



Article

Prediction of Significant Wave Heights with Engineered Features from GNSS Reflectometry

Jan M. Becker * and Ole Roggenbuck

Department of Geodesy, Federal Agency for Cartography and Geodesy, Richard-Strauss-Allee 11, 60598 Frankfurt am Main, Germany

* Correspondence: jan.becker@bkg.bund.de

Abstract: Estimating reflector heights at stationary GNSS sites with interferometric reflectometry (IR) is a well-established technique in ocean remote sensing. Additionally, IR offers the possibility to estimate the significant wave height (SWH) with a linear model using the damping coefficient from an inverse modelling applied to GNSS signal-to-noise ratio (SNR) observations. Such a linear model serves as a benchmark in the present study, where an alternative approach for the estimation of both SWH and reflector height is presented that is based on kernel regression and clustering techniques. In this alternative approach, the reflector height is estimated by analyzing local extrema occurring in the interference pattern that is present in GNSS SNR observations. Various predictors are derived from clustering statistics and the estimated reflector heights. These predictors are used for the SWH determination with supervised machine learning. Linear models, bagged regression trees, and artificial neural networks are applied and respective results are compared for various predictor sets. In a second step, damping coefficients obtained from the inverse modelling mentioned above are additionally taken into account as predictors. The usability of the alternative approach is demonstrated. Compared to the benchmark, a significant improvement in terms of accuracy is found for an artificial neural network with predictors from both the alternative and the inverse modelling approach.

Keywords: GNSS reflectometry; significant wave height; feature engineering; machine learning



Citation: Becker, J.M.; Roggenbuck, O. Prediction of Significant Wave Heights with Engineered Features from GNSS Reflectometry. *Remote Sens.* **2023**, *15*, 822. <https://doi.org/10.3390/rs15030822>

Academic Editor: Chung-yen Kuo

Received: 19 December 2022

Revised: 20 January 2023

Accepted: 28 January 2023

Published: 31 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sea surface characteristics have a direct or indirect effect on various areas of life and the environment, such as atmospheric dynamics [1] and weather forecasting, as well as coastal protection and shipping. Therefore, the ability to retrieve information on the local sea surface state is of great value. In this context, GNSS reflectometry (GNSS-R) has become a popular remote sensing technique that can be used for assessing changes in sea level [2,3], for estimating wave directions [4], and for characterizing the local roughness of the sea surface in terms of the significant wave height (SWH) [5,6]. The latter is defined in a statistical manner by observing heights of individual waves over a certain period of time and taking the mean value of the highest one-third of these wave height values as SWH [7].

The core principle of GNSS-R is to evaluate the superposition of the directly and indirectly received GNSS signal, where the indirect signal emerges from reflection at the surface to be probed (see Figure 1). The receiving antenna can be mounted as a static device [8,9] but there are also applications with moving antennas, such as ship-based [10] and spaceborne GNSS-R methods like [11], where in the latter case the SWH is determined by means of artificial neural networks using delay Doppler maps as input.

Another example for static and ship-based GNSS-R analyses is given in [6]. There, a linear relationship between the SWH and a damping coefficient is stated. The latter results from a least squares fit of a functional model describing the signal-to-noise ratio (SNR) interference pattern as a damped oscillation with additive trend. Such an estimation

method for the SWH, which stems from an inverse modelling, is considered as a benchmark in our analysis presented here.

Besides GNSS-R, there are other techniques for measurements of the SWH. Marine sensors like wave buoys or acoustic sensors can be used for in situ measurements. In addition to these local sensors, remote sensing techniques like HF radar systems [12,13] and satellite altimetry [14] can be used for SWH observation. There is a broad spectrum of usage of electromagnetic systems in remote sensing of the environment. This includes change detection for the sea surface [15], which is to a certain extent also addressed in the present work by introducing and processing a binary label related to the sea surface state.

In earlier work [16], we investigated alternative GNSS-R methods for the estimation of the reflector height, i.e., the distance h between the antenna phase center and the reflecting surface; see Figure 1. There, kernel regression and clustering techniques were used for analysing GNSS-R interference patterns in SNR observations made at the research platform FINO2 in the Baltic Sea. Furthermore, a binary reliability label based on statistics of the clustering was introduced, indicating whether the associated reflector height estimate can be considered trustworthy or not. A comparison of SWH observations with the estimated reflector heights and associated reliability labels at FINO2 suggested that these quantities are somehow related.

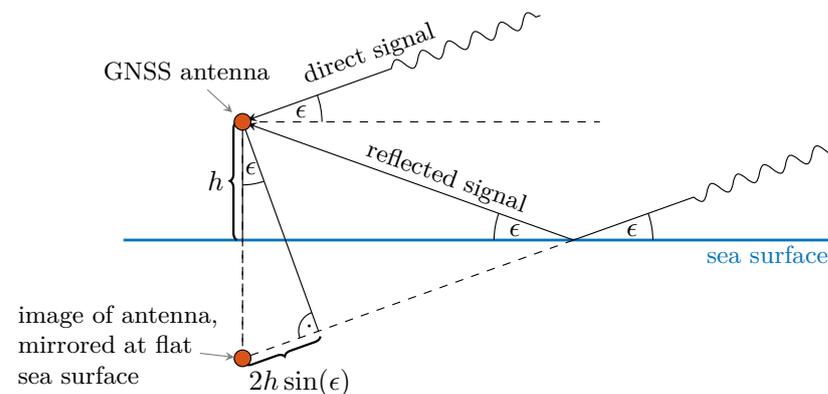


Figure 1. Schematic depiction of basic reflectometric scenario; see [17]. ϵ is the elevation angle of the satellite emitting the signal and h the reflector height. The delay between direct and reflected signal amounts to $2h \sin(\epsilon)$.

Motivated by these insights, we here provide a supervised machine learning approach for the prediction of the SWH at FINO2 based on engineered features obtained from the kernel regression and clustering analyses developed in [16]. In particular, statistics of estimated reflector heights and appendant reliability labels are introduced as derived predictors for the SWH. In the following, we use the terms *engineered feature* (commonly used in the machine learning community) and *derived predictor* (used in geosciences) as synonyms. The same holds for the terms *feature* and *predictor*. Linear models, artificial neural networks, and bagged regression trees are applied for predictions of the SWH, using different subsets of derived predictors as input. To this end, the number of engineered features used for prediction is increased successively. It should be noted that the order of the features used is somewhat arbitrary and that this type of feature selection offers the potential for optimization. In the present work, the key aspect is to demonstrate the general usability and the benefit of the new engineered features for estimation of SWH by assessing the predictive performance of various combinations of models and predictors. One potential flaw of successively increasing the number of predictors can be anticipated from Figure 2: at a certain point, the further increase will lead to overfitting for larger numbers of predictors. To a certain extent, this typical behavior is also observed in our analysis presented here.

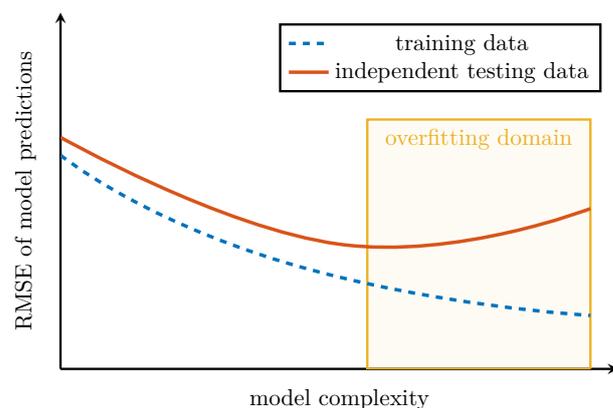


Figure 2. Typical behavior of the root mean square error (RMSE) of model predictions with increasing number of predictors (which amounts to an increase in model complexity); see [18] (pp. 37–38), [19] (pp. 244–252).

The remainder of this work is structured as follows. The methodological foundation of our SWH estimation is given in Section 2. The reflector height estimation with kernel regression and clustering techniques and the associated reliability labeling [16] are presented there, as well as the feature engineering. Different sets of engineered features are composed. Some of these include features based on damping coefficients from the inverse modeling, which we furthermore denote as damping predictors. Moreover, the different supervised machine learning models applied to various subsets of the engineered features are introduced in Section 2. Details on the research platform FINO2, from which the SNR and SWH observations for the present study originated, are presented in Section 3. The GNSS and SWH measurement equipment for raw data acquisition is specified there, and the derived data sets that are used for our machine learning analysis are characterized. Moreover, details on data preprocessing and the configuration involved in the application of the estimation methods are given in Section 3. A comparative analysis of predictive performance of various combinations of models and predictor settings follows in Section 4. Finally, we give some concluding remarks and an outlook on possible further work in Section 5.

2. Methodology

2.1. Introductory Remarks

The basic reflectometric scenario depicted in Figure 1 adheres to a strong idealization of sea state conditions in that the sea surface is considered to be flat. In reality, the sea surface can undergo large and rapid changes, and therefore it is typically treated as a random rough surface when scattering of electromagnetic microwaves at the air–sea interface is under consideration [20], as it is the case for GNSS-R.

An electromagnetic roughness can be attributed to the sea surface for individual spectral components of an electromagnetic signal. For surfaces that exhibit a sufficiently large electromagnetic roughness, there is no global specular reflection into a certain direction (as is the case for the ideally flat sea surface), but diffuse reflection where the incoming electromagnetic signal is scattered in several directions [17].

When the surface curvature radius in a certain region of the air–sea interface is large compared to the wavelength of a spectral component of the signal, the surface appears to be approximately locally flat for this spectral component, and a local specular reflection at a tangent plane can be assumed in the so-called Kirchhoff approximation [20]. Therefore, in the presence of crests, there can occur multiple reflections of an incoming GNSS signal at the sea surface, which are of course not found in the basic scenario of Figure 1.

Despite the aforementioned shortcomings and limitations of the basic scenario of Figure 1, this scenario is successfully used in combination with the neglect of spectral

components other than the carrier frequency of a GNSS signal for the estimation of reflector heights and the SWH with GNSS-R [2,3,6,9]. In the present work, we adopt these common simplifications for our GNSS-R approach and focus on an empirical evaluation of GNSS-R and SWH observation data with machine learning.

2.2. Estimation of Reflector Height

In the idealized scenario depicted in Figure 1, a satellite's GNSS signal is reflected at the sea surface, and it interferes with the directly received signal at the GNSS antenna. With varying elevation angle ϵ of the signal-emitting satellite, the delay between the direct and reflected signal also changes, resulting in constructive or destructive interference at the antenna (with a respective local extremum in the observed SNR), if the subsequent condition is met for a $k \in \mathbb{N}$.

$$\frac{k\lambda}{2} = 2h \sin(\hat{\epsilon}_k) \quad (1)$$

Here, $\hat{\epsilon}_k$ denotes the elevation angle of the satellite emitting the signal at the k th local extremum. λ is the carrier wavelength of the signal. For adjacent local extrema with elevation angles $\hat{\epsilon}_k$ and $\hat{\epsilon}_{k+1}$, differencing Equation (1) yields

$$\frac{\lambda}{4h} = \sin(\hat{\epsilon}_{k+1}) - \sin(\hat{\epsilon}_k) \quad (2)$$

Hence, in this basic reflectometric scenario, the difference in the sine of the elevation angle is constant for adjacent local extrema and equals $\lambda/4h$. From this expression, one can infer the reflector height h if the carrier wavelength is given.

Hereafter, we apply the notation $\underline{L} := \{l \mid l \in \mathbb{N}, l \leq L\}$ for $L \in \mathbb{N}$.

The preceding considerations motivate the following algorithm for estimating the reflector height h from scattered data $\{(\sin(\epsilon_l), \text{SNR}_l) \mid l \in \underline{L}\}$ consisting of pairs of the sine of the observed satellite elevation angle ϵ_l and associated observed value SNR_l of the signal-to-noise ratio.

1. A Nadaraya–Watson kernel regression [21] with a Gaussian kernel function (4) is performed for the scattered data, as exemplified in Figures 3 and 4. The regression function as expectation value of the SNR conditional on $\sin(\epsilon)$ reads as follows:

$$E[\text{SNR} \mid \sin(\epsilon)] = \frac{\sum_{l=1}^L \text{SNR}_l \cdot K(\sin(\epsilon) - \sin(\epsilon_l); \beta)}{\sum_{l=1}^L K(\sin(\epsilon) - \sin(\epsilon_l); \beta)} \quad (3)$$

$$K(x; \beta) := \frac{1}{\sqrt{2\pi\beta^2}} \exp\left(-\frac{x^2}{2\beta^2}\right), \quad x \in \mathbb{R}, \quad \beta \in \mathbb{R}_{>0} \quad (4)$$

Here, β denotes the bandwidth of the kernel function.

2. Local extrema in the kernel regression are determined by numerically detecting changes of the sign of the derivative of the kernel regression function (3). To this end, the derivative is calculated analytically and is evaluated on an equidistant grid on the horizontal axis. If a change of the sign of the derivative is detected for two consecutive gridpoints, the abscissa value of a local extremum is set as the average value of these two gridpoints. Let $(\sin(\tilde{\epsilon}_k))_{k \in \underline{N}}$ be the strictly increasing sequence of these abscissa values of the local extrema.
3. A mean shift clustering [22,23] is applied to the set

$$\mathcal{D} := \{\sin(\tilde{\epsilon}_k) - \sin(\tilde{\epsilon}_{k-1}) \mid k \in \underline{N}, k > 1\} \quad (5)$$

of differences of abscissa values of adjacent local extrema detected by the algorithm. Under idealized conditions, each element of this set would equal the left-hand side of (2). Let $\mathcal{C} \subset \mathcal{D}$ be the largest cluster (in terms of the number of elements contained,

uniqueness is tacitly assumed). The clustering is used for selecting those differences \mathcal{C} that are supposed to be associated with the reflector height according to Equation (2) and for discarding other differences in \mathcal{D} that can originate due to certain distortions (see Figure 4).

4. Averaging over the cluster \mathcal{C} yields the estimate Λ for the right-hand side of (2). Then, the reflector height h can be calculated from (2), as stated in Equation (6).

$$h = \frac{\lambda}{4\Lambda} \text{ with } \Lambda := \frac{1}{|\mathcal{C}|} \sum_{l \in \mathcal{C}} l \tag{6}$$

For low elevations, pronounced interference patterns like that in Figure 3 can occur. Therefore, the estimation procedure outlined above is restricted to low elevation angles with $\sin(\epsilon) < 0.15$. Moreover, either the rise or the set of the respective satellite is used for the procedure. The center of the time interval that is used for observing SNR and satellite elevation (with $\sin(\epsilon) < 0.15$) for a single analysis is used as nominal reference time associated with the estimated reflector height h and the results of the clustering as well as the underlying data \mathcal{D} .

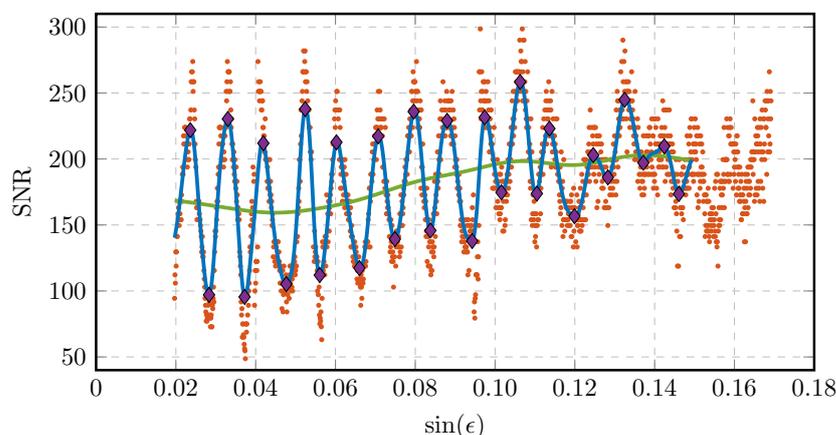


Figure 3. Pronounced interference pattern in SNR observations. A scatter plot of observed SNR data in dependence of the sine of the satellite elevation ϵ is depicted in red. A kernel regression with appropriate bandwidth resolving the regular interference pattern is shown in blue. Local extrema, as detected by the algorithm, are displayed in purple. A kernel regression with larger bandwidth is shown in green.

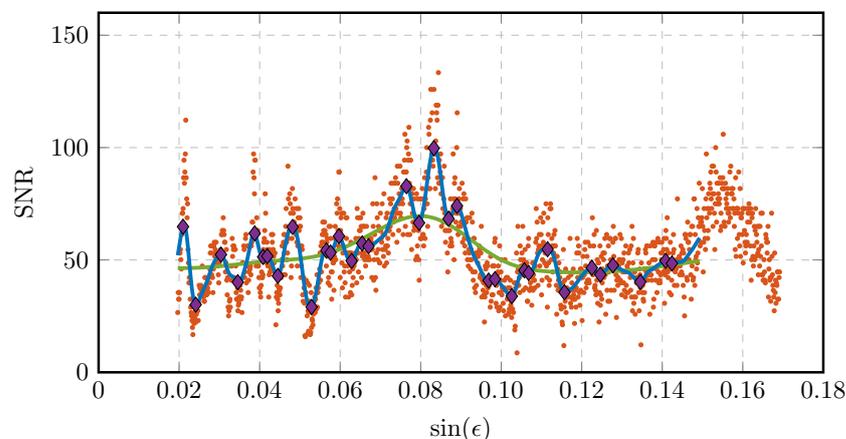


Figure 4. Depiction of kernel regressions and detection of local extrema for scattered data that do not exhibit a pronounced interference pattern. The meaning of the colors is similar to that in Figure 3.

2.3. Reliability Label for Estimated Reflector Height

Under certain circumstances, the estimation procedure for the reflector height h presented in Section 2.2 may lead to unreliable results. Under rough sea state conditions, when prerequisites of the basic reflectometric scenario (stationary and flat reflecting sea surface) are violated more or less severely, one cannot expect to observe regular interference patterns in the SNR data and the algorithm to yield proper results. The meaning of the reflector height h becomes questionable anyway under these conditions. An irregular looking scatter plot without appreciable oscillation, as shown in Figure 4, can be processed by the algorithm, but this typically yields improper values for the reflector height since, in contrast to Figure 3, there is no pronounced interference pattern related to the reflector height via Equation (2). Then, the abscissa values $\sin(\tilde{\epsilon}_k)$ of the local extrema found by the algorithm typically do not match the values $\sin(\hat{\epsilon}_k)$ of the idealized scenario with Equations (1) and (2).

In order to identify estimates of the reflector height that appear unreliable, a respective binary reliability label r is defined with regress to the results of the underlying clustering as follows.

$$r(\mathcal{C}, \mathcal{D}; N_{\min}, P_{\min}) := \begin{cases} 1 \\ 0 \end{cases} \text{ if } \begin{cases} |\mathcal{C}| > N_{\min} \text{ and } |\mathcal{C}|/|\mathcal{D}| > P_{\min} \\ \text{else} \end{cases} \quad (7)$$

If $r(\mathcal{C}, \mathcal{D}; N_{\min}, P_{\min}) = 1$, the estimated reflector height resulting from the clustering is considered to be reliable, otherwise unreliable. The definition (7) reflects the notion that, for a reliable estimate, there should be a large fraction of elements of \mathcal{D} pertaining to the largest cluster \mathcal{C} , which moreover contains a sufficient absolute number of elements. N_{\min} is this minimal number of differences pertaining to the largest cluster \mathcal{C} , and P_{\min} is the minimal fraction of elements contained in \mathcal{C} required for the estimate (6) of h to be considered reliable. $|\mathcal{D}|$ is the total number of available differences to which the clustering is applied; see Equation (5).

Figure 5 resumes the estimation and labeling procedure for the reflector height and leads to the engineering of features for the prediction of the SWH.

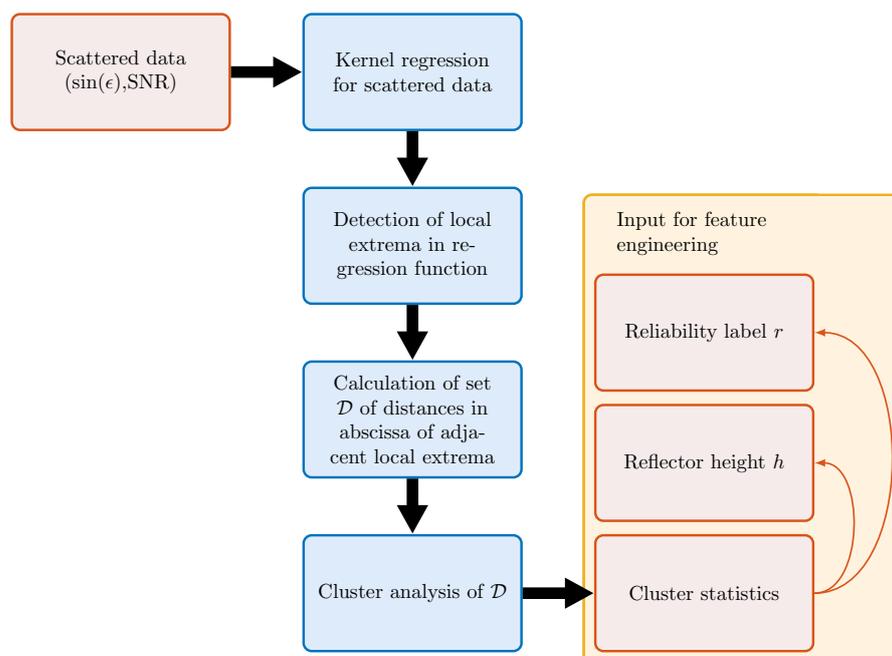


Figure 5. Flow illustrating the estimation of the reflector height with associated reliability label. The quantities used for the engineering of features for the prediction of the SWH are highlighted in the yellow box.

2.4. Engineering of Features for Prediction of Significant Wave Heights

As pointed out above, a relationship between the reliability of a reflector height estimate (characterized by r as defined in (7)) and the sea surface conditions appears plausible. A quantity characterizing the latter is the SWH, and an application of the estimation and labeling procedure presented in Sections 2.2 and 2.3 for data from March 2021 verified that there apparently is a relationship between reliability and SWH [16].

An evaluation of the reliability label (7) for various choices of P_{\min} and N_{\min} may provide additional information related to the SWH, and the computation of the aggregated measure (8) is used as an intermediate step (reducing complexity) for the derivation of predictors for the SWH that are based on the reliability label.

$$R(\mathcal{C}, \mathcal{D}; \mathcal{P}, \mathcal{N}) := \frac{1}{|\mathcal{P}| \cdot |\mathcal{N}|} \sum_{P_{\min} \in \mathcal{P}} \sum_{N_{\min} \in \mathcal{N}} (1 - r(\mathcal{C}, \mathcal{D}; N_{\min}, P_{\min})) \quad (8)$$

For the present analysis, we used $\mathcal{N} = \underline{10} \cup \{0\}$ and $\mathcal{P} = \{n/10 \mid n \in \mathcal{N}\}$.

The quantities defined so far refer to a single analysis with associated reflector height estimate and reliability labelling. In the following, we use the additional index μ to identify and enumerate different analyses, each based on scattered data as depicted in Figure 3 and Figure 4, with an associated kernel regression and clustering analysis. μ indicates the affiliation of a quantity to a particular analysis:

\mathcal{D}_μ : set of differences of μ th analysis; see Equation (5)

\mathcal{C}_μ : largest cluster of μ th analysis

ν_μ : number of clusters obtained in μ th analysis

h_μ : reflector height estimate of μ th analysis

$R_\mu := R(\mathcal{C}_\mu, \mathcal{D}_\mu; \mathcal{P}, \mathcal{N})$

t_μ : reference epoch of μ th analysis

In order to reduce noise and incorporate persistency for relevant quantities, a moving average with window length $2T_w$ can be performed, e.g.:

$$\hat{R}(t; T_w) := \left(\sum_{\mu} R_\mu \chi(t_\mu, t, T_w) \right) / \left(\sum_{\mu} \chi(t_\mu, t, T_w) \right) \quad (9)$$

$$\chi(t_\mu, t, T_w) := \begin{cases} 1 \\ 0 \end{cases} \text{ if } \begin{cases} |t - t_\mu| \leq T_w \\ \text{else} \end{cases} \quad (10)$$

If not stated otherwise explicitly, a similar notation is applied for averages of other quantities.

The analysis in [16] also suggests that for a higher SWH, the reflector height estimates become more noisy and, on average, skewed to higher values. This motivates the subsequent derivation of predictors for the SWH based on statistics of estimated reflector heights.

$\hat{s}_h(t; T_w)$: standard deviation of estimates $\mathcal{H}(t; T_w)$ (11)

$\hat{m}_h(t; T_w)$: maximum of estimates $\mathcal{H}(t; T_w)$ (12)

$\hat{q}_h(t; T_w)$: upper quartile of estimates $\mathcal{H}(t; T_w)$ (13)

where $\mathcal{H}(t; T_w) := \{h_\mu \mid \chi(t_\mu, t, T_w) = 1\}$ (14)

For $M \in \mathbb{N}$ prescribed window lengths $T_w = \tau_1, \dots, \tau_M$ we define tuples $\pi_a(t; T_w)$ of predictors that are concatenated to a comprehensive tuple $\Pi_a(t)$ of predictors for the SWH at time t .

$$\pi_a(t; T_w) := \left(\hat{R}(t; T_w), \hat{\nu}(t; T_w), \hat{h}(t; T_w), \hat{s}_h(t; T_w), \hat{m}_h(t; T_w), \hat{q}_h(t; T_w) \right) \quad (15)$$

$$\Pi_a(t) := \left(\pi_a(t; T_w = \tau_1), \dots, \pi_a(t; T_w = \tau_M) \right) \quad (16)$$

This arrangement of predictors, which solely stems from the kernel regression and clustering techniques outlined in Sections 2.2 and 2.3, is furthermore referred to as predictor setting a . This setting serves as one source of input for predictions of the SWH with different supervised machine learning models at different levels of complexity. Starting with $n = 1$, we choose the first n predictors of the tuple $\mathbf{\Pi}_a(t)$ for setting up supervised machine learning. This procedure is then repeated with successively increased n , until n equals the full length of the predictor tuple $\mathbf{\Pi}_a(t)$. In this context, we use the following terminology: for $n = 7$, e.g., we denote the predictors used as predictor setting $7a$.

2.5. Additional Predictors from Inverse Modeling

A common approach for an estimation of the SWH with scattered data similar to those in Figure 3 is based on an inverse modeling [6], where the parametric model

$$\text{SNR} = \text{trend} + \text{attenuation} \cdot \text{oscillation} \quad (17)$$

$$\text{trend} = \sum_{j=0}^n c_j t^j \quad (18)$$

$$\text{attenuation} = \exp\left(-4\pi^2 \sin^2(\epsilon) \delta^2 / \lambda^2\right) \quad (19)$$

$$\text{oscillation} = A \cos\left(4\pi h \sin(\epsilon) / \lambda + \phi_0\right) \quad (20)$$

is fitted to scattered data consisting of tuples $(\text{SNR}(t), \sin(\epsilon(t)), t)$ for varying time t . Besides the reflector height h , the phase offset ϕ_0 , the amplitude A , the polynomial coefficients c_j , and the damping coefficient δ appear as unknown parameters. By resorting to several estimations with this model, a linear relationship between the hourly average of the estimated damping coefficients δ and the SWH is stated.

Here, for the purpose of comparison, the damping coefficient is similarly calculated for the scattered data of the μ th analysis, and averages like (9) are computed for the damping coefficient for the time windows used in (16). This provides complementary predictors for an estimation of the SWH and allows for the definition of the following additional predictor settings b (Equations (21) and (22)) and c (Equation (23)), similar to predictor setting a .

$$\boldsymbol{\pi}_b(t; T_w) := \left(\hat{\delta}(t; T_w), \hat{R}(t; T_w), \hat{v}(t; T_w), \hat{h}(t; T_w), \hat{s}_h(t; T_w), \hat{m}_h(t; T_w), \hat{q}_h(t; T_w) \right) \quad (21)$$

$$\mathbf{\Pi}_b(t) := \left(\boldsymbol{\pi}_b(t; T_w = \tau_1), \dots, \boldsymbol{\pi}_b(t; T_w = \tau_M) \right) \quad (22)$$

$$\mathbf{\Pi}_c(t) := \left(\hat{\delta}(t; T_w = \tau_1), \dots, \hat{\delta}(t; T_w = \tau_M) \right) \quad (23)$$

In the following, we apply the same terminology for predictor settings b and c as that stated for predictor setting a at the end of Section 2.4.

2.6. Models for Prediction of the Significant Wave Height

2.6.1. Introductory Notes

We applied the scikit-learn library for Python [24–26]. For predictions of the SWH with features from predictor settings a , b , and c , we used a linear model, an artificial neural network, and a bagged regression tree, which we refer to as LinReg, ANN, and BaggedRT, respectively. The model definition and configuration stated below are intimately connected to the data used for model training. In particular, this holds for the configuration governing the model complexity and hyperparameters that can be used to prevent overfitting. Details on the data sets used follow in Section 3. An overview of terminology concerning model types and predictor settings is given in Tables 1 and 2.

Table 1. Overview of predictor settings used as input for model types listed in Table 2.

Predictor Setting	Explanation
<i>a</i>	Features only based on kernel regression, clustering; see Equations (15) and (16).
<i>b</i>	Features based on kernel regression, clustering, as well as damping coefficients; see Equations (21) and (22).
<i>c</i>	Features only based on damping coefficients; see Equation (23).

Table 2. Overview of model types used in supervised machine learning for prediction of the SWH.

Model Type Abbreviation	Explanation
ANN	Artificial Neural Network
LinReg	Linear Regression
BaggedRT	Bagged Regression Tree

For the estimation of model parameters, here concisely denoted by θ , the objective function (24) is used in the model training.

$$\mathcal{L}(\theta) = \frac{1}{K} \left(\sum_{k=1}^K (p_k(\theta) - o_k)^2 + \frac{\alpha}{2} \|\theta\|_2^2 \right) \quad (24)$$

Here, $p_k(\theta)$ are the model predictions and o_k the associated actual observations of the SWH for the training data set used. The dependence on the predictors is omitted in $p_k(\theta)$. For the training of the linear model, $\mathcal{L}(\theta)$ is used with $\alpha = 0$. In this case, $\mathcal{L}(\theta)$ reduces to the mean squared error (MSE) to be minimized in the training. For the ANN, $\mathcal{L}(\theta)$ is applied with $\alpha = 0.0001$. This L2-regularization serves as a measure to prevent overfitting. Further peculiarities of the training of the ANN and the BaggedRT are stated below.

2.6.2. Linear Model

The linear model with parameters $\theta_j, j \in \underline{J} \cup \{0\}$, reads as follows:

$$\text{SWH}(t) = \theta_0 + \sum_{j=1}^J \theta_j \Pi_{a,j}(t) \quad (25)$$

Here, $\Pi_{a,j}(t)$ denotes the j th component of the concatenated predictor tuple $\Pi_a(t)$, which means a component occurring in $\pi_a(t; T_w)$ for a particular choice of T_w . By increasing $J \in \mathbb{N}$ up to the length of $\Pi_a(t)$, the model complexity is increased successively. Likewise, the linear model is used for predictor settings *b* and *c*. In particular, for $J = 1$, the application of the linear model for predictor settings 1*b* and 1*c* means using only $\hat{\delta}(t; \tau_1)$ as predictor. This particular model, which is similar to the determination of the SWH in [6], is used as benchmark in the present study.

2.6.3. Neural Network

The ANN used has fully connected layers, in which 5 are hidden layers. Denoting the number of neurons in the i th layer with $S^{(i)} \in \mathbb{N}$, the layer sizes explicitly read as follows: $S^{(0)}$ is the variable length of the feature vector; $S^{(1)} = 80$; $S^{(2)} = 40$; $S^{(3)} = 10$; $S^{(4)} = 25$; $S^{(5)} = 25$; $S^{(6)} = 1$. The architecture of the network is illustrated in Figure 6.

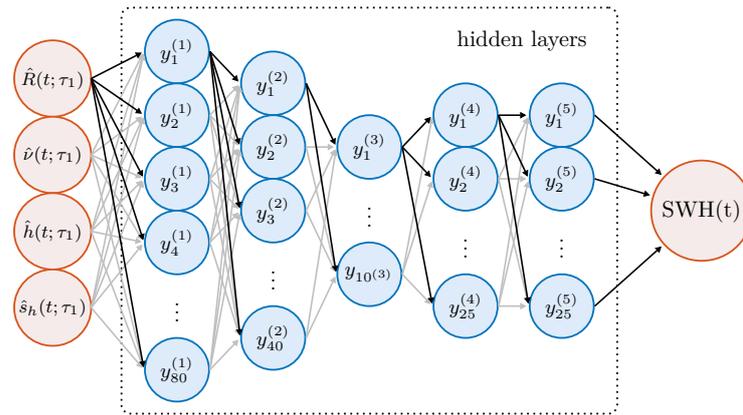


Figure 6. Depiction of the architecture of the ANN for predictor setting 4a. The neurons of the hidden layers are colored in blue. Input and output layer are displayed in red. Predictor information is passed from left to right through the network, as symbolized by the arrows.

We used the rectified linear unit (26) as an activation function.

$$ReLU : \mathbb{R} \rightarrow \mathbb{R}_+ : x \mapsto \max(0, x) \tag{26}$$

The output value $y_k^{(i+1)}$ of the k th neuron in layer $i + 1$ is given in Equation (27).

$$y_k^{(i+1)} = ReLU \left(\theta_k^{(i+1)} + \sum_{s=1}^{S^{(i)}} \theta_{ks}^{(i+1)} y_s^{(i)} \right) \quad i \in \{0\} \cup \underline{5}, \quad k \in \underline{S^{(i+1)}} \tag{27}$$

Here, $y_k^{(0)}$ denotes the k th component of the feature vector, and $y_1^{(6)}$ is the value for the SWH as output of the ANN. The parameters $\theta_k^{(i+1)}, \theta_{ks}^{(i+1)}$ of the ANN are determined in the course of the training, which was performed with the stochastic gradient-based *Adam* optimizer [27] with exponential decay rates $\beta_1 = 0.9$ and $\beta_2 = 0.999$ and an initial learning rate of 0.001. The maximum number of epochs is set to 1000 in order to prevent overfitting. The size of minibatches was set to 200.

With the small hidden layer containing 10 neurons in the middle of the network, surrounded by larger layers, the network structure depicted in Figure 6 exhibits similarities to an autoencoder, which is commonly used for nonlinear dimension reduction [28]. The notion behind this architecture is as follows. Relevant information for the prediction task is extracted from a potentially large or redundant set of predictors and is condensed into quite a small number of intermediate features in the smallest hidden layer. These intermediate features then are fed into the remainder of the network for the prediction task. Similar autoencoder-like network architectures used for supervised machine learning tasks can be found in [29] or [30].

2.6.4. Bagged Regression Tree

With the bagged (bootstrap aggregated) regression tree [31], the prediction of the SWH results in the average of predictions from an ensemble of regression trees, as illustrated in Figure 7. The ensemble used here consists of 3000 individual trees, each one trained on an individual set of regression cases that are randomly drawn from the training data set with a bootstrap [19] (pp. 172–178). The size of the tree-specific bootstrapped training sets equals that of the original training data set.

An individual tree is formed from recursive binary classifications that are successively applied to the instances of the respective training data. Starting with the whole tree-specific training data set, an instance thereof is attributed to a class depending on whether the instance’s value for a particularly chosen feature exceeds an appendant threshold. Classes

can then analogously be partitioned into subclasses. Any class or subclass involved in a binary classification is a node in the graph representation of the tree, as exemplified in Figure 7, and a binary classification providing subsequent two nodes is denoted as split. An optimized choice of the feature and the appendant threshold for a split is performed during training. Depending on certain stopping criteria, a node may not be further split and is then called a leaf, from which the prediction of the SWH results in the average SWH of its class instances in the training.

In our study, the maximal depth of individual trees, i.e., the maximal number of successive splits, is set to 4. As further measure to prevent overfitting, the minimum required number of training samples contained in a node allowing the node to be split is set to 30 and for a node to be considered as leaf to 10, respectively.

The MSE is used during the growth of individual trees with the particular bootstrapped training data in order to decide whether and how, respectively, to split a node in the tree. A split is only performed if it results in a sufficiently large reduction of the MSE and if the stopping rules concerning the growth of the tree are not violated. A more detailed treatment of regression trees can be found in [18] (pp. 305–317).

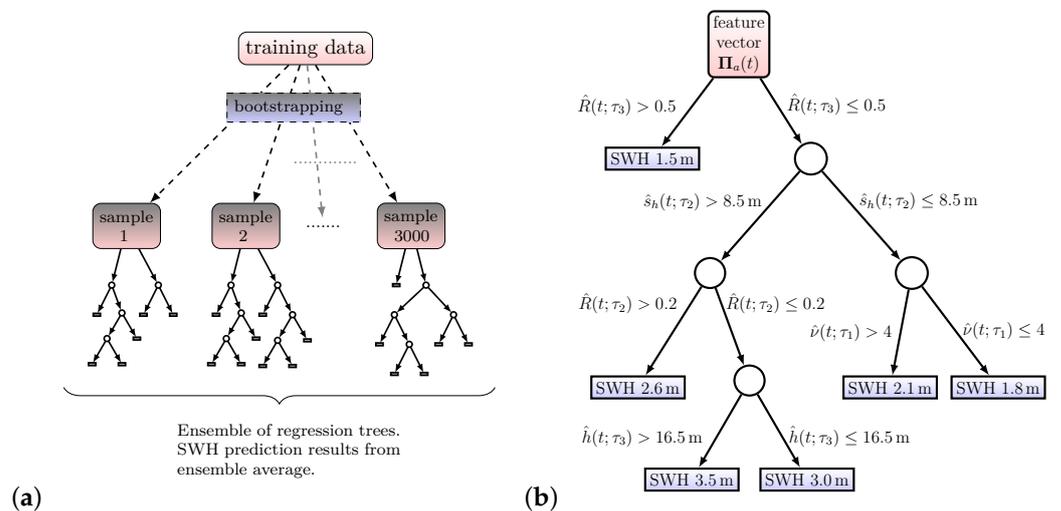


Figure 7. Simplified depiction of bagging for regression trees (a) and a fictive example (b) of a regression tree in the bagged ensemble for predictions with feature vector Π_q .

3. Sensors, Data Sets, and Preprocessing

For our analysis, we used 1 Hz SNR data of GPS L5 (S5X, carrier frequency 1176.45 MHz) observed with a JAVAD TRE 3 DELTA receiver and a LEIAR25.R3 (before 24 June 2021) or LEIAR25.R4 antenna (after 13 July 2021) located at the research station FINO2 in the Baltic Sea. The antenna is mounted at the southwestern corner of the platform built on FINO2; see Figure 8. The large mast of FINO2 and wind turbines in the surrounding of the platform may cause a shielding of GNSS signals. The location of the antenna relative to the tower and an investigation on the azimuth dependency of the quality of SNR signals motivated the introduction of an azimuth mask, discarding observations in the azimuth range from 0° to 90° . SWH time series data with a sampling interval of 1 min were gathered with a radar sensor at FINO2. These SWH observations provided the ground truth for training and testing of models in our supervised machine learning approach. To a limited extent, there were additional SWH measurements available that were made with an acoustic Doppler current profiler (ADCP) sensor at FINO2 with a sampling interval of 1 h. These ADCP measurements of the SWH were used for an additional validation of the SWH predictions resulting from supervised machine learning. IGS precise orbits from the Center for Orbit Determination in Europe (CODE) [32] were applied for computation of elevation angles, which were corrected for tropospheric refraction with a standard method [33] by using

atmospheric pressure and air temperature data from FINO2.

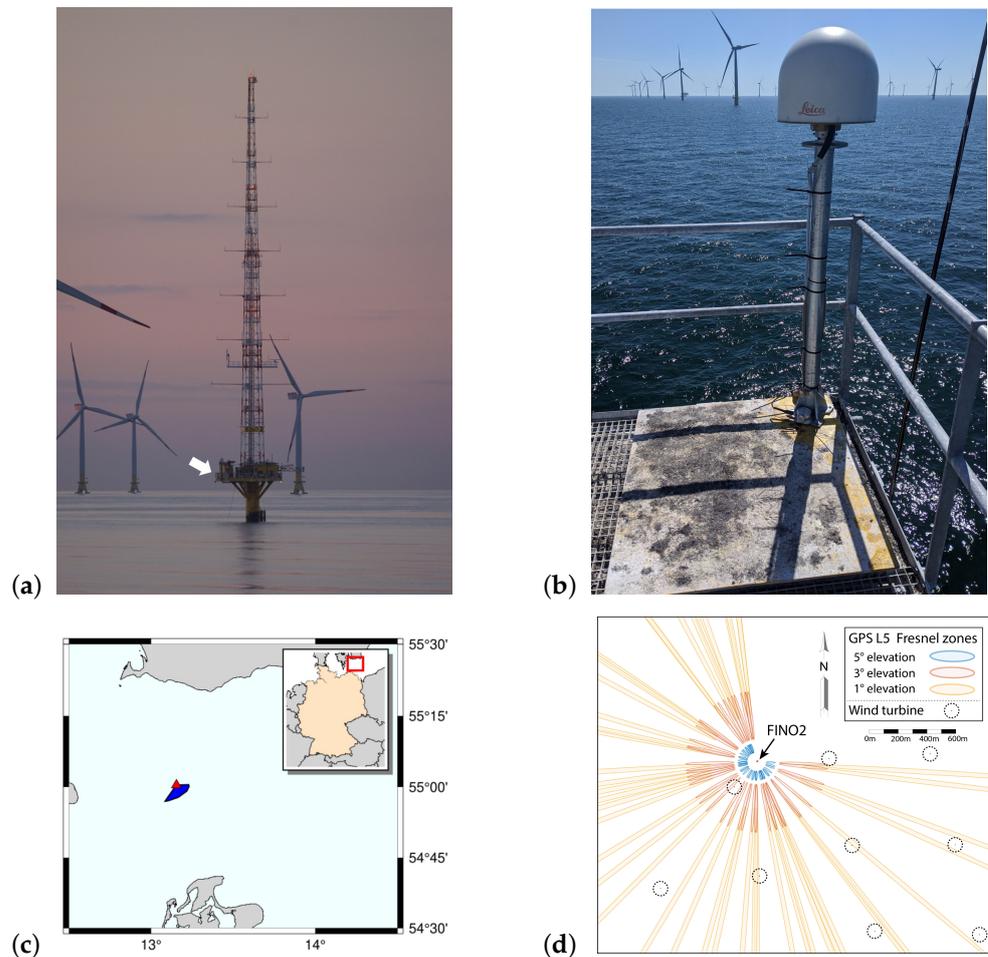


Figure 8. (a) Research station FINO2 in the Baltic Sea. The GNSS antenna is mounted on the platform marked with the white arrow. (b) The GNSS antenna. (c) Map with location of FINO2. The position of FINO2 is highlighted with the red triangle, and the area where wind turbines are located is displayed in blue. (d) Detailed map with location of individual wind turbines and reflection zones for different satellite elevation angles. (Photos: Federal Agency for Cartography and Geodesy, Germany).

We used the reflectometry preprocessing tool [34] for preparing the SNR and elevation data, based on the aforementioned IGS orbit data and RINEX observation files containing the SNR values as transformed, dimensionless quantity $S = 20 \cdot \log_{10}(\text{SNR})$. In the course of feature engineering (see Section 2), we used a bandwidth of 0.001 for the kernel regression and the mean shift clustering. Moreover, an equidistant discretization of the analysis interval $\sin(\epsilon_0^\mu) < \sin(\epsilon) < 0.15$ with 250 gridpoints was chosen for the evaluation of the derivative of the kernel regression, where ϵ_0^μ is the smallest elevation angle available for the individual analysis μ . These specifications of bandwidth and discretization increment enabled a sufficient resolution of the location of the local extrema in the kernel regression. This resolution influences the outcome for the set \mathcal{D} in the first instance as well as the results of the clustering, which also depends on the bandwidth. Our analysis in [16] showed that the final results for the reflector height estimate obtained with the aforementioned configuration were on average in good agreement with the nominal reference value for the reflector height at FINO2, which is why we resorted to this configuration.

We chose the temporal periods described in Table 3 for the data setup of our machine learning analysis with the predictor settings a , b , and c , as defined in Equations (16), (22) and (23). This choice ensures independence of training and testing data. In June and July

2021, two antenna changes were conducted, which is why these months are chosen as one gap between training and testing data.

Table 3. Temporal arrangement and size of datasets used for supervised machine learning.

Dataset	Periods Used	Regression Cases
training	January 2021–May 2021	4914
testing	November 2020, August 2021, September 2021	3402

For the generation of the training, testing, and validation data, the temporal matching of values of SWH and predictors was achieved as follows. The SWH is linearly interpolated for the reference epochs t_μ of individual reflectometric analyses, and the predictors in Equations (16), (22) and (23) are also evaluated at these reference epochs, i.e., for $t = t_\mu$. We use the following five window lengths $T_w = \tau_1, \dots, \tau_5$ in the preceding equations for concretely defining the predictor settings a, b , and c : $\tau_1 = 0.05d$, $\tau_2 = 0.08d$, $\tau_3 = 0.11d$, $\tau_4 = 0.15d$, $\tau_5 = 0.2d$. Here d denotes the length of a solar day.

4. Analysis and Results

For the three types of models described in Section 2.6 and the training and testing periods as characterized in Table 3, we performed multiple supervised machine learning analyses, i.e., an estimation of the SWH based on a successively increased number of predictors chosen from either predictor setting a, b , or c , respectively. Besides the root mean square error (RMSE), we used the mean error (ME) and the coefficient of determination R^2 for characterizing the predictive performance of the machine learning models for a set $\{(p_k, o_k)\}_{k \in K}$ of pairs of predictions p_k and associated observations o_k of the SWH:

$$\text{RMSE} = \sqrt{\frac{1}{K} \sum_{k=1}^K (p_k - o_k)^2} \quad (28)$$

$$\text{ME} = \frac{1}{K} \sum_{k=1}^K (p_k - o_k) \quad (29)$$

$$R^2 = 1 - \left(\sum_{k=1}^K (p_k - o_k)^2 \right) / \left(\sum_{i=1}^K (\bar{o} - o_k)^2 \right) \quad \text{with} \quad \bar{o} = \frac{1}{K} \sum_{k=1}^K o_k \quad (30)$$

As pointed out in Section 3, the ground truth observations o_k of the SWH used for training and testing stem from radar sensor measurements at FINO2. The RMSE is used as a measure of the accuracy of the SWH estimations. With respect to accuracy, a detailed overview of results of the analyses is depicted in Figure 9. In order to assess the sampling uncertainty of the testing RMSE, a block bootstrap, see [19] (pp. 172–178), was applied to the testing data by repeatedly and randomly drawing blocks of consecutive data (each block comprising 300 cases) from the original testing data with replacement, until the accumulated number of cases of the concatenated data blocks reached that of the original testing data. For each randomly generated synthetic sample of testing data (we used 1000), the RMSE was calculated, and the 0.05 and 0.95 percentiles of the resulting RMSE values are used as bounds of a confidence region for the testing RMSE, as indicated with error bars in Figure 9.

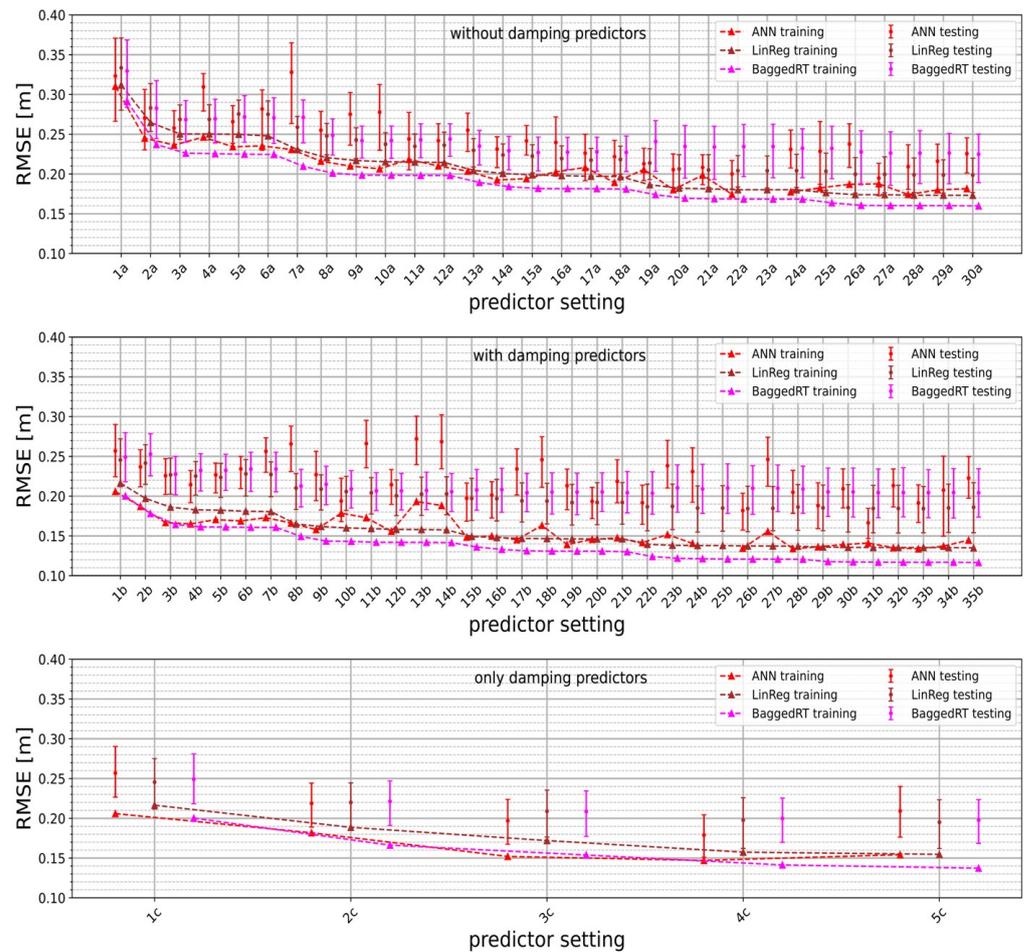


Figure 9. Accuracy obtained in training and testing for different models and predictor settings (top: setting *a*, middle: setting *b*, bottom: setting *c*). Poor results of the ANN for settings 23*a* and 25*b* with training and testing RMSE of approximately 0.5 m are not displayed.

With an increasing number of predictors, the testing RMSE of the linear model more or less decreases in alignment with the corresponding training RMSE. However, there are exceptions concerning the monotony, and the decrease stagnates for a higher number of predictors. For the ANN, there are some erratic jumps in RMSE values, which might occur due to the model complexity in combination with the stochastic *Adam* optimization. For the bagged regression tree, the testing RMSE does not improve for increasing predictor settings larger than 15*a* and 11*b*, respectively, while the training RMSE further decreases. This indicates a tendency of overfitting for the bagged regression tree for larger predictor settings. For predictor setting *a*, the training RMSE is frequently in or near the confidence range of the testing RMSE, at least for the linear model and the ANN. In contrast to this, for predictor setting *b*, the accuracy score obtained with the training data is typically well separated from the confidence range of the appendant testing results (in particular for the bagged regression tree), which may also be interpreted as a sign of overfitting. Exceptions are found for the ANN with predictor settings 10*b* and 31*b*.

Selected results of the analyses depicted in Figure 9 are stated explicitly in Table 4. The application of predictor setting *a* with quite a large number of predictors (ANN with predictor setting 27*a*, linear model with predictor setting 29*a*) yields a remarkable improvement (concerning testing RMSE) over the usage of a single damping predictor with small averaging time window (predictor setting 1*c*).

Table 4. RMSE values and appendant confidence intervals obtained with the testing data for selected predictor settings and models. Results for the benchmark model are given in the first row.

Model Type	Predictor Setting	RMSE	Confidence Interval
LinReg	1c	0.246 m	(0.212 m, 0.275 m)
LinReg	29a	0.199 m	(0.176 m, 0.218 m)
LinReg	5c	0.195 m	(0.162 m, 0.223 m)
ANN	27a	0.195 m	(0.173 m, 0.213 m)
LinReg	26b	0.184 m	(0.155 m, 0.212 m)
ANN	4c	0.179 m	(0.151 m, 0.204 m)
ANN	31b	0.167 m	(0.147 m, 0.185 m)

However, by applying additional damping predictors originating from larger averaging time windows, testing RMSE values similar to those of the ANN with setting 27a and the linear regression with setting 29a can be obtained with only a small number of predictors (settings 4c, 5c). Applying both the alternative and the damping predictors (setting b) can provide better results for the testing RMSE compared to either using only damping predictors (setting c) or using only alternative predictors (setting a). The best results are obtained for the ANN with predictor setting 31b, yielding a significant improvement compared to the benchmark model (linear regression with predictor setting 1c). The testing RMSE is reduced by 32%. The improvement resulting from the application of predictor setting a (in the first instance) and the extended setting b in combination with the use of the ANN is also illustrated in Figure 10.

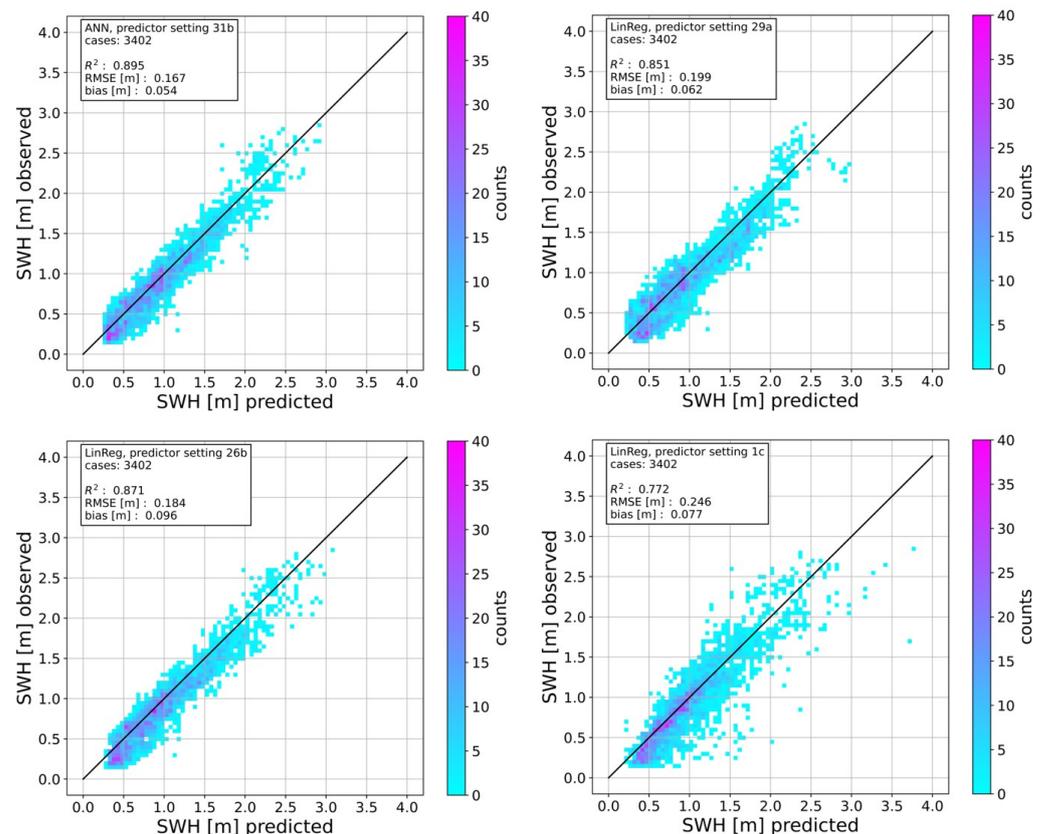


Figure 10. Observed SWH plotted against predicted SWH for the whole testing dataset for four different combinations of model and predictor setting stated in the individual plots. The number of regression cases used and measures for the quality of the estimations are also given.

Despite the complexity of the ANN model with predictor setting 31b, the improvement over the linear model with predictor setting 26b (or setting 31b as well) is not that large,

suggesting that the ANN in essence mimics a linear behavior, with some amelioration due to certain nonlinear deviations from the linear model. As also stated in Figure 10, all estimations exhibit a positive bias, i.e., ME. With respect to this, the estimation with the ANN also slightly outperforms the other ones. The coefficient of determination R^2 (the larger, the better; 1 is the optimal value) is also given as a measure for the quality of the predictions for the testing data.

Time series of the estimated and observed SWH for the testing data available for August and September 2021 are shown in Figure 11, allowing for the assessment of particular error characteristics.

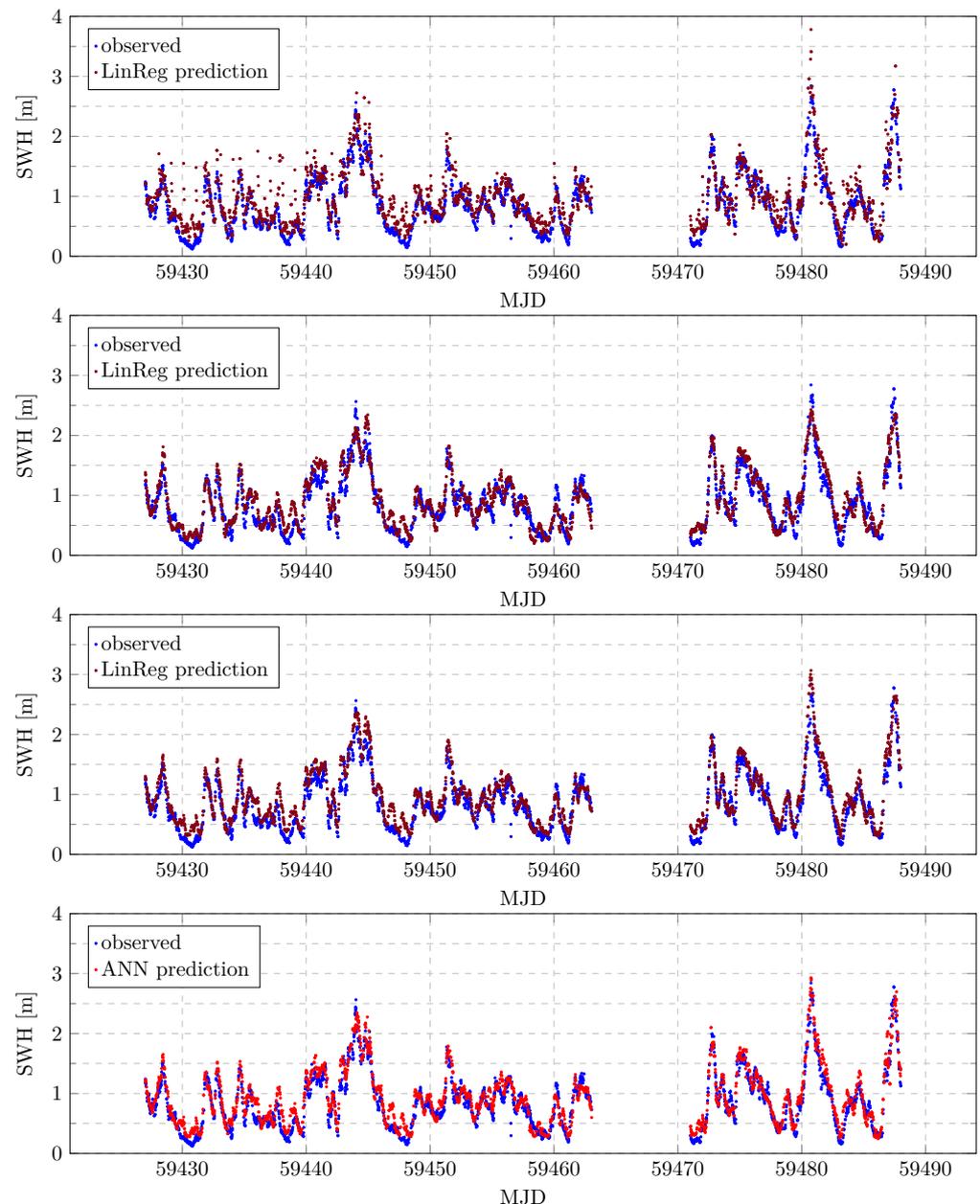


Figure 11. Time series of observed and estimated SWH for testing data available in August and September 2021 for selected estimates with the linear regression or the ANN. Predictor settings used, from top to bottom: 1c, 29a, 26b, 31b. For the intermittent part, the SWH radar sensor measurements were not available. We used the modified Julian date (MJD) as time reference on the horizontal axis.

In Figure 11, the four combinations of predictor settings and models used for the estimation of the SWH agree with those used in Figure 10. As is evident from Figures 10 and 11,

the testing dataset covers a large range of observed SWH values. Of these values, 157 exceed 2 m. While the estimation with the linear regression for predictor setting 1c exhibits an overshooting of the peaks with observed SWH > 2 m in the time series displayed in Figure 11, and the linear regression for predictor setting 29a mostly shows an undershooting, respectively, the estimation with the ANN based on predictor setting 31b can resolve these peaks quite well. Nevertheless, all estimations shown in Figure 11 still exhibit the weakness that there are small peaks in the estimated SWH around local minima of the observed SWH where the observed values are small. This deficiency similarly shows up in the comprehensive overview of conditional biases for the whole testing data given Figure 12. However, the time series depicted in Figure 11 also demonstrate the improvement over the linear regression with predictor setting 1c, which is obtained by using other predictors from settings *a* and *b* or models (i.e., the ANN).

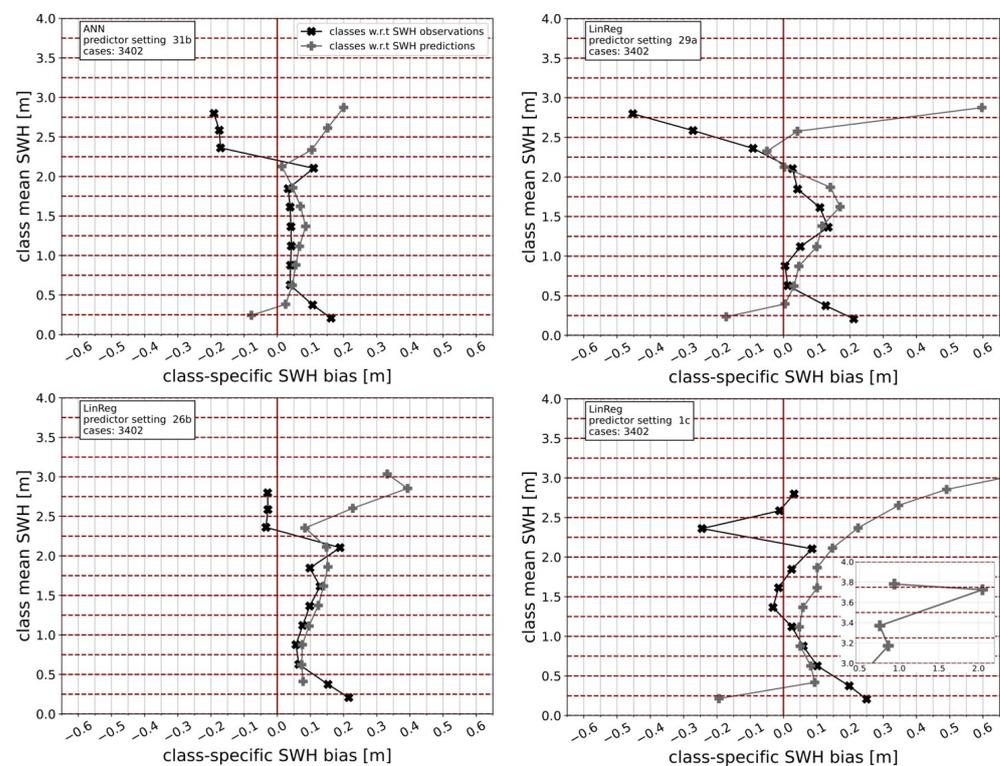


Figure 12. Conditional biases resulting from the whole testing dataset with the model predictions associated with scatter plots from Figure 10. The biases are conditioned on the observed SWH and the predicted SWH, respectively, as follows. The dataset consisting of pairs of predictions and associated observations is partitioned into classes according to the value of the observed (predicted) SWH. For each class, the conditional bias is displayed as a black (grey) cross with mean error of the respective class as abscissa and mean observed (predicted) SWH as ordinate. The partitioning into classes is made with an equidistant binning of the interval [0 m, 4 m) with a bin width of 0.25 m (lower bin value inclusive, upper value exclusive). The boundaries of the bins are shown as red, dashed horizontal lines. Larger conditional biases of the benchmark model exceeding values of 0.5 m and resulting from a vast overshooting of predictions are displayed in the lower-right pane with an additional inserted plot.

One should be aware of the fact that larger predictor settings, such as setting 31b, contain redundant information. Some pairs of predictors exhibit substantial positive correlations (see Figure 13), which can cause undesirable effects in particular for linear regressions; see [35] (pp. 278–293, pp. 406–407).

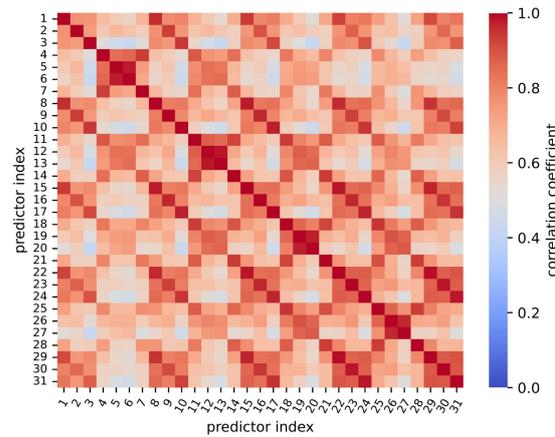


Figure 13. Visualization of the correlation matrix for predictor setting 31b for the testing dataset. The predictor index refers to the position of a predictor variable as occurring in (21) within the concatenated tuple (22). In particular, higher correlations occur for pairs of predictors of the same type, but different time windows T_w used for averaging.

Moreover, the additional use of certain features, e.g., $\hat{s}_h(t, \tau_1)$, $\hat{m}_h(t, \tau_1)$, $\hat{q}_h(t, \tau_1)$, seemingly does not lead to an improvement in accuracy. This becomes evident from certain transient stagnations in the decline of the training RMSE in Figure 9, e.g., for predictor settings 3a to 6a. The design of the ANN with the autoencoder-like bottleneck (see Figure 6) is intended to remedy these flaws and to extract the most valuable information from large predictor settings for predictions of the SWH.

For the training and testing of the supervised machine learning models, we used SWH observations made with the radar sensor at FINO2 as ground truth. For the purpose of additional validation, in Figure 14, predictions of the SWH made with the linear regression model with predictor setting 26b are compared to respective SWH measurements stemming from the ADCP sensor at FINO2. The predictions result from predictor data of a period in May 2021 consecutive to that of the training data. Therefore, neither training nor testing data are involved in this validation. There is good agreement between the SWH predictions and associated SWH observations from ADCP.

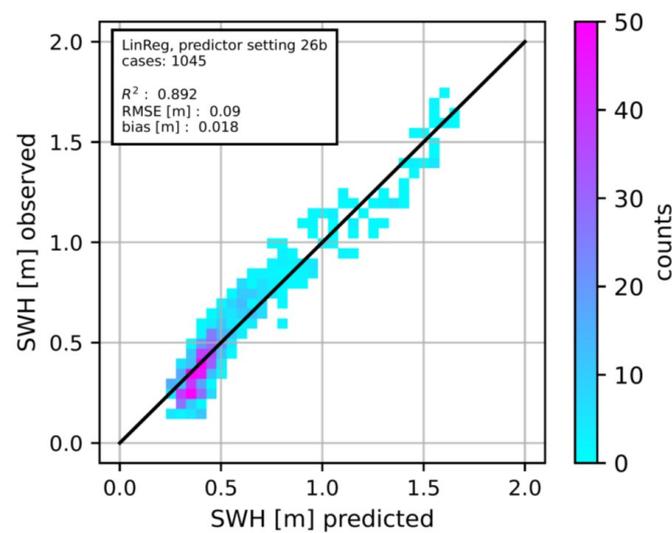


Figure 14. Validation of the GNSS-based SWH predictions against SWH measurements made with the ADCP sensor at FINO2.

5. Summary, Concluding Remarks, and Outlook

In this work, methods for the estimation of the SWH at the stationary sea site FINO2 with supervised machine learning were proposed and evaluated. To this end, engineered features were applied that stem from a reflectometric analysis of GNSS SNR interference patterns with kernel regression and clustering techniques. The engineered features are statistics of samples of the following quantities resulting from individual reflectometric analyses: the estimated reflector height, an aggregated measure calculated from a reliability label for the estimated reflector height, and the number of clusters. An individual sample consists of values of the respective quantity encompassed by a moving time window centered around the epoch for which the SWH is estimated. Appendant features were constructed for five different lengths of the window. For evaluation purposes, subsets of these features with systematically increased size served as input for the estimation of the SWH with three different models: a linear regression, an artificial neural network, and a bagged regression tree. The usability of the estimations was demonstrated, and a remarkable improvement in terms of accuracy for the independent testing data was found in comparison to a common benchmark model, i.e., a linear regression using a time-averaged damping coefficient from a reflectometric analysis with an inverse modeling.

The latter time-averaged damping coefficients were also calculated for the aforementioned five time windows, yielding complementary features (referred to as damping predictors) for the evaluation procedure described above. The additional use of the complementary damping predictors led to a further improvement of accuracy for the testing data, with a significant improvement (compared to the benchmark model) for a particular estimation with a neural network model.

It moreover turned out that the sole use of several damping predictors for the estimation of the SWH with the models mentioned above also leads to a remarkable reduction of RMSE for the testing data in comparison to the benchmark model, and in particular for the ANN.

It should be emphasized that the way of selecting predictors for the SWH estimation from the large predictor settings a and b can be enhanced, as the choice of setting c already demonstrates. Redundant predictor data are certainly provided to the models by the successive increase in the number of predictors underlying the analysis depicted in Figure 9. When the models become unnecessarily complex, the improvement in accuracy for independent testing data is rather small (if there is any) with growing complexity, and overfitting can be induced. The calculation of the correlation matrix (Figure 13) showed that there are substantial correlations between certain pairs of predictor time series, leading to unwanted collinearity in respective linear regressions. A dimension reduction with a principal component analysis may be helpful for the generation of a set of uncorrelated predictors without losing valuable predictor information. To a certain extent, the ANN used in the present study already performs an implicit, nonlinear dimension reduction due to its autoencoder-like architecture. Another way of optimizing the choice of predictors for a linear model is a stepwise multilinear regression scheme, as used in model output statistics forecasts [36].

Nevertheless, despite the redundancy in predictor data, the present analysis demonstrates the benefit and the improvements that can be obtained from the engineered features in combination with certain model types (in particular the ANN). Although the application of an ANN provided the best results in terms of RMSE (and ME), it is remarkable that linear models also performed quite well with certain predictor sets, which may be interpreted as a result of reasonable feature engineering. In contrast to ANNs, linear models are easy to understand and to interpret, as long as there is no problem with collinearity. This would be an argument to envisage the use of linear models in our context, bearing in mind that the ANNs did not perform much better. Referring to this, it should be noted that we did not perform a comprehensive tuning of the configuration of the complex models, i.e., the ANNs and the bagged regression trees. As a possible measure for improving the predictive performance of the latter, a random feature selection for the split of nodes could be intro-

duced for larger predictor settings, which effectively means a transition to a random forest model [37].

Another means for improving SWH predictions may be applicable, if the biases conditional on the predictions (see Figure 12) turn out to be a general, characteristic property of the respective model. Then, a prediction can be calibrated by subtracting the conditional bias for that prediction, if the conditional bias is known from investigations with independent data. A similar technique known as *reliability diagram statistics method* is used for calibration purposes in the field of probabilistic weather forecasting [38].

Moreover, in future work, the transfer of the methods presented here to other stationary sea sites could be envisaged, taking into account peculiarities of the stations such as pronounced tides, which are not found at FINO2 [39]. To this end, it could be helpful to increase the actual, physical reflector height in order to reduce the time that is necessary to observe an evaluable GNSS SNR interference pattern (see Equation (1)).

Due to the use of the (partially quite large) time windows for feature engineering, which also comprise future observations (with respect to the epoch for which the SWH is estimated), the methods developed here are not suited for real-time applications or for a forecast of SWH values in the future. The impact of a restriction on respective time windows in the past could be an objective for future work. This also holds for the incorporation of additional features related to the roughness of the sea surface, e.g., meteorological data like wind speed and wind direction, which could help to further improve the estimations. Apart from the meteorological data used for correction of tropospheric refraction, the predictions of SWH with supervised machine learning in the present work solely rely on features derived from GNSS-related data.

Author Contributions: Conceptualization, J.M.B. and O.R.; methodology, J.M.B.; software, J.M.B. and O.R.; validation, J.M.B. and O.R.; formal analysis, J.M.B. and O.R.; investigation, J.M.B. and O.R.; data curation, O.R. and J.M.B.; writing—original draft preparation, J.M.B.; writing—review and editing, O.R. and J.M.B.; visualization, J.M.B. and O.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Since the GNSS site on FINO2 is still in the experimental phase, the SNR data are only available for internal use. Precise orbits from the CODE Analysis Center can be freely downloaded from the Astronomical Institute of the University of Bern via <http://www.aiub.unibe.ch/download/CODE> (accessed on 29 June 2022). The meteorological data, the SWH radar, and the SWH ADCP observations were downloaded from the FINO observation database <https://www.fino2.de/de/projekt/fino-datenbank.html> (accessed on 29 June 2022). Data access requires registration but is free of charge.

Acknowledgments: We gratefully acknowledge the provision of the SNR observation data by our colleagues from Unit G3 at the Federal Agency for Cartography and Geodesy, Germany. Moreover, we gratefully acknowledge the provision of the SWH radar and ADCP sensor data as well as the meteorological data from FINO2 by the FINO initiative (Research Platforms in the North and Baltic Seas), which was organized by the Federal Ministry for Economic Affairs and Climate Action (BMWK) on the basis of a resolution of the German Bundestag, by Project Management Jülich (PTJ) and coordinated by the BSH.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADCP	Acoustic Doppler Current Profiler
ANN	Artificial Neural Network
BaggedRT	Bagged Regression Tree
BMWK	Bundesministerium für Wirtschaft und Klimaschutz
BSH	Bundesamt für Seeschifffahrt und Hydrographie
CODE	Center for Orbit Determination in Europe

FINO	Forschung in Nord- und Ostsee
GNSS	Global Navigation Satellite System
GNSS-R	GNSS Reflectometry
GPS	Global Positioning System
IGS	International GNSS Service
IR	Interferometric Reflectometry
LinReg	Linear Regression
ME	Mean Error
MJD	Modified Julian Date
MSE	Mean Squared Error
PTJ	Projekträger Jülich
RINEX	Receiver Independent Exchange Format
RMSE	Root Mean Square Error
SNR	Signal-to-Noise Ratio
SWH	Significant Wave Height

References

1. Atlas, R.; Hoffman, R.N.; Ardizzone, J.; Leidner, S.M.; Jusem, J.C.; Smith, D.K.; Gombos, D. A Cross-calibrated, Multiplatform Ocean Surface Wind Velocity Product for Meteorological and Oceanographic Applications. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 157–174. [\[CrossRef\]](#)
2. Larson, K.M.; Lay, T.; Yamazaki, Y.; Cheung, K.F.; Ye, L.; Williams, S.D.P.; Davis, J.L. Dynamic Sea Level Variation from GNSS: 2020 Shumagin Earthquake Tsunami Resonance and Hurricane Laura. *Geophys. Res. Lett.* **2020**, *48*, e2020GL09137. [\[CrossRef\]](#)
3. Peng, D.; Feng, L.; Larson, K.M.; Hill, E.M. Measuring Coastal Absolute Sea-Level Changes Using GNSS Interferometric Reflectometry. *Remote Sens.* **2021**, *13*, 4319. [\[CrossRef\]](#)
4. Reinking, J.; Roggenbuck, O.; Even-Tzur, G. Estimating Wave Direction Using Terrestrial GNSS Reflectometry. *Remote Sens.* **2019**, *11*, 1027. [\[CrossRef\]](#)
5. Alonso-Arroyo, A.; Camps, A.; Park, H.; Pascual, D.; Onrubia, R.; Martín, F. Retrieval of Significant Wave Height and Mean Sea Surface Level Using the GNSS-R Interference Pattern Technique: Results From a Three-Month Field Campaign. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3198–3209. [\[CrossRef\]](#)
6. Roggenbuck, O.; Reinking, J.; Lambertus, T. Determination of Significant Wave Heights Using Damping Coefficients of Attenuated GNSS SNR Data from Static and Kinematic Observations. *Remote Sens.* **2019**, *11*, 409. [\[CrossRef\]](#)
7. Holthuisen, L. *Waves in Oceanic and Coastal Waters*; Cambridge University Press: Cambridge, UK, 2010.
8. Soulat, F.; Caparrini, M.; Germain, O.; Lopez-Dekker, P.; Taani, M.; Ruffini, G. Sea state monitoring using coastal GNSS-R. *Geophys. Res. Lett.* **2004**, *31*, 1–4. [\[CrossRef\]](#)
9. Löfgren, J.; Haas, R. Sea level measurements using multi-frequency GPS and GLONASS observations. *EURASIP J. Adv. Signal Process.* **2014**, *2014*, 50, 1–13. [\[CrossRef\]](#)
10. Roggenbuck, O.; Reinking, J. Sea Surface Heights Retrieval from Ship-Based Measurements Assisted by GNSS Signal Reflections. *Mar. Geod.* **2019**, *42*, 1–24. [\[CrossRef\]](#)
11. Wang, F.; Yang, D.; Yang, L. Retrieval and Assessment of Significant Wave Height from CYGNSS Mission Using Neural Network. *Remote Sens.* **2022**, *14*, 3666. [\[CrossRef\]](#)
12. Bué, I.; Semedo, A.; Catalão, J. Evaluation of HF Radar Wave Measurements in Iberian Peninsula by Comparison with Satellite Altimetry and in Situ Wave Buoy Observations. *Remote Sens.* **2020**, *12*, 3623. [\[CrossRef\]](#)
13. Ludeno, G.; Serafino, F. Estimation of the Significant Wave Height from Marine Radar Images without External Reference. *J. Mar. Sci. Eng.* **2019**, *7*, 432. [\[CrossRef\]](#)
14. Passaro, M.; Fenoglio-Marc, L.; Cipollini, P. Validation of Significant Wave Height From Improved Satellite Altimetry in the German Bight. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2146–2156. [\[CrossRef\]](#)
15. Bayindir, C.; Frost, J.; Barnes, C. Assessment and Enhancement of SAR Noncoherent Change Detection of Sea-Surface Oil Spills. *IEEE J. Ocean. Eng.* **2018**, *43*, 211–220. [\[CrossRef\]](#)
16. Becker, J.; Roggenbuck, O. Sea level monitoring with GNSS reflectometry based on non-parametric modelling. In Proceedings of the 1st Workshop on Data Science for GNSS Remote Sensing, Potsdam, Germany, 13–15 June 2022.
17. Braasch, M. Multipath. In *Springer Handbook of Global Navigation Satellite Systems*; Teunissen, P., Montenbruck, O., Eds.; Springer: Berlin/Heidelberg, Germany, 2017. [\[CrossRef\]](#)
18. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer Series in Statistics; Springer: Berlin/Heidelberg, Germany, 2009.
19. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*, 3rd ed.; Academic Press Elsevier: Oxford, UK, 2011.
20. Thompson, D. Microwave Scattering from the Sea. In *Synthetic Aperture Radar Marine User's Manual*; Jackson, C., Apel, J., Eds.; U.S. Department of Commerce, National Oceanic and Atmospheric Administration: Washington, DC, USA, 2004.
21. Watson, G.S. Smooth Regression Analysis. *Sankhyā Indian J. Stat. Ser. A* **1964**, *26*, 359–372.

22. Mean Shift Clustering Using a Flat Kernel; Scikit-Learn Machine Learning Library for Python. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.cluster.MeanShift.html> (accessed on 6 December 2022).
23. Comaniciu, D.; Meer, P. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 603–619. [CrossRef]
24. Ordinary Least Squares Linear Regression; Scikit-Learn Machine Learning Library for Python. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html (accessed on 6 December 2022).
25. Multi-Layer Perceptron Regressor; Scikit-Learn Machine Learning Library for Python. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html (accessed on 6 December 2022).
26. A Random Forest Regressor; Scikit-Learn Machine Learning Library for Python. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html> (accessed on 6 December 2022).
27. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2015**, arXiv:1412.6980v8.
28. Hinton, G.; Salakhutdinov, R.R. Reducing the Dimensionality of Data with Neural Networks. *Science* **2006**, *313*, 504. [CrossRef]
29. Zhang, S.; Yao, J.; Hu, J.; Zhao, Y.; Li, S.; Hu, J. Deep Autoencoder Neural Networks for Short-Term Traffic Congestion Prediction of Transportation Networks. *Sensors* **2019**, *19*, 2229. [CrossRef]
30. Nam, K.; Wang, F. The performance of using an autoencoder for prediction and susceptibility assessment of landslides: A case study on landslides triggered by the 2018 Hokkaido Eastern Iburi earthquake in Japan. *Geoenviron. Disasters* **2019**, *6*, 19. [CrossRef]
31. Breiman, L. Bagging Predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
32. Dach, R.; Schaer, S.; Arnold, D.; Kalarus, M.S.; Prange, L.; Stebler, P.; Villiger, A.; Jäggi, A. *CODE Final Product Series for the IGS*; Astronomical Institute, University of Bern: Bern, Switzerland, 2020. Available online: <http://www.aiub.unibe.ch/download/CODE> (accessed on 29 June 2022).
33. Bennett, G.G. The Calculation of Astronomical Refraction in Marine Navigation. *J. Navig.* **1982**, *35*, 255–259. [CrossRef]
34. Reflectometry Preprocessing Tool. Available online: <https://github.com/kristinemlarsen/gnssrefl> (accessed on 30 May 2022).
35. Kutner, M.H.; Nachtsheim, C.J.; Neter, J.; Li, W. *Applied Linear Statistical Models*, 5th ed.; McGraw-Hill Irwin: New York, NY, USA, 2005.
36. Glahn, H.R.; Lowry, D.A. The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *J. Appl. Meteorol. Climatol.* **1972**, *11*, 1203–1211. [CrossRef]
37. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
38. Kober, K.; Craig, G.C.; Keil, C.; Dörnbrack, A. Blending a probabilistic nowcasting method with a high-resolution numerical weather prediction ensemble for convective precipitation forecasts. *Q. J. R. Meteorol. Soc.* **2012**, *138*, 755–768. [CrossRef]
39. Medvedev, I.P.; Rabinovich, A.B.; Kulikov, E.A. Tides in Three Enclosed Basins: The Baltic, Black, and Caspian Seas. *Front. Mar. Sci.* **2016**, *3*, 46. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.