. and Osher, S. (1988). Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.*, 77:439–471.

Siegel, A., Weiss, J., Toomre, J., McWilliams, J., Berloff, P., and Yavneh, I. (2001). Eddies and vortices in ocean basin dynamics. *Geophys. Res. Lett.*, 28:3183–3186.

Woodward, P. and Colella, P. (1984). The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comput. Phys.*, 54:115–173.

Xie, X., Mohebujjaman, M., Rebholz, L., and Iliescu, T. (2018). Data-driven filtered reduced order modeling of fluid flows. *SIAM J. Sci. Comput.*, 40:B834–B857.