The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Jonas Mellin
jonas.mellin@his.se

Sten F. Andler
sten.f.andler@his.se

University of Skövde
P.O. Box 408
SE-541 28 Skövde

Abstract

We use data-generated models based on data from experiments of an ocean-going vessel to study the effect of optimizing fuel consumption. The optimization is an add-on module to the existing diesel-engine fuel-injection control built by Q-TAGG R&D AB. The work is mainly a validation of knowledge-based models based on a priori knowledge from physics. The results from a simulation-based analysis of the predictive models built on data agree with the results based on knowledge-based models in a companion study. This indicates that the optimization algorithm saves fuel.

We also address specific problems of adapting data to existing machine learning methods. It turns out that we can simplify the problem by ignoring the auto-correlative effects in the time series by employing low-pass filters and resampling techniques. Thereby we can use mature and robust classification techniques with less requirements on the data to demonstrate that fuel is saved compared to the full-fledged time series analysis techniques which are harder to use. The trade-off is the accuracy of the result, that is, it is hard to tell exactly how much fuel is saved. In essence, however, this process can be automated due to its simplicity.

1 This report is the result of the project “System för bränslebesparing på stora fartyg” with Diarienummer 2013-00301 funded by Vinnova Forska & Väx 2013 program. The authors would like to thank Alexander Karlsson for his valuable input in terms of a proposal how to use random forest classification for this work as well as code implementing part of the solutions.

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels:
An experimental study

Introduction

The standard deviations on fuel consumption on ships in transit is normally in the range of 30-50%, while the effect of typical operational fuel efficiency methods might be well under 10% of the fuel average. Thus measuring the actual fuel saving effects of any fuel optimization method is a challenging task, due to a great number of uncontrollable variables (wind, waves, sea currents, ship load variations) that extensively affect fuel consumption. In cases where fuel optimization algorithms can be turned on and off at reasonably short intervals, statistical methods can be used to evaluate fuel consumption savings when using optimization against the baseline level when not using the optimization. These methods can then give independent support to optimization-based fuel saving estimations obtained with the physical model-based methods.

This report makes a statistical analysis of a fuel saving optimization using an anti-rolling principle developed by Q-TAGG R&D AB. The experiments have been done on the Maersk Kithira ship in two instances, in 2013 and 2014. This analysis was done by the Information Fusion Research Program, School of Informatics, at the University of Skövde.

2 Experimental Conditions

The fuel consumption of a ship depends on three types of factors: (a) technical factors such as engine efficiency and hull + propeller state, (b) environmental conditions such as wind, waves and sea currents, and (c) operational factors such as ship acceleration and heading control. This study assumes all technical factors to be constant during the experiment. This assumption is true for shorter time measurements and under the assumption that engines, propellers and hull have not suffered damage or breakdown, and have not been serviced or repaired, during the experiment. For experiments with longer durations, the hull and propeller may decrease ship performance due to sea fouling (accumulation of biologic material on the hull and on the propeller); however, in the addressed experiments these effects are known to be small and have not been considered. Environmental conditions are strongly influencing operational factors, for example, a ship sailing in strong waves might reduce its sailing speed, the ship can roll and pitch, and the heading control might work harder to compensate for deviations from wind and waves. The propeller slip measures the relative difference between the distances that the propeller would have covered if turned in a solid medium, relative to the actual distance of the propeller in water. A higher propeller slip indicates heavier propulsion conditions and thus higher fuel consumption. Further, ship rolling affects the ship heading, and thus the automatic heading control will increase the movement of the rudder.
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

During the experiments, the fuel optimization algorithm was turned on and off at regular time intervals. Besides fuel consumption and ship rolling, the optimization is expected to have effects on the propeller slip and on the rudder movements.

2.1 Optimization principle
The patented optimization principle employed by Q-TAGG R&D AB is based on influencing operational factors by reducing ship rolling via controlled engine-speed changes. These changes modify the hydrodynamics of the ship around its longitudinal direction and thus influence rolling in a way that reduces ship heading changes, rudder movement and hull friction with water. The commercial name of the product is Q-TAGG RollOut. When the optimization is activated, it is expected that ship rolling, fuel consumption, propeller slip and rudder movement are reduced compared with the case when the optimization is not activated.

2.2 Experimental data
An automated data collection device (Q-TAGG Energy Evaluation Device) samples data on the ship. The resulting log files are of two types: raw data: sampled at 250 ms intervals (high frequency sampling); and consolidated log data with samples at 15-minute interval (obtained from raw data, consolidated for manageable size).

The logged data is organized in columns. Each column has data originating from one of the following sources: raw data from physical sensors used by the energy optimization equipment; data from the ship's NMEA data network (the National Marine Electronics Association standard); and indirect measurements computed from raw data, such as propeller slip. The reliability of the data sources varies. For example, the sensors for sea depth below the ship and speed through water occasionally generate values that are implausible; these values have been corrected or removed from the log. In general, the measurement data range was verified for errors and physical correctness. In addition, due to measurement delays, the raw data on propeller slip has a change dynamics that is implausible in some samples. As shown below, the candidate methods in this report can handle uncertainty and missing values as long as the errors are distributed according to a Gaussian distribution; thus, implausible values are removed prior to applying the methods.

2.3 Models and propeller slip
Q-TAGG has developed models for fuel consumption using physical principles and statistical identification of the ship's parameters (MAN Report n.d.). The main physical principle is the Bernoulli law for dynamic pressure, \( p = 0.5 \cdot \text{density} \cdot \text{speed}^2 \), where the density in this case is for water or air. The power of the engine is proportional to the pressure and the fuel consumption is proportional to power times ship speed. Thus, the fuel consumption of the ship over time (in l/h) is proportional to the cube of

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

The speed and over distance (l/nm) it is proportional to the square of the speed. This report analyzes the experimental, described in section 2.2, independently of the underlying physical model. The results of this report are then compared to those predicted by the physical model.

2.4 Basis for models and measurements
The following documents have been used as a basis:

1. Basic principles of ship propulsion (MAN Report n.d.)
2. Using propeller for hull monitoring (Logan 2012)
3. Assessing effective power and fuel consumption (Borkowski, Kasyk & Kowalak 2011)
4. Measuring fuel efficiency (Baumel 2012)
5. Total solution approach to ship energy efficiency management (Ballou 2013)
7. Benchmarks and measures for fuel efficiency (Wigforss 2012)

3 Problem definition
Our aim is to employ statistical methods and machine learning techniques to evaluate if an optimization technique, Q-TAGG Rollout, improves fuel consumption, roll and propeller slip. For this evaluation, we use data-generated prediction models as a complement to the knowledge-based models briefly introduced in section 2.3.

3.1 Motivation
The motivation for the use of statistical methods and data-generated models based on machine learning techniques is that they rely on fewer assumptions on a priori knowledge and can thus be used to increase the validity of previous knowledge-based results as well as increase the credibility of the optimization technique. An additional benefit is that this can be viewed as a pre-study to an automated data collection and model generation to improve the knowledge of fuel consumption, roll and propeller slip during different environmental conditions.

3.2 Objectives
The objectives of this study are:

1. To resample data with sampling periods to fit the methods. All methods require equidistant sampling to be used and missing values be estimated (e.g., interpolated).

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

2. To analyze the experimental data by auto and cross correlation to see if the experimental data demonstrates expected features (e.g., less variability when the optimization technique is turned on compared to when it is turned off) or if there are significant deviations.

3. To generate predictive models by a machine learning technique (random forests).

4. To employ the predictive models to evaluate whether the experiments indicate that the optimization improves fuel consumption, roll, and propeller slip.

4 Method of Analysis

4.1 Overall approach

The overall approach is to analyze and preprocess experimental data so that it does not include implausible measurements, to handle missing values (e.g., use average, previous measurement etc.) and to make the format of the time series fit existing methods.

We employ the R tool (Kuhn 2008) (http://r-project.org), since it contains a lot of packages for statistics, machine learning etc. It is a mature, open source, community where researchers actively support packages.

4.2 Resampling of data

To resample the data, an R program was built. This program is configurable based on column so that each column can be resampled in a specific way. For example, to sample the speed for the resampling interval, it is possible to use the average of each measurement within the resampling interval; we assume that the process is sufficiently slow so that interpolation in the resampling interval is unnecessary. Another example is for the optimization on/off column, where the resampling can be based on a majority of the measurements in the resampling interval. In “Appendix B: Resampling code”, the R code for the resampling is found.

An example of resampling is equidistant samples is shown in Table 1 as computed from the raw data (Table 2). Here, engine speed (Eng.Spd) is computed as the average of readings in the interval and roll control optimization active (Roll.Active) is the majority of indicators in the interval.

---

2 If there are no measurements in the resampling interval, then we mark this with “no value” and process it later.

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

4.3 Analysis of auto and cross correlation

To analyze auto and cross correlation in the experimental data, the R functions “acf” (auto correlation function) and “ccf” (cross correlation function) were employed. See Appendix C: Auto and cross correlation code for example code.

Table 1: Resample data example

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-06-25 18:15:01</td>
<td>68.125</td>
<td>0.6</td>
<td>221.75</td>
<td>3856.4525</td>
<td>2.275</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2014-06-25 18:15:02</td>
<td>67.7</td>
<td>0.6</td>
<td>212.85</td>
<td>3703.535</td>
<td>2.85</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2014-06-25 18:15:03</td>
<td>66.7</td>
<td>0.6</td>
<td>202.3</td>
<td>3520.1</td>
<td>3.025</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2014-06-25 18:15:04</td>
<td>66.125</td>
<td>0.6</td>
<td>208.6</td>
<td>3629.9525</td>
<td>3.175</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2014-06-25 18:15:05</td>
<td>66.525</td>
<td>0.6</td>
<td>210.15</td>
<td>3656.4925</td>
<td>3.3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2014-06-25 18:15:06</td>
<td>66.9</td>
<td>0.6</td>
<td>214.066666666667</td>
<td>3703.23</td>
<td>3.066666666667</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2014-06-25 18:15:07</td>
<td>67.15</td>
<td>0.6</td>
<td>208.8</td>
<td>3612.3875</td>
<td>2.825</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Briefly, in auto correlation we correlate a column with itself with a lag starting from 0 and increasing. For each lag we consider how much the value of an attribute at time t correlates to its value at time t+lag.

As can be seen in Illustration 1, at lag 0, we initially have perfect correlation, and then the auto correlation falls quickly beneath 0.5 after 4 seconds of lag. That is, after 4 seconds, there is less correlation between values at time t and t+lag. Note that since the roll period of the ship is 15 seconds, we also see an expected increase in the correlation at a lag corresponding to the period. As can be seen in Illustration 2, without optimization, the auto-correlation drops even faster. After 4 seconds of lag, there is a negative correlation.


Table 2: Sample raw data

<table>
<thead>
<tr>
<th>Time</th>
<th>Eng Spd</th>
<th>Scav Air Pres</th>
<th>Fuel Cons l per nm</th>
<th>Fuel Cons l per h</th>
<th>Rudder Pos</th>
<th>Roll State</th>
<th>Roll Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>18:15:00.080</td>
<td>67.8</td>
<td>0.60</td>
<td>222.0</td>
<td>3862.96</td>
<td>1.9</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>18:15:00.331</td>
<td>68.2</td>
<td>0.60</td>
<td>224.6</td>
<td>3908.61</td>
<td>2.2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:00.587</td>
<td>68.2</td>
<td>0.60</td>
<td>223.0</td>
<td>3880.08</td>
<td>2.4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:00.852</td>
<td>68.3</td>
<td>0.60</td>
<td>217.4</td>
<td>3782.16</td>
<td>2.6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:01.103</td>
<td>68.3</td>
<td>0.60</td>
<td>216.1</td>
<td>3759.91</td>
<td>2.8</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:01.355</td>
<td>67.5</td>
<td>0.60</td>
<td>216.1</td>
<td>3759.91</td>
<td>2.8</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:01.617</td>
<td>67.5</td>
<td>0.60</td>
<td>215.1</td>
<td>3743.23</td>
<td>2.8</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:01.883</td>
<td>67.5</td>
<td>0.60</td>
<td>204.1</td>
<td>3551.09</td>
<td>3.0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:02.133</td>
<td>66.9</td>
<td>0.60</td>
<td>202.3</td>
<td>3519.19</td>
<td>3.1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:02.386</td>
<td>66.9</td>
<td>0.60</td>
<td>201.3</td>
<td>3503.24</td>
<td>3.0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:02.649</td>
<td>66.5</td>
<td>0.60</td>
<td>201.0</td>
<td>3497.93</td>
<td>3.0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:02.914</td>
<td>66.5</td>
<td>0.60</td>
<td>204.6</td>
<td>3560.04</td>
<td>3.0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:03.164</td>
<td>66.1</td>
<td>0.60</td>
<td>205.2</td>
<td>3570.83</td>
<td>3.1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:03.422</td>
<td>66.1</td>
<td>0.60</td>
<td>210.9</td>
<td>3669.83</td>
<td>3.2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:03.687</td>
<td>66.1</td>
<td>0.60</td>
<td>209.6</td>
<td>3647.83</td>
<td>3.2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:03.952</td>
<td>66.2</td>
<td>0.60</td>
<td>208.7</td>
<td>3631.32</td>
<td>3.2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:04.203</td>
<td>66.2</td>
<td>0.60</td>
<td>208.1</td>
<td>3620.32</td>
<td>3.4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:04.465</td>
<td>66.6</td>
<td>0.60</td>
<td>209.9</td>
<td>3651.90</td>
<td>3.2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:04.730</td>
<td>66.6</td>
<td>0.60</td>
<td>210.8</td>
<td>3668.55</td>
<td>3.3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:04.995</td>
<td>66.7</td>
<td>0.60</td>
<td>211.8</td>
<td>3685.20</td>
<td>3.3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:05.261</td>
<td>66.7</td>
<td>0.60</td>
<td>214.0</td>
<td>3701.85</td>
<td>3.2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:05.526</td>
<td>67.0</td>
<td>0.60</td>
<td>213.8</td>
<td>3698.40</td>
<td>3.1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:05.791</td>
<td>67.0</td>
<td>0.60</td>
<td>214.4</td>
<td>3709.44</td>
<td>2.9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:06.056</td>
<td>67.0</td>
<td>0.60</td>
<td>209.1</td>
<td>3617.81</td>
<td>2.9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:06.321</td>
<td>67.2</td>
<td>0.60</td>
<td>208.2</td>
<td>3601.54</td>
<td>2.8</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:06.587</td>
<td>67.2</td>
<td>0.60</td>
<td>208.2</td>
<td>3601.54</td>
<td>2.8</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18:15:06.852</td>
<td>67.2</td>
<td>0.60</td>
<td>209.7</td>
<td>3628.66</td>
<td>2.8</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

3 In the illustration, the x axis is labeled with number of 15 second period. So, 1 means 15 seconds, 2 means 30 seconds.
If we consider the features of these two auto-correlations (from time series 5), when optimization is turned on, the fuel consumption l/nm is smoother than if the optimization is turned off. That is, there are more positive correlation of values when optimization is turned on compared to when it is turned off, which implies that there are fewer or less radical changes in the fuel consumption in term of liter per nm.

Illustration 1: Auto correlation of time series 5 when optimization is turned on.
The cross-correlation between, for example, the fuel consumption in l/nm and speed over ground is negative, weak and not changing when optimization is turned on in time series 5 (see Illustration 3). This means that when speed over ground increases, in general, it is more likely that the fuel consumption l/nm decreases. This may seem to be counter-intuitive, but we conjecture that the reason is that the data we use is collected at speeds where increases in speed can reduce the fuel consumption in terms of l/nm since the measurements are assumed to be sufficiently accurate\textsuperscript{4} in this case. In contrast, without optimization, it is not as smooth (cf. Illustration 4).

\textsuperscript{4} The speed over ground is measured by using GPS and the fuel consumption in terms of liter per nautic mile are based on fuel consumption in terms of liter per second. Both of these sensors are assumed to be sufficiently reliable.

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Illustration 3: Cross-correlation between speed over ground and fuel consumption (l/nm) in time series 5 when optimization is turned on.

Illustration 4: Cross-correlation between speed over ground and fuel consumption (l/nm) when optimization is turned off.
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

In Appendix D: Auto and cross correlation results, you can find numerous interesting examples of auto and cross correlation. In general, when optimization is turned on, the auto-correlation and cross-correlation indicate that the time series are smoother. That is, there are fewer or less radical variations in the data when optimization is turned on compared to when it is turned off. In essence, optimization provides smoother time series that is likely to be beneficial for fuel saving.

The label of the graph denotes if it is auto-correlation (ACF) or cross-correlation. Then comes a number denoting what time series we are addressing (2, 5, 6 or 7) followed by optimization (yes/no), where any number but 0 mean that optimization is turned on and that 0 means that optimization is turned off. Finally, the name of the variable is specified (e.g., “Fuel.cons.l.per.nm”, which represents the fuel consumption in liter per nautical mile.

5 Data-generated prediction models: Application of random forest

In this approach, we produce data-generated prediction models based on random forest classification. These data-generated prediction models can be analyzed as well as used for statistical analysis of what we can expect under different environmental conditions. So even if an environmental condition has not been encountered in the experiment, it is possible to approximate it within reasonable variations; for example, since we have not encountered a storm, it is impossible to know how well the prediction model performs. To avoid invalidating the results, we only consider predictions based on input parameters within the minimum and maximum values encountered in the experiment. However, we do not consider various relations between input parameters to see if and how constraints affects the results.

5.1 Random forest briefer

A random forest model (Breiman 2001) is an ensemble classification model which means that it consists of a set of classifiers. During the training phase, the goal is to achieve a set of classifiers that make as independent errors as possible. By training the classifiers so that the degree of independence is as high as possible, it is possible to achieve good classification results by letting the classifiers vote for the predicted class, that is, the likelihood that the majority is right is high. In the case of random forests, the classifiers consist of decision trees (or regression trees in the case of continuous variables). We can employ the generated model for prediction purposes.

---

5 The reason is that there are different configurations of the optimization algorithm, where each configuration is identified by an integer. In time series 2, 5 and 6 configuration ‘3’ has been used and in time series 7 configuration ‘4’ has been used.

6 Since we need to train the classifiers on the same data, it is impossible to achieve completely independent errors.

One problem is that random forests handle each sample as an individual, which is not really the case since it is a time series. In a time series, there are auto-correlations (that is, there is reduced serial independence). We conjecture from the analysis based on auto-correlation and cross-correlation that the auto-correlative effects can be reduced if we employ a low-pass filter. We can employ a low-pass filter, since we are interested in long-term effects, not transient changes. So we try out different low-pass filter by resampling the data to larger samples (intervals longer than 1 second) prior to application of the generating models based on data.

In Appendix E: Random forest: data-generated prediction models and Monte Carlo simulation for evaluation of prediction models, the code for generating data-generated prediction models, computing the variable influence as well as evaluating the prediction model by using Monte Carlo simulation are found. The caret package has been employed (Kuhn 2008) and we follow the advice by Liaw and Wiener (2002).

5.2 Analysis of random forests

It is possible to determine the most important variable with respect to the predictions of, for example, fuel consumption. A specific technique for variable importance is employed, which in random forests is a measure for the degree of importance of the involved variable with respect to the prediction made by the model. The variable importance is determined by estimating how much the prediction error increases when the value of a specific variable in a set of instance is permuted while the other are not (Liaw & Wiener, 2002). To be more specific, given a set of instances (a set of tuples) to be classified, permute index i in this tuple, at random, and classify each instance by each tree in the ensemble. For each instance you obtain a collection of votes. The average difference between the number of votes for the true class based on the original tuple and the permuted tuple defines the variable importance. For the case of regression trees, the mean squared error is used instead as the error measure.

In Illustration 5, the “Roll.Active2” denotes that optimization is active. As can be seen, there is some influence in this case where we have resampled at 2 seconds, used 3 decision trees in the building of random forests as well as used seed 11 to partition the data set randomly. Due to that there are no distinct variations in the data, the results vary with the randomness of various part of the method (e.g., data partitioning, inherent random steps in random forest technique). Therefore, several prediction models are generated based on different resampling intervals, number of decision trees as well as seeds to partititon to check if they concur. No single configuration has been discovered that is guaranteed to provide results that point in the same direction concerning, for example, variable influence. For example, for time series 5 and 6, the sea current is important in this case, whereas in time series 6 and 7 the

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

speed over ground is important. These differences may be due to that we do not take auto correlative effects into account.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth.Below.Keel</td>
<td>0,39</td>
<td>0,00</td>
<td>36,18</td>
<td>3,15</td>
</tr>
<tr>
<td>Eng.Spd</td>
<td>5,24</td>
<td>13,68</td>
<td>89,34</td>
<td>1,61</td>
</tr>
<tr>
<td>Prop.Ship.Spd</td>
<td>4,04</td>
<td>22,40</td>
<td>81,86</td>
<td>2,73</td>
</tr>
<tr>
<td>Roll.Active2</td>
<td>0,00</td>
<td>29,45</td>
<td>18,02</td>
<td>1,78</td>
</tr>
<tr>
<td>Roll.Amp</td>
<td>8,90</td>
<td>15,96</td>
<td>18,55</td>
<td>4,61</td>
</tr>
<tr>
<td>Roll.Period</td>
<td>7,42</td>
<td>10,56</td>
<td>17,64</td>
<td>6,78</td>
</tr>
<tr>
<td>Rudder.Pos</td>
<td>100,00</td>
<td>100,00</td>
<td>41,83</td>
<td>62,94</td>
</tr>
<tr>
<td>Scav.Air.Pres</td>
<td>1,98</td>
<td>0,70</td>
<td>0,00</td>
<td>2,51</td>
</tr>
<tr>
<td>Sea.Current</td>
<td>18,24</td>
<td>86,17</td>
<td>100,00</td>
<td>5,14</td>
</tr>
<tr>
<td>Speed.Over.Ground</td>
<td>18,76</td>
<td>27,14</td>
<td>96,16</td>
<td>100,00</td>
</tr>
<tr>
<td>Wind.Head</td>
<td>1,58</td>
<td>16,59</td>
<td>22,17</td>
<td>4,68</td>
</tr>
<tr>
<td>Wind.Side</td>
<td>3,14</td>
<td>10,79</td>
<td>73,14</td>
<td>0,00</td>
</tr>
</tbody>
</table>

Illustration 5: Variable importance for resampling period of 2 seconds, 3 decisions tree and seed 11 to partition the data

In cases where we do not receive any influence of the optimization, we analyze the model generation. For example, if the training set contains no variations for optimization, then the prediction model does not react to optimization.

5.3 Analysis of predictions

To analyze the prediction models, we used Monte Carlo simulation to generate attributes and then divided the effect of having optimization turned on with optimization turned off. We generate a probability density function to check if we can find expected effects. For example, Illustration 6 is the probability density function of the predicted fuel savings based on the Monte Carlo simulations of the time series 2 when it is resampled with an interval of 2 seconds, with 3 decision trees and seed 11 to partition the time series. As can be seen, the majority of predictions indicate that we save fuel since the quotient is less than 1.0.

This was done for all possible combinations of resampling intervals, number of decisions trees and seeds, since there is no configuration that is guaranteed to include influence for the optimization. In the

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

compressed archive, files named density-{2,5,6,7}-{Fuel.Cons.l.per.nm,Prop.Slip,Roll.Amp}-all-Window={2,4,8,16}-DecisionsTrees={1,3,5,7,11}-all_seeds.png contains the distribution based on different kinds of moving average windows as well as number of decision trees.

5.3.1 Prediction of fuel consumption optimization

Almost all combinations of resampling period, number of decision trees, and data partitioning seeds generate predictive models which show an influence from the use of optimization algorithm. In all cases where there is an influence from the optimization, the result of the Monte Carlo simulation is similar to the example (in Illustration 6). That is, there are savings with a significant major part of the probability density functions beneath the quotient 1.0 which indicates that the use of the optimization algorithm saves fuel.

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

By using summary statistics and checking 20 vigintiles (1/20th) of the probability density functions for all prediction models we obtained the following results. In 77 out of 80 possible prediction models, the major part\(^7\) of the probability density function is beneath 1.0 and in 11 out of the 80 95%\(^8\) or more of the probability density function is beneath 1.0. In 74 models, less than 25%\(^9\) of the probability density function is above 1.0. In the 3 remaining prediction models, optimization does not have any influence and, thus, no effect on the prediction.

**Prediction sample: 2-Fuel.Cons.l.per.nm-all-2:3seed:11: on/off**

![Probability density function of Monte Carlo simulation of prediction based on prediction model generated from resampling interval 2, 3 decision trees and seed 11 for partitioning for time series 2.](image)

Illustration 6: Probability density function of Monte Carlo simulation of prediction based on prediction model generated from resampling interval 2, 3 decision trees and seed 11 for partitioning for time series 2.

---

7 That is, 11 or more vigintiles of the probability density function is beneath 1.0 indicating that a majority of the predictions indicate fuel saving when optimization is turned on.

8 That is, 19 vigintiles or more.

9 That is, less than 5 vigintiles.

5.3.2 Prediction of roll amplitude

Concerning roll amplitude, no evidence of improvement has been discovered in the prediction models. The result is a parametric distribution of the quotient of the prediction with and without optimization with an average of approximately 1.0 (i.e., no effect). For example, Illustration 7 is an example of this.

Illustration 7: Example of probability density function of Monte Carlo simulation of predictions by models generated from for time series 2, all resampling intervals, 4 decision trees and seed 11 to partition the data set.

5.3.3 Prediction of propeller slip

Concerning propeller slip, the models indicate an improvement. For example, Illustration 8 is an example of the quotient of propeller slip between the optimized case and the unoptimized case for time series 5.

Again, by using summary statistics, we obtained the following results. In 59 out of 80 generated prediction models, the major part of the probability density function is beneath 1.0 and in 43 out of the 80 95% or more of the probability density function is beneath 1.0. In 79 cases, less than 25% of the probability density function is greater than 1.0. In the 21 remaining prediction models, optimization does not have any influence and, thus, no effect on the prediction.

6 Discussion

6.1 Avoidance of multivariate time series analysis
We have successfully demonstrated that the anti-rolling optimization algorithm has a positive effect on fuel consumption as well as propeller slip, when first applying a low-pass filter to the data. In this way,

17 (74)
we ignore fast, spurious changes as well as noise and only emphasize slower processes in terms of fractions or multiples of roll periods of the ship. The auto-correlative effects in the time series are sufficiently weak and smooth to obtain results that agree with model-generated predictions. Note, however, that the variations in the influences in the prediction models may be due to auto correlative effects.

6.2 Related work
The results of this report is in agreement with Q-TAGG report based on simulations and knowledge-based analysis of the data (Q-TAGG R&D, 2014) in general terms. That is, we have not discovered the level of fuel saving with the same accuracy, but there are general tendencies that a majority of predictions based on a Monte Carlo simulation where the data-generated prediction models are employed indicate fuel savings. To be able to compute the level of fuel savings with sufficient accuracy, further analysis is required as well as proving our conjecture that the lack of influence from optimization in some prediction models are due to the random forest technique itself.

7 Conclusions
To summarize, we have found evidence that the anti-roll engine control optimization algorithm saves fuel as well as improves the propeller slip. This agrees with the results from the knowledge-generated models. However, we have found no evidence that the roll amplitude is reduced. This can be caused by the particular configurations of the optimization algorithm that has been used during the experiments.

We conjecture that since the effect of the optimization is comparatively small to other factors (e.g., sea current) and the method we applied contains randomness (e.g., the feature bagging of random forest technique imply that data in the training set can be replaced by random), the effect of optimization is not always visible in the prediction models based on random forests. We therefore used Monte Carlo simulation on the data-generated prediction models and, in a majority of these prediction models, we found that fuel consumption in terms of liter per nautical mile is reduce as well as propeller slip is reduced.

Since random forest treats each sample as an individual, ignoring auto correlative effects, we also did basic auto correlative and cross correlative analysis on the time series of each attribute. The qualitative features were compared and we found that with optimization, the analysis indicates a smoother engine performance when optimization is turned on. Further, the cross correlative effects in time are monotonically decreasing with increased lag in all the time series that we analyzed.
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

7.1 Future work

It seems promising to employ multivariate time series analysis on the data and see if the effect of optimization is visible in these models as well. There are analytical state-space methods are based on Gaussian distribution and linear relationships among variables. There are also simulation-based state-space methods that are not based on Gaussian distributions and linear relationships.

As an alternative to multivariate time series analysis, we could employ particle filtering. For example, Osgood and Liu (2014) demonstrate this for evaluation of spread of diseases, where the prediction can be recalibrated as new data appears. In essence, analysis of the spread of diseases faces similar problems that appears for ships: noisy data, non-Gaussian distributions, and non-linear relationships. For example, if we want continuous improvement of fuel savings data, with an opportunity to further dave increase fuel savings, then we could have a data collection system where ships are scheduled to turn on and off optimization under various conditions. As new data appears, running predictions of fuel consumption monitoring and control can be recalibrated onboard ships in the same way as for the analysis and control of the spread of diseases.

8 REFERENCES


The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study


Wigforss, J 2012, Benchmarks and measures for better fuel efficiency., Chalmers University of Technology, Göteborg, Sweden.

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

APPENDIX A: Signals used for logs

Signals for high-frequency samples:

Raw data signals:

<table>
<thead>
<tr>
<th>Name</th>
<th>R Name</th>
<th>Unit</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eng Spd</td>
<td>Eng.Spdd</td>
<td>RPM</td>
<td>Engine RPM</td>
</tr>
<tr>
<td>Spd Set</td>
<td>Spd.Set</td>
<td>RPM</td>
<td>Engine Speed Set Value for the control loop</td>
</tr>
<tr>
<td>Scav Air Pres</td>
<td>Scav.Air.Pres</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rudd Pos</td>
<td>Rudd.Pos</td>
<td>Degrees</td>
<td>Actual position of rudder</td>
</tr>
<tr>
<td>Roll Pos</td>
<td>Roll.Pos</td>
<td>Degrees</td>
<td>Ship Roll Position. + to Starboard and - to port</td>
</tr>
<tr>
<td>Roll Ctrl</td>
<td>Roll.Ctrl</td>
<td>%</td>
<td>Roll Ctrl Delta Fuel value</td>
</tr>
<tr>
<td>Roll Pitch</td>
<td>Roll.Pitch</td>
<td>degrees</td>
<td>Ship Pitch Position</td>
</tr>
<tr>
<td>DI Activate Roll Ctril</td>
<td>DI.Activate.Roll.Ctrl</td>
<td>Integer</td>
<td>Roll Control Activated from Digital Input</td>
</tr>
</tbody>
</table>

NMEA signals:

<table>
<thead>
<tr>
<th>Name</th>
<th>R Name</th>
<th>Unit</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc Fuel</td>
<td>Acc.Fuel</td>
<td>m³</td>
<td>Accumulated m³ fuel</td>
</tr>
<tr>
<td>Fuel Cons [/h]</td>
<td>Fuel.Cons.l.per.h</td>
<td>l/h</td>
<td>Fuel consumption in liter per hour</td>
</tr>
<tr>
<td>Fuel Cons [/nm]</td>
<td>Fuel.Cons.l.per.nm</td>
<td>l/nm</td>
<td>Fuel consumption in liter per nautic mile</td>
</tr>
<tr>
<td>Speed Through Water</td>
<td>Speed.Through.Water</td>
<td>Knots</td>
<td>Speed through water from the logs</td>
</tr>
<tr>
<td>Speed Over Ground</td>
<td>Speed.Over.Ground</td>
<td>Knots</td>
<td>Speed over ground from GPS</td>
</tr>
<tr>
<td>Log Dist Trip</td>
<td>Log.Dist.Trip</td>
<td>nautic miles</td>
<td>Logged distance of this trip</td>
</tr>
<tr>
<td>Log Dist Tot</td>
<td>Log.Dist.Tot</td>
<td>nautic miles</td>
<td>Logged distance in total</td>
</tr>
<tr>
<td>Heading</td>
<td>Heading</td>
<td>degrees</td>
<td>Compass course</td>
</tr>
<tr>
<td>Wind Rel Spd</td>
<td>Wind.Rel.Spd</td>
<td>m/s</td>
<td>Relative wind speed (to the ship)</td>
</tr>
<tr>
<td>Wind Rel Angle</td>
<td>Wind.Rel.Angle</td>
<td>Degrees</td>
<td>Relative wind direction (to the ship)</td>
</tr>
</tbody>
</table>

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

<table>
<thead>
<tr>
<th>Name</th>
<th>R Name</th>
<th>Unit</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth Below Keel</td>
<td>Depth.Below.Keel</td>
<td>m</td>
<td>Depth below keel in meters</td>
</tr>
<tr>
<td>GPS Long</td>
<td>GPS.Long</td>
<td></td>
<td>GPS Longitude Position (decimal format)</td>
</tr>
<tr>
<td>GPS Lat</td>
<td>GPS.Lat</td>
<td></td>
<td>GPS Latitude Position (decimal format)</td>
</tr>
<tr>
<td>ME Power</td>
<td>ME.Power</td>
<td>MW</td>
<td>Total Power Output from Engine</td>
</tr>
<tr>
<td>Fuel Oil Flow</td>
<td>Fuel.Oil.Flow</td>
<td>l/h</td>
<td>Fuel consumption in liter per hour from meter</td>
</tr>
</tbody>
</table>

Indirect measurements:

<table>
<thead>
<tr>
<th>Name</th>
<th>R Name</th>
<th>Unit</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel Cons per s</td>
<td>Fuel.Cons.per.s</td>
<td>ml/s</td>
<td>Average fuel rack position last second</td>
</tr>
<tr>
<td>Ave Fuel Last Turn</td>
<td>Ave.Fuel.Last.Turn</td>
<td>ml</td>
<td>Milliliter fuel injected last engine turn</td>
</tr>
<tr>
<td>ml fuel</td>
<td>ml.fuel</td>
<td>ml</td>
<td>Actual ml injected fuel injected per revolution</td>
</tr>
<tr>
<td>Eng Rev</td>
<td>Eng.Rev</td>
<td>integer</td>
<td>number of revolutions of the engine [0 to 65535]</td>
</tr>
<tr>
<td>Prop Dist</td>
<td>Prop.Dist</td>
<td>nautic miles</td>
<td>Total Propeller Distance</td>
</tr>
<tr>
<td>Prop Ship Spd</td>
<td>Prop.Ship.Spd</td>
<td>knots</td>
<td>Water Speed through propeller in knots</td>
</tr>
<tr>
<td>Roll Amp</td>
<td>Roll.Amp</td>
<td>degrees</td>
<td>Amplitude for ship roll</td>
</tr>
<tr>
<td>Roll Period</td>
<td>Roll.Period</td>
<td>seconds</td>
<td>Roll ship period</td>
</tr>
<tr>
<td>Roll Rudd+Phase</td>
<td>Roll.Rudd.Phase</td>
<td>integer</td>
<td>Combined state for Roll and Rudder</td>
</tr>
<tr>
<td>Roll Fuel</td>
<td>Roll.Fuel</td>
<td>%</td>
<td>Roll Ctrl Fuel value to actuator</td>
</tr>
<tr>
<td>Roll Ave SpdErr</td>
<td>Roll.Ave.SpErr</td>
<td>%</td>
<td>Average Speed Error when in roll control mode</td>
</tr>
<tr>
<td>Roll State</td>
<td>Roll.State</td>
<td>integer</td>
<td>Roll Control State (^{10})</td>
</tr>
</tbody>
</table>

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

<table>
<thead>
<tr>
<th>Name</th>
<th>R Name</th>
<th>Unit</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roll Active</td>
<td>Roll.Active</td>
<td>integer</td>
<td>Roll Control Active</td>
</tr>
<tr>
<td>Roll Active2</td>
<td>Roll.Active2^{11}</td>
<td>integer</td>
<td>Optimization turned on (If Roll.State&gt;0 &amp;&amp; Roll.Active&gt;0)</td>
</tr>
</tbody>
</table>

10 In the experiments, 3 and 4 were used.

11 Not part of the log file, introduced in the analysis.

Appendix B: Resampling code

The resampling (Table 5) employ the resampling functions (Table 6) according to a configuration file (a sample is available in Table 8). The other functions are there to support reading raw log data or zoo\textsuperscript{12} files and writing to zoo files.

\textsuperscript{12} Zoo is an R format for equidistant time series that support various forms of aggregation, resampling etc.

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels:
An experimental study


25 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

# close file
close(.file)

# get rid of additional empty columns
# get rid of the configuration row (1) marked with a "T"

.tmpTableOnlyDataRows <- .tmpTable[2:length(.tmpTable$Time),1:numberOfSrcColumnNames]

# correct the time string by adding prefix '0's to the
# fraction of the second for entries where the number of
# digits in the fraction is less than three. For example
# "2013-01-01 20:57:23.23" is converted to "2013-01-01 20:57:23.023"
# If the number of digits in the fraction is 3 or more, then
# no conversion is made. So, it will work for erroneous log files
# as well as correct log files.
.tmpTableOnlyDataRows$CorrectedTimeString <- sapply(.tmpTableOnlyDataRows$Time,FUN=.correctTimeString)

# compute index based on time string
.timeIndex <- as.POSIXct(.tmpTableOnlyDataRows$CorrectedTimeString,format=.timeFormat)

# convert into a zoo object, where the index is the time stamp
# and thus we do not need to have a column with the time stamp
# In fact, zoo cannot handle columns with different types well
# so removing the "Time" attribute, this problem is removed.
.zooTable <- zoo(.tmpTableOnlyDataRows[,2:numberOfSrcColumnNames],order.by=timeIndex)
if (any(sort(names(.zooTable))!=sort(.zooColumnNames))) {
    warning("Incorrect number of columns in the result")
}
return(.zooTable) # raw, irregular time series

Table 3: QTAGGReadLogIntoZoo: Reads log data into zoo format

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```
.getTimeString <- function(s) {
    part <- unlist(strsplit(s,"\."))
    return(paste(part[1],sprintf("%03d",as.numeric(part[2])),sep="."))
}
QTAGGReadZoo <- function(.filePath = stop("A file path must be provided"),.orgColumnNames = .QTAGGOrgColumns,.zooColumnNames=.QTAGGZooColumns,.header=TRUE,.skip=0,.sep="\t",.samplingInterval=1.0,.checkColumnNames=TRUE) {
    require(zoo) # specify that this function required 'zoo' to work
    .orgLength <- length(.orgColumnNames)
    # set up coercion rules via .cls
    .cls <- rep(NA,.orgLength)
    .cls[1]="character" # do not perform automatic coercion the first column (Time attribute)
    # read the raw data
    .tmpTable <- read.table(.filePath,header=.header,skip=.skip,sep=.sep,colClasses=.cls)
    # get rid of additional empty columns
    # get rid of the configuration row (1) marked with a "T"
    .tmpTableOnlyDataRows <- .tmpTable[1:length(.tmpTable$Index),1:.orgLength]
    # compute index based on time string
    .timeIndex <- as.POSIXct(.tmpTableOnlyDataRows$Index,format=.QTAGGTimeFormat)
    # convert into a zoo object, where the index is the time stamp
    # and thus we do not need to have a column with the time stamp
    # In fact, zoo cannot handle columns with different types well
    # so removing the "Time" attribute, this problem is removed.
    .zooTable <- zoo(.tmpTableOnlyDataRows[,2:.orgLength],order.by=.timeIndex,frequency=.samplingInterval)
    if (.checkColumnNames) {
        if (.checkColumnNames & any(names(.zooTable)!=.zooColumnNames)) {
            stop("Incorrect number of columns in the result")
        }
    }
    return (.zooTable) # raw, irregular time series
}
```

Table 4: QTAGGReadZoo: Reads a csv file as if it is in zoo format.

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
QTAGGResample <- function(.QTAGGZooTable, .samplingInterval=1, .aggregationCfgTable=QTAGGAttrAggregationCfg) {
  require(plyr, quietly=TRUE)
  if (!is.zoo(.QTAGGZooTable)) {
    stop("First argument should be 'zoo' object")
  }
  .names <- names(.QTAGGZooTable)
  .nAttributes <- length(.names)
  .resampledTimeIndex <- as.POSIXct(ceiling(as.double(time(.QTAGGZooTable))/.samplingInterval)*.samplingInterval, origin=.QTAGGOrigin)

  # iterate over the attributes, transform them one by one
  # according to the aggregation configuration table
  # unfortunately, the aggregate function can only

  for (.index in 1:.nAttributes) {
    .name <- .names[.index]
    .tmpAttr <- .QTAGGZooTable[, .index]

    .aggrCmd <- paste0(.name, " <- aggregate(.tmpAttr, .resampledTimeIndex, function(x) ");

    # build the aggregation command
    .aggrCmd <- paste0(.aggrCmd, ", return(" .aggregationCfgTable$Resampling[.index], "(x)"
    if (!is.na(.aggregationCfgTable$AdditionalArgument[.index])) {
      .aggrCmd <- paste0(.aggrCmd, ", " .aggregationCfgTable$AdditionalArgument[.index])
    }
    .aggrCmd <- paste0(.aggrCmd, "), regular=TRUE")

    # parse and evaluate the string
    eval(parse(file="", text=.aggrCmd)[[1]])
  }

  # iterate over the resampled variables and bind them together
  # into a result and return it.
  .cbindCmd <- "cbind(
  for (.index in 1:.nAttributes) {
    
  .cbindCmd <- "cbind("
  for (.index in 1:.nAttributes) {
```

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Table 5: QTAGGResample: Resampling of logged data

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

# computes the mean of Depth below the keel attribute.
# it handles the >300m represented as zeros

QTAGGdbkMean <- function(x) {
    .tmp <- x[x!=0] # get rid of 0's, since they imply >300 depth
    if (is.null(.tmp) || length(.tmp)==0) {
        return(NA)
    } else {
        return(mean(.tmp))
    }
}

# computes the frequency of values within an interval
# the majority is chosen. Is used for, for example, Roll.Active.
# if there is a tie, the lowest value is chosen.

QTAGGmajority <- function(x) {
    .tabular <- table(x)
    .maximumCountVector <- .tabular[.tabular==max(.tabular)]
    return(min(index(.maximumCountVector))) # use the lowest value, if there is a tie
}

# Used to select the last value in the interval to represent the interval.
# Used for, for example, Roll.State

QTAGGlast <- function(x) {
    tail(x,1)
}

QTAGGsum <- function(x,.noOfValPerInterval=4) {
    .tmp <- x[!is.na(x)]
    if (length(.tmp)<1) {
        return(0);
    } else {
        return(sum(.tmp)/length(.tmp)*.noOfValPerInterval)
    }
}

QTAGGmean <- function(x) {
    return(mean(x[is.na(x)]))
}

QTAGGmax <- function(x) {
    return(max(x[is.na(x)]))
}

QTAGGmin <- function(x) {
    return(min(x[is.na(x)]))
}


30 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Table 6: Resampling functions, x is a set of measurements in the resampling interval.

```
.qtaggorgcolumns<-(c(
  "Time",
  "Eng.Spd",
  "Spd.Set",
  "Fuel.Set",
  "Fuel.Cons.per.s",
  "Ave.Fuel.Last.Turn",
  "ml.fuel",
  "Acc.Fuel",
  "Eng.Rev",
  "Prop.Dist",
  "Prop.Ship.Spd",
  "Rudder.Pos",
  "Roll.Pos",
  "Roll.Ctrl",
  "Roll.Amp",
  "Roll.Period",
  "Roll.Rudd.Phase",
  "Roll.Fuel",
  "Roll.Ave.SpdErr",
  "Roll.Pitch",
  "DI.Activate.Roll.Ctrl",
  "Roll.State",
  "Roll.Active",
  "Speed.Through.Water",
  "Speed.Over.Ground",
  "Log.Dist.Trip",
  "Log.Dist.Tot",
  "Heading",
  "Wind.Rel.Speed",
  "Wind.Rel.Angle",
  "Depth.Below.Keel",
  "GPS.Long",
  "GPS.Lat",
  "ME.Power",
  "Shaft.Power",
  "Fuel.Oil.Flow"
));

.qtaggzoocolumns <- .qtaggorgcolumns[2:length(.qtaggorgcolumns)]
```

Table 7: QTAGG Default columns

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Index | Attribute | Resampling | AdditionalArgument
--- | --- | --- | ---
1 | Eng.Sp | QTAGGmean | NA
2 | Spd.Set | QTAGGmean | NA
3 | Fuel.Set | QTAGGmean | NA
4 | Fuel.Cons.per.s | QTAGGmean | NA
5 | Ave.Fuel.Last.Turn | QTAGGsum | NA
6 | ml.Fuel | QTAGGmean | NA
7 | Acc.Fuel | QTAGGmean | NA
8 | Eng.Rev | QTAGGmean | NA
9 | Prop.Dist | QTAGGmean | NA
10 | Prop.Ship.Sp | QTAGGmean | NA
11 | Rudder.Pos | QTAGGmean | NA
12 | Roll.Pos | QTAGGmean | NA
13 | Roll.Ctrl | QTAGGmean | NA
14 | Roll.Amp | QTAGGmax | NA
15 | Roll.Period.QTAGGmax | NA
16 | Roll.Rudd.Phase | QTAGGmean | NA
17 | Roll.Fuel | QTAGGmean | NA
18 | Roll.Ave.Sp | QTAGGmean | NA
19 | DI.Activate.Roll.Ctrl | QTAGGmajority | NA
20 | Roll.State | QTAGGlastNA | NA
21 | Roll.Active | QTAGGmajority | NA
22 | Roll.Pitch | QTAGGmean | NA
23 | Speed.Through.Water | QTAGGmean | NA
24 | Speed.Over.Ground | QTAGGmax | NA
25 | Log.Dist.Trip | QTAGGmax | NA
26 | Log.Dist.Tot | QTAGGmean | NA
27 | Heading | QTAGGmean | NA
28 | Wind.Rel.Speed | QTAGGmean | NA
29 | Wind.Rel.Angle | QTAGGdbklMean | NA
30 | Depth.Below.Keel | QTAGGmean | NA
31 | GPS.Long | QTAGGmean | NA
32 | GPS.Lat | QTAGGmean | NA
33 | ME.Power | QTAGGmean | NA
34 | Shaft.Power | QTAGGmean | NA
35 | Fuel.Oil.Flow | QTAGGmean | NA

Table 8: Sample configuration file for resampling

Appendix C: Auto and cross correlation code

```r
library(QTAGGData)

library(forecast)

if (.Platform$OS.type=="windows") {
  logDataDirPath <- "C:/Dropbox/AREA/HiS - Skövde/Jonas/Loggs 2013-11-07";
} else {
  logDataDirPath <- "~/Dropbox/AREA/HiS - Skövde/Jonas/Loggs 2013-11-07";
}

tests <- c("2","5","6","7")
tparams <- c("Fuel.Cons.l.per.nm","Roll.Amp")
params <-


32 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels:
An experimental study

der.Pos","Wind.Head","Wind.Side"

arima.list <- list()

evaluate <- function(string) {
    return(eval(parse(file="",text=string)))
}

getZoo <- function(zooData,index) {
    return(evaluate(paste0("zooData","$",index)))
}

for (tst in tests) {
    t2<-list()
    fileName<-paste0("Kithira 20131106 Biscaya ",tst,".zoo")
    x <- QTAGGReadZoo(fileName,.checkColumnNames=FALSE)
    x <- zooReg(coredata(x),frequency=15)
    Wind.Head <- x[, "Wind.Rel.Speed"] * cos(x[, "Wind.Rel.Angle"] * pi)/180
    Wind.Side <- abs(x[, "Wind.Rel.Speed"] * sin(x[, "Wind.Rel.Angle"] * pi)/180)
    x<-merge(x,Wind.Head)
    x<-merge(x,Wind.Side)
    if (tst!="7") {
        index=c("all",0,3)
    } else {
        index=c("all",0,4)
    }
    # produce 3 different views, one for without control, one with control, and one with all
    t2 <- list()
    for (i in index) {
        if (i!="all") {
            df <- list(name=i,value=x|xRoll.State==i)
        } else {
            df <- list(name=i,value=x)
        }
        t2 <- append(t2,list(df))
    }
    for (nm in t2) {
        m<-nm$value
        print("ACF1")
        for (tp in tparams) {
            fname <- paste0(logDataDirPath,"/ACF- ",tst,"-",nm$name,"-" TP,".png")
            png(filename = fname)
            a1<-acf(eval(parse(file="",text=paste0("m","$",tp))),lag.max=60,type="correlation",na.action=na.approx,plot=TRUE,main=paste0("ACF: ",tst,"-",nm$name,"-" TP),ylm=(c(-1,1))
            dev.off()
            idx <- paste0(tst,"-",nm$name,"-" TP)
            print(idx)
            arima.list[[idx]]<-auto.arima(eval(parse(file="",text=paste0("m","$",tp))),stepwise=FALSE,parallel=TRUE,num.cores=12)
            print("Hello")
        }
    }
}

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels:
An experimental study

for (p in params) {
    if (tp==p) {
        next;
    }
    print(p)
    fname <- paste0(logDataDirPath,"/ACF- ",tst,"-",nm$name,"-",p,".png")
    png(filename = fname,height=10,width=10,units="cm",res=72)
    dev.off()
    a2<-acf(eval(parse(file="",text=paste0("m","$,",p))),lag.max=60,,type="correlation",na.action=na.approx,plot=TRUE,main=paste0("ACF: ",tst,"-",nm$name,"-",p),ylim=c(-1.0,1.0))

    ccf<-
ccf(eval(parse(file="",text=paste0("m","$,",p))),eval(parse(file="",text=paste0("m","$,tp"))),lag.max=60,type="correlation",na.action=na.approx,plot=TRUE,main=paste0("CCF; ",tst,"-",nm$name,"-",p," -> ",tp),ylim=c(-1.0,1.0))
    dev.off()
}
fname <- paste0(logDataDirPath,"/arimaModels.txt")
sink(file=fname)
for (n in names(arima.list)) {
    cat("Series: ",n)
    ar <- arima.list[[n]]
    print(ar)
    print("")
}
sink()
fname <- paste0(logDataDirPath,"/arimaModels-R-format.txt")
save(arima.list,file=fname,ascii=TRUE)

Appendix D: Auto and cross correlation results
In Table 9 and Table 10, we can see that with optimization, the time series is smoother than without optimization. Further, there is a stronger auto-correlative effect with optimization than without, which

34 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

can be interpreted as that there is a state concerning the environment conditions that is handled and this corresponds to the optimization algorithm. As can be seen in Table 11, the speed is rather stable with small variations for both with and without optimization, an interesting fact when considering the cross-correlation.

Further, concerning cross-correlation in Table 12 and Table 13, we can see that with optimization, there is less correlation between the speed over ground and the fuel consumption (l/nm) as well as the roll amplitude. Further, again, with optimization, the curves indicate a more smooth operation. Finally, there is a shift in lag, with optimization, the effect seems to be delayed compared to without optimization. This latter corresponds to that the optimization keeps a state of the current environmental conditions and make decisions about that.

The rest of the diagrams are found in compressed archives for your convenience. The naming of the models are as follows: ACF-{2,5,6,7}-{0,3,4}-Variables={Depth.Below.Keel,Eng.Spd,Fuel.Cons.l.per.nm,Roll.amp,Rudder.Pos,Scav.Air.Pres,Speed.Over.Ground,Wind.Head,Wind.Side} is the auto-correlative analysis where {2,5,6,7} refers to the log, {0,3,4} refers to no optimization (0) or optimization (either 3 or 4), {Depth....} are the variables in question; CCF-{2,5,6,7}-{0,3,4,All}-Variables-{Roll.Amp,Prop.Slip,Fuel.Cons.l.per.nm} which are the cross-correlative analysis.

13 Not only the set speed and the current speed, but also other factors such as roll amplitude etc.

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

<table>
<thead>
<tr>
<th>With optimization</th>
<th>Without optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Graph 1" /></td>
<td><img src="image2" alt="Graph 2" /></td>
</tr>
<tr>
<td><img src="image3" alt="Graph 3" /></td>
<td><img src="image4" alt="Graph 4" /></td>
</tr>
</tbody>
</table>


36 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

With optimization

Without optimization

Table 9: Auto-correlation of fuel consumption l/nm

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

<table>
<thead>
<tr>
<th>With optimization</th>
<th>Without optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="ACF: 2-3-Roll.Amp" /></td>
<td><img src="image2" alt="ACF: 2-4-Roll.Amp" /></td>
</tr>
<tr>
<td><img src="image3" alt="ACF: 5-3-Roll.Amp" /></td>
<td><img src="image4" alt="ACF: 5-4-Roll.Amp" /></td>
</tr>
</tbody>
</table>

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

<table>
<thead>
<tr>
<th>With optimization</th>
<th>Without optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="ACF: 6-3-Roll.Amp" /></td>
<td><img src="image" alt="ACF: 6-0-Roll.Amp" /></td>
</tr>
<tr>
<td><img src="image" alt="ACF: 7-4-Roll.Amp" /></td>
<td><img src="image" alt="ACF: 7-0-Roll.Amp" /></td>
</tr>
</tbody>
</table>

Table 10: Auto-correlation of roll amplitude

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

<table>
<thead>
<tr>
<th>Without optimization</th>
<th>With optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph 1" /></td>
<td><img src="image2.png" alt="Graph 2" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Graph 3" /></td>
<td><img src="image4.png" alt="Graph 4" /></td>
</tr>
</tbody>
</table>

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Table 11: Auto-correlation of speed over ground

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

<table>
<thead>
<tr>
<th>With optimization</th>
<th>Without optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

**The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study**

<table>
<thead>
<tr>
<th>With optimization</th>
<th>Without optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td><img src="image3" alt="Graph" /></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
</tbody>
</table>

**Table 12: Cross-correlation between speed over ground and fuel consumption (l/nm)**

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Without optimization

With optimization


44 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Without optimization | With optimization
---|---

Table 13: Cross-correlation between speed over ground and roll amplitude


45 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Appendix E: Random forest: data-generated prediction models and Monte Carlo simulation for evaluation of prediction models

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Library("QTAGGData")
library("parallel")

## check platform
if (.Platform$OS.type=="windows") {
  logDataDirPath <- "C:/Dropbox/AREA/HiS - Skövde/Jonas/Loggs 2013-11-07";
} else {
  logDataDirPath <- "~/QTAGGFiles/AREA/HiS - Skövde/Jonas/Loggs 2013-11-07";
}
cluster<-makePSOCKcluster(12,outfile=paste0(logDataDirPath,"/banzai.txt")) # 12 cores

isIn <- function(value,agg) {
  return(agg[agg==value]!=0);
}

analyzeTestCase <- function(runParameter=stop("A sun parameter must be provided")) {
  if (!require("caret")) {
    stop("Library caret is unavailable")
  }
  if (!require("QTAGGData")) {
    stop("Library QTAGGData is unavailable")
  }
  file <- runParameter["filePath"]
  targetVariableName <- runParameter["target"]
  state<-runParameter["state"]
  rollMeanPreprocessing <- runParameter["rollMeanPreprocessing"]

47 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels:
An experimental study

rollMeanWindow <- runParameter[["rollMeanWindow"]]
randomForestPartition <- runParameter[["randomForestPartition"]]
randomForestSeed <- runParameter[["randomForestSeed"]]

print(file)
print(targetVariableName)
print(state)
print(rollMeanPreprocessing)

## read log data into a data frame
tab <- QTAGGReadZoo(file, .checkColumnNames = FALSE)

## only consider rows with the appropriate state
if (state != "all") {
  tab <- tab[tab$Roll.State == state,]
} else {
  tab$Roll.Active2 <- tab$Roll.State != 0
}

## attributes to remove
  "Roll.Ctrl", "Fuel.Cons.l.per.h")

## potentially influencing variables per target
potentialTargetInfluencers <- list()
potentialTargetInfluencers[["Fuel.Cons.l.per.nm"]]<-
  c(
    "Depth.Below.Keel",
    "Eng.Spd",
    "Prop.Ship.Spd",
    "Roll.Active2",
  )

48 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

| "Roll.Amp", |
| "Roll.Period", |
| "Rudder.Pos", |
| "Scav.Air.Pres", |
| "Sea.Current", |
| "Speed.Over.Ground", |
| "Wind.Head", |
| "Wind.Side" |

potentialTargetInfluencers["Roll.Amp"] <-
c("Depth.Below.Keel",
"Eng.Spd",
"Prop.Ship.Spd",
"Prop.Slip",
"Roll.Active2",
"Roll.Period",
"Rudder.Pos",
"Scav.Air.Pres",
"Sea.Current",
"Speed.Over.Ground",
"Wind.Head",
"Wind.Side"
)

potentialTargetInfluencers["Prop.Slip"] <-
c("Depth.Below.Keel",
"Fuel.Cons.l.per.nm",
"Prop.Ship.Spd",
"Roll.Active2",
"Roll.Amp",
"Roll.Period",
"Rudder.Pos",
"Sea.Current",
"Speed.Over.Ground",

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
"Wind.Head",
"Wind.Side"

)

# remove

tab <- tab[, !colnames(tab) %in% c("Speed.Through.Water")]

# calculate ml/sec
tab[, "ml.fuel"] <- (tab[, "ml.fuel"] * tab[, "Eng.Spd"]) / 60

# rename
names(tab)[which(names(tab) == "ml.fuel")] <- "ml.fuel.sec"

# divide Wind.Rel.Angle into components
Wind.Head <- tab[, "Wind.Rel.Speed"] * cos((tab[, "Wind.Rel.Angle"] * pi)/180)
Wind.Side <- abs(tab[, "Wind.Rel.Speed"] * sin((tab[, "Wind.Rel.Angle"] * pi)/180))

# rename
names(tab)[which(names(tab) == "Wind.Rel.Speed")] <- "Wind.Head"
names(tab)[which(names(tab) == "Wind.Rel.Angle")] <- "Wind.Side"

# assign new values

tab[, "Wind.Head"] <- Wind.Head

# add inverse + inverse square + square

print("Banzai")
```

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
for (i in 0:1) {
    # skip i == 1 or 0
    if (i %in% c(0,1)) {
        next
    }
        if (i<0) {
            vName <- paste0(col,".m",abs(i))
        } else {
            vName <- paste0(col,".",i)
        }
        dynCmd <- paste0(vName,"<-tab[,"col","^",i]
        print(dynCmd)
        eval(parse(file="",text=dynCmd))
        dynCmd2<-paste0("merge(tab,"vName,")")
        print(dynCmd2)
        tab<-eval(parse(file="",text=dynCmd2))
    }
}
print("Yeehaw")

if (rollMeanPreprocessing==TRUE) {
    #tab <- rollmean(x=tab,k=rollMeanWindow)
    ## apply roll mean
    tab <- rollmean(x=tab,k=rollMeanWindow)
    ## only select values
    tab <- tab[unlist(lapply(1:length(tab),function(i) i%%rollMeanWindow==0))]}
    # remove NA rows, corrected: melj
    tab <- tab[which(!is.na(tab), arr.ind=TRUE)][,1]

    # divide into training and testing data
    set.seed(randomForestSeed)
```

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
# indices for training data
print(names(tab))

# targetVariableName <- "Fuel.Cons.l.per.nm"
inTrain <- createDataPartition(tab[, targetVariableName], p=1/randomForestPartition, list=FALSE)

# target variable
trainVar <- tab[inTrain, targetVariableName]
testVar <- tab[-inTrain, targetVariableName]

# remove target variable from data frame
tab <- tab[-which(names(tab) == targetVariableName)]

# training data
trainDescr <- tab[inTrain, ]
# testing data
testDescr <- tab[-inTrain, ]

# run a random forest
model <- train(trainDescr, trainVar, method = "rf", ntree=randomForestPartition, importance=TRUE)

return(list(name=paste0(runParameter$testCase,"-",targetVariableName,"-",state,"-",rollMeanWindow,"-",randomForestPartition,"-",randomForestSeed),testCase=runParameter$testCase,target=targetVariableName,state=state,model=model,rollMeanWindow=rollMeanWindow,randomForestPartition=randomForestPartition,randomForestSeed=randomForestSeed))
}

parAnalyzeTestCases <- function(tc=stop("A list of test cases must be provided"),targets=stop("A list of target must be provided")) {
  # check that it is a list
  if (!is.list(tc)) {
    stop(paste0("Content in ",tc," is not a list of test cases"))
  }

52 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

# check that it is a list of test cases
lapply(tc,function(t) if (!(any("id"==names(t)) & & any("optimizationIndex"==names(t)))))
stop(paste0("\"","tc","\" is not a list of test cases")))

# translate test cases to run parameters for "run"
runParameters <- list()
for (t in tc) {
  for (target in targets) {
    for (state in c(0,t["optimizationIndex"],"all")) {
      for (rmw in t$rollMeanWindows) {
        if (rmw<=1) {
          rollMeanPreprocessing=FALSE
        } else {
          rollMeanPreprocessing=TRUE
        }
        for (rfpar in t$randomForestPartitioningFactors) {
          for (rfseed in t$randomForestSeeds) {
            rp <- list(testCase=t$id,target=target,filePath=paste0(logDataDirPath,"/","Kithira 20131106 Biscaya ",t["id"],".zoo"),state=state,rollMeanPreprocessing=rollMeanPreprocessing,rollMeanWindow=rmw,randomForestPartition=rfpar,randomForestSeed=rfseed)
            runParameters <- append(runParameters,list(rp))
          }
        }
      }
    }
  }
}
result <- parLapply(cluster,runParameters,analyzeTestCase)
#result <- lapply(runParameters,analyzeTestCase)
return(result)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

arrOfMatches <- function(lst,attr,valSet) unlist(lapply(lst,function(z) z[[attr]] %in% valSet))

summarizeResults <- function(models) {
  if (!require("caret")) {
    stop("Library caret is unavailable")
  }

  relImpOfModelsSortedOnRows<-list()
  relImpOfModelsNoSort<-list()
  nameOfData <- names(models) #c("2-0","2-3","2-all","5-0","5-3","5-all","6-0","6-3","6-all","7-0","7-4","7-all")
  print(nameOfData)
  indexOfSeparated <- which(arrOfMatches(models,"state",c(0,3,4))) #c(1,2,4,5,7,8,10,11)
  indexOfAggregated <- which(arrOfMatches(models,"state",c("all")))
  print(nameOfData)
  # obtain the variable influence
  for (m in models) {
    # get the model
    vm <- varImp(m$model)
    # figure out the influence
    ri <- vm$importance
    # sort it on the names of the rows
    sortedRi <- ri$Overall[order(row.names(ri))]
    #ensure that dimensions are correct, lost in sorting
    dim(sortedRi)<-c(length(sortedRi),1)
    # set the row names on the sorted vector
    row.names(sortedRi) <- row.names(ri)[order(row.names(ri))]
    #append it to the list of sorted element
    relImpOfModelsSortedOnRows <- append(relImpOfModelsSortedOnRows,list(sortedRi))
    # append unsorted to the list of unsorted
    relImpOfModelsNoSort <- append(relImpOfModelsNoSort,list(ri))
  }
}

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

# combine into tables (vectors with row & column names)

if (length(relImpOfModelsSortedOnRows)>0) {

  # get the first element
  resSriSeparated <- relImpOfModelsSortedOnRows[indexOfSeparated[1]][[1]]
  # iterate over the following, possibly empty sequence of elements
  dim(resSriSeparated)<-c(length(resSriSeparated),1)
  i<-2
  sri<-c()
  while (i<=length(indexOfSeparated)) {
    sriIndex <- indexOfSeparated[i]
    # get the actual object
    sri <- relImpOfModelsSortedOnRows[[sriIndex]]
    # bind it columnwise
    resSriSeparated <- cbind(resSriSeparated,sri)
    i<-i+1
  }
  # set the row names on the vector
  rownames(resSriSeparated)<-rownames(sri)
  # set the column names on the vector
  print(indexOfSeparated)
  print(length(colnames(resSriSeparated)))
  cat("N=",Length(nameOfData),"\n")
  print(dim(colnames(resSriSeparated)))
  print(dim(nameOfData))
  print(dim(resSriSeparated))
  print(colnames(resSriSeparated))
  print(resSriSeparated)
  print(nameOfData)
  colnames(resSriSeparated)<-nameOfData[indexOfSeparated]
    # set the names
    #names(resSriSeparated)<-nameOfData[indexOfSeparated]
  } else {
    resSriSeparated <- c()
  }

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

if (length(relImpOfModelsSortedOnRows)>0) {
    # get the first element
    resSriAggregated <- relImpOfModelsSortedOnRows[indexOfAggregated[1]][[1]]
    if (length(resSriAggregated) >0) {
        print("A")
        dim(resSriAggregated) <- c(length(resSriAggregated),1)
        print("B")
        ## iterate over the following, possibly empty sequence of elements

        i <- 2
        while ( i<=length(indexOfAggregated)) {
            sriIndex <- indexOfAggregated[i]
            ## get the actual object
            sri <- relImpOfModelsSortedOnRows[sriIndex][[1]]
            ## bind it columnwise
            resSriAggregated <- cbind(resSriAggregated,sri)
            i<-i+1
        }

        # set the row names on the vector
        rownames(resSriAggregated)<-rownames(sri)
        # set the column names on the vector

        cat("Index of aggregated=",indexOfAggregated,"\n")
        cat("Names",nameOfData[indexOfAggregated],"\n")
        cat("Dim(resSriAggregated)=",dim(resSriAggregated),"\n")
        colnames(resSriAggregated)<-nameOfData[indexOfAggregated]
        # set the names

        #names(resSriAggregated)<-nameOfData[indexOfAggregated]
    } else {
        resSriAggregated <- c()
    }
} else {
    resSriAggregated<-c()


56 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

rollMeanWindows=c(2,4,8,16,32)
randomForestPartitioningFactors <- c(3,5,7,11)
randomForestSeeds <- c(1,3,5,7,11)

# test cases for Kithira November: 3, 5, 6 & 7
# 2, 5 & 6 alternates between no optimization & optimization level 3
# 7 alternates between no optimization & optimization level 4

rollMeanWindows=c(1,3,5,7,11)
randomForestPartitioningFactors <- c(1,3,5,7)
randomForestSeeds <- c(1,3,5)

# testCases<- list(list(id=2,optimizationIndex=3,rollMeanWindows=rollMeanWindows,randomForestPartitioningFactors=randomForestPartitioningFactors,randomForestSeeds=randomForestSeeds),
# list(id=5,optimizationIndex=3,rollMeanWindows=rollMeanWindows,randomForestPartitioningFactors=randomForestPartitioningFactors,randomForestSeeds=randomForestSeeds),
# list(id=6,optimizationIndex=3,rollMeanWindows=rollMeanWindows,randomForestPartitioningFactors=randomForestPartitioningFactors,randomForestSeeds=randomForestSeeds))

# testCases<- list(list(id=2,optimizationIndex=3,rollMeanWindows=rollMeanWindows,randomForestPartitioningFactors=randomForestPartitioningFactors,randomForestSeeds=randomForestSeeds))

Funded by: Vinnova Forska & Väx 2013, “System för bränslebesparing på stora fartyg”, 2013-00301  57 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
dim(modelTargets)<-c(length(modelTargets),1)
dim(modelWindows)<-c(length(modelWindows),1)
modelProperties <- cbind(modelTargets,modelWindows,modelRandomForestPartitions,modelRandomForestSeeds)
```

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

colnames(modelProperties) <- c("target","window","randomForestPartition","randomForestSeed")

## aggregate the variable influence into sorted tables
## based on partitions

modelView <- list()
varImpSummary <- list()
for (target in targets) {
  modelView[[target]] <- list()
  varImpSummary[[target]] <- list()
  for (rollMeanWindow in rollMeanWindows) {
    modelView[[target]][[rollMeanWindow]] <- models[which(modelTargets==target && modelWindows==rollMeanWindow)]
    varImpSummary[[target]][[rollMeanWindow]] <- list()
    for (rfpar in randomForestPartitioningFactors) {
      modelView[[target]][[rollMeanWindow]][[rfpar]] <- list()
      varImpSummary[[target]][[rollMeanWindow]][[rfpar]] <- list()
      for (rfseed in randomForestSeeds) {
        idx <- unlist(lapply(1:dim(modelProperties)[1],function (i) modelProperties[i,"target"]==target && modelProperties[i,"window"]==rollMeanWindow && modelProperties[i,"randomForestPartition"]==rfpar && modelProperties[i,"randomForestSeed"]==rfseed))
        modelView[[target]][[rollMeanWindow]][[rfpar]][[rfseed]] <- models[idx]
        varImpSummary[[target]][[rollMeanWindow]][[rfpar]][[rfseed]] <- summarizeResults(modelView[[target]][[rollMeanWindow]][[rfpar]][[rfseed]])
      }
    }
  }
}

print(varImpSummary)
for (target in targets) {


The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```
##print("Target=",target)
for (rollMeanWindow in rollMeanWindows) {
  ##print("Roll mean window =", rollMeanWindow)
  for (partition in partitions) {
    ##print("Partition=",partition)
    for (randomForestPartition in randomForestPartitioningFactors) {
      for (randomForestSeed in randomForestSeeds) {
        write.csv(varImpSummary[[target]][[rollMeanWindow]][[randomForestPartition]][[randomForestSeed]][[partition]],paste0(logDataDirPath,"/","VarImp-",target,"-",rollMeanWindow,"-",randomForestPartition,"-",randomForestSeed,"-",partition,".csv"))
      }
    }
  }
}

for (m in models) {
  save(m,file=paste0(logDataDirPath,"/model-",m$name,".dat"))
}
```

Table 14: Random forest prediction model generation

(AREA_Alex_Jonas_nm_rollmean_parallel_seeds_and_ntrees.R)

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

library("parallel")

## check platform
if (.Platform$OS.type=="windows") {
  logDataDirPath <- "C:/Dropbox/AREA/HiS - Skövde/Jonas/Loggs 2013-11-07";
} else {
  logDataDirPath <- "~/QTAGGFiles/AREA/HiS - Skövde/Jonas/Loggs 2013-11-07";
}

cluster<-makePSOCKcluster(12,outfile=paste0(logDataDirPath,"/banzai.txt")) # 12 cores

processAttempt <- function(attemptParameter=stop("Provide attemptParameters")) {

  ## map of target to potential influencing variables
  potentialTargetInfluencers <- list()
  potentialTargetInfluencers[["Fuel.Cons.l.per.nm"]]
  c(
    "Depth.Below.Keel",
    "Eng.Spd",
    "Prop.Ship.Spd",
    "Roll.Active2",
    "Roll.Amp",
    "Roll.Period",
    "Rudder.Pos",
    "Scav.Air.Pres",
    "Sea.Current",
    "Speed.Over.Ground",
    "Wind.Head",
    "Wind.Side"
  )

  potentialTargetInfluencers[["Roll.Amp"]]

  # Function to process data
  processData <- function(data) {
    # Data processing logic here
  }

  # Process data for each attempt
  for (i in 1:length(attemptParameter)) {
    attemptParameters <- attemptParameter[i]
    data <- read.csv(paste0(logDataDirPath,"/data_",as.character(i),".csv"))
    processedData <- processData(data)
    # Save processed data
  }
}

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

c("Depth.Below.Keel",
"Eng.Spd",
"Prop.Ship.Spd",
"Prop.Slip",
"Roll.Active2",
"Roll.Period",
"Rudder.Pos",
"Scav.Air.Pres",
"Sea.Current",
"Speed.Over.Ground",
"Wind.Head",
"Wind.Side"
)
potentialTargetInfluencers["Prop.Slip"] <-
c("Depth.Below.Keel",
"Fuel.Cons.l.per.nm",
"Prop.Ship.Spd",
"Roll.Active2",
"Roll.Amp",
"Roll.Period",
"Rudder.Pos",
"Sea.Current",
"Speed.Over.Ground",
"Wind.Head",
"Wind.Side"
)

## bounds on different kind of potential influencing variables
## N.B., sample() require integers, min and max are integers
## after a sample has been drawn, it is divided by the divisor

bounds<--list()
## depth below keel 0.8 to 159.4
bounds["Depth.Below.Keel"]<-list(min=8,max=1594,divisor=10)


63 (74)
## engine speed 60.95 to 71.85 rpm
bounds(["Eng.Spd"])<-list(min=6095,max=7185,divisor=100)

## fuel consumption liter per nautic mile: 15.835 to 27.052
bounds(["Fuel.Cons.l.per.nm"])<-list(min=15835,max=27052,divisor=1000)

## propeller ship speed 15.00 to 20.00
bounds(["Prop.Ship.Spd"])<-list(min=1500,max=2000,divisor=100)

## propeller slip: 3.00 to 12.00
bounds(["Prop.Slip"])<-list(min=300,max=1200,divisor=100)

## roll amplitude: 0.1 to 10.9
bounds(["Roll.Amp"])<-list(min=1,max=109,divisor=10)

## roll period: 10.00 to 20.00
bounds(["Roll.Period"])<-list(min=1000,max=2000,divisor=100)

## rudder position: 0.00 to 5.00
bounds(["Rudder.Pos"])<-list(min=0,max=500,divisor=100)

## scav air pressure: 0.46 to 0.70
bounds(["Scav.Air.Pres"])<- list(min=46,max=70,divisor=100)

## sea current: -3.000 to 3.000
bounds(["Sea.Current"])<- list(min=-3000,max=3000,divisor=1000)

## speed over ground: 0.0 to 20.0
bounds(["Speed.Over.Ground"])<-list(min=0,max=200,divisor=10)

## head wind: -8.00 to 8.00
bounds(["Wind.Head"])<- list(min=-800,max=800,divisor=100)

## side wind: -8.00 to 8.00
bounds(["Wind.Side"])<- list(min=-800,max=800,divisor=100)

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

configuration<-c("Roll.Active2")

## generate uniform sample for potential influencers of
## the specified target variable
genSample <- function(target=stop("A target must be specified")) {
  result <- list()
  if (! target %in% names(potentialTargetInfluencers)) {
    stop(paste0(target,": No such target"))
  }
  for (n in potentialTargetInfluencers[[target]]) {
    if (! n %in% configuration) {
      result[[n]]<-sample(bounds[[n]]"min",bounds[[n]]"max",1)/bounds[[n]]"divisor"
    }
  }
  return(result)
}

print(attemptParameter)

## check platform
if (.Platform$OS.type=="windows") {
  logDataDirPath <- "C:/Dropbox/AREA/HiS - Skövde/Jonas/Loggs 2013-11-07";
} else {
  logDataDirPath <- "~/QTAGGFiles/AREA/HiS - Skövde/Jonas/Loggs 2013-11-07";
}

## iterate over configurations
## for each model associated with a random seed for the partitioning
## of training/testing data, run through a number of samples
## generate predictions with optimization turned on and optimization
## turned off. Save to a file.
outFileName <- paste0(logDataDirPath,"/prediction-",attemptParameter["attemptName",].dat)
count<0
result<-list()
for (rsf in attemptParameter["randomForestSeeds"])) {
  fileName <- paste0(logDataDirPath,"/model-",attemptParameter["attemptName",","",rsf,.dat")
}

65 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
load(fileName)
cat("Loaded ",fileName,"\n")
predOn=c()
predOff=c()
for (i in 1:attemptParameter["noOfSamples"]) {
  s<-genSample(attemptParameter["target"])
  son<-s
  soff<-s
  son["Roll.Active2"]<-1
  soff["Roll.Active2"]<-0
  pon<-predict(m["model"],son)
  poff<-predict(m["model"],soff)
predOn<-c(predOn,pon)
predOff<-c(predOff,poff)
}
result[[rsf]]<-cbind(predOn,predOff)
}
save(result,file=outFileName)

return(result)
}

## test cases
testcase <- c("2","5","6","7")

## target variables
targets<- c("Fuel.Cons.l.per.nm","Roll.Amp","Prop.Slip")

## roll mean sampling window
rollMeanWindows=c(2,4,8,16,32)

## partitioning factors for random forest
randomForestPartitioningFactors <- c(3,5,7,11)
```


66 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

## seed values for partitioning the data set into training/testing data
randomForestSeeds <- c(1,3,5,7,11)

## build list of parameter configurations
attemptParameter<-list()
noOfSamples<-1000
for (tc in testcase) {
  for (t in targets) {
    for (rmw in rollMeanWindows) {
      for (rfpf in randomForestPartitioningFactors) {
        ## name without seed
        attemptName <- paste0(tc,"-",t,"-","all","-",rmw,"-",rfpf)
        ap<-list(
          attemptName=attemptName,
          testCase=tc,
          target=t,
          rollMeanWindow=rmw,
          randomForestPartitioningFactor=rfpf,
          randomForestSeeds=randomForestSeeds,
          noOfSamples=noOfSamples
        )
        attemptParameter<-append(attemptParameter,list(ap))
      }
    }
  }
}

## apply processing on list of parameter configurations
parLapply(cluster,attemptParameter,processAttempt)

Table 15: Prediction evaluation (genRandomValues2.R)

67 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
## test cases
testcase <- c("2","5","6","7")

## target variables
targets<- c("Fuel.Cons.l.per.nm","Roll.Amp","Prop.Slip")

## roll mean sampling window
rollMeanWindows=c(2,4,8,16,32)

## partitioning factors for random forest
randomForestPartitioningFactors <- c(3,5,7,11)

## seed values for partitioning the data set into training/testing data
randomForestSeeds <- c(1,3,5,7,11)

library("parallel")

## check platform
if (.Platform$OS.type=="windows") {
  logDataDirPath <- "C:/Dropbox/AREA/HiS - Skövde/Jonas/Loggs 2013-11-07";
}
```

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study


69 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

Funded by: Vinnova Forska & Väx 2013, “System för bränslebesparing på stora fartyg”, 2013-00301 70 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

\[
\text{rate} \leftarrow \frac{\text{result}[\text{rsf}][1]}{\text{result}[\text{rsf}][2]}
\]

\[
\text{fileName} \leftarrow \text{paste0}(\logDataDirPath,"/density-",\text{attemptParameter["attemptName"]}, ",","\text{rsf","-optimization_on_divided_by_off"}, ",".png")
\]

\[
\text{png(file=fileName)}
\]

\[
\text{plot(density(rate),main=paste0("Prediction sample:","\text{attemptParameter["attemptName"]},","\text{seed:","rsf","-on/off"})}
\]

\[
\text{dev.off()}
\]

\[
\text{allRate} \leftarrow \text{append(allRate,rate)}
\]

The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
q <- quantile(allRate,probs=seq(0,1,0.05))
lessThanOne <- (length(q[which(q<1.0)])-1)*0.05
if (lessThanOne<0) {
  lessThanOne<-0
}
greaterThanOne <- (length(q[which(q>1.0)])-1)*0.05
if (greaterThanOne<0) {
  greaterThanOne<-0
}
res <- list(configuration=attemptParameter,c(lessThanOne,greaterThanOne))
cat("<RESULT>
")
print(attemptParameter)
print(res)
cat("</RESULT>
")
return(res)
```

### build list of parameter configurations

attemptParameter<-list()
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
noOfSamples<-1000
for (tc in testcase) {
    for (t in targets) {
        for (rmw in rollMeanWindows) {
            for (rfpf in randomForestPartitioningFactors) {
                ## name without seed
                attemptName <- paste0(tc, "\_\_\_" , t, "\_\_\_" , "all\_\_\_" , rmw, "\_\_\_" , rfpf)
                ap<-list(
                    attemptName=attemptName,
                    testCase=tc,
                    target=t,
                    rollMeanWindow=rmw,
                    randomForestPartitioningFactor=rfpf,
                    randomForestSeeds=randomForestSeeds,
                    noOfSamples=noOfSamples
                )
                attemptParameter<-append(attemptParameter,list(ap))
            }
        }
    }
}
```


73 (74)
The effect of optimizing engine control on fuel consumption and roll amplitude in ocean-going vessels: An experimental study

```r
## apply processing on list of parameter configurations
proportions <- parLapply(cluster, attemptParameter, generateDensityGraphs)
## proportions <- lapply(attemptParameter, generateDensityGraphs)
stopCluster(cluster)
```

Table 16: Generate probability density graphs based on Monte Carlo simulation and perform summary statistics analysis based on vigintiles