frontiers in Marine Science

TYPE Original Research
PUBLISHED 11 March 2026
DOI 10.3389/fmars.2026.1739607

OPEN ACCESS

EDITED BY
Pavel Berloff,
Imperial College London,
United Kingdom

REVIEWED BY
Sergey Karabasov,
Queen Mary University of London,
United Kingdom
Dmtry Sein,
Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI), Germany

*CORRESPONDENCE
Andrew C. Poje
andrewpoje@gmail.com

RECEIVED 04 November 2025
REVISED 12 January 2026
ACCEPTED 13 February 2026
PUBLISHED 11 March 2026

CITATION
Poje AC, Uchida T, Jamet Q, Sun L, Penduff T, Deremble B, Schoonover J, Trapanese M, Wienders N and Dewar WK (2026) Thoughts on prognostically modeling an eddying double-gyre ensemble mean.
Front. Mar. Sci. 13:1739607.
doi: 10.3389/fmars.2026.1739607

Thoughts on prognostically modeling an eddying double-gyre ensemble mean

Andrew C. Poje[1]*, Takaya Uchida[2,3,4], Quentin Jamet[5,6], Luolin Sun[2,3,7], Thierry Penduff[3], Bruno Deremble[3], Joseph Schoonover[8], Megan Trapanese[1], Nicolas Wienders[2] and William K. Dewar[3,7]

[1]Department of Mathematics, College of Staten Island, The City University of New York, Staten Island, NY, United States, [2]Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, FL, United States, [3]Université Grenoble Alpes, CNRS, IRD, Grenoble-INP, Institut des Géosciences de l'Environnement, Grenoble, France, [4]Climate Dynamics Laboratory, Center for Earth Sciences, Moscow Institute of Physics and Technology (MFTI), Dolgoprudny, Russia, [5]Service Hydrographique et Oceanographique de la Marine (SHOM), Brest, France, [6]Institut National de Recherche en Sciences et Technologies du Numerique (INRIA), ODYSSEY Group, Plouzane, France, [7]Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, FL, United States, [8]Fluid Numerics, Hickory, NC, United States

We address the question of separating the ocean's deterministic response to time-dependent forcing from its intrinsic chaotic variability. Ideally, one could compute the ensemble mean directly without performing numerous realizations, but this requires knowledge or closure of the second-order statistics — the classical turbulent-closure problem, here recast for a non-equilibrium, geophysical setting. Building on the ideas of nonlinear midlatitude ocean adjustment, we examine this problem using idealized quasi-geostrophic (QG) double-gyre ensembles subjected to episodic temporal variations in wind forcing. Our objective here is not to develop a subgrid parameterization of unresolved eddies, but rather to construct and test prognostic equations for the ensemble mean itself, using the simplest possible closure assumptions. We find that the performance of ensemble mean closures is highly dependent on the spatiotemporal structure of the forcing. Under slowly varying forcing, approximate closures reproduce the mean evolution reasonably well; under rapidly varying, near-zero-mean forcing, the simplest ensemble-mean closures fail, even at the level of basin-averaged total energy and enstrophy. In both regimes, the ensemble-mean response is not simply the accumulated imprint of the applied forcing, but instead appears as a continuing, non-equilibrated dialogue between the mean and eddy fields.

KEYWORDS
eddy parameterization, ensemble mean, ensemble simulation, mesoscale eddies, quasi-geostrophy, wind-driven gyre

1 Introduction

Understanding how large-scale ocean circulations adjust to changes in external forcing remains one of the central challenges of geophysical fluid dynamics. At climate scales, the difficulty is not simply that the governing equations are nonlinear and chaotic, but that the ocean's mean state and its intrinsic variability are dynamically entangled. Even under steady forcing, mesoscale turbulence continually feeds back on the large-scale flow, producing a fluctuating equilibrium that is only statistically stationary. When the forcing itself varies in time, this balance is disturbed and the ocean's adjustment reflects both deterministic and stochastic elements of the dynamics. Predicting that adjustment—and, in particular, predicting the evolution of the *ensemble mean* circulation—is the subject of this paper.

In modern climate modeling, such questions are typically recast as problems of parameterization. Global circulation models cannot resolve the full spectrum of mesoscale and submesoscale motions, so their collective influence must be represented through effective diffusivities or flux laws (e.g., Gent and McWilliams, 1990; Gent et al., 1995; Marshall et al., 2012; Mak et al., 2017; Wei et al., 2024). Despite their success in stabilizing coarse-resolution models, these schemes rest on heuristic assumptions—most notably that eddy fluxes act downgradient with respect to mean quantities (Cessi, 2008; Eden and Greatbatch, 2008; Ferrari et al., 2010; Eaves et al., 2025)]—whose physical justification remains limited. In particular, eddy–mean flow interactions involve both reversible exchanges and irreversible transfers of energy and potential vorticity, allowing energy to flow in either direction between the mean circulation and the fluctuating field rather than strictly downscale in a diffusive sense (Vallis, 2006; Arbic et al., 2014; Uchida et al., 2022b, 2024a; b).

While many turbulent flows admit well-defined temporal or spatial averaging—owing to statistical stationarity, periodicity, or the presence of homogeneous directions—such structure is absent in most oceanic settings. The forcing is neither steady nor periodic, and both the forcing pattern and boundary conditions preclude spatial homogeneity. As a result, neither time nor spatial averages yield a meaningful evolving mean field. Statistical analysis must therefore rely on ensemble averaging, defined here as the arithmetic mean over realizations with equally probable initial conditions. Unlike time or spatial means, the ensemble mean retains the full space–time dimensionality of the instantaneous flow and evolves in both space and time. The same is true of higher-order ensemble statistics, including the Reynolds stresses that couple the mean and fluctuating fields.

As such, ensemble averaging (Smagorinsky, 1963; Kraichnan, 1976; Young, 2012; Maddison and Marshall, 2013; Serazin et al., 2015; Leroux et al., 2018; Romanou et al., 2023; Gu et al., 2024) offers a convenient—though computationally demanding—way to separate the response to the imposed forcing from the intrinsic variability arising from nonlinear instability and sensitivity to initial conditions. In contrast to spatial or temporal filtering, which generally partitions any applied forcing into both the mean and fluctuation equations, applying any identical forcing across all ensemble members implies that the external forcing enters only the ensemble–mean equation, leaving the fluctuations unforced in a precise algebraic sense.

Here we shift focus from individual realizations of a turbulent flow to the ensemble mean itself as a dynamical object. For non-autonomous forcing, the ensemble mean defines a time-dependent statistical state that is reproducible across realizations under identical forcing, but not in general reducible to a single trajectory. The question we address is whether one can construct a stable, closed prognostic equation for this ensemble mean using only minimal and physically motivated information about the fluctuating field, whose dominant scales are determined by the dynamics rather than prescribed *a priori*. The aim is not to predict individual realizations or the detailed evolution of the fluctuations, but to determine under what conditions the ensemble-mean evolution can be captured without explicitly simulating a large ensemble of realizations.

As a minimal model for the wind-driven gyres of the North Atlantic and Pacific, the quasi-geostrophic (QG) double-gyre model has long served as a paradigm for studying the large-scale response of the ocean to wind forcing (Veronis, 1963). Despite its simplicity, the system captures many of the essential ingredients of the midlatitude ocean circulation: basin-scale recirculations, an energetic eastward jet reminiscent of the Gulf Stream or Kuroshio. In this idealized setting, nonlinear eddy–mean flow interactions can be isolated and examined without the confounding influences of complex topography or buoyancy forcing. The double-gyre model thus provides a compact, easily-computed, and physically interpretable test bed for exploring how large-scale oceanic jets adjust to changes in external forcing.

A substantial body of work has shown that such flows behave as weakly nonlinear oscillators, capable of multiple equilibria, regime transitions, and intrinsic low-frequency variability even under steady forcing (Berloff and McWilliams, 1999; Simonnet, 2005; Berloff et al., 2007). In a broader sense, this behavior exemplifies the idea that the ocean's mean circulation should be viewed as a continuously forced, weakly non-equilibrium system, in perpetual dialogue with its own turbulence—a perspective articulated in earlier studies of midlatitude adjustment and energetics (Dewar, 2003). Here, we revisit that viewpoint using ensembles of QG double-gyre simulations subjected to episodic and oscillatory wind forcing. The QG framework retains the essential baroclinic dynamics and nonlocal eddy feedbacks of the mesoscale ocean (Grooms et al., 2015; Uchida et al., 2021; 2022a; Deremble et al., 2023), while remaining simple enough to permit fully controlled ensemble experiments over climatically relevant timescales.

Our approach is deliberately simple. We first establish a statistically stationary, eddying double-gyre circulation under steady wind forcing and diagnose its time-mean and fluctuating properties, including the dominant space–time modes obtained via Spectral Proper Orthogonal Decomposition (SPOD; Towne et al., 2018). We then generate large ensembles of simulations subjected to two distinct classes of time-dependent forcing. The first (Case 1) modulates only the amplitude of the large-scale wind stress, producing a basin-scale "pulse" that perturbs the mean jet but

leaves its spatial structure unchanged. The second (Case 2) imposes forcing patterns with the same space–time scales as the most energetic eddy modes, effectively driving the system at the fluctuation scales. These two experiments bracket the spectrum of possible eddy–mean interactions, from slowly varying, quasi-equilibrium adjustment to rapidly varying, near-resonant excitation.

To clarify the dynamical controls governing the evolution of the ensemble mean, we derive a simple prognostic closure model and compare its behavior with fully simulated, reference ensembles and with two idealized dynamical-response models. The prognostic *steady-stress* (or *frozen-turbulence*) model replaces the instantaneous Reynolds stresses with their long-time means, providing a minimal closure that predicts the mean field from fixed eddy statistics. The two response models, by contrast, completely neglect the influence of ensemble fluctuations, describing how the mean field adjusts to prescribed forcing in the absence of any eddy contributions. The nonlinear version retains advection of the mean flow but omits coupling to the background state, while the linear version further simplifies the dynamics to a classical β-plane response. Comparing these simplified models to the fully diagnosed ensemble evolution allows us to assess when, and under what forcing regimes, the ensemble mean can be accurately predicted from limited statistical information.

The results demonstrate that the viability of simple prognostic closures of the ensemble mean depends critically on the temporal and spatial structure of the forcing. The forcing configurations examined here are deliberately chosen diagnostic cases. When the forcing acts on large, basin-scale structures (Case 1), the ensemble mean responds coherently and the frozen-turbulence closure performs well. When the forcing projects onto the dynamically active eddy scales—identified here using SPOD of the unforced turbulent state (Case 2)—energy is rapidly transferred from the mean to the fluctuating field, and all simplified closures fail, even at the level of total energy. In both cases, the ensemble mean is not merely the accumulated imprint of the external forcing, but a dynamically active field engaged in continuous exchange with the underlying turbulence.

These findings highlight the non-equilibrium nature of the oceanic mean state and provide a controlled framework for evaluating closure assumptions in more complex models. While our experiments are idealized, they reveal general principles likely to extend to the real ocean: namely, that successful prognostic modeling of the ensemble mean depends not only on the amplitude of the forcing, but also on its spatial and temporal structure relative to the intrinsic variability of the system.

The remainder of this paper is organized as follows. Section 2 describes the QG model and ensemble methodology, including the SPOD analysis used to identify dominant eddy modes. Section 3 presents the results from the two forcing experiments and evaluates the performance of the three prognostic models. The focus of this paper is on the ensemble mean, referring to past results that emphasize the importance of capturing an accurate mean state in order to accurately capture the eddies (Hallberg, 2013; Mak et al., 2023), but we shall end with some discussion of the eddies in Section 4 and on the broader problem of 124 predicting the large-scale ocean response to time-dependent forcing.

2 Data and method

Throughout our study, we document the characteristics of an ensemble-mean oceanic jet using the QG wind-driven gyre system. We numerically solve the canonical QG potential vorticity (PV) using the Fast-Fourier Transform (FFT)-based qgw solver Deremble et al. (2024),

$$\frac{\partial q_i}{\partial t} + J(\psi_i, q_i) = \frac{\tau_0}{H_1} F(\boldsymbol{x}, t)\delta_{i,1} + \nu\nabla^2 q_i \,. \tag{1}$$

For two layers, the layer PV, q_i, and stream function, ψ_i, are non-locally related by:

$$q_1 = \nabla^2\psi_1 + \frac{f_0^2}{g' H_1}(\psi_2 - \psi_1) + \beta y,$$
$$q_2 = \nabla^2\psi_2 + \frac{f_0^2}{g' H_2}(\psi_1 - \psi_2) + \beta y \,.$$

The total energy (per unit density and area) in the domain is given by

$$\begin{aligned} \mathrm{TE}(t) &= \mathrm{KE}(t) + \mathrm{APE}(t) \\ &= -\frac{1}{2A}\sum_{i=1}^{2}\int H_i \psi_i | q_i \mathrm{d}x\mathrm{d}y := -\{q\psi\} \,, \end{aligned} \tag{2}$$

with units $\frac{\mathrm{m}^3}{\mathrm{s}^2}$.

2.1 Base case

We take the simplest two-layer configuration, $(H_1, H_2) = (400, 2600)$ m, in a square domain $L_x = L_y = L = 3840$ km on an interior grid 512^2 ($\Delta x = 7.5$ km). The parameters are $f_0 = 9.4 \times 10^{-5}$ s^{-1}, $\beta = 1.7 \times 10^{-11}$ m^{-1} s^{-1}, and $g' = 0.045$ m s^{-2} with a resulting Rossby deformation radius $R_d = 42$ km.

The steady, asymmetric wind-stress curl follows Berloff and McWilliams (1999):

$$\begin{aligned} F(x,t) &= F(y) \\ &= \frac{\pi}{L}\left\{\sin\left[2\pi\frac{(y - L/2)}{L}\right] - \lambda_0 \cos\left[\pi\frac{(y - L/2)}{L}\right]\right\}, \end{aligned} \tag{3}$$

with $\tau_0 = 4.0 \times 10^{-5}$ m^2 s^{-2} and asymmetry parameter $\lambda_0 = 0.25$. This choice places the zero wind-stress curl at $y \approx 2000$ km, slightly breaking the meridional symmetry.

The numerical solution is computed with free-slip boundary conditions for 1,000 years following a 50 year spin-up from rest. The Laplacian diffusivity is $\nu = 75$ m^2 s^{-1}. With $\Delta x = 7.5$ km and $R_d = 42$ km, the deformation radius is resolved by $R_d/\Delta x \approx 5.6$ grid points, placing the reference configuration in the commonly used eddy–resolving regime for two–layer QG double–gyre flows. The domain–scale separation is also large, with $L/R_d \approx 91$. While classical Munk–layer estimates are sometimes invoked as resolution criteria, such viscous boundary layers arise under no–slip conditions and are absent in the present free–slip formulation. Appendix A demonstrates that the computed flows and statistics are robust across spatial resolutions.

The results of steady wind forcing at these parameters show a statistically stationary eddying flow with ~10-15% fluctuations of

10.3389/fmars.2026.1739607

the total energy about the time-mean state (Figure 1A). As shown in the spatial plots of the upper layer PV, the flow is strongly eddying. With the time-mean defined by $\bar{g}(x) \stackrel{\text{def}}{=} \lim_{T\to\infty} \frac{1}{T}\int_0^T g(x,\, t)$, the time-averaged total energy (Equation 2) can be decomposed into mean and eddy reservoirs

$$\overline{\text{TE}} = -\{\bar{q}\bar{\psi}\} - \overline{\{q'\psi'\}} = \text{ME} + \text{EE}$$

and we find a roughly even split, Total energy = 89 units = 53 (mean) + 46 (eddy), between the two components.

To quantify the dominant coherent space–time structures within the statistically stationary eddying regime, we apply Spectral Proper Orthogonal Decomposition (SPOD; Towne et al., 2018) to the potential vorticity and streamfunction fields. SPOD provides an energy-optimal modal basis in the joint spatial–temporal sense: each mode represents a coherent structure oscillating at a single frequency, ranked by its contribution to the total fluctuation energy. In contrast to the traditional spatial POD, which diagonalizes the spatial covariance at zero time lag, or Singular Spectrum Analysis (Ghil et al., 2002) that considers time-lagged covariances, SPOD diagonalizes the cross-spectral density tensor—the Fourier transform of the temporal correlation operator—thereby isolating physically meaningful, frequency-resolved modes. The mathematical details of the implementation, including the energy inner product appropriate to the quasi-geostrophic system, are provided in Appendix A.

For the present analysis, we use the final 1,000 years of statistically steady data. To estimate the cross-spectral statistics with adequate frequency resolution and statistical convergence, the full time series is divided into ten non-overlapping segments of 100 years each. Each block is windowed, Fourier transformed, and used to compute a segment-averaged cross-spectral density matrix from which the SPOD eigenvalues and modes are obtained.

Figure 2 summarizes the results. Panel (a) shows the SPOD eigenvalue spectrum for the leading two modes, $\lambda_1(f)$ (black) and $\lambda_2(f)$ (red), as functions of the frequency. The eigenvalue curves indicate how the total fluctuation energy is distributed over frequency and between coherent spatial patterns. Three frequency bands, highlighted by dashed vertical lines, correspond to distinct dynamical regimes of the flow: a low-frequency, large-scale meandering of the mean jet ($f \approx 0.21$), an intermediate-frequency mode associated with gyre-scale recirculation variability ($f \approx 0.65$), and a higher-frequency, mesoscale wave-like pattern ($f \approx 0.95$).

Panels (b–d) show the real parts of the leading SPOD potential vorticity modes, $\Re[\phi_1(x;f)]$, normalized by their spatial amplitude for the three distinct time-scales. The low-frequency mode is equivalent to the 'gyre-mode' examined in Berloff et al. (2007) encapsulating slow jet migration and changes in the intergyre boundary. The intermediate-frequency mode captures coherent eddy-shedding fluctuations along the western boundary current extensions, while the selected high-frequency modeshows compact, oscillatory vortical features localized in the jet core and recirculation

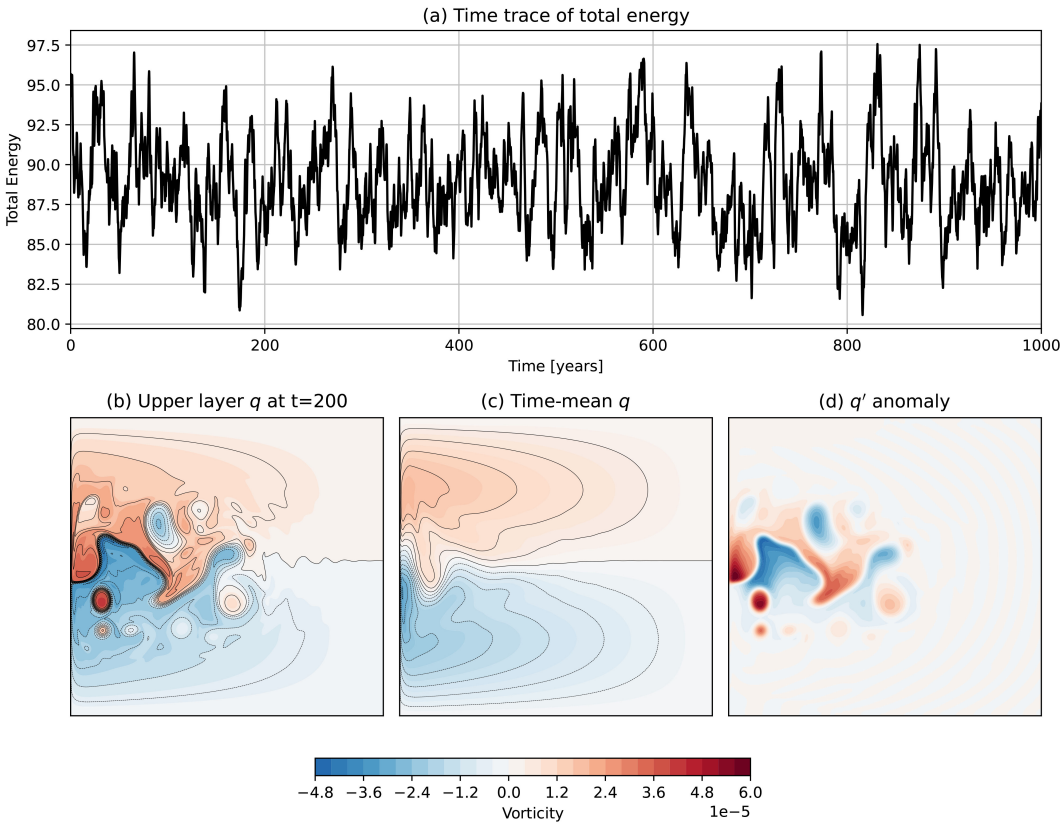

FIGURE 1
(A) 1,000-year time trace of the volume-averaged energy under steady forcing. **(B–D)** Snapshots of the upper-layer PV at time 200 years **(B)**, the time-mean upper-layer PV, $\bar{q}(x)$ **(C)**, and the fluctuation upper-layer PV, $q'(x,\, t)$ at time 50 years **(D)**.

10.3389/fmars.2026.1739607

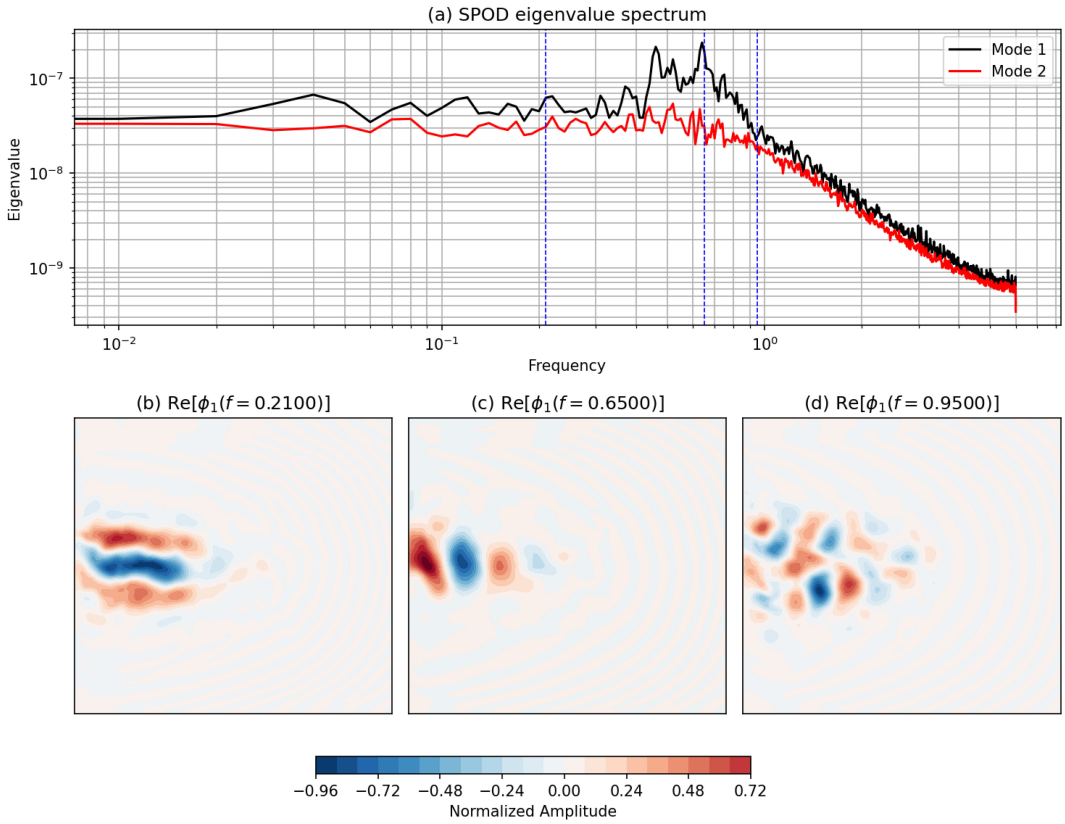

FIGURE 2
Spectral proper orthogonal decomposition from 1,000 years of steady wind forcing. **(A)** Temporal spectra of first two POD modes [frequencies in units of year $^{-1}$). **(B–D)** Spatial patterns of the real part of the leading POD pv-mode at three selected frequencies (corresponding to blue lines in **(A)**].

zones. Together, these modes form a natural energetic hierarchy of the flow variability, separating slowly varying gyre–jet adjustment from the more rapid eddy motions that potentially modulate it.

2.2 Modeling evolution of the ensemble mean

Statistical stationarity implies, via the ergodic theorem (Birkhoff, 1931; Frisch, 1995; Galanti and Tsinober, 2004; McWilliams, 2006), the equivalence of time and ensemble averages. If the forcing is independent of time (Equation 3), then the time average of (Equation 1) is

$$J(\bar{\psi}_i, \bar{q}_i) = F(y)\delta_{i,1} + \nu\nabla^2\bar{q}_i - \overline{J(\psi_i', q_i')}\,, \tag{4}$$

where $\bar{f}(x) \stackrel{\text{def}}{=} \lim_{T\to\infty} \frac{1}{T}\int_0^T f(x,\ t)$ and, for bounded vorticity, the contributions from the time derivative can be made vanishingly small for large enough T. Importantly, under steady forcing, the evolution equation for the fluctuations, $q'(\mathrm{x}, t) = q(x, t) - \bar{q}(x)$,

$$\frac{\partial q_i'}{\partial t} + [J(\bar{\psi}_i + \psi_i', q_i') + J(\psi_i', \bar{q}_i + q_i')] = \nu\nabla^2 q_i' + \overline{J(\psi_i', q_i')}\,, \tag{5}$$

contains no direct contribution from the forcing. The fluctuations feel the effect of steady forcing only through its imprint on the time-averaged stream function, potential vorticity, and Reynolds stress terms. This is distinctly different from other situations forced to statistical stationary, perhaps stochastically, by forcing of the form $F_i(x, \mathrm{t}) = \bar{\mathrm{F}}(x) + \mathrm{F}_i'(x, \mathrm{t})$ where the fluctuations are directly driven. The time-mean decomposition of the flow field shown in the lower panels of Figure 1 clearly indicates that the details of the instantaneous interior PV at any time is dominated by the intrinsic variability represented by q', but its large-scale structure, of weaker amplitude, is captured by the time mean and mostly reflects the imposed forcing.

Equivalent results hold for temporally periodic forcing with given period, $\mathcal{T}$,

$$F(x, t) = F(x, t + \mathcal{T})\,.$$

In this case, the forcing is invariant under phase averaging,

$$\bar{\bar{g}}(x, \tau) \stackrel{\text{def}}{=} \lim_{M\to\infty} \frac{1}{M}\sum_{j=0}^{M} g(x, \tau + j\mathcal{T}),$$

with $\tau \in [0, \mathcal{T})$. Since $\bar{\bar{F}}(x, \tau) = F(x, \tau)$, (Equation 5) holds for the phase-averaged fluctuations which receive no direct input from the periodic forcing.

In the following, we shall develop ensembles of two-layer QG double-gyre simulations which only differ by their initial conditions. Before presenting the results from the numerical experiments, we consider the consequences of aperiodic temporal forcing. Under these conditions, time is no longer a homogeneous direction and the only rigorous averaging operator in the statistical sense is the ensemble average,

$$\langle f(x,t)\rangle \overset{\text{def}}{=} \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} f_n(x,t)\,.$$

Here the individual members of the ensemble, f_n, are assumed to be drawn from the set of equally probable flow states at some initial time. By construction, the ensemble mean commutes with any linear spatio-temporal operator and the ensemble mean retains the full dimensionality of any single ensemble member (Chen and Flierl, 2015; Sérazin et al., 2017; Jamet et al., 2022; Uchida et al., 2021, 2023, 2024a). Mathematical notations are summarized in Table 1.

Although the forcing in (Equation 1) is now time dependent, the fact that it is simply additive implies that $\langle F(x,t)\rangle = F(x,t)$. The evolution of the ensemble-mean PV,

$$\frac{\partial\langle q\rangle_i}{\partial t} + J(\langle\psi\rangle_i,\langle q\rangle_i) = F(y,t)\delta_{i,1} + \nu\nabla^2\langle q\rangle_i - \left\langle J(\psi_i^\dagger, q_i^\dagger)\right\rangle\,, \qquad (6)$$

differs from that of the time mean (Equation 4) only by the retention of the time derivative and, critically, by the statistical definition of the fluctuations comprising the eddy stress term. The corresponding evolution of the ensemble mean energy, $\mathrm{ME} = -\{\langle q\rangle\langle\psi\rangle\}$ is given by,

$$\frac{d}{dt}\mathrm{ME} = -H_1\iint\langle\psi_1\rangle F\,\mathrm{d}A - \mathcal{D} + \Pi \qquad (7)$$

where the transfer from mean to eddy reservoirs is

$$T_{M\to E} = -\Pi \overset{\text{def}}{=} -\frac{1}{2}\sum_{i=1}^{2} H_i\iint\langle\psi_i\rangle\left\langle J(\psi^\dagger,q^\dagger)\right\rangle\mathrm{d}A\,. \qquad (8)$$

By construction, for identically forced ensemble members, there is no direct contribution of such forcing to fluctuations about the ensemble mean (i.e., the forcing does not appear in the equation corresponding to fluctuations; $q^\dagger \overset{\text{def}}{=} q - \langle q\rangle, F^\dagger = 0$), and the ensemble fluctuations evolve in analogy to (Equation 5).

The parameterization problem involves modeling the final expression on the right-hand side of (Equation 6) in terms of known mean quantities: $\left\langle J(\psi^\dagger,q^\dagger)\right\rangle = \mathcal{G}(\langle q\rangle)$. Here we consider the ensemble response to episodic changes in the forcing about some steady reference state. The goal is to predict the deviation of the ensemble mean from the given reference steady state whose time-mean statistics are known (based, for example, on our 1,000-year simulation; Figure 1).

TABLE 1 Definition of mathematical notations.

Notation	Description
$\overline{(\cdot)}$	Time mean
$(\cdot)' = (\cdot) - \overline{(\cdot)}$	Temporal fluctuation
$\langle\cdot\rangle$	Ensemble mean
$(\cdot)^\dagger = (\cdot) - \langle\cdot\rangle$	Ensemble fluctuation
$\widetilde{(\cdot)} = \langle\cdot\rangle - \overline{(\cdot)}$	Temporal deviation of ensemble mean from its steady state
$\tilde{q}_{\mathrm{FT}}$	Steady-stress (frozen turbulence) model (11)
$\tilde{q}_{\mathrm{LR}}$	Linear response model (12)
$\tilde{q}_{\mathrm{NLR}}$	Nonlinear response model (13)

The equation for the deviation of the ensemble mean from the temporal mean, $\tilde{q}_i(x,t) = \langle q_i\rangle(x,t) - \bar{q}_i(x)$ (dropping the layer index i for brevity) is

$$\frac{\partial\tilde{q}}{\partial t} + J(\tilde{\psi},\tilde{q}) + J(\tilde{\psi},\bar{q}) + J(\bar{\psi},\tilde{q}) = \mathcal{L}(\tilde{q}) + \tilde{F} + \Big[\overline{J(\psi',q')} - \underbrace{\left\langle J(\psi^\dagger,q^\dagger)\right\rangle}_{=\mathcal{G}(\langle q\rangle)}\Big] \qquad (9)$$

$\tilde{F} = F(y,t) - \bar{F}(y)$ is the temporal deviation of the imposed forcing from its steady value, and $\mathcal{L}(\tilde{q})$ is the linear diffusion term.

Assuming knowledge of the first-order, time-mean statistics $(\bar{q},\bar{\psi})$, of Equation (9) can be closed by adopting a quasi-equilibrium, or frozen-turbulence, approximation (Taylor, 1938; Farrell and Ioannou, 2003; Marston et al., 2016) and simply ignoring any temporal variations in the second-order fluctuation terms. Setting

$$\left\langle J(\psi^\dagger,q^\dagger)\right\rangle = \mathcal{G}(\langle q\rangle) \approx \overline{J(\psi',q')}\,, \qquad (10)$$

produces a closed, nonlinear model for $\tilde{q}_{\mathrm{FT}}$ that includes interaction with the underlying time-mean fields,

$$\frac{\partial\tilde{q}_{FT}}{\partial t} + J(\tilde{\psi}_{\mathrm{FT}},\tilde{q}_{\mathrm{FT}}) + J(\tilde{\psi}_{\mathrm{FT}},\bar{q}) + J(\bar{\psi},\tilde{q}_{\mathrm{FT}}) = \mathcal{L}(\tilde{q}_{\mathrm{FT}}) + \tilde{F}\,. \qquad (11)$$

Although this closure (Equation 10) is by no means "perfect," it should be viewed as a baseline rather than an optimal scheme. In the steady-stress (frozen-turbulence) model, the Reynolds stresses are fixed at their statistically steady values under the control forcing and cannot evolve in response to changes in the ensemble-mean flow. As a result, the model cannot represent the time-dependent transfer of energy (Equations 7, 8) or potential vorticity between the ensemble mean and the eddy field. Within that limitation, it nevertheless provides a simple and physically interpretable benchmark, likely competitive with existing prognostic mesoscale parameterizations so long as the modeled regime does not cross a bifurcation point (Simonnet et al., 2003).

Given the common interpretation of the ensemble mean as the forced response of the system (Penduff et al., 2018; Zhao et al., 2021; Narinc et al., 2024; Uchida et al., 2024a; Sane et al., 2024; Takasuka et al., 2025), we also consider simple force-response models that completely ignore the effects of the fluctuations; both the fully linear response model, $\tilde{q} = \tilde{q}_{\mathrm{LR}}$,

$$\frac{\partial\tilde{q}_{LR}}{\partial t} + \beta\frac{\partial\tilde{\psi}_{\mathrm{LR}}}{\partial x} = \mathcal{L}(\tilde{q}_{\mathrm{NLR}}) + \tilde{F} \qquad (12)$$

as well as the nonlinear response, $\tilde{q} = \tilde{q}_{\mathrm{NLR}}$,

$$\frac{\partial\tilde{q}_{NLR}}{\partial t} + J(\tilde{\psi}_{\mathrm{NLR}},\tilde{q}_{\mathrm{NLR}}) = \mathcal{L}(\tilde{q}_{\mathrm{NLR}}) + \tilde{F}\,. \qquad (13)$$

Comparison of the full ensemble-mean response, $\langle q\rangle$, to $\tilde{q}_{\mathrm{FT}}$, $\tilde{q}_{\mathrm{NLR}}$ and $\tilde{q}_{\mathrm{LR}}$ quantifies the relative importance of nonlinearity, time-mean interactions and temporal variations in the Reynolds stresses in response to changes in the forcing.

All three response models are integrated using the same numerical discretization, time stepping, spatial grid, and boundary conditions as the fully nonlinear QG simulations described in Section 2.1. Differences among the frozen-

turbulence, nonlinear-response, and linear-response models therefore reflect only their respective dynamical assumptions regarding eddy–mean interactions, rather than differences in numerical configuration or resolution.

3 Results

In what follows, we investigate how the statistically stationary eddying flow responds to episodic changes in the imposed forcing. As shown above, for identically forced ensemble members, deviations from the steady forcing act only on the ensemble mean. The spatial–temporal structure of the mean field determines the pathways through which energy and potential vorticity anomalies are communicated to, and subsequently redistributed by, the eddy field. Here we consider two extreme cases.

In Case 1, the spatial form of the wind stress is fixed, but the amplitude varies in time as

$$\tau_0 \rightarrow \tau_0(1 + b(t; a, t_0, \sigma)),$$

where $b(t)$ is the Gaussian

$$b(t; a, t_0, \sigma) = a \exp\left(\frac{(t - t_0)^2}{2\sigma^2}\right).$$

In Case 2, by contrast, the spatial and temporal scales of $\tilde{F}(x, t)$ are chosen to roughly match those of the most energetic eddy structures identified through the SPOD analysis of the statistically steady flow. This formulation allows us to explore how the ensemble-mean circulation responds when externally driven by patterns that resemble the internal, energetic modes of the turbulent flow, rather than by a basin-scale modulation of the steady wind stress.

This forcing framework—examined here in the ensemble-mean context—is closely related to that introduced by Dewar (2003) in his study of nonlinear midlatitude ocean adjustment. Our Case 1 is essentially a "turn-on/turn-off" experiment, while Case 2 corresponds conceptually to his periodically forced regime, in which the spatial and temporal structure of the forcing is tuned to resonate with the intrinsic variability of the system.

3.1 Case 1: change in wind-stress amplitude

We begin with the simpler case of a time-dependent modulation of the large-scale wind stress amplitude. Each ensemble member is integrated for 20 years under this forcing history, which represents a short-duration, basin-wide amplification of the same spatial forcing pattern that maintains the steady double-gyre circulation. The temporal modulation is prescribed as a Gaussian pulse with parameters $a = 4.0$, $t_0 = 3$, and $\sigma = 1/2$ years.

An ensemble of initial conditions is constructed by randomly sampling the statistically stationary 1,000-year reference simulation shown in Figure 1. A total of 120 initial states are selected, with a minimum separation of five years between samples to ensure statistical independence. Each ensemble member differs only in its initial condition; all model parameters and forcing histories are otherwise identical. For the diagnostics emphasized in this study, ensemble-mean statistics are not sampling-limited: the standard error of the basin-integrated ensemble-mean energy at the initial time is approximately 0.3% of the mean, and subsampling the ensemble indicates that the ensemble-mean energy differs by less than 1% when comparing 80-member and 120-member ensembles.

Figure 3 summarizes the ensemble-mean energetics. Panel (a) shows the evolution of the total energy for all ensemble members (light gray) together with the ensemble-mean total energy (black) and the ensemble-mean eddy energy (red). The blue curve indicates the temporal shape of the imposed forcing anomaly $\tilde{F}(t)$.

As shown in Figures 3A, the fourfold increase in wind stress produces a rapid, roughly 350% rise in total energy, peaking about one year after the maximum forcing. This response is followed by a slower relaxation back toward the equilibrium value. Decomposition of the energy into mean and eddy components (Figure 3A) indicates that the injected energy is directly fed into the ensemble-mean component, with little short-term change in the eddy reservoir during the brief forcing pulse.

Panel (b) shows the diagnosed terms of the ensemble–mean energy budget,

$$\frac{d}{dt}\mathrm{ME}(t) = P_{\mathrm{in}}(t) + \Pi(t) + D(t),$$

where $\mathrm{ME}(t) \overset{\mathrm{def}}{=} -\{\langle q\rangle\langle\psi\rangle\}$ is the basin–mean energy, $P_{\mathrm{in}}(t)$ is the wind–work input, $\Pi(t)$ is the reversible exchange between the mean and eddy reservoirs (negative values denote mean $\rightarrow$eddy transfer), and $D(t) \leq 0$ is the mean–field dissipation, diagnosed as the residual $D = d\mathrm{ME}/dt - P_{\mathrm{in}} - \Pi$. During the pulse P_{in} rises sharply and accounts for nearly all of the positive $d\mathrm{ME}/dt$, while Π remains small: the wind energizes the mean circulation directly. After the pulse, P_{in} relaxes toward its control value but Π becomes strongly negative, indicating a delayed mean $\rightarrow$eddy cascade; $d\mathrm{ME}/dt$ crosses zero when $|\Pi| + |D|$ exceeds P_{in}, and the mean energy decays on the basin–adjustment timescale. This two–stage sequence—direct wind input to the mean followed by a lagged transfer to the eddies—explains the phase offset between the input forcing and the peaks in $\mathrm{TE}(t)$ and eddy energy seen in Figure 3A. The magnitude of the direct mean dissipation is small and in phase with the mean response.

To assess model performance, we track the basin-integrated energy of the ensemble mean, $\mathrm{ME}(t) \overset{\mathrm{def}}{=} -\{\langle q\rangle\langle\psi\rangle\}$, and two field-level skill metrics between modeled and reference upper-layer ensemble-mean PV: anarea-weighted RMS error and a spatial pattern correlation $r(t)$,

$$\mathrm{RMS}(t) = \left[\frac{1}{A}\iint (\langle q\rangle_{\mathrm{mod}} - \langle q\rangle_{\mathrm{ref}})^2 \, dA\right]^{1/2},$$

$$\mathrm{r(t)} = \frac{\iint \left(\langle q\rangle_{\mathrm{mod}} - \overline{\langle q\rangle_{\mathrm{mod}}}\right)\left(\langle q\rangle_{\mathrm{ref}} - \overline{\langle q\rangle_{\mathrm{ref}}}\right) dA}{\sqrt{\iint (\langle q\rangle_{\mathrm{mod}} - \langle q\rangle_{\mathrm{mod}})^2 \, dA}\sqrt{\iint \left(\langle q\rangle_{\mathrm{ref}} - \overline{\langle q\rangle_{\mathrm{ref}}}\right)^2 dA}},$$

where $\langle q\rangle_{\mathrm{mod}} \overset{\mathrm{def}}{=} \tilde{q}_{\mathrm{mod}} + \bar{q}$ represents the modeled ensemble mean from (11)–(13) and $\langle q\rangle_{\mathrm{ref}}$ is the reference ensemble mean diagnosed

10.3389/fmars.2026.1739607

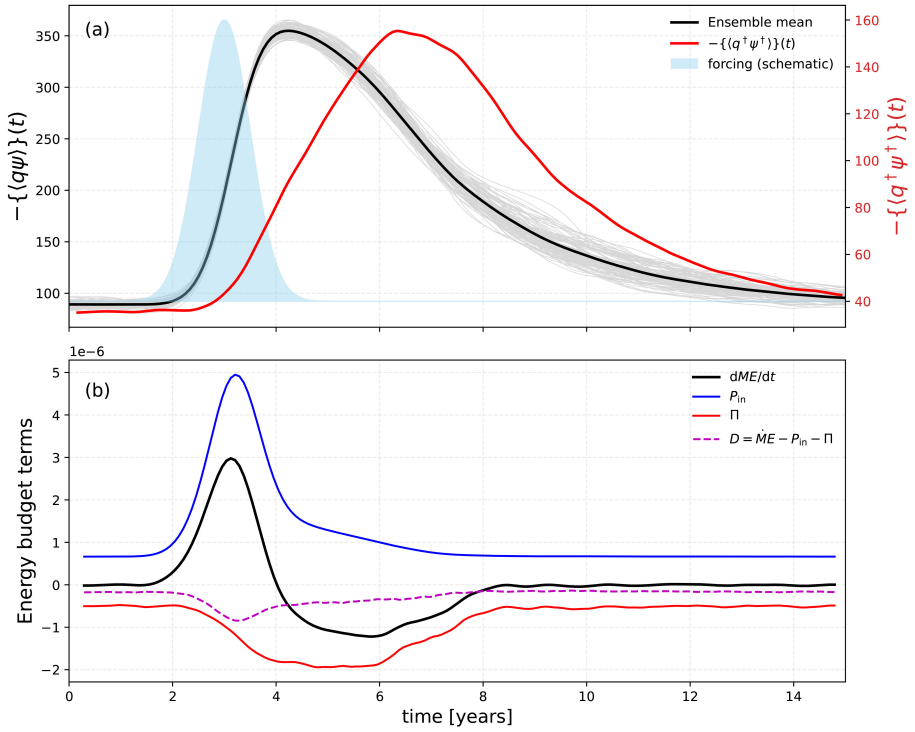

FIGURE 3
Case 1: **(A)** Left *y* axis shows evolution of the total energy for 120 individual ensemble members (light gray curves) and the ensemble mean of total energy (black). Right *y* axis shows evolution of the ensemble mean eddy energy (red). Temporal dependence of $\tilde{F}(t)$ shown schematically in blue. **(B)** Evolution of the components of reference ensemble mean energy budget.

from the actual ensemble computations. Comparisons of the three models are shown in Figure 4.

All three models reproduce the rapid, one-year-lagged rise of ME(t) and its slower decay (Figure 4A), as expected for basin-scale forcing that projects directly onto the steady mean state. Larger differences are seen in the field-level metrics (Figures 4B, C): FT has the smallest RMS and the highest $r(t)$ throughout most of the 20-year window; LR shows intermediate skill—tracking the early adjustment but losing correlation during the recovery as phase and amplitude drift; NLR performs worst, overshooting the peak response and diverging thereafter. The largest decline in $r(t)$ for steady-stress model occurs during the period of strong mean-eddy coupling shown in Figure 3B.

Snapshots of the upper-layer ensemble-mean *PV anomaly* fields (Figure 5) clarify the skill ranking. Shown are deviations from the long-time mean, $\tilde{q} = \langle q \rangle - \bar{q}$, which isolate the dynamically relevant response to the transient forcing. The diagnosed ensemble mean anomaly ($\tilde{q}_{\text{ref}}$, first column) exhibits a coherent intensification and northward displacement of the jet during the forcing pulse, followed by a gradual relaxation back toward zero anomaly. This evolution reflects a large-scale, reversible adjustment of the background circulation rather than a permanent restructuring of the mean state.

The steady-stress model ($\tilde{q}_{\text{FT}}$, second column) captures this large-scale anomaly remarkably well. The spatial organization, amplitude, and timing of the response closely match the reference, although the jet anomaly is slightly broader and more diffuse. This reflects the neglect of time-dependent eddy feedbacks; nevertheless, because the model retains the time-mean Reynolds-stress divergence, it remains anchored to the correct background flow and produces a bounded, physically consistent anomaly evolution.

The steady-stress model ($\tilde{q}_{\text{FT}}$, second column) captures this large-scale anomaly remarkably well. The spatial organization, amplitude, and timing of the response closely match the reference, although the jet anomaly is slightly broader and more diffuse. This reflects the neglect of time-dependent eddy feedbacks; nevertheless, because the model retains the time-mean Reynolds-stress divergence, it remains anchored to the correct background flow and produces a bounded, physically consistent anomaly evolution.

The nonlinear response model ($\langle q_{NLR} \rangle$, third column) behaves very differently. Lacking any reference to the background mean state or its stabilizing Reynolds-stress divergence, it reacts to the large forcing amplitude by developing its own mesoscale instabilities. Within a few years the field becomes fully eddying, effectively spinning up a new chaotic turbulent circulation unrelated to the directly computed ensemble-mean adjustment. In this sense, its apparent realism—nonlinear eddy activity—is misplaced, as it arises from the absence of stabilizing effects of the Reynolds stresses that anchor the mean-flow.

The linear response model ($\tilde{q}_{\text{LR}}$, fourth column) remains stable but produces overly smooth and symmetric PV anomalies. Because the β-term is retained, even weak zonal variations excite barotropic Rossby waves. The Gaussian temporal pulse generates a westward-propagating packet with (barotropic) phase speed $c_R = \beta/(k^2 + l^2)$. In the absence of nonlinear redistribution or mean-flow absorption, this wave activity reflects at the western boundary where $c_R \to 0$, leading to excessive PV anomaly accumulation along the western wall. In contrast, both the fully eddying ensemble and the steady-

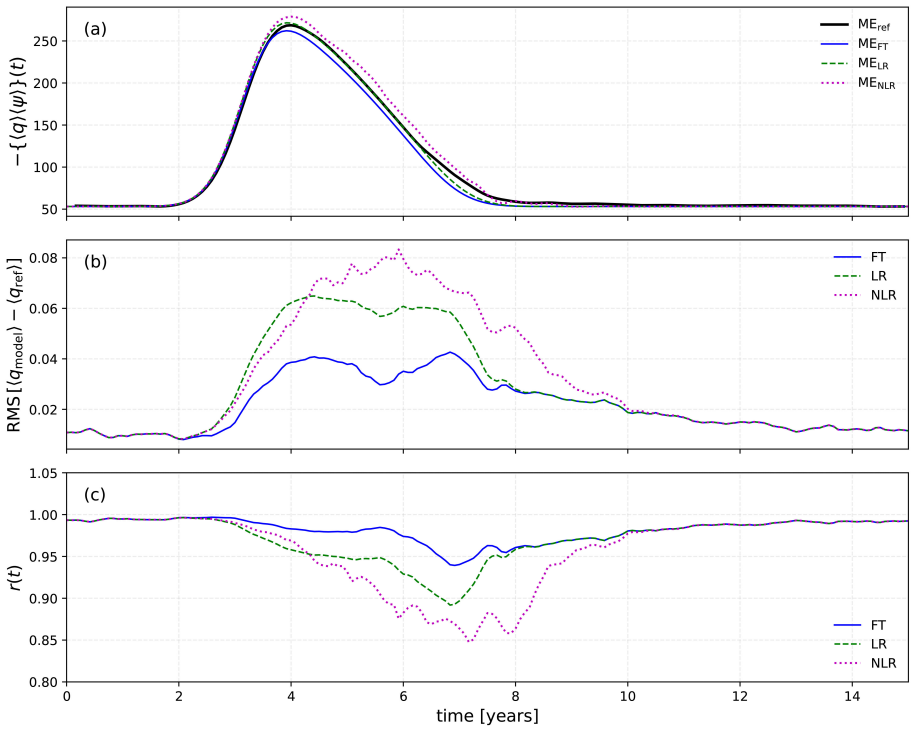

FIGURE 4
Case 1: Comparison between the directly computed and modeled upper-layer potential vorticity response to time-dependent forcing. **(A)** Mean energy $-\{\langle q\rangle\langle\psi\rangle\}(t)$ from the reference ensemble (black) and from the frozen-turbulence (FT, blue), nonlinear-response (NLR, magenta), and linear-response (LR, green) models. **(B)** Root-mean-square difference $RMS[\langle q\rangle_{model}-\langle q\rangle_{obs}]$ between the modeled and reference upper-layer ensemble mean fields. **(C)** Spatial pattern correlation $r(t)$ between modeled and reference ensemble mean fields.

stress model absorb and redistribute westward energy fluxes through eddy stresses, preventing this unphysical buildup.

Although all three models produce comparable basin-integrated energy curves, only those that maintain an explicit coupling to the time-mean flow reproduce the observed, coherent, and reversible ensemble-mean response. The linear model, dominated by β-plane Rossby adjustment and diffusion, remains overly smooth and accumulates PV at the western boundary, while the unconstrained nonlinear model destabilizes and spawns its own eddy field. The steady-stress closure, though neglecting temporal feedbacks, best captures the large-scale, bounded evolution of the ensemble mean.

The response in Case 1 thus reflects how the ensemble mean adjusts when the external perturbation acts at the same spatial scales as the time-mean forcing. In the next experiment, we shift perspective: the forcing itself is designed to operate on eddy scales, with spatial and temporal organization patterned after the dominant SPOD mode of the turbulent flow. This coherent eddy-scale (SPOD-mode) forcing provides a complementary test of how the ensemble mean responds when driven not by basin-wide modulation but by forcing at scales that characterize its intrinsic eddy variability.

3.2 Case 2: coherent eddy-scale (SPOD-mode) forcing

We next consider a perturbation whose spatial and temporal structure is patterned after the most energetic eddy variability of the statistically steady flow. As before, an ensemble of 120 identically forced integrations is performed, differing only by their initial conditions drawn from the long steady-state record. In this case, $\tilde{F}(x,t)$ is chosen to act on time-space scales consistent with the leading SPOD mode at $f = 0.65$ year^{-1}.

As shown in Figures 6A, C, the leading SPOD mode represents an oscillatory jet-centered structure. The mode exhibits a Gaussian-like envelope in the meridional direction, centered on the inter-gyre jet, and a quasi-sinusoidal variation in the zonal direction with an effective wavelength of about one quarter of the basin width. Roughly two complete zonal oscillations are visible within the first half of the domain, with the amplitude decaying downstream along the jet. To obtain a compact and analytically tractable representation, we idealize this pattern as a simple dipole in x modulated by a Gaussian in y giving the fields $A(x,y)$ and $B(x,y)$ shown in Figures 6B, D. This abstraction retains the dominant spatial phase relationship and scale of the coherent mode.

These two fields are then combined to define the time-dependent forcing,

$$\tilde{F}(x,t) = \tau_1 h(t - t_0)\{A(x,y)\cos(2\pi\omega(t-t_0)) + B(x,y)\sin(2\pi\omega(t-t_0))\}\ .$$

As shown in the insets on panels (b, d), the forcing oscillates with period 1.5 years, and $h(t)$ is a smooth bump-function envelope active for $t \in [1,9]$ years, giving 6 cycles of the forcing. The amplitude, τ_1 is set to $10\tau_0$.

Figure 7A shows the basin-averaged energetics for the full ensemble. Because all ensemble members experience identical

10.3389/fmars.2026.1739607

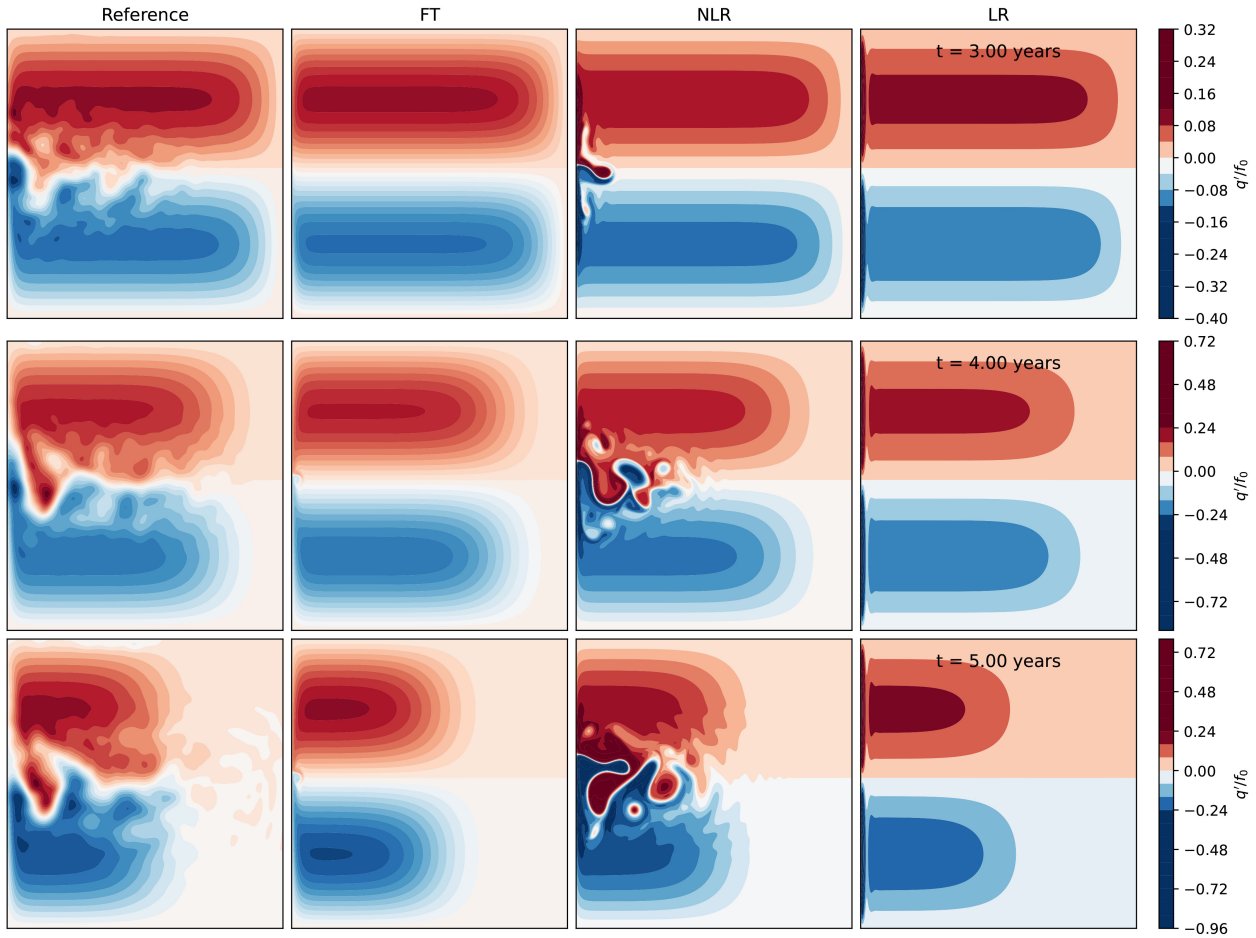

FIGURE 5
Case 1: Instantaneous snapshots of reference and modeled upper layer PV anomaly, $\tilde{q}_{ref} = \langle q \rangle_{ref} - \bar{q}$ and $\tilde{q}_{FT}, \tilde{q}_{LR}, \tilde{q}_{NLR}$ respectively, at years three, four and five; (cf. Figure 4). Colormaps and contour levels are identical across each row.

forcing, external work enters only through the ensemble-mean equations. Nevertheless, this energy is almost immediately transferred to the eddy component through nonlinear eddy–mean interactions. The modest growth in total energy over the forcing period is due to the increase in the eddy component, not the mean. Panel (b) decomposes the anomaly energy budget into its principal components—rate of change, wind-work, and mean–eddy transfer—revealing that the externally supplied power is almost immediately

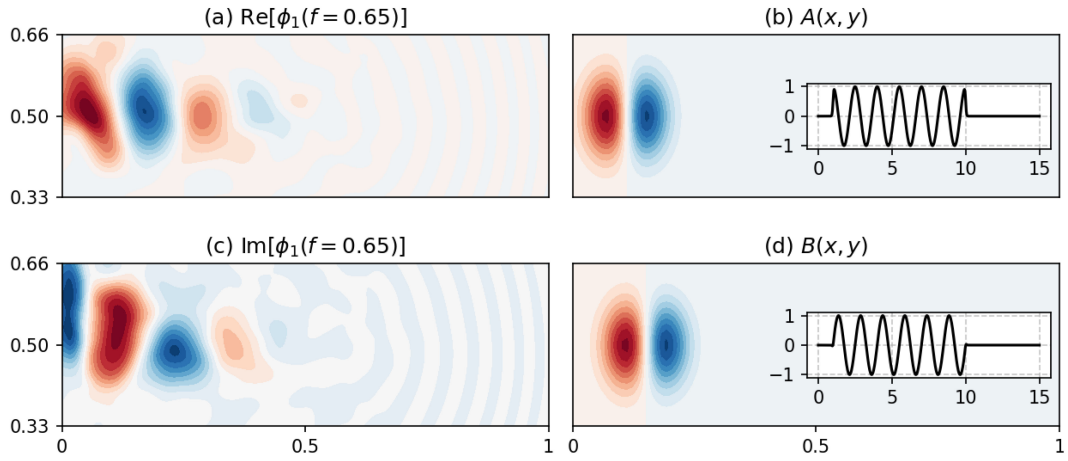

FIGURE 6
Forcing function for Case 2. **(A, C)** Real and Imaginary parts of the most energetic spatial SPOD modes (only $y \in [1/3, 2/3]L_y$ shown). **(B, D)** Idealized $A(x, y), B(x, y)$ with temporal dependence inset.

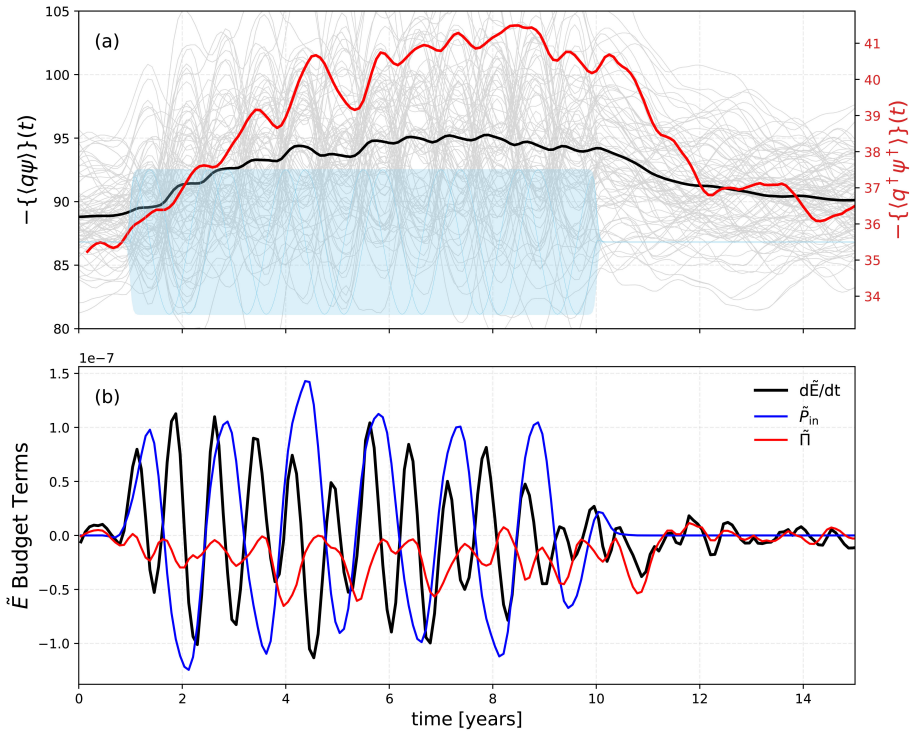

FIGURE 7
Case 2: Energetic response to the time-dependent wind-stress forcing. **(A)** Reference ensemble-mean total energy, $-\{\langle q\psi\rangle\}(t)$ (black) and eddy energy $-\{\langle q'\psi'\rangle\}(t)$ (red), computed from 120 ensemble members (thin gray lines). The shaded region denotes the temporal envelope of the oscillatory wind forcing. **(B)** Corresponding anomaly energy budget terms showing the rate of change $d\tilde{E}/dt$ (black), the wind-work input $\tilde{P}_{in}$ (blue), and the mean-to-eddy energy transfer˜(red). Both panels share the same time axis; the forcing period in **(A)** is vertically aligned with the response in **(B)**.

exported to the eddy field. In contrast to Case 1—where injected energy first accumulated in the mean reservoir and only later cascaded into the fluctuating reservoir, the response here exhibits a persistent and efficient mean →eddy energy transfer operating continuously during the forcing window. As a result, the anomaly mean response is not merely a phase-lagged copy of the forcing but instead cycles at a higher, internally determined, frequency.

As shown in Figure 8, the simple prognostic models for $\tilde{q}$ are incapable of capturing such dynamics. Each absorbs the applied power directly into the mean equation but lacks any mechanism to transfer that energy to the evolving eddy field. Consequently, all three clearly overpredict the ensemble-mean energetics:

1. The linear response model (Equation 12) integrates the oscillatory forcing into a large, smooth rise of mean energy, tracking the envelope $h(t)$. Without nonlinear redistribution, the β-term merely drives westward-propagating barotropic and baroclinic Rossby packets that reflect at the western wall, amplifying the mean field still further.
2. The nonlinear response model (Equation 13), though containing self-advection, still lacks coupling to the underlying steady eddy stresses. It therefore traps the injected energy within the mean until perturbations reach finite amplitude, at which point new mesoscale instabilities emerge and the model spins up a spurious, self-sustained eddy field.
3. The steady-stress prognostic model (Equation 11) performs best. Because it includes the spatial structure of the statistically steady Reynolds-stress divergence, it can redistribute the injected anomalies along mean-flow pathways, exporting energy toward regions of enhanced dissipation.

While RMS comparisons of $\tilde{q}_{\mathrm{ref}}(=\langle q\rangle_{\mathrm{ref}}-\bar{q})$ and $\tilde{q}_{\mathrm{mod}}$ (Figure 8B) are roughly consistent with those for the energy, all three models yield very similar $r(t)$ curves that oscillate between positive and negative values ($\approx \pm 0.5$). As seen in Figure 7B, the directly computed ensemble responds to the periodic forcing at its own internal frequency. By contrast, the models lack a time-dependent mean *row*eddy pathway, and their responses are inherently locked to the forcing frequency, producing alternating spatial alignment of the model and reference ensemble-mean fields (Figure 9).

Case 2 highlights the qualitative change in the energy pathways when the external forcing acts at eddy scales. Even though the external work formally enters only in the ensemble-mean equation, in the reference ensemble solution that energy is rapidly and continuously transferred to the eddy field through nonlinear eddy–mean coupling. Forced perturbations of the ensemble mean remain weakly energetic, serving primarily as an intermediary through which energy is injected and immediately exported to the eddies. All three prognostic models fail to capture this redistribution: Forced to absorb the injected energy directly, they

10.3389/fmars.2026.1739607

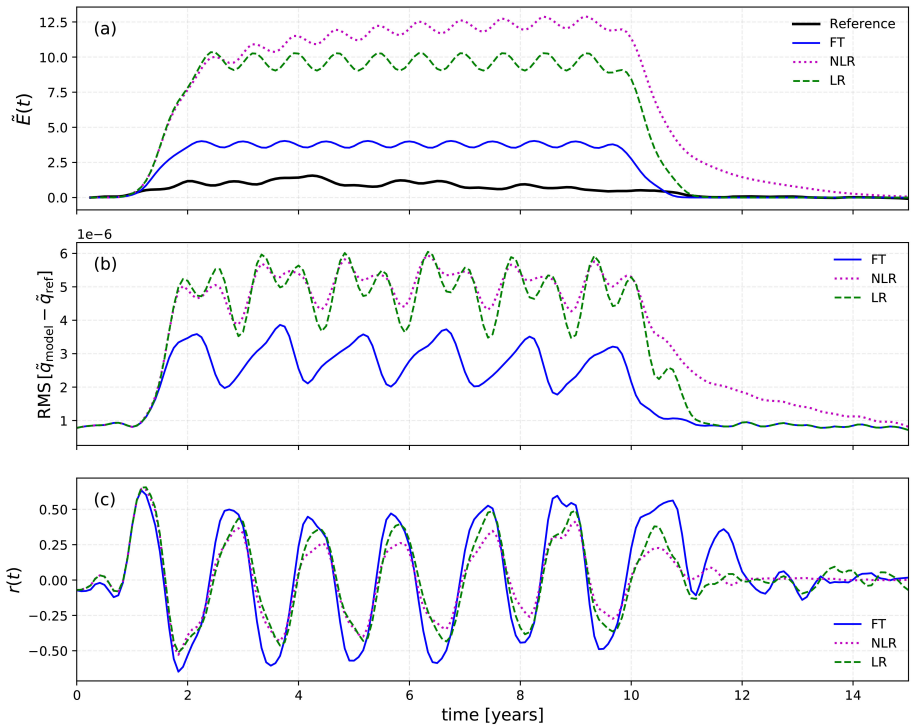

FIGURE 8
Case 2: **(A)** Mean anomaly energy $-\{\tilde{q}\tilde{\psi}\}(t)$ from the reference ensemble (black) and from the frozen-turbulence (FT, blue), nonlinear-response (NLR, magenta), and linear-response (LR, green) models. **(B)** Root-mean-square difference $RMS[\tilde{q}_{model} - \tilde{q}_{ref}]$ between the modeled and reference upper-layer anomaly fields. **(C)** Spatial pattern correlation $r(t)$ between modeled and reference upper layer anomalies.

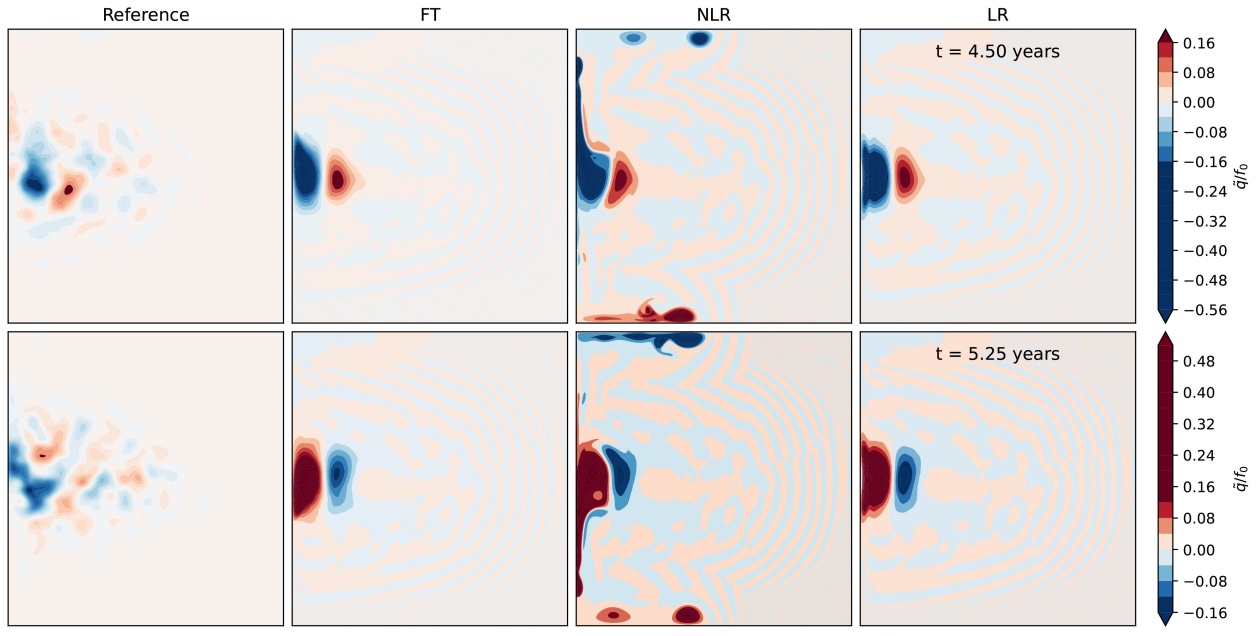

FIGURE 9
Case 2: Reference and modeled upper layer PV anomaly, $\tilde{q}_{ref} = \langle q \rangle_{ref} - \bar{q}$ and $\tilde{q}_{FT}, \tilde{q}_{LR}, \tilde{q}_{NLR}$ respectively, at opposite phases of forcing cycle (4.5 and 5.25 years). Colormaps and contour levels are identical across each row.

overpredict the amplitude of the mean response by large factors. The steady-stress model, by retaining the stationary eddy-stress divergence, is able to advect the perturbations away and limit their accumulation, but it still lacks the explicit feedback necessary to reproduce the observed eddy energy growth.

These results emphasize that the dominant pathway of adjustment at eddy scales is the continual mean $\rightarrow$eddy transfer, not storage within the mean circulation. Accurately capturing this behavior requires explicit representation of time-dependent eddy feedbacks, which are absent in all simplified closures tested here.

4 Conclusions and discussion

In studying parameterizations of oceanic turbulence, one must decide what is meant by the "mean" and by the "eddies" whose effects are to be represented. A common practice in coarse–grained modeling is to define these quantities through spatial or temporal filtering, leading to filtered–residual equations that depend explicitly on the filter scale. This procedure inevitably introduces ambiguities in how forcing terms appear in the mean and residual equations. For example, applying a spatial filter with length scale ℓ to an external forcing F produces both filtered and residual components, $\bar{F}^{\ell}$ and $F - \bar{F}^{\ell}$ respectively, implying that forced variability is injected directly into the residual flow.

By contrast, the ensemble formulation used here provides a mathematically consistent decomposition: because all ensemble members are identically forced, the external forcing enters only the ensemble–mean equation, $\langle F \rangle = F$, while the fluctuations evolve without direct forcing, i.e., $F^{\dagger} = 0$. In this precise sense, the ensemble mean is the only component of the system that is directly forced. The dynamical response to that forcing, however, need not be confined to the ensemble mean. Through instability and nonlinear interaction, the forced mean flow can transfer energy and potential vorticity to the fluctuating field, rendering the overall ensemble statistics intrinsically non–equilibrated (Pierini, 2020; Fedele et al., 2021).

We have explored these ideas in perhaps the simplest nonlinear "ocean" model—a two–layer QG double–gyre circulation—by constructing ensembles of identically forced simulations subject to episodic changes in the wind stress. Two experiments were designed to isolate contrasting regimes of mean–eddy interaction. In Case 1, a basin–scale modulation of the wind–stress amplitude produces a clear and intuitive response: the ensemble mean adjusts directly to the imposed change in forcing, and energy subsequently flows from the mean reservoir into the fluctuations as the system relaxes. In Case 2, the forcing varies at the observed eddy scales, patterned after the dominant SPOD mode of the turbulent state. In this regime, the ensemble–mean response is highly muted—the injected power is transferred almost immediately to the fluctuating fields, and the mean acts primarily as an intermediary in a continuous mean–to–eddy energy exchange. These contrasting cases emphasize that even in nominally forced–dissipative systems, ensemble statistics can remain far from equilibrium when the forcing varies in time or projects onto dynamically active scales.

To interpret these results, we compared three simple prognostic models for the ensemble mean: (i) a "steady–stress" model in which the instantaneous Reynolds stresses are replaced by their steady, time–mean values; (ii) a purely linear response model that neglects nonlinear and mean–flow interactions; and (iii) a nonlinear model that evolves with self–advection but omits the stabilizing influence of the background eddy field. The results demonstrate that, while the ensemble mean is the only directly forced component of the system, its evolution cannot in general be obtained by linear superposition of the forcing. The linear model reproduces integrated energy variations but fails to capture the spatial evolution of the mean fields. The nonlinear response model quickly develops its own chaotic dynamics, behaving more like an individual ensemble member than an ensemble mean. Only the frozen–turbulence model yields a bounded and physically plausible evolution: it performs reasonably well for the large–scale forcing of Case 1, but substantially overpredicts the ensemble–mean energy when the system is driven at eddy scales in Case 2. The absence of feedback between evolving mean and eddy statistics thus limits its applicability to regimes in which energy exchange is slow and effectively one–way.

Our findings on the ensemble mean reinforce the view that parameterizations based solely on steady or time–mean statistics will be inadequate in systems where eddy–mean exchange is intrinsically time dependent. In such non–equilibrium regimes, the mean flow and fluctuations co–evolve, and their mutual adjustment must be represented if the correct amplitude and phasing of the ensemble mean are to be captured. From a practical standpoint, our experiments provide an *a posteriori* test of closure ideas: the steady–stress approximation works surprisingly well when the forcing acts at large scales, but breaks down when variability is injected at the scales of the intrinsic turbulence. We argue that such *a posteriori* tests are essential for assessing closure robustness and avoiding overfitting parameterizations to specific configurations (cf. Uchida et al., 2025a).

Placed in a broader context, these results are consistent with earlier work emphasizing that the oceanic response to wind forcing depends sensitively on the structure of that forcing, rather than on its amplitude alone (Dewar, 2001; Berloff, 2005; Berloff et al., 2007; Penduff et al., 2018; Uchida et al., 2022b). The present study recasts this perspective in explicitly ensemble–mean terms: even when averaged over many realizations, the oceanic response reflects a continuing, dynamically active interaction between mean flow and eddies. The central lesson is therefore not that the ensemble mean isolates the full forced response, but that it retains only those aspects of the response that remain coherent across realizations.

Finally, we emphasize that our numerical experiments are highly idealized. The flat–bottom two–layer QG configuration omits bathymetry (Deremble et al., 2023; Sterl et al., 2025), vertical gradients of interior potential vorticity (Lobo et al., 2025), vertical velocity, and thermodynamics, all of which will modify the balance between forced and intrinsic variability in more realistic settings (Penduff et al., 2018; Griffies et al., 2015, 2025; Uchida et al., 2025b; Sun et al., 2025). Nevertheless, we expect the qualitative lessons identified here to carry over to primitive–equation models, as QG dynamics comprise an important component of the full

oceanic dynamics (Vallis, 2006; Deremble et al., 2023; Jamet et al., 2024). In particular, the dependence of ensemble–mean evolution on how energy is exchanged between mean and fluctuating fields appears to be a robust feature of forced, turbulent flows. Exploring these issues in stacked shallow–water and primitive–equation settings remains an important direction for future work.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

AP: Formal analysis, Methodology, Writing – original draft, Project administration, Data curation, Resources, Visualization, Investigation, Supervision, Writing – review & editing, Conceptualization, Validation, Funding acquisition, Software. TU: Writing – review & editing, Investigation, Funding acquisition, Writing – original draft. QJ: Writing – review & editing, Investigation. LS: Writing – review & editing, Investigation. TP: Funding acquisition, Investigation, Writing – review & editing. BD: Writing – review & editing, Investigation, Software. JS: Investigation, Writing – review & editing. MT: Investigation, Writing – review & editing, Formal analysis. NW: Writing – review & editing, Funding acquisition, Investigation. WD: Funding acquisition, Methodology, Writing – original draft, Conceptualization, Investigation.

Funding

The author(s) declared that financial support was received for this work and/or its publication. W. Dewar was supported through National Science Foundation (NSF) grants OCE-1829856, OCE-1941963, OCE-2123632, and OCE-2023585, and the French 'Make Our Planet Great Again' (MOPGA) program managed by the Agence Nationale de la Recherche under the Programme d'Investissement d'Avenir, reference ANR-18-MPGA-0002. This MOPGA program, led by W. Dewar and T. Penduff between mid 2019 and late 2024, funded the postdoctoral activities by Q. Jamet, T. Uchida and L. Sun at l'Institut des Géosciences de l'Environnement (IGE) and several stays of W. Dewar at IGE over this period. T. Uchida also acknowledges support from the NSF grant OCE-2123632 and NASA award 80NSSC24K1649 during his time in the U.S. and the Moscow Institute of Physics and Technology Development Program (Priority-2030) upon moving to Russia. A. Poje acknowledges support from the NSF grant OCE-2123633. L. Sun is now supported from the NSF grant OCE-2123632. This research was also supported, in part, by a grant of computer time from the City University of New York High Performance Computing Center under NSF Grants CNS-0855217, CNS-0958379 and ACI-1126113.

Acknowledgments

Computational resources for producing the QG simulations and analyzing the outputs were provided by the CUNY High Performance Computing Center. We would like to thank Edward Peirce and Kelly Hirai for maintaining the FSU cluster on which some of the data were also analyzed. This study grew out from the many years of fruitful and exciting discussions with the late William Kurt Dewar. He is dearly missed but we hope he is enjoying many pints of IPA, bike tours, the Scottish snare and some ultimate Frisbee in the afterlife.

Conflict of interest

Author JS was employed by Fluid Numerics.

The remaining author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Arbic, B., Muller, M., Richman, J., Shriver, J., Morten, A., Scott, R., et al. (2014). Geostrophic turbulence in the frequency-wavenumber domain: Eddy-driven low-frequency variability. *J. Phys. Oceanography* 44, 2050–2069. doi: 10.1175/JPO–D–13–054.1

Berloff, P. S. (2005). Random forcing of oceanic eddies. *J. Phys. Oceanography* 35, 1008–1024. doi: 10.1175/JPO2730.1

Berloff, P. S., Hogg, A. M. C., and Dewar, W. K. (2007). The turbulent oscillator: A mechanism of low-frequency variability of the wind-driven ocean gyres. *J. Phys. Oceanography* 37, 2363–22386. doi: 10.1175/JPO3118.1

Berloff, P. S., and McWilliams, J. C. (1999). Large-scale, low-frequency variability in wind-driven ocean gyres. *J. Phys. Oceanography* 29, 1925–1949. doi: 10.1175/1520-0485(1999)029<1925:LSLFVI>2.0.CO;2

Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proc. Natl. Acad. Sci.* 17, 656–660. doi: 10.1073/pnas.17.12.656

Cessi, P. (2008). An energy-constrained parameterization of eddy buoyancy flux. *J. Phys. oceanography* 38, 1807–1819. doi: 10.1175/2007JPO3812.1

Chen, R., and Flierl, G. R. (2015). The contribution of striations to the eddy energy budget and mixing: Diagnostic frameworks and results in a quasigeostrophic barotropic system with mean flow. *J. Phys. Oceanography* 45, 2095–2113. doi: 10.1175/JPO-D-14-0199.1

Deremble, B., Miller, L., and Albert, A. (2024). *qgw: Finite-differences QG model that relies on FFTW for the elliptic solver [Software]*.

Deremble, B., Uchida, T., Dewar, W. K., and Samelson, R. M. (2023). Eddy-mean flow interaction with a Multiple-Scale Quasi Geostrophic model. *J. Adv. Modeling Earth Syst.* 15, e2022MS003572. doi: 10.1029/2022MS003572

Dewar, W. K. (2001). On ocean dynamics in midlatitude climate. (AGU: Washington, D.C., USA) *J. Climate* 14, 4380–4397. doi: 10.1175/1520-0442(2001)014<4380:OODIMC>2.0.CO;2

Dewar, W. K. (2003). Nonlinear midlatitude ocean adjustment. *J. Phys. oceanography* 33, 1057–1082. doi: 10.1175/1520-0485(2003)033<1057:NMOA>2.0.CO;2

Eaves, R. E., Maddison, J. R., Marshall, D. P., and Waterman, S. (2025). An energy-and enstrophy-constrained parameterization of barotropic eddy potential vorticity fluxes. *J. Phys. Oceanography* 55, 573–591. doi: 10.1175/JPO-D-24-0027.1

Eden, C., and Greatbatch, R. J. (2008). Towards a mesoscale eddy closure. *Ocean Model.* 20, 223–239. doi: 10.1016/j.ocemod.2007.09.002

Farrell, B. F., and Ioannou, P. J. (2003). Structural stability of turbulent jets. *J. Atmospheric Sci.* 60, 2101–2118. doi: 10.1175/1520-0469(2003)060<2101:SSOTJ>2.0.CO;2

Fedele, G., Penduff, T., Pierini, S., Alvarez-Castro, M. C., Bellucci, A., and Masina, S. (2021). Interannual to decadal variability of the kuroshio extension: analyzing an ensemble of global hindcasts from a dynamical system viewpoint. *Climate Dynamics* 57, 975–992. doi: 10.1007/s00382-021-05751-7

Ferrari, R., Griffies, S. M., Nurser, A. G., and Vallis, G. K. (2010). A boundary-value problem for the parameterized mesoscale eddy transport. *Ocean Model.* 32, 143–156. doi: 10.1016/j.ocemod.2010.01.004

Frisch, U. (1995). *Turbulence: the legacy of A. N. Kolmogorov* (Cambridge: Cambridge University Press).

Galanti, B., and Tsinober, A. (2004). Is turbulence ergodic? *Phys. Lett. A* 330, 173–180. doi: 10.1016/j.physleta.2004.07.009

Gent, P. R., and McWilliams, J. C. (1990). Isopycnal mixing in ocean circulation models. *J. Phys. Oceanography* 20, 150–155. doi: 10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2

Gent, P. R., Willebrand, J., McDougall, T. J., and McWilliams, J. C. (1995). Parameterizing eddy-induced tracer transports in ocean circulation models. *J. Phys. oceanography* 25, 463–474. doi: 10.1175/1520-0485(1995)025<0463:PEITTI>2.0.CO;2

Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., et al. (2002). Advanced spectral methods for climatic time series. *Rev. Geophysics* 40, 3–1–3–41. doi: 10.1029/2000RG000092

Griffies, S. M., Adcroft, A., Beadling, R. L., Bushuk, M., Chang, C.-Y., Drake, H. F., et al. (2025). The GFDL-CM4X climate model hierarchy, Part I: model description and thermal properties. *J. Adv. Modeling Earth Syst.* 17, e2024MS004861. doi: 10.1029/2024MS004861

Griffies, S. M., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C. O., et al. (2015). Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *J. Climate* 28, 952–977. doi: 10.1175/JCLI-D-14-00353.1

Grooms, I., Majda, A. J., and Smith, K. S. (2015). Stochastic superparameterization in a quasigeostrophic model of the Antarctic Circumpolar Current. *Ocean Model.* 85, 1–15. doi: 10.1016/j.ocemod.2014.10.001

Gu, Q., Gervais, M., Danabasoglu, G., Kim, W. M., Castruccio, F., Maroon, E., et al. (2024). Wide range of possible trajectories of North Atlantic climate in a warming world. *Nat. Commun.* 15, 4221. doi: 10.1038/s41467-024-48401-2

Hallberg, R. (2013). Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects. *Ocean Model.* 72, 92–103. doi: 10.1016/j.ocemod.2013.08.007

Jamet, Q., Berger, A., Deremble, B., and Penduff, T. (2024). Thermodynamical effects of ocean current feedback in a quasigeostrophic coupled model. *J. Phys. Oceanography* 54, 1691–1704. doi: 10.1175/JPO-D-23-0159.1

Jamet, Q., Leroux, S., Dewar, W. K., Penduff, T., Le Sommer, J., Molines, J.-M., et al. (2022). Non651 local eddy-mean kinetic energy transfers in submesoscale-permitting ensemble simulations. *J. Adv. Modeling Earth Syst.* 14, e2022MS003057. doi: 10.1029/2022MS003057

Kraichnan, R. H. (1976). Eddy viscosity in two and three dimensions. *J. Atmospheric Sci.* 33, 1521–1536. doi: 10.1175/1520-0469(1976)033<1521:EVITAT>2.0.CO;2

Leroux, S., Penduff, T., Bessières, L., Molines, J.-M., Brankart, J.-M., Sérazin, G., et al. (2018). Intrinsic and atmospherically forced variability of the AMOC: Insights from a large-ensemble ocean hindcast. *J. Climate* 31, 1183–1203. doi: 10.1175/JCLI-D-17-0168.1

Lobo, M., Griffies, S. M., and Zhang, W. (2025). Vertical structure of baroclinic instability in a three-layer quasigeostrophic model over a sloping bottom. *J. Phys. Oceanography* 55, 341–359. doi: 10.1175/JPO-D-24-0130.1

Maddison, J. R., and Marshall, D. P. (2013). The Eliassen–Palm flux tensor. *J. Fluid Mechanics* 729, 69–102. doi: 10.1017/jfm.2013.259

Mak, J., Maddison, J. R., Marshall, D. P., Ruan, X., Wang, Y., and Yeow, L. (2023). Scale-awareness in an eddy energy constrained mesoscale eddy parameterization. *J. Adv. Modeling Earth Syst.* 15, e2023MS003886. doi: 10.1029/2023MS003886

Mak, J., Marshall, D. P., Maddison, J. R., and Bachman, S. D. (2017). Emergent eddy saturation from an energy constrained eddy parameterisation. *Ocean Model.* 112, 125–138. doi: 10.1016/j.ocemod.2017.02.007

Marshall, D. P., Maddison, J. R., and Berloff, P. S. (2012). A framework for parameterizing eddy potential vorticity fluxes. *J. Phys. Oceanography* 42, 539–557. doi: 10.1175/JPO-D-11-048.1

Marston, J. B., Chini, G. P., and Tobias, S. M. (2016). Generalized quasilinear approximation: Application to zonal jets. *Phys. Rev. Lett.* 116, 214501. doi: 10.1103/PhysRevLett.116.214501

McWilliams, J. C. (2006). *Fundamentals of geophysical fluid dynamics* (Cambridge: Cambridge University Press).

Narinc, O., Penduff, T., Maze, G., Leroux, S., and Molines, J.-M. (2024). North Atlantic Subtropical Mode Water properties: intrinsic and atmospherically forced interannual variability. *Ocean Sci.* 20, 1351–1365. doi: 10.5194/os-20-1351-2024

Penduff, T., Sérazin, G., Leroux, S., Close, S., Molines, J.-M., Barnier, B., et al. (2018). Chaotic variability of ocean heat content: Climate-relevant features and observational implications. *Oceanography* 31, 63–71. doi: 10.5670/oceanog.2018.210

Pierini, S. (2020). Statistical significance of small ensembles of simulations and detection of the internal climate variability: An excitable ocean system case study. *J. Stat. Phys.* 179, 1475–1495. doi: 10.1007/s10955-019-02409-x

Romanou, A., Rind, D., Jonas, J., Miller, R., Kelley, M., Russell, G., et al. (2023). Stochastic bifurcation of the North Atlantic Circulation under a mid-range future climate scenario with the NASA-GISS modelE. *J. Climate*, 36, 1–49. doi: 10.1175/JCLI-D-22-0536.1

Sane, A., Fox-Kemper, B., and Ullman, D. S. (2024). Internal versus forced variability metrics for general circulation models using information theory. *J. Geophysical Research: Oceans* 129, e2023JC020101. doi: 10.1029/2023JC020101

Sérazin, G., Jaymond, A., Leroux, S., Penduff, T., Bessières, L., Llovel, W., et al. (2017). A global probabilistic study of the ocean heat content low-frequency variability: Atmospheric forcing versus oceanic chaos. *Geophysical Res. Lett.* 44, 5580–5589. doi: 10.1002/2017GL073026

Serazin, G., Penduff, T., Gregorio, S., Barnier, B., Molines, J., and Terray, L. (2015). Intrinsic variability of sea level from global ocean simulations: Spatiotemporal scales. *J. Climate* 10, 4279–4292. doi: 10.1175/JCLI-D-14-00554.1

Simonnet, E. (2005). Quantization of the low-frequency variability of the double-gyre circulation. *J. Phys. Oceanography* 35, 2268–2290. doi: 10.1175/JPO2806.1

Simonnet, E., Ghil, M., Ide, K., Temam, R., and Wang, S. (2003). Low-frequency variability in shallow-water models of the wind-driven ocean circulation. Part II: Time-dependent solutions. *J. Phys. oceanography* 33, 729–752. doi: 10.1175/1520-0485(2003)33<729:LVISMO>2.0.CO;2

Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. @ the basic experiment. *Monthly weather Rev.* 91, 99–164. doi: 10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2

Sterl, M. F., Palóczy, A., Groeskamp, S., Baatsen, M. L., LaCasce, J. H., and Isachsen, P. E. (2025). The joint effects of planetary, topography and friction on baroclinic instability in a two-layer quasi-geostrophic model. *J. Fluid Mechanics* 1012, A1. doi: 10.1017/jfm.2025.10172

Sun, L., Uchida, T., Penduff, T., Dewar, W. K., Deremble, B., Poje, A. C., et al. (2025). *On the dynamics of the subtropical mode water from an ensemble view* (Authorea Preprints: Washington, D.C., USA). doi: 10.22541/essoar.174802928.80389595/v1

Takasuka, D., Suematsu, T., Miura, H., and Nakano, M. (2025). *Propagation of the Madden-Julian oscillation as a deterministic chaotic phenomenon* (ArXiv: Ithaca, USA). doi: 10.48550/arXiv.2506.23195

Taylor, G. I. (1938). The spectrum of turbulence. Proceedings of the Royal Society of London. *Ser. A-Mathematical Phys. Sci.* 164, 476–490. doi: 10.1098/rspa.1938.0032

Towne, A., Meidner, M., and Schmid, P. J. (2018). Spectral proper orthogonal decomposition of turbulent flows. *J. Fluid Mechanics* 847, 821–857. doi: 10.1017/jfm.2018.328

Uchida, T., Balwada, D., Jamet, Q., Dewar, W. K., Deremble, B., Penduff, T., et al. (2023). Cautionary tales from the mesoscale eddy transport tensor. *Ocean Model.* 182, 102172. doi: 10.1016/j.ocemod.2023.102172

Uchida, T., Bodner, A. S., Reichl, B. G., Adcroft, A., Fox-Kemper, B., Ilicak, M., et al. (2025a). Representation of surface mixed-layer eddies affects the large-scale ventilation of the global ocean. *Geophysical Res. Lett.* 53, e2025GL116872. doi: 10.1029/2025GL116872

Uchida, T., Deremble, B., and Penduff, T. (2021). The seasonal variability of the ocean energy cycle from a quasi-geostrophic double gyre ensemble. *Fluids* 6, 206. doi: 10.3390/fluids6060206

Uchida, T., Deremble, B., and Popinet, S. (2022a). Deterministic model of the eddy dynamics for a midlatitude ocean model. *J. Phys. Oceanography* 52, 1133–1154. doi: 10.1175/725JPO-D-21-0217.1

Uchida, T., Jamet, Q., Dewar, W. K., Balwada, D., Le Sommer, J., and Penduff, T. (2022b). Diagnosing the thickness-weighted averaged eddy-mean flow interaction in an eddying North Atlantic ensemble: The Eliassen–Palm flux. *J. Adv. Modeling Earth Syst.* 14, e2021MS002866. doi: 10.1029/2021MS002866

Uchida, T., Jamet, Q., Dewar, W. K., Deremble, B., Poje, A. C., and Sun, L. (2024a). Imprint of chaos on the ocean energy cycle from an eddying North Atlantic ensemble. *J. Phys. Oceanography* 54, 679–696. doi: 10.1175/JPO-D-23-0176.1

Uchida, T., Jamet, Q., Poje, A. C., Wienders, N., and Dewar, W. K. (2024b). Wavelet-based wavenumber spectral estimate of eddy kinetic energy: Application to the North Atlantic. *Ocean Model.* 190, 102392. doi: 10.1016/j.ocemod.2024.102392

Uchida, T., Jamet, Q., Poje, A. C., Wienders, N., Sun, L., and Dewar, W. K. (2025b). Dynamics and thermodynamics of the Boussinesq North Atlantic eddy kinetic energy spectral budget. *J. Adv. Modeling Earth Syst.* 17, e2024MS004781. doi: 10.1029/2024MS004781

Vallis, G. (2006). *Atmospheric and oceanic fluid dynamics* (Cambridge University Press, Cambridge, UK).

Veronis, G. (1963). An analysis of wind-driven ocean circulation with a limited number of Fourier components. *J. Atmospheric Sci.* 20, 577–593. doi: 10.1175/1520-0469(1963)020<0577:AAOWDO>2.0.CO;2

Wei, H., Wang, Y., and Mak, J. (2024). Parameterizing eddy buoyancy fluxes across prograde shelf/slope fronts using a slope-aware Geometric closure. *J. Phys. Oceanography* 54, 359–377. doi: 10.1175/JPO-D-23-0152.1

Young, W. (2012). An exact thickness-weighted average formulation of the Boussinesq equations. *J. Phys. Oceanography* 42, 692–707. doi: 10.1175/JPO-D-11-0102.1

Zhao, M., Ponte, R. M., Penduff, T., Close, S., Llovel, W., and Molines, J.-M. (2021). Imprints of ocean chaotic intrinsic variability on bottom pressure and implications for data and model analyses. *Geophysical Res. Lett.* 48, e2021GL096341. doi: 10.1029/2021GL096341

Appendix A: Resolution dependence

To assess sensitivity to spatial resolution, we compare simulations at 385^2, 513^2, and 1025^2 grid points, corresponding to grid spacings $\Delta x \approx 10.0$, 7.5, and 3.75 km, respectively, with all other parameters held fixed.

Figure 10 (top panel) shows the evolution of total domain–integrated energy for the three resolutions. After an initial spin–up period, the total energy equilibrates to statistically steady levels that differ by less than $\mathcal{O}(10\,\%)$ instantaneously across resolutions, with comparable temporal variability. No systematic drift with resolution is observed over the 75–year averaging window shown. The lower panels of Figure 10 show the upper–layer potential vorticity mean (25–100 yr) and instantaneous fluctuations at t = 50 yr. The large–scale gyre structure and mean jet position are visually indistinguishable across resolutions. The deformation–scale and larger interior dynamics that dominate the ensemble–mean statistics are robust to resolution changes over this range.

In Case 2, the time dependent forcing directly drives relatively small spatial scales on sub-annual temporal scales. The ability of the

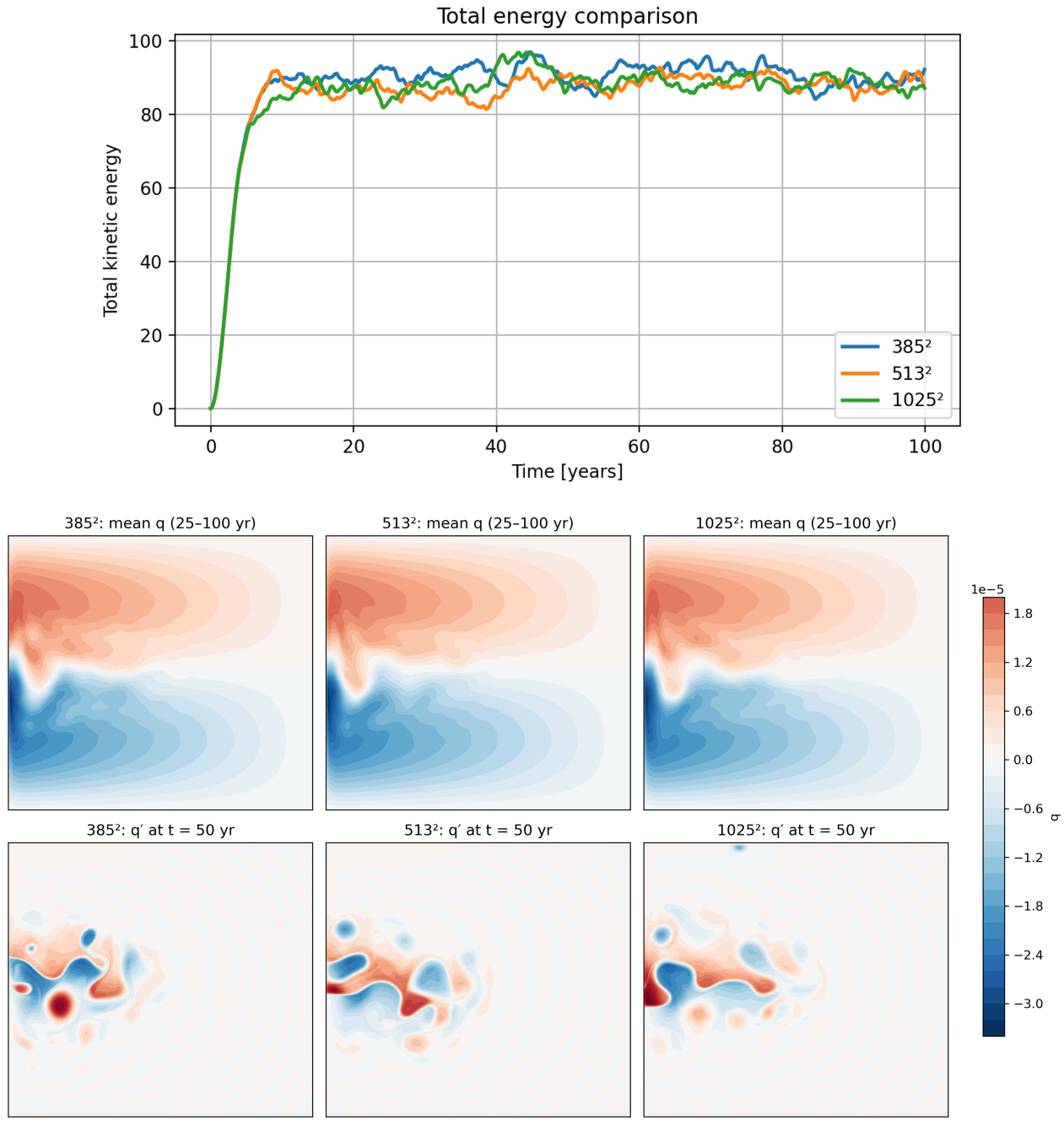

FIGURE 10
Resolution dependence of the steady wind–driven double–gyre simulations. (*Top*) Time series of total domain–integrated energy for 385^2, 513^2, and 1025^2 grid resolutions. After spin–up, all resolutions equilibrate to comparable statistically steady energy levels with similar temporal variability. (Bottom) Upper–layer potential vorticity: 25–100 yr time means (top row) and instantaneous fluctuations at t = 50 yr (bottom row) for the same three resolutions.

computational model to resolve this forcing can be tested in the context of the nonlinear response model that requires solution of the fully nonlinear QG system initialized from rest and driven by the perturbation forcing alone (i.e., the difference between the full forcing and the steady mean forcing). All other parameters identical to the reference configuration. The frozen–turbulence, linear–response, and nonlinear–response models all employ the same numerical discretization, grid, time stepping, and boundary conditions as the reference QG system. The linear and nonlinear response models and the frozen turbulence models differ only in their treatment of the Reynolds stress contributions and introduce no additional numerical parameters.

As shown in Figure 11, the numerics of the nonlinear response is robust to spatial resolution. The maximum difference in total kinetic energy between the 385^2 and 1025^2 simulations relative to the 513^2 reference remains below 1 % of the peak energy at all times. Resolution effects are confined to small–scale variance and do not modify the large–scale structure or timing of the response, confirming that the dynamically important scales excited by the Case 2 forcing are resolved.

Appendix B: spectral proper orthogonal decomposition (SPOD) with QG energy norm

The analysis of the multi-layer quasi-geostrophic (QG) model's spatio-temporal dynamics was performed using a modified version of the Spectral Proper Orthogonal Decomposition (SPOD) procedure. This approach, while following the general framework of SPOD, was adapted to incorporate a physically relevant energy norm. This ensures that the resulting modes and their ranking are directly tied to the energetics of the system, providing a physically meaningful decomposition.

General SPOD procedure

The SPOD method, as formulated by []?, provides a rigorous framework for identifying and ranking statistically stationary coherent structures by frequency. The procedure is based on the method of snapshots, but instead of analyzing instantaneous fields,

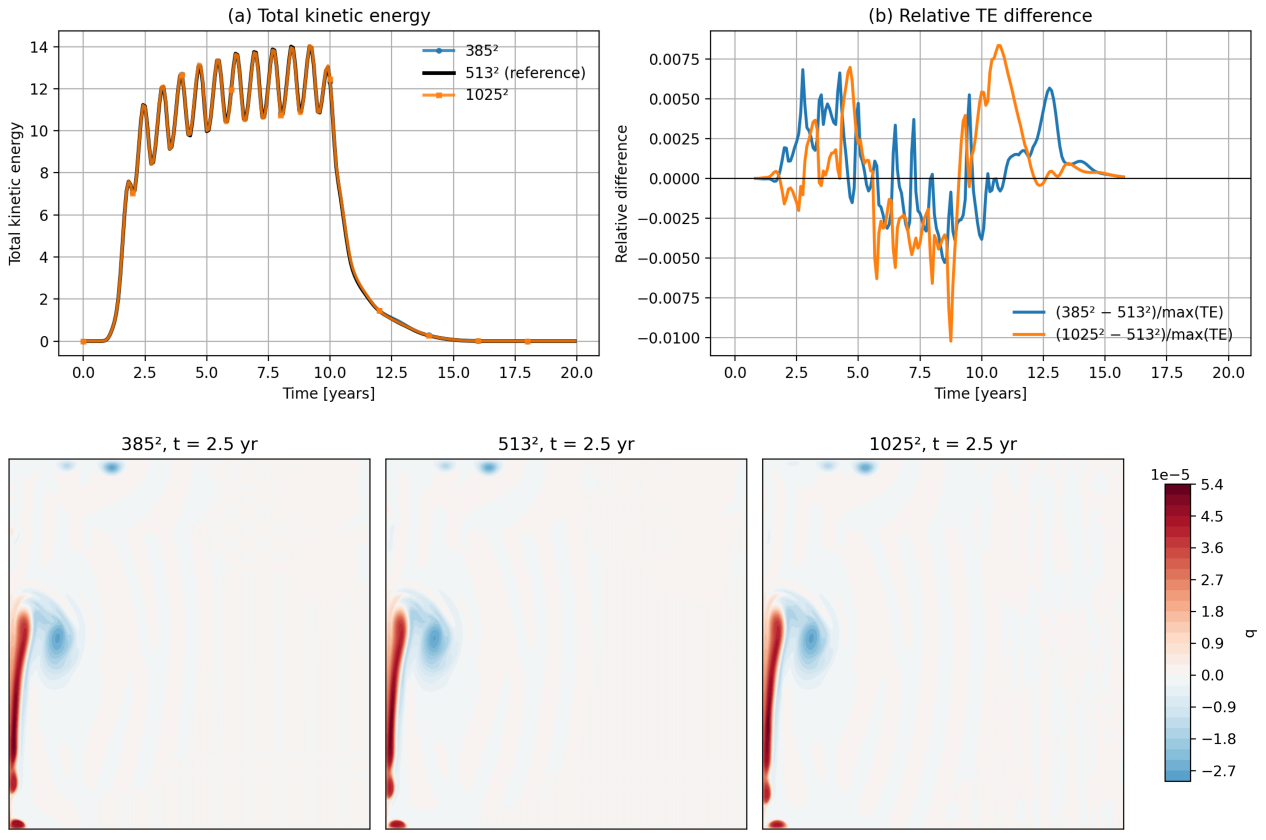

FIGURE 11
Resolution robustness of the nonlinear response model for Case 2 forcing. (*Top left*) Total ensemble mean perturbation energy evolution for three spatial resolutions. (*Top right*) Relative kinetic energy differences with respect to the 513^2 reference simulation. (*Bottom*) Instantaneous upper layer q' (x, y) fields at t = 2.5 years for Δx = 10, 7.5 and 3.75 km.

it operates on the Fourier-transformed data. The core of the method is the computation and eigendecomposition of the cross-spectral density (CSD) matrix, which captures the average two-point correlations in the frequency domain.

Our implementation, adapted for the multi-layer quasi-geostrophic (QG) model, follows these steps:

1. Data Segmentation: A time series of N snapshots, represented by the state vector $u(x,t)$, is segmented into N_b overlapping blocks. Each block, $u_k(x,t)$, contains N_t snapshots. This blocking allows for the use of a time-averaging ensemble, crucial for statistical convergence. In our two-layer model with $N_s = N_x \times N_y$ spatial points, the state vector at each time step, $u(t)$, is a column vector of $4N_s$ points, specifically containing the values of the potential vorticity q_i and streamfunction ψ_i for both layers:

$$u(t) = [q_1(t), \psi_1(t), q_2(t), \psi_2(t)]^T$$

where the fields are stacked in column-vector form.

2 Windowing: A windowing function, such as a Hanning window, is applied to each block in time to minimize spectral leakage and improve the frequency resolution.

3 Fourier Transform: A discrete Fourier transform (DFT) is applied to each windowed block to transform the data into the frequency domain. This yields a set of Fourier-transformed snapshots, $\hat{u}_k(x,\omega)$, for each block k and frequency ω.

The Physically-based energy norm and weighting matrix L

A critical deviation from the standard SPOD formulation is our use of a basin-averaged energy norm to define the inner product. This is essential for ensuring that the resulting modes are orthonormal with respect to the system's energy. In the context of our multi-layer QG model, the total energy of the system is the sum of the kinetic and potential energies, which can be expressed in terms of the streamfunction ψ_i and potential vorticity q_i for each layer i:

$$E = \sum_i \frac{1}{A} \iint_A -\{\psi_i q_i\} dA$$

To incorporate this energy norm into the discrete SPOD procedure, we define a weighting matrix, L, such that the inner product between two state vectors, u and v, is given by $\langle u, v \rangle = u^* Lv$. The matrix L is a block-diagonal matrix whose structure is determined by the energy equation. For our two-layer, uniform grid system, L is a $4N_s \times 4N_s$ matrix:

$$L = \begin{bmatrix} 0 & -\Delta A I & 0 & 0 \\ -\Delta A I & 0 & 0 & 0 \\ 0 & 0 & 0 & -\Delta A I \\ 0 & 0 & -\Delta A I & 0 \end{bmatrix}$$

Here, ΔA is the constant area of a single grid cell and I is the $N_s \times N_s$ identity matrix. This matrix ensures that the SPOD analysis identifies and ranks modes based on their contribution to the total energy of the system, rather than on a generic mathematical norm.

CSD matrix and eigenvalue decomposition

4 CSD Matrix Construction: For each frequency ω, the CSD matrix, $S(\omega)$, is computed as an ensemble average of the outer product of the Fourier-transformed data from all blocks, weighted by the matrix L:

$$S(\omega) = \frac{1}{N_b} \sum_{k=1}^{N_b} \hat{u}_k(x,\omega) L \hat{u}_k^*(x,\omega)$$

This explicit form of the CSD matrix calculation directly incorporates our energy norm.

5 Eigenvalue Decomposition: The CSD matrix is then decomposed into its eigenvalues and eigenvectors:

$$\mathbf{S}(\omega)\phi_j(x,\omega) = \lambda_j(\omega)\phi_j(x,\omega)$$

The eigenvectors, $\phi_j(x,\omega)$, are the SPOD modes—spatially coherent structures that oscillate at frequency ω. The corresponding eigenvalues, $\lambda_j(\omega)$, represent the energy of each mode at that frequency. By ranking the eigenvalues, we can identify the most energetic and dynamically significant structures in the system. The time-dependence of the modes is harmonic, given by $\Phi_j(\mathbf{x}, t) = \mathrm{Re}[\phi_j(\mathbf{x}, \omega) e^{i\omega t}]$.

Appendix B: spectral proper orthogonal decomposition (SPOD) with QG energy norm

The analysis of the multi-layer quasi-geostrophic (QG) model's spatio-temporal dynamics was performed using a modified version of the Spectral Proper Orthogonal Decomposition (SPOD) procedure. This approach, while following the general framework of SPOD, was adapted to incorporate a physically relevant energy norm. This ensures that the resulting modes and their ranking are directly tied to the energetics of the system, providing a physically meaningful decomposition.