https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


# Deep Learning for Super-Resolution of Mediterranean Sea Surface Temperature Fields

Claudia Fanelli[1*], Daniele Ciani[2], Andrea Pisano[2], and Bruno Buongiorno Nardelli[1]

[1]Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine (CNR-ISMAR), Calata Porta di Massa, 80133, Naples, Italy
[2]Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine (CNR-ISMAR), Via del Fosso del Cavaliere 100, 00133, Rome, Italy

**Correspondence:** Claudia Fanelli (claudia.fanelli@cnr.it)

**Abstract.**

Sea surface temperature (SST) is one of the essential variables of the Earth climate system. Being at the interface with the atmosphere, SST modulates heat fluxes in and out of the ocean, provides insight on several upper/interior ocean dynamical processes, and it is a fundamental indicator of climate variability potentially impacting marine ecosystems' health. Its accurate estimation and regular monitoring from space is therefore crucial. However, even if satellite infrared/microwave measurements provide a much better coverage than what achievable from in situ platforms, they cannot sense the sea surface under cloudy/rainy conditions. Large gaps are present even in merged multi-sensor satellite products and different statistical strategies have thus been proposed to obtain gap-free (L4) images, mostly based on the Optimal Interpolation algorithms. This kind of techniques, however, filter out the signals below the space-time decorrelation scales considered, significantly smoothing most of the small mesoscale and submesoscale features. Here, deep learning models, originally designed for single image Super Resolution (SR), are applied to enhance the effective resolution of SST products and the accuracy of SST gradients. SR schemes include a set of computer vision techniques leveraging Convolutional Neural Networks to retrieve high-resolution data from low-resolution images. A dilated convolutional multi-scale learning network, which includes an adaptive residual strategy and implements a channel attention mechanism, is used to reconstruct features in SST data at 1/100° spatial resolution starting from 1/16° data over the Mediterranean Sea. The application of this technique shows a remarkable improvement in the high resolution reconstruction, being able to capture small scale features and providing a root-mean-squared-difference improvement of 0.02°C with respect to the L3 ground-truth data.

## 1 Introduction

Investigating ocean dynamics and climate variability requires accurate, regular and systematic observations of the Sea Surface Temperature (SST). SST plays indeed a key role in air-sea interaction and upper ocean circulation processes (Warner et al., 1990; Deser et al., 2010; Chang and Cornillon, 2015), it is used to track climate variability and change (Jha et al., 2014; Pisano et al., 2020), and it is at the base of various chemical and biological processes (MacKenzie and Schiedek, 2007; Dong et al., 2022a). SST and the estimate of its gradients have also been proven to be a powerful tool to assess and investigate mesoscale and submesoscale variability (e.g., Bowen et al., 2002; Isern-Fontanet et al., 2006; González-Haro and Isern-Fontanet, 2014;

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


Rio et al., 2016; Castro et al., 2017; Ciani et al., 2020). Therefore, the availability of high resolution SST fields is crucial, since they serve as primary data source for many scientific and operational applications. However, their reliability is hindered by the limitations of infrared (IR) and microwave-based (MW) measurements (Minnett et al., 2019). In fact, thermal IR instruments are able to provide SST images at kilometer to sub-kilometer scale resolution, although their application is limited by cloud cover, aerosols radiation emission/absorption and scattering. Conversely, SST retrieval in the microwave is hampered only due to sunglint, rain or proximity to land, but the lower spatial resolution ($\simeq$ 25 km) achievable with present platforms represents a significant disadvantage. Higher resolution MW SSTs ($\simeq$ 15 km resolution) will only be available after the launch of the Copernicus Imaging Microwave Radiometer (Pearson et al., 2019), expected during 2029. Consequently, SST fields at high resolution are generally affected by several data voids. For this reason, a few statistical techniques have been developed to obtain gap-free SST images, mostly based on optimal interpolation (OI) (Bretherton et al., 1976). However, as a result of the temporal and spatial averaging applied during the interpolation, the effective resolution of the interpolated products can be significantly coarser than the nominal grid resolution, rarely getting down to less than a few tens of kilometres (Chin et al., 2017; Ciani et al., 2020; Yang et al., 2021). As such, providing interpolated data increases the accessibility of sea surface temperature fields for a wide community of users, but this improvement comes with a trade-off, as statistical interpolation leads to a strong smoothing of small scale ocean features.

In this context, we investigate here the potential of applying deep learning models to improve the effective resolution of a gap-free SST field, providing those small scale features even when direct measurements are missing. We exploit techniques generally used in the field of computer vision, which have proven very successful especially for processing gridded data, managing large-scale datasets while controlling the computational efficiency. In the field of oceanography, the research community started only recently to explore the applicability of machine learning (ML) methods to ocean remote sensing images (Dong et al., 2022b). The applications range from oil spill (Singha et al., 2013) to eddy detection (Lguensat et al., 2018; Duo et al., 2019) and parametrization (Bolton and Zanna, 2019), to marine algae species discrimination (Balado et al., 2021; Cui et al., 2022), to forecasting of ocean variables (Deo and Naidu, 1998; Ham et al., 2019) and estimation of meteorological parameters (Krasnopolsky et al., 2013; Zanna et al., 2019). Moreover, good results have been achieved from applying Neural Networks (NNs) to space-time interpolation and short-term forecasting issues with satellite altimetry data (Fablet et al., 2021), to high-performance description of turbulence processes (Mohan et al., 2020; Zanna and Bolton, 2020) and to SST reconstruction (Meng et al., 2021; Lloyd et al., 2021).

Among the image processing techniques, impressive performances have been obtained by Convolutional Neural Networks (CNNs) due to the high ability of this kind of networks to extract the most important information from two-dimensional spatial fields. Recently, the application of CNN architectures in the process of reconstructing High-Resolution (HR) images from Low-Resolution (LR) ones, the so-called single image Super Resolution (SR) problem, has attracted much attention in a wide range of scientific challenges. The idea is to implement a network that directly learns the end-to-end mapping between low and high resolution images. One of the simplest attempts made by Dong et al. (2015) was the construction of a network for image restoration composed by three 2D convolutional layers with different kernel size. The three layers might be seen as the three conceptual phases of this NN algorithm: a first extraction of overlapping tiles from the input images representing them

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


into features maps; the non-linear mapping of these maps onto one high-dimensional vector of high-resolution feature maps; the final reconstruction aggregating the above representations to generate the final high-resolution image. This final image is expected to be similar to the ground truth one. Despite the simplicity of the architecture, the SRCNN developed by Dong et al. (2015) achieved excellent performances with respect to more traditional methods and it has already been applied to reconstruct satellite-derived SST data, with promising results (Ducournau and Fablet, 2016). Building on Dong's work, several more complicated structures have been developed to tackle the Super Resolution problem. In the subsequent years, a few intents to develop residual networks have shown the convergence improvement of deeper architectures, mainly given by the introduction of skip-connections and recursive convolutions (He et al., 2016; Kim et al., 2016a, b). Similarly, a step forward has been made by Lim et al. (2017) with the development of an Enhanced Deep Residual Network for Super Resolution (EDSR) which makes use of residual blocks with constant scaling layers after the last convolutional layer, in order to stabilize the training even in presence of a large number of filters. This modification brought to a significantly better accuracy using much deeper networks, while controlling the computational cost of the training phase. A further step has been proposed by Liu et al. (2019) with the Adaptive Residual Blocks (ARBs), which replace the constant factor with adaptive residual factors, increasing the adaptability of the network. Recently, to further push these networks to efficiently handle different spatial scales in a multichannel input (each channel including a different variable with characteristic feature scales), Buongiorno Nardelli et al. (2022) introduced dilated convolutional multi-scale learning modules in the network developed by Liu et al. (2019), expanding the network receptive fields while still controlling its computational cost. In that work, the deep learning network was designed to super-resolve absolute dynamic topography (ADT) learning from both low resolution ADT and high resolution SST data, through an observing system simulation experiment (namely simulating all observations through an ocean general circulation numerical model).

Our aim is to exploit the ability of the dilated SR network to increase gap-free SST effective resolution, directly training our network on remote sensing SST data for both the input and target datasets. We make use of the data produced by the Italian National Research Council - Institute of Marine Sciences (CNR-ISMAR), within the Copernicus Marine Service, consisting in merged multi-sensor (L3S) and gap-free (L4) Sea Surface Temperature products over the Mediterranean Sea at high (HR, nominal 1/16° resolution) and ultra-high (UHR, nominal 1/100° resolution) spatial resolution (Buongiorno Nardelli et al., 2013). Considering that our UHR interpolation accounts for space-time decorrelation scales of 5 km and 1 day, in presence of valid UHR L3 observations the resulting L4 data can be considered as submesoscale resolving (Kurkin et al., 2020). UHR OI processing makes use of upsized HR L4 data as a background field. This first guess is left unchanged in absence of UHR L3 data and L4 effective resolution is thus by definition lower than 1/100° there. Our goal is thus to enhance the effective resolution of the upsized 1/16° background field.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


## 2 Materials and Methods

### 2.1 Training and test datasets

When dealing with deep learning methods it is very important how to construct the training and the test datasets, to ensure a sufficient generalization capability and, more specifically, to avoid under and over-fitting problems. Under-fitting occurs when the model fails to achieve a suitably low error on the training set, while over-fitting occurs when the gap between the training error and test error becomes excessively wide (Goodfellow et al., 2016). Remote sensing data are a very suitable resource to prevent the occurrence of these problems, due to the wide availability of large-scale gridded datasets which are complex enough to encapsulate an extensive variety of features.

The suite of products considered for this project provides the foundation SST (i.e., the temperature free of diurnal variability) over the Mediterranean Sea from 2008 to present (https://doi.org/10.48670/moi-00172) at Near Real Time (NRT). These data are built from level 2 (L2) infrared measurements (i.e., data in satellite native grid/swath coordinates) from several instruments on board of both geostationary and polar orbiting satellites, after the execution of a series of processing steps divided in different modules, as detailed in Buongiorno Nardelli et al. (2013). The final daily (night-time) gap-free field is provided by applying a space-time Optimal Interpolation algorithm to the L3 SST field for both HR and UHR products. As introduced before, in the case of the finer resolution product due to the small decorrelation scales assumed in the UHR interpolation, small scale features are correctly represented only when valid UHR L3 observations are present close to the interpolation point within a short temporal window. In the absence of those observations, the background field used by the OI algorithm is the L4 HR SST field, preliminary upsized onto a 1/100° regular grid (through a thin plate spline) but having an effective resolution equal or even lower than 1/16°.

Therefore, we train the network to improve the satellite SST effective resolution introducing realistic small scale features in the interpolated and upsized gap-free L4 HR SST images. The target data are derived from a ground-truth super-collated UHR SST dataset (L3S, namely merged multi-sensor data) specifically built for this purpose. The dataset is obtained by applying the CNR SST processing chain to acquisitions from the Sea and Land Surface Temperature Radiometer (SLSTR) on board of the Sentinel 3A and 3B missions, due to their high radiometric accuracy and km-scale resolving capabilities (Coppo et al., 2020).

Both input and target datasets are mapped on a regular grid at 1/100° spatial resolution over the Mediterranean Sea for the year 2020. Since our goal is to retrieve small scale features, a first moving average high pass filter (with a kernel radius of 200 km) is applied to remove the large scale dynamics in both input and target images (Kurkin et al., 2020). The data are then selected considering overlapping patches of dimensions 100 × 100 km, chosen extracting all the tiles containing at least 95% of valid pixels. SST values are then transformed into anomalies to avoid seasonal variability and scaled between -1 and 1 through a classical min-max normalization technique. The test dataset is finally selected separating the 15% of the tiles available after the preprocessing, chosen in order to be able to reconstruct the full geographical coverage of four days which are representative of different seasons. These data are separated in a fully independent dataset which is composed by 18576 tiles (374Mb). The training dataset finally consists of 94110 pairs of tiles (6.4Gb), of which one part (properly separated from the original training dataset at the beginning of the training phase, following a 85:15 ratio between the two) is used by the network in the validation

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


phase. In Figure 1 an example of the images used to construct the two datasets is shown. The top left panel depicts the gap-free first guess map on the 4th January 2020 and below we show the corresponding L3 image derived from merged Sentinel 3A and 3B data. On the right, we provide three examples of the pairs of tiles extracted from the related low and high resolution SST, fed as input to train the network.

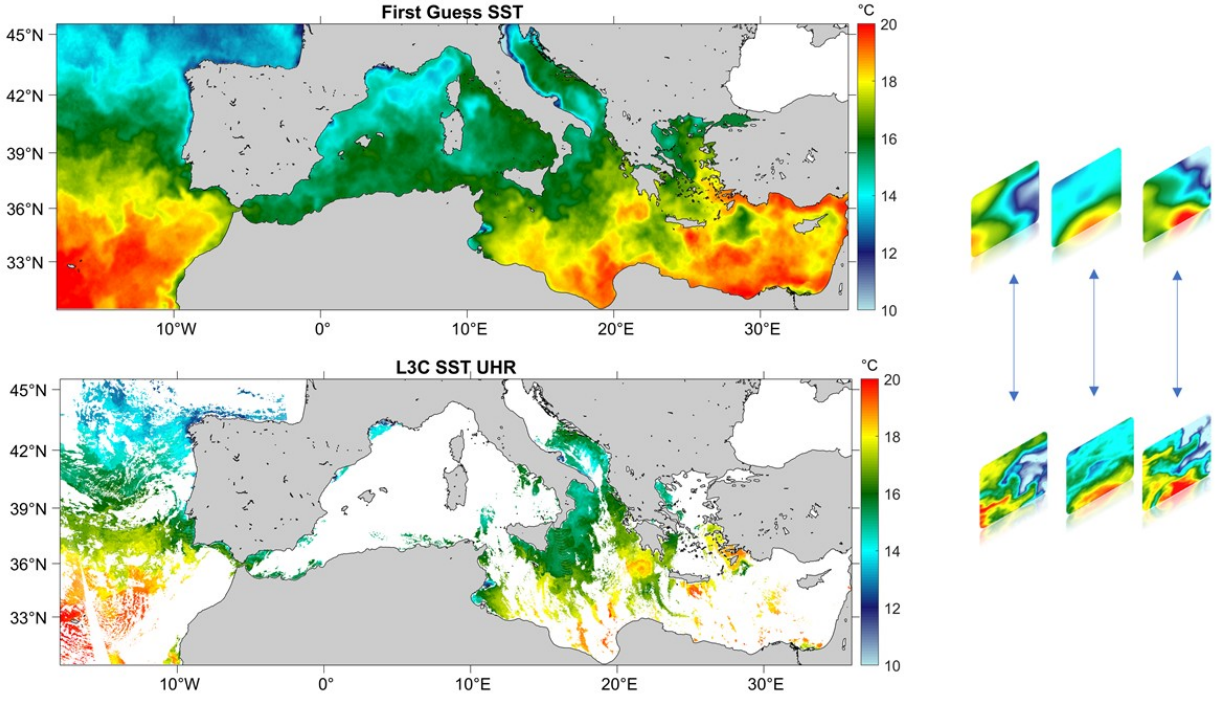


**Figure 1.** On the left the SST First Guess map used to extract the input tiles (top) and the SST L3C target image (bottom) on 4th January 2020. On the right three examples of extracted tiles used for the training and test datasets.

### 2.2 Super Resolution Convolutional Neural Network

In deep learning, Super Resolution algorithms are example-based methods which generate exemplar patches from the input image. As mentioned before, the application of Convolutional Neural Networks to the Super Resolution problem is based on networks that directly learn an end-to-end mapping between low and high resolution images. These networks consist in a series of interconnected layers which make use of the convolution operator simulating the connectivity pattern between neurons that we can find in the organization of the animal visual cortex. Formally, the output $Y$ of each layer $i$ is a function of a transformation of the previous layer output $X$, i.e.

$$Y = F(W_i * X + B_i), \tag{1}$$

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


where $F$ is the non-linear activation function (in this case the Rectified Linear Unit or ReLU), $W_i$ the weights, $B_i$ the biases and $*$ the convolution operator. The array of weights, generally called filters or kernel, is able to detect a specific type of feature in the input (generating what is called a features map) and might include for instance Gaussian-like filters or edge detectors along different directions, or any other kind of filter learned during the training. Having its own functionality, each layer will contain different structures. The main idea is that the network learns the weights of the filters and updates the parameters of the system through an optimization process based on minimizing the error between the output and the data from the validation set. This simple yet powerful idea can be complicated as much as desired, especially exploiting the potential of deep networks. The CNN implemented here was originally developed by Buongiorno Nardelli et al. (2022) and called dilated Adaptive Deep Residual Network for Super-Resolution (dADR-SR). In the dADR-SR network (schematized in Fig. 2) the low resolution input dataset is initially fed to three parallel dilated convolutional layers with the same number of filters (equal to $3 \times 3$) but increasing dilation factor (1, 3 and 5, respectively), which allows to have a larger receptive field, extracting information at different scales without increasing the number of parameters. After this first stage, the data pass through a sequence of twelve Multiscale Adaptive Residual Blocks (M-ARB), each one including two sets of parallel dilated convolutional layers (with 120 and 10 filters, respectively) and a Squeeze-and-Excitation (SE) module able to improve channel interdependencies at almost no computational cost (Hu et al., 2018), before being summed up to produce the final high resolution output. The SE block is a channel attention mechanism that weights each channel adaptively. It captures the global importance of each channel by initially squeezing the feature maps to a single numeric value (therefore obtaining a vector of size equal to the number of channels) and, finally, feeding this output to a two-layer "bottleneck" network which will produce new features maps, scaling each channel based on its importance. The training algorithm follows an early stopping rule which terminates the iterations as soon as the validation loss function increases for a previously chosen number of epochs (defined by the *patience* parameter which is set here equal to 20). An adaptive learning rate (initialized at $l_r = 10^{-4}$) is given by the Adam optimizer (Kingma and Ba, 2014), where the hyperparameters are set following the values found in most of the recent literature (Lim et al., 2017; Liu et al., 2019; Buongiorno Nardelli et al., 2022): the numerical stability constant is $\varepsilon = 10^{-8}$, the exponential decay rates for the first and the second moment estimates are set to $\beta_1 = 0.9$ and $\beta_2 = 0.999$, respectively. Instead of using the classical dropout regularization technique, a DropBlock strategy is implemented where contiguous region of a feature map are dropped together, which has been shown to increase the accuracy of the network for convolutional layers (Ghiasi et al., 2018). To evaluate the accuracy, the mean-squared error is used as reference in the loss function. The dADR-SR training model finally uses almost 1.6 M trainable parameters. All codes are written in Python using the deep learning framework *Keras* and the training was performed on a single NVIDIA T4 GPU in almost 4 days.

### 2.3 Evaluation of model performances

Three different error measures are calculated to evaluate the network reconstruction performance between the ground-truth image $x$ and the network output $y$.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


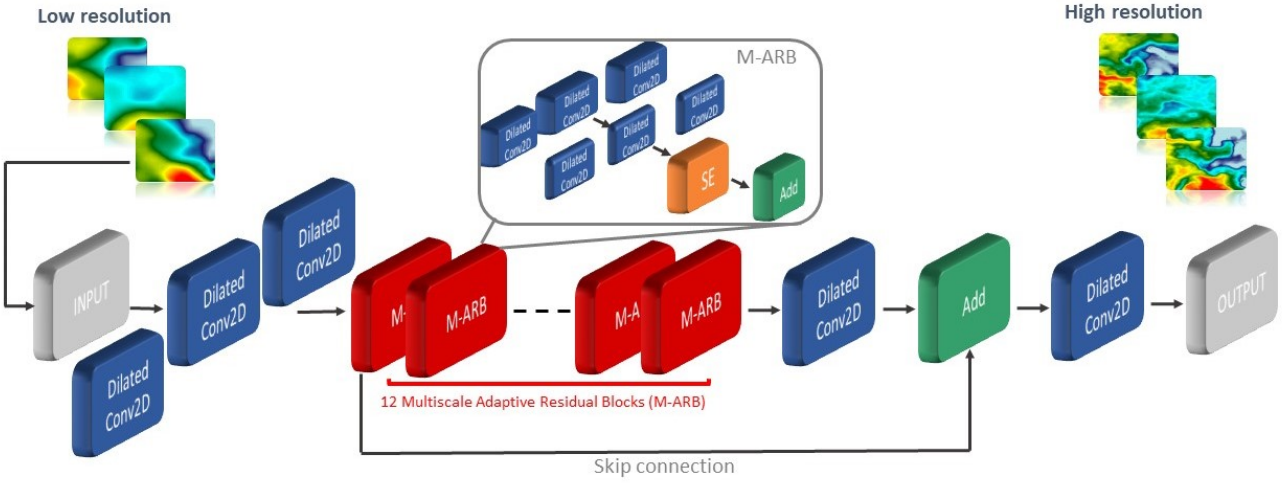


**Figure 2.** Schematic of the dilated Adaptive Deep Residual Network for Super-Resolution developed by Buongiorno Nardelli et al. (2022).

Firstly, we consider the classical Root Mean Squared Error (RMSE) given by:

$$RMSE(x,y) = \sqrt{\frac{\sum_i^N (x_i - y_i)}{N}}, \tag{2}$$

where $x_i$ and $y_i$ are the pixel $i$ of the images $x$ and $y$, respectively, and $N$ is the total number of pixels. This measure can be useful to evaluate the accuracy of a reconstructed value pixel by pixel, but it can be misleading in the assessment of the ocean state reconstruction, depending on the objectives of the application considered. For instance, if the NN is able to reproduce an ocean structure in a position which is slightly misplaced with respect to the ground-truth measurement, the RMSE will be high indicating a poor reconstruction, even worse than if it would have entirely missed the structure. This issue is often referred to as the double penalty issue, since point-matching measures will penalize the misplacement twice (where the structure should actually be and where is wrongly predicted). However, in some cases it is possible that capturing an ocean phenomenon, even if in a slightly wrong position, is better than missing it completely. For this reason, we also consider two additional measures usually considered in image processing.

The most commonly used measure for reconstructed image quality is the Peak Signal to Noise Ratio (PSNR), representing the ratio between the maximum possible pixel value of the image $I$ and the power of distorting noise that affects the quality of its representation, usually represented by the RMSE itself:

$$PSNR(x,y) = 20\log_{10}\left(\frac{\max(I)}{RMSE(x,y)}\right). \tag{3}$$

The PSNR can be seen as an approximation to human perception of reconstruction quality, where the higher the value the better the quality of the image.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


The third error measure is the structural similarity index measure (SSIM) proposed by Wang et al. (2004), widely used for measuring image quality and especially the similarity between two images. It is based on the idea that, if the PSNR estimates perceived errors to quantify the image degradation, the SSIM can capture perceived changes in the structural information variation. That is, if a reconstructed image is altered with different type of degradation (for instance, mean-shifted, blurred or with a salt-pepper impulsive noise effect), while the MSE will come out nearly identical for all the cases, the SSIM will capture the different perceptual quality, being a weighted combination of luminance, contrast and structure measurements. The mathematical formulation of the SSIM between two images $x$ and $y$ is given by:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \tag{4}$$

where $\mu_x$ and $\mu_y$ are the mean values of $x$ and $y$, respectively, $\sigma_{xy}$ is the cross-correlation of $x$ and $y$, $\sigma_x^2$ and $\sigma_y^2$ are the variance of $x$ and $y$, respectively, and $c_1$ and $c_2$ are the regularization constants for the luminance, contrast, and structural terms.

## 3 Results and discussion

Our aim is to verify whether the dADR-SR network, trained by means of satellite-derived observations, is able to improve the effective resolution of the SST fields in the areas where our interpolation technique removes most of the spatial variability associated with mesoscale and submesoscale processes.

Figure 3 shows the comparison of the result obtained on one of the SST fields included in the test dataset which corresponds to the SST daily map of 1st August 2020. In the top panel we can observe the reconstruction of the dADR-SR network at 1/100° spatial resolution over the Mediterranean Sea. In order to visually evaluate the reconstruction of the network, the corresponding L3 SST merged field observed by Sentinel 3A and 3B (central panel) and the First Guess map used in the optimal interpolation algorithm (bottom panel) are shown below the network output. We can observe that the SST features estimated by the CNN appears much sharper than the ones approximated by the low resolution map, showing promising capabilities to effectively reconstruct dynamical features.

To highlight the ability of the CNN method to more accurately capture small scale features with respect to the statistical algorithm, in Figure 4 we can observe in details smaller panels corresponding to a zoomed area of the same map of Fig. 3 delimited by the coordinates [30, 40]°N and [19, 36]°E, with the correspondent SST spatial gradients (calculated using the Sobel operator) for the dADR-SR reconstruction, L3S ground-truth data and the first guess approximation (from top to bottom, respectively). This particularly structure-rich area is an excellent example to demonstrate the ability of the network to capture dynamical processes which are quite clear in the high resolution SLSTR data (shown in the central panels). In fact, SST fronts are strongly connected with the surface dynamics and they are generally associated with energetic motions at the mesoscale and submesoscale. While the CNN is able to capture at least the most energetic structures (as shown in the top panels of Fig. 4), the optimal interpolation algorithm produces an extremely smooth field, where all the small scale features have been filtered out even when high resolution data are present.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


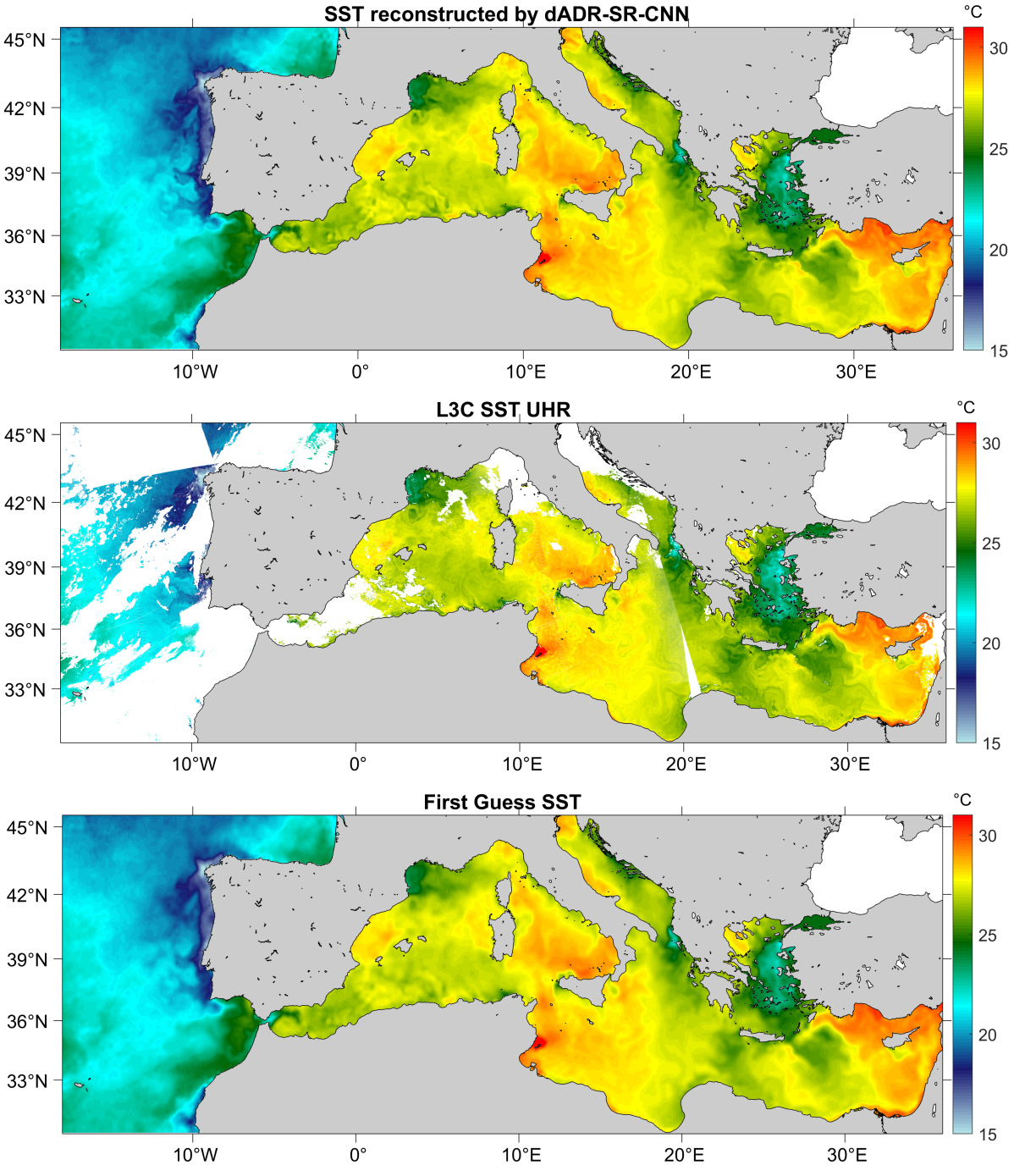


**Figure 3.** Comparison of the SST fields of 1st August 2020 provided by the reconstruction of the dADR-SR-CNN (top panel), the collated L3 data measured by Sentinel 3A and 3B (central panel) and the Optimal Interpolated First Guess (bottom panel).

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


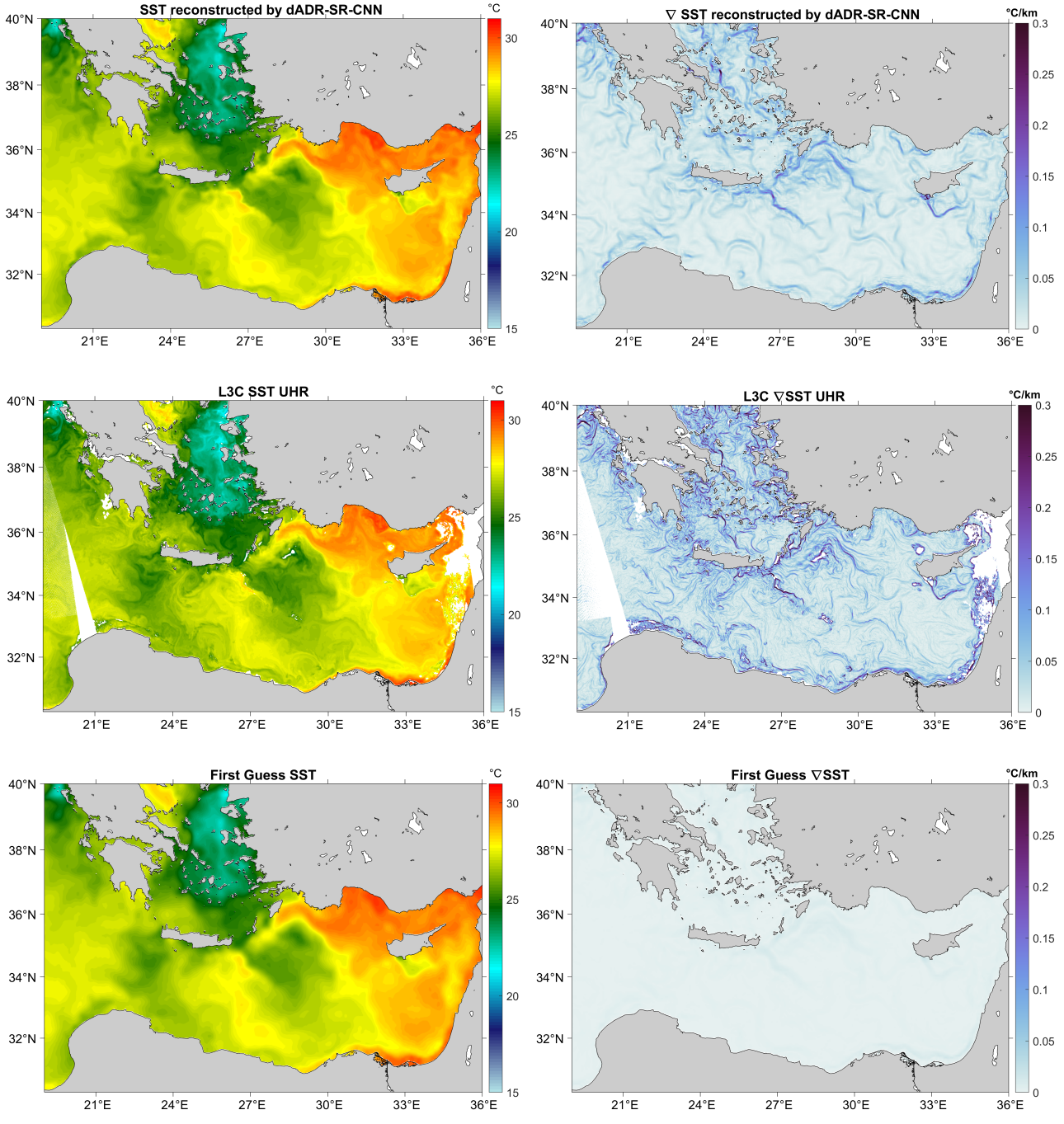


**Figure 4.** Comparison of the SST (on the left) and SST gradients (on the right) provided by the reconstruction of the dADR-SR-CNN (top panels), the collated L3 data measured by Sentinel 3A and 3B (central panels) and the Optimal Interpolated First Guess (bottom panels) in the selected region delimited by the coordinates [30, 40]°N in latitude and [19, 36]°E in longitude.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


This visual analysis is quantitatively confirmed by the maps in Figure 5, displaying the difference between the error made by the low resolution approximation and the dADR-SR model, respectively, with respect to the original L3 image, averaged on 1° × 1° boxes. Here, red indicates an improvement of the CNN with respect to the low resolution image and blue a degradation. A clear predominance of red boxes is found both in the SST and the SST gradients error maps.

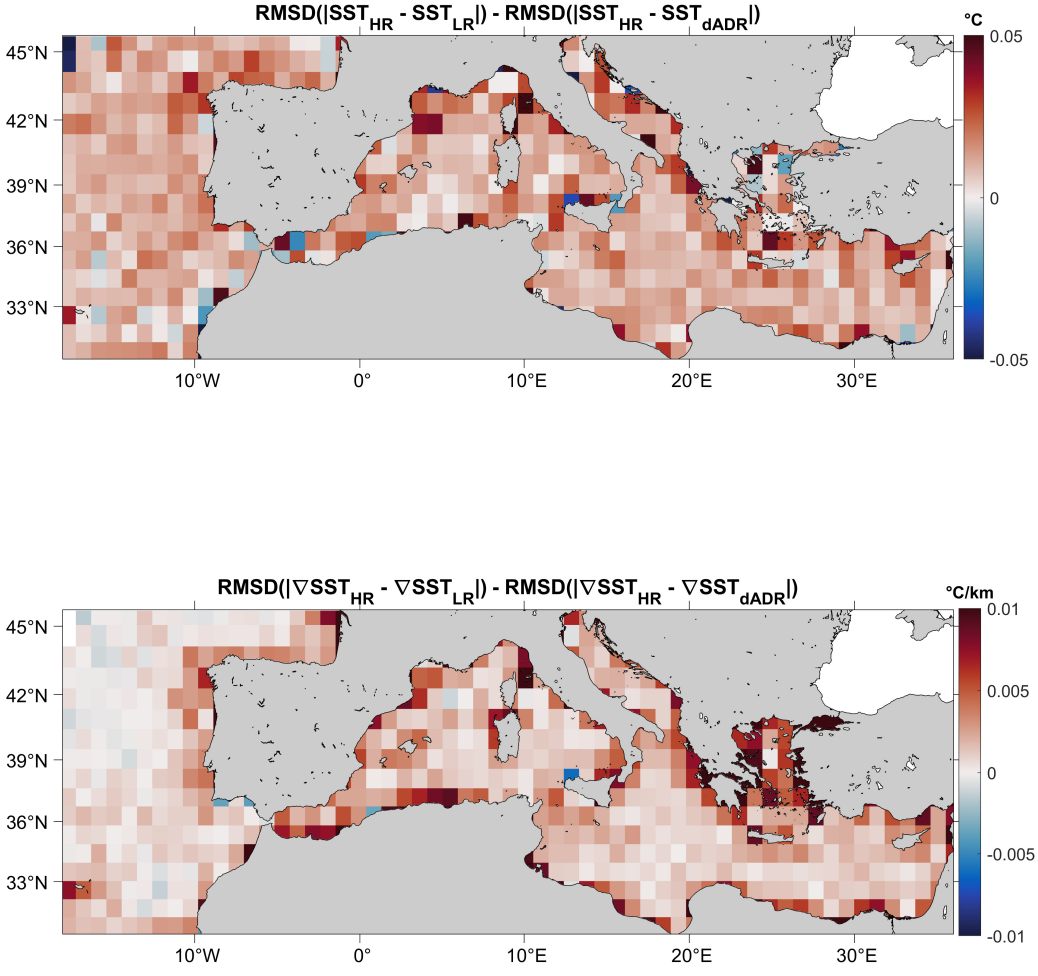


**Figure 5.** Comparison of the performance of the SST (top) and $\nabla SST$ (bottom) dADR-SR reconstruction and the L4 First Guess with respect to the L3 data measured by Sentinel 3A and 3B satellites. Red positive values show an improvement of the network reconstruction with respect to the optimal interpolated First Guess.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


**Table 1.** Error estimations of the SST dADR-SR output and the First Guess map with respect to the L3S ground-truth: the RMSE given by Equations (2) and the corresponding confidence interval calculated by the bootstrapping procedure, the Peak-Signal to Noise Ratio obtained by (3) and the Structural Similarity Index Measure given by (4).

| Model | RMSE ($^{\circ}C$) | PSNR | SSIM |
|---|---|---|---|
| **dADRSR** | $0.31 \pm 7 \times 10^{-5}$ | 37.9 | 0.54 |
| **Low Resolution** | $0.33 \pm 7 \times 10^{-5}$ | 37.5 | 0.53 |

In Table 1, the comparison is summarized quantitatively, with the network reconstruction presenting a RMSE = 0.31°C (against the 0.33°C obtained from the OI approximation), a mean PSNR equal to 37.9 and a SSIM equal to 0.54, both larger than the low resolution result. We recall that while the RMSE should be low to ensure a good approximation, for the other two quantities high values indicate an improvement.

To analyse the effectiveness of the high resolution reconstruction, we compare the Power Spectral Density (PSD) of the network output with the L3 product and the First Guess map over three selected zones. We compute the PSD via Fast Fourier Transform (FFT) with a Blackman–Harris window over the three areas corresponding to the boxes with labels *a*, *b* and *c* in Figure 6. The zones are chosen in order to have the maximum number of valid pixels available in the L3 observations and represent different dynamical regimes:

a - A region over the Sea of Sardinia with a smooth spatial variability.

b - A region over the Ionian Sea with an important SST variability between the eastern and the western part of the area.

c - A region over the Levantine Sea characterized by small scale structures.

The three central panels of Figure 6 show the PSD, presented as a function of the wavenumber, of the SST reconstructed by the dADR-SR (in yellow), the first guess map (in red) and the high resolution observations (in blue) over the three zones delimited by the black rectangles over the L3 SST field on the top panel. In all the cases we can observe that the PSD follows the same behavior for wavenumbers smaller than 1 $\text{deg}^{-1}$, i.e. for scales higher than 100 km. This means that for such scales the SST field reconstructed by both the CNN and the OI algorithm equally characterize the large mesoscale features. Conversely, for the regions over the boxes *a* and *b*, both the Low Resolution and the network reconstruction exhibit an important PSD decrease for wavenumbers higher than 1 $\text{deg}^{-1}$, but starting from approximately 10 $\text{deg}^{-1}$, the LR spectrum separates from the CNN spectrum (and consequently increases the distance from the HR one), indicating a poorer reconstruction of spatial features below 10 km. The CNN spectrum, on the other hand, shows that the machine learning algorithm is able to capture those small scale features in all the cases. An analogous behaviour is found for the spectra of the SST gradients over the same three regions (three bottom panels). The abrupt decreases that we can observe are probably due to the artifacts introduced by the re-gridding, as already mentioned in Liberti et al. (2023), leading to the lack of physical meaning for the spectra from the "bumps" onwards (i.e. for the highest wavenumbers).

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


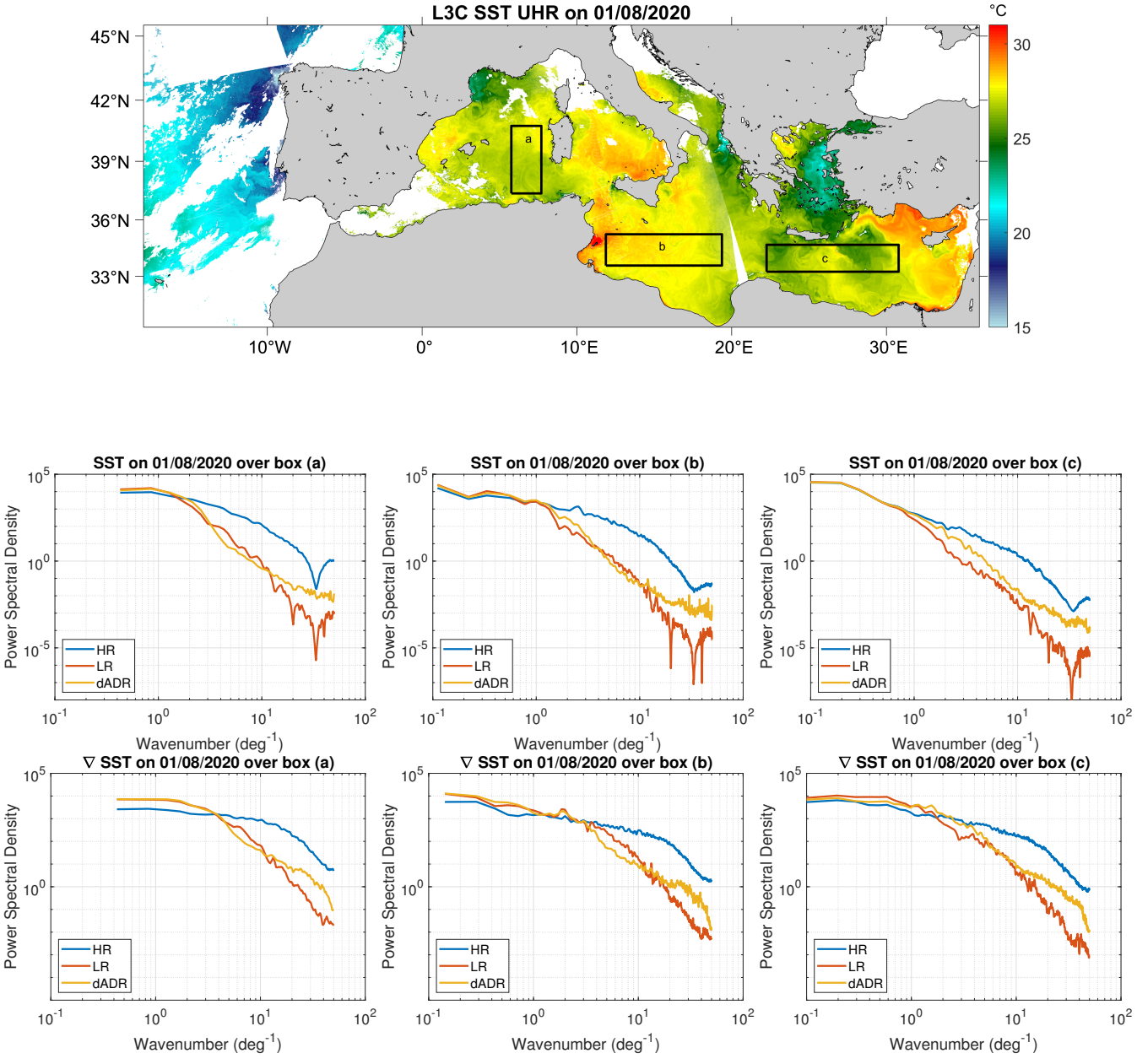


**Figure 6.** The PSD of the SST (central panels) and SST gradients (bottom panels) reconstructed by the dADR-SR (in yellow), the first guess map (in red) and the high resolution observations (in blue) over the three zones delimited by black rectangles on the L3 SST field on 1st August 2020 in the top panel.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


## 4 Conclusions

The advancements obtained by the application of novel machine learning-based techniques for the improvement of the effective resolution of remote sensing observations have recently opened a new way to approach to satellite-derived data processing. The great advantages provided by making high resolution gap-free images available for a wide range of scientific users are severely limited by the number of valid L3 observations. In the case of sea surface temperature measurements, infrared data are commonly contaminated by cloud cover, reducing the quality of the L4 data that can be obtained via statistical interpolation techniques. The machine learning approach used here exploits the progresses made in the field of computer vision for extrapolating high resolution features even when a direct measurement is missing. Learning directly from ground-truth data, and taking advantage of both dilated convolution and attention mechanisms, the deep neural network employed here proved able to reproduce small scale signals generally smoothed out by Optimal Interpolation algorithms. The strong variability of the SST in the Mediterranean Sea allowed us to obtain excellent results even considering just one year of data during the training phase. However, it would be important to investigate whether using longer time series may help to improve the network ability to reconstruct SST fields, as well as to rely on more robust statistics. Moreover, given the inhomogeneity of the spatial error distribution related to the interpolation technique, it would be interesting to expand the present investigation in order to take into account the OI error field as an additional predictor and to consider the contribution of the error of the SST gradients in the loss function. Another aspect that would deserve further investigation concerns the applicability of the dADR-SR network to different sea/ocean areas, even though a fine-tuning of the model would probably be needed.

The results achieved here, however, may already benefit a wide range of applications. Super-resolved SST fields would facilitate the challenging task of 2D/3D ocean dynamics reconstruction in synergy with other variables (e.g., Buongiorno Nardelli et al., 2022; Fablet et al., 2023) or the monitoring of ocean fronts in areas of particular interest (e.g., areas affected by vertical exchange and upwelling regions). To enhance the effective resolution of SST data and especially SST gradients may also benefit data assimilation in forecast modelling, given their proven sensibility to small structures of sea surface temperature (Woollings et al., 2010). We also plan to validate our results exploiting the high resolution SST data derived by the CNN reconstruction within the operational SST chain in the framework of the Copernicus Marine Service.

*Author contributions.* Conceptualization, C.F., A.P. and B.B.N; methodology, C.F., D.C., A.P. and B.B.N.; investigation, C.F., D.C., A.P. and B.B.N.; visualization, C.F.; writing—original draft, C.F.; writing—review and editing, C.F., D.C., A.P. and B.B.N. All authors have read and agreed to the published version of the manuscript.

*Financial support.* This work has been supported by the Copernicus Marine Service, funded through contract agreement no. 21001L03-COP-TAC SST-2300– Lot 3: Provision of Sea Surface Temperature Observation Products (SST-TAC).

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


*Competing interests.* The authors declare no conflict of interest.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


## References

Balado, J., Olabarria, C., Martínez-Sánchez, J., Rodríguez-Pérez, J. R., and Pedro, A.: Semantic segmentation of major macroalgae in coastal environments using high-resolution ground imagery and deep learning, International Journal of Remote Sensing, 42, 1785–1800, 2021.

Bolton, T. and Zanna, L.: Applications of deep learning to ocean data inference and subgrid parameterization, Journal of Advances in Modeling Earth Systems, 11, 376–399, 2019.

Bowen, M. M., Emery, W. J., Wilkin, J. L., Tildesley, P. C., Barton, I. J., and Knewtson, R.: Extracting multiyear surface currents from sequential thermal imagery using the maximum cross-correlation technique, Journal of Atmospheric and Oceanic Technology, 19, 1665–1676, 2002.

Bretherton, F. P., Davis, R. E., and Fandry, C.: A technique for objective analysis and design of oceanographic experiments applied to MODE-73, in: Deep Sea Research and Oceanographic Abstracts, vol. 23, pp. 559–582, Elsevier, 1976.

Buongiorno Nardelli, B., Tronconi, C., Pisano, A., and Santoleri, R.: High and Ultra-High resolution processing of satellite Sea Surface Temperature data over Southern European Seas in the framework of MyOcean project, Remote Sensing of Environment, 129, 1–16, 2013.

Buongiorno Nardelli, B., Cavaliere, D., Charles, E., and Ciani, D.: Super-resolving ocean dynamics from space with computer vision algorithms, Remote Sensing, 14, 1159, 2022.

Castro, S. L., Emery, W. J., Wick, G. A., and Tandy, W.: Submesoscale sea surface temperature variability from UAV and satellite measurements, Remote Sensing, 9, 1089, 2017.

Chang, Y. and Cornillon, P.: A comparison of satellite-derived sea surface temperature fronts using two edge detection algorithms, Deep Sea Research Part II: Topical Studies in Oceanography, 119, 40–47, 2015.

Chin, T. M., Vazquez-Cuervo, J., and Armstrong, E. M.: A multi-scale high-resolution analysis of global sea surface temperature, Remote sensing of environment, 200, 154–169, 2017.

Ciani, D., Rio, M.-H., Nardelli, B. B., Etienne, H., and Santoleri, R.: Improving the altimeter-derived surface currents using sea surface temperature (SST) data: A sensitivity study to SST products, Remote Sensing, 12, 1601, 2020.

Coppo, P., Brandani, F., Faraci, M., Sarti, F., Dami, M., Chiarantini, L., Ponticelli, B., Giunti, L., Fossati, E., and Cosi, M.: Leonardo spaceborne infrared payloads for Earth observation: SLSTRs for Copernicus Sentinel 3 and PRISMA hyperspectral camera for PRISMA satellite, Applied Optics, 59, 6888–6901, 2020.

Cui, B., Zhang, H., Jing, W., Liu, H., and Cui, J.: SRSe-net: Super-resolution-based semantic segmentation network for green tide extraction, Remote Sensing, 14, 710, 2022.

Deo, M. and Naidu, C. S.: Real time wave forecasting using neural networks, Ocean engineering, 26, 191–203, 1998.

Deser, C., Alexander, M. A., Xie, S.-P., and Phillips, A. S.: Sea surface temperature variability: Patterns and mechanisms, Annual review of marine science, 2, 115–143, 2010.

Dong, C., Loy, C. C., He, K., and Tang, X.: Image super-resolution using deep convolutional networks, IEEE transactions on pattern analysis and machine intelligence, 38, 295–307, 2015.

Dong, C., Liu, L., Nencioli, F., Bethel, B. J., Liu, Y., Xu, G., Ma, J., Ji, J., Sun, W., Shan, H., et al.: The near-global ocean mesoscale eddy atmospheric-oceanic-biological interaction observational dataset, Scientific Data, 9, 436, 2022a.

Dong, C., Xu, G., Han, G., Bethel, B. J., Xie, W., and Zhou, S.: Recent developments in artificial intelligence in oceanography, Ocean-Land-Atmosphere Research, 2022, 2022b.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


Ducournau, A. and Fablet, R.: Deep learning for ocean remote sensing: an application of convolutional neural networks for super-resolution on satellite-derived SST data, in: 2016 9th IAPR Workshop on Pattern Recogniton in Remote Sensing (PRRS), pp. 1–6, IEEE, 2016.

Duo, Z., Wang, W., and Wang, H.: Oceanic mesoscale eddy detection method based on deep learning, Remote Sensing, 11, 1921, 2019.

Fablet, R., Amar, M., Febvre, Q., Beauchamp, M., and Chapron, B.: End-to-end physics-informed representation learning for satellite ocean remote sensing data: Applications to satellite altimetry and sea surface currents, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 3, 295–302, 2021.

Fablet, R., Febvre, Q., and Chapron, B.: Multimodal 4DVarNets for the reconstruction of sea surface dynamics from SST-SSH synergies, IEEE Transactions on Geoscience and Remote Sensing, 2023.

Ghiasi, G., Lin, T.-Y., and Le, Q. V.: Dropblock: A regularization method for convolutional networks, Advances in neural information processing systems, 31, 2018.

González-Haro, C. and Isern-Fontanet, J.: Global ocean current reconstruction from altimetric and microwave SST measurements, Journal of Geophysical Research: Oceans, 119, 3378–3391, 2014.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep learning, MIT press, 2016.

Ham, Y.-G., Kim, J.-H., and Luo, J.-J.: Deep learning for multi-year ENSO forecasts, Nature, 573, 568–572, 2019.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, 2016.

Hu, J., Shen, L., and Sun, G.: Squeeze-and-excitation networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7132–7141, 2018.

Isern-Fontanet, J., Chapron, B., Lapeyre, G., and Klein, P.: Potential use of microwave sea surface temperatures for the estimation of ocean currents, Geophysical research letters, 33, 2006.

Jha, B., Hu, Z.-Z., and Kumar, A.: SST and ENSO variability and change simulated in historical experiments of CMIP5 models, Climate dynamics, 42, 2113–2124, 2014.

Kim, J., Lee, J. K., and Lee, K. M.: Accurate image super-resolution using very deep convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1646–1654, 2016a.

Kim, J., Lee, J. K., and Lee, K. M.: Deeply-recursive convolutional network for image super-resolution, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1637–1645, 2016b.

Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980, 2014.

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., and Belochitski, A. A.: Using ensemble of neural networks to learn stochastic convection parameterizations for climate and numerical weather prediction models from data simulated by a cloud resolving model, Advances in Artificial Neural Systems, 2013, 5–5, 2013.

Kurkin, A., Kurkina, O., Rybin, A., and Talipova, T.: Comparative analysis of the first baroclinic Rossby radius in the Baltic, Black, Okhotsk, and Mediterranean seas, Russian Journal of Earth Sciences, 20, 8, 2020.

Lguensat, R., Sun, M., Fablet, R., Tandeo, P., Mason, E., and Chen, G.: EddyNet: A deep neural network for pixel-wise classification of oceanic eddies, in: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, pp. 1764–1767, IEEE, 2018.

Liberti, G. L., Sabatini, M., Wethey, D. S., and Ciani, D.: A Multi-Pixel Split-Window Approach to Sea Surface Temperature Retrieval from Thermal Imagers with Relatively High Radiometric Noise: Preliminary Studies, Remote Sensing, 15, 2453, 2023.

Lim, B., Son, S., Kim, H., Nah, S., and Mu Lee, K.: Enhanced deep residual networks for single image super-resolution, in: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 136–144, 2017.

https://doi.org/10.5194/egusphere-2024-455
Preprint. Discussion started: 20 February 2024


Liu, S., Gang, R., Li, C., and Song, R.: Adaptive deep residual network for single image super-resolution, Computational Visual Media, 5, 391–401, 2019.

Lloyd, D. T., Abela, A., Farrugia, R. A., Galea, A., and Valentino, G.: Optically enhanced super-resolution of sea surface temperature using deep learning, IEEE Transactions on Geoscience and Remote Sensing, 60, 1–14, 2021.

MacKenzie, B. R. and Schiedek, D.: Long-term sea surface temperature baselines—time series, spatial covariation and implications for biological processes, Journal of Marine Systems, 68, 405–420, 2007.

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

Minnett, P., Alvera-Azcárate, A., Chin, T., Corlett, G., Gentemann, C., Karagali, I., Li, X., Marsouin, A., Marullo, S., Maturi, E., et al.: Half a century of satellite remote sensing of sea-surface temperature, Remote Sensing of Environment, 233, 111 366, 2019.

Mohan, A. T., Lubbers, N., Livescu, D., and Chertkov, M.: Embedding hard physical constraints in neural network coarse-graining of 3D turbulence, arXiv preprint arXiv:2002.00021, 2020.

Pearson, K., Good, S., Merchant, C. J., Prigent, C., Embury, O., and Donlon, C.: Sea surface temperature in global analyses: Gains from the Copernicus Imaging Microwave Radiometer, Remote Sensing, 11, 2362, 2019.

Pisano, A., Marullo, S., Artale, V., Falcini, F., Yang, C., Leonelli, F. E., Santoleri, R., and Buongiorno Nardelli, B.: New evidence of Mediterranean climate change and variability from sea surface temperature observations, Remote Sensing, 12, 132, 2020.

Rio, M.-H., Santoleri, R., Bourdalle-Badie, R., Griffa, A., Piterbarg, L., and Taburet, G.: 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, 2016.

Singha, S., Bellerby, T. J., and Trieschmann, O.: Satellite oil spill detection using artificial neural networks, IEEE Journal of selected topics in applied earth observations and remote sensing, 6, 2355–2363, 2013.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P.: Image quality assessment: from error visibility to structural similarity, IEEE transactions on image processing, 13, 600–612, 2004.

Warner, T. T., Lakhtakia, M. N., Doyle, J. D., and Pearson, R. A.: Marine atmospheric boundary layer circulations forced by Gulf Stream sea surface temperature gradients, Monthly Weather Review, 118, 309–323, 1990.

Woollings, T., Hoskins, B., Blackburn, M., Hassell, D., and Hodges, K.: Storm track sensitivity to sea surface temperature resolution in a regional atmosphere model, Climate dynamics, 35, 341–353, 2010.

Yang, C., Leonelli, F. E., Marullo, S., Artale, V., Beggs, H., Buongiorno Nardelli, B., Chin, T. M., De Toma, V., Good, S., Huang, B., et al.: Sea surface temperature intercomparison in the framework of the Copernicus Climate Change Service (C3S), Journal of Climate, 34, 5257–5283, 2021.

Zanna, L. and Bolton, T.: Data-driven equation discovery of ocean mesoscale closures, Geophysical Research Letters, 47, e2020GL088 376, 2020.

Zanna, L., Brankart, J., Huber, M., Leroux, S., Penduff, T., and Williams, P.: Uncertainty and scale interactions in ocean ensembles: From seasonal forecasts to multidecadal climate predictions, Quarterly Journal of the Royal Meteorological Society, 145, 160–175, 2019.