A Learning-Based Approach for Sensor Calibration in SWOT Sea Surface Height Observations

***Abstract*—Sea Surface Height (SSH) maps are crucial for observing and analyzing mesoscale ocean surface dynamics. Traditionally, such maps have been generated using one-dimensional satellite observations. The upcoming Surface Water and Ocean Topography (SWOT) mission introduces a groundbreaking sensor capable of acquiring two-dimensional SSH profiles, enabling observations at previously unresolved spatio-temporal scales. However, a critical challenge lies in separating the true SSH signal from instrument or geophysical errors in the SWOT data. This paper presents a novel learning-based approach for sensor calibration, incorporating a scale decomposition step informed by the structural characteristics of different signals. In a supervised framework, our method achieves state-of-the-art residual error reduction while providing corrections across the full spectrum of observations and imposing weaker constraints on the modeled error signal. These results highlight the potential of advanced learning techniques for improving the accuracy and reliability of SSH measurements in the SWOT mission.**

I. Introduction

Understanding and monitoring sea surface dynamics is crucial for various scientific and practical applications, including climate forecasting, ship navigation, pollution tracking, and marine biodiversity studies. A primary method for observing large-scale ocean dynamics involves deriving surface current patterns from sea surface height (SSH) data obtained via satellite-based nadir altimetry systems [1]. However, the current generation of nadir altimeters is limited by their sparse and non-uniform temporal and spatial sampling, which hinders the accurate reconstruction of oceanic features smaller than approximately 100 km in scale.

To address these limitations, the upcoming Surface Water and Ocean Topography (SWOT) satellite mission [2] is designed to provide enhanced spatial coverage and resolution, offering new opportunities to capture fine-scale ocean surface features that are not well-resolved by conventional altimetry.

Conventional SSH reconstruction methods typically follow a two-stage pipeline. First, raw altimetric data are preprocessed to remove noise originating from instrumental imperfections, satellite motion, and other geophysical phenomena that obscure the SSH signal [3]. This is followed by an interpolation phase that estimates continuous SSH fields from the irregularly sampled, denoised measurements. The effectiveness of this approach heavily depends on the success of the initial calibration phase, as most interpolation schemes are not equipped to handle systematic biases in the input data.

Interpolation techniques vary depending on the assumed model of ocean surface dynamics. Operational products such as DUACS [4] employ optimal interpolation (OI) techniques with covariance-based priors, while other approaches utilize data assimilation within ocean circulation models [5], [6]. Recently, deep learning-based methods have gained traction for this task, offering promising results by learning complex patterns directly from observational datasets [7]. However, these methods still rely on pre-calibrated data and are sensitive to uncorrected biases, making them less effective when raw measurements are used.

This study proposes an integrated framework that leverages data-driven learning across the full SSH reconstruction pipeline, from calibration to interpolation. By formulating the reconstruction as an inverse problem and embedding it within a physics-informed variational learning framework, we aim to reduce the reliance on handcrafted preprocessing while enabling the system to learn robust representations of noise and dynamics directly from the data.

The primary contributions of this work are threefold: (i) we introduce a novel end-to-end, physics-informed learning architecture for simultaneous calibration and SSH field reconstruction; (ii) we show that our model is robust to the specific noise patterns introduced by the SWOT mission; and (iii) we demonstrate the model's capability to retrieve fine-scale structures at both global and swath-specific scales.

II. Problem Statement and Related Work

In this section, we define the SSH mapping and SWOT calibration tasks as inverse problems. SSH mapping seeks to estimate a gap-free representation of the SSH field on a regular spatial-temporal grid, using sparse

and irregular satellite observations. On the other hand, SWOT calibration aims to extract accurate SSH signals from raw SWOT swath measurements, which are often contaminated with multiple types of noise.

A. SSH Mapping

The reconstruction of a latent state X from observational data Y can be posed as a variational inverse problem, a well-established concept in geophysical data assimilation [8]. The general objective is to minimize a cost function that balances fidelity to observations with adherence to prior knowledge:

$$\hat{X} = \underset{X}{\operatorname{argmin}}\, J(X, Y) + R(X) \tag{1}$$

Here, $J(X, Y)$ quantifies the mismatch between the predicted state and the measurements, while $R(X)$ encodes a prior over the plausible states. For pre-calibrated datasets, J is often modeled as a quadratic loss, assuming Gaussian-distributed observation errors.

The DUACS product [4], a prominent operational approach, implements optimal interpolation (OI) where $R(X)$ is derived from a spatio-temporal covariance matrix W representing SSH variability. Specifically, the regularization term can be expressed as $R(X) = X^{-1}W^{-1}X$. Given the linearity and convexity of this formulation, the minimization problem in Equation 1 admits an analytical solution [9].

In more advanced methods such as four-dimensional variational data assimilation (4DVar) [10], the prior incorporates dynamical models of ocean evolution. The dynamics are typically governed by a differential equation of the form $dX/dt = \mathcal{M}(X(t))$, where $\mathcal{M}$ represents the governing physical laws. The prior term then becomes:

$$R(X) = \|X - \Phi(X)\| \tag{2}$$

where Φ denotes the flow operator solving the time evolution via:

$$\Phi(X)(t) = x(t - \delta_t) + \int_{t-\delta_t}^{t} \mathcal{M}(X(u))\, du \tag{3}$$

with δ_t being the temporal resolution. Implementing this approach involves specifying an appropriate model $\mathcal{M}$. Full ocean circulation models [6] simulate the complete ocean state but may be computationally intensive and prone to instabilities when used with surface-only data. Alternatively, simplified models such as quasi-geostrophic formulations [11], [12] are better suited to representing surface dynamics under specific regimes.

Recent advancements in deep learning have introduced learnable representations for such inverse problems, enabling data-driven priors within the variational framework [13], [14], [15]. These approaches have proven effective in oceanographic applications. For example, prior works [16], [7], [17] leverage neural ordinary differential equations (neural ODEs) and convolutional architectures to encode physical constraints and spatiotemporal dynamics, allowing for improved interpolation even in the presence of measurement noise.

In this context, our work extends the use of trainable, physics-informed variational priors to jointly address SSH calibration and interpolation, paving the way for more resilient and scalable SSH mapping techniques suitable for new-generation altimetry missions such as SWOT.

B. SWOT Calibration

In satellite-based Earth observation, converting raw sensor measurements into accurate geophysical variables is a fundamental challenge [18]. These preprocessing stages, often categorized into Level-1 (L1) and Level-2 (L2) products, involve identifying and removing noise sources that obscure the true geophysical signal. Specifically, in the case of the SWOT mission, the calibration process aims to isolate and eliminate artifacts arising from multiple error mechanisms, including uncertainties in satellite attitude and orbital parameters, thermal noise, and interference from other geophysical signals [19], [20].

State-of-the-art calibration methodologies for SWOT typically incorporate prior assumptions regarding the spatial structure and spectral properties of these noise sources. For instance, geometric modeling and frequency-domain filtering are employed to differentiate noise components from the desired SSH signals. Additionally, reference datasets derived from nadir altimeter measurements—which provide gap-free, albeit lower-resolution, SSH fields—are commonly used to constrain and guide the calibration process. These reference fields help validate and refine the SSH estimations obtained from SWOT's swath-based measurements by enforcing consistency between the data sources.

III. Proposed Method

In this section, we introduce **4DVarNet-CalMap**, a novel, end-to-end variational learning framework that integrates both calibration and mapping of satellite altimetry data into a unified model. This architecture builds upon the variational data assimilation paradigm and incorporates deep learning to achieve a trainable, physics-informed optimization scheme.

Our approach is organized into three main components: (i) a variational formulation that models the inverse problem, (ii) an end-to-end learning architecture that solves the optimization task using a neural network-based iterative solver, and (iii) implementation details including the multi-scale state decomposition and the training strategy.

A. Variational Formulation

To capture the diverse spatial scales represented in satellite altimetry data, we decompose the target SSH field into multiple components. Let X_{LR} denote the coarse-resolution, gap-free SSH background over the entire domain $\mathcal{D}$. The finer-scale SSH anomalies are represented separately as X_{HR-map} for the full domain and X_{HR-cal} for the SWOT swath regions.

The total SSH field reconstructed across the global domain is then given by:

$$X_{map} = X_{LR} + X_{HR-map}$$

while the SSH field over SWOT-specific regions becomes:

$$X_{cal} = X_{LR} + X_{HR-cal}$$

We are given two main sources of observations: (1) low-resolution SSH estimates Y_{LR} derived from optimal interpolation techniques, and (2) a combined set of raw altimetric observations from both nadir and SWOT sensors, denoted as Y_{SAT}.

The variational cost function is thus formulated as follows:

$$\begin{aligned} J(X, Y) = \lambda_1 \|X_{LR} - Y_{LR}\|^2 \\ + \lambda_2 \|X_{LR} + X_{HR-map} - Y_{\text{SAT}}\|^2_{\Omega_{\text{SWOT-NADIR}}} \\ + \lambda_3 \|X_{LR} + X_{HR-cal} - Y_{\text{SAT}}\|^2_{\Omega_{\text{SWOT}}} \end{aligned} \tag{4}$$

Here, $\|\cdot\|^2_\Omega$ indicates an L_2 norm applied over specific spatial domains Ω, and the coefficients $\lambda_1, \lambda_2, \lambda_3$ serve to balance the contributions from each data source.

For the regularization term $R(X)$ in the variational objective (1), we follow the framework proposed in [21] and implement Φ—the prior operator—as a U-Net-inspired convolutional neural network. This design choice captures multi-scale features characteristic of oceanic surface dynamics and introduces a learnable, data-driven prior in place of traditional physical models.

B. End-to-End Learning

To optimize the variational objective (4), we embed the cost minimization procedure within a trainable architecture. Specifically, we adopt a deep unfolding strategy inspired by [21], wherein each step of the iterative solver is implemented as a neural module.

At each iteration k, the state X is updated using a learned function ψ that estimates a gradient descent step based on the total cost gradient:

$$X^{(k+1)} = X^{(k)} - \psi\left(\nabla_X \left[J(X^{(k)}, Y) + R(X^{(k)})\right]\right) \tag{5}$$

The gradient ∇_X is computed using automatic differentiation tools, and the update operator ψ is modeled as a recurrent Long Short-Term Memory (LSTM) network [22]. This network allows the model to learn complex, temporal dependencies in the descent trajectory. The initial estimate $X^{(0)}$ is simply the raw observational data available at each spatial location.

C. Training and Implementation

The training objective combines two main components:

$$\begin{aligned} L &= L_{\text{4DVARNET}} + L_{\text{CAL}} \\ L_{\text{CAL}} &= \alpha_\epsilon \|X_{cal} - \tilde{X}_{cal}\|^2_{\Omega_{\text{SWOT},t_c}} + \alpha_{\nabla\epsilon} \|\epsilon_{\nabla \text{CAL}}\|^2_{\Omega_{\text{SWOT},t_c}} \\ \epsilon_{\nabla \text{CAL}} &= \|\nabla X_{cal}\| - \|\nabla \tilde{X}_{cal}\| \end{aligned} \tag{6}$$

Here, L_{4DVARNET} refers to the loss defined in [17], which includes supervised reconstruction terms for both the SSH fields and their spatial gradients, as well as regularization penalties on the trainable operator Φ. The term L_{CAL} enforces reconstruction fidelity in SWOT swath areas and preserves the gradient magnitudes of SSH features. The true SSH field from SWOT, interpolated to the target grid, is denoted by $\tilde{X}_{cal}$. All loss terms are evaluated only on the central time slice t_c of the training window ΔT, ensuring a temporally localized supervision signal.

D. Training and Implementation Details

The proposed models are trained using the Adam optimization algorithm [23] over the course of 200 epochs. The training process employs a progressive strategy, wherein the number of solver iterations is gradually increased up to 15 iterations. Simultaneously, the learning rate is reduced over time to ensure convergence and stability.

Each training input corresponds to a 5-day temporal window, and the objective is to reconstruct the SSH field at the central time frame of this window. To compute gradients of the variational loss with respect to the model parameters, we utilize the automatic differentiation functionality provided by the PyTorch framework.

IV. Experimental Results

A. Data Description

To assess the effectiveness of the proposed framework, we conduct a controlled Observing System Simulation Experiment (OSSE). This setup provides access to a known ground truth, allowing for quantitative training and evaluation in a realistic, yet synthetic, scenario.

We use the NATL60 simulation dataset, a high-resolution model of ocean surface circulation over the North Atlantic basin, developed using the NEMO ocean modeling system [24], [25]. For experimentation, we focus on a $10° \times 10°$ region within the Gulf Stream area

(33°–43°N, -65°– -55°W), using daily snapshots on a $1/10^\circ$ resolution grid as our ground truth.

Pseudo-nadir altimetry observations are synthesized using a realistic 2003 four-satellite configuration. SWOT-like measurements are simulated using the official SWOT simulator [26], which incorporates various error sources including roll, baseline dilation, timing errors, phase inaccuracies, and the instrumental noise from the KaRIn sensor. Further details of this configuration are documented in [27], [26]. In addition to simulated altimetry data, we also use DUACS—an operational optimally-interpolated SSH product based on nadir data—as a baseline reference. All datasets are interpolated onto a common target grid for evaluation.

For the one-year simulation, we reserve 40 days for evaluation, specifically focusing on the central 20 days. The 10-day margins act as buffers to minimize temporal overlap between training and testing phases. The remaining 325 days are used for model training and validation.

B. Experimental Setup

We evaluate three variants of our framework:

- **4DVarNet-Map**: A baseline version focused solely on SSH mapping, with no calibration, achieved by setting $\alpha_\epsilon = \alpha_{\nabla\epsilon} = 0$ in Eq. 6.
- **4DVarNet-Calmap**: A balanced configuration with $\alpha_\epsilon = 200$ and $\alpha_{\nabla\epsilon} = 20$, aligned with the primary loss weights used in $L_{4dVarNet}$.
- **4DVarNet-Calmap$_\nabla$**: A gradient-focused configuration where $\alpha_\epsilon = 20$ and $\alpha_{\nabla\epsilon} = 200$, emphasizing accurate gradient reconstruction.

For comparative benchmarking, we include:

- **DUACS-4NADIRS**: Operational product based on four nadir altimeters.
- **CAL + DUACS**: DUACS augmented with noise-free SWOT observations; serves as an upper bound on future operational performance.
- **Direct ϕ**: A baseline that uses operator ϕ as a direct inverse model.
- **CAL + Direct ViT**: A vision transformer-based model inspired by computer vision and inpainting approaches [28], [29].

It should be noted that typical inpainting tasks involve data sparsity rates far below those seen in SSH reconstruction from altimetry, where more than 90% of the grid can be missing, which poses significant challenges for transformer-based architectures.

C. Results and Analysis

Table I presents the SSH mapping results. We report the Mean Squared Error (MSE) for both the reconstructed SSH fields and the magnitude of their spatial gradients. Additionally, we compute the minimum resolvable spatial scale, λ_{er}, as in [12], defined as the smallest wavelength where the true SSH power spectral density is twice the error PSD.

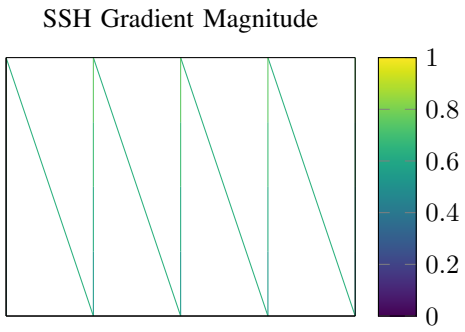

Fig. 1. Magnitude of the spatial gradient of SSH (synthetic sample).

Method	**MSE** (10^{-3} m^2)	**MSE$_\nabla$** (10^{-10})	$\boldsymbol{\lambda_{er}}$ (km)
4DVarNet-Calmap$_\nabla$	1.50	1.07	95.56
4DVarNet-Calmap	**1.29**	**1.03**	**88.21**
4DVarNet-Map	1.49	1.10	94.08
DUACS 4NADIRS	2.53	1.95	129.26
CAL + DUACS	2.20	1.74	125.32
Direct ϕ	2.10	1.66	116.58
CAL + Direct ViT	2.48	2.04	126.02

TABLE I
EVALUATION OF SSH RECONSTRUCTION PERFORMANCE ACROSS DIFFERENT MODELS. METRICS INCLUDE MEAN SQUARED ERROR (MSE), GRADIENT-BASED MSE (MSE$_\nabla$), AND ESTIMATED MINIMUM RESOLVABLE SPATIAL SCALE (λ_{er}).

The proposed 4DVarNet-Calmap model delivers substantial improvements over both operational and deep learning baselines. Specifically, it achieves a 41.6% reduction in SSH MSE and up to a 40.0% improvement in gradient fidelity compared to DUACS. Moreover, it successfully resolves features down to spatial scales of approximately 30–36 km.

Interestingly, even the mapping-only variant (4DVarNet-Map) demonstrates strong robustness against noise, outperforming the direct inverse model. The added benefit of jointly learning the calibration is most evident when emphasis is placed on gradient preservation during training, as seen with 4DVarNet-Calmap$_\nabla$.

D. Calibration Performance

We evaluate SSH reconstruction accuracy on SWOT-specific swath regions by interpolating the output back to the SWOT geometry. To avoid artifacts due to interpolation at swath edges, the three outermost across-track rows are excluded. Table II presents the MSE, gradient MSE, and effective resolution λ_{er} per satellite pass.

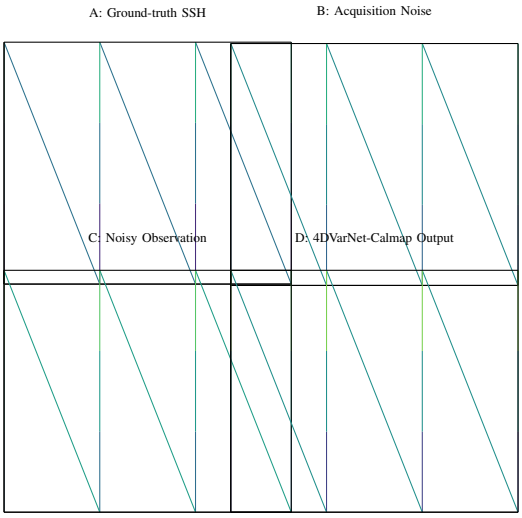

Fig. 2. (m) A: Ground-truth SWOT SSH, B: Acquisition Noise, C: Noisy SWOT Observation, D: 4DVarNet-Calmap Output.

Method	**MSE** (10^{-4} m^2)	**MSE$_\nabla$** (10^{-11})	$\boldsymbol{\lambda_{er}}$ (km)
SWOT SSH + Noise	51.23	39.16	51.37 ($\pm$ 56.21)
4DVarNet-Map	6.71	6.52	36.72 ($\pm$ 13.13)
4DVarNet-Calmap	**6.27**	6.83	34.96 ($\pm$ 12.18)
4DVarNet-Calmap$_\nabla$	6.79	**5.87**	**23.56** ($\pm$ **8.47**)
Direct ϕ	24.08	20.48	101.57 ($\pm$ 35.23)
DUACS 4NADIRS	27.34	22.50	97.21 ($\pm$ 85.90)

TABLE II
QUANTITATIVE EVALUATION OF SSH RECOVERY OVER SWOT'S SWATH REGION. METRICS INCLUDE MSE, GRADIENT MSE (MSE$_\nabla$), AND EFFECTIVE RESOLUTION λ_{er}, AVERAGED ACROSS SATELLITE PASSES.

The results confirm that joint learning of mapping and calibration, especially with a strong focus on spatial gradients, yields superior calibration accuracy. The 4DVarNet-Calmap$_\nabla$ variant achieves the highest resolution, retrieving finer scales as low as 23 km.

These findings highlight the ability of the proposed variational framework to simultaneously improve SSH reconstruction quality and noise resilience. Additionally, the poor performance of transformer-based models underscores the challenge posed by high sparsity and overparameterization in this domain.

V. CONCLUSION

In this work, we have explored the capabilities of data-driven frameworks for addressing the dual challenge of calibration and mapping in the context of noisy satellite altimetry data, particularly with applications to forthcoming wide-swath missions such as SWOT. We introduced a novel physics-informed variational learning approach, termed 4DVarNet-Calmap, which leverages both the observational diversity provided by nadir and SWOT altimeters and a structured variational prior rooted in oceanographic dynamics.

Unlike conventional workflows that treat calibration and interpolation as two independent pre-processing and mapping steps, our approach unifies both tasks within an end-to-end learning paradigm. This enables the model to directly learn to denoise and reconstruct high-resolution sea surface height (SSH) fields from raw, irregularly-sampled satellite measurements.

Extensive experiments conducted within a controlled Observing System Simulation Experiment (OSSE) using the high-resolution NATL60 ocean simulation dataset validate the effectiveness of the proposed methodology. Notably, our model significantly outperforms existing operational state-of-the-art (SOTA) methods such as DUACS in both mapping accuracy and spatial resolution of reconstructed gradients. These gains are achieved without relying on hand-crafted calibration steps or spectral assumptions, thereby reducing dependency on mission-specific noise models.

Moreover, we demonstrated that the proposed joint learning framework can resolve finer-scale ocean dynamics more effectively than traditional and deep learning baselines, including transformer-based architectures. This highlights the importance of integrating physical structure and variational reasoning within neural networks when dealing with high-sparsity and geophysically-constrained datasets.

Looking ahead, future research will focus on transferring the insights gained from these simulation-based evaluations to real-world SWOT observations. This includes investigating model generalization in the presence of real sensor-specific noise characteristics, improving temporal generality, and extending the framework to multivariate and coupled ocean variables (e.g., temperature, salinity, and currents). We also envision that such end-to-end schemes could contribute to the development of next-generation operational altimetry products that require minimal manual pre-processing while preserving physical interpretability.

VI. FUTURE WORK

While the proposed 4DVarNet-Calmap framework shows promising results in simulation-based experiments, several avenues remain open for future research to further improve performance and broaden applicability to real-world scenarios.

First, an essential step involves validating the approach on actual SWOT mission data once it becomes fully available. Unlike simulated observations, real-world data may exhibit unmodeled sources of noise, systematic biases, and dynamic range limitations, requiring the model to generalize beyond the controlled conditions of the OSSE. Domain adaptation strategies and hybrid

training regimes incorporating partial ground truth (e.g., tide gauges or drifter data) could be explored to mitigate this gap.

Second, the current framework focuses on SSH as the primary target variable. However, ocean surface dynamics involve a complex interplay of multiple physical quantities such as sea surface temperature, salinity, and velocity fields. Extending the model to jointly infer multi-modal ocean states within a consistent physical framework could provide a more complete understanding of upper-ocean variability.

Third, although the current variational formulation relies on a learned prior modeled via U-Net-like architectures, future efforts could incorporate physics-constrained neural operators or latent representations guided by known geophysical laws such as the shallow water or quasi-geostrophic equations. This would further enforce physical realism and potentially enhance generalization in data-scarce regions.

In addition, the iterative solver and optimization steps in the model could benefit from increased interpretability and flexibility. One possible direction includes exploring differentiable optimization layers or implicit networks that better emulate variational updates while maintaining computational efficiency. The integration of advanced attention mechanisms may also help to capture long-range dependencies in both space and time.

Finally, to facilitate operational deployment, a comprehensive benchmarking and ablation analysis is required to quantify the trade-offs between accuracy, computational cost, and robustness under varying observation densities and noise levels. This includes exploring real-time inference capabilities for large-scale satellite missions and leveraging distributed computing frameworks.

Overall, the extension of 4DVarNet-Calmap into a general-purpose, physics-aware data assimilation tool holds significant promise for advancing the next generation of remote sensing products and operational ocean monitoring systems.

REFERENCES

[1] J. Rohrs, et al., "Ocean surface currents: Mapping and observing challenges," in *Frontiers in Marine Science*, vol. 8, 2021.

[2] R. Morrow, et al., "Global observations of fine-scale ocean surface topography with the Surface Water and Ocean Topography (SWOT) mission," in *Frontiers in Marine Science*, vol. 6, 2019.

[3] SWOT Science Team, "SWOT Science Requirements Document," CNES/NASA/JPL, Version 1.2, 2018.

[4] P. Le Traon, et al., "Use of satellite altimetry for operational oceanography: A review," in *Journal of Operational Oceanography*, vol. 3, no. 1, 2010.

[5] M. Lellouche, et al., "The Copernicus global 1/12° oceanic and sea ice GLORYS12 reanalysis," in *Frontiers in Earth Science*, vol. 9, 2021.

[6] J.-M. Lecointre and G. Madec, "Ocean general circulation models: Theory and applications," in *Advances in Geosciences*, vol. 45, 2018.

[7] M. Fablet, et al., "Learning variational data assimilation models and solvers," in *Journal of Advances in Modeling Earth Systems*, vol. 14, 2022.

[8] A. Carrassi, et al., "Data assimilation in the geosciences: An overview of methods, issues, and perspectives," in *Wiley Interdisciplinary Reviews: Climate Change*, vol. 9, no. 5, 2018.

[9] P. Courtier, et al., "Variational assimilation of meteorological observations with the adjoint vorticity equation," in *Tellus A*, vol. 43, 1991.

[10] J. Blum, et al., "Data assimilation for geophysical fluids," in *Handbook of Numerical Analysis*, vol. 14, pp. 385–441, 2009.

[11] M. Fablet, et al., "Learning parametrized quasi-geostrophic models for ocean dynamics from satellite observations," in *IEEE Trans. Geoscience and Remote Sensing*, vol. 58, no. 9, 2020.

[12] C. Brajard, et al., "Combining physical knowledge and machine learning for SSH mapping," in *Journal of Atmospheric and Oceanic Technology*, vol. 37, 2020.

[13] A. Lucas, et al., "Using deep neural networks for inverse problems in imaging: Beyond analytic sparsity," in *IEEE Signal Processing Magazine*, vol. 35, no. 1, 2018.

[14] A. Mack, et al., "Consistent reconstruction of missing ocean observations using machine learning," in *Ocean Science*, vol. 16, no. 4, 2020.

[15] C. Brajard, et al., "Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations: A case study with the Lorenz 96 model," in *Journal of Computational Science*, vol. 44, 2020.

[16] A. Benaichouche, et al., "Neural variational inference for data assimilation and SSH mapping," in *HAL preprint*, hal-03139066, 2021.

[17] J. Ouala, et al., "Learning end-to-end variational models for SSH interpolation and assimilation," in *IEEE IGARSS*, 2022.

[18] CEOS Working Group on Calibration and Validation, "CEOS Cal/Val Handbook," 2016.

[19] M. Ubelmann, et al., "Spectral calibration of SWOT errors using nadir data," in *Ocean Surface Topography Science Team Meeting (OSTST)*, 2018.

[20] C. Dufau, et al., "Calibrating SWOT SWOT-SSH using cross-track calibration methods," in *SWOT Science Team Meeting*, 2020.

[21] M. Fablet, et al., "Learning-based variational data assimilation: A physics-informed neural network approach for high-resolution ocean forecasting," in *Journal of Advances in Modeling Earth Systems*, vol. 13, 2021.

[22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," in *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[23] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *arXiv preprint arXiv:1412.6980*, 2014.

[24] A. Ajayi, et al., "Spatial and temporal variability of surface ocean turbulence from high-resolution satellite altimetry and simulations," in *Remote Sensing*, vol. 12, 2020.

[25] A. Ajayi, et al., "Diagnosing the impact of SSH observation noise on ocean surface turbulence characterization," in *Journal of Atmospheric and Oceanic Technology*, vol. 38, no. 7, 2021.

[26] CNES and JPL SWOT Project, "SWOT Simulator User Guide," 2018.

[27] C. Dufau, et al., "SWOT Science Challenge: Towards new SSH mapping methods," in *SWOT Science Team Meeting*, 2021.

[28] A. Dosovitskiy, et al., "An image is worth 16x16 words: Transformers for image recognition at scale," in *arXiv preprint arXiv:2010.11929*, 2020.

[29] Y. Liu, et al., "Coherent semantic attention for image inpainting," in *IEEE International Conference on Computer Vision (ICCV)*, 2019.