Ocean Engineering 162 (2018) 316–330

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Ship speed prediction based on full scale sensor measurements of shaft thrust and environmental conditions

Andreas Brandsæter[a,b,*], Erik Vanem[a,b]

[a] DNV GL, Veritasveien 1, N-1363, Høvik, Norway
[b] Department of Mathematics, University of Oslo, P.B.1053, Blindern, N-0316, Oslo, Norway

ARTICLE INFO

Keywords:
Ship speed estimation
Computational methods/numerical analysis
Sensor data analytics
Ship resistance and propulsion
Performance measure quantification
Statistical modelling
Energy efficiency

ABSTRACT

The primary goal of this study is to adapt and validate various regression models to predict a ship's speed through water based on relevant and available full scale sensor measurements from a ship, including measurements of external environmental forces. The wind is measured by on-board wind sensors, and the effect of the waves is measured by motion reference units (MRUs) installed on the ship, measuring motions in six degrees of freedom; three translational motions and rotations about these. Accurate speed estimates, which relate directly to the estimates of the propulsion efficiency, fuel efficiency and pollution, are vital to be able to optimize ship design and operation. We demonstrate how regression models such as linear regression, projection pursuit (PPT) and generalized additive models (GAM) can be easily implemented for this application. A simple regression model based on the well-established relationship between ship speed and shaft thrust represent a benchmark model towards which the other models are compared.

1. Introduction

Accurate estimates of ship propulsion and fuel efficiency are important to be able to optimize ship design and operation. Deviations between the measured ship speed and the speed predictions based on propulsion power and other internal and external conditions can be indicative of an anomaly, such as e.g. hull, propeller or engine damage. Furthermore, the effect of modifications can be quantified. This can include modifications of the ship hull, such as for example hull cleaning or bow optimization, installation of new equipment such as kites, fixed sails or batteries for machinery optimization, propeller optimization such as contra rotating propeller and various efficiency improvement devices, and operational optimization measures such as weather routing and trim and draft optimization. The logistics planning can also be optimized with accurate of time of arrival estimation.

Due to the complexity of a modern ship and its exposure to external factors such as wind, waves and currents, estimating the ship efficiency accurately is not an easy task. Various methods are described in literature and used by the industry. The methods can be divided into four main groups as suggested by (Petersen et al., 2012):

1. Traditional and standard series methods which typically rely on a set of parameters describing the hull (Savitsky, 1964; Øyan, 2012; Holtrop and Mennen, 1982),
2. regression based methods based on a set of sensor measurements (Petersen et al., 2012; Bocchetti et al., 2015; Mao et al., 2016),
3. direct model tests in test tanks (Chuang and Steen, 2011), and
4. computational fluid dynamics (CFD) (Peri et al., 2001; Sadat-Hosseini et al., 2013; Ozdemir and Barlas, 2016).

The methods are often ordered on a scale between methods that are governed by physical laws and empirical or data driven methods (sometimes referred to as black box methods) that are based on statistical inference of historical data. Due to the fact that the data driven methods require little knowledge of the physical system (Coraddu et al., 2017) and there is no need to manually build a model of the data, these methods can be easily implemented in marine operations; making such technologies a lean alternative to complex tailor-made analytics (Brandsæter et al., 2016). At the same time, the data driven methods can be unsatisfactory in terms of physical explanation and it might require a significant amount of data to be sufficiently accurate (Vanem et al., 2017; Petersen et al., 2012).

The primary goal is to survey various regression models to estimate the ship speed through water based on relevant and available sensor measurements of the shaft thrust and external environmental forces from wind, waves and currents. The wind is measured by on-board wind sensors, and the effect of the waves is measured by motion reference units (MRUs) installed on the ship, measuring motions in six

* Corresponding author. DNV GL, Veritasveien 1, N-1363, Høvik, Norway.
E-mail address: andreas.brandsaeter@dnvgl.com (A. Brandsæter).

https://doi.org/10.1016/j.oceaneng.2018.05.029
Received 8 March 2017; Received in revised form 23 April 2018; Accepted 14 May 2018
Available online 29 May 2018
0029-8018/

degrees of freedom; three translational motions and rotations about these, and the sea currents are incorporated in the measure of ship speed through water. The speed through water y relates directly to the propulsion efficiency which is commonly defined as $e_{prop} = \frac{y}{p}$, where p is the propulsion power. It is also linked directly to the fuel efficiency $e_{fuel} = \frac{y}{f}$, where f is the energy in the consumed fuel, as defined by (Petersen et al., 2012).

We work towards a better understanding of how the external conditions affect the ship's speed, propulsion and fuel efficiency, and aim to be able to quantify these effects. Several case studies using different methods are described in the recent literature, both with main focus on propulsion efficiency estimation (Petersen et al., 2012; Vanem et al., 2017; Chuang and Steen, 2011; Øyan, 2012; Holtrop and Mennen, 1982; Mao et al., 2016) and fuel efficiency and emission estimation (Trodden et al., 2015; Bialystocki and Konovessis, 2016; Coraddu et al., 2017; Rakke, 2016; Bocchetti et al., 2015).

2. Data description

This study is based on an extended version of the dataset used in (Vanem et al., 2017). For completeness, parts of the data description provided in (Vanem et al., 2017) is rendered in the following with minor modifications.

The dataset contains variables associated with the efficiency of the ship machinery system, such as the speed through water (knots), propulsion power [kW] and shaft thrust (N). The shaft thrust is assumed to be proportional to the propulsion power over speed over ground. Other variables included in the dataset are related to the ship's motions, wind speed relative to the ship and trim and draft. These variables represent external factors and are used to explain variation in the efficiency and ship speed.

The ship is installed with two motion reference units (MRUs) measuring the ship's motion in all six degrees of freedom (heave, surge, sway, roll, yaw and pitch). From the raw motion data recorded at 5 Hz various integrated parameters are stored every 30 s, calculated from the preceding 15 min time record of the motions. The integrated parameters reported by the system include the first five spectral moments of each motion signal (m_0, m_1, , m_3 and m_4), the mean, standard deviation, skewness and kurtosis of the signal as well as the maximum and minimum values during the time window. Also the spectral peak period T_p and zero crossing period T_z were recorded.

Out of these parameters many are not relevant for the present analysis and are not considered here. Besides, some of the parameters carry redundant information and can be derived from other parameters, such as the standard deviation of the signal $\sigma = \sqrt{m_0}$ and $T_z = T_{02} = \sqrt{m_0/m_2}$. We therefore limit ourselves to consider m_0 and T_z for each of the degrees of freedom. The zeroth spectral moment $m_0 = \sigma^2$ is the total energy of the motion spectrum and indicates the magnitude of the ship motion. Note that for a wave record the significant wave height is usually defined as $H_s = 4\sqrt{m_0}$. Likewise, T_z indicates the typical period of the different motions. Since the periodic ship motions are primarily an effect of the waves, both m_0 and T_z can be considered as proxies for the real wave conditions in the sense that m_0 will be proportional to the real significant wave height and T_z will be similar to the typical period of the wave field. Moreover, the ship response in the different degrees of freedom will to a certain extent reflect the wave direction relative to the ship.

In addition to the ship motions, representing the effect of the waves, the wind speed relative to the ship is recorded; the wind component perpendicular to the ship (Wind-y) and the wind parallel to the ship (Wind-x). Wind-x is defined so that a positive value means wind blowing in the same direction as the ship speed. Two other parameters that are important for the hydrodynamic resistance of a ship is the draft and trim, which have also been recorded and included in the present analyses. The draft is defined as the vertical distance between the waterline and the bottom of the hull and is naturally related to the cargo level of the ship, while the trim is the difference between the forward and aft drafts.

For the analysis presented in this paper the original 30 s data were down-sampled to 5 min resolution by calculating the average values within each 5 min window. This makes the dataset smaller and easier to handle, and reduces the time dependency.

The data have been collected from an ocean-going ship over approximately 10 months in normal operation. Due to data quality issues, a large fraction of the data were removed in initial cleaning and outlier removal. For example when the difference between the measured speed through water and speed over ground is significantly larger than reasonable current, at least one of the measured speed sensors must report wrong values. Although the measurement of the speed through water is known not to be most reliable (van den Boom and Hasselaar, 2014), it is difficult to know which reading is wrong, hence we remove the data point. After the initial filtering and outlier removal, the dataset used in the analysis contains about 33000 data points, which correspond to about 115 days of data.

Initially, 18 selected variables are included in the analysis. Trace plots of the data are shown in Fig. 1. The upper plot shows the speed and thrust data series, the next shows all the ship motion data and the two lower plots show the wind and the trim and draft data, respectively. Each point on the horizontal axis represents the average sensor value in the previous 5 min.

A correlation plot showing the linear correlation between the various variables in the data set is shown in Fig. 2. It is seen that the a highly correlated structure is present in the dataset.

3. Methodology

3.1. Regression models

We employ various regression models to describe and predict the data, i.e., linear regression models, generalized additive models (GAM) and projection pursuit regression (PPR) models. Other models including various regression trees and kernel density estimation were also explored to some extent. We were not able to tune these models to provide accurate predictions, hence they are omitted. In the following, a brief introduction to each of the models will be provided. Reference is made to textbooks such as (Hastie et al., 2009) for a more thorough introduction.

The response variable will be denoted Y and the explanatory variables will be denoted $\mathbf{X} = (X_1, X_2, \ldots, X_P)$. The basic problem is to construct a prediction rule for predicting Y conditioned on the explanatory variables based on a stochastic model on the form

$$Y = f(\mathbf{X}) + \varepsilon, \tag{1}$$

where ε represents stochastic white noise and is often modelled as a zero-mean Gaussian variable. Different models for the $f(\cdot)$ function give rise to different regression models. Assuming N observations, the observed values are y_j and $\mathbf{x}_j$ for $j = 1, \ldots, N$.

3.1.1. Linear regression models

In linear regression one assumes a linear model on the following form:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_P X_P + \varepsilon. \tag{2}$$

The error made in such predictions are referred to as the residuals, and the residual sum of squares (RSS) will be a function of the model parameters. It is defined as

$$RSS(\hat{\beta}) = \sum_{j=1}^{N} (y_j - \hat{y}_j)^2 \tag{3}$$

and the fitted model parameters $\hat{\beta}_p$, $p = 0, \ldots, P$ are estimated from the