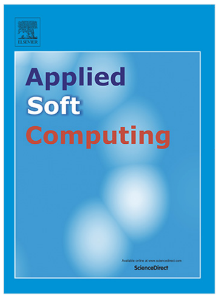
Accepted Manuscript



Modelling of a surface marine vehicle with kernel ridge regression confidence machine

David Moreno-Salinas, Raul Moreno, Augusto Pereira, Joaquin Aranda, Jesus M. de la Cruz

PII: S1568-4946(18)30684-7
DOI: https://doi.org/10.1016/j.asoc.2018.12.002
Reference: ASOC 5228

To appear in: *Applied Soft Computing Journal*

Received date : 14 November 2017
Revised date : 18 October 2018
Accepted date : 4 December 2018

Please cite this article as: D. Moreno-Salinas, R. Moreno, A. Pereira et al., Modelling of a surface marine vehicle with kernel ridge regression confidence machine, *Applied Soft Computing Journal* (2018), https://doi.org/10.1016/j.asoc.2018.12.002

Highlights

HIGHLIGHTS

A new black box identification method based on Conformal Predictors is proposed for modelling marine vehicles.

A continuous-time mathematical model is trained and tested with data obtained from real experiments with a ship developing some classical manoeuvres used for system identification on marine vehicles.

Kernel Ridge Regression and Kernel Ridge Regression Confidence Machine are the machine learning techniques used for the model computation.

A confidence margin is proposed to ensure that the model behaviour is within some limits where the real behaviour of the vehicle should lie.

Modelling of a Surface Marine Vehicle with Kernel Ridge Regression Confidence Machine

David Moreno-Salinas dmoreno@dia.uned.es
Department of Computer Science and Automatic Control
University of Distance Learning Education, UNED
Madrid

Raul Moreno raul.salinas@insight-centre.org
Insight Centre for Data Analytics
University College Dublin
Dublin

Augusto Pereira augusto.pereira@ciemat.es
National Fusion Laboratory
CIEMAT
Madrid

Joaquin Aranda jaranda@dia.uned.es
Department of Computer Science and Automatic Control
University of Distance Learning Education, UNED
Madrid

Jesus M. de la Cruz †
Department of Computers Architecture and Automatic Control. Faculty of Physics
Complutense University of Madrid
Madrid

Editor: XXXX

Abstract

This paper describes the use of Kernel Ridge Regression (KRR) and Kernel Ridge Regression Confidence Machine (KRRCM) for black box identification of a surface marine vehicle. Data for training and test have been obtained from several manoeuvres typically used for marine system identification. Thus, a 20/20 degrees Zig-Zag, a 10/10 degrees Zig-Zag, and different evolution circles have been employed for the computation and validation of the model. Results show that the application of conformal prediction provides an accurate model that reproduces with large accuracy the actual behaviour of the ship with confidence margins that ensure that the model response is within these margins, making it a suitable tool for system identification.

Keywords: System Identification, Marine Systems, Kernel Ridge Regression (KRR), Conformal Predictors (CP), Kernel Ridge Regression Confidence Machine (KRRCM)

. †Prof. Jesus Manuel de la Cruz passed away on October 11, 2017.

1. Introduction

System identification, commonly known as *surrogate modelling* in industrial design, is one of the most important steps of the engineering process, where important efforts are driven to design and construct reliable mathematical models for a wide range of applications, Forrester et. al. (2008). In this sense, there is also a great interest in developing tools for practitioners that provide multiple modelling methods to construct and use surrogate models in an easy and practical manner, see for example Belyaev et. al. (2016), where a detailed review on surrogate modelling and optimization methods is given and a tool, GPApprox (Generic Tool for Approximation), with multiple approximation algorithms is described.

In marine systems, the computation of accurate models is of utmost importance due to the costly (in time and money) operations at sea, which usually need an important group of people and infrastructures to be involved, apart from the multiple technical problems that could arise in any experimental task. In this sense, the availability of mathematical models would allow to test in simulation a number of theoretical developments, from control systems to path planning or trajectory tracking strategies, before testing the real system at sea. In the knowledge that the behavior of the model in simulation is correct, the experimental part could be approached with additional confidence.

There is a plethora of identification techniques to compute mathematical models of physical systems; for a short survey on some essential features in the identification area and a classification of methods, the reader is referred to Ljung (1999) and Ljung (2006). Concerning marine vehicles, there exists a large number of works that deal with different algorithms and propose models for different kinds of marine vehicles. Some of these models are widely used in practical applications for simulation of surface marine vehicles such as the Nomoto models, the Abkowitz model, or the Blanke model, to name but a few. There are different methods to estimate these models, for instance, in Källström & Åström (1981) the autopilots of different surface vehicles are designed with several parametric identification algorithms, in Abkowitz (1980) a Kalman Filter (KF) is used to estimate the hydrodynamic characteristics of a ship, in Fossen et al. (1996) ship dynamics identification for dynamic positioning is made with an Extended Kalman Filter (EKF), or in Velasco et al. (2013), where the authors obtain a linear second order Nomoto heading model with an added AutoRegressive Moving Average (ARMA) disturbance model for an autonomous in-scale physical model of a fast-ferry. The computation of a reliable model usually needs an important amount of experimental data to characterize the hydrodynamics of the vehicle in addition to an important computation effort. In this sense, the identification of an accurate model may be a very complex task. For some other interesting related works the reader is referred to Caccia et al. (2008), Perez et al. (2006), Fossen (2011), De la Cruz et al. (2012), and the references therein.

Among the different techniques that can be found in the literature, those that come from the Artificial Intelligence (AI) field have demonstrated to be very promising, showing good results in several scenarios and problems in identification and control. On the one hand, neural networks have shown to be robust and effective in multiple problems since they are universal function estimators, see for example Narendra & Parthasarathy (1990), where the multilayer perceptron (MLP) is used for system identification, Pan & Yu (2016) where neural networks are used to implement a biomimetic hybrid feedback feedforward learning control based on the human learning control for non-linear systems, or Rubio (2017) in which a modified Kalman filter is used for the adaptation of a neural network and the algorithm applied for the identification of chaotic systems. However, they also show some inherent drawbacks, since they usually require a high number of parameters, and the function approximation is computed by an iterative non-linear optimization process usually implemented by back-propagation. This process is usually slow and the convergence to the optimal solution is not guaranteed due to the problem of multiple local minima, Boyd & Vandenberghe (2004). Moreover, neural networks may fall on the so-called *curse of dimensionality* that leads to a large complexity on the model structure and the need of a large amount of data for the system description. On the other hand, Kernel methods are based on the Statistical Learning Theory of Vapnik (1995), such as Support Vector Machines (SVM), Gaussian Processes (GP) or Kernel Ridge Regression (KRR). The basic idea of kernel methods is to map input data into a high dimensional feature Hilbert space using a non-linear mapping technique, i.e., the kernel dot product trick, Aizerman et al. (1964), and to carry out linear classification or regression in feature space. The Kernel functions replace a possibly very high dimensional Hilbert space without explicitly increasing the feature space, Schölkopf & Smola (2002). Kernel methods have the ability to simultaneously minimize the estimation error in the training data (the empirical risk) and the model complexity (the structural risk), and a unique global solution can be found by solving the resulting convex optimization problem.

The main difference between Kernel Ridge Regression (KRR) and Support Vector Regression (SVR) is that they employ different loss functions (ridge versus epsilon-insensitive loss). In contrast to SVR, fitting a KRR can be done in closed-form and is typically faster for medium-sized datasets. On the other hand, the learned model is non-sparse and thus slower than SVR at prediction-time. GP models and KRR models provide a similar mean result in the prediction, with a faster convergence for KRR for medium-sized training sets. The hyperparameters on GP are computed using a gradient-ascent algorithm while for KRR a grid-search is needed. The grid-search for hyperparameter optimization scales exponentially with the number of hyperparameters, so the parameter selection on GP may be faster as it does not suffer an exponential scaling. Besides, GP provides a posterior distribution for the prediction. This prediction is considered to be Gaussian, with a mean and a variance. This variance can be seen as a confidence level of the predicted mean. However, the disadvantages of GP are the loss of efficiency in high dimensional spaces and the computational complexity for the inverse of a high dimensional covariance matrix during the training. See Kocijan (2016) and Rasmussen & Williams (2006), and the references therein for more detailed information. KRR only provides predictions, for this reason in this work we incorporate Conformal Predictors to KKR so as to provide confidence margins in a similar way as GP. In addition, GPs are a Bayesian approach that provides optimal solutions when the prior knowledge is known to be correct but as in real problems it is unknown, it becomes necessary to make a priori assumptions that reflect the real distribution. Another important feature in the present work is that the experiments must usually be carefully chosen to provide an accurate model, while we propose a mathematical model construction with large generalization performance from a small amount of data and a limited input data range. The details are given in Section 5.

We can find some interesting works using SVM for system identification as, for example, the work in Drezet & Harrison (1998), where the authors study the possible use of SVM for system identification, in Adachi & Ogawa (2001), where an identification method based on SVR is proposed for linear regression models, in Jemwa & Aldrich (2003), in which the application of SVM to time series modelling is considered by means of simulated data from an autocatalytic reactor, or Wang & Ye (2004), where Least Squares Support Vector Machines (LS-SVM) is used for non-linear system identification for some simple examples of Non-linear AutoRegressive with eXogenous input (NARX) input-output models, to name but a few. Interesting works on system identification using GP are Kocijan (2007), where GP are proposed and used for the modelling of non-linear dynamic systems, Azman & Kocijan (2017), in which GP model identification is applied including prior knowledge as local linear models, or Belyaev et. al. (2015), where a new approach of GP regression is described to handle large datasets for aproximation. For a survey of GP identification see Kocijan (2016), and for an interesting survey on kernel methods for system identification, the reader is referred to Pilloneto et al. (2014). With respect to the Kernel Ridge Regression and Conformal Predictors methods, as far as the authors know, there is a lack of works dealing with the modelling of system dynamics using such methods, with some notable exception such as Burnaev & Nazarov (2011), where a computationally efficient conformal procedure for KRR is given and tested against Bayesian confidence tests.

In the field of marine systems, we can find some works that employ neural networks to define the dynamics of a surface marine vehicle, such as Haddara & Wang (1999), Haddara & Xu (1999), Hornick et al. (1989), or Mahfouz (2004). We can also find some interesting works that deal with the identification of marine vehicles by using SVM, for example Zhang & Zou (2009), where an Abkowitz model for ship manoeuvring is identified using LS-SVM, and Zhang & Zou (2011), where ϵ-SVM is employed for the computation of the same model. These two above works search to determine the hydrodynamic coefficients of a Mariner Class Vessel with simple training manoeuvres, however, the identification of the mathematical model is made with data obtained from simulation, and the prediction ability of the model is also tested in simulation. Furthermore, as far as the authors know, most of the works that deal with system identification using some SVR technique employ simulation data and numerical examples, where the models obtained are not tested on an experimental set-up. Some exceptions are the works Moreno-Salinas et al. (2013a), in which the steering equations of a Nomoto second order linear model with constant surge speed are identified using LS-SVM and tested in an experimental set-up with a scale ship model, Moreno-Salinas et al. (2013b) where the Blanke model of a surface marine vehicle is estimated using LS-SVM, Xu et al. (2013), in which an identification method based on SVM is proposed for modelling non-linear dynamics of a torpedo AUV, or Moreno-Salinas et al. (2015) where a black box identification of a ship is made using Genetic Programming based Symbolic Regression.

Following this trend, in this paper we propose a new black box identification of a marine vehicle using Conformal Predictors (CP), firstly introduced in Gammerman et al. (1998). Unlike other methods in machine learning that predict labels with no confidence intervals, CP are able to produce a confidence boundary for each prediction. There are different methods where confidence intervals can be calculated such as Probably Approximate Correct (PAC) learning and Bayesian approaches; however PAC can produce boundaries greater than 1, and Bayesian theory can generate confidence intervals but they require to make a priori assumptions that reflect the real distribution. The only assumption in CP is that data should be independently identically distributed (i.i.d.) but, according to Shafer & Vovk (2008), the *exchangeability assumption* is enough, which means that a sufficient assumption is that data do not follow any particular order instead of being independent. A detail analysis is given in Vovk (2005), and Shafer & Vovk (2008).

As far as the authors know, this kind of algorithms has not been previously tested and applied for marine system identification. Therefore, the aim of the work at hand is to obtain an accurate model of the vehicle without knowing the structure of the model a priori, i.e., a black box identification where the only a priori knowledge are the inputs and outputs of the system. Thus, we want to determine the model that better describes the relationship between inputs and outputs without any constraint or condition on the mathematical model structure. The model structure is given by the identification process itself as the best structure to fit the experimental data used to train and test the model.

The main contributions of the present paper are i) a new black box identification method based on Conformal Predictors is proposed for modelling of marine vehicles, ii) a continuous-time mathematical model is trained and tested with data obtained from real experiments with a ship developing some classical manoeuvres used for system identification on marine vehicles, iii) a confidence margin is proposed to ensure that the model behaviour is within some limits where the real behaviour of the vehicle should lie.

The paper is organized as follows. In Section 2 the problem formulation is presented. In Section 3 the techniques used for the identification of the vehicle are introduced and explained in detail. The model is computed in Section 4 and the results are shown in Section 5. Finally, the conclusions and future lines of work are given in Section 6.

2. Problem Formulation

In marine systems, the high cost of time and money in carrying out experiments pushes researchers to compute mathematical models of the physical systems for testing them in simulation before making the experiments in the open sea. There are multiple models described in the literature, see Fossen (2011) and Perez (2005) for an overview on common models in marine systems. In this paper we compute a 3 degrees of freedom (DOF) continuous-time black box model of the differential equations of the speeds of a marine surface vehicle, namely, the surge speed, the sway speed and the yaw rate.

A continuous-time black box model is selected to avoid the predefined behaviour given by parametric models that can be found in literature, which can fail in representing the actual behaviour of the system due to unmodelled components. In this sense, the black box identification allows to compute the model that better fits the experimental data, regardless of the model structure or its complexity, and taking into account behaviours that could not be considered by pre-defined models.

The continuous-time model considered in this paper has the following form in differential equations:

$$\begin{aligned} \dot{u} &= f_1(u, v, r, T, \delta) \\ \dot{v} &= f_2(u, v, r, T, \delta) \\ \dot{r} &= f_3(u, v, r, T, \delta), \end{aligned} \tag{1}$$

where u is surge speed, v is sway speed, r is yaw rate, T is tension applied to the motor that controls the turning speed of the propeller, and δ is the rudder angle. The right sides of equations $\dot{u}$, $\dot{v}$, and $\dot{r}$ are represented by the unknown functions f_1, f_2 and f_3 that will depend directly on the present u, v, and r of the ship, and the commanded T, and δ.

Two different approaches have been considered for the model computation: a first model constructed with Kernel Ridge Regression (KRR) and a second model constructed applying CP using KRRCM in order to provide confidence margins to the response of the estimated model. For the sake of clarity, these algorithms are briefly described in the next sections.

3. Machine Learning Techniques

This section provides an overview of the techniques employed for the identification of the vehicle. Firstly, in 3.1 Kernel Ridge Regression (KRR) theory, which will be used to train the first model, is briefly introduced. Secondly, in 3.2, CP framework is briefly described, and in 3.3 the Kernel Ridge Regression Confidence Machine (KRRCM), which will be used to compute the second model, is explained. Notice that, as both models are constructed using KRR, their mathematical representations and predictions will be the same. However, in the second case, and additionally to the predictions, confidence margins for these predictions are also computed, providing a more useful and informative model.

3.1 Kernel Ridge Regression

The well known regression shrinkage method Ridge Regression (RR) was proposed in Hoerl & Kennard (1970). Define the regression problem as follows: given a set of n vectors, $x_1, \ldots, x_n$ in R^m, where m is the number of attributes, and the dependent variable $y_i \in R, i = 1, \ldots, n$, the objective is to minimize the loss function (square error), i.e., the discrepancy between the real values y_i and the predicted values $\hat{y}_i = w \cdot x$:

$$L(w) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - w \cdot x_i)^2, \tag{2}$$

where $w \in R^m$, and the optimal value $\hat{w}_0$ will be used for labelling the new incoming samples: $y_{n+1} = \hat{w}_0 \cdot x_{n+1}$.

Ridge Regression modifies least squares by adding to the attributes a regularization l_2 to reduce the variance:

$$L(w) = \sum_{i=1}^{n} (y_i - w \cdot x_i)^2 + \lambda ||w||_2^2, \tag{3}$$

where the shrinkage parameter $\lambda > 0$ controls the penalization ($\lambda = 0$, there is no penalization and formulation is equivalent to least squares; large values for λ, the attributes are heavily constrained).

Saunders et al. (1998) proposed the "dual version" of the Ridge Regression using kernel functions, which represent dot products in the feature space. Then, the kernel trick can be applied using an algorithm in the feature space without the need of any computation within the feature space. Following Saunders et al. (1998), which partially follows Vapnik (1995) in the derivation, Kernel Ridge Regression (KRR) yields

$$\hat{y} = f(x) = \sum_{i=1}^{n} \beta_i K(x, x_i), \tag{4}$$

where $K(x, x_i)$ is the kernel function, which can be seen as a measure of similarity between features, and β_i are the weights. These weights are learnt by minimizing the cost function:

$$C(\beta) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda \beta^T K \beta, \tag{5}$$

where λ is the same variable introduced previously, the regularization parameter, K is the kernel matrix $K(x_i, x_j)$, y_i is the target value and $\hat{y}_i$ is the predicted value. Therefore the expression is similar to RR Eq. (3) but the kernel trick is applied to substitute all the dot products by the kernel function. The exact solution for coefficients β can be expressed as:

$$\beta = (K + \lambda I)Y, \tag{6}$$

where I is the identity matrix $(n \times n)$, and $Y = (y_1, ..., y_n)^T$. Then, the general expression for KRR in Eq. (4) can be used for the identification as exact solution of the algorithm.

3.2 Conformal Prediction

Conformal Prediction (CP) was firstly introduced in Gammerman et al. (1998). A detailed analysis is given in Vovk (2005); Shafer & Vovk (2008). There are two main aproaches in CP, Transductive Conformal Prediction (TCP) and Inductive Conformal Prediction (ICP). TCP follows a transductive approach and it is computationally inefficient for large datasets. All the computations are repeated for each new sample, the training of the model and prediction. Then, the new prediction is based on all the previous samples. On the contrary, ICP for regression and for classification uses a batch of old samples to train a model which is applied to all new samples. As mentioned in Papadopoulos (2011), ICP replaces the transductive inference followed in TCP with the inductive inference. Consequently, ICP is almost as computationally efficient as the underlying algorithms but the cost is some loss in the quality of the produced confidence measures. However, this loss is negligible, especially for large datasets. In the paper at hand, we have chosen a TCP approach for the modelling of the marine vehicle since the experimental data are not too large and we look for better quality on the confidence boundaries.

Following Papadopoulos (2011), given a dataset $z_i \in Z$ with $z_1, ..., z_n$ as training set, formed by pairs $z_i = (x_i, y_i)$, where $x_i \in R^m$ is the vector of attributes in a regression problem and $y_i \in R$ is the label of each sample, the task, given a new unlabelled sample x_{n+1}, is to determine the confidence values $\tilde{y}$ for label y_{n+1}. At this point, it is necessary to introduce and define a *nonconformity measure*, which can be seen as the family of functions $A_n : Z^{n-1} \times Z \rightarrow R, n = 1, 2, ...$ (where Z^{n-1} is the set of all multisets of size $n-1$), giving a measure or score of the difference of a new sample from the old samples:

$$\alpha_i = A_n(\{z_1, ..., z_{i-1}, z_{i+1}, ..., z_n\}, z_i), \tag{7}$$

where α_i indicates how different z_i is from $\{z_1, ..., z_{i-1}, z_{i+1}, ..., z_n\}$. This score, that may be any function, is usually a function related to the underlying machine algorithm used in the problem. A prediction rule $D_{\{z_1,..z_{n+1}\}}$ can be trained and the nonconformity measure of $z_i \in \{z_1, .., z_{n+1}\}$ will be calculated as the disagreement between the real output y_i and the predicted value $\hat{y}_i = D_{\{z_1,..z_{n+1}\}}(x_i)$. Therefore, if we want to predict the value $\tilde{y}$ for a new sample x_{n+1}, we can form the dataset:

$$\{z_1, ..., z_{n+1}\} = \{(x_1, y_1), ..., (x_{n+1}, \tilde{y})\} \tag{8}$$

and measure the nonconformity score for each sample z_i:

$$\alpha_i = A_{n+1}(\{z_1, ..., z_{i-1}, z_{i+1}, ..., z_{n+1}\}, z_i). \tag{9}$$

Since α_i is just a score, we can compare all the values α_i with α_{n+1} to compute the p-value:

$$p(\tilde{y}) = \frac{\#\{i = 1, ..., n+1 : \alpha_i \geq \alpha_{n+1}\}}{n+1}. \tag{10}$$

It is important to note that it is not the statistical p-value. An important property shown in Nouretdinov et al. (2001a) is that $\forall \delta \in [0, 1]$ and for all probability distributions P on Z:

$$P\{\{z_1, ..., z_{n+1}\} : p(y_{n+1}) \leq \delta\} \leq \delta. \tag{11}$$

It means that if the p-value is below some very low threshold (for example 0.05) this label is highly unlikely since these sets will only be generated at most 5% of the time by any i.i.d. process. If we can calculate p-values for every possible label $\tilde{y}$, given a significance level δ, i.e., a confidence level $1-\delta$, a regression using CP would output a set containing a predictive region with the possible labels $\tilde{y}$ that satisfy the confidence level:

$$\{\tilde{y} : p(\tilde{y}) > \delta\}. \tag{12}$$

Since it is not possible to calculate the p-values for all the possible labels $\tilde{y} \in R$, the predictive region in Eq. (12) is estimated following the Ridge Regression Confidence Machine (RR-CM) algorithm in Nouretdinov et al. (2001b).

3.3 Kernel Ridge Regression Confidence Machine

This section gives a quick review on Kernel Ridge Regression Confidence Machine (KR-RCM), please see Nouretdinov et al. (2001b) for a detailed description. Given the training examples $(x_1, x_2, ..., x_n)$ with labels $(y_1, y_2, ..., y_n)$, the purpose is finding all possible values $\tilde{y}$ for the new sample x_{n+1} with label y_{n+1}. Let define $X = (x_1, x_2, ..., x_n, x_{n+1})'$ and $Y = (y_1, y_2, ..., y_n, y_{n+1})'$. The loss function of RR, see Eq. (3), can be expressed in matrix form as follows:

$$\lambda||w||^2 + ||Y - Xw||^2 = Y'Y - 2w'X'Y + w'(X'X + \lambda I)w. \tag{13}$$

Taking derivatives the solution is:

$$w = (X'X + \lambda I)^{-1}X'Y. \tag{14}$$

Defining the nonconformity measure as the residuals:

$$e_i = |y_i - \hat{y}_i|, i = 1, ..., n+1, \tag{15}$$

we have the vector $\Delta = |Y - Xw|$ of residuals $e_1, ..., e_{n+1}$ as:

$$\Delta = |(I - X(X'X + \lambda I)^{-1}X')Y| \tag{16}$$

where $Y = (y_1, ..., y_n, 0)' + (0, ..., 0, \tilde{y})'$ and we can write $\Delta = |A + B\tilde{y}|$.

Using the dual formulation of ridge regression, the kernel trick can be applied to $\Delta = |Y' - wX'|$ obtaining the KRR form:

$$w = Y'(XX' + \lambda I)^{-1}XX') = Y'(K + \lambda I)^{-1}K. \tag{17}$$

where K is the matrix from the kernel function $K(x_i, x_j)$, and vectors $A = (a_1, ..., a_{n+1})$ and $B = (b_1, ..., b_{n+1})$ can be written as:

$$A = (y_1, ..., y_n, 0)(I - (K + \lambda I)^{-1}K), \tag{18}$$

$$B = (0, ..., 0, 1)(I - (K + \lambda I)^{-1}K). \tag{19}$$

The p-value will change only at points where $e_i = e_{n+1}$, so we only calculate those values greater than δ rather than all possible values. For each training sample the set of all possible values for $\tilde{y}$ yields

$$S_i = \{\tilde{y} : e_i \geq e_{n+1}\} = \{\tilde{y} : |a_i + b_i\tilde{y}| \geq |a_{n+1} + b_{n+1}\tilde{y}|\}. \tag{20}$$

According to Nouretdinov et al. (2001b) S_i can be given as shown in Eq. (21), where

- If $b_i \neq b_{n+1}$: u_i and v_i are the minimum and maximum respectively of the following values $\frac{a_i - a_{n+1}}{b_{n+1} - b_i}, -\frac{a_i + a_{n+1}}{b_{n+1} + b_1}$.
- If $b_i = b_{n+1} \neq 0$: $u_i = v_i = -\frac{a_i + a_{n+1}}{2b_1}$

$$S_i = \begin{cases} [u_i, v_i] & \text{if } b_{n+1} > b_i \\ (-\infty, u_i] \cup [v_i, \inf) & \text{if } b_{n+1} < b_i \\ (u_i, \infty] & \text{if } b_{n+1} = b_i \text{ and } a_{n+1} < a_i \\ (-\infty, u_i] & \text{if } b_{n+1} = b_i \text{ and } a_{n+1} > a_i \\ R & \text{if } b_{n+1} = b_i = 0 \text{ and } |a_{n+1}| \leq |a_i| \\ \emptyset & \text{if } b_{n+1} = b_i = 0 \text{ and } |a_{n+1}| > |a_i| \end{cases} \tag{21}$$

The p-value only changes at points u_i and v_i. If $\tilde{y}_1, ..., \tilde{y}_{2n} = (u_1, ..., u_n, v_1, ..., v_n)$ in ascending order, p is constant on any interval $(\tilde{y}_0, \tilde{y}_1), ..., (\tilde{y}_{2n}, \tilde{y}_{2n+1})$, where $\tilde{y}_0 = -\infty$ and $\tilde{y}_{2n+1} = \infty$

Then we can calculate p for any interval counting the number of S_i that include such p-value and dividing by $n+1$. If $N(\tilde{y}_i)$ is defined as:

$$N(\tilde{y}_i) = \#\{S_j : [\tilde{y}_i, \tilde{y}_{i+1}] \subseteq S_j\},\ i = 0, ..., 2n;\ j = 1, ..., n+1; \tag{22}$$

the p-value for $[\tilde{y}_i, \tilde{y}_{i+1}]$ is $\frac{N(\tilde{y}_i)}{n+1}$. We are interested in finding a region of labels where the probability of the true label outside that region is δ or less. This is exactly what is done in the KRRCM algorithm fully described in Nouretdinov et al. (2001b), given a training set and the new sample, the outputs are the predicted value and the confidence region for this new sample.

4. Surface marine vehicle identification

Once a brief overview on the core algorithms used in this work has been given, the modelling of the surface marine vehicle is carried out. The experimental system is firstly introduced and, afterwards, the computation of the model is developed.

4.1 Experimental system

The ship used in the experimental tests is a scale model of an operational vessel, in a 1/16.95 scale. The dimensions of the vessel and the scale model are detailed in Table 1.

Table 1: Main parameters and dimensions of the vessel and the scale ship.

Parameter	Vessel	Scale Ship
Length between perpendiculars (L_{pp})	74.40 m	4.389 m
Maximum beam (B)	14.20 m	0.838 m
Mean depth to the top deck (H)	9.05 m	0.534 m
Design draught (T_m)	6.30 m	0.372 m

The scale ship model, hereinafter referred to as the ship, is an underactuated vehicle controlled by a DC electric motor, connected to a single propeller, and a servo attached to the rudder. Then, the surge speed is controlled by the turning speed of the propeller, which is directly proportional to the tension applied to the motor, and the yaw rate is controlled by the servo. The desired rudder angle and surge speed are commanded through a long-range Wi-Fi connection between the ship and the control station at land, or they may be computed by a control law programmed in the onboard computer. The ship used for the experimental tests is shown in Fig. 1.

The data used for the model computation are obtained from a typical and simple manoeuvre used for marine vehicles identification, a 20/20 degrees Zig-Zag, which will be described in the next section. Then, the aim is to show that with a simple experiment and relatively few data, a reliable mathematical model can be computed, which will be tested with additional manoeuvres, such as a 10/10 degrees Zig-Zag and evolution circles.

The experiments were carried out in the reservoir of Valmayor, in the north of Madrid, Spain.

Figure 1: Ship used in the tests.

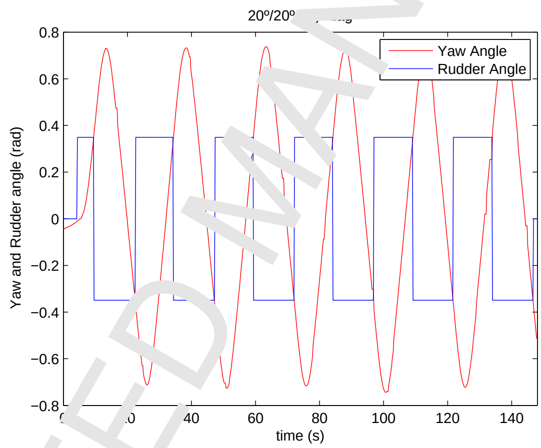


Figure 2: 20/20 degrees Zig-Zag manoeuvre. Yaw angle in red and rudder angle in blue.

4.2 Model computation

As mentioned above, the training data used for the system identification are obtained from a 20/20 degrees Zig-Zag manoeuvre with a sample time of $\Delta t = 0.2$ seconds, and a nominal commanded surge speed of 2 m/s. This manoeuvre consists in the following steps: i) start with constant speed with yaw and rudder angle fixed at 0 degrees; ii) change the rudder angle to 20 degrees and wait until the yaw angle reaches 20 degrees; iii) change the rudder to -20 degrees, and keep this rudder angle until the yaw angle reaches -20 degrees; iv) go to step ii) and repeat the process as many times as appropriate. The same steps apply for any Zig-Zag manoeuvre with any other angle.

A set of 740 samples is taken during 140 seconds of the 20/20 degrees Zig-Zag manoeuvre. The commanded rudder angle (blue line) and the yaw angle (red line) defined by the ship in the Zig-Zag manoeuvre are shown in Figure 2.

The commanded surge speed will be the same for all the tests, so it can be discarded as input for the model computation, and only the rudder angle will be considered as control input. Thus, the commanded control signal (rudder angle) and the data measured from the IMU on board the ship (surge speed, yaw rate and sway speed) from the 20 to 100 seconds of Fig.2 are considered as the training data. The first 20 seconds of the manoeuvre are not considered for the model computation to avoid possible disturbances on data during the beginning of the manoeuvre. The rest of the manoeuvre, from 100 to 140 seconds, will be used as test data together with the additional manoeuvres trough free running simulations, i.e., only the actual initial conditions are given to the model, and with the same control commands given to the real vehicle, the model executes the manoeuvre that corresponds to these control actions. Notice that the model will be computed just with 80 seconds of a simple manoeuvre, in contrast to the multiple and time costly manoeuvres usually employed for marine vehicle modelling. The behaviour of the model is then compared with that of the real ship to evaluate the accuracy of the model. The additional manoeuvres are a 10/10 degrees Zig-Zag and some evolution circles. This latter manoeuvre consists in the following steps: i) start with constant speed, with rudder angle at 0 degrees; ii) fix the rudder to the angle considered for the manoeuvre (in our case 20 degrees); iii) let the vehicle describe several consecutive circles with the same commanded rudder angle, iv) come back to step i) and select a new rudder angle.

It is important to remark at this point that all the experiences where carried out on an open environment and in different days, so that they were subjected to different perturbations such as currents and/or wind. In fact, the two Zig-Zags were carried out on a very calm day, so the perturbations could be neglected. However, during the day when the evolution circle trials were carried out, the wind was strong enough to affect the manoeuvre. This can be noticed in Figure 3 (a) and (b), where the surge speed and sway speed for these trials are shown in red. During this manoeuvre the surge speed and sway speed should be almost constant since the rudder is fixed at a given angle and the commanded advance speed is constant, as mentioned above; however, there is a periodic perturbation (wind) that modifies the speed at the same period at which the vehicle is turning and with an amplitude of 0.1 m/s. Moreover, the period of the perturbation in the first part (85 seconds) is a bit larger than in the second one (69 seconds) because the ship turns at a larger angular rate in the latter direction, see also Fig. 5(f) and Fig. 7(f) in the results of Section 5.1 and 5.2. Probably, this effect is caused by a slightly asymmetrical ballast, which makes the ship turn faster in one direction once the steady state is reached for a fixed rudder angle, but that it is not enough to be appreciated in the Zig-Zag manoeuvres, where the ship does not reach a steady state. For the sake of clarity on the results shown in the next section, we have tried to eliminate the periodic perturbation caused by wind from the speeds, shown in Fig. 3 (a) and (b) in blue, so that the model computed and the simulation results can be correctly compared with the experimental results.

As mentioned on Eq. (1), the differential equations of surge speed, sway speed and yaw rate are estimated, so the regression is made over the accelerations of the ship. For KRR a model will be obtained for each differential equation according to Eq. (4):

$$\begin{array}{l} \dot{u} = \sum_{i=1}^{n} \beta_i^u K(x, x_i) \\ \dot{v} = \sum_{i=1}^{n} \beta_i^v K(x, x_i) \\ \dot{r} = \sum_{i=1}^{n} \beta_i^r K(x, x_i) \end{array} \tag{23}$$

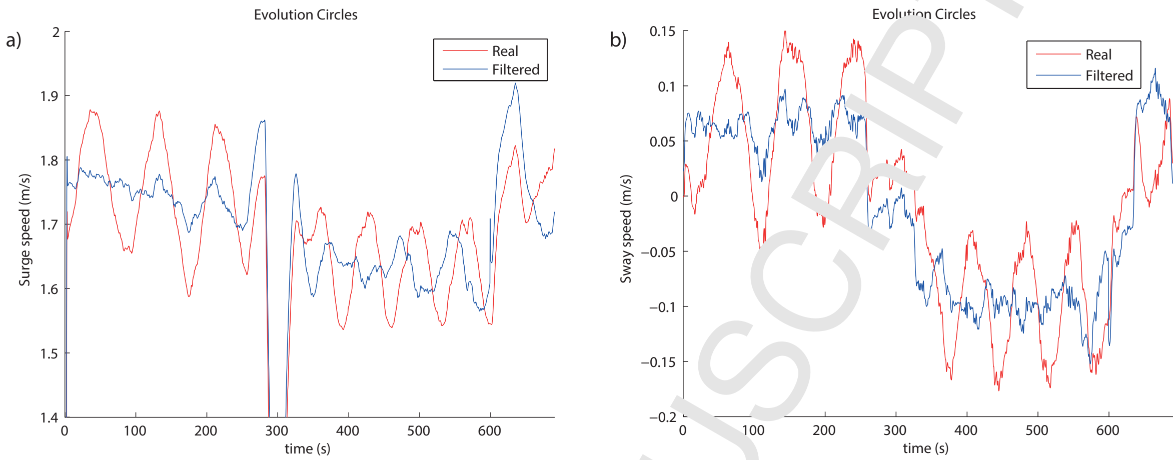


Figure 3: Surge (a) and sway (b) speeds for the evolution circle trials, in red the experimental values and in blue after filtering the periodic perturbation.

where a RBF kernel has been used, $K(x, x_i) = exp\left(-\frac{||x-x_i||^2}{2\sigma^2}\right)$, and each sample $x = (u, v, r, T, \delta)$ and each set of β's will be different for each equation. Applying CP with KRRCM a similar expression is obtained but, as mentioned before, the purpose is to obtain a confidence margin around the predicted value following the procedure explained in Section 3.3.

5. Results

5.1 Kernel Ridge Regression (KRR)

First we evaluate the results obtained with KRR. Notice that for the computation of the model the advance speed was considered constant and the same for all the tests.

There are two parameters to be tuned in the model, the regularization parameter λ from ridge regression and σ from the RBF kernel selected. A grid-search has been carried out for the selection of these parameters, σ_u, σ_v, σ_r, and λ_u, λ_v, λ_r, in Eq. (23), see Table 2 with the parameters scan, and Table 3 where the parameters that provide a more accurate result for the training and test data are shown. Two evaluation measures have been used to measure the performance on the training set and to select the optimal values in the parameters search: mean absolute error and R^2 coefficient.

Table 2: Parameters search for KRR models

# permutations	Range of σ	Range of λ
73	0.001 – 0.5	0.01 – 100
152	1	0.01 – 1
112	0.8 – 2	0.01

Table 3: Parameters for Kernel Ridge Regression

σ_u	σ_v	σ_r	λ_u	λ_v	λ_r
1	1	1	0.0313	0.0313	0.0313

In Fig. 4 the actual response of the ship (red line) is shown together with the one developed by the computed mathematical model (green and blue lines, for train and test, respectively), showing both a large similarity. In this figure we can see the surge acceleration in Fig. 4(a) and the surge speed in Fig. 4(b), the sway acceleration and sway speed in Fig. 4(c) and (d) respectively, and the yaw acceleration and yaw rate in Fig. 4(e) and (f), respectively. Notice how the prediction of the model is very accurate both for accelerations and speeds (where the errors are integrated), making the model very appropriate for control simulation purposes. These results correspond to the 20/20 degrees Zig-Zag manoeuvre, where almost the whole manoeuvre has been used for training and only the last part for testing.

For additional testing of the model, a 10/10 degrees Zig-Zag and some evolution circles of ± 10 degrees are considered, with their corresponding results shown in Figure 5. For the sake of clarity, only the speeds will be shown as outputs of the model, as the errors are more evident than in the accelerations. In addition, for simplicity reasons, in the simulations all the evolution circles manoeuvres are considered as consecutive manoeuvres instead of independent experiments. This will also allow to check the behaviour of the model when the manoeuvre is changed. Notice also that the motor is stopped at the beginning of the 10/10 degrees Zig-Zag and when the turning direction is changed in the evolution circles, so we can check how the model behaves with these changes on the motor, that have not been taken into account for the model computation. In Fig. 5 (a), (c) and (e) the surge speed, sway speed and yaw rate for the 10/10 degrees Zig-Zag are shown. In Fig. 5 (b), (d) and (f) the surge speed, sway speed and yaw rate for the evolution circles are given. We can notice in this example how the model response is very similar to the actual one of the ship.

It is important to recall that the experimental data obtained from the evolution circles manoeuvres were obtained in a different day than the Zig-Zags, and thus, with different environmental conditions, where disturbances were present, such as currents or wind. These environmental conditions have been smoothed by the filtering shown in Section 4.2, Fig. 3, but not completely eliminated. These unconsidered environmental conditions may be the cause of the slightly deviation of the results in the evolution circles compared with the Zig-Zags. In any case, the response for all the manoeuvres is very accurate.

Therefore, for the training data the behaviour is almost the same, as expected, but for the test data it is also very similar, showing how the model reproduces accurately the manoeuvres developed by the ship. Moreover, in the gap produced in the real data at the connection of two different evolution circles, where the surge speed starts from 0 m/s, we can see how the model continues its tendency with a very natural behaviour and providing a result very close to the real one once the ship has reached its nominal speed.

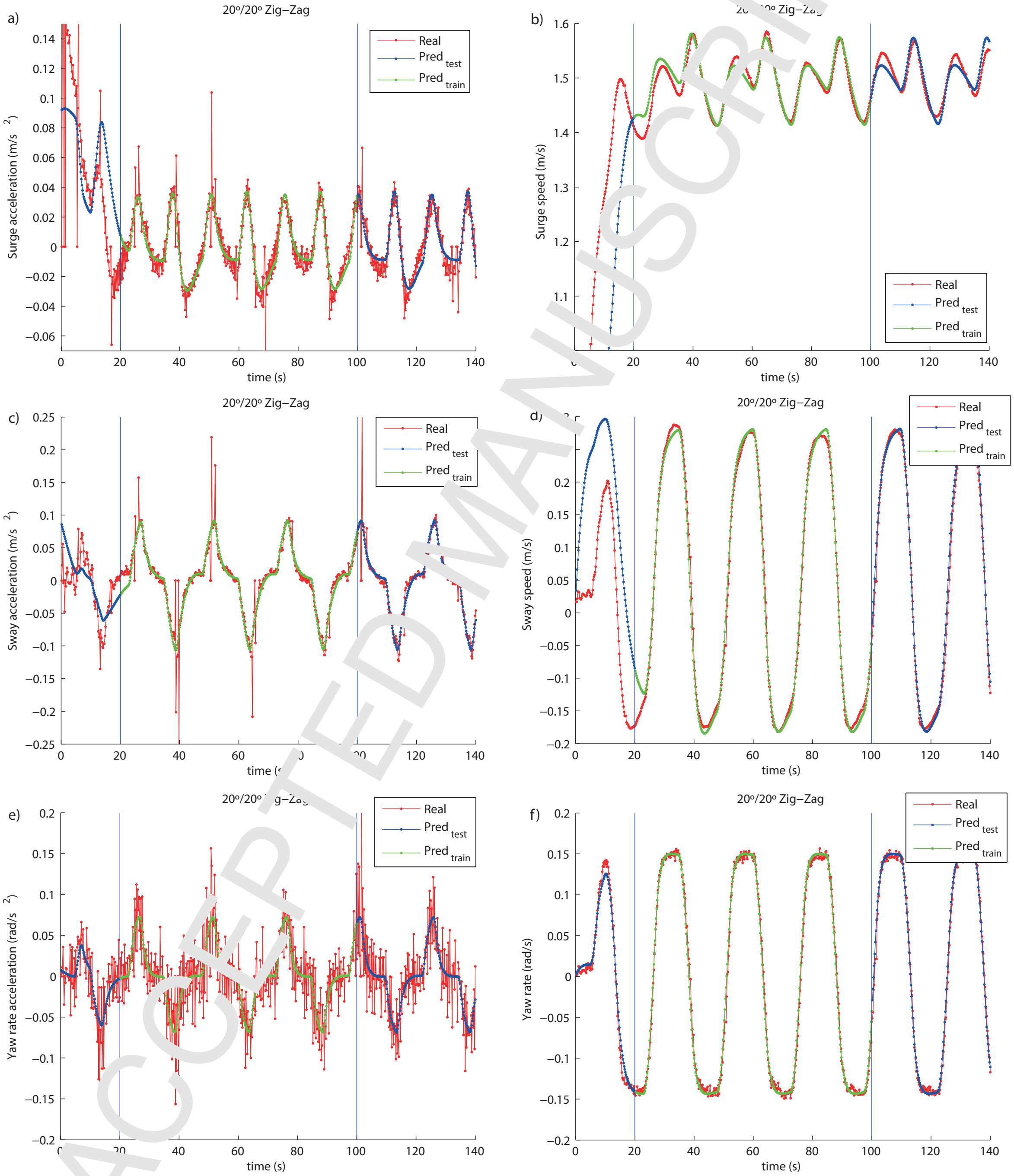


Figure 4: KRR: Accelerations and speeds of the ship and model for the training.

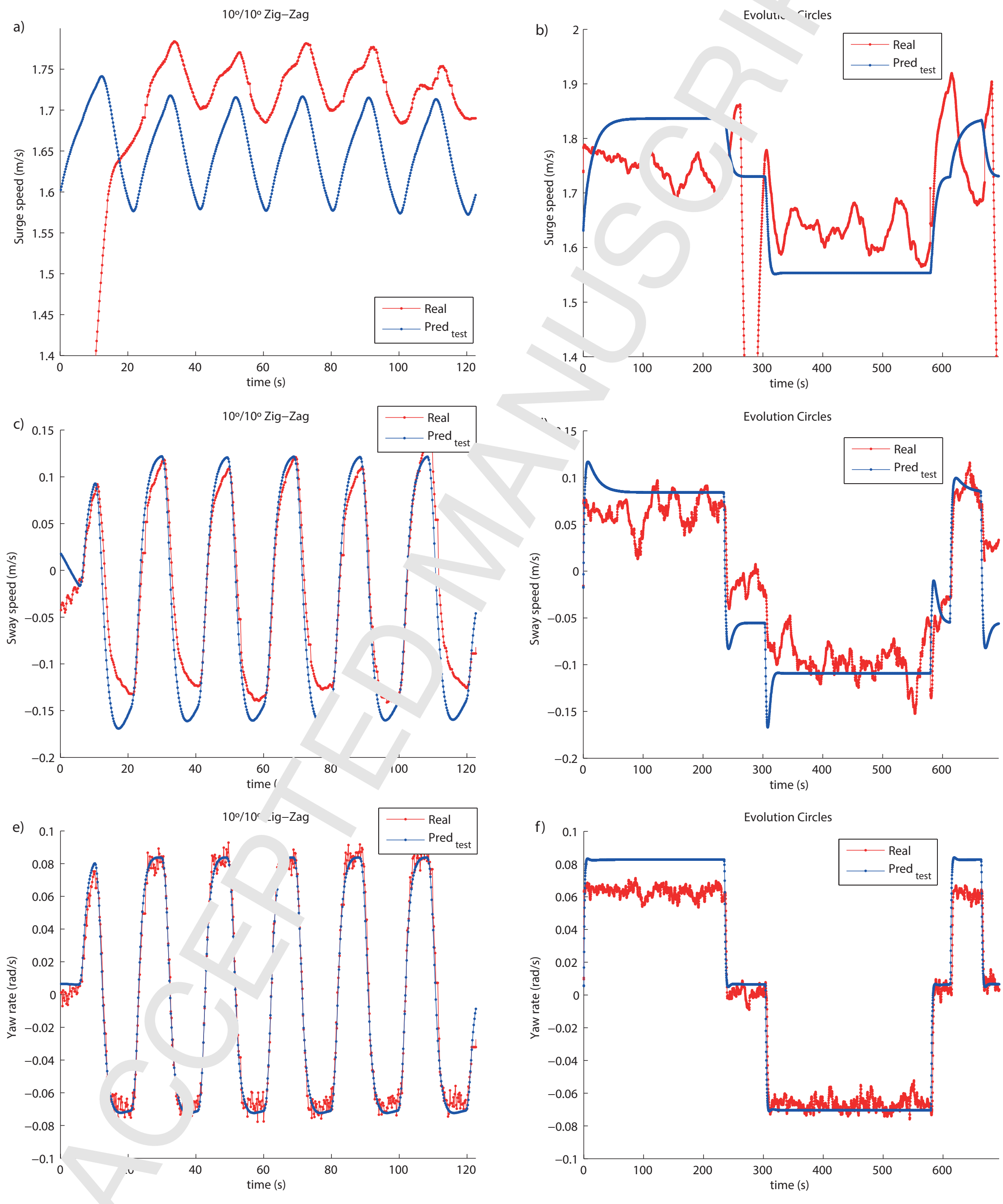


Figure 5: KRR: Testing of the model.

5.2 Kernel Ridge Regression Confidence Machine (KRRCM)

Now, the computation of the model is carried out considering KRRCM, i.e., we will employ Conformal Prediction for the system identification process. Therefore, with the computation we can ensure that the output model will be within some confidence margins where the real output of the system should also lie. We proceed in the same way as in the previous case, using the first part of the 20/20 degrees Zig-Zag as training data and the rest of the manoeuvre, the 10/10 degrees Zig-Zag and the evolution circles as test data. For the sake of simplicity, the two evolution circles will be plotted together and considered as a single manoeuvre for the model.

The parameters used for the computation of the model in this case are shown in Table 4. These parameters have been obtained following the same procedure that in Section 5.1.

Table 4: Parameters for Kernel Ridge Regression with Confidence Margins

σ_u	σ_v	σ_r	λ_u	λ_v	λ_r
1	1	1	0.0313	0.0313	0.0313

Similarly to the previous case, the model is trained with the 20/20 degrees Zig-Zag, keeping the last 40 seconds for testing, as shown in Figure 6, where the train (green line) and test (blue line) are divided by the vertical line. In Figure 6 (a) and (b) we can see the surge acceleration and the surge speed, respectively, in Fig. 6 (c) and (d) the sway acceleration and sway speed, respectively, and in Fig. 6 (e) and (f) the yaw acceleration and yaw rate. For the training data, notice how the speeds and accelerations are practically the same for model and ship. Notice also the confidence margins given by the magenta lines, that ensure that the response of the model is within these margins with a probability of 95 percent. It is also important to notice that most of the real response of the ship is also within these margins, so we ensure that the model will predict the output of the system with large accuracy.

To confirm the above conclusion, we compare the system dynamics of ship and model with the 10/10 degrees Zig-Zag and the evolution circles. In Figures 7 (a), (c) and (e) the surge speed, sway speed and yaw rate are shown for the 10/10 degrees Zig-Zag, respectively, and in Fig. 7 (b), (d) and (f) for the evolution circles. Notice how the behaviour is very accurate in the Zig-Zag. In the evolution circles, although less accurate than in the Zig-Zag, it is still very accurate and very precise for a simulation model and within the confidence margins, or very close to them, ensuring a correct system model.

It is interesting to notice, as shown in Figure 7, how the results are also very accurate for the sway speed. This fact is of special interest since we are considering an underactuated vehicle with 3 degrees of freedom (DOF) but only two control actions (advance speed and rudder angle). Therefore, it is not possible to act over the sway speed directly, only over the surge speed and yaw rate, making it more complex to emulate. However, as aforesaid, the results obtained with the computed model using Conformal Predictors are more than satisfactory.

The small discrepancy between model confidence intervals and ship in the surge speeds may be a consequence of the few data used for training and these latter being restricted

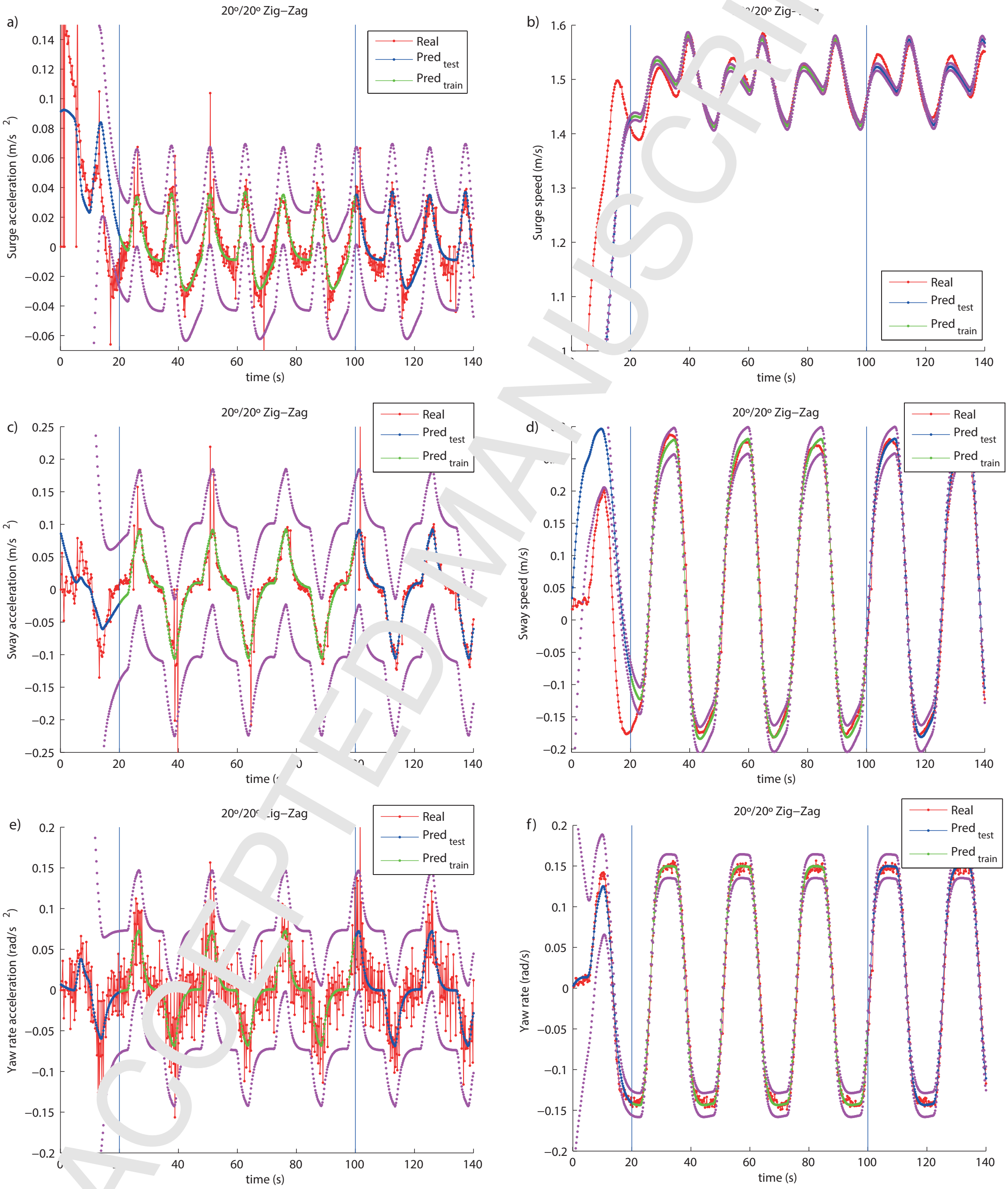


Figure 6: KRRCM: Training of the model.

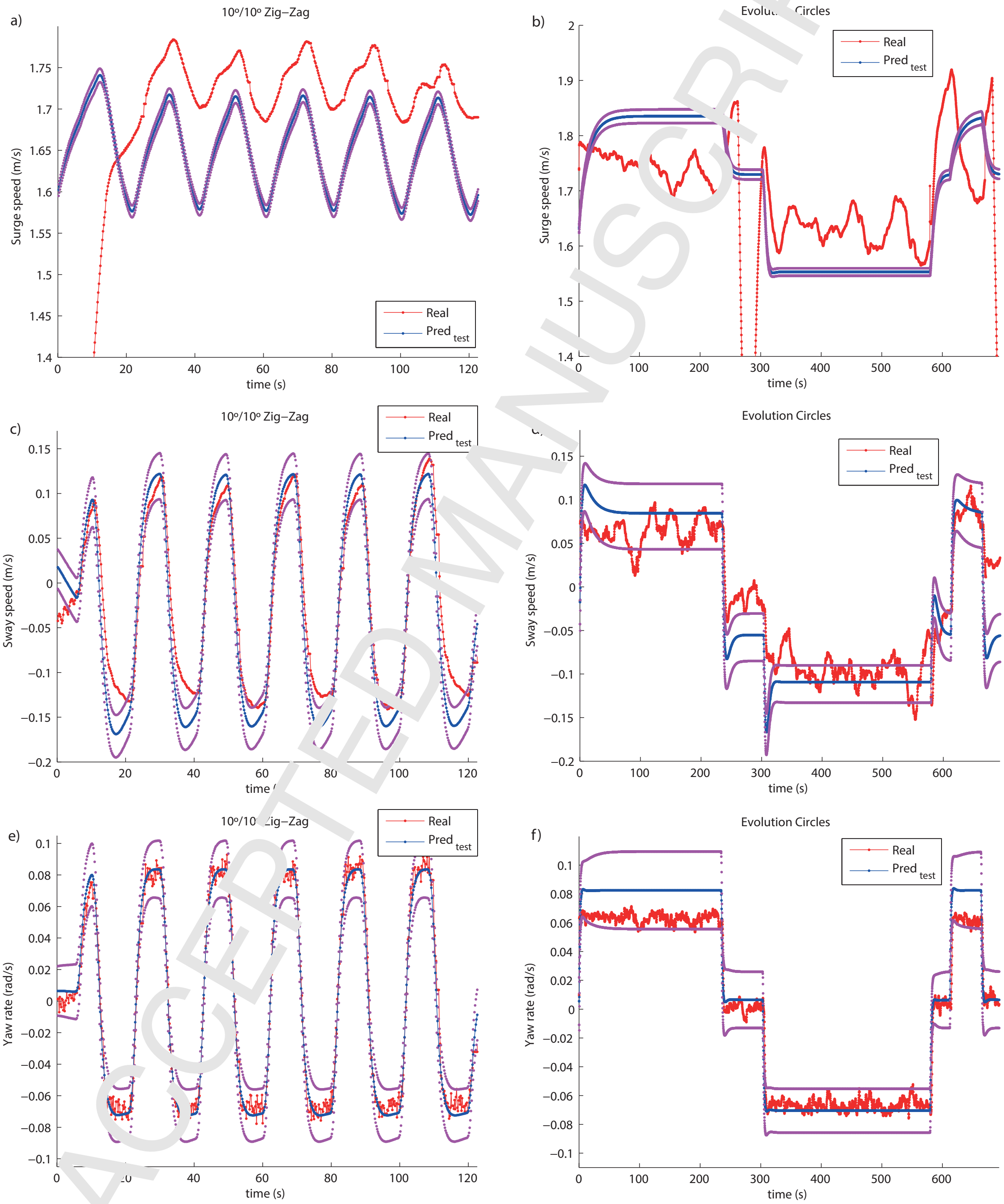


Figure 7: KRRCM: Testing of the model.

to a single manoeuvre. Another possible source of uncertainty may be the possible perturbations not eliminated from the experimental data, and that the tests were carried out in different days with different environmental conditions. However, it can be noticed how the discrepancy in the surge speeds between real and simulated is around 5 %, giving a very accurate simulation result.

For a model computation to be used in a real application of control design, the training data should be richer and obtained from different manoeuvres, in order to make the model obtained even more accurate. However, the aim of this work is to show the applicability of Conformal Predictors to estimate the model of a dynamic system as a surface vessel using relatively few data (only 80 seconds of a single manoeuvre were used for the training). Thus, it has been shown that with few experimental data, and using Conformal Predictors, an accurate model of the vehicle can be obtained. It has been shown that the behaviour of the model is very similar to the real one for different tests and manoeuvres, and it also lies within given confidence margins that ensure our model behaviour. Therefore, Conformal Predictors are a valuable tool to be employed and further studied in the identification of dynamic systems.

5.3 Stability analysis

In this section the stability of the model derived is shown, and how the outputs (surge speed, sway speed, and yaw rate) converge to a stable equilibrium point for constant and bounded inputs (constant commanded rudder angle and advance speed).

The physical system used for the experiments has been designed, constructed and tested before deployment to ensure its adequate behaviour in open-loop, and that it reproduces accurately the behaviour in scale of a real ship. Then, the physical system corresponds to a Bounded Input Bounded Output (BIBO) system, where the output of the system does not diverge for bounded input commands, producing a stable and well-defined output response. In order to show the correct behaviour of the model computed and that the intrinsic stability of the physical system is also carried out by the model, a brief stability analysis based on phase-portraits is given next.

The phase-portraits analysis provides an easy visualization of the qualitative behaviour of the model in open-loop. The main advantage of the phase-portraits is that we can analyse the system for a wide range of initial conditions without solving the nonlinear equations, and they can also be applied equally to any kind of nonlinearity of the system. These phase-portraits are computed considering constant commanded advance speed (the same of the experiments) and constant rudder angle for different initial conditions. Therefore, for different initial conditions and constant inputs, the model is simulated showing the evolution of the state variables, and how the model reaches a stability point regardless of the initial conditions. Since there are three state variables (surge speed, sway speed and yaw rate), simulations are performed for different initial conditions that vary between 1.2 and 1.8 m/s for the surge speed, from -0.5 to 0.5 m/s for the sway speed, and from -30 to 30 degrees per second for the yaw rate. Two different examples are studied.

In the first example, shown in Fig. 8, we consider a commanded rudder of -30 degrees, that is the limit angle of the real vehicle rudder. In Fig. 8 (a) the phase-portrait in 3D is shown. Notice how for the different initial conditions the state variables converge to a single

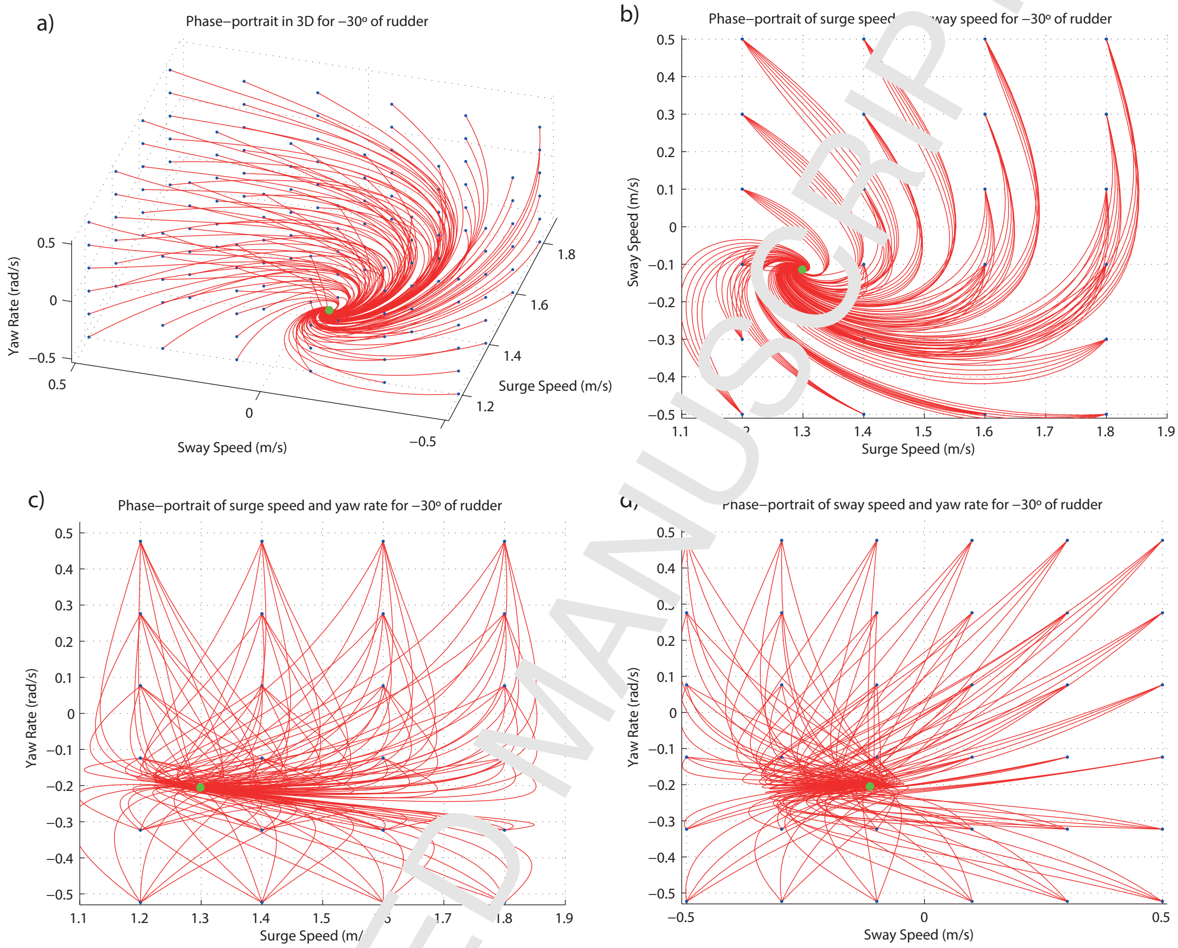


Figure 8: Phase-portraits for a commanded rudder of −30 degrees and commanded advance speed of 2 m/s; the stability point is shown in green. a) Phase-portrait in 3D, b) Phase-portrait for surge speed and sway speed, c) Phase-portrait for surge speed and yaw rate, d) Phase-portrait for sway speed and yaw rate.

and well-defined equilibrium or stable point. The behaviour of the model for the different initial conditions can be seen in detail in Fig. 8 (b), Fig. 8 (c), and Fig. 8 (d) where the projections of Fig. 8 (a) in the surge-speed/sway-speed plane, surge-speed/yaw-rate plane, and sway-speed/yaw-rate plane are shown, respectively. Notice how the vehicle reaches a stable surge speed of approximately 1.3 m/s and a sway speed of -0.1 m/s. The yaw rate takes a value around -0.2 rad/s.

In the second example of Fig. 9, a rudder angle of 5 degrees is used as commanded rudder. Similar plots to the previous example are shown. Fig. 9 (a) shows the phase-portrait in 3D and how for the different initial conditions the model converges to a single stable equilibrium point. The details of the trajectories followed and the equilibrium point reached can be seen in Fig. 9 (b), Fig. 9 (c), and Fig. 9 (d) where the projections of Fig. 9

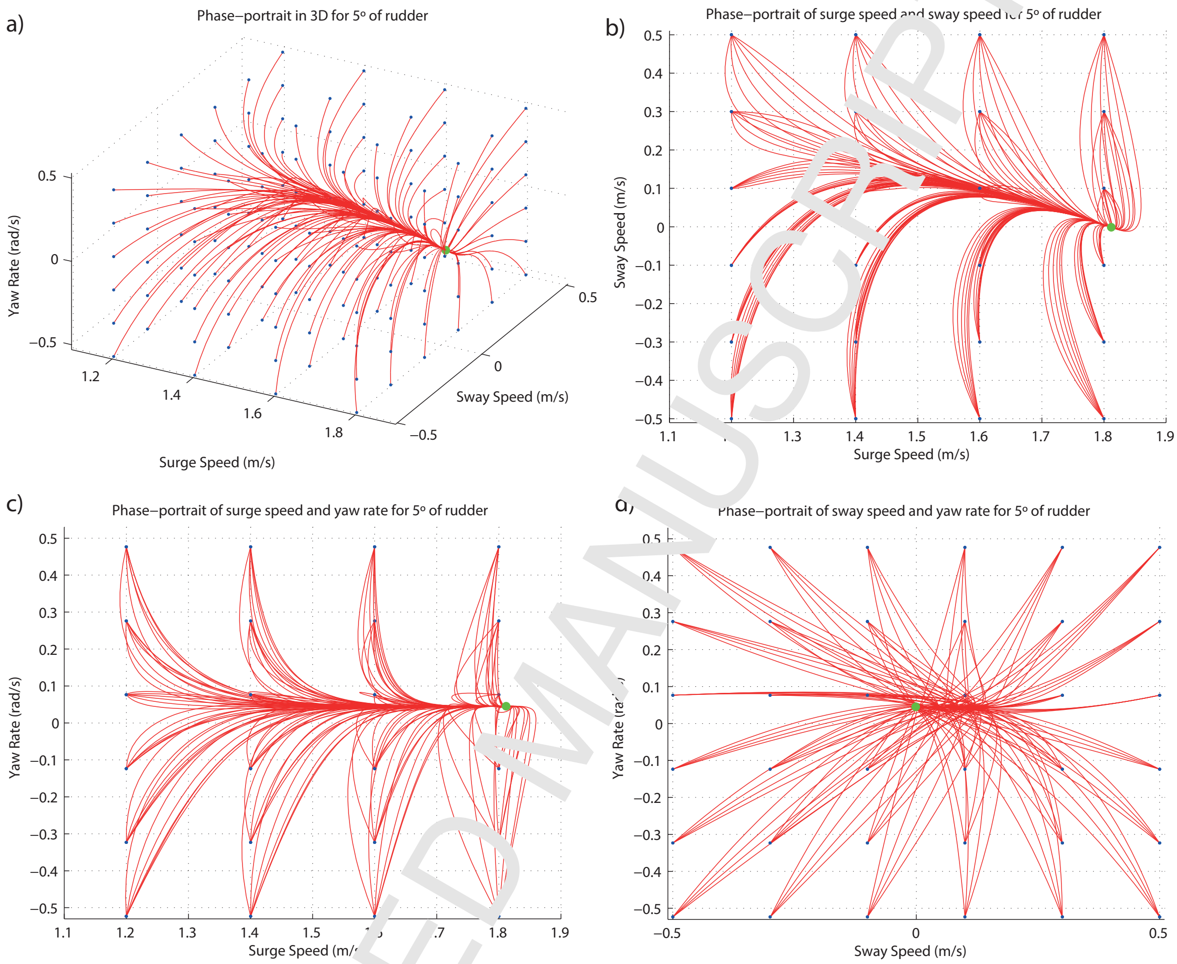


Figure 9: Phase-portraits for a commanded rudder of 5 degrees and commanded advance speed of 2 m/s; the stability point is shown in green. a) Phase-portrait in 3D, b) Phase-portrait for surge speed and sway speed, c) Phase-portrait for surge speed and yaw rate, d) Phase-portrait for sway speed and yaw rate.

(a) in the surge-speed/sway-speed plane, surge-speed/yaw-rate plane, and sway-speed/yaw-rate plane are shown respectively. The values reached in the equilibrium point are a surge speed of approximately 1.8 m/s, a sway speed of 0.04 m/s, and a yaw rate of 0.01 rad/s. Notice how in this case, due to the small rudder angle, the sway speed and yaw rate are small too, demonstrating that the model is consistent with the behaviour supposed to a real ship.

The values for the rudder angle in the above examples have been chosen to show that the performance of the model is consistent with the real behaviour of a ship even for commanded rudder angles quite different to those used for training and testing the model. For other possible rudder angles, such as 10, -10, 20 and -20 degrees, the train and test plots of Subsection 5.1 and Subsection 5.2 show the good response of the model and also its similarity to the real ship. Then, the train and test plots, together with these phase-portraits of the model, show the good behaviour of the model, its accurate response compared with the physical system, and its utility for simulation and control design.

6. Conclusions

In this work, Conformal Predictors have been employed for the identification of a surface marine vehicle. Namely, Kernel Ridge Regression and Kernel Ridge Regression Confidence Machine have been used for the identification process. Both techniques have shown a great performance, providing models that predict with large accuracy the behaviour of the real ship using relatively few data for the training obtained from a single simple manoeuvre. In addition, KRRCM provides confidence margins that ensure the output of the system to be within these limits, making it a more appropriate model for simulation and control design. The accuracy of the proposed models has been tested with different typical manoeuvres used in marine systems for the identification of surface vehicles. Namely, the performance of the models has been tested with two Zig-Zag manoeuvres of 20/20 and 10/10 degrees, respectively, and some evolution circles, showing large accuracy for all the tests. Therefore, it has been shown that Conformal Predictors are a valuable tool for system identification, computing an accurate model of a dynamical system with few experimental data.

As future work, the identification of the ship will be developed with random manoeuvres, to have a richer training data set, and also will be compared and tested with different emerging techniques, so that an appropriate model may be computed for control design purposes.

Acknowledgments

We express our appreciation to the late Prof. Jesus Manuel de la Cruz, whose contribution to this work was of great significance.
The work of D. Moreno-Salinas was supported by "Ministerio de Economia y Competitividad" under project CICYT DPI2014-55932-C2-2-R.
The work of R. Moreno has been supported by Science Foundation Ireland under Grant No. SFI/12/RC/2289

References

Abkowitz, M.A. Measurements of hydrodynamic characteristic from ship maneuvering trials by system identification. *Transactions of the Society of Naval Architects and Marine Engineers*, 88, pp. 283-318, 1980.

Adachi, S. & Ogawa, T. A new system identification method based on support vector machines. *IFAC Workshop Adaptation and Learning in Control and Signal Processing*, Italy, 2001.

Aizerman, M.; Braverman, E. & Rozonoer, L. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, Volume 25: 821837, 1964.

Aman, K. & Kocijan, J. Dynamical systems identification using Gaussian process models with incorporated local models. *Engineering Applications of Artificial Intelligence*, Volume 24, Issue 2, March 2011, Pages 398-408. https://doi.org/10.1016/j.engappai.2010.10.010

Belyaev M., Burnaev E., Kapushev Y. Gaussian Process Regression for Structured Data Sets. In: Gammerman A., Vovk V., Papadopoulos H. (eds) Statistical Learning and Data Sciences. SLDS 2015. Lecture Notes in Computer Science, vol 9047. Springer, Cham.

Belyaev, M.; Burnaev, E.V., Kapushev, E.; Panov, M.; Prikhodko, P.; Vetrov, D. & Yarotsky, D. GTApprox: Surrogate modeling for industrial design. *Advances in Engineering Software*, 102 (2016) 29-39. http://dx.doi.org/10.1016/j.advengsoft.2016.09.001

Boyd,S. & Vandenberghe, L. Convex Optimization. Cambridge University Press New York, NY, USA 2004 ISBN:0521833787

Burnaev, E.V. & Nazarov, I.N. Conformalized Kernel Ridge Regression. *Machine Learning and Applications (ICMLA), 2016 15th IEEE International Conference on*, Anaheim, CA, USA, 18-20 Dec. 2016. DOI: 10.1109/ICMLA.2016.0017

Caccia, M., Bruzzone, G., & Bono, R. A practical approach to modelling and identification of small autonomous surface craft. *IEEE Journal of Oceanic Engineering*, Volume 33, number 2, April 2008.

De la Cruz, J.M., Aranda, J., & Girón, J.M. Automática Marina: una revisión desde el punto de vista de control. *Revista Iberoamericana de Automática e Informática Industrial*, 9 (2012), pp. 205-218.

Drezet, P.M.L. & Harrison, R.F. Support Vector Machines for system identification. In *Proceedings of the International Conference on Control '98*, September 1998.

Forrester, A.; Sobester, A.; Kane, A. Engineering design via surrogate modelling: a practical guide. Jhon Wiley & and Sons; 2008.

Fossen, T.I., Sagatun, S.I., & Sorensen, A.J. Identification of dynamically positioned ships. *Identification and Control*, vol. 17, 2, pp. 153-165, 1996.

Fossen, T. I. *Handbook of Marine Craft Hydrodynamics and Motion Control.* Wiley, 2011, ISBN: 978-1119991496.

Gammerman, A., Vovk V., and Vapnik V. *Learning by transduction, in Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence.* Morgan Kaufmann Publishers Inc.: Madison, Wisconsin. pp. 148-155. 1998

Haddara, M., & Wang. Y. Parametric identification of manoeuvring models for ships. *International Shipbuilding Progress*, Volume 46, 445, pp 5-27, 1999.

Haddara, M.R., & Xu, J.S. On the identification of ship coupled heave-pitch motions using neural networks. *Ocean Engineering*, 26(5), pp 381-400, 1999.

Hoerl, A. E. and Kennard, R. W. *Ridge Regression - Biased Estimation for Nonorthogonal Problems.* Technometrics, 12(1): p. 55, 1970.

Hornik, K., Stinchcombe, M., & White, H. Multilayer feed-forward networks are universal approximations. *Neural networks*, 2, pp 359-366, 1989.

Jemwa, G.T. & Aldricht, C. Non-linear system identification of an autocatalytic reactor using least squares support vector machines. *The Journal of the South African Institute of Mining and Metallurgy*, pp. 119-126, March 2005.

Källström, C.G. & Åström, K.J. Experiences of system identification applied to ship steering. *Automatica*, 17, 1, pp. 187-198, 1981.

Kocijan, J.; Girard, A.; Banko, B. & Murray-Smith, R. Dynamic systems identification with Gaussian processes. *Mathematical and Computer Modelling of Dynamical Systems*, 11:4, 411-424, 2007. DOI: 10.1080/13873950500068567

Kocijan, J. Modelling and Control of Dynamic Systems Using Gaussian Process Models. Advances in Industrial Control, Springer International Publishing, 2016, DOI 10.1007/978-3-319-21021-6_2.

Ljung, L. *System Identification: Theory for the User.* [Ljung, 1999] Upper Saddle River, New Jersey: Prentice-Hall. ISBN 0-13-656695-2, 1999.

Ljung, L. Identification of Nonlinear Systems. In *Proceedings of the International Conference on Control, Automation, Robotics and Vision*, 2006.

Mahfouz, A.B. Identification of the nonlinear ship rolling motion equation using the measured response at sea. *Ocean Engineering*, 31, pp. 2139-2156, 2004.

Moreno-Salinas, D.; Chaos, D.; de la Cruz, J.M.; & Aranda, J. Identification of a Surface Marine Vessel Using LS-SVM. *Journal of Applied Mathematics*, vol. 2013, Article ID 803548, 11 pages, 2013. doi:10.1155/2013/803548.

Moreno-Salinas, D.; Chaos, D.; Besada-Portas, E.; Lopez-Orozco, J. A.; de la Cruz, J.M.; & Aranda, J. Semiphysical Modelling of the Nonlinear Dynamics of a Surface Craft with LS-SVM. *Mathematical Problems in Engineering*, Volume 2013 (2013), Article ID 890120, 13 pages http://dx.doi.org/10.1155/2013/890120.

Moreno-Salinas, D.; Chaos, D.; Besada-Portas, E.; Lopez-Orozco, J. A.; de la Cruz, J.M.; & Aranda, J.. Symbolic Regression for Marine Vehicles Identification. *10th IFAC Conference on Manoeuvring and Control of Marine Craft MCMC 2015*, Volume 48, Issue 16, 2015, Pages 210-216, https://doi.org/10.1016/j.ifacol.2015.10.282.

Narendra; K.S. & Parthasarathy, K. Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, vol.1, pp. 4-26, January 1990.

Nouretdinov, I., et al. Nouretdinov, I., Vovk, V., Vyugin M., and Gammerman, A. *Pattern recognition and density estimation under the general i.i.d assumption.* Computational Learning Theory, Proceedings, 2001. 2111: pp. 337-353.

Nouretdinov, I., Melluish, R., and Vovk, V. *Ridge Regression Confidence Machine.* In Proceedings of the Eighteenth International Conference on Machine Learning, 2001: pp. 385-392.

Pan, Yongping & Yu, Haoyong Biomimetic Hybrid Feedback Feedforward Neural-Network Learning Control. *IEEE Transactions on Neural Networks and Learning Systems.*, 28. 1-7. 10.1109/TNNLS.2016.2527501.

Papadopoulos, H., Vovk, V., and Gammerman, A. Regression Conformal Prediction with Nearest Neighbours. *Journal of Artificial Intelligence Research*, 2011. 40: pp. 815-840.

Perez, T. Ship Motion Control: Course Keeping and Roll Stabilisation Using Rudder and Fins. Springer-Verlag London, 2005, DOI:10.1007/1-84628-157-1

Perez, T., Sørensen, A.J., & Blanke, M. Marine Vessel Models in Changing Operational Conditions - A Tutorial. *14th IFAC Symposium on System Identification*, Newcastle, Australia, 2006.

Pillonetto, G.; Dinuzzo, F.; Chen, T.; De Nicolao, G. and Ljung, L. *Kernel methods in system identification, machine learning and function estimation: A survey.* 2014, Automatica, (50), 3, 657-682.

Rasmussen, C.E. & Williams, C.K.I. Gaussian Processes for Machine Learning. MIT Press 2006.

Rubio, J. de J. Stable Kalman filter and neural network for the chaotic system identification. *Journal of Franklin Institute*, Volume 354, Issue 16, November 2017, Pages 7444-7462. DOI: https://doi.org/10.1016/j.jfranklin.2017.08.038

Saunders, C., Gammerman, A. and Vovk, V. Ridge Regression Learning Algorithm in Dual Variables *Proceedings of the Fifteenth International Conference on Machine Learning.* 1998, Morgan Kaufmann Publishers Inc. pp. 515-521.

Schölkopf, B. & Smola, A. J. *Learning with Kernels.* Cambridge, Massachusetts: MIT press. ISBN 0-262-19475-9, 2002.

Shafer, G. and Vovk, V. *A Tutorial on Conformal Prediction.* J. Mach. Learn. Res., 2008. 9: pp. 371-421.

Vapnik, V. N. *The Nature of Statistical Learning Theory* Springer, 1995.

Velasco, F.J.; Revestido, E.; López, E. & Moyano, E. Identification for a Heading Autopilot of an Autonomous In-Scale Fast Ferry. *IEEE Journal of Oceanic Engineering*, vol 38, n 2, pp. 263-274. April 2013.

Vovk, V., Gammerman A. and Shafer G. *Algorithmic Learning in a Random World.* Springer, 2005.

Wang, X.D. & Ye, M.D. Nonlinear dynamic system identification using Least Squares Support Vector Machine Regression. In *Proceedings of the Third International Conference on Machine Learning and Cybernetics*, Shanghai, 2004.

Xu, F., Zou, Z.J., Yin, J.C., Cao, J. Identification modeling of underwater vehicles' nonlinear dynamics based on support vector machines. *Ocean Engineering*, Volume 67, 15 July 2013, Pages 68-76, ISSN 0029-8018, http://dx.doi.org/10.1016/j.oceaneng.2013.02.006. (http://www.sciencedirect.com/science/article/pii/S0029801813000863)

Luo, W.L. & Zou Z.J. Parametric identification of ship manoeuvring models by using Support Vector Machines. *Journal of Ship Research*, Vol. 53, No. 1, pp. 19-30, 2009.

Zhang, X.G. & Zou Z.J. Identification of Abkowitz model for ship manoeuvring motion using ϵ-Support Vector Regression. *Journal of Hydrodynamics*, 23(3):353-360, 2011.