

Data-Driven Stochastic Parameterization for Unresolved Ocean Dynamics

Workshop Machine Learning + Simulation

Long Li

Institut de Physique de Nice, CNRS | CALISTO, Inria Côte d'Azur

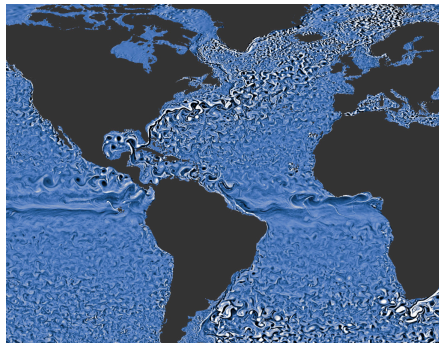
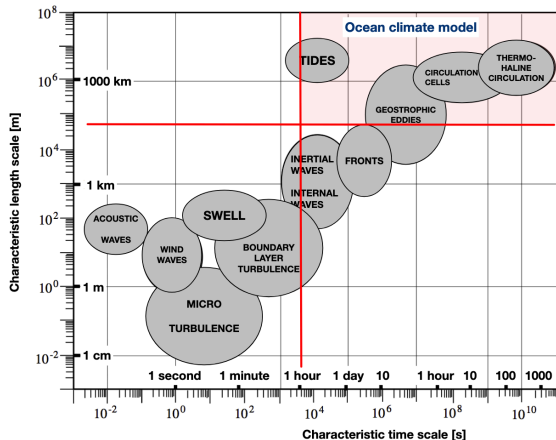
Joint work with F. Tucciarone (AIRSEA, Inria) and E. Mémin (ODYSSEY, Inria)

13 October 2026 | Laboratoire J.-A. Dieudonné, Nice

1. Introduction

2. Stochastic transport modeling
3. Data-driven calibration
4. Numerical validation in NEMO
5. Conclusion & perspectives

The ocean is a multiscale system

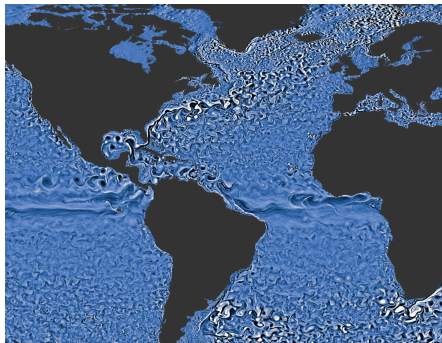
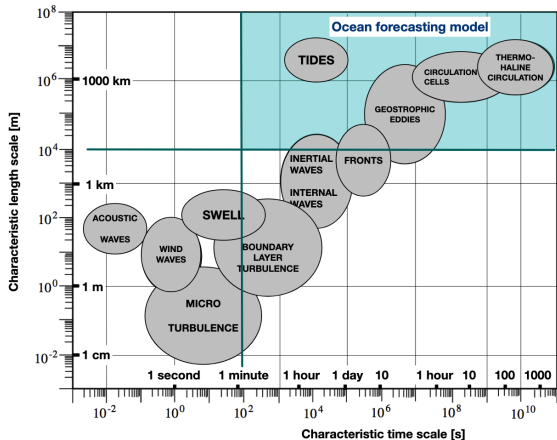


Surface relative vorticity, North Atlantic

- Processes span 10^{-3} to 10^7 m and seconds to millennia, **coupled across scales**
- A **climate** model resolves only the top-right corner: everything below ~ 50 km is lost

Whatever the grid, the unresolved scales must be **homogenized** $\Rightarrow$ **parameterized**

The ocean is a multiscale system

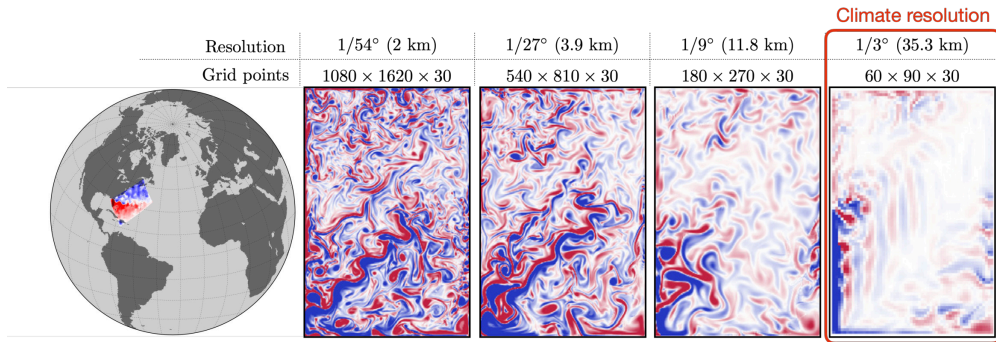


Surface relative vorticity, North Atlantic

- Processes span 10^{-3} to 10^7 m and seconds to millennia, **coupled across scales**
- A **climate** model resolves only the top-right corner: everything below ~ 50 km is lost
- A **forecasting** model pushes the cut to ~ 10 km $\rightarrow$ the window moves, the problem does not

Whatever the grid, the unresolved scales must be **homogenized** $\Rightarrow$ **parameterized**

Mesoscale eddies: the missing scale



The same NEMO double-gyre flow seen at four resolutions

Why they matter

- Carry most of the ocean's kinetic energy
- Set the **jets**, the transport of heat and tracers, and the intrinsic variability

Why we cannot simply resolve them

- Halving Δx costs $\sim 16\times$ more
- At climate resolution the eddy field is **gone**, not merely blurred

→ The eddies must be **modelled** — and their effect is *not* purely dissipative

Parameterizing the unresolved eddies?

Deterministic closures

e.g., eddy-induced diffusion

[Redi, 1982] & advection

[Gent and McWilliams, 1990]

- + robust, energetically sound
- mostly dissipative, no backscatter, too little variability

Parameterizing the unresolved eddies?

Deterministic closures

e.g., eddy-induced diffusion
[Redi, 1982] & advection
[Gent and McWilliams, 1990]

- + robust, energetically sound
- mostly dissipative, no backscatter, too little variability

Stochastic closures

e.g., random forcing
[Hasselmann, 1976], perturbed
tendencies [Palmer, 2019]

- + **distribution**, variability, UQ
- often added *a posteriori*, may break conservation / stability

Parameterizing the unresolved eddies?

Deterministic closures

e.g., eddy-induced diffusion
[Redi, 1982] & advection
[Gent and McWilliams, 1990]

- + robust, energetically sound
- mostly dissipative, no backscatter, too little variability

Stochastic closures

e.g., random forcing
[Hasselmann, 1976], perturbed
tendencies [Palmer, 2019]

- + **distribution**, variability, UQ
- often added *a posteriori*, may break conservation / stability

Learned closures

from hybrid-ML to full emulators
e.g., [Bolton and Zanna, 2019,
Guillaumin and Zanna, 2021]

- + expressive, learned from data
- conservation, on-line stability, generalization, interpretability

Parameterizing the unresolved eddies?

Deterministic closures

e.g., eddy-induced diffusion
[Redi, 1982] & advection
[Gent and McWilliams, 1990]

- + robust, energetically sound
- mostly dissipative, no backscatter, too little variability

Stochastic closures

e.g., random forcing
[Hasselmann, 1976], perturbed
tendencies [Palmer, 2019]

- + **distribution**, variability, UQ
- often added *a posteriori*, may break conservation / stability

Learned closures

from hybrid-ML to full emulators
e.g., [Bolton and Zanna, 2019,
Guillaumin and Zanna, 2021]

- + expressive, learned from data
- conservation, on-line stability, generalization, interpretability

Our route: physically consistent and data-driven

Derive the stochastic model *from the conservation laws*, so that energy consistency holds by construction; then **learn** the unresolved statistics from high-resolution data or from observations.

ERC Synergy Grant STUOD
Stochastic Transport in Upper Ocean Dynamics



Imperial College
London



1. Introduction
- 2. Stochastic transport modeling**
3. Data-driven calibration
4. Numerical validation in NEMO
5. Conclusion & perspectives

Stochastic transport [Bauer et al., 2020a]

1. The flow map becomes random

$$d\mathbf{X}_t = \underbrace{\mathbf{v}(\mathbf{X}_t, t) dt}_{\text{resolved}} + \underbrace{d\mathbf{W}_t(\mathbf{X}_t)}_{\text{unresolved}}$$



Stochastic transport [Bauer et al., 2020a]

1. The flow map becomes random

$$d\mathbf{X}_t = \underbrace{\mathbf{v}(\mathbf{X}_t, t) dt}_{\text{resolved}} + \underbrace{d\mathbf{W}_t(\mathbf{X}_t)}_{\text{unresolved}}$$



2. The noise: modes $\times$ Brownian motions

$$d\mathbf{W}_t(\mathbf{x}) = \sum_n \phi_n(\mathbf{x}) d\beta_t^n$$

$$\boxed{\text{wavy}}_{\phi_n} \times \mathcal{N} \boxed{\text{wavy}}_{\phi_n} \times \mathcal{N} \boxed{\text{wavy}}_{\phi_n} \times \mathcal{N} \dots$$

$\phi_n \quad d\beta^n \quad \phi_n \quad d\beta^n \quad \phi_n \quad d\beta^n$

Stochastic transport [Bauer et al., 2020a]

1. The flow map becomes random

$$d\mathbf{X}_t = \underbrace{\mathbf{v}(\mathbf{X}_t, t) dt}_{\text{resolved}} + \underbrace{d\mathbf{W}_t(\mathbf{X}_t)}_{\text{unresolved}}$$

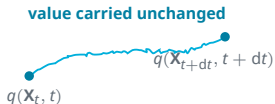


2. The noise: modes $\times$ Brownian motions

$$d\mathbf{W}_t(\mathbf{x}) = \sum_n \phi_n(\mathbf{x}) d\beta_t^n$$

$$\boxed{\text{wavy lines}}_{\phi_n} \times \mathcal{N}_{d\beta^n} \quad \boxed{\text{wavy lines}}_{\phi_n} \times \mathcal{N}_{d\beta^n} \quad \boxed{\text{wavy lines}}_{\phi_n} \times \mathcal{N}_{d\beta^n} \quad \dots$$

3. A quantity carried by the random flow



$$q(\mathbf{X}_{t+dt}, t+dt) = q(\mathbf{X}_t, t) \xrightarrow{\text{It\^o-Wentzell}} \text{Eulerian form}$$

Stochastic transport [Bauer et al., 2020a]

1. The flow map becomes random

$$d\mathbf{X}_t = \underbrace{\mathbf{v}(\mathbf{X}_t, t) dt}_{\text{resolved}} + \underbrace{d\mathbf{W}_t(\mathbf{X}_t)}_{\text{unresolved}}$$



2. The noise: modes $\times$ Brownian motions

$$d\mathbf{W}_t(\mathbf{x}) = \sum_n \phi_n(\mathbf{x}) d\beta_t^n$$

$$\boxed{\text{wavy lines}}_{\phi_n} \times \mathcal{N}_{d\beta^n} \quad \boxed{\text{wavy lines}}_{\phi_n} \times \mathcal{N}_{d\beta^n} \quad \boxed{\text{wavy lines}}_{\phi_n} \times \mathcal{N}_{d\beta^n} \quad \dots$$

3. A quantity carried by the random flow

value carried unchanged



$$q(\mathbf{X}_{t+dt}, t+dt) = q(\mathbf{X}_t, t) \xrightarrow{\text{It\^o-Wentzell}} \text{Eulerian form}$$

Stochastic transport operator

$$\mathbb{D}_t q := dq + \underbrace{(\mathbf{v}^* dt + d\mathbf{W}_t) \cdot \nabla q}_{\text{backscatter}} - \underbrace{\nabla \cdot (\mathbf{a} \nabla q)}_{\text{dissipation}} dt = 0$$

- **advection by the noise** $\rightarrow$ energy backscatter
- **inhomogeneous diffusivity:** $\mathbf{a} = \frac{1}{2} \mathbb{E}[d\mathbf{W}_t d\mathbf{W}_t^T] / dt$

Stochastic transport [Bauer et al., 2020a]

1. The flow map becomes random

$$d\mathbf{X}_t = \underbrace{\mathbf{v}(\mathbf{X}_t, t) dt}_{\text{resolved}} + \underbrace{d\mathbf{W}_t(\mathbf{X}_t)}_{\text{unresolved}}$$

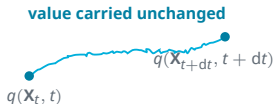


2. The noise: modes $\times$ Brownian motions

$$d\mathbf{W}_t(\mathbf{x}) = \sum_n \phi_n(\mathbf{x}) d\beta_t^n$$

$$\underbrace{\left[\begin{array}{|c|} \hline \text{wavy lines} \\ \hline \end{array} \right]}_{\phi_n} \times \underbrace{\mathcal{N}}_{d\beta^n} \underbrace{\left[\begin{array}{|c|} \hline \text{wavy lines} \\ \hline \end{array} \right]}_{\phi_n} \times \underbrace{\mathcal{N}}_{d\beta^n} \underbrace{\left[\begin{array}{|c|} \hline \text{wavy lines} \\ \hline \end{array} \right]}_{\phi_n} \times \underbrace{\mathcal{N}}_{d\beta^n} \dots$$

3. A quantity carried by the random flow



$$q(\mathbf{X}_{t+dt}, t+dt) = q(\mathbf{X}_t, t) \xrightarrow{\text{It\^o-Wentzell}} \text{Eulerian form}$$

Stochastic transport operator

$$\mathbb{D}_t q := dq + \underbrace{(\mathbf{v}^* dt + d\mathbf{W}_t) \cdot \nabla q}_{\text{backscatter}} - \underbrace{\nabla \cdot (\mathbf{a} \nabla q)}_{\text{dissipation}} dt = 0$$

- **advection by the noise** $\rightarrow$ energy backscatter
- **inhomogeneous diffusivity**: $\mathbf{a} = \frac{1}{2} \mathbb{E}[d\mathbf{W}_t d\mathbf{W}_t^T] / dt$

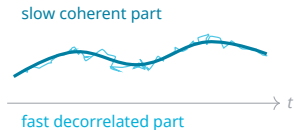
Structure for free

- ✓ fluctuation and dissipation are **exactly balanced**, so energy is conserved pathwise [Li et al., 2023a]
- ✓ solutions **converge** to the deterministic model as the noise vanishes, so the closure is consistent [Debussche et al., 2023]

Beyond white noise

Why white noise is not enough here? The unresolved ocean flow is **not zero-mean** — jets meander, eddy trains recur — and **not white in time**: eddies decorrelate over days to weeks, not instantly.

(i) **Give the slow part a drift** [Li et al., 2023a]



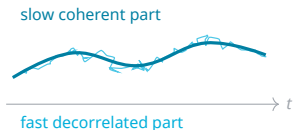
$$d\mathbf{W}_t \longrightarrow \underbrace{\boldsymbol{\Gamma}_t dt}_{\text{drift}} + \underbrace{d\tilde{\mathbf{W}}_t}_{\text{martingale}}$$

- **Girsanov's theorem** makes this legitimate:
 - under a new measure, $\tilde{\mathbf{W}}$ is again a Wiener field
- ✓ the martingale structure — hence every conservation property of $\mathbb{D}_t$ — is untouched

Beyond white noise

Why white noise is not enough here? The unresolved ocean flow is **not zero-mean** — jets meander, eddy trains recur — and **not white in time**: eddies decorrelate over days to weeks, not instantly.

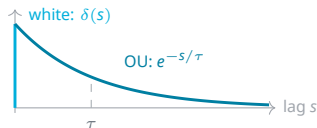
(i) Give the slow part a drift [Li et al., 2023a]



$$d\mathbf{W}_t \longrightarrow \underbrace{\boldsymbol{\Gamma}_t dt}_{\text{drift}} + \underbrace{d\tilde{\mathbf{W}}_t}_{\text{martingale}}$$

- **Girsanov's theorem** makes this legitimate:
under a new measure, $\tilde{\mathbf{W}}$ is again a Wiener field
- ✓ the martingale structure — hence every conservation property of $\mathbb{D}_t$ — is untouched

(ii) Give the fast part a memory [Debussche and Mémín, 2025]

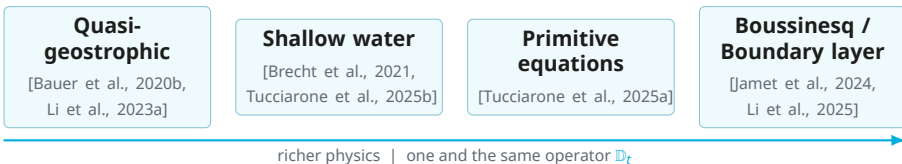


Replace $d\mathbf{W}_t$ by a **time-correlated** field $\boldsymbol{\eta}_t$:

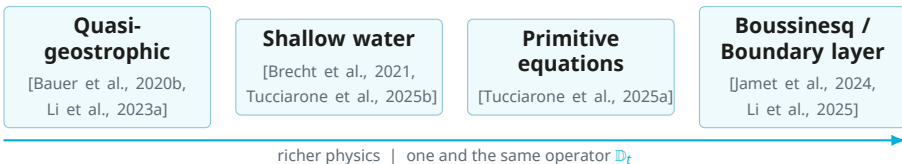
$$d\mathbf{X}_t = (\mathbf{v} + \boldsymbol{\Gamma}_t)dt + \boldsymbol{\eta}_t dt, \quad d\boldsymbol{\eta}_t = -\frac{1}{\tau}\boldsymbol{\eta}_t dt + \frac{1}{\tau}d\mathbf{W}_t$$

- $\boldsymbol{\eta}_t$ is an Ornstein-Uhlenbeck field:
 $\mathbb{E}[\boldsymbol{\eta}_t \otimes \boldsymbol{\eta}_s] \propto e^{-|t-s|/\tau}$
 - ✓ **Wong-Zakai**: as $\tau \rightarrow 0$, $\int_0^t \boldsymbol{\eta}_s ds \rightarrow \mathbf{W}_t$, the model converges back to the white-noise formulation
- ⇒ a single family, with τ interpolating between smooth and white

A hierarchy of stochastic models



A hierarchy of stochastic models



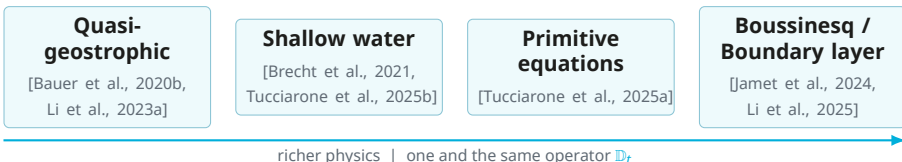
Stochastic primitive equations

Key ocean state variables: momentum $\mathbf{v} = (u, v, w)^T$, pressure p , density ρ , temperature T , salinity S

$$\begin{aligned}\mathbb{D}_t u &= (fv - \partial_x p + \mathcal{F}_u) dt \\ \mathbb{D}_t v &= (-fu - \partial_y p + \mathcal{F}_v) dt \\ \partial_z p &= -\rho g \\ \nabla \cdot \mathbf{v} &= 0, \quad \nabla \cdot d\mathbf{W}_t = 0 \\ \mathbb{D}_t T &= \mathcal{F}_T dt, \quad \mathbb{D}_t S = \mathcal{F}_S dt \\ \rho &= \rho_{\text{eos}}(T, S, p)\end{aligned}$$

- Each model derived from an **asymptotic expansion**, with consistent noise scaling
- Material derivative $\rightarrow$ **stochastic transport**, implemented in NEMO [Tucciarone et al., 2025a]
- **Well-posed** even for the full stochastic PE system [Debussche et al., 2025]

A hierarchy of stochastic models



Stochastic primitive equations

Key ocean state variables: momentum $\mathbf{v} = (u, v, w)^T$, pressure p , density ρ , temperature T , salinity S

$$\mathbb{D}_t u = (fv - \partial_x p + \mathcal{F}_u) dt$$

$$\mathbb{D}_t v = (-fu - \partial_y p + \mathcal{F}_v) dt$$

$$\partial_z p = -\rho g$$

$$\nabla \cdot \mathbf{v} = 0, \quad \nabla \cdot d\mathbf{W}_t = 0$$

$$\mathbb{D}_t T = \mathcal{F}_T dt, \quad \mathbb{D}_t S = \mathcal{F}_S dt$$

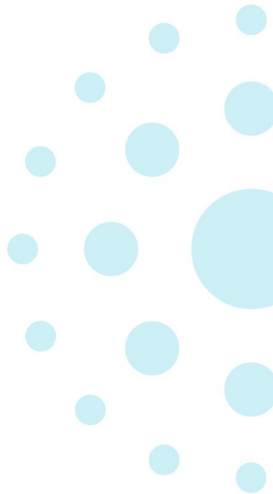
$$\rho = \rho_{\text{eos}}(T, S, p)$$

- Each model derived from an **asymptotic expansion**, with consistent noise scaling
- Material derivative $\rightarrow$ **stochastic transport**, implemented in NEMO [Tucciarone et al., 2025a]
- **Well-posed** even for the full stochastic PE system [Debussche et al., 2025]

Everything is determined **except the modes** $\{\phi_n\}$ **and the drift** Γ_t .

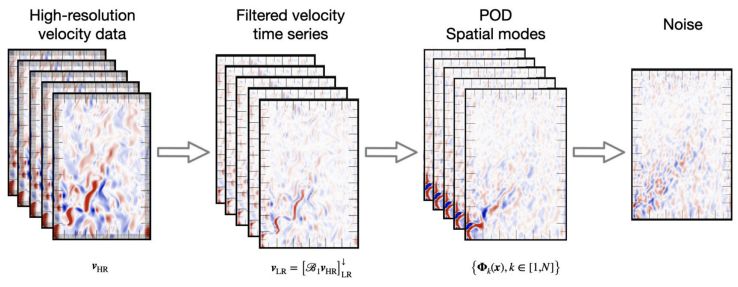
That is what we learn from data.

1. Introduction
2. Stochastic transport modeling
- 3. Data-driven calibration**
4. Numerical validation in NEMO
5. Conclusion & perspectives



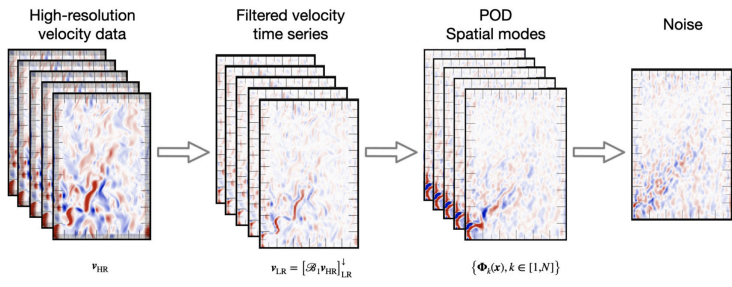
(a) Proper orthogonal decomposition [Li et al., 2023a]

POD extracts the dominant **coherent structures** of the coarse-grained high-resolution flow, as orthogonal modes ranked by energy.



(a) Proper orthogonal decomposition [Li et al., 2023a]

POD extracts the dominant **coherent structures** of the coarse-grained high-resolution flow, as orthogonal modes ranked by energy.



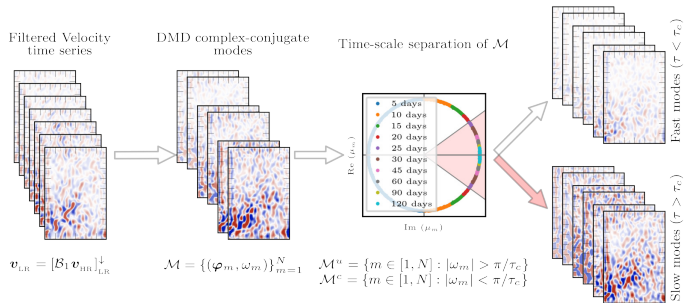
$$d\mathbf{W}_t(\mathbf{x}) = \sum_k \sqrt{\lambda_k} \Phi_k(\mathbf{x}) d\beta_t^k$$

(λ_k, Φ_k) : eigenpairs of the covariance

- ✓ Simple, robust, extensively tested — restores the jet and the low-frequency variability of coarse runs
- **Stationary**: the covariance is frozen once and for all, so the unresolved field carries *no temporal organization*

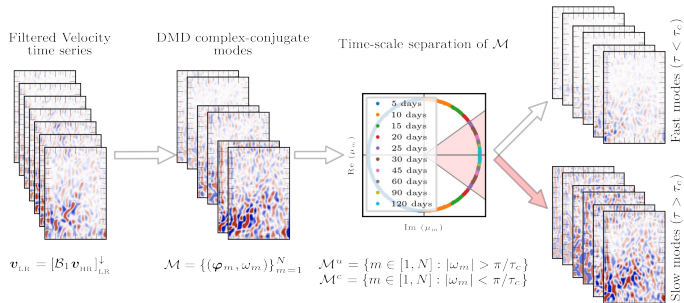
(b) Dynamic mode decomposition [Li et al., 2023b]

DMD fits a **linear dynamical system** to the same data, giving **spatio-temporal modes** (φ_m, ω_m) — each mode now carries its own frequency.



(b) Dynamic mode decomposition [Li et al., 2023b]

DMD fits a **linear dynamical system** to the same data, giving **spatio-temporal modes** (φ_m, ω_m) — each mode now carries its own frequency.



Split at a cut-off π/τ_c :

$$\mathcal{M}^c = \{|\omega_m| < \pi/\tau_c\}, \quad \mathcal{M}^u = \{|\omega_m| > \pi/\tau_c\}$$

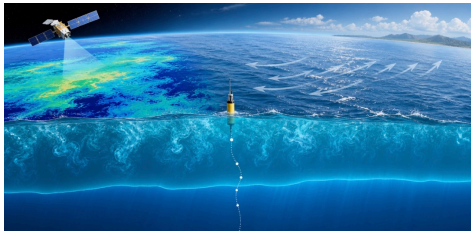
$$\Gamma_t = \sum_{m \in \mathcal{M}^c} e^{i\omega_m t} \varphi_m(\mathbf{x}),$$

$$d\tilde{\mathbf{W}}_t = \sum_{m \in \mathcal{M}^u} e^{i\omega_m t} \varphi_m(\mathbf{x}) d\beta_t^m$$

- ✓ **Natural split** by frequency: slow $\rightarrow$ drift, fast $\rightarrow$ noise
- ✓ Better variability and backscatter, with *fewer* modes than POD
- Needs a long, well-sampled high-resolution record; τ_c is a tuning knob

(c) Towards learning from observations [Tucciarone et al., 2025a]

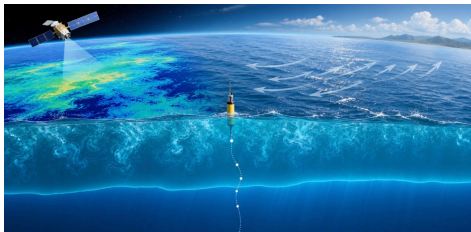
Problem: POD and DMD both need *full 3-D* high-resolution velocity — available from a reference simulation, *never from observations*.



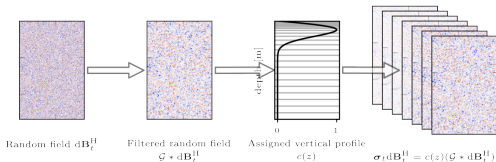
What we *do* observe: **sea surface height** from altimetry (SWOT), **sparse profiles** from Argo

(c) Towards learning from observations [Tucciarone et al., 2025a]

Problem: POD and DMD both need *full 3-D* high-resolution velocity — available from a reference simulation, *never from observations*.



What we *do* observe: **sea surface height** from altimetry (SWOT), **sparse profiles** from Argo



Hybrid approach

- **2-D Drift** — DMD on the surface geostrophic velocity $\mathbf{u}_\eta = \frac{g}{f} \nabla^\perp \eta$:

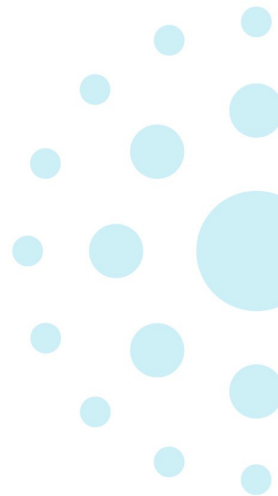
$$\Gamma_t = \bar{\mathbf{u}}_\eta + \sum_{m \in \mathcal{M}^c} e^{i\omega_m t} \varphi_m(\mathbf{x})$$

- **3-D noise** — a *homogeneous* 2-D field, modulated by a vertical profile:

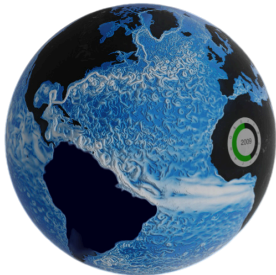
$$\text{d}\mathbf{W}_t = c(\mathbf{x}, z) (\mathcal{G} * \text{d}\mathbf{B}_t)$$

- **Now:** the only data entering is **SSH** — the vertical profile is model-driven (from a resolved variable)
- **Next:** let **Argo** profiles set it from data

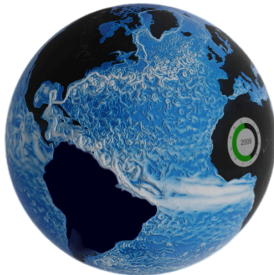
1. Introduction
2. Stochastic transport modeling
3. Data-driven calibration
- 4. Numerical validation in NEMO**
5. Conclusion & perspectives



Implementation in NEMO



- Operational ocean model developed by the European community
- Finite differences on an Arakawa C grid, non-linear free surface, MPI
- Noise advection and Itô diffusion integrated with a **Milstein-type** scheme

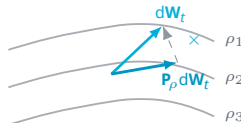


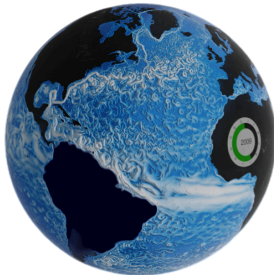
- Operational ocean model developed by the European community
- Finite differences on an Arakawa C grid, non-linear free surface, MPI
- Noise advection and Itô diffusion integrated with a **Milstein-type** scheme

Isopycnal projection of noises

$$dW_t^\rho = P_\rho dW_t, \quad P_\rho = I - \frac{\nabla \rho (\nabla \rho)^T}{|\nabla \rho|^2}$$

offline noises $\rightarrow$ online noises



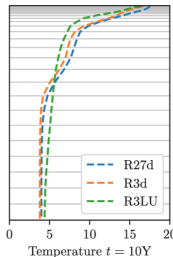
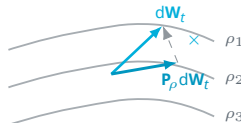


- Operational ocean model developed by the European community
- Finite differences on an Arakawa C grid, non-linear free surface, MPI
- Noise advection and Itô diffusion integrated with a **Milstein-type** scheme

Isopycnal projection of noises

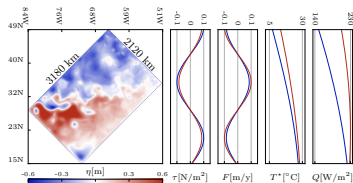
$$dW_t^\rho = P_\rho dW_t, \quad P_\rho = I - \frac{\nabla \rho (\nabla \rho)^T}{|\nabla \rho|^2}$$

offline noises $\rightarrow$ online noises



- ✓ Along-isopycnal mixing exceeds diapycnal mixing by **seven orders of magnitude**
- ! Unconstrained noise stirs *across* density surfaces and erodes the thermocline — a numerical artefact, not physics
- ✓ Model-driven noise needs no projection: it is built from resolved fields that already carry the stratification

Configuration and simulations [Tucciarone et al., 2025a]

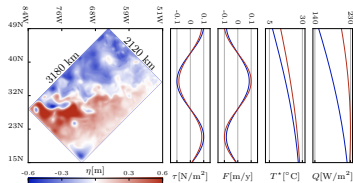


basin, SSH and seasonal forcing; 31 levels

	R27	R3
grid	$1/27^\circ$	$1/3^\circ$
Δx	3.9 km	35.3 km
role	eddy-resolving	eddy-permitting

($R_d \approx 40$ km in the gyre centre)

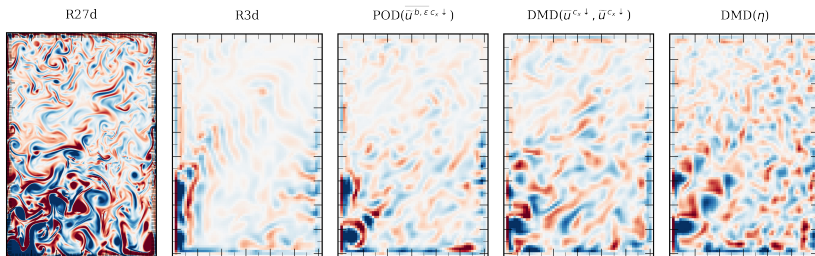
Configuration and simulations [Tucciarone et al., 2025a]



basin, SSH and seasonal forcing; 31 levels

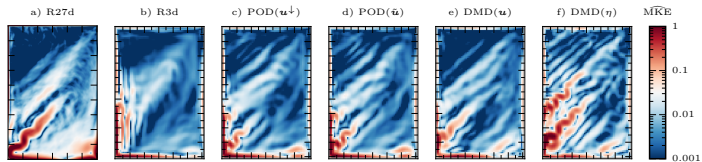
	R27	R3
grid	$1/27^\circ$	$1/3^\circ$
Δx	3.9 km	35.3 km
role	eddy-resolving	eddy-permitting

($R_d \approx 40$ km in the gyre centre)

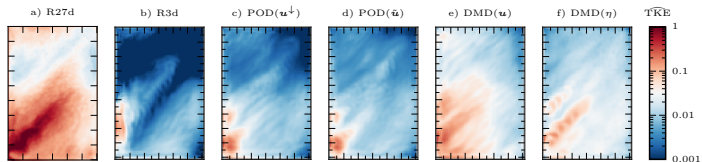


Instantaneous surface relative vorticity of different runs

Mean structure and variability

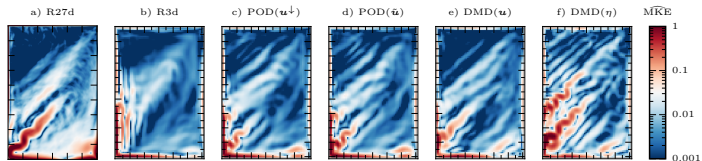


- deterministic R3: no jet at all

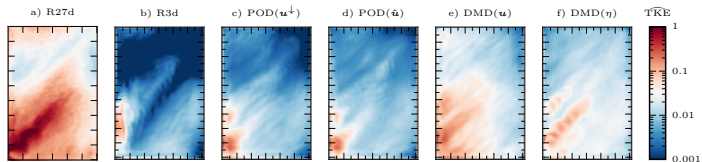


10-year mean KE (top) and turbulent KE (bottom), normalized by R27

Mean structure and variability

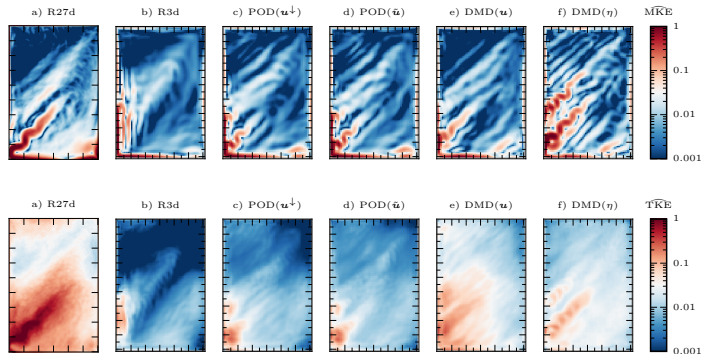


- deterministic R3: **no jet at all**
- *every* stochastic model imprints the **eastward jet** in the long-term mean



10-year mean KE (top) and turbulent KE (bottom), normalized by R27

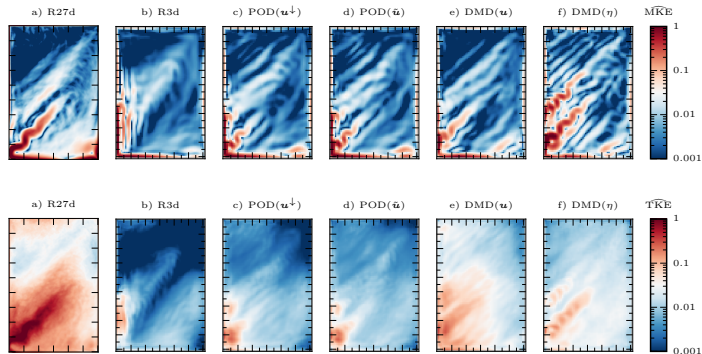
Mean structure and variability



- deterministic R3: **no jet at all**
- *every* stochastic model imprints the **eastward jet** in the long-term mean
- POD: turbulent KE confined near the western boundary

10-year mean KE (top) and turbulent KE (bottom), normalized by R27

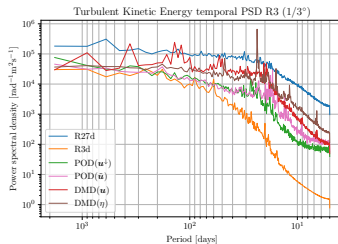
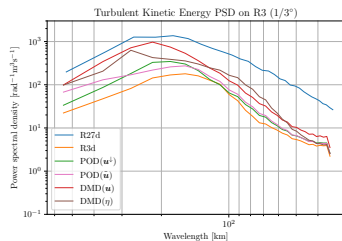
Mean structure and variability



- deterministic R3: **no jet at all**
- *every* stochastic model imprints the **eastward jet** in the long-term mean
- POD: turbulent KE confined near the western boundary
- DMD: turbulent KE **spread over the basin** — closest to R27

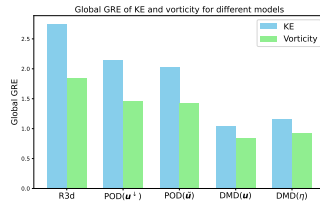
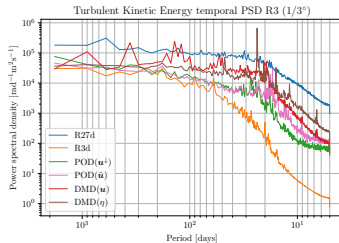
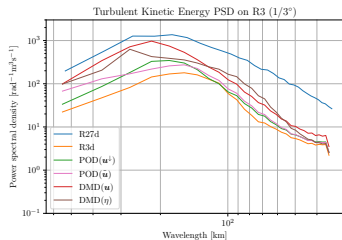
10-year mean KE (top) and turbulent KE (bottom), normalized by R27

Spectra and scores



- **Spatial** (left): the deterministic run is depleted at *all* wavelengths
- **Temporal** (right): 2–3 *decades* of energy missing below 100 d
- The DMD noises sit on top of R27 in the 20–100 d band → the right eddy time scales

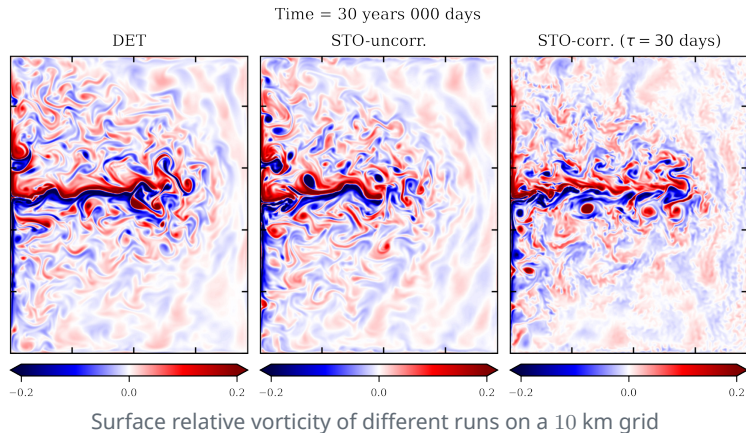
Spectra and scores



- **Spatial** (left): the deterministic run is depleted at *all* wavelengths
- **Temporal** (right): 2–3 *decades* of energy missing below 100 d
- The DMD noises sit on top of R27 in the 20–100 d band → the right eddy time scales

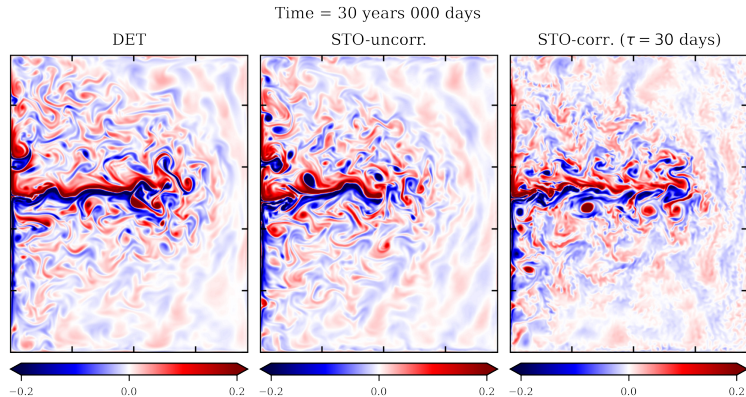
Gaussian relative entropy against R27 — **lower is better**. The DMD noise **halves** the distance to the reference statistics, and $\text{DMD}(\eta)$ **nearly matches** $\text{DMD}(\mathbf{u})$.

Preliminary results: OU noise in a QG model



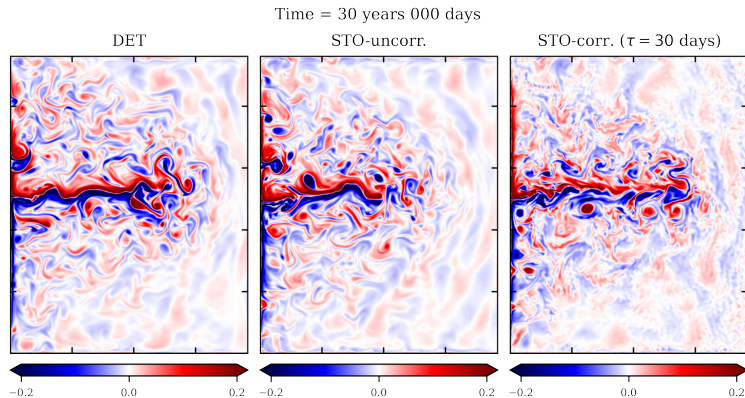
- ✓ In both stochastic runs the noise is simply **homogeneous** — nothing is calibrated on data here

Preliminary results: OU noise in a QG model



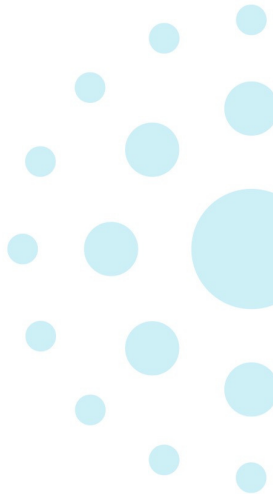
- ✓ In both stochastic runs the noise is simply **homogeneous** — nothing is calibrated on data here
- ✓ **Time correlation** sharpens the flow: more coherent jet, stronger vorticity gradients, finer filaments

Preliminary results: OU noise in a QG model



- ✓ In both stochastic runs the noise is simply **homogeneous** — nothing is calibrated on data here
- ✓ **Time correlation** sharpens the flow: more coherent jet, stronger vorticity gradients, finer filaments
- **Ongoing**: the same noise is being validated in **NEMO**, in a realistic global configuration

1. Introduction
2. Stochastic transport modeling
3. Data-driven calibration
4. Numerical validation in NEMO
- 5. Conclusion & perspectives**



Take-home messages

- **Stochastic transport is a physically consistent container for learned closures.**
Derived from the conservation laws, energy-consistent by construction, and the same operator serves the whole hierarchy.

Take-home messages

- **Stochastic transport is a physically consistent container for learned closures.**
Derived from the conservation laws, energy-consistent by construction, and the same operator serves the whole hierarchy.
- **What is learned from data is the noise, not the physics.**
POD gives a stationary noise; DMD splits slow coherent (drift) from fast decorrelated (noise), and does better with fewer modes.

- **Stochastic transport is a physically consistent container for learned closures.**
Derived from the conservation laws, energy-consistent by construction, and the same operator serves the whole hierarchy.
- **What is learned from data is the noise, not the physics.**
POD gives a stationary noise; DMD splits slow coherent (drift) from fast decorrelated (noise), and does better with fewer modes.
- **It works where it matters.**
On a coarse $1/3^\circ$ grid the eddy-driven jet is recovered, the spectra are restored over 2–3 decades, and the distance to the $1/27^\circ$ reference is halved.

- **Stochastic transport is a physically consistent container for learned closures.**
Derived from the conservation laws, energy-consistent by construction, and the same operator serves the whole hierarchy.
- **What is learned from data is the noise, not the physics.**
POD gives a stationary noise; DMD splits slow coherent (drift) from fast decorrelated (noise), and does better with fewer modes.
- **It works where it matters.**
On a coarse $1/3^\circ$ grid the eddy-driven jet is recovered, the spectra are restored over 2–3 decades, and the distance to the $1/27^\circ$ reference is halved.
- **Surface data works might be enough.**
 $\text{DMD}(\eta)$ learns the drift from sea-surface height alone and nearly matches the full 3-D calibration, but could be further improved by combining observation profiles.

ANR PINOT — Physics-Informed Neural networks for Ocean Turbulence



with E. Simonnet (INPHYNI), A. Venaille (ENS de Lyon), B. Deremble (IGE)

Two directions

1. **3-D reconstruction from observations.** Condition on surface fields (**SSH**, **SST**) and sparse **Argo** profiles, then sample the full 3-D state → the full DMD requires, and an **improved prior** for data assimilation [Rozet and Louppe, 2023]
2. **Calibration of the stochastic model.** Learn the law of the unresolved field given the resolved one → a **diffusion model** [Ho et al., 2020] samples $d\mathbf{W}_t$, while $\mathbb{D}_t$ enforces the physics

⇒ How do we keep a **learned, non-Gaussian** noise **physically admissible**?

Thank you for your attention!

Questions and discussion

Long Li | long.li@univ-cotedazur.fr | <https://matlong.github.io>

References I

-  **Bauer, W., Chandramouli, P., Chapron, B., Li, L., and Mémin, E. (2020a).**
Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach.
J. Phys. Oceanogr., 50(4):983–1003.
-  **Bauer, W., Chandramouli, P., Li, L., and Mémin, E. (2020b).**
Stochastic representation of mesoscale eddy effects in coarse-resolution barotropic models.
Ocean Model., 151:101646.
-  **Bolton, T. and Zanna, L. (2019).**
Applications of deep learning to ocean data inference and subgrid parameterization.
J. Adv. Model. Earth Syst., 11(1):376–399.
-  **Brecht, R., Li, L., Bauer, W., and Mémin, E. (2021).**
Rotating shallow water flow under location uncertainty with a structure-preserving discretization.
J. Adv. Model. Earth Syst., 13(12):e2021MS002492.
-  **Debussche, A., Hug, B., and Mémin, E. (2023).**
A consistent stochastic large-scale representation of the Navier-Stokes equations.
Journal of Mathematical Fluid Mechanics, 25(1):19.

References II

-  **Debussche, A. and Mémin, E. (2025).**
Variational principles for fully coupled stochastic fluid dynamics across scales.
Physica D: Nonlinear Phenomena, 481:134777.
-  **Debussche, A., Mémin, E., and Moneyron, A. (2025).**
Some properties of a non-hydrostatic stochastic oceanic primitive equations model.
In Stochastic Transport in Upper Ocean Dynamics III, volume 13 of *Mathematics of Planet Earth*, pages 161–182. Springer, Cham.
-  **Gent, P. R. and McWilliams, J. C. (1990).**
Isopycnal mixing in ocean circulation models.
J. Phys. Oceanogr., 20(1):150–155.
-  **Guillaumin, A. P. and Zanna, L. (2021).**
Stochastic-deep learning parameterization of ocean momentum forcing.
J. Adv. Model. Earth Syst., 13(9):e2021MS002534.
-  **Hasselmann, K. (1976).**
Stochastic climate models Part I. Theory.
Tellus, 28(6):473–485.

References III

-  Ho, J., Jain, A., and Abbeel, P. (2020).
Denoising diffusion probabilistic models.
In Adv. Neural Inf. Process. Syst. (NeurIPS), volume 33, pages 6840–6851.
-  Jamet, Q., Mémin, E., Dumas, F., Li, L., and Garreau, P. (2024).
Toward a stochastic parameterization for oceanic deep convection.
In Stochastic Transport in Upper Ocean Dynamics II, volume 11 of *Mathematics of Planet Earth*, page 143–157. Springer.
-  Li, L., Deremble, B., Lahaye, N., and Mémin, E. (2023a).
Stochastic data-driven parameterization of unresolved eddy effects in a baroclinic quasi-geostrophic model.
J. Adv. Model. Earth Syst., 15(2):e2022MS003297.
-  Li, L., Mémin, E., and Chapron, B. (2025).
A generalized stochastic formulation of the Ekman-Stokes model with statistical analyses.
J. Phys. Oceanogr.
<https://hal.science/hal-04672188v3/document>.

References IV

-  Li, L., Mémin, E., and Tissot, G. (2023b).
Stochastic parameterization with dynamic mode decomposition.
In Stochastic Transport in Upper Ocean Dynamics, volume 10 of *Mathematics of Planet Earth*, pages 179–193. Springer.
-  Madec, G. and NEMO System Team (2023).
NEMO ocean engine.
Scientific Notes of Climate Modelling Center, Institut Pierre-Simon Laplace.
-  Palmer, T. N. (2019).
Stochastic weather and climate models.
Nat. Rev. Phys., 1:463–471.
-  Redi, M. H. (1982).
Oceanic isopycnal mixing by coordinate rotation.
J. Phys. Oceanogr., 12(10):1154–1158.
-  Rozet, F. and Louppe, G. (2023).
Score-based data assimilation.
In Adv. Neural Inf. Process. Syst. (NeurIPS), volume 36, pages 40521–40541.

 Tucciarone, F., Li, L., Mémin, E., and Chandramouli, P. (2025a).
Derivation and numerical assessment of a stochastic large-scale hydrostatic primitive model.
J. Adv. Model. Earth Syst., 17:e2024MS004783.

 Tucciarone, F. L., Li, L., , Mémin, E., and Thiry, L. (2025b).
Transport noise defined from wavelet transform for model-based stochastic ocean models.
In Stochastic Transport in Upper Ocean Dynamics III, volume 13 of *Mathematics of Planet Earth*, page
287–303. Springer.