CAMBRIDGE
UNIVERSITY PRESS

APPLICATION PAPER

Neural network approaches for sea surface height predictability using sea surface temperature

Luther Ollier[1] , Sylvie Thiria[1], Carlos E. Mejia[1], Michel Crépon[1] and Anastase Charantonis[2]

[1]LOCEAN, IPSL, Paris, France
[2]Anastase Charantonis, INRIA, IPSL, Paris, France
Corresponding author: Luther Ollier; Email: luther.ollier@locean.ipsl.fr

Abstract

Sea Surface Height Anomaly (SLA) is a signature of the mesoscale dynamics of the upper ocean. Sea surface temperature (SST) is driven by these dynamics and can be used to improve the spatial interpolation of SLA fields. In this study, we focused on the temporal evolution of SLA fields. We explored the capacity of deep learning (DL) methods to predict short-term SLA fields using SST fields. We used simulated daily SLA and SST data from the Mercator Global Analysis and Forecasting System, with a resolution of $(1/12)°$ in the North Atlantic Ocean $(26.5–44.42°N, −64.25–41.83°E)$, covering the period from 1993 to 2019. Using a slightly modified image-to-image convolutional DL architecture, we demonstrated that SST is a relevant variable for controlling the SLA prediction. With a learning process inspired by the teaching-forcing method, we managed to improve the SLA forecast at 5 days by using the SST fields as additional information. We obtained predictions of 12 cm (20 cm) error of SLA evolution for scales smaller than mesoscales and at time scales of 5 days (20 days) respectively. Moreover, the information provided by the SST allows us to limit the SLA error to 16 cm at 20 days when learning the trajectory.

Impact Statement

This study uses deep learning to enhance short-term predictions of sea surface height anomaly (SLA) by incorporating sea surface temperature (SST) data. Using simulated data from the Mercator Global Analysis and Forecasting System, our approach demonstrated that SST significantly improves SLA predictions. With a modified image-to-image convolutional architecture and a teaching-forcing inspired learning process, we achieved substantial error reductions in SLA forecasts over 5 and 20-day periods. These findings highlight the potential of SST as a crucial variable in oceanographic models, offering improved accuracy in forecasting mesoscale ocean dynamics. This advancement supports a better understanding and prediction of ocean behavior, which is essential for navigation, climate studies, and marine resource management.

The findings presented in this article have significant implications for understanding and predicting ocean dynamics, especially in regions influenced by strong eddies and rapid variability. By employing a novel approach that samples ocean eddies at multiple time steps and utilizing deep learning models, the study enhances our ability to integrate information from sea level anomaly (SLA) and SST patterns.

1. Introduction

Improving the knowledge of upper ocean dynamics is crucial for understanding the ocean's role in climate and for various applications such as enhancing navigation trajectory or marine engineering. To this end, multi-satellite observations have provided a vast quantity of information on different oceanic parameters such as sea level anomaly (SLA) that is the difference between the mean Sea Surface Height and the height measured by altimeter, sea surface temperature (SST), Chlorophylle-a, or Sea Surface Salinity. SLA allows us to capture many oceanic features and learn about global ocean circulation. Using the geostrophic balance, it can retrieve Sea Surface currents and capture mesoscale structures such as eddies or ocean fronts (Moschos, 2020; Zeng, 2015). However, the altimetry products come from interpolation between different satellites at different times (Pujol, 2016;Taburet et al., 2019; Dufau et al., 2016). Therefore, improving SLA and surface currents resolution or predicting short-term surface evolution is an active research topic (Immas, 2021; Manucharyan, 2021;Isern-Fontanet, 2017; González-Haro, 2020). To this end, other satellite-measured parameters can be used, such as SST. SST is observed from multi-sensor measurements with high spatiotemporal resolution. Furthermore, SST is a good tracer for advection, a key phenomenon of SLA evolution. Several studies have shown that SLA and SST are strongly associated with complex functions (Isern-Fontanet, 2006; Martin et al., 2023). SLA observations are obtained with a delay of 15 days (Taburet et al., 2019; Dufau et al., 2016) due to the necessity of processing altimeter tracks, while SST is acquired daily. How can we ensure accurate estimates of SLA observations during this delay period? Mathematical methods based on physical knowledge have been built to improve the interpolation of SLA fields (Rio, 2016) from physical knowledge. In contrast, data-driven approaches such as deep learning (DL) algorithms can infer the underlying state of the system driving ocean surface parameters. Neural networks can efficiently process this vast amount of ocean data. It has already been proved that DL methods can improve SLA resolution using SST (Guan, 2023; Liu, 2020) but without using the temporal evolution of SLA and SST. Besides, predictive neural networks have been built to forecast SLA at a small resolution (¼°) without including other parameters like SST (Archambault, 2022; Thiria et al., 2023; Braakmann-Folgmann, 2017; Liu, 2022). Our study considers the correlated temporal evolution of SLA and SST to show that it is possible to refine SLA prediction using SST at high resolution (1/12°). To this end, a convolutional network algorithm was implemented for its robustness and stability in processing images and a specific training methodology has been designed as described in Section 2. Our article presents two main results in Section 3: the forecast performance we achieve thanks to our training methodology, and the performance in projecting SLA using SST as a control.

2. Material and methods

The CopernicusGlobal12 dataset (Global Ocean Physics Reanalysis, 2022) provides various daily parameters located on a 1/12° grid from 26.5° to 44.42° latitude North and − 64.25° to −53.58° longitude East. Our main study used the dynamic zone shown in Figure 1. SLA and SST selected spanned from 1993 to 2019, corresponding to approximately 10,000 daily images.

The Gulf Stream traverses this region, resulting in a more pronounced dynamical evolution of SLA and SST, particularly in the northern part of the area. In contrast, the southern region experiences less extreme variations and smaller eddies. Therefore, we chose to study only the northern area of latitudes described below. (See supplementary materials for generalization to the southern area.):

- lat N from 33.7 to 44.42° representing the dynamic area with the Gulf Stream flow

In Figure 3 (left panel of the second row), we present the standard deviation of the SLA estimated during the 20 years of the study period for our domain.

2.1. Times series decomposition

Since the focus of our forecast study is on short-time predictions, it is crucial that the data used does not exhibit periodicity due to seasonal variability. Additionally, since the data covers a large time period,
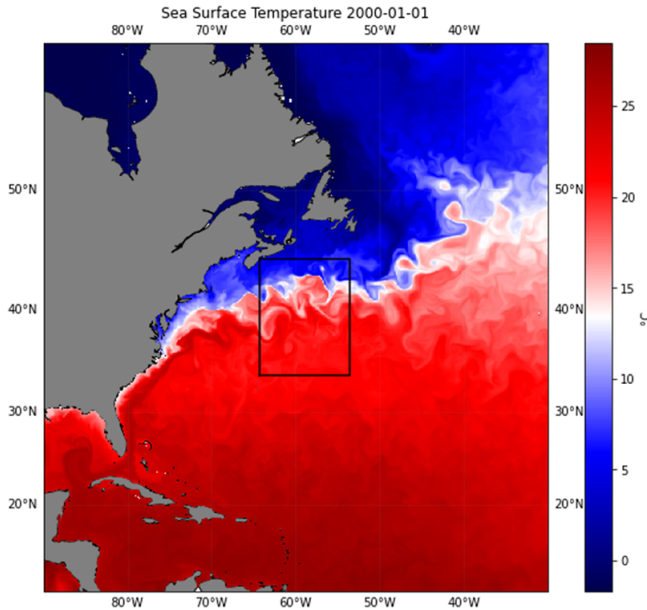
Figure 1. *The map shows the sea surface temperature (SST) of the North Atlantic region under study, with the black box indicating the dynamic area of focus.*

we removed the underlying trend corresponding to the interannual variability to only keep the high-frequency variability, the SLA residuals (Mann, 1945; Kendall, 1975) see supplementary materials). As we got rid of the seasonal and multi-annual variability here, SLA dynamics will now refer to the sub-seasonal spatial patterns and gradients of SLA variability across our ocean regions.

In Figure 2 we present the image of SLA residuals at time t (t = 05/09/1993 Figure 2a), its evolution for two steps (t + 5 and t + 10 Figure 2bc); the difference SLA(t)−SLA (t + dt) (Figure 2ef) represents the transformation of the SLA field due to the dynamic.

2.2. Methodology

We sought a neural approach that would enable us to predict SLA residuals given the knowledge of SLA and SST residuals time series. To accomplish this, we separated our data into three independent datasets for training, validation and testing. The same datasets were used for all experiments presented in this study. Twenty days were removed from the learning dataset at the beginning of 1994 and at the end of 2017 to separate the test from the learning phase. The three datasets cover different periods and are normalized separately:

- training set: 1994–2014 (7412 images)
- validation set: 2015–2017 (1275 images)
- test set: 1993 2018 2019 (873 images)

Deep convolutional networks are designed for image processing, using convolution operations to extract spatial context. In this study, we have chosen the U-net architecture (Ronneberger, 2015) for its robustness and flexibility, which make it suitable for a wide range of image processing problems. The aim is to reduce the input spatial resolution by passing it through various layers to get image patterns. As the first part reduces the input matrix size, it filters unnecessary information like noise to focus on patterns. The input information is propagated through these layers without losses via skipping connections. Moreover, we added Batch Normalization (Ioffe, 2015) to the convolution layers of the original U-net to limit the
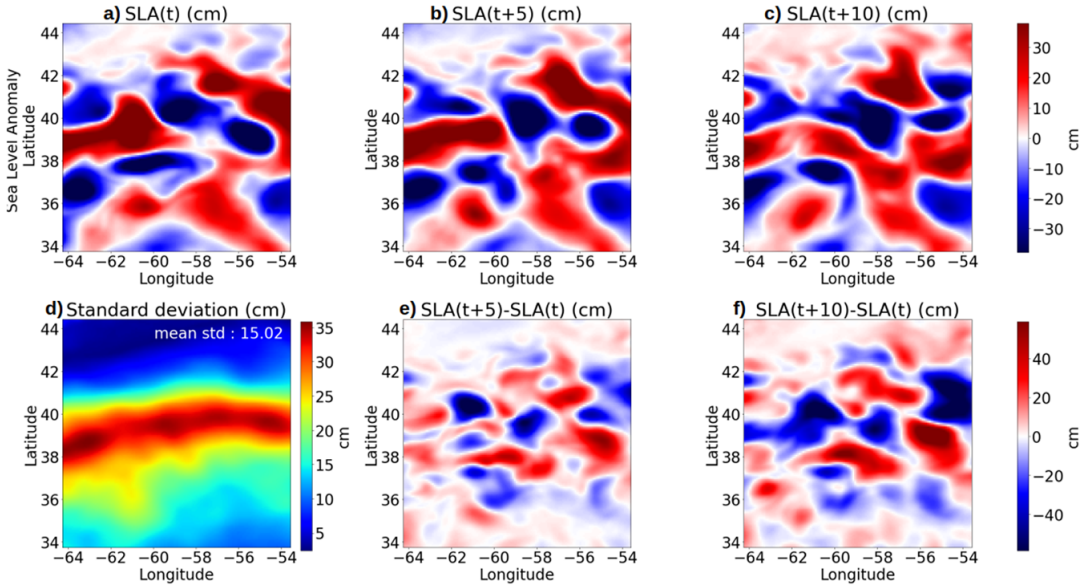
Figure 2. *(a) SLA residuals (t = 24/10/1993) and its evolution 5 (b) and 10 (c) days after. SLA evolution at 5 (e) and 10 (f) days compared to the previous state: $SLA(t+k) - SLA(t), k = 5(10)$. (d) Standard deviation of the SLA over the Gulf Stream region estimated using the 20 years.*
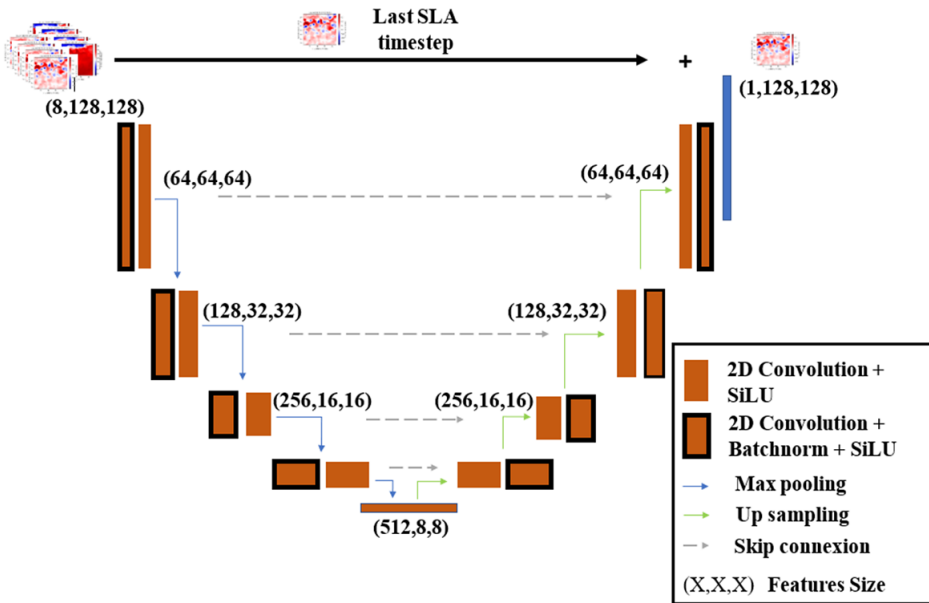

Figure 3. *SLA-Res-U-net Architecture: A typical U-net architecture comprises two main components. Initially, the encoder reduces spatial resolution to capture patterns and incorporates feature channels for context propagation. Then, the decoder expands the resolution features from the encoder, and its output is combined with the input image to act as a residual unit. Batch normalization is added between the 2D convolutions during descent and after SiLU activation during ascent. Hyper-parameters like activation, depth, and learning rate are determined via Bayesian optimization (S0).*

overflow issues. Finally, we have made a slight modification inspired by the residual connections process (Goodfellow, 2016). Essentially, this involves an identity mapping of the input *X*, added to the output *F*(*X*) of U-net transformations, resulting in *F*(*X*) + *X* (refer to Figure 3).

The nature of the task is changed such that instead of predicting an image, the network predicts an additive transformation over the last time step. This method has already been successfully applied in various domains, such as medical image processing (Guan, 2020). In the following, we denote this architecture as SLA-Res-U-net. We tried different architectures, including the classical U-net and found that SLA-Res-U-net outperforms them in every case (Zeng, 2015).

Due to the high correlation between SLA and SST (González-Haro, 2020), the SST images are provided as inputs along with the associated SLA images. Thus, the inputs are the images of SLA and SST residuals taken at successive times equally spaced, and the targets are the following SLA and SST residual images1.

$$IM(t+dt) = G[IM(t), IM(t-dt), IM(t-2dt), \ldots IM(t-ndt)] \tag{1}$$

where IM represents the (SLA, SST) images and dt the time lag between two images, the function G is our architecture (He, 2015).

Accurate predictions must provide sufficient information on the phenomenon under study and its dynamic. We performed preliminary tests to determine the optimal number of time steps and the time lag between two-time steps. These were selected based on various criteria, such as the accuracy of the prediction, geophysical considerations, and computation time.

The time evolution of the SLA is driven by ocean eddies whose return time is approximately 20 days in the Gulf Stream area (Kang and Enrique, 2013; Martin and Synthesizing, 2023; Chelton et al., 1998). Therefore, we sampled the eddy at four time steps (t, t−dt, t−2dt, t−3dt) separated by a time lag dt of dt = 5 days, resulting in a total sampling period of 20 days. These four-time step images of (SLA, SST) were chosen as inputs of the U-net model 3. The objective was to predict the SLA and SST images at the next time step forward (a time step corresponds to 5 days). We denote SLA-SST/SLA-SST the SLA-Res-U-net model obtained at the end of the learning phase. To understand the contribution of each variable, we trained two other SLA-Res-U-net by modifying the inputs and the outputs. The first model (called SLA/SLA) uses as input the SLAs at four specific time steps backward and estimates SLA at the next time step forward (5 days later). The second model (denoted SLA-SST/SLA) uses the two-time series (SLA, SST) to estimate the next SLA only, the cost function minimizes only the prediction of the SLA.

We focused the study on the determination of the SLA only since SLA altimetric resolution is much weaker than SST (Dufau et al., 2016). SST helps to retrieve SLA at high resolution. The second part of the study analyzes the number of time steps for which the forecast is accurate (prediction horizon).

We first tested the prediction given by the three models (SLA/SLA, SLA-SST/SLA, SLA-SST/SLA-SST,) by iterating each model in time. This involved introducing the estimated outputs at time t + ndt as input to estimate the image (SLA, SST) at time t + (*n* + 1)dt without new learning. For the SLA-SST/SLA, the true SST was introduced as input at each iteration since no prediction of SST was made. Therefore the model SLA-SST/SLA is not a forecast but rather a model of SLA evolution driven by SST. The SLA predicted by the model is reintroduced as an input to obtain the SLA values 10, 15, and 20 days later.

The second experiment consists in forcing the learning of the dynamics of the SLA fields with the model. First, we computed the output of the model at time t + dt. The model then received the preceding output as input and computed the output at the next step t + 2dt. This process was repeated, training the network to build the outputs at time t + 2dt from inputs at time t. We only iterated this process once more to estimate t + 3dt using the estimations at time t + 2dt, due to computation time constraints. In this way, the network can learn input conditions not seen during normal training.

The performances of the predictions were estimated at times t + dt, t + 2dt, t + 3dt, and t + 4dt (note that the prediction at t + 4dt was not learned). The flow chart of the procedure is shown in Figure 4. We denote this new architecture as DY-SLA-SST/SLA-SST. All the different configurations tested are shown in Table 1.

Moreover, to show the impact of SST on the trajectory of SLA, we did the same procedure with SLA-SST/SLA and trained the network by introducing the true SST at each step. We denote DY-SLA-SST/SLA

Table 1. *Input/output configurations and naming conventions for the various setups. Noted that for each setup, there are four input fields spaced five days apart, used to generate the subsequent five-day field*

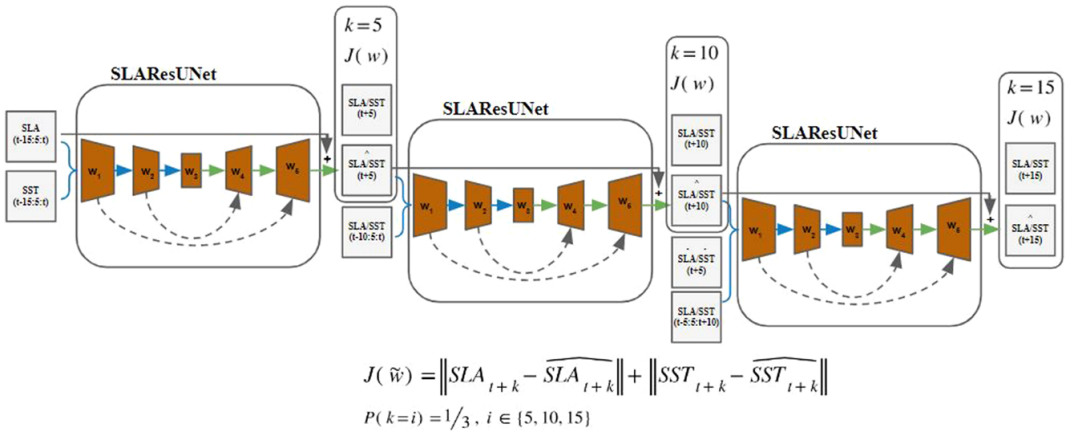| Name	Inputs	Outputs	Training method
SLA/SLA	SLA	SLA	Step-by-step iteration
DY SLA/SLA	SLA	SLA	Multi-timesteps iteration
SLA-SST/SLA	SLA-SST	SLA	Step-by-step iteration
DY SLA-SST/SLA	SLA-SST	SLA	Multi-timesteps iteration
SLA-SST/SLA-SST	SLA-SST	SLA-SST	Step-by-step iteration
DY SLA-SST/SLA-SST	SLA-SST	SLA-SST	Multi-timesteps iteration

$$J(\tilde{w}) = \left\| SLA_{t+k} - \widehat{SLA_{t+k}} \right\| + \left\| SST_{t+k} - \widehat{SST_{t+k}} \right\|$$

$$P(k=i) = 1/3 , i \in \{5, 10, 15\}$$

Figure 4. *Training scheme: During the training phase, the network is fed with both SLA prediction and SST observations which extends the horizon of SLA. By performing backpropagation over several time steps, the model is forced to learn the dynamics of SLA.*

in this model. In doing so, we learned the trajectory of the SLA image with the true SST information and consequently improved the SLA prediction at t + 4dt. This procedure allowed us to estimate SLA, whose observations are difficult to obtain in the near future, using SST which can be obtained almost instantaneously. The method was inspired by teacher-forcing (Goodfellow, 2016).

The learning process is performed (as shown in Figure 4) using classic backpropagation. During training, we randomly select whether to compute outputs at time t + dt, t + 2dt, or t + 3dt. The backpropagation is performed over the loop, meaning that we update the weights for all time step predictions. It is computationally heavier because for each time step the amount of calculation to run back the propagation increases proportionally for each time step.

2.3. Evaluation metrics

We consider the following metric to evaluate SLA predictions: the root-mean-square error computed on the test set (see Equation 2). Additionally, we use persistence performances as a threshold to assess our model's performances. Persistence refers to a prediction that assumes either zero uniform velocities or velocities that cancel each other out. Additional metrics can be found in the supplementary materials.

$$RMSE = \sqrt{\frac{\Sigma_n \left(y_{pred} - y_{target}\right)^2}{N}}, n \in [testpixel] \tag{2}$$

Table 2. RMSE according to predictions in time estimated from the Test set

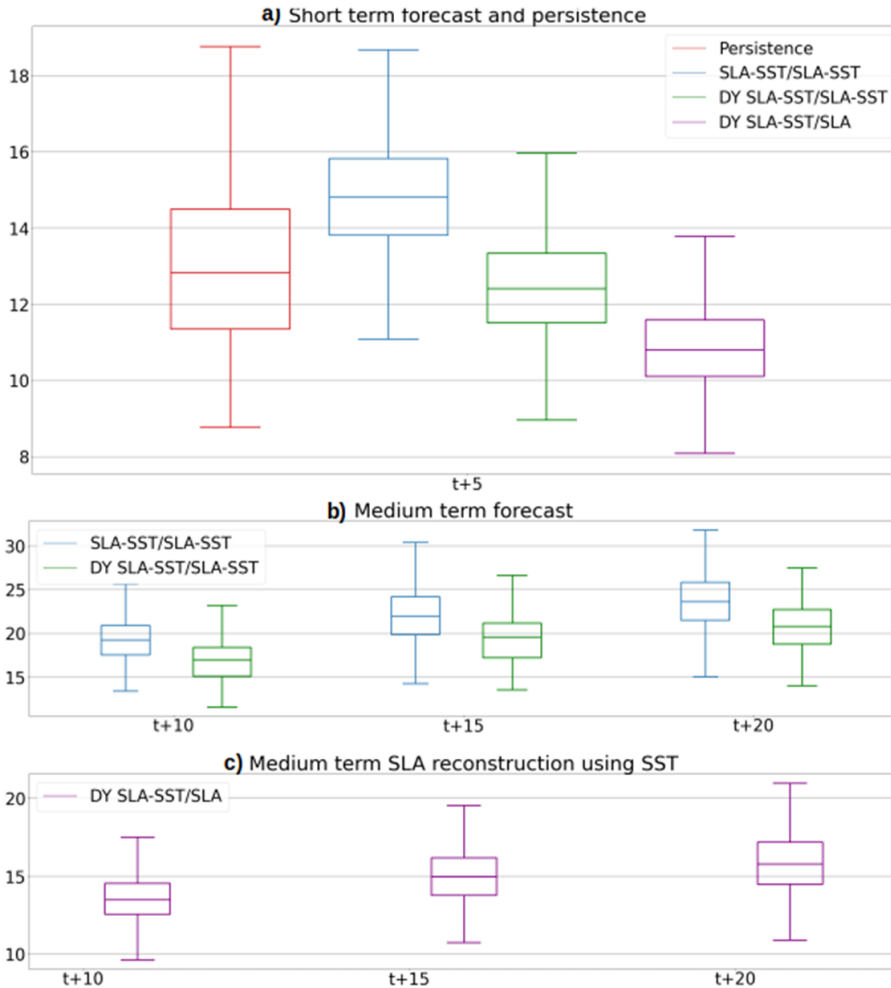| Inputs/predictions	Evaluate	$n = 5$	$n = 10$	$n = 15$	$n = 20$
SLA/SLA	RMSE (cm)	13.27	17.86	21.73	23.66
SLA-SST/SLA	RMSE (cm)	13.13	18.53	22.07	24.59
SLA-SST/SLA-SST	RMSE (cm)	14.81	18.73	21.88	23.76
Persistence	RMSE (cm)	13.00			

Figure 5. RMSE for different time steps and models estimated on the whole test set. The red box is the persistence RMSE, the blue box corresponds to the SLA-SST/SLA-SST model, the green one to the DY-SLA-SST/SLA-SST, and the purple one to the DY-SLA-SST/SLA.

3. Results

In this section, we examine the various performance statistics of the models (see Table 2). For the prediction at $t + 5$, the RMSE computed on the entire image is worse than that of persistence, with only a slight advantage on SLA-SST/SLA-SST. The RMSE values, while still large, are small in comparison to the SLA standard deviation of approximately 0.2 m (see Figure 2 and supplementary).
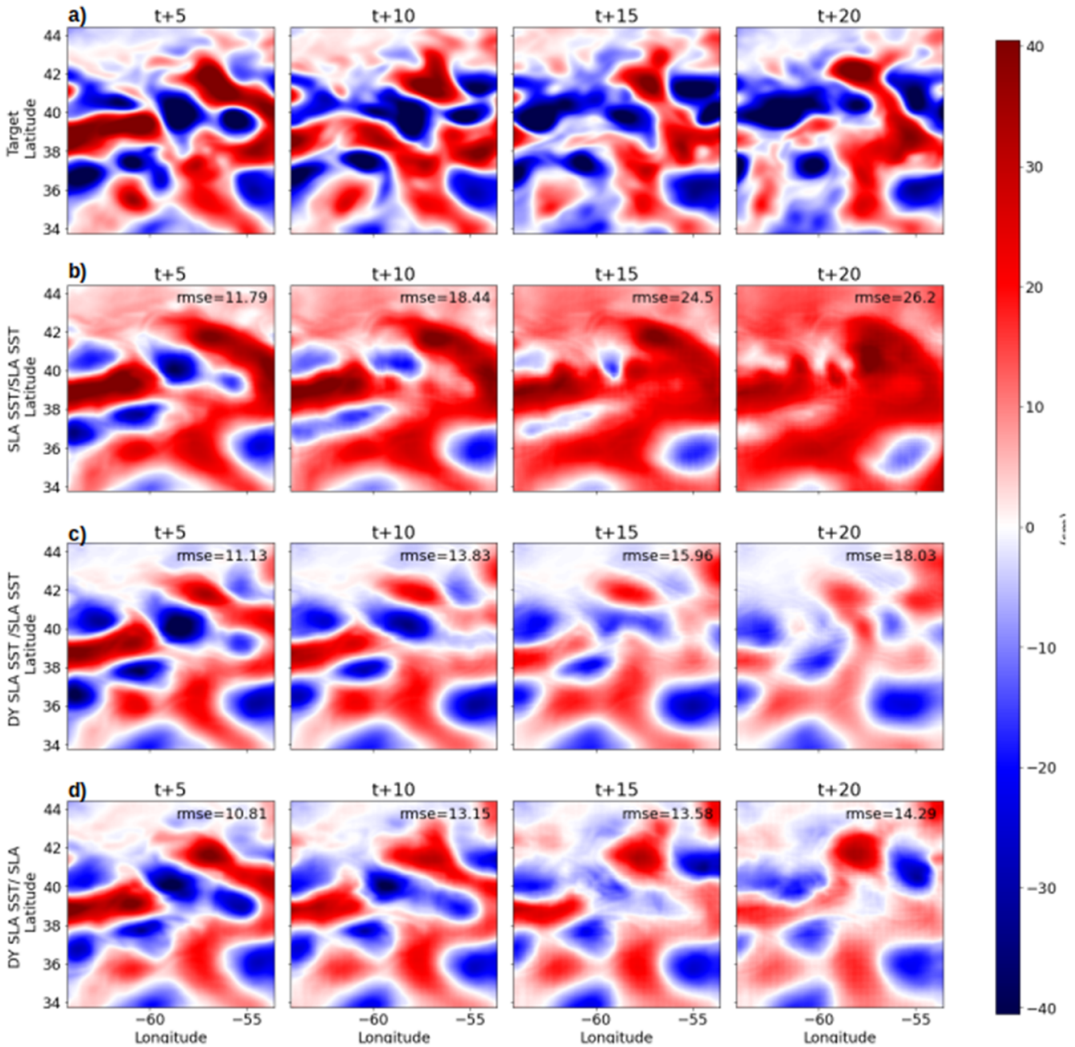
Figure 6. *SLA image predictions at time t + 5 day; t + 10 days; t + 15 days; t + 20 days (t = 24/10/1993). (a) correspond to the ground truth results, (b) to the SLA-SST/SLA-SST method, (c) correspond to the DY SLA-SST/SLA-SST method, and (d) to the DY-SLA-SST/SLA method.*

The training routine described in 4 was applied with different input and output configurations. Notice that DY SLA-SST/SLA is not a forecast, but DY SLA-SST/SLA-SST is one because the SST predicted can be reintroduced. Figure 5 provides a comprehensive comparison of SLA-SST/SLA and DY-SLA-SST/SLA-SST in terms of RMSE and standard deviation. While the RMSE is comparable to persistence at t + 5, the standard deviations are more reliable when the model is trained. This effect is even more pronounced for subsequent time steps.

The performances of the DY-SLA-SST/SLA model improve and do not get deteriorated with time as much as with the SLA-SST/SLA-SST model. By providing the model with the true SST at each step, the DY-SLA-SST/SLA model is better able to keep close to the true SLA compared to the DY-SLA-SST/SLA-SST. The most striking point is the improvement of the first step prediction with the DY-SLA-SST/SLA and DY-SLA-SST/SLA-SST (see t + 5 in Figure 5). This suggests that forcing the network to make consistent predictions across different steps allows for a deeper representation of the underlying dynamics of SLA.

4. Conclusions

The efficiency of the different methods we have developed is illustrated in Figure 6, with predictions starting at t = 24/10/1993.

Figure 6b shows the predictions in time of SLA-SST/SLA-SST and compares them to the targets (Figure 6a). Despite, the low RMSE (11.79 cm for t + 5), we can observe that the global dynamic is not accurately reproduced. Figure 6c,d presents the predictions in time of respectively DY-SLA-SST/SLA-SST and DY-SLA-SST/SLA models, respectively.

We note that SLA images given by the two teacher-forcing methods (DY SLA-SST/SLA-SST and DY SLA-SST/SLA, Figure 6cd) perform better than SLA images given by the SLA-SST/SLA-SST method (Figure 6b). The RMSEs of the two DY methods are much smaller than those of the SLA-SST/SLA-SST method. Besides the SLA patterns of the two DY methods well reproduce these of the target. The number and position of troughs and bumps that characterize ocean circulation via the geostrophic relationship (troughs are associated with a cyclonic circulation in the upper ocean and bumps with an anticyclonic one) are similar in both the DY predictions and ocean model images. The two DY methods give a satisfactory prediction up to t + 20 while the SLA-SST/SLA-SST method begins to degrade at t + 10 and gives a bad prediction at t + 20 both in RMSE and in SLA patterns.

Our architecture builds the transformation over the image, trained with a teacher-forcing method which helps the network to learn the trajectory of the SLA fields using the information provided by the SST fields. Future research on temporal-oriented architecture such as LSTM or Attention-based method will be conducted to enhance this first results on SST importance. Complementary results on the teacher-forcing method will be produced with another tracer of the SLA advection, the Sea Surface Salinity (SSS). It will allow us to assess the capability of the method of learning SLA trajectory using information from other correlated fields.

Since the SLA observations are obtained with some delay (15 days) due to the need to process the altimeter observations during the altimeter repeat period and the SSTs are obtained almost instantaneously, the DL algorithm we developed, can provide good estimates of the SLA observations during this delay period.

This study is a proof of concept that DL methods can effectively model the state of the ocean surface for the near future and hence the associated time-dependent PDEs. An interesting question would be to test how well the DL model reproduces the ocean heat transport which is a key variable for climate studies. This study will be useful to predict the generation of storms and cyclones which depend on the ocean surface SST, and for ship routine to optimize ship trajectories.

Finally, the high-resolution SLA data provided by the SWOT (Surface Water and Ocean Topography) mission offers unprecedented insights into ocean dynamics and variability. With its enhanced spatial resolution, SWOT data enable more accurate monitoring of oceanic phenomena such as mesoscale eddies, coastal currents, and sea level rise. However, its return time is approximately 20 days (see (Rosemary Morrow1, Denis Blumstein, and Gerald Dibarboure, 2018), Figure 8.1), which poses a challenge for temporal resolution. Therefore, improved forecasting capabilities, as facilitated by our current and future works, are essential for obtaining accurate real-time data.

Author contribution. Conceptualization: S.Thiria, M.Crépon; Methodology: A.Charantonis; Implementation: L.Ollier, C.E. Mejia; Data curation: C.E.Mejia; Data visualisation: L.Ollier; Writing original draft: L.Ollier, S.Thiria, M. Crépon; All authors approved the final submitted draft.

References

Archambault T, **Charantonis A**, **Béréziat D,**, **Mejia C**, **Thiria S.** Sea surface height super-resolution using high-resolution sea surface temperature with a subpixel convolutional residual network. *Environmental Data Science.* 2022;*1*:e26. doi:10.1017/eds.2022.28

Braakmann-Folgmann, Anne et al. "Sea Level Anomaly Prediction using Recurrent Neural Networks." ArXiv abs/1710.07099 (2017).

Chelton DB, **deSzoeke RA**, **Schlax MG**, **Naggar KE and Siwertz N** (1998) Geographical variability of the first Baroclinic Rossby radius of deformation. *Journal of Physical Oceanography vol 28 issue 3, pages 433-460* https://doi.org/10.1175/1520-0485(1998)028<0433:GVOTFB>2.0.CO;2.

Dufau C, **Orsztynowicz M**, **Dibarboure G**, **Morrow R and Traon P-YL** (2016) Mesoscale resolution capability of altimetry: present and future. *Journal of Geophysical Research: Oceans 121*(7), 4910–4927 https://doi.org/10.1002/2015JC010904.

Global Ocean Physics Reanalysis. Available at https://doi.org/10.48670/moi-00021 (accessed on 01 August 2022).

González-Haro C (2020) Ocean surface currents reconstruction: spectral characterization of the transfer function between SST and SSH. *Journal of Geophysical Research: Oceans 125* e2019JC015958.

Goodfellow I (2016) *Deep Learning*. MIT Press.

Guan S (2020) Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal. *IEEE Journal of Biomedical and Health Informatics 24* 568–576.

Guan S (2023) Downscaling of ocean fields by fusion of heterogeneous observations using deep learning algorithms. *Ocean Modelling 182* 102174.

He K (2015) Deep residual learning for image recognition. *ARXIV* https://doi.org/10.48550/ARXIV.1512.0.

Immas A (2021) Real-time in situ prediction of ocean currents. *Ocean Engineering 228*, 108922.

Ioffe S (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. *MICCAI.* arXiv:cs.LG/1502.03167.

Isern-Fontanet J (2006) Potential use of microwave sea surface temperatures for the estimation of ocean currents. *Geophysical Research Letters 33*, L24608.

Isern-Fontanet J (2017) Remote sensing of ocean surface currents: a review of what is being observed and what is being assimilated. *Nonlinear Processes in Geophysics 24*, 613–643.

Kang D and Enrique N (2013)Curchitser Gulf Stream Eddy characteristics in a high-resolution ocean model. *Journal of Geophysical Research 118*

Kendall M (1975) Rank Correlation Methods. Griffin editions.

Liu J (2020) A deep learning method with merged LSTM neural networks for SSHA prediction. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13*, 2853–2860.

Liu J (2022) Sea surface height prediction with deep learning based on attention mechanism. *IEEE Geoscience and Remote Sensing Letters 19*, 1–5.

Mann HB (1945) *Nonparametric Tests Against Trend*. Econometrica.

Manucharyan, G. E., **Siegelman, L.**, & **Klein, P.** (2021). A deep learning approach to spatiotemporal sea surface height interpolation and estimation of deep currents in geostrophic ocean turbulence. *Journal of Advances in Modeling Earth Systems*, *13*, e2019MS001965. https://doi.org/10.1029/2019MS001965

Martin, S. A., **Manucharyan, G. E.**, & **Klein, P.** (2023). Synthesizing sea surface temperature and satellite altimetry observations using deep learning improves the accuracy and resolution of gridded sea surface height anomalies. *Journal of Advances in Modeling Earth Systems*, *15*, e2022MS003589. https://doi.org/10.1029/2022MS003589

Moschos E (2020) Deep-SST-Eddies: a deep learning framework to detect oceanic eddies in sea surface temperature images. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2020, Barcelona,Spain, pp. 4307–4311.

Pujol MI (2016) DUACS DT2014: the new multi-mission altimeter data set reprocessed over 20 years. *Ocean Science 12*, 1067–1090.

Rio MH (2016) Improving the altimeter-derived surface currents using high-resolution sea surface temperature data: a feasability study based on model outputs. *Journal of Atmospheric and Oceanic Technology 33*, 2769–2784.

Ronneberger, O., **Fischer, P.**, & **Brox, T.** (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference*, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234–241). Springer International Publishing.

Morrow R, **Blumstein D and Dibarboure G** (2018) Fine-scale altimetry and the future SWOT mission. *New Frontiers in Operational Oceanography*, *6*, 2296–7745.

Taburet G, **Sanchez-Roman A**, **Ballarotta M**, **Pujol M-I**, **Legeais J-F**, **Fournier F**, **Faugere Y and Dibarboure G** (2019) DUACS DT2018: 25 years of reprocessed sea level altimetry products *Ocean Science 15*(5), 1207–1224 https://doi.org/10.5194/os-15-1207-2019.

Thiria S, **Sorror C**, **Archambault T**, **Charantonis AA**, **Béréziat D**, **Mejia C**, **Molines J-M and Crépon M** (2023) Downscaling of ocean fields by fusion of heterogeneous observations using deep learning algorithms. *Ocean Modelling*, Volume *182*, 1463–5003,

Zeng X (2015) Predictability of the loop current variation and Eddy shedding process in the Gulf of Mexico using an artificial neural network approach. *American Meteorological Society Section: Journal of Atmospheric and Oceanic Technology 32*, 1098–1111.