Article

Locally Weighted Non-Parametric Modeling of Ship Maneuvering Motion Based on Sparse Gaussian Process

Zhao Zhang ^{1,2} and Junsheng Ren ^{1,2,*}¹ Nautical Dynamic Simulation and Control Laboratory, Dalian Maritime University, Dalian 116026, China; zhang_z@dlmu.edu.cn² Navigation College, Dalian Maritime University, Dalian 116026, China

* Correspondence: jsren@dlmu.edu.cn

Abstract: This paper explores a fast and efficient method for identifying and modeling ship maneuvering motion, and conducts a comprehensive experiment. Through the ship maneuvering test, the dynamics interaction between ship and the environment is obtained. Then, the LWL (Locally Weighted Learning algorithm) underlying architecture is constructed by sparse Gaussian Process to reduce the data requirements of LWL-based ship maneuvering motion modeling and to improve the performance for LWL. On this basis, a non-parametric model of ship maneuvering motion is established based on the locally weighted sparse Gaussian Process, and the traditional mathematical model of ship maneuvering motion is replaced by the generative model. This generative model considers the hydrodynamic effects of ships, and reduces the sensitivity of local weighted learning to sample data. In addition, matrix operations are transferred to the auxiliary platform to optimize the calculation performance of the method. Finally, the simulation results of ship maneuvering motion indicate that this method has the characteristics of efficiency, rapidity and universality, and its accuracy conforms to engineering practice.

Citation: Zhang, Z.; Ren, J. Locally Weighted Non-Parametric Modeling of Ship Maneuvering Motion Based on Sparse Gaussian Process. *J. Mar. Sci. Eng.* **2021**, *9*, 606. <https://doi.org/10.3390/jmse9060606>

Academic Editor: Alessandro Ridolfi

Received: 6 May 2021

Accepted: 27 May 2021

Published: 1 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 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/>).

Keywords: ship maneuvering system; ship motion modeling; Gaussian process; locally weighted learning; identification modeling

1. Introduction

The prediction of ship motion attitude is the prerequisite for ship maneuvering decision and motion planning. How to accurately construct a ship motion mathematical model to accurately predict the trend of ship motion has always been a popular research direction for vessels, semisubmersibles/submersible platforms, and unmanned underwater vehicles. Furthermore, due to the time delays and errors in obtaining ships' dynamic data which will cause ship maneuvering errors and out-of-control drift, there are some great threats to the ships itself and shore-based facilities. Therefore, an accurate, rapid and economical ship maneuvering motion model is indispensable during sailing and berthing operations.

Compared with ship parametric modeling such as Abkowitz or MMG (Manoeuvring Mathematical Model Group) [1], which is obtained through dynamics theoretical analysis after determining each parameter in the known model structure, non-parametric modeling [2] is the response obtained directly or indirectly from the data analysis of the actual system, which avoids the problem of parameter drift of parametric modeling and the inaccuracy caused by the unmodeled dynamics. In addition, the construction of non-parametric modeling is also easier than parametric modeling. With the rapid development of machine learning and artificial intelligence, intelligent methods provide an effective way for non-parametric modeling, which has become popular in the ship motion modeling field from SVM (Support Vector Machines) [2,3] to ANN (Artificial neural networks) [4,5], from Bionic intelligence method [6,7] to deep learning [8]. However, they are not invulnerable. For example, the optimal number of hidden layer neurons cannot be determined, and their

generalization ability in ANN is limited. The size of training data is not easy to determine and model hyperparameters are not easy to grasp for SVM.

In ship motion identification modeling, LWL [9,10] is used as a non-parametric method [11] like the above methods, to construct a non-parametric model of ship maneuvering motion. The difference of this method is that whenever it predicts a new sample value, it will solve the new parameter value according to the weight of the local model training set, so it can learn from a large amount of data and add data incrementally. And LWL can effectively overcome the problems of unmodeled dynamics and parameter drift [12]. However, it does not provide a generative model for the actual system, and requires training data and test data to be independently generated in the same way. Moreover, the local model underlying architecture is in the least square form, i.e., a large effective data set is required to obtain a model with high accuracy, which seriously affects modeling efficiency. With the massive amount of information brought by big data, how to expand data-driven ship motion modeling is a question worth pondering in the LWL's future development. The existing LWL algorithm for ship maneuvering motion modeling has tricky parameters that need to be adjusted [7,10] and these parameters strongly affect the deviation-variance and learning speed of the ship maneuvering motion mathematical model.

In order to overcome these issues, a method for LWL to identify ship maneuvering motion model based on Gaussian process (GP) [13] is proposed. GP is also a non-parametric method. It combines the characteristics of the kernel method and Bayesian inference. It has the advantages of the above two machine learning methods and strict statistical theoretical basis. Meanwhile, the method is suitable for complex learning problems such as nonlinearity, small samples, and high dimensionality etc. and with strong generalization ability [14]. Compared with ANN, SVM and other methods, it has the advantages of easy implementation, self-adapting hyperparameters, flexible non-parametric inference, and statistics significance of prediction results. On the other hand, the GP provides a generative model with black box automatic parameter adjustment. It has no significant requirements for the training data distribution of and has high learning capabilities.

Recently, GP has shown promising prospects in robot system modeling [15–17], system identification [18], time Series [19] and ocean engineering applications [20,21] etc., which benefited from the continuous reduction of the computational complexity of the method. Simultaneously, its intuitive structure and Bayesian framework are very tolerant of sample data and hyperparameter. On the contrary, although the LWL can directly analyze the relationship between the input and output of the ship motion system [9], this approach has high requirements on sample data and hyperparameter. Importantly, it is hard to make a compromise between modeling accuracy and computational efficiency [11,22]. The development of GP theory provides a favorable basis for rapid non-parametric modeling of ship maneuvering motion.

In this article, we propose Sparse Gaussian process based locally weighted learning (LSGP) for ship maneuvering motion non-parametric modeling. In this way, we retain the advantages of the two methods. Inspired by LWL, the training data is divided into local domains, and a GP is learned independently in the local domains during training and prediction. Using approximate reasoning technology, a sparse GP model is constructed in local domains without losing the advantages of LWL. The covariance matrix is solved by C++ on auxiliary platform to improve computational efficiency. Meanwhile, the generative model of LSGP can be better applied to engineering simulation.

The arrangement of this article is as follows. Section 2 summarizes the background knowledge related to this content. The algorithm of LSGP is described in Section 3. In Section 4, an overview of the ship dynamics model, as well as ship identification scheme and simulation examples are given to verify the effectiveness of the proposed method. Section 5 presents the main conclusions and further prospects.

2. Related Background Review of Non-Parametric Modeling Methods

The main content of this section is a background review of the non-parametric modeling knowledge required in this paper, which is about locally weighted learning and Gaussian Processes.

Generalized non-parametric modeling methods are about approximating a system (or function) $f(x)$ through a number of n training points (or observations) $\{x_i, y_i\}_{i=1}^N$. Here, x denotes the training input point and y the system output. We assume that the training outputs are corrupted by noise, $y_i = f(x_i) + \epsilon$, with $\epsilon \sim N(0, \sigma_n^2)$, $f(x_i) \rightarrow x_i\theta$ is a linear model usually. As shorthand notation, we merge all the training points x_i into a training set $\mathbf{X}$ and all corresponding output values y_i into an output vector $\mathbf{y}$. Such that $\mathbf{y} = \mathbf{f} + \boldsymbol{\epsilon} \rightarrow \mathbf{X}^T\boldsymbol{\theta} + \boldsymbol{\epsilon}$, with $\boldsymbol{\epsilon} \sim \mathbf{N}(\mathbf{0}, \boldsymbol{\Sigma}_n)$ and $\boldsymbol{\Sigma}_n = \sigma_n^2\mathbf{I}$, $\mathbf{I}$ is the identity matrix, N and $\mathbf{N}$ represents normal distribution.

2.1. Locally Weighted Learning

LWL uses the idea of locally learning to solve the task of compressing a large amount of data into a small number of local models. In the spirit of Taylor's expansion, LWL's idea is that simple models may be accurate in the local range, and it may be difficult to find good nonlinear features to capture the entire system globally, i.e., many good local models may constitute a good global model.

LWL trains L local models, assuming that each local model has a local prediction in linear form $\hat{\mathbf{y}}_l = \mathbf{f} + \boldsymbol{\epsilon} \rightarrow \mathbf{X}_l^T\boldsymbol{\theta}_l$. When it works, the local predictions $\hat{\mathbf{y}}_l$ are combined into a joint prediction at the query input q as a normalized weighted sum. Hence, the prediction can be obtained according to Bayesian theory

$$\hat{\mathbf{y}}(\mathbf{X}) = \frac{\sum_{l=1}^L w_l \hat{\mathbf{y}}_l(\mathbf{X})}{\sum_{l=1}^L w_l} = \sum_l \Phi_l^T \boldsymbol{\theta}_l \quad (1)$$

with

$$\Phi_l = p(l | \mathbf{X}) = \frac{p(l, \mathbf{X})}{p(\mathbf{X})} = \frac{p(l, \mathbf{X})}{\sum_{l=1}^L p(l, \mathbf{X})} = \frac{w_l \mathbf{X}_l}{\sum_{l=1}^L w_l} \quad (2)$$

with w_l is the weighted or attributed of the local model. The weight w_l determines whether data point $\mathbf{X}_l$ falls into the region of validity of local model l , similar to a receptive field, and is usually characterized with a Gaussian kernel

$$w_l = \exp\left(-1/2(\mathbf{X}_l - q)^T h(\mathbf{X}_l - q)\right) \quad (3)$$

where h is the distance measure and $\boldsymbol{\theta}_l$ is the regression parameter. In the learning process, the shape of h and the regression parameters of the local model $\boldsymbol{\theta}_l$ are adjusted by minimize the loss function between the predicted value and the observed target

$$\text{Loss} = \sum_l w_l \left(\hat{\mathbf{y}}_l - \mathbf{X}_l^T \boldsymbol{\theta}_l\right)^2 \quad (4)$$

The regression parameter $\boldsymbol{\theta}_l$ can be calculated incrementally using the least square method [9,11]. The distance measure h determines the size and shape of each local model; it can be updated using cross-validation [9], gradient method [22], and intelligence methods [7,10].

An important observation is that Equation (4) cannot be interpreted as least-squares estimation of the linear model Φ from Equation (1). Thus, LWL cannot be converted probabilistically as one generative model for training and test simultaneously. Meanwhile, LWL requires several tuning hyperparameters, and its optimal value may be deeply dependent on the training data. Also, this high dependence makes this approach inefficient and time-consuming for learning and prediction. Here, we explored a probabilistic method to replace

LWL. This method reduces the need for parameter adjustment and data requirements, but retains the characteristics of LWL.

2.2. Gaussian Process

A Gaussian process is a set of random variables, in which any finite-dimensional random variable obeys a uniform joint Gaussian distribution. It is a powerful alternative method for accurate function approximation in high-dimensional space.

The Gaussian likelihood can be obtained as

$$p(\mathbf{y} | \mathbf{X}, \boldsymbol{\theta}) \sim \mathcal{GP}(\mathbf{y} | \mathbf{X}^T \boldsymbol{\theta}, \boldsymbol{\Sigma}_n) \quad (5)$$

For the regression parameters $\boldsymbol{\theta}$, suppose it is a Gaussian prior distribution with a mean value $\boldsymbol{\mu}$ of 0 and a covariance matrix of $\boldsymbol{\Sigma}_{\text{prior}}$, i.e., $p(\boldsymbol{\theta}) \sim \mathcal{GP}(\boldsymbol{\theta} | \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\text{prior}})$. The mean function is the a priori expectation of the unknown function. The covariance function, or kernel, is typically selected by design as a measure of similarity between data points. Therefore, Combining the Gaussian prior distribution $p(\boldsymbol{\theta})$ and the likelihood function Equation (5), the posterior distribution of the regression parameters can be calculated

$$p(\boldsymbol{\theta} | \mathbf{y}, \mathbf{X}) = [p(\mathbf{y} | \mathbf{X}, \boldsymbol{\theta}) p(\boldsymbol{\theta})] / p(\mathbf{y} | \mathbf{X}) \sim \mathcal{GP}(\mathbf{m}, \boldsymbol{\Sigma}_{\text{post}}) \quad (6)$$

The available mean and covariance matrix are

$$\begin{aligned} \mathbf{m} &= (\boldsymbol{\Sigma}_n^{-1} \mathbf{X}^T \mathbf{X} + \boldsymbol{\Sigma}_{\text{prior}}^{-1})^{-1} (\boldsymbol{\Sigma}_n^{-1} \mathbf{X}^T \mathbf{y} - \boldsymbol{\Sigma}_{\text{prior}}^{-1} \boldsymbol{\mu}) \\ \boldsymbol{\Sigma}_{\text{post}} &= (\boldsymbol{\Sigma}_n^{-1} \mathbf{X}^T \mathbf{X} + \boldsymbol{\Sigma}_{\text{prior}}^{-1})^{-1} \end{aligned} \quad (7)$$

But the probabilistic interpretation of Equation (6) has additional value over Equation (4) of LWL because it is a generative model for all (training and prediction) data points, which can be used to learn hyperparameters conveniently.

Once we have the training data, we want to predict the function output $f(x_*)$ at a specific test input x_* . To accomplish this using GP method, the distribution of $f_* = f(x_*)$ can be predicted

$$p(f_* | \mathbf{X}, \mathbf{y}, x_*) \sim \mathcal{GP}(f_* | x_* \mathbf{m}, x_*^T \boldsymbol{\Sigma}_{\text{post}} x_*) \quad (8)$$

Remark 1. Usually, by preprocessing the data, it is assumed that the prior distribution satisfies the mean equal to zero, here we use $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}_{\text{prior}}$ to denote the prior distribution and $\mathbf{m}$ and $\boldsymbol{\Sigma}_{\text{post}}$ to denote the posterior distribution to distinguish them.

Notation 1. Definition $k(x, x') = \phi(x) \boldsymbol{\Sigma}_{\text{prior}} \phi(x') \rightarrow k_x$, $k(x_*, x'_*) \rightarrow k_*$, $\mathbf{K}(\mathbf{X}, \mathbf{X}') \rightarrow \mathbf{K}_X$, $\mathbf{K}(x_*, \mathbf{X}) \rightarrow \mathbf{K}_*$ and so on, where $\phi(x)$ is a non-linear function. Take them into Equation (8) we have

$$\begin{aligned} x_* \mathbf{m} &\rightarrow \mathbf{K}_* (\mathbf{K}_X + \boldsymbol{\Sigma}_n)^{-1} \mathbf{y} \\ x_*^T \boldsymbol{\Sigma}_{\text{post}} x_* &\rightarrow k_{**} - \mathbf{K}_*^T (\mathbf{K}_X + \boldsymbol{\Sigma}_n) \mathbf{K}_* \end{aligned} \quad (9)$$

So, we call $k(x, x')$ the covariance function.

In the GP model, the covariance function $k(x, x')$ can measure similarity between system samples and determine the characteristics of the function $f(x)$. The most commonly used covariance function is the Squared Exponential (SE) covariance function

$$k_{SE}(x, x') = \sigma_f^2 \exp\left(-(x - x')^T l^{-1} (x - x')\right) + \sigma_n^2 \delta \quad (10)$$

Among them, l is called a relevance determination parameter, which measures the relevance of sample input and output. When $f(x)$ has a strong dependence on x , l is

smaller, and $f(x)$ is more tortuous. When the dependence of $f(x)$ and x is weak, l is larger, and $f(x)$ is smoother. σ_f is the signal standard deviation of the covariance function, which adjusts the amplitude of $f(x)$. σ_n is the noise standard deviation, δ is the Kronecker symbol.

3. Approximation Algorithm for LSGP

For ship maneuvering motion modeling, this paper has constructed and modified LWL based on sparse GP to perform effective modeling, including sparsification, localization. These core tasks are described below.

3.1. Sparsification

The sparse Gaussian process reduces the computational complexity of regression (and optimization). Quiñonero-Candela et al. [23] proposed the first unified framework to describe various methods at the time from the perspective of “approximate model, accurate inference” [24,25]. The framework of “approximate model, accurate inference” changes the assumptions made by the original model on sample data, and support points into the model in the form of parameters [26]. Bui et al. [27] also proposed a new unified framework, expounded from the perspective of “accurate model, approximate inference”. Under the new framework, the original model remains unchanged, and different approximate inference methods are used to reduce the time complexity. Using these methods, especially the FITC method [28] that we will apply in this article, the runtime can be reduced to being linear with respect to N , with only a limited reduction in how well the available data is being used, where N is the number of training points.

The FITC method regards the ‘Support points’ (or inducing points) as a virtual point, or a newly introduced parameter of the model, and optimizes the position of the induction point and other hyperparameters at the same time by means of parameter optimization. Suppose there are M virtual sample points (not interfered by observation noise), called ‘support points’, the input feature is $\mathbf{X}_s = \{x_{s,i}\}_{i=1}^M$, and the corresponding function value $\mathbf{f}_s = \{f(x_{s,i})\}_{i=1}^M$. Assuming that after these ‘support points’ are given, the prediction points are independent of the training sample conditions, i.e.,

$$p(f_* | \mathbf{y}) \approx \int p(f_* | \mathbf{f}_s) p(\mathbf{f}_s | \mathbf{y}) d\mathbf{f}_s \quad (11)$$

Under the condition of a given ‘support point’, the training sample function values are conditionally independent, i.e.,

$$p(\mathbf{f} | \mathbf{f}_s) \approx q(\mathbf{f} | \mathbf{f}_s) = \prod_i p(f_i | \mathbf{f}_s) \quad (12)$$

Approximate $q(\mathbf{y} | \mathbf{X}, \mathbf{X}_s, \mathbf{f}_s)$ of $p(\mathbf{y} | \mathbf{X}, \mathbf{X}_s, \mathbf{f}_s)$ can be calculated

$$p(\mathbf{y} | \mathbf{X}, \mathbf{X}_s, \mathbf{f}_s) \approx q(\mathbf{y} | \mathbf{X}, \mathbf{X}_s, \mathbf{f}_s) \sim \mathcal{GP}(\mathbf{K}_{NM} \mathbf{K}_M^{-1} \mathbf{f}_s, \mathbf{\Lambda} + \mathbf{\Sigma}_n) \quad (13)$$

where, $\mathbf{\Lambda} = \text{diag}(\lambda)$, $\lambda = k_i - \mathbf{K}_{M,i}^T \mathbf{K}_M^{-1} \mathbf{K}_{M,i}$, $\mathbf{K}_{NM} = \mathbf{K}(\mathbf{X}, \mathbf{X}_s)$, $\mathbf{K}_M = \mathbf{K}(\mathbf{X}_s, \mathbf{X}_s)$. For the ‘support point’ set M , we make it obey the Gaussian prior

$$p(\mathbf{f}_s | \mathbf{X}_s) \sim \mathcal{GP}(\mathbf{f}_s | \mathbf{0}, \mathbf{K}_M) \quad (14)$$

So, we can get the marginal likelihood by Equations (13) and (14)

$$p(\mathbf{y} | \mathbf{X}, \mathbf{X}_s, \Theta) = \int p(\mathbf{y} | \mathbf{X}, \mathbf{X}_s, \mathbf{f}_s) p(\mathbf{f}_s | \mathbf{X}_s) d\mathbf{f}_s \sim \mathcal{GP}(\mathbf{0}, \mathbf{K}_{MM} \mathbf{K}_M^{-1} \mathbf{K}_{NM}^T + \mathbf{\Lambda} + \mathbf{\Sigma}_n) \quad (15)$$

During training, the general method is maximizing the logarithmic marginal likelihood $L(\Theta) = \log p(\mathbf{y} | \mathbf{X}, \mathbf{X}_s, \Theta)$, the hyperparameter Θ and the set of 'support points' M are optimized through optimization algorithms.

When predicting through the GP model, first, we find the posterior distribution over 'support point' using Equations (14) and (15)

$$p(\mathbf{f}_s | \mathbf{X}_s, \mathbf{y}) \sim \mathcal{GP}(\mathbf{K}_M \mathbf{Q}_M^{-1} \mathbf{K}_{NM} (\mathbf{\Lambda} + \mathbf{\Sigma}_n)^{-1} \mathbf{y}, \mathbf{K}_M \mathbf{Q}_M^{-1} \mathbf{K}_M) \quad (16)$$

where, $\mathbf{Q}_M = \mathbf{K}_M + \mathbf{K}_{NM}^T (\mathbf{\Lambda} + \mathbf{\Sigma}_n)^{-1} \mathbf{K}_{NM}$.

Given a new input x_* , the predicted distribution is obtained by integrating Equations (14) and (16)

$$p(f_* | x_*, \mathbf{X}, \mathbf{X}_s, \mathbf{y}) = \int p(f_* | \mathbf{f}_s) p(\mathbf{f}_s | \mathbf{X}_s, \mathbf{y}) d\mathbf{f}_s \sim \mathcal{GP}(m_*, s_*^2) \quad (17)$$

The mean and covariance can be calculated as

$$\begin{aligned} m_* &= \mathbf{k}_{*M} \mathbf{Q}_M^{-1} \mathbf{K}_{NM}^T (\mathbf{\Lambda} + \mathbf{\Sigma}_n)^{-1} \mathbf{y} \\ s_* &= k_* - \mathbf{k}_{*M}^T (\mathbf{K}_M^{-1} - \mathbf{Q}_M^{-1}) \mathbf{k}_{*M} \end{aligned} \quad (18)$$

where, $\mathbf{k}_{*M} = \mathbf{k}(x_*, \mathbf{X}_s)$, $k_* = k(x_*, x_*)$.

3.2. Localization

Based on the idea of local learning, the similarity between the query point q and the local model is exploited to effectively cluster the system input data. Assumption: nearby query points may have similar target values. The local model is updated in a Gaussian process, as shown in Figure 1.

Figure 1. Structure of LSGP identification model.

In this study, the prediction of the system output $\mathbf{y}_l = \{y_{i,l}\}_{i=1}^{N_l}$ in local model $\{l_i\}_{i=1}^L$ and their interpretation are carried out through an easy-to-interpret linear model, i.e., system output y_i for the i th sample is assumed to be obtained using linear system in local model

$$f_l(\mathbf{X}_l) \stackrel{w}{\leftarrow} \mathbf{X}_l^T \boldsymbol{\theta}_l + \varepsilon \quad (19)$$

where $\mathbf{X}_l = \{X_{i,l}\}_{i=1}^{N_l}$ is the system input in local model l , the number N_l is determined by w ; $\boldsymbol{\theta}_l = \{\theta_{l,d}\}_{d=1}^D$, $\boldsymbol{\theta}_d = \{\theta_i\}_{i=1}^{N_l \in l}$ is a D -dimensional weight vector for the l th local model; and ε is a Gaussian noise. Estimating $\boldsymbol{\theta}_l$ without any constraints is an ill-posed problem, we assume that functions determining $\boldsymbol{\theta}_l$ are generated from a multivariate GP. More specifically, the weight vector for the i th sample for the D -dimensional in l th local model is obtained as follows

$$\theta_i = \mathbf{g}(\mathbf{x}_i) + \varepsilon_\theta \quad (20)$$

where, ε_θ is a Gaussian noise. Here, vector-valued function $\mathbf{g}(\bullet)$ that determines the regression parameters vector for each sample, and each element of $\mathbf{g}(\bullet)$ is generated from a univariate GP independently:

$$\begin{aligned} \mathbf{g}(\mathbf{x}) &= [g_1(x), g_2(x), \dots, g_D(x)] \\ g_d(x) &\sim \mathcal{GP}(m(x), k(x, x')) \end{aligned} \quad (21)$$

$m(x)$ is the mean function, and $k(x, x')$ is the covariance function.

Use the kernel function to learn local models, the similarity measure will be $w_i = \text{Kernel}(d(x_i, q), h)$, where $d(\bullet)$ represents the distance between the query point q and the training point $\{x_i\}_{i=1}^N$, and h represents the hyperparameter, which is called the distance measure. It should be emphasized that in order to better perform local domain partitioning, any allowable kernel function can be used. Thus, the kernel can be varied in many ways. Here, the Gaussian kernel is used to cluster the local model

$$w_i = \exp\left(-\frac{1}{2}(x_i - q)^T h (x_i - q)\right) \quad (22)$$

After sample data is partitioned, the kernel matrix is updated for each local model. The regression and prediction under these local models is the weighted average of the predicted mean y_k of each local model for a query point. The optimization of the hyperparameters h of the kernel function is usually performed on sub-samples of the full training data. Here, Algorithm 1 gives the relevant process of the LSGP method.

Algorithm 1 Localization and model learning of LSGP.

Input: $\mathbf{X} \in \mathbb{R}$: input sample points; $\mathbf{y} \in \mathbb{R}$: output sample points; q : query points; h_0 : initial distance measure; Θ_0 : initial covariance parameter; M : initial support point set;

Output: $\mathcal{GP}_*$: generative model; h_* , M_* , Θ_* : related parameters;

- 1: Calculate the distance d between the query point q and the sample point $\mathbf{X}$ in the training set, such as Euclidean distance, Mahalanobis distance, etc.;
 - 2: Use Equation (22) to cluster data to obtain the measure (or weight) of the local model corresponding to the query point q ;
 - 3: The training data sample points are weighted to construct a local domain centered on the query point q ;
 - 4: In the local domain, through Equations (10), (17), (18) obtain the covariance matrix and the Gaussian posterior in the Gaussian process, i.e., the construction of the local model;
 - 5: Calculate the marginal likelihood by Equation (15), with logarithmic marginal likelihood $L(\bullet) = \log p(\mathbf{y} | \mathbf{X}, \mathbf{X}_s, \Theta)$ as the loss function. Use the conjugate gradient to maximize the loss function to optimize the relevant parameters, $\Theta, h, M = \{\mathbf{X}_s, \mathbf{f}_s\}$;
 - 6: Then, output the optimal relevant parameters, i.e., the distance measure of the local model h_* , the support points $M_* = \{\mathbf{X}_s, \mathbf{f}_s\}$, and the hyperparameters Θ_* of the Gaussian process to obtain the optimal local model;
 - 7: Obtain the generative model of ship maneuvering motion from optimal local model, and perform simulation prediction through the generative model.
-

Notation 2. When building a local model, calculating the Gaussian posterior and marginal likelihood of the model requires the calculation of the covariance matrix and its inverse. For the ship motion system, it is easy to produce singular data when performing steady manoeuvring. Therefore, LU-decomposition is used to process the covariance matrix to reduce calculation errors. For solving the inverse covariance matrix, Cholesky-decomposition can be used, i.e., $\mathbf{C} = \text{cholesky}(\mathbf{K})$, to improve calculation efficiency.

4. LSGP for Identifying the Ship Maneuvering Motion System Model

In this section, the relevant content of ship motion identification modeling based on the LSGP method is explained, and secondly, explanatory learning examples will be provided to verify the performance of the proposed scheme. In order to achieve this goal, taking the plane motion model of the ‘Mariner’ class vessel as the research object. In addition, two LWL modeling schemes are proposed, including Global LWL [9] and Local LWL [22] for comparison. These simulations are all carried out on a computer with 3.2 GHz CPU and 8 GB RAM.

4.1. Ship Dynamics Overview of Ship Maneuvering Motion

Ships sailing at sea can be approximately regarded as a rigid body. In order to better describe the ship’s motion state, a space-fixed coordinate system $O - x_0y_0z_0$ and a body-fixed coordinate system $G - xyz$ are introduced, as shown in Figure 2. It can be seen that the actual motion of a ship is very complex, and it has 6-DOF generally. For most ship motions and their control problems, heave motion, pitch motion and roll motion can be neglected, only surge motion, sway motion and yaw motion can be discussed, which can satisfy enough engineering requirements. Therefore, simplifies the ship motion into a 3-DOF plane motion. At the same time, in the ship coordinate system it is more simple and straightforward to study the movement of the ship in the surrounding flow field and the interaction between fluid and hull.

Figure 2. Describing Ship Motion in Coordinate System.

According to the characteristics of the ship’s motion and the effects of the forces on the hull, the forces acting on the ship can be divided into three categories [29]: (1) the main power F_c , the force generated by the ship’s power system, including the propeller thrust and the rudder force. (2) Environmental interference force F_i , i.e., the force acting

on the ship by external factors such as wind, waves and current. (3) Hydrodynamic force F_h , i.e., the movement of the ship in the fluid under the action of the main power and the environmental inference force, and the reaction force generated by the fluid on the movement of the ship, including the fluid inertial force and the fluid viscous force etc. Therefore, the composition of the force and moment of the ship can be expressed as

$$\begin{cases} F = F_c + F_i + F_h \\ M = M_c + M_i + M_h \end{cases} \quad (23)$$

According to the rigid body momentum theorem, the momentum theorem of rigid body centroid motion and Newton's law of motion, ship motion can be expressed as follows

$$\begin{cases} m(\dot{u} - vr - x_G r^2) = F_X \\ m(\dot{v} + ur + x_G \dot{r}) = F_Y \\ I_z \dot{r} + m x_G (\dot{v} + ur) = M_N \end{cases} \quad (24)$$

where, m is the mass of ship; x_G is the longitudinal coordinate of the ship gravity center; I_z is the yaw moment of inertia about z axis; $\dot{u}$, $\dot{v}$ and $\dot{r}$ denote the surge acceleration, sway acceleration and yaw angular acceleration corresponds to u the surge velocity, v the sway velocity, r the yaw angular velocity; F_X , F_Y and M_N is the projection of the total forces and moments F , M in different directions.

Generally, in the study of ship motion, the motion law of ship's gravity center is mainly considered. Therefore, the origin of the body-fixed coordinate system can be taken at the center of gravity, i.e., $x_G = 0$. In addition, the forces and moments in ship motion are related to the characteristics of the ship motion (velocity, acceleration etc.), the hull characteristics and the maneuvering factors (rudder angle δ , propeller speed n , etc.). The Equation (24) can be written as follows

$$\begin{cases} m(\dot{u} - vr) = f_X(u, v, r, \delta, n) \\ m(\dot{v} + ur) = f_Y(u, v, r, \delta, n) \\ I_z \dot{r} = f_N(u, v, r, \delta, n) \end{cases} \quad (25)$$

The main task of identification modeling is to find a suitable and accurate f_X , f_Y and f_N to better describe the ship maneuvering motion system.

4.2. System Identification Modeling Experiment

Although it is also possible to use the strategy of directly considering the relationship between ship's motion states [2,5,9,10] in non-parametric identification modeling (i.e., to study the relationship between ship motion $u(k)$, $v(k)$, $r(k)$ at k th time and $u(k+1)$, $v(k+1)$, $r(k+1)$ at $(k+1)$ th time), here, we adopt a strategy that considers the relationship between the ship's motion states and the ship's external forces, (i.e., we care about the relationship between the ship's motion $u(k)$, $v(k)$, $r(k)$ at k th time and f_X , f_Y , f_N , at k th time) as shown in Figure 3, the relevant content in the green dashed box. This way, getting rid of the discrete processing mode of the ship motion system in the time domain is a continuous-time system mode. At the same time, during model training and simulation prediction, the LSGP-based modeling method has matrix operations, so in addition to the LU-decomposition and Cholesky-decomposition processing of the matrix mentioned in the third section above, the correlation matrix operations are transplanted to the C++ language Environment, as shown in Figure 3, to improve computing efficiency.

Figure 3. Structural modules for identification and modeling of ship maneuvering motion based on LSGP.

Data is the first element of recognition, and the quality of the data will directly affect the recognition result. For the LWL method, it is extremely important for data collection and processing. Therefore, in references [5,9,10], several ship maneuvering tests (8-shaped test, sinusoidal signal) are designed and processed for this purpose to make the collected data as much as possible. Information, most of the design tests are more complicated and are not suitable for actual ship maneuvering. Here, we only use the zigzag test (see Table 1) mentioned in the IMO MSC.137 (76) and ITTC Guidelines (2017) for data collection and model identification learning.

The well-known 'Mariner' class vessel is taken to establish the experimental maneuvering motion model [2,9]. Simulation of 25°/25° zigzag test, which is conducted with initial sailing speed of 15 knots (7.72 m/s), sampling time is 700 s and the interval is 1 s, i.e., 700 measurement pairs of u , v , r , δ , as the training set for identification of ship hydrodynamic model. In this paper, the propeller rotation change is not considered. The rudder angle is the system input regarded as the system excitation. Therefore, the rudder angle is a known variable. In order to verify the generalization and prediction capabilities of the proposed LSGP-based ship motion model, we considered a convincing maneuvering test as the inspection plan. For the sake of simplicity, the design speed $u_0 = 7.2$ m/s (15 knots) is the initial speed, and two typical zigzag maneuvers expressed by 'ZZ- θ (heading angle)- θ (rudder angle)', namely ZZ-20°-10° and ZZ-30°-5°; use turning test 'T- θ (rudder angle)' to verify the ship's turning test performance, and give experiments on T-25°; at the same time, the simulation experiments and the reference model were compared. The simulation experiment situation is shown in Table 1. It is pointed out here that the 'reference model' used in the article is based on the parameter model obtained through the PMM test, and the parameters and architecture of the reference model can be found in [2,29].

The initial parameter values about LSGP are generated randomly during training ship motion model, where, $h_0 \in \text{rand}(0, 2]$, $\Theta_0 \in \log(0, 2)$, The 'support point' is a random sample of $M = 200$ in the training set, the training step is set to 1000. The marginal likelihood is maximized by conjugate gradient to learn the best generative model and parameters.

Table 1. Training and simulation experiment settings.

Data Set	Groups	Maneuvering Test	Number of Samples
Training	No.1	ZZ-25°-25°	700
Test	No.1	ZZ-20°-10°	700
	No.2	ZZ-30°-5°	700
	No.3	T-25°	700

4.3. Simulation Results and Analysis

In order to better explain the pros and cons of the LSQP algorithm, we demonstrate how the proposed algorithm can be executed on full-scale experimental data for different manipulation scenarios. The proposed ship motion model based on LSQP can better approximate the original ship model. As shown in Figure 3, since the corresponding system input and output are different from the previous identification modeling, the relationship between the ship's motion state and the external force on the ship is considered. Figures 4–6 shows the results of three sets of simulation experiments in Table 1. It can be seen that the non-parametric model under LSQP method has good generalization, except for deviation of the Yaw moment N of the simulation prediction. This is because the ship Yaw moment is relatively complex, has strong nonlinearity and is difficult to decouple from the motion state in other DOF.

In the zigzag test, the nonlinearity of the ship motion state is more serious. The learning error of LSQP in Figures 4 and 5 is mainly manifested in the Yaw moment. This is because the Yaw moment is more complicated than other DOF, and its value is not in the same order of magnitude as other degrees of freedom. The prediction of the ship's turning test shows good performance in Figure 6, which avoids the divergence due to the singular value generated when the data is close to zero by the ship's steady motion.

Figure 4. Comparison results of non-dimensional with the simulated ZZ-20°-10° test. (a) is about the Surge force of the “mariner”; (b) is about the Sway force of the ship; (c) is the situation of Yaw moment of the ship; the figure shows one and two times the standard deviation bandwidth respectively.

Figure 5. Comparison results of non-dimensional with the simulated ZZ-30°-5° test. (a) is about the Surge force of the “mariner”; (b) is about the Sway force of the ship; (c) is the situation of Yaw moment of the ship; the figure shows one and two times the standard deviation bandwidth respectively.

Figure 6. Comparison results of non-dimensional with the simulated T-25° test. (a) is about the Surge force of the “mariner”; (b) is about the Sway force of the ship; (c) is the situation of Yaw moment of the ship; the figure shows one and two times the standard deviation bandwidth respectively.

Notation 3. Due to the large magnitude difference between the various physical quantities of ship motion, which are easily affected by ship size, speed, fluid physical parameters, etc., they have been processed in a dimensionless manner. The various physical quantities in Figures 4–6 are in dimensionless form.

From the above simulation results that the LSGP algorithm has a certain prediction simulation effects, and low requirements for training data. This property is determined by the nature of the LSGP algorithm. It is the main line of this paper to predict the external force of the ship through identification modeling, and to further calculate the trend of the ship’s motion at the next moment according to the Equation (25) in real time, and then to predict the state and trajectory of the ship’s motion. The simulation prediction results of the mariner test are shown in Figures 7–9, and a comparison based on the two LWL (global LWL [9], local LWL [22]) algorithms is also given.

Figure 7. Comparison results of ship motion states for the simulated ZZ-20°-10° test. It shows the changes of the ship's motion state on the three degrees of freedom, (a) is the Surge Velocity, (b) is the Sway Velocity, and (c) is the Yaw rate. (d) Scatter plot of simulation prediction and reference model, the abscissa is simulation prediction, and the ordinate is reference model in each sub-figure.

Figure 8. Cont.

Figure 8. Comparison results of ship motion states for the simulated ZZ-30°-5° test. It shows the changes of the ship's motion state on the three degrees of freedom, (a) is the Surge Velocity, (b) is the Sway Velocity, and (c) is the Yaw rate. (d) Scatter plot of simulation prediction and reference model, the abscissa is simulation prediction, and the ordinate is reference model in each sub-figure.

Figure 9. Comparison results of ship motion states for the simulated T-25° test. It shows the changes of the ship's motion state on the three degrees of freedom, (a) is the Surge Velocity, (b) is the Sway Velocity, and (c) is the Yaw rate. (d) Scatter plot of simulation prediction and reference model, the abscissa is simulation prediction, and the ordinate is reference model in each sub-figure.

As can be seen from Figure 7, the simulation of LWL does not work effectively. Compared to the Gaussian process, LWL cannot effectively build a generative model. For tests like ZZ-20°-10°, it is not included in the features of the LWL training set, so simulation predictions cannot be made well. It directly shows that LWL has high requirements for training data and is not easy to use in applications. On the contrary, LSGP has good generalization ability. In order to quantitatively describe the differences between different methods, the following Table 2 gives mean squared error (MSE) evaluation criteria for comparison

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (26)$$

where, y_i represents the reference model, $\hat{y}_i$ represents the simulation prediction, and n represents the data length.

Table 2. Comparison of the prediction accuracy for different methods.

Maneuvering Test		Global LWL	Local LWL	LSGP
ZZ-20-10	u	1.0743×10^{-2}	1.0754×10^{-2}	6.6611×10^{-4}
	v	1.7575×10^{-2}	1.4977×10^{-2}	5.1163×10^{-3}
	r	3.0272×10^{-2}	1.6836×10^{-2}	8.8154×10^{-3}
ZZ-30-5	u	1.6510×10^{-2}	1.5225×10^{-2}	1.2690×10^{-3}
	v	2.9816×10^{-2}	1.0276×10^{-1}	3.5295×10^{-2}
	r	6.8430×10^{-2}	6.5596×10^{-2}	5.1747×10^{-2}
T-25	u	8.8399×10^{-5}	2.3083×10^{-4}	1.0963×10^{-3}
	v	6.5412×10^{-4}	4.9372×10^{-4}	7.0614×10^{-5}
	r	1.1510×10^{-3}	1.0541×10^{-3}	3.8243×10^{-3}

It can be seen intuitively from Table 2 that LSGP has advantages over other methods in the zigzag test, generally, Global LWL < local LWL < LSGP, and the simulation accuracy in the turning test needs to be improved, which is related to the fact that our training data does not include turning data. At the same time, three algorithms are given about the time spent in training and simulation prediction, as shown in Figure 10.

Figure 10. Compare of time spent by different methods. The training part (filled with diagonal lines) refers to the left vertical axis, the variable unit is hours, especially the time unit of LSGP is minutes; the test part refers to the right vertical axis, the unit is second.

There are two methods for the GP to calculate the inverse matrix, one is to directly use the traditional Gaussian elimination method, and the other is matrix decomposition. When the dimension of the matrix is too large, the inverse calculation of the matrix by

Gaussian elimination method needs to be calculated in blocks. On the other hand, for the ill-conditioned matrix generated in the ship steady motion, coupled with the influence of the computer truncation error, the direct calculation using the Gaussian elimination method cannot obtain accurate results, and the time complexity is $O(N^3)$. We use Cholesky decomposition, i.e., $L = \text{cholesky}(K_{NM})$, its time complexity is $O(M^2N/6)$. Usually the dimension of input space is limited, so the transition matrix is defined on the auxiliary platform to store the distance between input data, that is, the elements of the covariance matrix: $\Gamma_k(i, j) = \left(x_i^{(k)} - x_j^{(k)}\right)^2$. x_i^k represents the k th dimension element of the i th dimension input of the data. These Γ matrices can be used to reduce the calculation amount, its time complexity is $O(M)$.

The time spent on training and simulation prediction for the LSGP algorithm decreases exponentially in Figure 10. In training part, LSGP is reduced by 100 times compared with Local LWL, and it is also reduced by 10 times in simulation prediction. At the same time, Table 3 gives different training data method requirements, among them, LSGP has very low requirements for training data. It only needs a set of zigzag tests as training data, which is also easy in engineering practice. For the 'reference model' based on mechanism modeling, LSGP can make the computational time consistent with the reference model while ensuring the modeling accuracy. This advantage is self-evident. On the one hand, LSGP is a data-driven model, which can provide considerable results based on limited data without the need to use physical tests, empirical formulas and CFD methods to determine model parameters, thus reducing operational difficulty and economic investment. On the other hand, non-parametric modeling can avoid the parameter drift, which is not limited by the model structure. It is a feasible method when the ship principal particulars are not easy to obtain. At the same time, decreasing the calculation time of non-parametric modeling to be consistent with that of the reference model can provide preparations for future online identification modeling.

Table 3. Data requirements of different methods.

Methods	Global LWL	Local LWL	LSGP
No. of data	15, 240	4835	700
Data requirement	8-shaped; zigzag	8-shaped; zigzag	zigzag

5. Conclusions and Prospects

This paper introduces the development of a new local weighting scheme based on sparse Gaussian process, which is used for ship maneuvering identification modeling in a full-scale ocean environment. The main advantages of this scheme are low requirements for training data, high computational efficiency, and no need to know the structure of the ship maneuvering model, which is easy for engineering practice. As far as the author knows, the existing research work rarely considers the choice of model structure, which is meaningful and inevitable in actual ship engineering. Application experiments show the engineering practicability, effectiveness and generalization of the scheme. In addition, the program can also be extended to other online identification or prediction systems in the field of ocean engineering.

Further work will focus on ship maneuvering identification modeling under the dual-rate measurement engineering environment considering the speed of the ship's main engine, and the influence of factors such as wind and waves should also be considered. At the same time, the ship maneuvering motion modeling method based on the LSGP is not perfect and limited to the Gaussian noise distribution assumption. So, the next step is to consider breaking the Gaussian noise distribution assumption and considering the influence of time factors, as well as to research the relationship between localization and sparseness in LSGP to improve ship motion modeling accuracy as much as possible.

Author Contributions: Conceptualization, Z.Z. and J.R.; methodology, Z.Z. and J.R.; software, Z.Z.; validation, Z.Z.; formal analysis, Z.Z.; investigation, J.R. and Z.Z.; resources, Z.Z.; data curation, Z.Z.; writing—original draft preparation, Z.Z.; writing—review and editing, Z.Z.; visualization, J.R. and Z.Z.; supervision, J.R.; project administration, J.R.; funding acquisition, J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 51779029 and 51679024 ; the Fundamental Research Funds for the Central Universities, grant number 3132019312 and Ministry of Industry and Information Technology Letter [2018] No. 473.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank Xiufeng Zhang for her comments that improved the manuscript, Weiwei Bai for his assistance with the implementation of ship motion modeling with LWL and for comments, Lu Bai for her Proofreading the manuscript, the anonymous reviewers for their constructive suggestions which comprehensively improve the quality of the paper.

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

Abbreviations

The following abbreviations are used in this manuscript:

MMG	Manoeuvring Mathematical Model Group
LWL	Locally Weighted Learning
DOF	Degree of freedom
SVM	Support Vector Machines
ANN	Artificial Neural Networks
FITC	Fully Independent Training Conditional Approximation
Local LWL	Locally optimal-based Locally Weighted Learning
Global LWL	Global optimal-based Locally Weighted Learning
PMM	planar motion mechanism

References

1. Liu, Y.; Li, B.; Zou, Z.J. Manoeuvrability Prediction for Container Ships in Deep and Shallow Waters. *Chuan Bo Li Xue/J. Ship Mech.* **2019**, *23*, 267–282.
2. Wang, X.G.; Zou, Z.J.; Hou, X.R.; Xu, F. System identification modelling of ship manoeuvring motion based on ϵ —Support vector regression. *J. Hydrodyn.* **2015**, *27*, 502–512. [\[CrossRef\]](#)
3. Wang, Z.; Xu, H.; Xia, L.; Zou, Z.; Soares, C.G. Kernel-based support vector regression for nonparametric modeling of ship maneuvering motion. *Ocean Eng.* **2020**, *216*, 107994. [\[CrossRef\]](#)
4. Luo, W.; Zhang, Z. Modeling of ship maneuvering motion using neural networks. *J. Mar. Sci. Appl.* **2016**, *15*, 426–432. [\[CrossRef\]](#)
5. Wang, N.; Er, M.J.; Han, M. Large Tanker Motion Model Identification Using Generalized Ellipsoidal Basis Function-Based Fuzzy Neural Networks. *IEEE Trans. Cybern.* **2015**, *45*, 2732–2743. [\[CrossRef\]](#)
6. Zhao, X.; Wang, X.; Tang, H. Modeling and prediction of disturbing force and moment based on wavelet neural network optimized by GA. *J. Ship Mech.* **2008**, *12*, 25–30.
7. Zhang, Z.; Ren, J.; Wang, G. Multi-dimensional local weighted regression ship motion identification modeling based on particle swarm optimization. In Proceedings of the 38th Chinese Control Conference, Guangzhou, China, 27–30 July 2019; pp. 1520–1525.
8. Woo, J.; Park, J.; Yu, C.; Kim, N. Dynamic model identification of unmanned surface vehicles using deep learning network. *Appl. Ocean Res.* **2018**, *78*, 123–133. [\[CrossRef\]](#)
9. Bai, W.; Ren, J.; Li, T.; Zhang, X. Global-Optimal-based Locally Weighted Learning for Ship Maneuvering Motion Identification. *China Navig.* **2017**, *40*, 37–41.
10. Bai, W.; Ren, J.; Li, T. Modified genetic optimization-based locally weighted learning identification modeling of ship maneuvering with full scale trial. *Future Gener. Comput. Syst.* **2019**, *93*, 1036–1045. [\[CrossRef\]](#)
11. Sigaud, O.; Salan, C.; Padois, V. On-line regression algorithms for learning mechanical models of robots: A survey. *Rob. Auton. Syst.* **2011**, *59*, 1115–1129. [\[CrossRef\]](#)
12. Sutulo, S.; Guedes Soares, C. On the application of empiric methods for prediction of ship manoeuvring properties and associated uncertainties. *Ocean Eng.* **2019**, *186*, 106111. [\[CrossRef\]](#)
13. Rasmussen, C.E.; Williams, C.K. *Gaussian Processes for Machine Learning*; The MIT Press: Cambridge, MA, USA, 2006.

14. Atkinson, S.; Zabarar, N. Structured Bayesian Gaussian process latent variable model: Applications to data-driven dimensionality reduction and high-dimensional inversion. *J. Comput. Phys.* **2019**, *383*, 166–195. [[CrossRef](#)]
15. Quann, M.; Ojeda, L.; Smith, W.; Rizzo, D.; Castanier, M.; Barton, K. Power Prediction for Heterogeneous Ground Robots through Spatial Mapping and Sharing of Terrain Data. *IEEE Robot. Autom. Lett.* **2020**, *5*, 1579–1586. [[CrossRef](#)]
16. Wilcox, B.; Yip, M.C. SOLAR-GP: Sparse Online Locally Adaptive Regression Using Gaussian Processes for Bayesian Robot Model Learning and Control. *IEEE Robot. Autom. Lett.* **2020**, *5*, 2832–2839. [[CrossRef](#)]
17. Sandzimir, R.J.; Asada, H.H. A Data-Driven Approach to Prediction and Optimal Bucket-Filling Control for Autonomous Excavators. *IEEE Robot. Autom. Lett.* **2020**, *5*, 2682–2689. [[CrossRef](#)]
18. Shoukat, H.; Tahir, M.; Ali, K. Approximate GP Inference for Nonlinear Dynamical System Identification Using Data-Driven Basis Set. *IEEE Access* **2020**, *8*, 90665–90675. [[CrossRef](#)]
19. Cheng, L.F.; Dumitrascu, B.; Darnell, G.; Chivers, C.; Draugelis, M.E.; Li, K.; Engelhardt, B.E. Sparse multi-output Gaussian processes for medical time series prediction. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 152–175. [[CrossRef](#)]
20. Aftab, W.; Mihaylova, L. A Learning Gaussian Process Approach for Maneuvering Target Tracking and Smoothing. *IEEE Trans. Aerosp. Electron. Syst.* **2021**, *57*, 278–292. [[CrossRef](#)]
21. Liu, Y.; Xue, Y.; Huang, S.; Xue, G.; Jing, Q. Dynamic model identification of ships and wave energy converters based on semi-conjugate linear regression and noisy input gaussian process. *J. Mar. Sci. Eng.* **2021**, *9*, 194. [[CrossRef](#)]
22. Bai, W.; Ren, J.; Li, T.; Chen, C.L.P. Grid index subspace constructed locally weighted learning identification modeling for high dimensional ship maneuvering system. *ISA Trans.* **2019**, *86*, 144–152. [[CrossRef](#)] [[PubMed](#)]
23. Quiñonero-Candela, J.; Rasmussen, C.E. A unifying view of sparse approximate Gaussian process regression. *J. Mach. Learn. Res.* **2005**, *6*, 1939–1959.
24. Mishra, A.; Dey, D.K.; Chen, Y.; Chen, K. Generalized co-sparse factor regression. *Comput. Stat. Data Anal.* **2021**, *157*, 107127. [[CrossRef](#)]
25. Khan, M.; Patel, A.; Chatterjee, A. Multi-sparse gaussian process: Learning based semi-parametric control. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25–29 October 2020; pp. 5327–5334.
26. Yang, L.; Wang, K.; Mihaylova, L. Online Sparse Multi-Output Gaussian Process Regression and Learning. *IEEE Trans. Signal Inf. Process. Netw.* **2019**, *5*, 258–272. [[CrossRef](#)]
27. Bui, T.D.; Yan, J.; Turner, R.E. A unifying framework for Gaussian process pseudo-point approximations using power expectation propagation. *J. Mach. Learn. Res.* **2017**, *18*, 3649–3720.
28. Bauer, M.; Van Der Wilk, M.; Rasmussen, C.E. Understanding probabilistic sparse Gaussian Process approximations. In Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain, 5–10 December 2016; pp. 1533–1541.
29. Fossen, T.I. *Handbook of Marine Craft Hydrodynamics and Motion Control*; John Wiley and Sons: New York, NY, USA, 2011.