A SVM-based Sensitivity Analysis Approach for Data-Driven Modeling of Ship Motion

Chunlin Wang, Xu Cheng, Shengyong Chen
School of Computer Communication and Engineering
Tianjin University of Technology
Tianjin, 300384 PR China
csy@tjut.edu.cn

Guoyuan Li, Houxiang Zhang
Department of Ocean Operations and Civil Engineering
Norwegian University of Science and Technology
Aalesund, 6009 Norway
{guoyuan.li & hozh}@ntnu.no

Abstract

This paper presents a novel method that combines support vector machine (SVM) with sensitivity analysis (SA) to analyze sensor data for ship motion modeling. In order to investigate how each model input contributes to the model's output, the PAWN method which is based on cumulative distribution function (CDF) is used. Considering the limitation of the PAWN method that it cannot be applied to sensor data directly, a surrogate model using SVM is integrated into PAWN method as it has a better solution of non-linearity and high-dimension problem as well as excellent generalization capability. Implementation of whole systematic workflow is elaborated in detail and related experiments are made such as comparison of four distance metric methods, benchmark test and SA of vessel data. The results show the proposed method can be used for analyzing vessel data to seek prime parameters that affect ship motion.

Index Terms

Sensitivity Analysis, PAWN method, Support Vector Machine, Ship Motion Modeling

I. Introduction

As an increase of maritime operations, the reliability and safety of ship maneuverability become of great concern in the maritime domain. The motion of ship on the open sea is determined by various factors, mainly including human factors, wind, waves and other external disturbances [1], [2]. In fact, to guarantee ship safety, a ship is equipped with many sensors, such as lidar, radar, sonar and GPS/INS for operation and navigation. The data collected by these sensors for a long time form the data set in terms of big data [3], involving various parameters of internal apparatus in ship and external environment. These raw data are characterized by high dimension and nonlinearity mixed up with noisy of various frequencies and redundant information. Given these characteristics of sensor data, it is necessary to analyze the ship data to find the most influential factors that affect the ship motion. Furthermore, modeling ship motion based on the most significant factors is beneficial for both control and prediction of ship motion, especially for ships operating in complex scenarios.

Sensitivity analysis (SA) method is an efficient tool to find that the variation of model output is apportioned to every model input factors in order to allow decision makers to sharpen their view of the problem [4]. With the help of SA, one can acquire useful information and decision support from complicated sensor data. SA has been widely used in the domain of maritime. Li et al. applied SA to find out how much influence input parameters on the output of ship position for ship motion prediction [5]. Cheng et al. utilized variance-based SA to find out the influence of input parameters on the output of ship heading [6]. Zhang et al. utilized the sum of square derivatives (SSD) to select the network inputs for NARMAX model, which was also used for ship motion prediction [7]. Shenoi et al. conducted SA of all hydrodynamic derivatives in a four-DOF of container ship on the basis of simulated data to study the hydrodynamic coefficients. The results are very important to the ship motion [8]. To investigate the various performance parameters affecting the propulsion and maneuvering abilities of a ship, a SA was proposed by Panagiotis [9].

In the previous work, SA is highly dependent on a model in practical application. In the case of only sensor data available, it is a challenge to employ the SA directly to the vessel data. The surrogate model provides us a potential solution to overcome the challenge. The surrogate model has some advantages with cheap computational cost and good prediction capabilities, compared with the original model. The surrogate model can learn from the historical sensor data to find the inherent relationship between model output and input. The widely used surrogate models including polynomials, splines, Gaussian Processing (GP) [10], SVM and Artificial Neural Network (ANN) [11], have been effectively used for practical SA. However, polynomials, spline, and GP are not always efficient, especially in the case of complex nonlinear phenomena. Compared with ANN, the SVM can avoid falling into a local optimal solution. Moreover, overfitting is unlikely to occur with SVM [12].

Our on-going project is to build an intelligent system for decision makers having a good command of ship operation. The project is supported by mechatronics lab, at Department of Ocean Operations and Civil Engineering (IHB), NTNU, Aalesund. This new platform, based on data analysis tool and data-driven modeling, can offer real-time autonomous and semi-autonomous control of ship maneuverability. In this paper, a fitted and validated surrogate model for representing ship motion is constructed on the basis of sensor data that are observed in simulated vessels. A hybrid approach to integrate an SVM and CDF-based PAWN method is proposed. The SVM is used for constructing a surrogate model. Compared to other SA methods, CDF-based method has distinctive advantages, including low computational cost of approximation of CDFs and SA indices. So PAWN method is selected to conduct SA on vessel data. The contribution of this paper includes: first, propose a more appropriate method that is capable of coping with the SA on the sensor data; second, the state of practice by providing a data-driven surrogate model that can learn new knowledge from historical sensor data to simulate the system behavior of ship motion.

This paper is structured as follows. In the next section, the system structure of the proposed method is introduced to understand the framework of SA on vessel model and concrete implementation of every part such as data pre-processing, modeling and data analysis. After that, related experiments are shown to confirm the idea in the next section and conclusions are given in the end.

II. System Structure

The whole work-flow includes three parts as shown in Fig. 1. The first part is data pre-processing. The purpose of pre-processing is to remove the effect of noisy data, which can ensure the precise of modeling and SA result. The meanings of data parameters are described in TABLE I and the three thrusters are fixed on the ship as shown in Fig. 2. The second part is modeling. When the raw data are processed after the first step, the SVM is selected to construct the vessel model, which is an important foundation for the following operation. The last part is data analysis in which sensitivity index of every input parameters can be calculated with the help of PAWN method. After that, one can have an intuitive view of how much effect of every parameter on the variation of model output is.

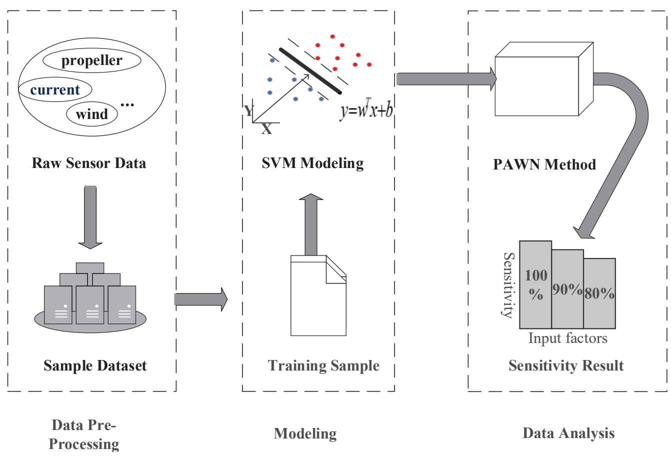

Fig. 1. The system structure of SA on vessel data

A. Data Pre-Processing

Data pre-processing involves data cleaning and normalization. Due to raw data collected from sensors containing noisy and redundant information, it is inevitable to generate errors if these data are directly utilized to construct a model. So hierarchical clustering which is widely used in data mining and statistics, is considered as a method of removing noisy. This is a "bottom-up" approach: each observation is seen as one cluster, and pairs of clusters are merged as one moves up the hierarchy. The result of clustering is shown in a dendrogram [13]. If the distance between one observation and its nearest neighbor cluster is below the threshold set by user, this observation will be seen as noisy data, then it is deleted. Trough repeated clustering, the number of noisy data will gradually decrease but not to zero. Then we can get sample data consisting of useful information. Further, the sample data need to be normalized in the range $[0, 1]$ by (1)

$$x_i = \frac{x_i - min(X)}{max(X) - min(X)} \quad i = 1 \dots n \tag{1}$$

where $X = [x_1, x_2, \dots, x_n]$. After the procedure of data pre-processing, SVM can use these sample data to train vessel model.

B. Modeling

Our goal is to compute sensitivity index of every input parameter to construct a compact data-driven model. In this paper, SVM is selected as a tool to construct a surrogate regression model because of its high generalization, a better solution of non-linearity and high-dimensionality, avoidance of local optimal solution in ANN, which is explained in detail by [14].

TABLE I
THE MEANING OF INPUT FACTORS

	inputs	unit
Ship-environment status	surge_vel	[m/s]
	sway_vel	[m/s]
	yaw_vel	[deg/s]
	roll_vel	[deg/s]
	pitch_vel	[deg/s]
	roll	[deg]
	pitch	[deg]
	heading	[deg]
Thruster1	percent	[%]
	rmp	[RPM]
	pitch_angle	[deg]
	force	[N]
	yaw_moment	[Nm]
	consumed_power	[W]
Thruster2	percent	[%]
	rmp	[RPM]
	pitch_angle	[deg]
	force	[N]
	yaw_moment	[Nm]
	consumed_power	[W]
Thruster3	percent	[%]
	rmp	[RPM]
	pitch_angle	[deg]
	force	[N]
	yaw_moment	[Nm]
	consumed_power	[W]

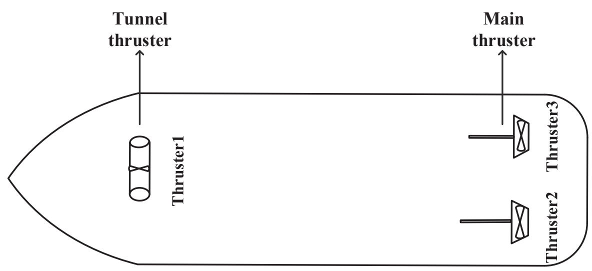

Fig. 2. The position of thrusters on ship

For *m* training samples (x_i, y_i) where $x_i \in R^k$ *and* $y_i \in R$ The basic form of SVM is as follows:

$$\begin{aligned} &\min_{\omega,b} \frac{1}{2}||\omega||^2 \\ &s.t.\ y_i(\omega^T\phi(x_i)+b) \geq 1,\ i=1,2,\dots,m \end{aligned} \tag{2}$$

Note that $\phi(x)$ is eigenvector by which x is mapped into high dimensional space, $\phi(x)^T$ is the transpose of $\phi(x)$, ω is normal vector of a plane, b is an offset. Lagrange multiplier method is used for transforming (2) to its dual problem

$$\begin{aligned} &\max_{\alpha} \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j \phi(x_i)^T\phi(x_j) \\ &s.t. \begin{cases} \sum_{i=1}^{m}\alpha_i y_i = 0 \\ \alpha_i \geq 0,\ i=1,2,\dots,m \end{cases} \end{aligned} \tag{3}$$

With the help of kernel function $K(x_i, x_j) = \phi(x_i)^T\phi(x_j)$, it makes a great reduction of computational cost owing to avoidance of gain the concrete feature vector and mapping function. After α is calculated from solving (3) , the model can be obtained as follows:

$$\begin{aligned} f(x) &= \omega^T\phi(x)+b \\ &= \sum_{i=1}^{m}\alpha_i y_i K(x, x_i) + b \end{aligned} \tag{4}$$

In addition to regression and time series prediction applications, excellent performances were received by [15], [16]. Given these advantages, SVM will be a preferable choice to construct vessel model.

The sensor data which after pre-processed, are divided into train samples (80%) and test samples (20%). The Gaussian function is the kernel function in this section. The best parameters of Gaussian function would be selected through trial and error methods.

C. Data Analysis

SA involves variance-based and density-based method [17]. The variance-based method does not completely represent the output uncertainty when the model output is highly-skewed or multimodal [18]. To overcome these defects, a new method called moment-independent GSA method, also known as the density-based method, was proposed by [19], [20]. Sensitivity indices computed by the density-based method in the view of whole density distribution function instead of a specific moment of distribution, which is independent of the shape of output distribution. Although the density-based method has made up for defect of the variance-based method, the development of it has been limited because of the high computational cost of estimating probability density function (PDF).

To find a better solution of high computational cost in estimating PDF, a new GSA method, called PAWN method, was proposed that characterizes the distribution function by CDF instead of density distribution function in order to solve the problem in the density-based method. One merit is that the approximation of CDF from data sample has no need to compute unknown parameters, therefore, achieve a big reduction of computational cost. Another merit is that sensitivity indices can be easily acquired either considering the entire range of variation of the model output or just a sub-range. So it is very liable to implement and facilitates the practical application of data SA. Given the merits of PAWN against the defects of the variance-based and density-based method, the PAWN method is employed to conduct SA on vessel data.

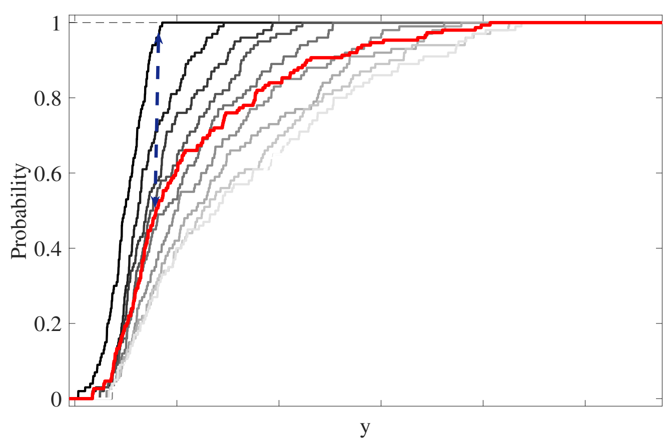

Fig. 3. Unconditional and conditional CDFs of output

The basic idea of the PAWN method is to quantify the discrepancy or compute the distance between unconditional and conditional CDFs. As is shown in Fig. 3, 'y' denotes model output, the red curve represents unconditional CDF and those grey curves indicate conditional CDFs where one of input parameters is fixed. The blue dash represents the discrepancy of the two CDFs, the longer it is, the larger discrepancy between two CDFs is, i.e. the bigger sensitivity index of x_i is. The pattern of calculating distance between two CDFs is various from different statistic methods. Four distance metric methods (Kolmogorov-Smirnov, K-sample Aderson-Darling, Friedman, Euclidean distance) those are widely used in statistic domain can be adopted to calculate the distance.

The Kolmogorov-Smirnov (KS) test is based on the largest difference between unconditional and conditional CDFs in PAWN. The formula is shown as follows:

$$KS(x_i) = \max_{y} |F_y(y) - F_{y|x_i}(y)| \tag{5}$$

where $F_y(y)$ is unconditional CDF and $F_{y|x_i}(y)$ is conditional CDF that x_i is fixed.

According to [21], K-sample Aderson-Darling (K-AD) test is the most powerful CDF tests, which verifies whether K distribution functions come from the same one. That is, The bigger value computed by AD test, the larger discrepancy is between two CDFs. As there are one unconditional CDF and conditional CDFs at ten points, These eleven CDFs are tested by AD to get a value of x_i, if the value is the biggest one, the discrepancy of CDFs is largest than any one, which means the sensitivity indices of x_i is maximum.

The Friedman test is a non-parametric statistical test developed by Milton Friedman. It uses ranks to detect whether there is a significant difference among on multiple overall distributions. So it is considered to compute PAWN sensitivity indices.

Euclidean distance is applied to computing the distance between two points in m-dimensional space. The CDF is linked by countless two-dimensional points. So the distance between two CDFs can be solved by computing that of all these points. Given two vector $p = (p_1, p_2, \dots, p_n)$ and $q = (q_1, q_2, \dots, q_n)$, the distance (D) from p to q is computed by (6).

$$D(p, q) = \sqrt{(p_1 - q_1)^2 + \cdots + (p_n - q_n)^2} \tag{6}$$

To the best knowledge of the author, There is no detailed comparison of the application of four distance metric methods to SA. In our paper, the four distance metric methods will be employed to calculate SA index, and the performance will be compared as well.

III. Experiments

This section mainly refers to three relevant experiments which are used in verifying the feasibility of the method we proposed. The four distance metric methods are compared on the basis of benchmark function to test the ability to compute the discrepancy between the unconditional and conditional distribution function, which is shown in the first experiment. The second experiment demonstrates that PAWN incorporating in SVM can obtain the desired result. The last experiment shows a case study of SA on ship sensor data for modeling of ship heading.

A. Comparison of Four Distance Metric Methods

This experiment is mainly to determined which statistic method is suitable to compute the discrepancy between unconditional and conditional CDFs. Ishigami function is selected as test function because it is widely used in many fields as benchmark [22]:

$$y = sin(x_1) + asin(x_2) + bx_3^4 sin(x_1) \tag{7}$$

where x_i follows a uniform distribution over$[0, 1]$, and $a = 7$, $b = 0.1$. The utilization of PAWN to compute sensitivity indices has been elaborated in II-C in detail. Owing to Kolmogorov-Smirnov, Friedman and Euclidean distance methods only computing the distance between two CFDs simultaneously, therefore, when the distance between unconditional CDF and each of conditional CDFs is computed by these three methods for ten times, the method of processing ten distance values of every input factors contains many manners such as min, max, mean. K-sample Aderson-Darling is no need to do this according to its principle of calculation. In the light of different manners, related result of SA on Ishigami function are listed in TABLE II. '$x_2 \to x_1 \to x_3$' denotes sensitivity index of x_2 is maximum, x_1 following, x_3 minimum. The result was concluded as same as benchmark except Friedman test when measuring manners is min.

TABLE II
THE RESULT OF FOUR STATISTIC TESTS USING DIFFERENT MEASURING MANNERS

	mean	max	min
K-AD	$x_2 \to x_1 \to x_3$		
KS	$x_2 \to x_1 \to x_3$		
Euclidean	$x_2 \to x_1 \to x_3$		
Friedman	$x_2 \to x_1 \to x_3$		$x_1 \to x_2 \to x_3$

From the comprehensive comparison of these statistic methods, these four distance metric methods can come to a correct conclusion of SA on Ishigami function in case of only considering rank of sensitivity indices of three factors. For Aderson-Darling test, as size of sample growing, the difference of curves of x_1 and x_3 is not conspicuous while Friedman test does not converge to a constant value. So these two methods can not be employed as measuring discrepancy among CDFs in PAWN. When training sample attains 1200, the convergence of sensitivity indices computed by Euclidean distance is ideal for us. But its training cost is high than others. On the one hand, sensitivity indices calculated by Kolmogorov-Smirnov method range from 0 to 1 and are directly computed from CDFs without any other computational cost. This method is easy to implement. On the other hand, the speed of PAWN index convergence is quicker than others when training sample attains 200. Yet there is still a minor flaw that the sum of sensitivity indices computed by Kolmogorov-Smirnov may be bigger than 1 because this is determined by its principle. Considering above comparison, Kolmogorov-Smirnov is widely applied to computing sensitivity indices of input factors.

Fig. 4 displays the sensitivity indices of input factors in Ishigami function which computed by Kolmogorov-Smirnov test when measuring manner is selected as max. In this figure, the value on X-axis denotes the logarithm of sample size. The size of sample ranges from 10 to 1400.

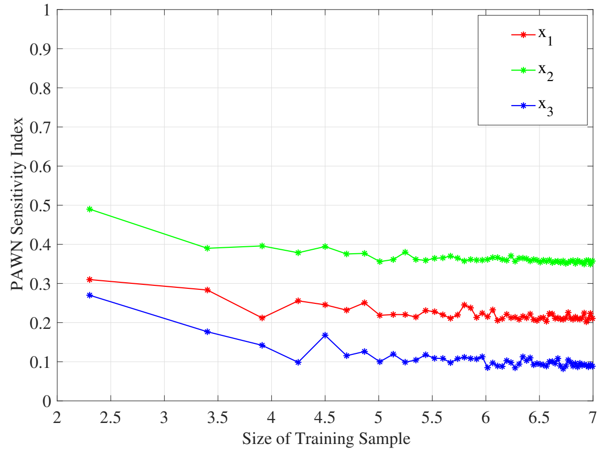

Fig. 4. PAWN sensitivity indices of Ishigami function computed by Kolmogorov-Smirnov test.

B. Benchmark Test

The feasibility of the proposed method is also tested with the Ishigami function. The benchmarks of three factors are sensitivity indices of Ishigami function computed by PAWN method. The size of sample, used for estimating unconditional and conditional CDFs, is fixed at 150 that has explained in III-A. Following the procedure of computing PAWN sensitivity indices in II-C, the result is acquired as three red dashes shown in Fig. 5. In this figure, the three lines, including triangle, square and circle, are obtained by prediction of SVM while three red dotted lines are true value computed by Ishigami function respectively. Then, SVM model is trained with input and output sample data which are previously used for computing benchmarks. In process of computing sensitivity indices with PAWN method, the size of sample is as much as that of computing benchmark and output value is obtained by model evaluation trained by SVM instead of Ishigami function. The result is shown in Fig. 5 whose three lines-triangle, square, circle- denote predicting value. After training sample attains 500, two value reach to consistency. That illustrates the same result can be achieved by SVM in contrast to Ishigami function. That means SVM can be used in training model with existing sample data and make a solid underpinning for SA on vessel data.

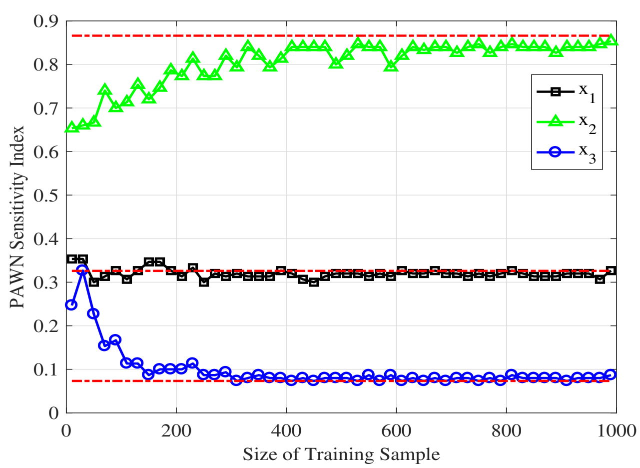

Fig. 5. Comparison of PAWN sensitivity index based on benchmark to that based on SVM.

C. SA of Vessel Data

Given feasibility of PAWN integrated with SVM to carry on SA, it is applied into executing SA on vessel model in which heading is chosen as a output parameter. The heading is within $[0°, 360°]$ originally. The remaining are model input parameters. The whole flow has been described in II-C. The result of SA of vessel data are achieved as in Fig. 6 shown. Note that these parameters with sensitivity index lower than 0.05 are not shown in Fig. 6. The size of conditional sample was fixed at 7800 at which the sensitivity index of each factor reaches to convergence. This figure displays ratio of roll_vel to total effect attains 12.91%. That means rolling velocity is the most influential parameter for variation of heading. Sway_vel is next biggest accounting for 11.15%. Other factors indicates the sum of proportion, because the each proportion of other 17 factors is so small to regard its effect. As we can see from Fig. 6, though every of other 17 parameters has little effect on heading angel, the sum of effect of these parameters is far bigger than any one parameter.

In addition, the proposed method in this paper is compared with Extended Fourier Amplitude Sensitivity Test (EFAST)

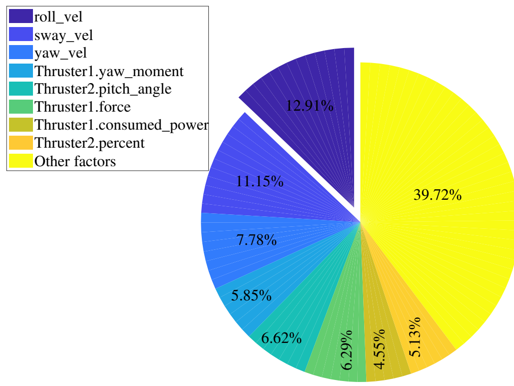

Fig. 6. The pie chart of SA on ship heading.

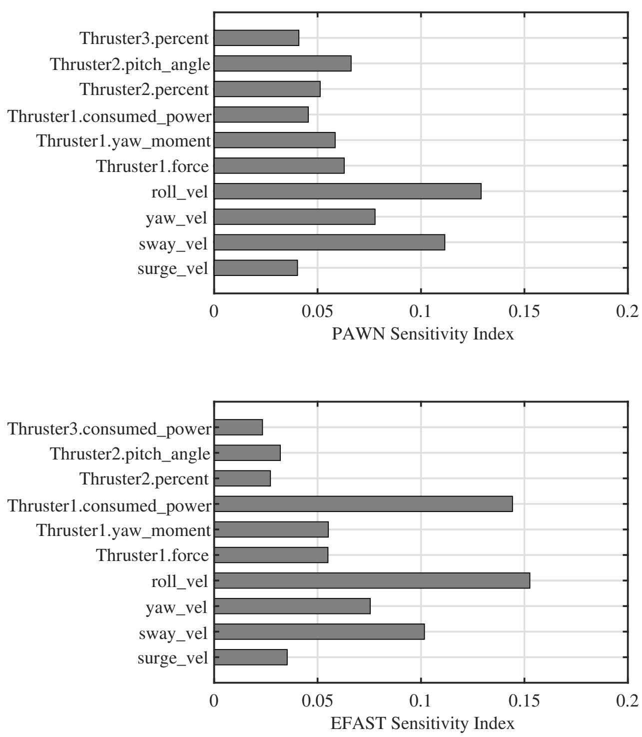

Fig. 7. Comparison of SA on ship heading based on PAWN to that based on EFAST.

integrated with SVM. The top ten significant parameters are only displayed in the Fig. 7. Vertical axis lists 10 input parameters while horizontal axis represents the degree of impact each input parameter on variation of heading. The top ten of influential parameters computed by PAWN method resemble them computed by EFAST method except Thruster3.percent, Thruster3.consumed_power and value of Thruster1.consumed_power. The most influential parameter computed by PAWN is the same as that by EFAST method. Although the rank of five inputs by PAWN including Thruster1.comsumed_power, Thruster1.yaw_moment, Thruster1.force, yaw_vel and sway_vel are different from the result from EFAST, they are found having similar values of sensitivity index. The differences may be caused by the different principles between PAWN and EFAST, i.e., the former is based on CDF, whereas the latter based on variance. Despite of the difference, PAWN method can correctly find influential parameters from all of them against EFAST method. From above comparison, as a whole, SA based on PAWN method can be applied for analyzing vessel data. In future work, the proposed method will be compared with ANN-EFAST method described in [6] to verify the excellent performance of conducting SA on vessel data.

IV. Conclusion

In this paper, the method that PAWN integrated with SVM is proposed to be applied for conducting SA on ship sensor data. In order to verify the feasibility of this method, three parts of work are made: (1) The best statistic method is selected to compute the difference of unconditional and conditional CDFs from the comparison of four distance metric methods. The result shows that the Kolmogorov-Smirnov test outperforms in the four methods. (2) The benchmark test validated the feasibility of the proposed method used for conducting SA. (3) The result of SA on the ship sensor data shows that this method can find those influential factors that have an effect on the heading of the ship. In addition, the results achieved by the SVM-EFAST method are almost the same as those from the proposed method, indicating that the rationality and feasibility of the proposed method is scientifically verified to a great extent.

Future work will be taken in two aspects. On the one hand, experiments comparison between ANN and SVM (like time consumption and analysis accuracy) on SA framework will be conducted for comparative study. On the other hand, the current work does not take environmental changes into account. Actually, with the changes of environmental effects, SA will become time-dependent. Therefore, our future work will focus on investigating how to extend the proposed method to adapt to such external disturbances.

References

[1] T. I. Fossen, *Marine Control Systems: Guidance, Navigation, and Control of Ships, Rigs and Underwater Vehicles*, 2002.

[2] A. J. Srensen, "A survey of dynamic positioning control systems," *Annual Reviews in Control*, vol. 35, no. 1, pp. 123–136, 2011.

[3] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, "Big data: Issues and challenges moving forward," in *Hawaii International Conference on System Sciences*, 2013, pp. 995–1004.

[4] A. Saltelli, *Global sensitivity analysis : the primer*. John Wiley, 2008.

[5] G. Li, H. Zhang, B. Kawan, H. Wang, O. L. Osen, and A. Styve, "Analysis and modeling of sensor data for ship motion prediction," in *Oceans*, 2016, pp. 1–7.

[6] X. Cheng, G. Li, S. Chen, C. Diao, M. Liu, and H. Zhang, "Simplifying neural network based model for ship motion prediction: a comparative study of sensitivity analysis," in *ASME 2017 International Conference on Ocean, Offshore and Arctic Engineering*, 2017.

[7] W. Zhang and Z. Liu, "Real-time ship motion prediction based on time delay wavelet neural network," *Journal of Applied Mathematics*, vol. 2014, 2014.

[8] R. R. Shenoi, P. Krishnankutty, and R. P. Selvam, "Sensitivity study of hydrodynamic derivative variations on the maneuverability prediction of a container ship," in *ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering*. American Society of Mechanical Engineers, 2015, pp. V007T06A008–V007T06A008.

[9] P. Mizythras, E. Boulougouris, A. Priftis, A. Incecik, O. Turan, and D. Reddy, "Sensitivity analysis of the tool for assessing safe manoeuvrability of ships in adverse sea conditions," in *International Conference on Shipping in Changing Climates 2016*, 2016, pp. 1–13.

[10] J. E. Oakley and A. O'Hagan, "Probabilistic sensitivity analysis of complex models: a bayesian approach," *Journal of the Royal Statistical Society*, vol. 66, no. 3, p. 751769, 2004.

[11] I. Guevara, M. Gutierrez, and P. Zuniga, "Identification of weak buses for proper placement of reactive compensation through sensitivity analysis using a neural network surrogate model," in *IEEE International Autumn Meeting on Power, Electronics and Computing*, 2016, pp. 1–6.

[12] N. Cristianini and J. Shawe-Taylor, "An introduction to support vector machines," 2000, pp. 1–28.

[13] G. J. Szekely and M. L. Rizzo, "Hierarchical clustering via joint between-within distances: Extending ward's minimum variance method," *Journal of Classification*, vol. 22, no. 2, pp. 151–183, 2005.

[14] B. Kawan, H. Wang, G. Li, and K. Chhantyal, "Data-driven modeling of ship motion prediction based on support vector regression," in *Conference on Simulation and Modelling*, 2017, pp. 350–354.

[15] M. O. Stitson, A. Gammerman, V. Vapnik, V. Vovk, C. Watkins, and J. Weston, "Support vector regression with anova decomposition kernels," in *Advances in kernel methods*, 1999, pp. 285–291.

[16] Mattera, Davide, Haykin, and Simon, "Support vector machines for dynamic reconstruction of a chaotic system," 2008.

[17] F. Pianosi, K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, and T. Wagener, "Sensitivity analysis of environmental models," *Environmental Modelling Software*, vol. 79, no. C, pp. 214–232, 2016.

[18] F. Pianosi, F. Sarrazin, and T. Wagener, *A Matlab toolbox for Global Sensitivity Analysis*. Elsevier Science Publishers B. V., 2015.

[19] E. Borgonovo, "A new uncertainty importance measure," *Reliability Engineering System Safety*, vol. 92, no. 6, pp. 771–784, 2007.

[20] E. Plischkeabc, "Global sensitivity measures from given data," *European Journal of Operational Research*, vol. 226, no. 3, pp. 536–550, 2013.

[21] M. Arshad, M. T. Rasool, and M. I. Ahmad, "Anderson darling and modified anderson darling tests for generalized pareto distribution," *Journal of Applied Sciences*, vol. 3, no. 2, pp. 85–88, 2003.

[22] K. Tang, P. M. Congedo, and R. Abgrall, "Adaptive surrogate modeling by anova and sparse polynomial dimensional decomposition for global sensitivity analysis in fluid simulation," *Journal of Computational Physics*, vol. 314, no. 3, pp. 557–589, 2016.