Graphical Abstract

Deep Learning for Ocean Parameters Prediction: A Systematic Literature Review

Wei Wang, Anqi Lu, Zheng Jiang, Gaowei Zhang, Yi Wang, Lingyu Xu, Suixiang Shi

Highlights

Deep Learning for Ocean Parameters Prediction: A Systematic Literature Review

Wei Wang, Anqi Lu, Zheng Jiang, Gaowei Zhang, Yi Wang, Lingyu Xu, Suixiang Shi

- Present a systematic literature review of DL for ocean parameter prediction.
- Analyze the types, formats and characteristics of marine data used in DL models.
- Categorize DL models by architecture and task type across 122 reviewed studies.
- Explore model optimization techniques and evaluation metrics that impact performance.
- Discuss challenges and future research directions for ocean parameter prediction.

Deep Learning for Ocean Parameters Prediction: A Systematic Literature Review

Wei Wang[a], Anqi Lu[a], Zheng Jiang[a], Gaowei Zhang[a,*], Yi Wang[a], Lingyu Xu[b], Suixiang Shi[c]

[a]*Beijing University of Posts and Telecommunications, Beijing, 100876, China*
[b]*Shanghai University, , Shanghai, 200444, China*
[c]*National Marine Data and Information Service, Tianjin, 100876, China*

Abstract

Accurately predicting ocean temperature, salinity, and velocity is fundamental to capture the dynamics of subsurface thermohaline structure and understanding its impacts on marine ecosystems and the Earth's climate system. With the rise of Deep Learning (DL), researchers have explored its potential to capture complex spatiotemporal dependencies in ocean parameter prediction. However, a comprehensive understanding of the application of DL in this field remains in its early stage. To address this gap, we conduct a systematic review of 122 peer-reviewed articles published between 2013 and 2024, aiming to answer four key Research Questions (RQs). In RQ1, we examine how marine datasets are categorized and represented for ocean parameter prediction. RQ2 explores the spatial, temporal, and input-level patterns of these datasets as utilized across various DL models. In RQ3, we analyze the DL architectures adopted for forecasting temperature, salinity, and velocity, highlighting trends across model types. Finally, RQ4 investigates the optimization strategies and evaluation metrics employed to assess predictive performance. Based on our findings, we highlight underexplored challenges and outline promising directions for future research.

Keywords: deep learning, ocean parameter prediction, thermohaline structure, survey

*corresponding author
Email address: zhanggaowei@bupt.edu.cn (Gaowei Zhang)

1. Introduction

The ocean covers three-quarters of the Earth's surface, yet much about it remains unexplored and mysterious. Understanding the movement characteristics and patterns of **ocean thermohaline structure** is crucial for comprehending the exchange of energy and substances in the atmosphere and the water cycle (Rahmstorf, 2006). This, in turn, is closely linked to climate change, predicting weather processes, and maintaining marine ecological balance, and thus (Whitehead, 1995). The primary driving force of subsurface thermohaline is the density differences caused by variations in salinity and temperature (Rahmstorf, 2006). Therefore, accurately predicting three fundamental dynamic variables, i.e., **ocean temperature**, **ocean salinity** and **ocean velocity**, is important for estimating ocean thermohaline structure and 3D dynamic processes in the ocean's interior beforehand.

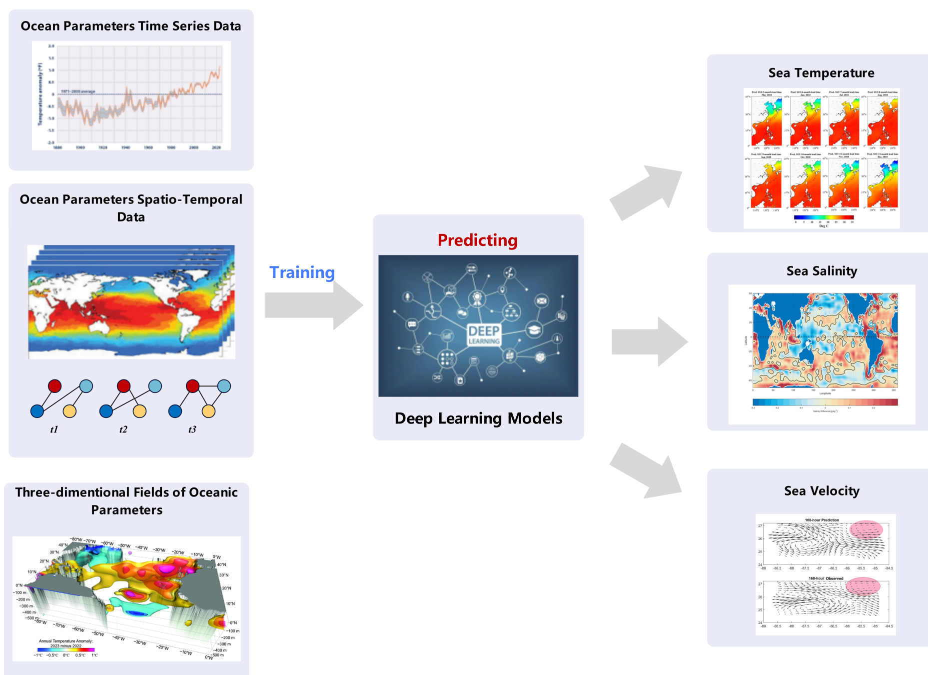


Figure 1: Deep learning models can be tailored to predict fundamental ocean parameters, including temperature, salinity, and velocity.

Historically, dynamic theoretical approaches have been proposed for ocean variables prediction (Olbers et al., 2012; Berliner et al., 2000). These methods aim at modeling the physical environment. These physical models, grounded in the principles of fluid dynamics and thermodynamics, excel in simulating large-scale processes but often struggle with fine-grained spatial and temporal variability due to their reliance on computationally intensive numerical simulations (Wang et al., 2024b). Benefiting from available data from floating buoys and satellite observations, empirical statistical models have been well developed (Vanem et al., 2022), including linear and geographically weighted regression models (Fotheringham et al., 2009), as well as traditional machine learning methods such as support vector machine(Cortes and Vapnik, 1995) and XGBoost(Chen and Guestrin, 2016). In the era of big ocean data, efficiently mining the multi-scale, highly dynamic rules hidden in massive data has become a new challenge in oceanographic research (Zhang et al., 2021a). In recent years, a new paradigm Deep Learning (DL) has emerged that offers unprecedented opportunities for improving the prediction of ocean parameters (Li et al., 2020). As illustrated in Figure 1, DL models can be tailored to predict fundamental ocean parameters, including temperature, salinity, and velocity.

In the specific context of estimating ocean thermohaline structure, Deep Learning models have shown promising results in particularly capturing the dynamic relationships between various oceanographic variables (Meng and Yan, 2022). For instance, **in temperature prediction**, DL models are increasingly able to incorporate not only temperature data but also related factors such as sea surface height, wind speed, solar radiation, and atmospheric pressure to generate more accurate forecasts. Specifically, in Sea Surface Temperature (SST) forecasting, LSTM-based architectures (Xiao et al., 2019a,b) have been used to model time series SST data and have achieved high accuracy in short- and long-term predictions over coastal regions. Beyond surface-level forecasting, advanced models such as multilayer convolutional LSTM networks (Zhang et al., 2020) have been proposed to predict 3D ocean temperature profiles, capturing temperature dynamics from the sea surface down to depths of 2000 meters. Similarly, **in salinity prediction**, DL models have been employed to capture the complex interplay between freshwater input, evaporation, and ocean circulation, enabling more accurate estimates. For example, in the Mid-Atlantic coastal and estuarine region, neural networks trained on over 9,000 paired in-situ and MODIS-Aqua observations achieve surface salinity predictions with RMS errors as low as 1.40 psu, significantly outperforming baseline models (Geiger et al., 2013). **In the case of velocity**, DL models incorporate diverse factors such as bathymetry, wind stress, and pressure gradients to estimate ocean current dynamics more accurately. For example, Thongniran et al. (2019) propose a spatio-temporal deep learning model that combines Convolutional Neural Networks (CNNs) and Gated Recurrent Units (GRUs). Trained on HF radar data collected from the Gulf of Thailand, their model significantly outperforms traditional methods, showing an 11.21% and 27.01% average improvement on the U and V components of surface current velocity, respectively. **The use of DL models is particularly impactful when predicting temperature, salinity, and velocity simultaneously.** For example, a convolutional LSTM-based model has been developed to jointly forecast 3D ocean temperature, salinity, and current fields (Jin et al., 2021). By capturing spatiotemporal dependencies and inter-variable correlations, it outperforms single-variable models in predictive accuracy.

With the growing interest in DL-based ocean parameter prediction and its encouraging outcomes, several review articles have emerged (Aldini et al., 2024; Zhao et al., 2024; Hao et al., 2025; Haghbin et al., 2021; Ahmad, 2019). However, existing surveys are fragmented–limited either in the scope of literature and methodologies reviewed, the diversity of covered tasks, or in-depth discussion of future directions and emerging trends. In particular, Haghbin et al. (2021) exclusively focus on SST prediction and primarily survey soft traditional neural networks and fuzzy logic, overlooking recent deep learning architectures like GNNs, GANs, and Transformer-based models. Ahmad (2019) provides a broad overview of traditional machine learning applications in various oceanographic domains (e.g., oil spill detection, species identification), but lacks a task-specific focus on parameter prediction and does not systematically evaluate modern DL-based forecasting techniques. Hao et al. (2025) offer an analysis of the performance of DL in various weather and ocean prediction applications without conducting a Systematic Literature Review (SLR). Zhao et al. (2024) center on operational and real-time forecasting applications, but focus almost exclusively on SST prediction, with limited attention to other essential parameters such as salinity and velocity, as well as recent advances in interpretable modeling and multi-modal data fusion. In addition, the recent article Aldini et al. (2024) focus specifically on surface current modeling based on numerical methods, without systematic coverage of deep learning-based approaches. These limitations collectively highlight the need for a more focused and structured review that synthesizes recent advances in deep learning with diverse ocean parameter forecasting tasks–particularly temperature, salinity, and velocity–while also offering insights into data sources, model interpretability, and future research challenges.

To gain a deeper understanding of the relationship between deep learning models and their applications in ocean parameter prediction, this study conducts a systematic survey to review, summarize, classify, and analyze relevant research papers, offering valuable insights to the community. We collect, review, and analyze 122 papers published between 2013 and 2024 from multiple disciplines. Through a comprehensive analysis of existing studies, this paper provides valuable insights for researchers and practitioners aiming to advance the DL application in ocean parameter prediction. In this context, it seeks to foster interdisciplinary collaborations between oceanography and artificial intelligence, contributing to more reliable methods for understanding and predicting subsurface thermohaline dynamics in the face of global environmental change. This study makes the following contributions:

- We present the first systematic literature review focusing on the application of deep learning solutions to address ocean parameter prediction challenges, including temperature, salinity, and velocity. Our analysis encompasses 122 papers, examining publication trends and the distribution of publication venues.

- We conduct a comprehensive analysis of data characteristics, encompassing data collection from diverse oceanographic sensors and platforms, the geographical distribution of data across various marine regions, the temporal and spatial coverage of the data, and data representations tailored for deep learning models.

- We provide a categorization of DL models used in ocean parameter prediction based on their architectures and review representative studies by task type (e.g., temperature, salinity, velocity), offering a comparative analysis across different prediction types. Further, we discuss the trend of how DL models have been applied in these tasks.

- We investigate the key factors that impact the performance of DL models in ocean parameter prediction, including model optimization techniques and model evaluation metrics.

- We explore the unique technical challenges associated with applying deep learning to ocean parameter prediction, and propose key directions for future research, offering a comprehensive roadmap to enhance the accuracy and practicality of deep learning in this domain.

Section 2 presents the methodology of our Systematic Literature Review, including four Research Questions (RQs) that guide our investigation, search strategy, study selection process, and data extraction and analysis procedures. Section 3 addresses RQ1 by analyzing the categorization of marine datasets, their key characteristics, and the data representations commonly used in DL-based ocean prediction. Section 4 investigates RQ2 by analyzing the spatial, temporal, and input variable usage patterns of marine datasets in existing literature. Section 5 explores RQ3 by reviewing DL-based approaches applied to predict sea temperature, salinity, and velocity, and summarizes trends. Section 6 answers RQ4 by analyzing the learning algorithms used to optimize DL models and the evaluation metrics. Section 7 discusses the major challenges and outlines future research opportunities. Section 8 concludes this article and summarizes the main findings.

2. Methodology

We present the methodology used for conducting the survey based on general **Systematic Literature Review** (SLR) guidelines proposed by Keele et al. (2007). Following the guidelines, our methodology includes three main steps: planning the review (i.e., Sections 2.1), conducting the review (i.e., Sections 2.2 and 2.3), and analyzing the basic review results (i.e., Section 2.4).

2.1. Research Questions

To provide a comprehensive overview and identify future research opportunities, we address the following research questions:

RQ1: How are marine datasets categorized and represented for ocean parameter prediction? RQ1 examines data types, properties, and representation formats in DL models for ocean parameter prediction. This inquiry aims to clarify foundational data types, identify common spatial-temporal biases and resolution gaps, and summarize prevailing data representations.

RQ2: How are marine datasets utilized across spatial, temporal, and input-level patterns within DL models? RQ2 examines marine datasets utilization from three perspectives: spatial (geographic focus and regional biases), temporal (prediction horizons distinguishing nowcasting from forecasting), and input-level patterns (oceanographic parameter types and combinations used as model inputs).

RQ3: Which DL-based approaches have been applied to predict three different ocean parameters? This question systematically investigates deep learning architectures in oceanographic prediction, categorizing methodologies into six groups: (1) MLPs, (2) CNNs, (3) RNNs/LSTMs, (4) GNNs, (5) GANs, and (6) Transformers. We explore how these models are applied to specific prediction tasks and compare their performance trends and strengths across different applications.

RQ4: How are these DL models optimized and evaluated in predicting ocean parameters? RQ4 investigates optimization algorithms and evaluation metrics that include error-based measures, correlation-based metrics and computational efficiency scores. These elements provide a comprehensive understanding of model accuracy, robustness, and computational efficiency in ocean parameters forecasting.

2.2. Search Strategy

2.2.1. Search Datasets

We searched multiple databases which were likely to publish related studies. The databases searching was performed in October 2024. There were four databases included in the searching process, which were ACM Digital Library (Computer Science); Elsevier ScienceDirect (Multi-disciplinary); IEEE Xplore (Engineering Technology); and SpringerLink (Multi-disciplinary). In addition, several high-profile interdisciplinary journals might also publish related work. We selected Nature (including the affiliated journals of Nature), Science (including the affiliated journals of Science), Proceedings of the National Academy of Sciences of the United States of America (PNAS) (Multi-disciplinary). Given the long-standing academic tradition of recognizing conference papers as significant contributions in computing-related disciplines, we made two decisions, (1) keeping the conference papers retrieved from the ACM Digital Library and IEEE Xplore; (2) adding four top artificial intelligence conferences: AAAI Conference on Artificial Intelligence (AAAI), International Conference on Machine Learning (ICML), International Joint Conferences on Artificial Intelligence (IJCAI) and Annual Conference on Neural Information Processing Systems (NeurIPS), which were not covered in these databases.

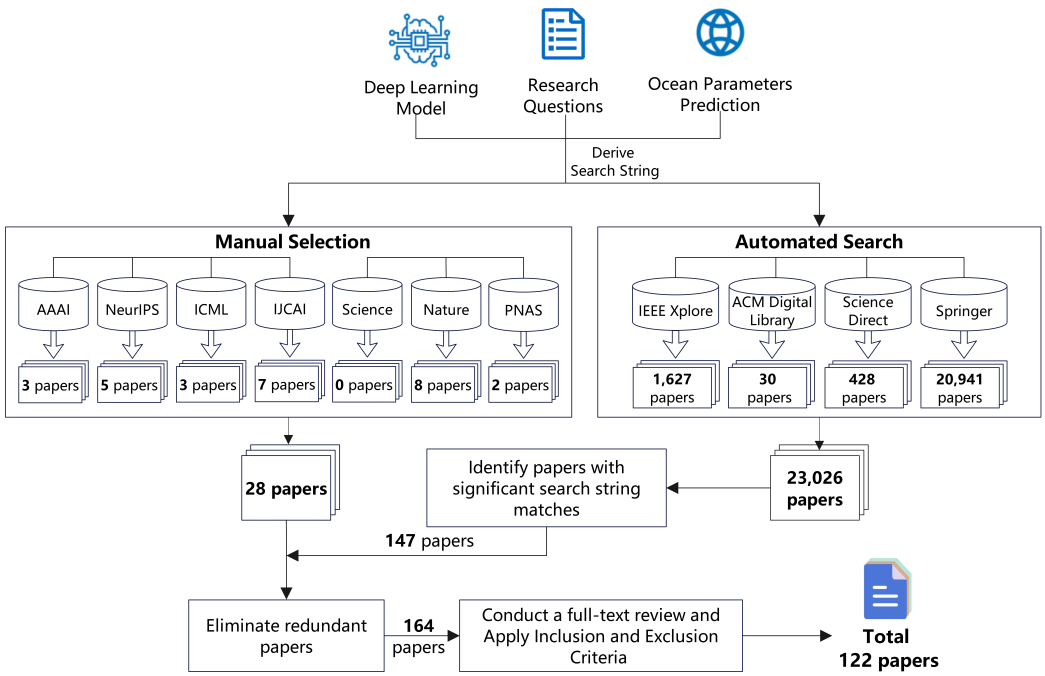
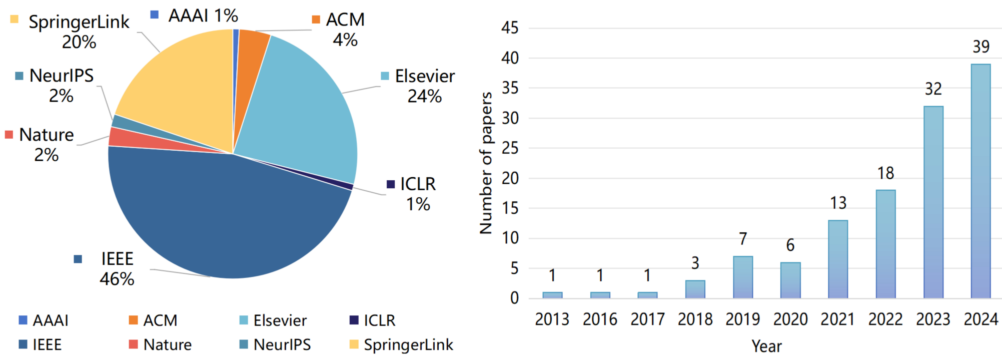


Figure 2: Study identification and selection process.

2.2.2. *Search Items*

This study aims to identify literature at the intersection of advanced computational methods and ocean parameter prediction. A targeted keyword search strategy was employed to ensure relevance, combining four categories: **Deep learning models:** Network, Deep Learning, LLM; **Oceanographic research:** Ocean, Sea, Marine; **Physical parameters:** Velocity, Current, Temperature, Salinity; **Prediction tasks:** Forecasting, Predicting, Forecast, Prediction. The complete search string is: (Network OR Deep Learning OR LLM) AND (Ocean OR Sea OR Marine) AND (Velocity OR Current OR Temperature OR Salinity) AND (Forecasting OR Predicting OR Forecast OR Prediction). Articles are deemed highly relevant only if they included at least one term from each category. This approach ensures that the selected studies specifically focus on DL applications for oceanographic parameters prediction, directly supporting the aims of this review.



(a) Distribution of papers across venues.

(b) Distribution of papers over years.

Figure 3: Distribution of papers across venues.

2.3. Study Selection

As shown in Figure 2, we employed a systematic multistage approach combining automated and manual selection. **During the automated search phase**, our search string was applied to four digital libraries, initially retrieving 23,026 candidate articles. Due to imprecise database matching, many results were only loosely relevant. To enhance precision, we implemented a custom script that strictly matched the search string in titles and abstracts, reducing the pool to 147 articles. This refined set allowed for a more targeted analysis. **For the manual search**, we reviewed seven additional sources–Science, PNAS, Nature, AAAI, ICML, IJCAI, and NeurIPS. Manual screening was necessary because these conferences lack unified search interfaces, and for the selected journals, the expected number of relevant articles was low, making manual review more efficient. This process identified 28 further papers that met our criteria.

Table 1: Inclusion Criteria and Exclusion Criteria

Inclusion criteria
(1) The article claims that a deep learning-based approach is used.
(2) The article claims that the study involves at least one of the following aspects of ocean parameter prediction, e.g., temperature, salinity, or velocity.
(3) The article with accessible full text.
Exclusion criteria
(1) A survey, thesis or book.
(2) Comparing different forecasting methods but not proposing new stuff.
(3) Duplicate articles or similar studies with different versions from the same authors.
(4) Studies belonging to books, thesis, monographs, keynotes, panels, or venues not executing a full peer-review process.
(5) Non-English written literature.

After integrating the results from both automated and manual searches, we removed duplicate entries, yielding a final pool of 164 unique articles. To ensure a rational and objective screening process, we manually applied the inclusion and exclusion criteria in Table 1. Based on this evaluation, 122 articles were selected in the final dataset.

2.4. Data Extraction and Analysis

Figure 3 summarizes the publication distribution of the selected articles. As depicted in Figure 3a, IEEE is the leading publisher, representing 46% of the articles, followed by Elsevier (24%) and SpringerLink (20%). While these three publishers dominate, the inclusion of papers from venues such as NeurIPS and Nature reflects increasing cross-disciplinary engagement. Figure 3b illustrates a clear upward trend, with publications rising from fewer than 10 papers per year before 2020 to 32 in 2023, and reaching 39 by October 2024 . This trend underscores the rapidly growing interest in leveraging deep learning for oceanographic parameter prediction.

3. RQ1: How Are Marine Datasets Categorized and Represented for Ocean Parameter Prediction?

3.1. What Are the Main Types of Marine Datasets?

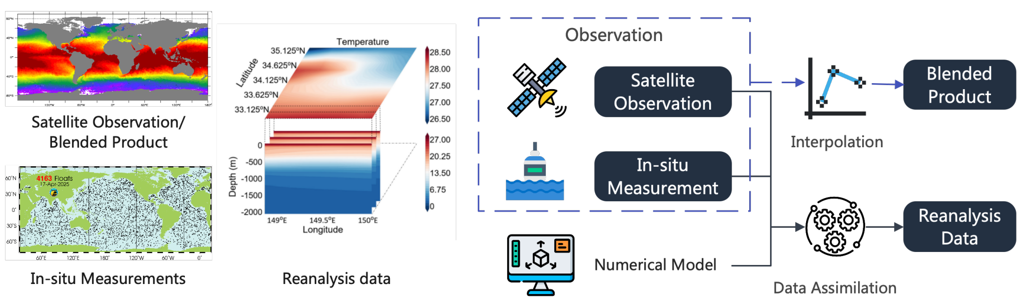


Figure 4: Overview of the main types of marine datasets and their interrelationships.

Marine datasets for ocean parameter prediction can be systematically categorized into four types based on acquisition methods and processing levels: **satellite observation, in-situ measurement, blended product, and reanalysis data**. This classification reflects both the data source and the extent of fusion or assimilation applied.

Satellite observations are derived from remote sensing instruments onboard earth-observing platforms, providing near real-time global coverage of sea surface parameters. For example, NASA's MODIS sensors aboard Terra and Aqua provide global SST observations (NASA MODIS Science Team, 2024), while ESA's SMOS mission delivers key Sea Surface Salinity (SSS) data as part of the Living Planet Programme (European Space Agency, 2025). Satellite observations offer global coverage and high temporal resolution (daily to sub-daily), enabling real-time ocean monitoring, but are constrained by cloud interference and limited to surface measurements, thereby lacking subsurface insights.

In-situ measurements refer to direct oceanic observations obtained through a network of distributed instruments, including buoys, ships, and autonomous floats. Notable examples include Argo floats, which autonomously profile temperature and salinity to 2,000m depth (Roemmich et al., 2009). PIRATA moorings, which are fixed observation platforms in the tropical Atlantic, continuously monitor air-sea parameters to support real-time prediction (Bourlès et al., 2008). In-situ measurements provide high-precision, localized observations, but suffer from sparse, uneven spatial coverage and potential temporal gaps due to deployment cost and operational constraints.

Blended products are datasets created by integrating satellite observations with in-situ measurements using statistical, mathematical, or machine learning techniques, resulting in continuous, gap-free fields optimized for operational applications. For example, NOAA's OISST (NOAA Physical Sciences Laboratory, 2025) and CMEMS's OSTIA (E.U. Copernicus Marine Service, 2025) blend satellite, ship, buoy, and Argo data into bias-corrected, gap-free daily sea surface temperature fields at 1/4° and 1/20° resolution, respectively. With consistent spatial and temporal coverage, these products are widely adopted for ocean parameter prediction. However, their surface-only scope and reliance on interpolation or multi-source fusion can smooth fine-scale patterns and introduce biases due to cross-platform inconsistencies.

Reanalysis data involves the integration of historical observations with numerical models using data assimilation techniques to produce comprehensive ocean parameter datasets. Widely used examples include ERA5 (Hersbach et al., 2020) and SODA (Carton and Giese, 2008), which provide 3D reconstructions of ocean states by assimilating multi-source observations into physical models. Reanalysis datasets encode physical relationships among ocean variables, providing continuous spatiotemporal estimates across multiple scales. However, they suffer from latency–often lagging real-time by days to weeks. Since reanalysis datasets rely on model-based data assimilation, their accuracy is shaped by both the underlying model assumptions and the quality and availability of observational inputs, particularly in sparsely sampled or rapidly evolving regions.

Figure 4 illustrates representative examples and the interrelationships among different types of marine datasets. Satellite observations and in-situ measurements serve as the primary types. These observations can be integrated using interpolation techniques to produce blended products. Furthermore, observational data may be assimilated into numerical models to generate reanalysis datasets, which provide temporally consistent and physically coherent representations of ocean states.

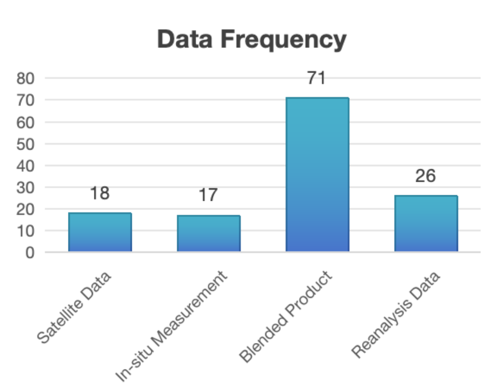
3.2. How Do Different Dataset Types Support Ocean Parameter Prediction?

Each type of marine dataset offers distinct spatial, temporal, and physical properties. To enable systematic comparison, we assess five key dimensions: **originality, spatial coverage, spatial and temporal resolution, and updated frequency**. Here, originality indicates whether the datasets are direct observed (high originality) or derived through interpolation, assimilation, or modeling (low originality). Table 2 summarizes the strengths and limitations of each data type in deep learning contexts, while Table 3 details representative datasets for practical implementation.

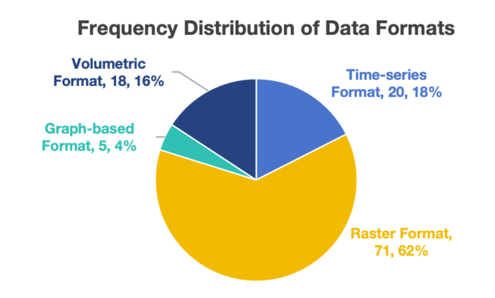
Table 2: Comparison of Different Ocean Data Sources

	Satellite Data	In-situ Measurement	Blended Product	Reanalysis Data
Originality	High	High	Medium	Medium
Spatial Coverage	Global	Regional/Local	Global/Regional	Global
Spatial Resolution	Medium-High	High	Medium	Low to Medium
Temporal Resolution	Hourly/Daily	Hourly to MultiDays	Daily/Monthly	Hourly to Monthly
Updated Frequency	Daily	Hourly to MultiDays	Daily/Monthly	Daily/Monthly

Each data type plays distinct yet complementary roles in ocean parameter prediction tasks. **Satellite observations** provide large-scale, near-real-time surface patterns, making them ideal for dynamic surface modeling, e.g., Meng et al. (2023) fuse GHRSST satellite observations with HYCOM outputs within a GAN framework, where satellite-derived physics guides the generation of enhanced SST fields. Geiger et al. (2013) combine satellite-based SST and ocean color data to estimate salinity patterns in Mid-Atlantic coastal and estuarine zones, leveraging broad spatial coverage. **In-situ measurement** provides high-accuracy, depth-resolved profiles critical for model calibration and validation. For example, Han and Hong (2023) integrate BOA-ARGO data to validate a conditional GAN for SSS prediction, while Meng et al. (2021) use Argo profiles to fine-tune a physics-guided GAN for subsurface temperature, bridging simulations with real-world observations. **Blended products** combine the strengths of satellite and in-situ data, offering gap-free, consistent datasets ideal for training DL models. Janmaijaya et al. (2024) utilize multi-source blended SST and heat content data to train a spatio-temporal dilated ConvLSTM network for long-range forecasting, while Xiao et al. (2024) leverage regional blended SST in a graph-based framework to improve coastal SST prediction. **Reanalysis data** offer physically consistent, multi-variable fields with temporal continuity and vertical structure, making them valuable for long-term forecasting. Their structured 3D representation of subsurface parameter fields enhances physical interpretability in data-driven models, e.g., Jin et al. (2021) jointly predict 3D temperature, salinity, and velocity fields using multi-source reanalysis data to capture cross-variable dependencies in ocean dynamics. Choi et al. (2023) leverage reanalysis time series to detect and predict anomalous high-temperature events near the Korean Peninsula.



(a) Frequency of different data types used. Note: A single study may incorporate multiple data types.



(b) Distribution of marine data formats. Note: Individual studies do not specify the data formats.

Figure 5: Marine data types used in ocean prediction: (a) by frequency, (b) by format.

Figure 5a shows the frequency of data usage across the surveyed studies. Blended products are the most frequently used, likely due to their comprehensive coverage and continuity. Reanalysis datasets follow, favored for their consistency and utility in historical trend analysis. However, integrating real-time satellite or in-situ measurements remains more challenging due to their noise, gaps, and limited depth coverage–yet essential for improving prediction timeliness and adaptability.

3.3. How Are Marine Datasets Represented for DL Prediction?

Marine data in ocean prediction research spans multiple formats, each tailored to specific tasks. These formats can be categorized into four formats: time-series, raster, graph-based, and volumetric data.

Time-series format tracks the temporal evolution of parameters at a single point, across multiple stations, supporting trend analysis and seasonal forecasting. For example, Hittawe et al. (2024) apply an ensemble transformer to predict time-series SST in the Red Sea. Tresnawati et al. (2024) employ multi-source satellite time-series data to predict long-term SST in the Indonesian seas.

Raster format encodes 2D discrete spatial grids, typically derived from satellite observations or blended products. This data format is widely used for surface ocean patterns. Xie et al. (2019) introduce an attention-based CNN model tailored for SST raster data to adaptively capture multi-scale spatial features. Xie et al. (2021) introduce a convolutional GRU model with MLP to model spatiotemporal dependencies in the Bohai Sea SST grids.

Graph-based format treats spatially distributed data as nodes in a graph, capturing spatial dependencies through edge structures. This format is increasingly used by integrating satellite data with land-ocean masks and in-situ float data. Zhang et al. (2021b) build SST graphs from satellite grids and proposes a memory-augmented GCN for spatiotemporal SST forecasting. Kim et al. (2023) dynamically build adjacency matrices from in-situ observations to train a spatiotemporal GNN for multi-step SST prediction.

Volumetric format extends surface raster data by incorporating the vertical (depth) dimension, forming 3D fields of temperature, salinity, and velocity, typically derived from reanalysis products and numerical simulations. Zuo et al. (2021) design a 3D CNN for learning deep spatial-temporal-vertical features in temperature fields. Zhang et al. (2020) apply a multilayer ConvLSTM to capture vertical structure and temporal evolution.

Figure 5b illustrates the frequency distribution of various data formats employed in ocean prediction research. Raster format accounts for the majority (62%), reflecting the prevalence of gridded satellite or blended datasets in spatially resolved surface modeling. Time-series (18%) and volumetric data

Table 3: Overview of commonly used marine data products categorized by data type.

	Product	Data Source	Variables	Spatial Resolution	Spatial Coverage	Temporal Resolution	Temporal Extent	Update Frequentcy
Satellite	MODIS	NASA	SST	4km	Global	Daily	1999 - present	Daily
	AVHRR	NOAA	SST	1.1km	Global	Daily	1981 - present	Daily
	SMOS	ESA	SSS	35–50 km	Global	Daily	2009–present	Daily
In-situ	Argo	Argo Program	T, S	Points	Global	10 days	2000–present	10 Day
	ICOADS	NOAA/NCAR	T, S	Points	Global	Daily-Monthly	1662–present	Daily
	WOD	NOAA/NCEI	T, S	Points	Global	Variable	1700s–present	Annual
	NDBC	NOAA	T, S	Points	Regional	Hourly	1970s–present	8 Hourly
	PIRATA	NOAA/INPE /IRD	T, S	Points	Tropical Atlantic	Hourly–Daily	1997–present	Real-time
	TOGA	NOAA/PMEL	T, S	Points	Tropical Pacific	Daily	1980s–present	Daily
Reanalysis	ORAS5	ECMWF	T, S, V	0.25deg	Global	Monthly	1958–present	Monthly
	ERA5	ECMWF	SST	0.25deg	Global	Hourly	1940–present	Daily
	HYCOM	US Navy/NRL	T, S, V	0.08deg	Global	3 Hourly	1994–2015	Daily
	CORTA	Chiese NMDC	T, S, V	0.1deg	WesternNorth Pacific	Daily	1990–present	Daily
	GLORYS12V1	Mercator Ocean	T, S, V	0.083deg	Global	Daily	1993–present	Monthly
	SODA	U. Maryland	T, S, V	0.25-0.5deg	Global	Monthly	1871-2010	N/A
	GODAS	NOAA	T, S, V	0.3x1deg	Global	Monthly	1980s–present	Monthly
Blended	OISST v2	NOAA	SST	0.25deg	Global	Daily	1981–present	Daily
	OSTIA	UK Met Office	SST	0.05deg	Global	Daily	1985–present	Daily
	BOA-Argo	CSIO, MNR	T,S	0.5deg	Global	Monthly	2004–present	Yearly
	Microwave OISST	RSS	SST	0.25deg	Global	Daily	2002-present	Daily
Numerical Model	ROMS	Community	T, S, V	1-10km	Regional	Hourly-Daily	Custom	Custom
	NEMO	ECMWF	SST	0.1-0.25deg	Global	6-hourly	Real-time	Real-time
	HYCOM Prediction Model	NECP/NRL /NOAA	T, S, V	1/12deg	Global	3 Hourly	2011-present	Real-time

(16%) are moderately used, typically in temporal forecasting or 3D subsurface modeling tasks. In contrast, graph-based formats are rarely employed (4%), suggesting their potential remains underutilized in current literature.

4. RQ2: How Are Marine Datasets Utilized across Spatial, Temporal, and Input-Level Patterns within DL Models?

To address RQ2, we conduct three levels of analysis to uncover imbalances in current ocean prediction research. First, we examine the geographic distribution of study areas, revealing significant regional disparities. Second, we analyze the temporal scope of dataset usage, highlighting a skew toward either short-term applications or long-term historical analyses. Third, we review the input parameter choices and combinations, identifying patterns that may influence model performance and generalizability.

4.1. What Are the Spatial Patterns within DL Models?

To investigate the geographical distribution of data utilization in current research, we analyze 122 research papers, categorizing marine regions based on the four major ocean basins: Atlantic, Pacific, Indian, and Arctic Oceans. Figure 6 illustrates the distribution of study areas across major oceanic regions. The analysis reveals a strong focus on the Pacific and Indian Oceans, with comparatively fewer studies in the Arctic, Atlantic, and deep-sea Pacific regions. This regional bias likely stems from the prominent role of currents like the Kuroshio and Agulhas in driving large-scale circulation and influencing regional climate and marine ecosystems (Lutjeharms et al., 1989).

We further examine the subregional geographical distribution of study areas. The results show that despite the availability of global datasets, global-scale research remains scarce due to the computational cost and efficiency constraints. In the Pacific Ocean, research is concentrated in the northwest and coastal regions. The regions with the highest number of studies are coastal areas, including the South China Sea, the Bohai Sea, and the East China Sea, driven by the North Pacific Gyre and their economic importance (Di Lorenzo et al., 2008). In the Indian Ocean, the Bay of Bengal is a key research focus due to its susceptibility to monsoon-driven changes in temperature and salinity, as well as the seasonal variability of ocean currents, which significantly affect coastal ecosystems and human livelihoods. In contrast, research in the Atlantic Ocean is more dispersed, covering subregions such as the Moroccan Atlantic Coast, Caribbean Sea, and South Atlantic.

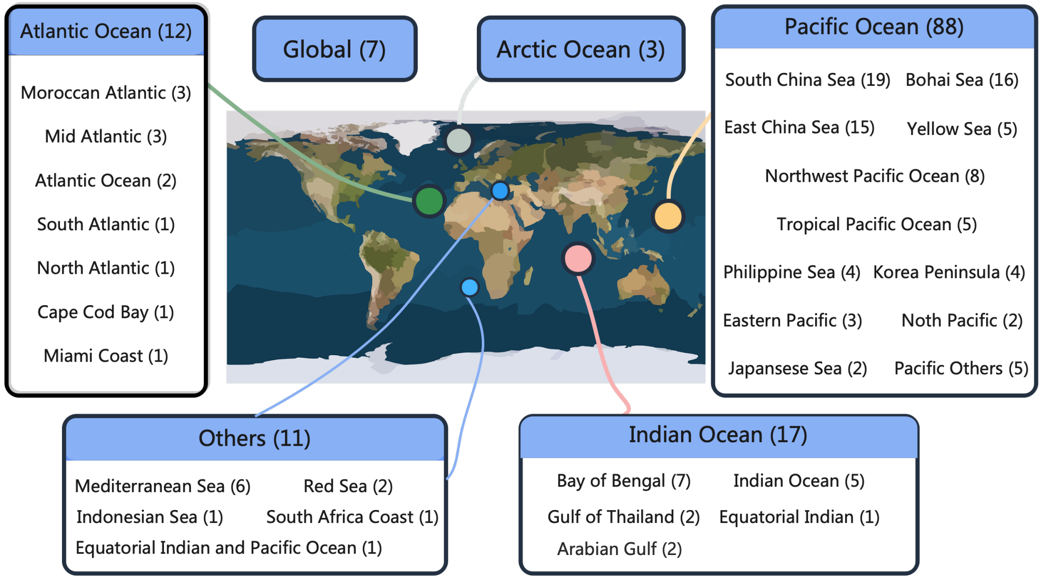


Figure 6: The statistics of study areas across different oceanic regions in the surveyed research papers. Note: One study may involve multiple study areas.

Marine regions like the Mediterranean Sea remain understudied, yet their subsurface dynamics and frequent heatwaves make them vital for ocean parameter prediction meeting national needs (Dayan et al., 2023). Meanwhile, polar regions, especially the Arctic, are the least studied, likely due to harsh environmental conditions and data accessibility challenges.

Beyond the major ocean basins and subregions, extending this analysis to ecological and geographical functional zones in coastal, open-ocean, tropical, and polar regions also reveals similarly substantial imbalances. Coastal areas in the Pacific and Indian Oceans are well-studied, while open-ocean regions like the central Atlantic and wider Pacific remain underexplored despite their role in large-scale heat transport (Bhagtani et al., 2024; Sutton et al., 2024). Tropical regions in the Pacific and Indian Oceans receive attention, but the South Atlantic and equatorial zones are notably underrepresented, highlighting the need for more balanced geographic coverage to close data gaps and support robust ocean forecasting.

4.2. What Are the Temporal Patterns within DL Models?

Our analysis of temporal patterns in ocean parameter prediction studies highlights differences in data resolution and time span across forecasting horizons. These studies are commonly grouped into four regimes: **real-time** (nowcasting to a few hours ahead), **short-term** (1–7 days), **subseasonal** (10–30 days), and **long-term** (from months to decades). Each temporal window aligns with different oceanic phenomena and thus requires tailored modeling strategies and data characteristics.

Most existing work focuses on **short-term forecasting**, often using daily-resolution data spanning 5–10 years to predict SST and surface velocity. This window is well-suited for capturing mesoscale dynamics such as eddies and coastal upwelling (Wyatt et al., 2023), with GHRSST SST and CMEMS velocity fields widely adopted in regional models (Chen and Dong, 2019). In contrast, **long-term predictions** target capturing low-frequency variability and climate-driven trends, often using monthly reanalysis datasets spanning multiple decades. For example, studies leveraging the ECMWF ORAS5 reanalysis have demonstrated improved skill in forecasting seasonal to interannual SST evolution (Ye et al., 2023). **Subseasonal prediction** remains underdeveloped, despite its crucial role in early warnings of marine heatwaves and anomalous current patterns. This forecast window poses unique challenges, as it requires capturing both fast evolving synoptic variability and slower intraseasonal signals (e.g., the Madden-Julian oscillation) (Zhang and Dong, 2004), which require dense observations and stable hybrid modeling frameworks. **Real-time prediction** is similarly constrained primarily due to scarce low-latency, high-resolution data and the high cost of frequent model updates. The challenge is particularly pronounced for velocity fields, which rely on indirect measurements like thermal marker velocimetry (Veron et al., 2008), leading to sparse and delayed availability.

These findings highlight not only the temporal imbalance in current research efforts but also the pressing need for integrative approaches that combine high-resolution, real-time satellite data with long-term historical records, thereby enhancing prediction skills across multiple timescales.

4.3. What Are the Input Feature Patterns within DL Models?

In parameter prediction, DL models utilize varied input configurations to address the complex, nonlinear dynamics of the marine environment. These configurations can be generally classified as: **Single-task, Single-variable; Single-task, Multi-variable; and Multi-task, Multi-variable**.

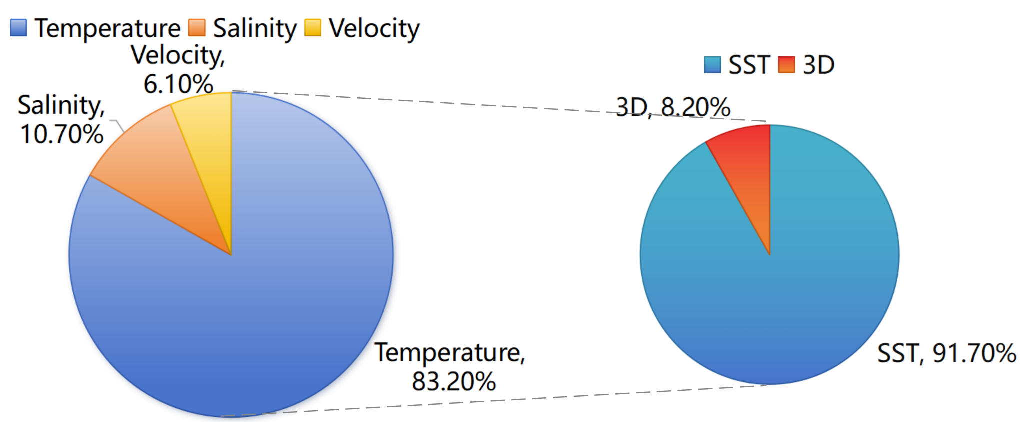
Single-task, Single-variable. In this framework, models predict a single target variable (e.g., SST) based solely on its historical data. Inputs typically include raw observations or features derived from the target variable, such as anomalies, seasonal cycles, or periodic signals (e.g., tropical instability waves). Although such features may undergo mathematical transformation, they remain within the same variable domain. Spatiotemporal inputs can be structured to capture temporal evolution (e.g., sliding time windows) and spatial patterns (e.g., gridded data with geographic or depth information). This approach balances simplicity with representational capacity, making it suitable when the target variable's own dynamics are the primary drivers of prediction accuracy.

Single-task, Multi-variable. Here, prediction focuses on a single target variable using multiple related environmental parameters as inputs. Incorporating interdependent variables–such as sea surface height anomalies, sea surface wind, air temperature, and atmospheric pressure–provides a more holistic view of the marine system, typically enhancing predictive performance. Advanced strategies may integrate diverse datasets (e.g., combining satellite and in-situ measurements) through data fusion, thereby capturing complex spatiotemporal dependencies.

Multi-task, Multi-variable. This framework involves the simultaneous prediction of several oceanographic parameters using a comprehensive set of input features. These models capture the physical and dynamical couplings among variables by integrating multiple environmental parameters (e.g., temperature, salinity, velocity components). Spatiotemporal features, such as gridded, time-series, or graph structures, are employed to reflect spatial variability and temporal dynamics. Data fusion from sources like satellite imagery, in-situ sensors, and reanalysis products further enhances the model's capacity to represent both short-term and long-term ocean processes.

5. RQ3: Which DL Models Have Been Applied to Predict Three Different Ocean Parameters?

To elucidate the current state of DL applications in ocean parameter prediction, we first examine the distribution of studies across three core prediction tasks. As depicted in Figure 7, temperature-related prediction tasks overwhelmingly dominate, comprising 83.2% of surveyed articles. In comparison, salinity and velocity prediction are notably less represented, accounting for only 10.7% and 6.1% of studies, respectively. Within temperature-focused research, Sea Surface Temperature (SST) is the primary target (91.7%), with 3D temperature prediction constituting just 8.2%. Given the limited num-



(a) Distribution of articles related to temperature, salinity and velocity.

(b) Distribution of SST and 3D temperature prediction within temperature-related tasks.

Figure 7: Distribution of research articles related to three fundamental ocean parameters.

ber of studies on salinity and velocity, no further subdivision into sea surface and 3D tasks is conducted. This distribution highlights the field's predominant emphasis on SST prediction. Based on this task-oriented overview, we categorize the DL approaches into six main groups: **MLPs, CNNs, RNNs/LSTMs, GNNs, GANs, and Transformers**. We classify the 122 studies (see supplementary material) by publication year and model category (Figure 8) to systematically examine research trends.

5.1. DL Models in Sea Temperature Prediction

Sea temperature prediction is typically divided into two subcategories: SST prediction and 3D temperature prediction. While SST prediction has received the majority of attention due to the abundance of satellite data, 3D temperature prediction addresses the more complex challenge of modeling subsurface and volumetric thermal structures. In the following sections, we review how various categories of DL models have been applied to these two types of temperature prediction tasks.

5.1.1. Sea Surface Temperature Prediction

MLPs constitute a fundamental class of deep learning models inspired by biological neural networks. Comprising an input layer, one or more hidden

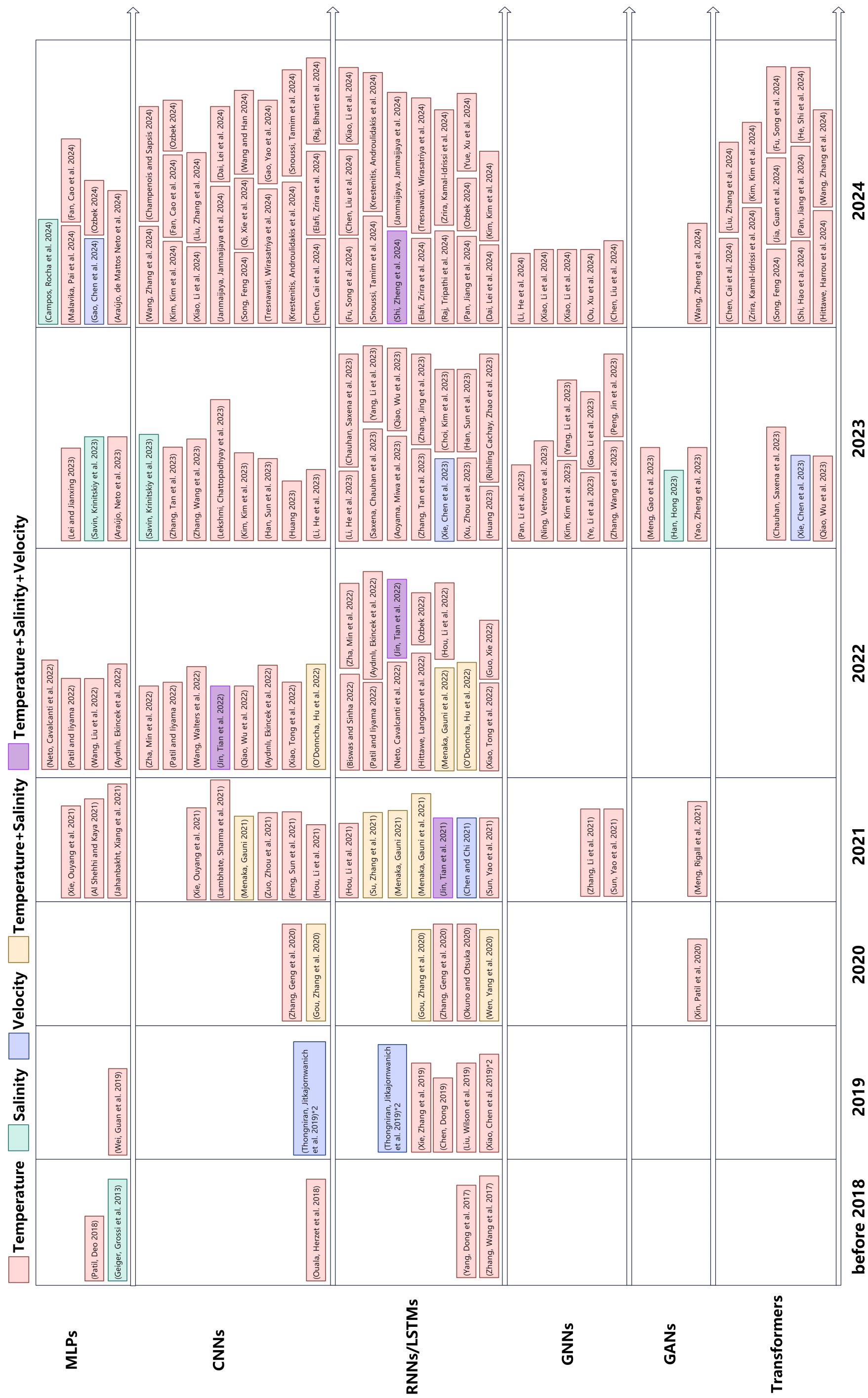


Figure 8: Distribution of the DL-based models applied in the collected articles.

layers, and an output layer, MLPs are capable of approximating nonlinear relationships through iterative weight optimization. In SST prediction, MLPs account for approximately 15.4% of reported models. Although less prevalent than recurrent architectures such as RNNs and LSTMs, MLPs are useful in cases where data exhibit relatively simple spatial or temporal patterns. MLPs are often chosen when simplicity and efficiency are priorities. For example, Patil and Deo (2018) apply MLPs to basin-scale SST forecasting, achieving strong performance in short-range predictions (1–3 days) and acceptable accuracy for medium-term horizons (4–7 days). In Wei et al. (2019), MLPs are employed to forecast SST using the OSTIA dataset by separately modeling climatological monthly means and anomalies with two neural networks and aggregating their outputs. This method achieves an average bias of -0.16°C and a standard deviation of 0.37°C between predicted and observed values, demonstrating the viability of MLPs for regional SST forecasting. Collectively, these studies establish MLPs as efficient baseline models, particularly advantageous for real-time SST prediction tasks.

CNNs account for approximately 38.1% of SST prediction. By using convolution and pooling layers, they efficiently capture spatial hierarchies from gridded data. Recent advances have extended CNNs to temporal domains for long-term forecasting. For example, a TCN-based model successfully predicts monthly SST in the Indian Ocean with a RMSE of 0.506°C over five years (Feng et al., 2021). A 3D CNN-LSTM model with attention captures both spatial and temporal dependencies, achieving higher accuracy with lower complexity (Qiao et al., 2022).

RNNs/LSTMs account for approximately 55.5% of SST prediction applications due to their strength in handling sequential data. LSTMs, with their gated memory cells, overcome the vanishing gradient issue of traditional RNNs and capture long-term temporal dependencies. LSTM-based models have successfully predicted SST across short and long timescales, including daily to monthly averages (Zhang et al., 2017). Integrating spatial correlations further improves SST forecasts in regional seas (Yang et al., 2017), and hybrid LSTM–AdaBoost frameworks help reduce overfitting while preserving predictive power (Xiao et al., 2019a). Overall, RNN/LSTM-based methods robustly capture SST's complex spatiotemporal dynamics.

GNNs account for about 13.6% of SST prediction applications. By learning node and edge features, they capture complex spatiotemporal dependencies. For example, MGCN integrates temporal and spatial components, outperforming traditional methods on SST datasets (Zhang et al., 2021b).

Likewise, TSGN uses LSTMs and graph structures to improve long-term SST forecasts (Sun et al., 2021). These examples highlight GNNs' effectiveness in ocean temperature modeling.

GANs consist of a generator and discriminator that compete to synthesize realistic data, making up about 4.5% of SST prediction applications. These models are particularly effective when combined with domain knowledge or advanced architectural designs. For example, a physics knowledge-enhanced GAN integrates observed data into numerical models to improve SST forecasts, significantly outperforming existing baselines (Meng et al., 2023). Similarly, Wang et al. (2024a) propose a DGAN model that utilizes composite generator layers, a patchGAN-based discriminator, and an improved loss function to generate future SST maps with higher reliability and accuracy. These studies demonstrate the potential of GANs to enhance both the accuracy and interpretability of SST forecasting.

Transformers account for about 11.7% of SST applications and leverage self-attention to capture long-range temporal dependencies. For example, StackPred (Hittawe et al., 2024) uses an ensemble Transformer with multiscale denoising to predict Red Sea SST, achieving R^2 of 99.83 and outperforming LSTM, GRU, and baseline Transformers. TL-iTransformer (Jia et al., 2024) adapts pre-trained models to data-sparse regions, attaining lower MAE and MSE than other time-series baselines. These examples show that Transformer-based methods can effectively model SST across varying spatial scales and levels of data availability.

5.1.2. 3D Temperature Prediction

MLPs are also used for 3D ocean temperature forecasting, often in combination with decomposition techniques. For example, the FDL model (Lei and Jianxing, 2023) applies Tucker decomposition to separate stable and dynamic components,followed by incremental learning and residual correction. Likewise, a hybrid ML approach (Malavika et al., 2024) combines empirical mode decomposition with parameter selection to improve Arabian Sea subsurface temperature predictions. These examples highlight that lightweight MLPs can perform well when paired with effective feature extraction.

CNNs are increasingly used for 3D ocean temperature forecasting due to their strength in capturing spatial features. For example, Liu et al. (2024) integrate CNNs with attention and residual layers to process inputs such as SST, salinity, sea surface height, and ocean velocity, aiming to predict sub-surface temperature variations in the South China Sea over a two-month horizon. The SST-4D-CNN model (Zuo et al., 2021) uses four-dimensional convolutions to capture both vertical and horizontal temperature variations, achieving over 98% accuracy on real-time ocean data.

RNNs/LSTMs are popular for 3D ocean temperature forecasting due to their ability to capture temporal dependencies. For example, SA-PredRNN (Yue et al., 2024) enhances long-range spatiotemporal modeling with self-attention. The M-convLSTM model(Zhang et al., 2020) combines CNNs and stacked ConvLSTM layers to capture both horizontal and vertical temperature variations, enabling accurate 3D ocean temperature prediction from the surface to 2000 meter depth using ARGO data.

GNNs are increasingly applied to 3D ocean temperature forecasting due to their flexibility in capturing irregular spatial dependencies. For example, ASTDG (Pan et al., 2023) combines adaptive graph learning with graph and temporal convolutions to model latent spatiotemporal relationships without predefined adjacency. Ou et al. (2024) integrate attention into the graph pipeline to further improve performance.

Transformers are emerging as powerful tools in 3D temperature prediction for capturing long-range dependencies. Liu et al. (2024) employ a 4D-OT-ATTN-Conv architecture that combines Transformer layers with convolutional structures to achieve improved performance in predicting deep ocean temperature.

5.2. DL Models in Sea Salinity Prediction

Sea salinity prediction is a relatively underexplored area in the field of ocean parameter forecasting. Among the various modeling approaches, RNNs–particularly LSTMs–are the most commonly used, accounting for approximately 61.3% of the models applied in salinity forecasting tasks. For instance, in Wen et al. (2020), the authors propose a semi-supervised prediction framework that combines an improved unsupervised clustering algorithm with a supervised LSTM network to perform time-series forecasting. This hybrid model is designed to address issues such as data deviation and redundancy typically found in marine datasets. The remaining studies primarily utilize CNNs, often in combination with RNNs, to enhance spatial-temporal feature extraction in sea salinity prediction. A representative example is the model proposed in Menaka and Gauni (2021), which integrates Multi-Variant Convolution (MVC), and a High-Speed Long Short-Term Memory (HM-LSTM) network. Unlike traditional time-series models that overlook spatial dependencies, this framework captures both horizontal and vertical parametric variations up to 2000 meters below the surface by leveraging spatial-temporal dependencies. Apart from the aforementioned methods based on RNNs and CNNs, a number of recent studies have adopted GANs for salinity prediction. For example, in Han and Hong (2023), the authors propose a conditional GAN-based data-to-data translation framework to retrieve and predict sea surface salinity. By leveraging satellite-based polarized brightness temperatures, sea surface temperature, and sea surface wind speed as conditional inputs, the model effectively learns the nonlinear mappings required for high-resolution salinity prediction.

5.3. DL Models in Sea Velocity Prediction

Sea velocity prediction has received even less attention compared to other ocean parameters. The most commonly used models for sea velocity prediction are the RNNs/LSTMs, which account for 87.5% of all models employed. For example, Chen and Chi (2021) proposes a novel model, STA-GRU, which integrates a spatial module to extract spatial features and an attention mechanism to capture nearest-neighbor temporal correlations, all built upon a GRU backbone. Applied to ocean velocity prediction, this model demonstrates enhanced accuracy in handling large-scale and complex ocean datasets. The remaining models include MLPs, which are applied in specific cases where simple architectures are sufficient, accounting for 12.5% of the models used. For instance, Gao et al. (2024) utilize High Frequency Surface Wave Radar sea state inversion data and proposes a Genetic Algorithm-optimized Backpropagation Neural Network (GA-BP) to forecast ocean current speeds during typhoons. Results show that while CNN, LSTM, and BP models exhibit MAPE values exceeding 20%, the GA-BP model achieves a MAPE below 20%, indicating moderate to high predictive accuracy.

5.4. Trend Analysis of Applied DL Models

The integration of deep learning into ocean parameter prediction research has undergone a transformative journey over the past decade, mirroring the broader adoption of artificial intelligence in scientific disciplines. Figure 9 illustrates the evolving landscape of DL model applications in three fundamental oceanographic prediction tasks from 2013 to 2024.

Evolution in 2013–2020. Between 2013 and 2019, the adoption of deep learning models in oceanographic prediction was limited. While AlexNet's 2012 breakthrough in image classification spurred rapid advances in computer

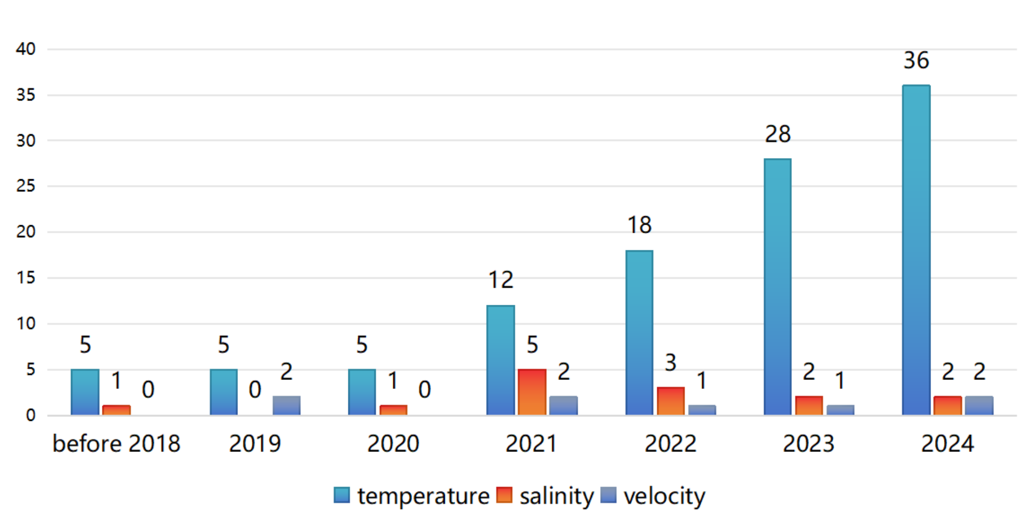
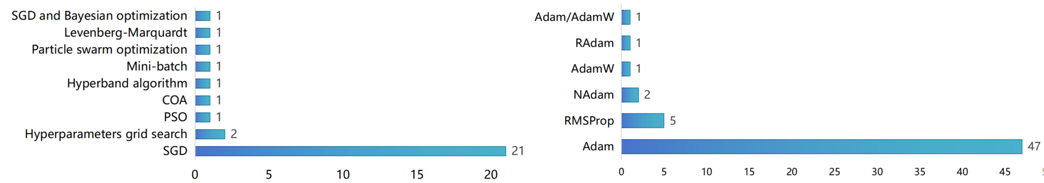


Figure 9: Trends of articles on deep learning models applied in three fundamental oceanographic prediction tasks over the years.

vision and related fields, oceanography saw only gradual uptake. Prior to 2018, just two studies employed MLPs and one used a CNN. Interest began to grow modestly in 2019, with six studies incorporating RNNs/LSTMs, and by 2020, applications of diverse DL models expanded. This slow early adoption reflects challenges in adapting general DL architectures to the high-dimensional, nonlinear, and spatiotemporal nature of oceanographic data.

Diversification in 2021–2022. The field diversified substantially in 2021 and 2022. RNNs/LSTMs remained prominent–appearing in 10 articles in 2021–reflecting their strength in sequence modeling. CNNs and MLPs also gained traction, each used in six studies that year. This period of growth was supported by increased computational resources and wider availability of DL frameworks. Notably, GNNs made their first appearance in oceanography in 2022, paralleling their broader adoption in computer science. Their introduction reflects a growing interest in explicitly modeling spatial dependencies and accommodating irregular sensor networks, such as buoy arrays.

Flourishing in 2023–2024. By 2023, RNNs/LSTMs led with 15 studies, and CNNs continued their strong presence. The upward trend persisted in 2024, with CNNs and RNNs/LSTMs rising to 19 and 16 articles, respectively. Transformers, introduced in 2017 and widely adopted in NLP by 2019, began appearing in oceanographic research in 2023. Their adoption



(a) Various parameter optimizers used in primary studies.

(b) Different learning rate schedulers used in primary studies.

Figure 10: Distribution of optimization algorithms used in primary studies

accelerated in 2024, with 12 articles leveraging their capacity for modeling long-range temporal dependencies and global spatial attention–abilities critical for large-scale ocean dynamics. GNNs and MLPs maintained steady but moderate usage (5 and 6 articles, respectively, in 2024), complementing other model types based on data characteristics and prediction tasks. Overall, the recent trajectory indicates a flourishing field converging toward architectures adept at capturing both spatial (CNNs, GNNs, Transformers) and temporal (RNNs, LSTMs, Transformers) dependencies, in line with the intrinsic properties of oceanographic processes.

Criteria for Model Selection. The choice of deep learning architecture for oceanographic prediction is shaped by the need to capture complex spatiotemporal dynamics. RNNs/LSTMs and CNNs are often preferred for their respective strengths in modeling temporal and spatial patterns. GNNs are increasingly adopted for their ability to represent relational structures, such as grid points and heat transfer pathways in prediction, enabling more accurate simulation of physical processes within the marine environment.

6. RQ4: How Are These DL Models Optimized and Evaluated in Predicting Ocean Parameters?

6.1. What Learning Algorithms Have Been Used to Optimize the Models?

The performance of ocean parameter prediction models is heavily influenced by optimization algorithms. Based on our analysis of the primary studies, these optimization techniques can be broadly divided into two categories. The first category includes optimizers that guide the convergence of deep learning models by updating model parameters during training. The second category encompasses learning rate schedulers that adjust the learning rate throughout training.

As illustrated in Figure 10, the distribution of optimizers exhibits clear patterns. Stochastic Gradient Descent (SGD) emerges as the most widely adopted, accounting for approximately 80.77%. Its dominance underscores its critical role in iteratively adjusting model parameters to ensure stability and convergence when processing complex oceanographic data. Grid search for hyperparameter tuning ranks second at 7.69%, while nature-inspired algorithms, including Particle Swarm Optimization and Coati Optimization Algorithm, each account for only 3.85%. For learning rate scheduler, Adam overwhelmingly dominates, comprising 87.04%. Its adaptive mechanism–adjusting learning rates for individual parameters based on gradient moments–is well-suited to the spatiotemporal complexity of oceanographic datasets. RMSProp follows with approximately 9.26%, while NAdam accounts for 3.7%. The minimal adoption of alternatives such as AdamW, RAdam, and hybrid variants (e.g., Adam/AdamW) further highlights Adam's superiority.

6.2. What Evaluation Metrics Have Been Used to Evaluate the models?

DL-based ocean parameter predictions are usually assessed with three main types of metrics: error-based, correlation-based, and efficiency scores. **Error-based metrics** quantify differences between predictions and observations. Common examples include RMSE, MAE, MAPE, NRMSE, and NSE. Combinations like RMSE+MAE+MAPE highlight both average and extreme errors. **Correlation-based metrics** assess the alignment of predicted and observed trends, often using Pearson's or Spearman's coefficients and R^2. For example, Xiao et al. (2019a) use Pearson's r with RMSE and MAE to evaluate SST predictions. **Efficiency scores** report computational cost and speed, such as training time, number of parameters, FLOPs, and FPS. For instance, REL-STPN (Shi et al., 2024) achieves inference speeds exceeding 1000 FPS, showing its suitability for real-time forecasting.

7. Challenges and Opportunities

7.1. Challenges

7.1.1. Challenges in Generalization

Generalization refers to a model's capacity to effectively apply learned patterns from a specific dataset to unseen datasets or new environmental conditions. In oceanographic parameter prediction, achieving generalization is particularly challenging due to the significant spatial and temporal variability inherent in ocean systems. Models trained on data from specific regions, time periods, or seasonal conditions often struggle to maintain performance when transferred to other geographical locations or temporal contexts. For instance, a model trained to forecast SST in the tropics may not perform as well when applied to polar regions due to the difference in environmental conditions. This difficulty is further amplified by the dynamic and non-stationary nature of oceanographic phenomena. Parameters such as ocean currents, temperature, and salinity exhibit substantial variability driven by seasonal cycles. Consequently, models that do not explicitly account for these variations may experience degraded performance when applied outside their training domains. Improving generalization requires training on diverse, multi-source datasets that span wide spatial and temporal ranges, as well as mechanisms that capture temporal dependencies and long-term trends. While conventional DL models struggle to meet these requirements, emerging large-scale pretrained models may offer new possibilities.

7.1.2. Challenges in Reproducibility

In the field of DL, reproducibility is critical for validating research findings and advancing scientific knowledge. Reproducibility refers to the ability to independently verify results by adhering to the original methodology, including the use of identical datasets, code implementations, and experimental protocols. However, in oceanographic research, the lack of transparent code sharing undermines this principle. While many studies publicly release their datasets, essential components such as model architectures, data preprocessing scripts, and training parameters are often withheld. This opacity prevents the scientific community from rigorously testing the validity of published results, as even minor deviations in implementation details can lead to inconsistent outcomes. Consequently, the inability to reproduce experiments slows the integration of DL technologies into oceanographic applications.

7.1.3. Challenges in Interpretability

In oceanographic parameter prediction, a key interpretability challenge lies in the lack of embedded physical knowledge within DL models. Ocean systems are governed by well-established physical principles, including temperature gradients, salinity-driven density stratification, and the dynamic influence of currents. Traditional numerical models are built upon these laws, ensuring that predictions remain consistent with physical reality. In contrast, DL models are typically data-driven and operate without explicit physical constraints. This absence raises concerns about the scientific reliability and interpretability of their predictions.

7.2. Opportunities

7.2.1. Foundation Models Design

The rise of Foundation Models presents promising opportunities for oceanographic parameter prediction. Trained on massive and diverse datasets, these models can learn hierarchical, multi-scale representations, making them well-suited for the spatiotemporal complexity of ocean systems. **Generalization Across Time and Space**: While traditional DL models struggle with generalization across spatial and temporal domains, foundation models pre-trained on large-scale, heterogeneous ocean datasets (e.g., reanalyses, satellite observations, Argo floats) can better capture transferable patterns, enabling broader applicability across diverse oceanographic contexts. **Few-Shot Learning**: Owing to their strong pre-training, foundation models can be fine-tuned with minimal labeled data–an advantage particularly important in data-sparse regions like the Southern Ocean. This few-shot learning capability addresses the limitations of conventional supervised models, which often require large annotated datasets to achieve acceptable performance in new domains. Despite these promising directions, the successful application of foundation models in oceanography will require careful consideration of several factors: the availability of sufficiently large and diverse pretraining datasets, the development of domain-specific architectures, and the management of the substantial computational resources required for training. Nonetheless, in an era increasingly defined by data richness, foundation models present a powerful new framework for advancing our capacity to model and predict complex ocean systems.

7.2.2. Multi-Source Data Fusion

Oceanographic systems are governed by a complex interplay of factors–including temperature, salinity, atmospheric forcing, marine ecosystems, and human activities. Effectively fusing multi-source data presents a critical opportunity to improve ocean parameter prediction. By integrating diverse data modalities–such as satellite remote sensing, in-situ sensors, and historical archives–models can leverage richer, more comprehensive inputs to enhance prediction accuracy. Recent advances in data fusion, such as deep multi-modal learning (Li et al., 2022a), spatiotemporal graph-based fusion

(Jiang et al., 2024), and embedding-driven feature integration (Li et al., 2022b), offer promising solutions for combining heterogeneous sources.

7.2.3. Cross-Disciplinary Collaboration

Achieving advancements in ocean parameter prediction requires close collaboration between oceanographers and experts in computer science. Oceanographers bring deep domain knowledge of marine systems, while computational scientists contribute advanced tools for managing complex datasets and building sophisticated models. This interdisciplinary synergy facilitates the development of deep learning methods capable of capturing intricate oceanographic patterns, which traditional methods may overlook. Importantly, such collaboration also paves the way for integrating physical constraints into DL frameworks. Oceanographers can help identify essential physical laws–such as conservation of mass, momentum, and energy–that must be preserved, while DL researchers can encode these constraints through physics-informed architectures, regularization terms, or hybrid modeling strategies. This integration improves model interpretability, enhances trustworthiness, and ensures scientific consistency in prediction outcomes.

8. Conclusion

In this study, we conduct a comprehensive systematic literature review on the application of DL techniques for predicting three fundamental ocean parameters: temperature, salinity, and velocity. By analyzing 122 peer-reviewed papers published between 2013 and 2024, we provide a structured overview of how DL models are used in this domain, identify gaps in current research, and suggest promising directions for future development.Our analysis reveals significant progress in leveraging models such as LSTMs, CNNs, and more recent architectures like GNNs and Transformers. However, the research remains heavily skewed toward temperature prediction, with much less attention paid to salinity and velocity. Spatially, most studies are concentrated in coastal and tropical zones, while deep-ocean and polar areas remain substantially underexplored. Methodologically, issues such as data sparsity, generalization, and model interpretability remain major bottlenecks.

Looking ahead, the advancement of ocean science will rely on unified multi-source data fusion frameworks capable of integrating satellite, in-situ, and reanalysis data–enabling DL models to leverage diverse and complementary information. Equally critical is the incorporation of physical oceanographic principles into model architectures, promoting hybrid architectures that balance empirical performance with physical consistency. Meanwhile, foundation models–large-scale pre-trained networks–offer transformative potential. When trained on extensive oceanographic datasets, they provide powerful, adaptable backbones for a wide range of downstream tasks.

References

Ahmad, H., 2019. Machine learning applications in oceanography. Aquatic Research 2, 161–169.

Aldini, I., Permanasari, A.E., Hidayat, R., Ramdhan, A., 2024. Prediction of ocean surface current: Research status, challenges, and opportunities. a review. Ocean Systems Engineering 14, 85–99.

Berliner, L.M., Wikle, C.K., Cressie, N., 2000. Long-lead prediction of pacific ssts via bayesian dynamic modeling. Journal of climate 13, 3953–3968.

Bhagtani, D., Hogg, A.M., Holmes, R.M., Constantinou, N.C., 2024. Unraveling how winds and surface heat fluxes control the atlantic ocean's meridional heat transport. arXiv preprint arXiv:2401.14230 .

Bourlès, B., Lumpkin, R., McPhaden, M.J., Hernandez, F., Nobre, P., Campos, E., Yu, L., Planton, S., Busalacchi, A., Moura, A.D., et al., 2008. The pirata program: History, accomplishments, and future directions. Bulletin of the American Meteorological Society 89, 1111–1126.

Carton, J.A., Giese, B.S., 2008. A reanalysis of ocean climate using simple ocean data assimilation (soda). Monthly weather review 136, 2999–3017.

Chen, P., Chi, M.Y., 2021. Stagru: ocean surface current spatio-temporal prediction based on deep learning, in: 2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI), IEEE. pp. 495–499.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785–794.

Chen, Z., Dong, J., 2019. Study of lstm model in sea surface temperature prediction of the yellow sea cold water mass area, in: 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), IEEE. pp. 367–371.

Choi, H.M., Kim, M.K., Yang, H., 2023. Deep-learning model for sea surface temperature prediction near the korean peninsula. Deep Sea Research Part II: Topical Studies in Oceanography 208, 105262.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine learning 20, 273–297.

Dayan, H., McAdam, R., Juza, M., Masina, S., Speich, S., 2023. Marine heat waves in the mediterranean sea: An assessment from the surface to the subsurface to meet national needs. Frontiers in Marine Science 10, 1045138.

Di Lorenzo, E., Schneider, N., Cobb, K.M., Franks, P., Chhak, K., Miller, A.J., McWilliams, J.C., Bograd, S., Arango, H., Curchitser, E., et al., 2008. North pacific gyre oscillation links ocean climate and ecosystem change. Geophysical research letters 35.

E.U. Copernicus Marine Service, 2025. Global Ocean OSTIA Sea Surface Temperature and Sea Ice Analysis. URL: https://data.marine.copernicus.eu/product/SST_GLO_SST_L4_NRT_OBSERVATIONS_010_001/description, doi:10.48670/moi-00165. accessed: 2025-05-16.

European Space Agency, 2025. SMOS Science Products. URL: https://earth.esa.int/eogateway/catalog/smos-science-products. accessed: 2025-05-16.

Feng, Y., Sun, T., Li, C., 2021. Study on long term sea surface temperature (sst) prediction based on temporal convolutional network (tcn) method, in: Proceedings of the ACM Turing Award Celebration Conference-China, pp. 28–32.

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2009. Geographically weighted regression. The Sage handbook of spatial analysis 1, 243–254.

Gao, J., Chen, X., Liu, A., Yu, C., 2024. Prediction of ocean velocity during typhoon using a ga-bp neural network model, in: 2024 Photonics & Electromagnetics Research Symposium (PIERS), IEEE. pp. 1–7.

Geiger, E.F., Grossi, M.D., Trembanis, A.C., Kohut, J.T., Oliver, M.J., 2013. Satellite-derived coastal ocean and estuarine salinity in the mid-atlantic. Continental Shelf Research 63, S235–S242.

Haghbin, M., Sharafati, A., Motta, D., Al-Ansari, N., Noghani, M.H.M., 2021. Applications of soft computing models for predicting sea surface temperature: a comprehensive review and assessment. Progress in earth and planetary science 8, 1–19.

Han, K.H., Hong, S., 2023. Ocean salinity retrieval and prediction for soil moisture active passive satellite using data-to-data translation. IEEE Transactions on Geoscience and Remote Sensing 61, 1–15.

Hao, R., Zhao, Y., Zhang, S., Deng, X., 2025. Deep learning for ocean forecasting: A comprehensive review of methods, applications, and datasets. IEEE Transactions on Cybernetics .

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al., 2020. The era5 global reanalysis. Quarterly journal of the royal meteorological society 146, 1999–2049.

Hittawe, M.M., Harrou, F., Togou, M.A., Sun, Y., Knio, O., 2024. Time-series weather prediction in the red sea using ensemble transformers. Applied Soft Computing 164, 111926.

Janmaijaya, M., Janmaijaya, M., Muhuri, P.K., 2024. A novel model based on spatio-temporal dilated convlstm networks for indian ocean dipole forecasting using multi-source global sea surface temperature and heat content data. IEEE Access .

Jia, W., Guan, S., Xue, Y., 2024. Tl-itransformer: Revolutionizing sea surface temperature prediction through itransformer and transfer learning. Earth Science Informatics 17, 4847–4857.

Jiang, J., Zuo, Y., Xiao, Y., Zhang, W., Li, T., 2024. Stmgf-net: a spatiotemporal multi-graph fusion network for vessel trajectory forecasting in intelligent maritime navigation. IEEE Transactions on Intelligent Transportation Systems .

Jin, Q., Tian, Y., Sang, Q., Liu, S., Yu, J., Wang, X., 2021. A deep learning model for joint prediction of three-dimensional ocean temperature, salinity and flow fields, in: 2021 6th International Conference on Automation, Control and Robotics Engineering (CACRE), IEEE. pp. 573–577.

Keele, S., et al., 2007. Guidelines for performing systematic literature reviews in software engineering. Technical Report. Technical report, ver. 2.3 ebse technical report. ebse.

Kim, J., Kim, T., Ryu, J.G., Kim, J., 2023. Spatiotemporal graph neural network for multivariate multi-step ahead time-series forecasting of sea temperature. Engineering Applications of Artificial Intelligence 126, 106854.

Lei, L., Jianxing, Z., 2023. Feature disentanglement learning model for ocean temperature field forecast. Applied Mathematical Modelling 117, 549–562.

Li, J., Hong, D., Gao, L., Yao, J., Zheng, K., Zhang, B., Chanussot, J., 2022a. Deep learning in multimodal remote sensing data fusion: A comprehensive review. arXiv preprint arXiv:2205.01380 .

Li, J., Wei, Y., Wang, X., Li, X., Zhao, Y., 2022b. Progressive multi-branch embedding fusion network for underwater image enhancement. Information Fusion 82, 1–12.

Li, X., Liu, B., Zheng, G., Ren, Y., Zhang, S., Liu, Y., Gao, L., Liu, Y., Zhang, B., Wang, F., 2020. Deep-learning-based information mining from ocean remote-sensing imagery. National Science Review 7, 1584–1605.

Liu, Y., Zhang, L., Hao, W., Zhang, L., Huang, L., 2024. Predicting temporal and spatial 4-d ocean temperature using satellite data based on a novel deep learning model. Ocean Modelling 188, 102333.

Lutjeharms, J., Catzel, R., Valentine, H., 1989. Eddies and other boundary phenomena of the agulhas current. Continental Shelf Research 9, 597–616.

Malavika, A., Pai, M.L., Johny, K., 2024. A hybrid decomposition-based machine learning approach for predicting subsurface temperature in the arabian sea. Modeling Earth Systems and Environment , 1–20.

Menaka, D., Gauni, S., 2021. Prediction of dominant ocean parameters for sustainable marine environment. IEEE Access 9, 146578–146591.

Meng, L., Yan, X.H., 2022. Remote sensing for subsurface and deeper oceans: An overview and a future outlook. IEEE Geoscience and remote sensing magazine 10, 72–92.

Meng, Y., Gao, F., Rigall, E., Dong, R., Dong, J., Du, Q., 2023. Physical knowledge-enhanced deep neural network for sea surface temperature prediction. IEEE Transactions on Geoscience and Remote Sensing 61, 1–13.

Meng, Y., Rigall, E., Chen, X., Gao, F., Dong, J., Chen, S., 2021. Physics-guided generative adversarial networks for sea subsurface temperature prediction. IEEE transactions on neural networks and learning systems 34, 3357–3370.

NASA MODIS Science Team, 2024. MODIS (Moderate Resolution Imaging Spectroradiometer). URL: `https://modis.gsfc.nasa.gov/`. accessed: 2025-05-15.

NOAA Physical Sciences Laboratory, 2025. Noaa optimum interpolation sea surface temperature v2 high resolution dataset. URL: `https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html`. accessed: 2025-05-16.

Olbers, D., Willebrand, J., Eden, C., 2012. Ocean dynamics. Springer Science & Business Media.

Ou, M., Xu, S., Luo, B., Zhou, H., Zhang, M., Xu, P., Zhu, H., 2024. 3d ocean temperature prediction via graph neural network with optimized attention mechanisms. IEEE Geoscience and Remote Sensing Letters .

Pan, J., Li, Z., Shi, S., Xu, L., Yu, J., Wu, X., 2023. Adaptive graph neural network based south china sea seawater temperature prediction and multivariate uncertainty correlation analysis. Stochastic Environmental Research and Risk Assessment 37, 1877–1896.

Patil, K., Deo, M., 2018. Basin-scale prediction of sea surface temperature with artificial neural networks. Journal of Atmospheric and Oceanic Technology 35, 1441–1455.

Qiao, B., Wu, Z., Tang, Z., Wu, G., 2022. Sea surface temperature prediction approach based on 3d cnn and lstm with attention mechanism, in: 2022 24th International Conference on Advanced Communication Technology (ICACT), IEEE. pp. 342–347.

Rahmstorf, S., 2006. Thermohaline ocean circulation. Encyclopedia of quaternary sciences 5, 191.

Roemmich, D., Johnson, G.C., Riser, S., Davis, R., Gilson, J., Owens, W.B., Garzoli, S.L., Schmid, C., Ignaszewski, M., 2009. The argo program: Observing the global ocean with profiling floats. Oceanography 22, 34–43.

Shi, Z., Zheng, H., Dong, J., 2024. Oceanvp: A hycom based benchmark dataset and a relational spatiotemporal predictive network for oceanic variable prediction. Ocean Engineering 304, 117748.

Sun, Y., Yao, X., Bi, X., Huang, X., Zhao, X., Qiao, B., 2021. Time-series graph network for sea surface temperature prediction. Big Data Research 25, 100237.

Sutton, P., Rickard, G., Roemmich, D., 2024. Southwest pacific ocean warming driven by circulation changes. Geophysical Research Letters 51, e2024GL109174.

Thongniran, N., Vateekul, P., Jitkajornwanich, K., Lawawirojwong, S., Srestasathiern, P., 2019. Spatio-temporal deep learning for ocean current prediction based on hf radar data, in: 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE), Ieee. pp. 254–259.

Tresnawati, R., Wirasatriya, A., Wibowo, A., Susanto, R.D., Widiaratih, R., Setiawan, J.D., Maro, J.F., Dollu, E.A., Fitria, S., Kurang, R.Y., 2024. Long term of sea surface temperature prediction for indonesia seas using multi time-series satellite data for upwelling dynamics projection. Remote Sensing Applications: Society and Environment 33, 101117.

Vanem, E., Zhu, T., Babanin, A., 2022. Statistical modelling of the ocean environment–a review of recent developments in theory and applications. Marine Structures 86, 103297.

Veron, F., Melville, W.K., Lenain, L., 2008. Infrared techniques for measuring ocean surface processes. Journal of Atmospheric and Oceanic Technology 25, 307–326.

Wang, J., Zheng, G., Yu, J., Shao, J., Zhou, Y., 2024a. Sea surface temperature prediction method based on deep generative adversarial network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing .

Wang, X., Xu, Y., Liu, S., Ren, B., Kosinka, J., Telea, A.C., Wang, J., Song, C., Chang, J., Li, C., et al., 2024b. Physics-based fluid simulation in computer graphics: Survey, research trends, and challenges. Computational Visual Media 10, 803–858.

Wei, L., Guan, L., Qu, L., 2019. Prediction of sea surface temperature in the south china sea by artificial neural networks. IEEE geoscience and remote sensing letters 17, 558–562.

Wen, J., Yang, J., Jiang, B., Song, H., Wang, H., 2020. Big data driven marine environment information forecasting: a time series prediction network. IEEE Transactions on Fuzzy Systems 29, 4–18.

Whitehead, J., 1995. Thermohaline ocean processes and models. Annual review of fluid mechanics 27, 89–113.

Wyatt, A.S., Leichter, J.J., Washburn, L., Kui, L., Edmunds, P.J., Burgess, S.C., 2023. Hidden heatwaves and severe coral bleaching linked to mesoscale eddies and thermocline dynamics. Nature communications 14, 25.

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., Chen, Z., 2019a. Short and mid-term sea surface temperature prediction using time-series satellite data and lstm-adaboost combination approach. Remote Sensing of Environment 233, 111358.

Xiao, C., Chen, N., Hu, C., Wang, K., Xu, Z., Cai, Y., Xu, L., Chen, Z., Gong, J., 2019b. A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data. Environmental Modelling & Software 120, 104502.

Xiao, L., Li, S., Chen, B., 2024. Gssa: A network for short to medium-term regional sea surface temperature prediction. IEEE Geoscience and Remote Sensing Letters .

Xie, J., Ouyang, J., Zhang, J., Jin, B., Shi, S., Xu, L., 2021. An evolving sea surface temperature predicting method based on multidimensional spatiotemporal influences. IEEE Geoscience and Remote Sensing Letters 19, 1–5.

Xie, J., Zhang, J., Yu, J., Xu, L., 2019. An adaptive scale sea surface temperature predicting method based on deep learning with attention mechanism. IEEE Geoscience and Remote Sensing Letters 17, 740–744.

Yang, Y., Dong, J., Sun, X., Lima, E., Mu, Q., Wang, X., 2017. A cfcc-lstm model for sea surface temperature prediction. IEEE Geoscience and Remote Sensing Letters 15, 207–211.

Ye, M., Li, B., Nie, J., Wen, Q., Wei, Z., Yang, L.L., 2023. Graph convolutional network-assisted sst and chl-a prediction with multicharacteristics modeling of spatio-temporal evolution. IEEE Transactions on Geoscience and Remote Sensing 61, 1–14.

Yue, W., Xu, Y., Xiang, L., Zhu, S., Huang, C., Zhang, Q., Zhang, L., Zhan, X., 2024. Prediction of 3-d ocean temperature based on self-attention and predictive rnn. IEEE Geoscience and Remote Sensing Letters 21, 1–5.

Zhang, B., Li, F., Zheng, G., Wang, Y., Tan, Z., Li, X., 2021a. Developing big ocean system in support of sustainable development goals: challenges and countermeasures. Big Earth Data 5, 557–575.

Zhang, C., Dong, M., 2004. Seasonality in the madden–julian oscillation. Journal of climate 17, 3169–3180.

Zhang, K., Geng, X., Yan, X.H., 2020. Prediction of 3-d ocean temperature by multilayer convolutional lstm. IEEE Geoscience and Remote Sensing Letters 17, 1303–1307.

Zhang, Q., Wang, H., Dong, J., Zhong, G., Sun, X., 2017. Prediction of sea surface temperature using long short-term memory. IEEE geoscience and remote sensing letters 14, 1745–1749.

Zhang, X., Li, Y., Frery, A.C., Ren, P., 2021b. Sea surface temperature prediction with memory graph convolutional networks. IEEE Geoscience and Remote Sensing Letters 19, 1–5.

Zhao, X., Qi, J., Yu, Y., Zhou, L., 2024. Deep learning for ocean temperature forecasting: a survey. Intelligent Marine Technology and Systems 2, 28.

Zuo, X., Zhou, X., Guo, D., Li, S., Liu, S., Xu, C., 2021. Ocean temperature prediction based on stereo spatial and temporal 4-d convolution model. IEEE Geoscience and Remote Sensing Letters 19, 1–5.