Proceedings of the ASME 2025
44th International Conference on Ocean, Offshore & Arctic Engineering
OMAE2025
June 22–27, 2025, Vancouver, British Columbia, Canada

OMAE2025-156721

DATA-DRIVEN PHASE-RESOLVED SEA SURFACE RECONSTRUCTION FROM SYNTHETIC X-BAND RADAR DATA

Mathies Wedler[1,*], Nicolas Desmars[1], Merten Stender[2], Sören Ehlers[1], Marco Klein[1]

[1]German Aerospace Center (DLR), Institute of Maritime Energy Systems, Düneberger Str. 108, 21502 Geesthacht, Germany
[2]Technische Universität Berlin, Cyber-Physical Systems in Mechanical Engineering, Straße des 17. Juni 135, Berlin, Germany

ABSTRACT

Accurate phase-resolved information on the sea surface can be crucial for well-founded decision making in the maritime sector. This information can be obtained by analyzing X-band radar measurements, which contain sparse information on the surrounding sea. This analysis, referred to as sea surface reconstruction, usually relies on simplified physical models or computationally expensive optimization procedures, thus creating a trade-off between accuracy and computational cost. This work proposes a purely data-driven approach, which aims at providing accurate sea surface reconstruction from X-band radar data in real-time. For this, state-of-the-art methods from computer vision and deep learning were combined to a model, which maps successive historic radar images to the 2D phase-resolved sea surface. The training data were generated using the high-order spectral method to model nonlinear hydrodynamic effects, and features a wide range of governing sea-state conditions. The synthetic radar images were derived from these simulations using a numeric radar model. The results show that the proposed data-driven approach is capable of faithfully reconstructing the 2D sea surface from sparse radar information over a wide range of governing sea state parameters. Moreover, the approach is able to extrapolate over the radar blind zone, yielding a complete reconstruction of the sea surface within the radar radius.

Keywords: Deep Learning, X-band radar, phase-resolved sea surface reconstruction, high-order spectral method

1. INTRODUCTION

Phase-resolved wave predictions have manifold applications in the maritime sector, ranging from decision support for safety-related operations like boarding, lifting or helicopter landing, to the active control of wave energy converters. A common pipeline for phase-resolved wave predictions involves two essential steps, the *reconstruction* of the phase-resolved sea surface from information on the surrounding wave field, followed by the spatio-temporal propagation or *prediction* based on the reconstructed sea surface [1–6]. The non-coherent X-band marine radar is a common remote sensing system to gather information on the surrounding wave field. The received signal or *backscatter* is a highly modulated map of the sea surface within a radius of 2–5 km [7]. The dominant modulation effects are caused by the alignment of the radar antenna relative to the illuminated wave facets, as well as the occlusion of waves further away. The challenge in sea surface reconstruction is the *accurate* and *ahead-of-time* translation of this sparse map to the phase-resolved wave field. The accuracy is of particular importance, as the reconstruction serves as the initial condition for the wave prediction and thus constraints its fidelity. While reconstruction approaches based on the 3D FFT reportedly achieved correlations above 90% w.r.t. the real wave [1, 7], studies have shown that the reconstruction phase remains the limiting factor for real-time capable phase-resolved wave predictions [2, 8].

This shortcoming in traditional methods has motivated the application of data-driven methods for the sea surface reconstruction from X-band marine radar [6, 9]. Machine learning (ML) resembles a class of data-driven methods, which are capable of approximating any continuous function to arbitrary accuracy [10, 11]. Classical ML approaches have no physical priors but learn descriptive pattern directly from data. This results in a highly flexible approach which does not rely on the existance of physical models, and eliminates the dependence on fine time steps known from the numeric world. The majority of computational cost is shifted to the training phase, the trained model is a fixed mapping operator with minimal online cost. Ehlers et al. [6] have presented a study on the data-driven phase-resolved sea surface reconstruction from upstream X-band radar information in 1D. The work has compared two ML models on the task, the U-Net [12] and the Fourier neural operator (FNO) [13]. The synthetic radar data has been generated using the nonlinear high-order spectral method [14, 15] and a numeric radar tool that applies geometric modulation effects. While both models have yielded highly accurate sea surface reconstructions, the FNO has been

*Corresponding author: mathies.wedler@dlr.de

identified to be slightly superior to the U-Net. Zhao et al. [9] have attempted the 2D phase-resolved sea surface reconstruction using a 2D U-Net [12] based on a single radar image. The underlying synthetic wave data has been generated by superposition of single harmonic waves, the synthetic radar data has been derived from the linear waves using a geometric radar model. The results have indicated high reconstruction accuracy within the range of the radar, excluding the inner radar blind zone.

This work builds on the prior data-driven attempts [6, 9] and aims at combining their strengths into a single powerful approach to the 2D phase-resolved sea surface reconstruction from synthetic X-band radar. In constrast to the prior approaches [6, 9], the sea surface reconstruction was attempted within the whole radar radius, i.e., the reconstruction involves the radar blind zone. The model consequently had to generalize over the unknown area around the radar location. The wave- and radar model described in Section 2 established a data basis, that reflects the real-world problem as closely as possible. We are confident, that a high quality data basis will result in a superior data-driven reconstruction of the sea surface. The wave model yielded fully evolved wind-induced nonlinear waves over a wide range of characteristic sea state parameters. The numeric radar was modeled based on the radar equation and augmented the commonly used geometric approach to a numeric radar with system losses. Section 3 subsequently introduces the ML framework that was used for the sea surface reconstruction. Section 4 starts with a comparison of 9 different training setups to estimate the optimal number of consecutive radar snapshots for the reconstruction. Afterwards, the best performing model is evaluated in detail. In Section 5, the conclusions and a brief outlook are provided.

2. DATA GENERATION

Supervised ML approaches attempt to learn patterns from data, and apply this knowledge to solve new problems within the same problem space [16, 17]. The quality of the training data is consequently of particular importantance, as it constraints the quality of the ML model itself. The most holistic representation of any physical problem is arguably measurement data from real-world processes. However, the collection of real-world data in sufficient quality and quantity to train an ML model is, depending on the problem, often not feasible. In those cases, numerical methods are common practice for the targeted generation of synthetic data.

The scope of this work is the phase-resolved sea surface reconstruction from synthetic X-band radar data. During the training, the model attempted to map the sparse X-band radar data to the corresponding sea surface data. Since the collection of ground truth data, i.e., sea surface elevations that are sampled at equidistant grid points over a large domain, is not feasible, nonlinear sea surface simulations substitute the absence real-world data. The synthetic X-band radar data were derived from the sea surface data using a numeric radar model, that aims at closely resembling the characteristics of a real X-band radar apparatus.

2.1 Nonlinear wave model

The sea surface data served as the baseline for this work, as it ultimately determined the quality of the synthetic radar data, and thus the ML model. To achieve the best possible results, the sea surface data has to faithfully depict the dynamics of water waves. In this work, the high-order spectral (HOS) method [14, 15] was used to simulate the time evolution of water waves. The HOS method takes into account physically existing (resonant and non-resonant) nonlinear interactions, and is thus capable of generating high-fidelity simulations of sea surface dynamics.

The underlying wave model is based on potential flow theory. The flow characteristcs are described by a velocity potential $\phi(x, y, z, t)$, which, under the assumption of irrotational flow and incompressible and inviscid fluid, satisfies the Laplace equation

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0, \tag{1}$$

where $z = 0$ is the mean free surface elevation. The boundary value problem can be expressed in terms of the kinematic and dynamic boundary conditions at the free surface $z = 0$

$$\phi_z - \phi_x\eta_x - \phi_y\eta_y - \eta_t = 0 \quad \text{and} \quad \phi_t + \frac{1}{2}(\nabla\phi)^2 + g\eta = 0, \tag{2}$$

as well as the bottom boundary condition $\phi_z = 0$ at the seabed. The HOS method estimates the kinematic and dynamic boundary conditions (2) as a pertubation series of arbitrary nonlinear order M around $z = 0$. Former studies have shown, that order $M \leq 4$ models all relevant nonlinear effects [2, 18].

In this work, the numerically efficient pseudo-spectral approach which has been proposed by West et al. [15] was used for the nonlinear simulation of water waves with order $M = 4$. The nonlinear HOS simulations were initialized with linear sea states, that were generated using superposition of harmonic waves [19] with random phase $\varphi \in [0, 2\pi]$, and amplitudes sampled from a JONSWAP-TMA [20, 21] spectrum. The space of characteristic sea state parameters was spanned by three wave spreadings $n \in \{4, 30, 300\}$ with the continuous directional spreading function [22]

$$G(\theta, n) = \cos(\theta - \theta_0)^n \left(\frac{\Gamma(n/2 + 1)}{\sqrt{\pi}\Gamma(n/2 + 0.5)} \right), \tag{3}$$

where higher n lead to less directional spreading, and thus more long-crested waves [19]. For each spreading, 11 peak wavelengths $L_\mathrm{p} \in \{100, 110, 120, \ldots, 200\}$ m, and 26 wave steepness values $\varepsilon \in \{0.01, 0.015, 0.02, \ldots, 0.135\}$ were defined. The peak wavelength corresponds to the wavenumber $k_\mathrm{p} = 2\pi/L_\mathrm{p}$, at which the peak of the wave spectrum is located. The steepness $\varepsilon = \pi H_\mathrm{S}/L_\mathrm{p}$ is proportional to the ratio of the significant wave height H_S and the peak wavelength, and serves as a dimensionless measures for the steepness of the wave slopes. The simulation grid resolved the square domain of size (4000×4000) m at 1024×1024 grid points. For numerical stability, the HOS simulations were performed with a relaxation period of $T_\mathrm{a} = 10T_\mathrm{p}$, throughout which the nonlinearities were gradually ramped up [23]. The fully developed nonlinear sea states were simulated for $t_\mathrm{sim} = 100$ s with a saving interval of $dt = 0.1$ s.

2.2 X-band radar model

The synthetic X-band radar data were derived from the HOS sea surface simulations using a numeric radar model. This radar

model was based on the radar equation [24]

$$P_\mathrm{r} = \frac{P_\mathrm{t} G^2 \lambda^2 \sigma}{(4\pi)^3 R^4 \prod \mathcal{L}_i}, \tag{4}$$

which estimates the received power P_r from the transmitted power P_t, antenna gain G, radar wavelength λ, radar cross section (RCS) σ, distance R, and various loss terms $\mathcal{L}_i$. X-band radar operates in the frequency range $f \in [8, 12]$ GHz [24, 25]. Here, an operating frequency of $f = 10\,\mathrm{GHz}$ was assumed, such that $\lambda \approx 0.03\,\mathrm{m}$. The RCS σ is a measure for the amount of information that is reflected by the target and received by the radar antenna. The amount of information reflected by the environment mainly depends on the relative orientation of the antenna and the illuminated wave facet [26]. This can be geometrically expressed in terms of the angle between the radar beam vector $\mathbf{u}(\mathbf{r}, t)$ and the sea surface normals $\mathbf{n}(\mathbf{r}, t)$, and is usually refered to as the tilt modulation

$$\mathcal{T}(\mathbf{r}, t) = \max\left(0, \frac{\langle \mathbf{u}, \mathbf{n} \rangle}{\|\mathbf{u}\| \cdot \|\mathbf{n}\|}\right) \in [0, 1], \tag{5}$$

where $\mathbf{r}$ is the range vector along the radius at mean free surface elevation $z = 0$. The shadowing modulation describes the loss of information due to occlusion. Shadowing occurs, as a wave facet at distance r' blocks the direct line of sight between the antenna and local parts of the sea surfaces at r with $r' < r$. The binary shadowing mask [26]

$$m(\mathbf{r}, t) = \begin{cases} 0 & \text{if } (r' < r) \wedge (\Theta_0(r', t) \geq \Theta_0(r, t)), \\ 1 & \text{otherwise,} \end{cases} \tag{6}$$

can be constructed from the local incidence angle

$$\Theta_0(\mathbf{r}, t) = \arccos\left(\frac{\langle \mathbf{e}_z, \mathbf{u} \rangle}{\|\mathbf{u}\|}\right), \tag{7}$$

with normal vector in z-direction $\mathbf{e}_z$. The radar blind zone was realized by $m(\mathbf{r}, t) := 0 \,\forall r \in \mathbf{r} : r \leq 200\,\mathrm{m}$. The amount of information received by the antenna depends on the clutter area [12, 25, 26]. Moreover, the reception is subject to speckle noise $n(\mathbf{r}, t)$, which is common in optical systems [26]. Here, the speckle noise was modeled by sampling random values from a exponential distribution with a scale of 1. Assuming a rectangular pulse, the clutter area is given by $A_\mathrm{c} = r \Delta\phi \Delta r$ [25, 26], and the RCS can be estimated by

$$\sigma(\mathbf{r}, t) \approx \mathcal{T}^4(\mathbf{r}, t) m(\mathbf{r}, t) n(\mathbf{r}, t) A_\mathrm{c}. \tag{8}$$

Radiowave propagation is subject to losses, for which the International Telecommunication Union (ITU) has provided in-depth modeling recommendations. The radar beam travels a certain distance through the earth's atmosphere [24, 25]. The loss over traveled distance $R = \|\mathbf{u}\|$ was implemented through the two-way path loss [27, 28]

$$\mathcal{L}_\mathrm{fsp,dB} = 20 \log_{10}\left(\frac{4\pi R}{\lambda}\right) \mathrm{dB}. \tag{9}$$

The atmospheric attenuation depends on the ambient temperature and pressure at sea level [29], for which the ITU has defined the reference values $T_\mathrm{amb} = 288.15\,\mathrm{K}$ and $P_\mathrm{amb} = 1013.25\,\mathrm{hPa}$ [28]. With an operating frequency $f = 10\,\mathrm{GHz}$, the atmospheric attenuation was modeled as $\mathcal{L}_\mathrm{aa,dB} = 0.0142\,\mathrm{dB\,km^{-1}}$ [29].

The transmitted power P_t and antenna gain G were unknown system constants. Instead of assuming values for P_t and G, the radar equation (4) was expressed in terms of dB,

$$P_\mathrm{r,dB} \approx 2\lambda_\mathrm{dB} + \sigma_\mathrm{dB} - 4R_\mathrm{dB} - \sum \mathcal{L}_{i,\mathrm{dB}} + C. \tag{10}$$

The conversion to logarithmic scale ($x_\mathrm{dB} = 10 \log_{10}(x)$ with arbitrary quantity in linear scale x) allowed to summarize all unknown system constants to an offset term C [26].

The radar model (10) with an installation height of $h = 15\,\mathrm{m}$ and a maximum range of $r_\mathrm{max} = 2000\,\mathrm{m}$ was used to generate synthetic X-band radar data from the HOS simulations data. The angular frequency of the radar was assumed to be $\Omega = 2\pi/1.5\,\mathrm{s^{-1}}$, which translated to a full revolution every $dt_\mathrm{rev} = 1.5\,\mathrm{s}$. With the HOS saving interval of $dt = 0.1\,\mathrm{s}$, a full radar revolution was build from circular sectors of $dt_\mathrm{rev}/dt = 15$ consecutive radar images. Consequently, each radar image exhibited multiple temporal gaps. The most pronounced gap of dt_rev was located along $y = 0$ for positive x. The final radar model resolved the spatial domain of $(4000 \times 4000)\,\mathrm{m}$ at 512×512 grid points.

3. MACHINE LEARNING FRAMEWORK

In this Section, the ML framework that was used in this study is presented. Section 3.1 introduces the proposed model architecture for the data-driven phase-resolved sea surface reconstruction. Afterwards, Section 3.2 briefly describes the data preprocessing, as well as the training process.

3.1 Machine learning model

From a data-driven perspective, the prediction of the phase-resolved sea surface from X-band radar data can be interpreted as an image-to-image prediction task. Generally speaking, an *image* is a special case of numerical data, that is structured in large multi-dimensional arrays. Convolutional neural networks (CNN) form a class of neural network architectures, that are designed to efficiently process such data [16, 17]. The weights are organized in multidimensional kernels. The size of these kernels is a design parameter, and is typically chosen way smaller than the size of the input data [17]. While most neural network archetypes dedicate one unique weight to each input, convolutional models move the kernel relative to the input. Hence, the same set of weights is applied to different parts of the input. This concept of *parameter sharing* allows CNNs to learn recurring local patterns at different locations in the data, and significantly reduces the number of trainable weights [17].

In this work, we propose a convolutional network architecture to attempt the sea surface reconstruction. The chosen model architecture followed the well-known U-Net topology [12], which is a convolutional auto-encoder with residual connections [30] between the encoding- and the decoding path, and augmented it with a Fourier Neural Operator (FNO) [13] in the bottleneck. FNOs excel at learning global features, and thus complement the convolutional nature of the U-Net. A FNO layer consists of two parallel paths. The first path transforms the input to the

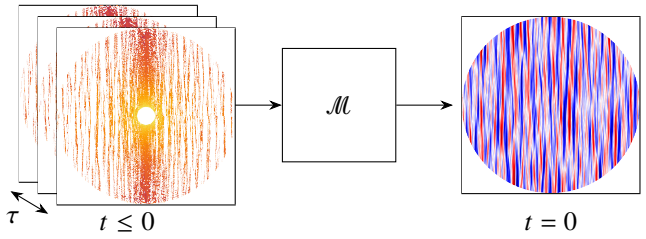

FIGURE 1: SCHEMATIC OF ML FRAMEWORK. THE MODEL $\mathscr{M}$ MAPS τ SUCCESSIVE HISTORIC RADAR IMAGES ($t \leq 0$) TO THE SEA SURFACE AT PRESENT TIME ($t = 0$).

complex Fourier space via a discrete Fourier transform (DFT). The resulting Fourier modes are truncated to the lowest m modes, which are then multiplied with the trainable weights of the FNO. Next, the modes are padded with zeros to match their initial shape, and transformed back to the original data space using an inverse DFT. The second path applies a linear transform, which accounts for non-periodic boundary conditions in the input. Moreover, it implicitly passes through higher order modes, which are truncated in the upper path. The results of both paths are added, and the result is passed through a GELU activation [31].

In the encoding path of the U-Net, each convolutional block bisected the spatial resolution of the data. This way, the model learned local features on different scales of the data. By reducing stepwise the spatial resolution throughout the forward path, only the most relevant features were maintained, i.e., the information density in the latent space was maximized. The decoder inflated the compressed information from the forward path, and combined it with the high resolution feature maps from the residual connections. Up- and downsampling of the data was realized using nearest neighbor interpolation followed by a convolution [32]. All convolutional layers except the linear output layer applied leaky ReLU activation. The final model had around 1.2 million trainable parameters; the exact number depends on the number of input channels τ, and therefore varied slightly throughout this study.

3.2 Training process

Overall, 858 HOS simulations with 1000 time steps each formed the data foundation. The radar was limited in range by $r_{\max} = 2000\,\mathrm{m}$. Since the reconstruction of the sea surface outside the radar range was not targeted, the HOS simulations were cut off at $r = r_{\max}$, i.e., $\eta(x, y, t) := 0$ for $\sqrt{x^2 + y^2} \geq r_{\max}$. To account for the mismatch in spatial resolution between radar images and sea surface data, the latter was undersampled to a 512×512 grid. The data were split into a training-, a validation-, and a test data set using a stratified split w.r.t. to the governing sea state parameters peak wavelength L_P, steepness ε, and spreading n. Splitting the data by sea state parameter prevented information leakage between the data sets, which is a common pitfall in ML [33]. The training set held 80% of all data (686 simulations), the validation and test set held 10% each (86 simulations). The data sets were processed with a sliding window, where the window size determined the number of consecutive radar snapshots τ. This way, $N = \left\lfloor t_\mathrm{sim} t_\mathrm{rev}^{-1} \right\rfloor - \tau + 1$ labeled samples ($\mathsf{x} \in \mathbb{R}^{512\times512\times\tau}$, $\mathsf{y} \in \mathbb{R}^{512\times512}$) were generated from a single simulation. Hence, with a context size of, e.g., $\tau = 16$, there were 34 986 samples in the training-, and 4 386 samples in the validation- and test set, respectively.

The training was conducted using the surface similiarity parameter (SSP) loss function [34, 35]

$$J_\mathrm{SSP}(\mathsf{y}, \hat{\mathsf{y}}) = \frac{\sqrt{\int |\mathsf{Y} - \hat{\mathsf{Y}}|^2 d\mathbf{k}}}{\sqrt{\int |\mathsf{Y}|^2 d\mathbf{k}} + \sqrt{\int |\hat{\mathsf{Y}}|^2 d\mathbf{k}}} \in [0, 1] \qquad (11)$$

where Y is the spatial DFT of y and $\mathbf{k} = [k_x, k_y]^\intercal$ the wavenumber vector. By comparing the ground truth y to the prediction $\hat{\mathsf{y}}$ in the complex Fourier space, the SSP explicitly penalizes deviations in magnitude and phase [35]. The SSP is normalized to the range $[0, 1]$; $\mathrm{SSP}(\mathsf{y}, \hat{\mathsf{y}}) = 0$ indicates perfect agreement among y and $\hat{\mathsf{y}}$, $\mathrm{SSP}(\mathsf{y}, \hat{\mathsf{y}}) = 1$ is only achieved for $\hat{\mathsf{y}} = -\mathsf{y}$ or $(\hat{\mathsf{y}} = \mathbf{0} \wedge \mathsf{y} \neq \mathbf{0})$. In this work, the SSP loss (11) was augmented with a lowpass filter, which operated only on the ground truth y, and thus pushed the model towards focussing on a certain frequency range of the wave spectrum [36]. The filter was designed to attentuate all frequecy components $k \geq 6.0 k_\mathrm{p}$, which, from an empirical standpoint, preserved all relevant information in y. In consequence, $\mathrm{SSP} = 0$ reflected perfect agreement among y and $\hat{\mathsf{y}}$ for $k < 6.0 k_\mathrm{p}$ only. Furthermore, the Adam optimizer [37] with learning rate $\alpha = 0.001$ was used to adapt the weights during the training. An early stopping routine monitored the optimization process, and terminated the training, once there were no improvements in the validation loss over 10 consecutive epochs. One training epoch with a batch size of 16 took about 10 minutes on a workstation with an NVIDIA® A100 Tensor Core GPU with 40GB VRAM and an Intel® Xeon® Platinum 8358 12 core CPU.

4. RESULTS

In the following, the results of the data-driven sea surface reconstruction are presented. First, the impact on the number of consecutive radar snapshots τ on the reconstruction accuracy of the model is discussed. Then, the best performing model set up is evaluated in depth. After a global perspective on the results, the best and worst model predictions on unseen data are inspected. All results are reported in terms of SSP. The SSP relies on the discrete Fourier transform, cf. (11), and inherently requires a square domain for evaluation. To exclude the impact of potential errors in the reconstruction for $R > 2000\,\mathrm{m}$, a mask was applied to both the true and reconstructed wave prior to the error assessment via SSP. This mask was implemented as a circular Tukey window [38] with $\alpha = 0.25$, which imposed a smooth transition to zero at $R = 2000\,\mathrm{m}$. Later on, the individual reconstructions are additionally evaluated using the point-wise absolute error.

4.1 Study on the context size

The *context* refers to the number of consecutive historic radar images ($t \leq 0$) that were mapped to the sea surface elevation at $t = 0$. Using a larger context was expected to improve the reconstruction result, as the model was given more information. In this study, 9 context sizes of $\tau \in \{4, 6, 8, \ldots, 20\}$ historic radar images were compared. One model training was conducted for

each context size. By setting a global random seed, all models were initialized with the same weights. The first layer formed an exception here, since the number of weights in a convolutional layer scales with the depth of the input, i.e., τ. As a result, the initial weights in the first layer were not identical across the models. To further maintain consistency, the same stratified data split was used across all model trainings.

Figure 2 shows a statistic comparison on the impact of the context size on the reconstruction accuracy on unseen data. The results are grouped by spreading $n \in \{4, 30, 300\}$, and reported twofold for each model: The left-hand side (dark blue) indicates the results when evalutating over the whole domain (i.e., within the radar radius $R \leq 2000\,\mathrm{m}$), the right-hand side (light blue) indicates the results for evaluation on the center (2000×2000) m. For convenience, the evaluation on the whole domain is refered to as case (i), and on the center cutout is referred to as case (ii). For each case and model, the reconstruction accuracy distribution is visualized using a box- and a violin-plot. While the box plot provides basic statistic information on the data like the median, the size of the interquartile range, and the existance of outliers (marked by triangles), the violin plot adds valuable information about the density curves of the underlying distributions.

The density curves exhibit a particular shape for each spreading n, and are shifted towards the higher (lower) SSP regime for smaller (larger) context size. The results for $\tau = 12$ break that trend by being slightly worse compared to adjacend values of τ. This may be attributed to unfavorable weight initialization for that particular model, and may be diminished by performing a K-fold cross validation, i.e., train the same models with different training distributions[1] [17]. For $n = 30$, the density curves for case (i) and (ii) are of similar shape, for $n \in \{4, 300\}$, the overall shape strongly depends on the evaluated spatial domain. Across all spreadings n, case (ii) generally exhibits higher median SSP and more outliers towards high SSP values. However, there are a few exceptions, where the median SSP for case (ii) is on par or lower than the median SSP for case (i), cf., e.g., $n \in \{30, 300\} \wedge \tau = 4$. Despite the generally worse performance observed for case (ii), the best reconstruction for $n = 4$ is always provided for the center cutout of the domain. In constrast, for $n \in \{30, 300\}$, there is an offset in favor of case (i) between the density curves, which gets more pronounced with increasing context size. While all models perform better on long-crested sea ($n = 300$) than short-crested sea ($n = 4$), the median SSP converges with increasing context size. In particular, there exist only minor differences between the results for a context size $\tau \in [14, 20]$, indicating that there is no benefit from using a context size greater than 14 radar snapshots. With regards to inference speed, there is no significant difference across the tested configurations. For instance, the inference of a single input on the GPU took 72 ± 7 ms for $\tau = 16$, and 73 ± 4 ms for $\tau = 4$ (CPU: 184 ± 13 ms, and 185 ± 6 ms, respectively; mean and standard deviation over 1000 iterations, batch size of 1).

4.2 Evaluation of the ML model

In Section 4.1, the sea surface reconstruction model with a context size of $\tau = 14$ historic radar snapshots is identified to provide the best balance between reconstruction accuracy and model complexity. This model, formally referred to as $\mathscr{M}_{14}$, is now evaluated in detail. First, a global picture of the reconstruction accuracies is presented. Then, the best and the worst reconstructions on unseen data are examined to highlight the models strengths and shortcomings.

Evaluation over parameter space. Figure 3 provides an in-depth look into the performance of $\mathscr{M}_{14}$ by resolving the reconstruction accuarcy in peak wavelength L_p and wave steepness ε. The upper row shows the results for case (i), the lower row shows the results for case (ii). The spreading increases from left to right. Cells marked with "•" indicate samples from the validation set, cells marked with "+" indicate samples from the test set. Since the model is fit to the training set, both the validation- and test set are considered unseen data and thus valid test cases for the model performance. In the center of each panel, the mean and standard deviation of the SSP is stated.

The results show a general performance degradation with increasing directionality. For case (i), the mean reconstruction accuracy drops from $\overline{\mathrm{SSP}} = 0.106$ for long-crested sea surfaces to $\overline{\mathrm{SSP}} = 0.144$ for short-crested sea surfaces. The results for $n = 30$ sit right in between with $\overline{\mathrm{SSP}} = 0.115$. The standard deviation ranges within $\sigma^2 = 0.0235 \pm 0.0015$ for all three spreadings n. Besides the difference in reconstruction accuracy, the resolved errors indicate a similar model behavior for all spreadings. The performance degrades for decreasing steepness and wavelength, culminating in the worst reconstruction accuracy for $L_\mathrm{p} = 100\,\mathrm{m}$ and $\varepsilon = 0.01$. The results on unseen data (marked with "+" and "•") integrate well into the overall picture, indicating that the model generalizes well. While the results for case (ii) exhibit the same characteristics as described before, there is a slight increase in mean accuracy and standard deviation.

The overall degradation in accuracy with decreasing wave spreading n can be attributed to the more complex wave dynamics with increased directionality. For $n = 4$, the directional spectrum is way broader, leading to manifold wave directions. This results in a more complex sea surface, which is characterized by many small-scale patterns. In contrast, $n = 300$ yields almost vertical wave crests with a more uniform wave propagation in positive x-direction. The dependence on wave steepness is most certainly linked to the correlation between wave height and slope for constant L_p. The wave steepness first and foremost steers the amplitude of a wave. This translates to steeper slopes, hence smaller angles of incidence between radar beam and wave surface normals, and thus stronger radar return signals for higher values of ε. The amount of shadowed area, i.e., missing values in the data, also correlates with the steepness. For a single radar snapshot, the amount of shadowed area within $R = 2000\,\mathrm{m}$ peaks at almost 85% for $n = 4$, $L_\mathrm{p} = 200\,\mathrm{m}$ and $\varepsilon = 0.135$. Nonetheless, the worst performace is observed for the exact opposite of the parameter space, namely $L_\mathrm{p} = 100\,\mathrm{m}$ and $\varepsilon = 0.01$, where less than 10% of the radar area is shadowed. This trend is most likely linked to less steep wave facets and thus weaker radar return signals for lower steepness values. In fact, the results for $n \in \{30, 300\}$ indicate a sweet spot for $\varepsilon = [0.04, 0.10]$ and $L_p \geq 150\,\mathrm{m}$, as this region is attributed the lowest reconstruction errors. Here, the return signal strength is sufficiently strong,

[1] Due to time constraints, K-fold cross validation is not performed in this study.

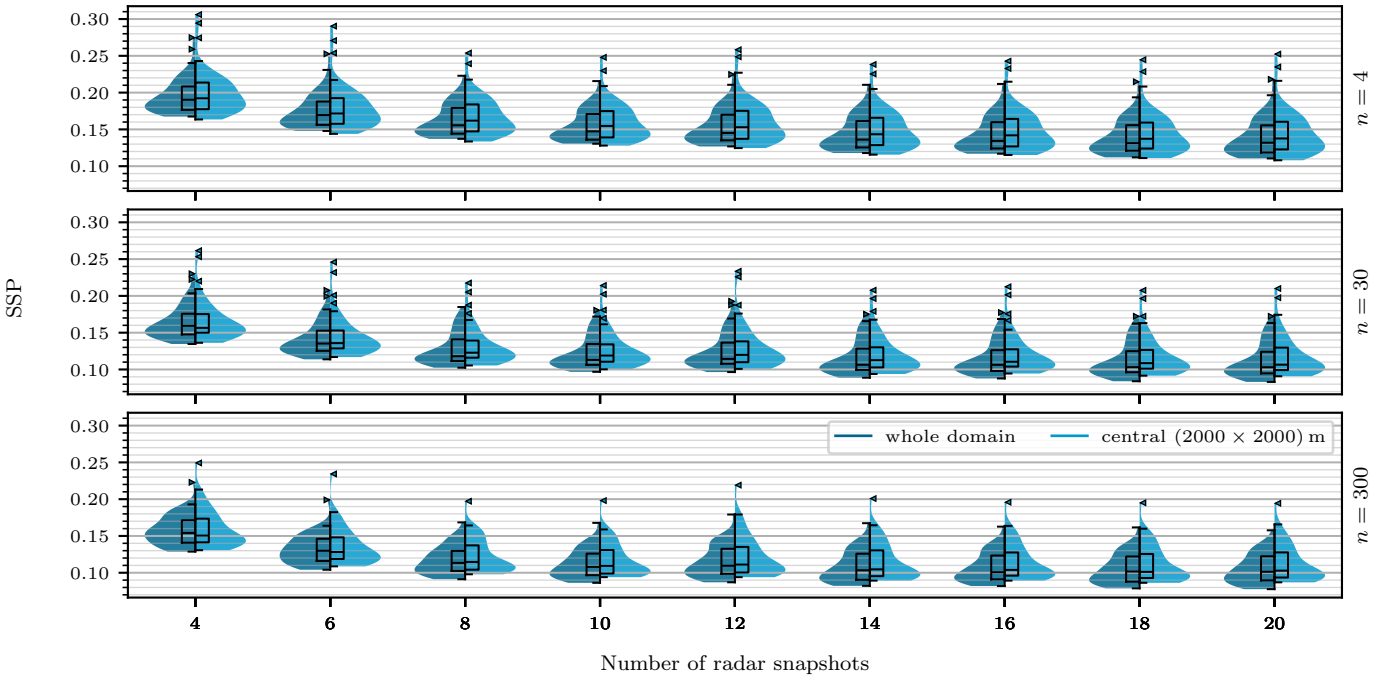

FIGURE 2: IMPACT OF THE CONTEXT SIZE ON THE RECONSTRUCTION ACCURACY EXPRESSED IN TERMS OF SSP. THE TRAINED MODELS ARE EVALUATED ON ALL UNSEEN DATA, I.E. THE VALIDATION- AND THE TEST SET. THE RESULTS ARE GROUPED BY WAVE SPREADING, RANGING FROM SHORT-CRESTED ($n = 4$) TO LONG-CRESTED ($n = 300$) SEA STATES. EACH RECONSTRUCTION IS EVALUATED OVER THE WHOLE SPATIAL DOMAIN, AND THE CENTRAL (2000×2000) m. THE DISTRIBUTIONS OF THE RECONSTRUCTION ERRORS ARE VISUALIZED WITH BOX- AND VIOLIN PLOTS. OUTLIERS ARE MARKED WITH A TRIANGLE.

while the wave height is small enough to minimize the effect of shadowing and nonlinear wave effects. The overall decreased reconstruction accuracy observed for case (ii) can be linked to the radar blind zone. The relative share of the radar blind zone on the area of the evaluated domain increases from $\frac{\pi}{4}\%$ for case (i) to $\pi\%$ for case (ii). Since there is no information available for the radar blind zone, the model has to generalize over it, which is expected to yield generally higher reconstruction errors.

Evaluation of best and worst reconstruction. In this section, the best- and the worst reconstruction over all unseen data are evaluated in detail. The best reconstruction is achieved for the reconstruction of a sea state with $L_\mathrm{p} = 200\,\mathrm{m}$, $\varepsilon = 0.055$ and $n = 300$. Figure 4 (a) shows the ground truth data from HOS simulation, (b) shows the reconstruction of $\mathscr{M}_{14}$. Here, the area of evaluation for case (ii), i.e., the central (2000×2000) m cutout, is indicated by a dashed rectangle. The SSP value within the rectangle states the reconstruction error for case (ii), the SSP value outside the rectangle states the reconstruction error for case (i), i.e., over the whole domain. Figure 4 (c) shows the normalized absolute error

$$\mathrm{NAE} = H_\mathrm{S}^{-1}|\mathrm{y} - \hat{\mathrm{y}}|, \tag{12}$$

which spatially resolves the reconstruction error over x and y. Note that the NAE is a point-wise comparison between the ground truth y and the prediction $\hat{\mathrm{y}}$, i.e. the sea surface reconstruction. The NAE, in constrast to the filtered SSP (cf. Section 3.1), inherently implicitly compares the signals over the whole frequency range. Therefore, the NAE is expected to exhibit some background noise, even for SSP = 0.

A qualitative comparison between the ground truth and the prediction shows, that the ML model $\mathscr{M}_{14}$ provides faithful sea surface reconstructions, cf. Figure 4 (a) and (b). At closer inspection, the prediction appears slightly smoother than the ground truth. This loss of detail is attributed to the filtered SSP forcing the model to attenuate the high frequency range. This is a design choice, and can be adjusted by a looser low-pass filter during training. The reconstruction error for case (i) quantifies to $\mathrm{SSP}_{(\mathrm{i})} = 0.082$, and $\mathrm{SSP}_{(\mathrm{ii})} = 0.095$ for case (ii). The spatially resolved error evaluation via NAE exhibits vertical patterns, which are most pronounced around $x = 0$. Hence, the largest deviations between y and $\hat{\mathrm{y}}$ occur, when the radar beam is, within some margin, parallel to the wave crests. In those cases, there are no wave fronts or -throughs, which can reflect the radar beam, and thus no backscattered information. Nonetheless, the ML approach manages to provide a faithful sea surface reconstruction in those areas.

The worst reconstruction is provided for a sea state with $L_\mathrm{p} = 130\,\mathrm{m}$, $\varepsilon = 0.01$ and $n = 4$, cf. Figure 5. Due to the wide wave spreading and short wavelength, the sea surface is characterized by manifold small-scale patterns. The reconstruction error for case (i) is $\mathrm{SSP}_{(\mathrm{i})} = 0.211$, the reconstruction error for case (ii) is $\mathrm{SSP}_{(\mathrm{ii})} = 0.238$. At closer inspection, the reconstruction exhibits a circle-shaped artifacts in the center, cf. Figure 5 (b). The spatially resolved error confirms, that the main cause of error is the less accurate sea surface reconstruction within the radar blind zone, i.e., all points $\{(x, y)|\sqrt{x^2 + y^2} \leq 200\,\mathrm{m}\}$, cf. Fig-

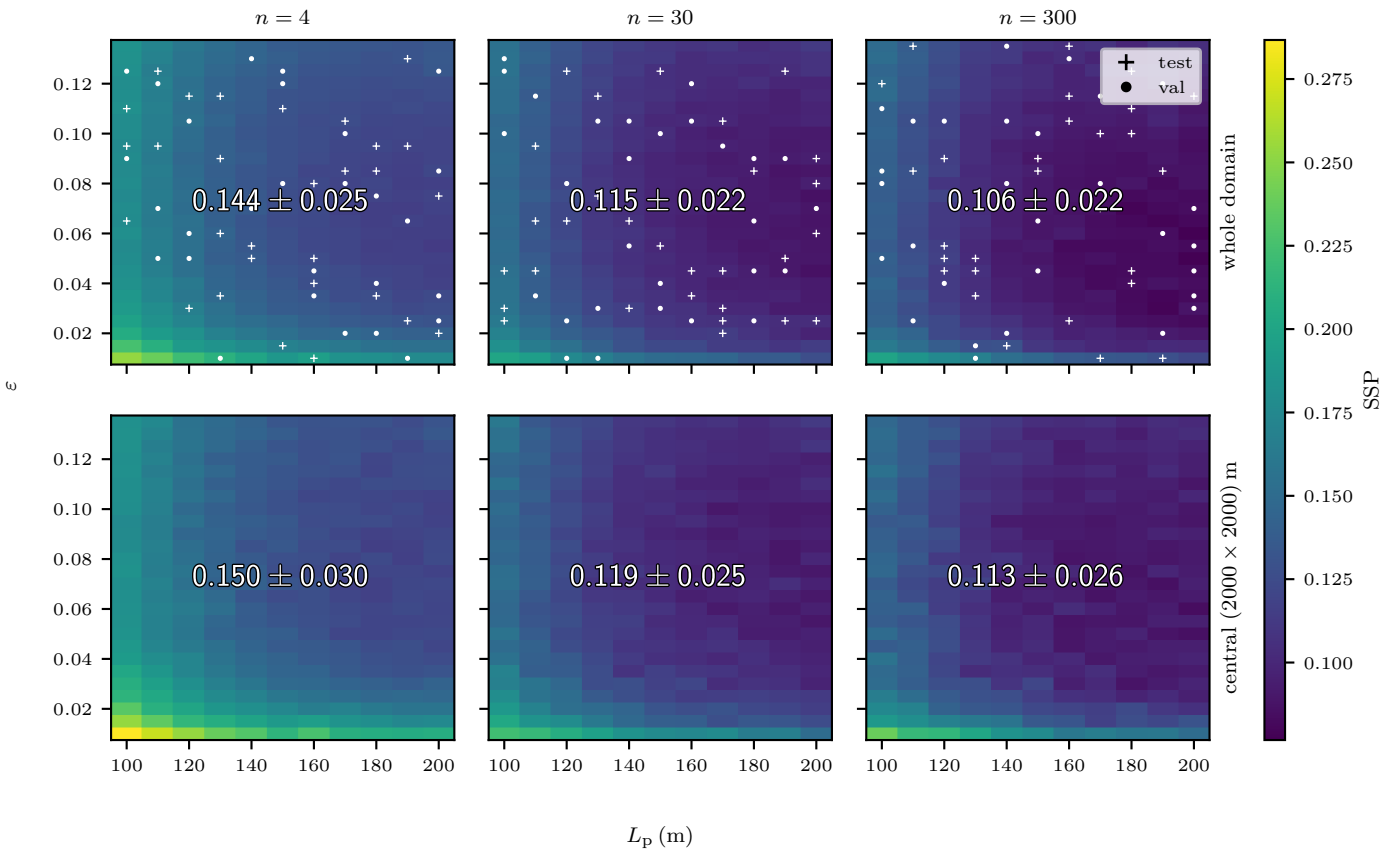

FIGURE 3: GLOBAL EVALUATION OF THE RECONSTRUCTION ACCURACY. THE UPPER THREE PANELS SHOW THE EVALUATION OVER THE WHOLE SPATIAL DOMAIN, THE LOWER THREE PANELS OVER THE CENTRAL (2000 × 2000) m, RESPECTIVELY. EACH PANEL STATES THE MEAN AND STANDARD DEVIATION OF THE SSP. MARKED CELLS INDICATE SAMPLES, THAT WERE NOT USED FOR TRAINING ("+": TEST SET, "•": VALIDATION SET). THE RESULTS DEMONSTRATE GOOD GENERALIZATION PROPERTIES SINCE THERE IS NO REMARKABLE DROP-OFF IN ACCURACY BETWEEN THE DATA THE MODEL WAS FIT TO AND THE VALIDATION- OR TEST SET, RESPECTIVELY.

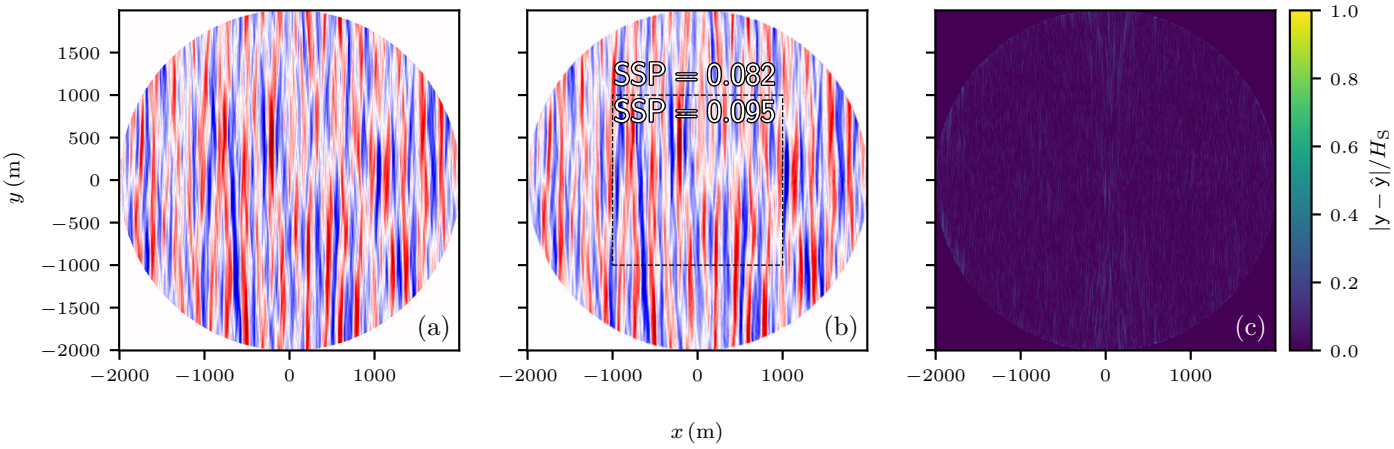

FIGURE 4: SEA SURFACE RECONSTRUCTION RESULTS FOR SEA STATE WITH CHARACTERISTIC PARAMETERS L_p = 200 m, ε = 0.055, AND n = 300. (a) SHOWS THE HOS SIMULATION WITH M = 4, WHICH SERVED AS GROUND TRUTH. (b) SHOWS THE RECONSTRUCTION FROM ML MODEL $\mathcal{M}_{14}$. (c) SHOWS THE ABSOLUTE ERROR BETWEEN (a) AND (b) NORMALIZED BY THE SIGNIFICANT WAVE HEIGHT H_S. THIS CASE REPRESENTS THE BEST ML PREDICTION ON UNSEEN DATA.

ure 5 (c). Moreover, the NAE reveals isolated radial patterns, which are least prominent within the upstream reconstruction, assuming a mean wave heading in positive x-direction. Similar to Figure 4 (c), these patterns most certainly originate from the alignment of the wave crests and the radar beam, leading to missing backscattered information. However, the higher frequency and radial alignment of the patterns in Figure 5 (c) indicate another factor that shapes the error. With decreasing wavelength L, all frequency components that contribute to the sea state undergo a larger share of their respective period $T = \sqrt{2\pi L g^{-1}}$ during one radar revolution dt_{rev}. Paired with the overall more complex spatio-temporal pattern for $n = 4$, this effect results in a larger spatial mismatch between the individual circular sectors that contribute to a single radar image. As mentioned at the end of Section 2, the largest temporal gap in each synthetic radar image exists along $y = 0$ for positive x, i.e., where the NAE exhibits a pronounced radial line. This suggests, that the radial lines observed in Figure 5 (c) are partly a product of the time discretization, and should decrease with finer dt. In a real-world scenario, the radar image can be thought of as a spatio-temporal spiral that is continuous in time, i.e., $dt \to 0$. This should in theory reduce all radial error pattern except for the necessary temporal gap of dt_{rev}, which was here located along $y = 0$ for positive x.

5. CONCLUSIONS AND OUTLOOK

The presented study proposed a data-driven approach to the 2D sea surface reconstruction from synthetic X-band radar data. The synthetic data were generated using the HOS method and a geometric radar tool with additional system losses. The results indicated sufficient reconstruction results across a wide parameter space. The reconstruction error scaled with the overall shape of the sea surface, which is defined by the spreading n. Long-crested sea surfaces were reconstructed with higher fidelity than short-crested sea surfaces. Moreover, a correlation between the wave height, steered by the steepness ε, and the reconstruction accuracy was identified. The area of the parameter space, where the steepness was large enough to guarantee a low angle of incidence between the surface normals and the radar beam, and low enough to minimize the effect of shadowing, marked a sweet spot, wherein the model reconstructions exhibited the best accuracy.

The presented study used a wide spectrum of characteristic sea state parameters, and proposed a unifying data-driven solution. In particular the spreading n significantly affected the diversity of the data as it shapes the dominant wave patterns. The ML model held a limited number of trainable parameters, and thus could learn a finite number of patterns. A future study may investigate, whether training on a reduced parameter space can further improve the reconstruction results. In particular, focussing the reduced parameter space around lower steepness values ε will most certainly improve the predictions within the calmer region of the parameter space, where marine operations are usually performed.

REFERENCES

[1] Wijaya, A. P., Naaijen, P., Andonowati and Groesen, E. van. "Reconstruction and future prediction of the sea surface from radar observations." *Ocean Engineering* Vol. 106 (2015): pp. 261–270. DOI 10.1016/j.oceaneng.2015.07.009.

[2] Desmars, N. "Real-time reconstruction and prediction of ocean wave fields from remote optical measurements." phdthesis, École centrale de Nantes. 2020.

[3] Zinchenko, V., Vasilyev, L., Halstensen, S. O. and Liu, Y. "Short-time deterministic prediction of individual waves based on space-time X-band Marine radar measurements." *Proceedings of the Institution of Mechanical Engineers Part M: Journal of Engineering for the Maritime Environment* Vol. 235 No. 2 (2021): pp. 356–371. DOI 10.1177/1475090220987401.

[4] Lee, J. H., Nam, Y. S., Liu, Y. and Yang, H. "Study on predictability of ocean wave fields based on marine radar measurement data." *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment* Vol. 238 No. 3 (2023): pp. 656–676. DOI 10.1177/14750902231184096.

[5] He, J., Chen, Z., Zhao, C. and Chen, X. "On the Reconstruction and Prediction Improvements of the Deterministic Sea Wave Predictable Zone Using Spatio-Temporal Coherent Radar Measurements." *IEEE Transactions on Geoscience and Remote Sensing* Vol. 61 (2023): pp. 1–14. DOI 10.1109/tgrs.2023.3314533. URL http://dx.doi.org/10.1109/TGRS.2023.3314533.

[6] Ehlers, S., Klein, M., Heinlein, A., Wedler, M., Desmars, N., Hoffmann, N. and Stender, M. "Machine learning for phase-resolved reconstruction of nonlinear ocean wave surface elevations from sparse remote sensing data." *Ocean Engineering* Vol. 288 (2023): p. 18. DOI 10.1016/j.oceaneng.2023.116059.

[7] Zinchenko, V., Vasilyev, L., Halstensen, S. O. and Liu, Y. "An improved algorithm for phase-resolved sea surface reconstruction from X-band marine radar images." *Journal of Ocean Engineering and Marine Energy* Vol. 7 No. 1 (2021): pp. 97–114. DOI 10.1007/s40722-021-00189-9.

[8] Köllisch, N., Behrendt, J., Klein, M. and Hoffmann, N. "Nonlinear real time prediction of ocean surface waves." *Ocean Engineering* Vol. 157 (2018): pp. 387–400. DOI 10.1016/j.oceaneng.2018.03.048.

[9] Zhao, M., Zheng, Y. and Lin, Z. "Sea surface reconstruction from marine radar images using deep convolutional neural networks." *Journal of Ocean Engineering and Science* Vol. 8 No. 6 (2023): pp. 647–661. DOI 10.1016/j.joes.2023.09.002.

[10] Cybenko, G. "Approximation by superpositions of a sigmoidal function." *Mathematics of Control, Signals and Systems* Vol. 2 No. 4 (1989): pp. 303–314. DOI 10.1007/BF02551274.

[11] Hornik, K., Stinchcombe, M. and White, H. "Multilayer feedforward networks are universal approximators." *Neural Networks* Vol. 2 No. 5 (1989): pp. 359–366. DOI 10.1016/0893-6080(89)90020-8. Accessed 2024-02-16, URL https://www.sciencedirect.com/science/article/pii/0893608089900208.

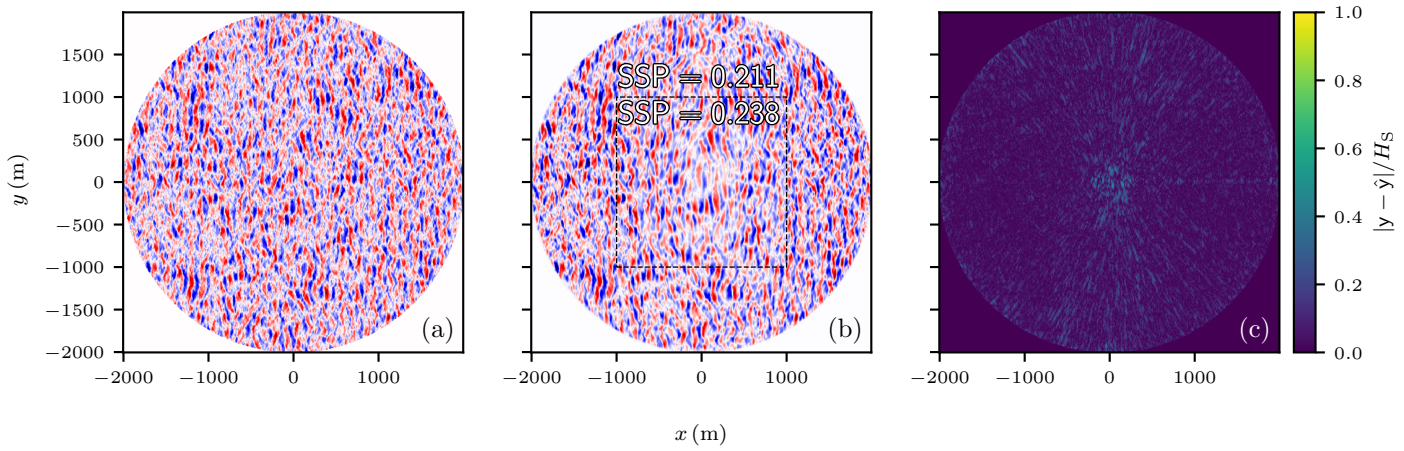

FIGURE 5: SEA SURFACE RECONSTRUCTION RESULTS FOR SEA STATE WITH CHARACTERISTIC PARAMETERS $L_p = 130\,\mathrm{m}$, $\varepsilon = 0.01$, AND $n = 4$. (a) SHOWS THE HOS SIMULATION WITH $M = 4$, WHICH SERVED AS GROUND TRUTH. (b) SHOWS THE RECONSTRUCTION FROM ML MODEL $\mathcal{M}_{14}$. (c) SHOWS THE ABSOLUTE ERROR BETWEEN (a) AND (b) NORMALIZED BY THE SIGNIFICANT WAVE HEIGHT H_S. THIS CASE REPRESENTS THE WORST ML PREDICTION ON UNSEEN DATA.

[12] Ronneberger, O., Fischer, P. and Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Springer International Publishing (2015): pp. 234–241. DOI 10.1007/978-3-319-24574-4_28.

[13] Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A. and Anandkumar, A. "Fourier Neural Operator for Parametric Partial Differential Equations." (2020). DOI 10.48550/ARXIV.2010.08895.

[14] Dommermuth, D. G. and Yue, D. K. P. "A high-order spectral method for the study of nonlinear gravity waves." *Journal of Fluid Mechanics* Vol. 184 (1987): pp. 267–288. DOI 10.1017/S002211208700288X.

[15] West, B. J., Brueckner, K. A., Janda, R. S., Milder, D. M. and Milton, R. L. "A new numerical method for surface hydrodynamics." *Journal of Geophysical Research* Vol. 92 No. C11 (1987): p. 11803. DOI 10.1029/JC092iC11p11803.

[16] LeCun, Y., Bengio, Y. and Hinton, G. "Deep learning." *Nature* Vol. 521 No. 7553 (2015): pp. 436–444. DOI 10.1038/nature14539.

[17] Goodfellow, I., Bengio, Y. and Courville, A. *Deep Learning*. MIT Press (2016). http://www.deeplearningbook.org.

[18] Lünser, H., Hartmann, M., Desmars, N., Behrendt, J., Hoffmann, N. and Klein, M. "The Influence of Characteristic Sea State Parameters on the Accuracy of Irregular Wave Field Simulations of Different Complexity." *Fluids* Vol. 7 No. 7 (2022): p. 243. DOI 10.3390/fluids7070243.

[19] Sobey, R. J. *Real Sea States: Advanced Short Course Notes*. Leichtweiß-Institut für Wasserbau, Technische Universität Braunschweig (1999).

[20] Hasselmann, K., Barnett, T., Bouws, E., Carlson, H., Cartwright, D., Enke, K., Ewing, J., Gienapp, H., Hasselmann, D., Kruseman, P., Meerburg, A., Muller, P., Olbers, D., Richter, K., Sell, W. and Walden, H. "Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP)." *Deut. Hydrogr. Z.* Vol. 8 (1973): pp. 1–95.

[21] Bouws, E., Günther, H., Rosenthal, W. and Vincent, C. L. "Similarity of the wind wave spectrum in finite depth water: 1. Spectral form." *Journal of Geophysical Research: Oceans* Vol. 90 No. C1 (1985): pp. 975–986. DOI 10.1029/JC090iC01p00975.

[22] Mitsuyasu, Hisashi, Tasai, Fukuzo, Suhara, Toshiko, Mizuno, Shinjiro, Ohkusu, Makoto, Honda, Tadao and Rikiishi, Kunio. "Observations of the Directional Spectrum of Ocean Waves Using a Cloverleaf Buoy." *Journal of Physical Oceanography* Vol. 5 No. 4 (1975): pp. 750–760. DOI 10.1175/1520-0485(1975)005<0750:ootdso>2.0.co;2.

[23] Dommermuth, D. "The initialization of nonlinear waves using an adjustment scheme." *Wave Motion* Vol. 32 (2000): pp. 307–317. DOI 10.1016/S0165-2125(00)00047-0.

[24] Richards, Mark A. *Fundamentals of radar signal processing*. McGraw-Hill electronic engineering, McGraw-Hill, New York, NY u.a. (2005). Includes index.

[25] Seybold, J. S. *Introduction to RF Propagation*. Wiley (2005). DOI 10.1002/0471743690.

[26] Salcedo-Sanz, S., Nieto Borge, J.C., Carro-Calvo, L., Cuadra, L., Hessner, K. and Alexandre, E. "Significant wave height estimation using SVR algorithms and shadowing information from simulated and real measured X-band radar images of the sea surface." *Ocean Engineering* Vol. 101 (2015): pp. 244–253. DOI 10.1016/j.oceaneng.2015.04.041.

[27] Recommendation ITU-R P.525-3. "Calculation of free-space attenuation." Technical report no. International Telecommunications Union, Geneva, Switzerland. 2016. URL https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.525-3-201611-S!!PDF-E.pdf.

[28] Recommendation ITU-R P.835-6. "Reference standard athmospheres." Technical report no. International Telecommunications Union, Geneva, Switzerland. 2017. URL https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.835-6-201712-S!!PDF-E.pdf.

[29] Recommendation ITU-R P.676-11. "Attentuation by atmospheric gases." Technical report no. International Telecommunications Union, Geneva, Switzerland. 2016. URL https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.676-11-201609-S!!PDF-E.pdf.

[30] He, K., Zhang, X., Ren, S. and Sun, J. "Deep Residual Learning for Image Recognition." (2015). DOI 10.48550/ARXIV.1512.03385.

[31] Hendrycks, D. and Gimpel, K. "Gaussian Error Linear Units (GELUs)." (2016). DOI 10.48550/ARXIV.1606.08415.

[32] Odena, A., Dumoulin, V. and Olah, C. "Deconvolution and Checkerboard Artifacts." (2016). DOI 10.23915/distill.00003.

[33] Hewamalage, H., Ackermann, K. and Bergmeir, C. "Forecast evaluation for data scientists: common pitfalls and best practices." *Data Mining and Knowledge Discovery* Vol. 37 No. 2 (2022): pp. 788–832. DOI 10.1007/s10618-022-00894-5. URL http://dx.doi.org/10.1007/s10618-022-00894-5.

[34] Perlin, M. and Bustamante, M. D. "A robust quantitative comparison criterion of two signals based on the Sobolev norm of their difference." *Journal of Engineering Mathematics* Vol. 101 (2014): pp. 115–124. URL https://api.semanticscholar.org/CorpusID:119299161.

[35] Wedler, M., Stender, M., Klein, M., Ehlers, S. and Hoffmann, N. "Surface similarity parameter: A new machine learning loss metric for oscillatory spatio-temporal data." *Neural Networks* Vol. 156 (2022): pp. 123–134. DOI 10.1016/j.neunet.2022.09.023.

[36] Wedler, M., Stender, M., Klein, M. and Hoffmann, N. "Machine learning simulation of one-dimensional deterministic water wave propagation." *Ocean Engineering* Vol. 284 (2023): p. 115222. DOI 10.1016/j.oceaneng.2023.115222.

[37] Kingma, D. P. and Ba, J. "Adam: A Method for Stochastic Optimization." (2014). DOI 10.48550/ARXIV.1412.6980.

[38] Harris, F. J. "On the use of windows for harmonic analysis with the discrete Fourier transform." *Proceedings of the IEEE* Vol. 66 No. 1 (1978): pp. 51–83. DOI 10.1109/PROC.1978.10837.