Davie, S., Minto, C., Officer, R., Lordan, C., Jackson, E. 2015. Modelling fuel consumption of fishing vessels for predictive use. ICES Journal of Marine Science, 72: 708-719.

Modelling fuel consumption of fishing vessels for predictive use

Sarah Davie[1*], Cóilín Minto[2], Rick Officer[2], Colm Lordan[1], Emmet Jackson[3]

[1] Fisheries Ecosystems Advisory Services, Marine Institute, Oranmore, Galway, Ireland

[2] Marine and Freshwater Research Centre, Galway-Mayo Institute of Technology, Dublin Road, Galway, Ireland

[3] Resource Development and Environment Section, Bord Iascaigh Mhara (BIM), Crofton Road, Dun Laoghaire, Co. Dublin, Ireland

[*] Corresponding author: tel: +353 91 387200; fax: +353 91 387201; email: sarah.davie@marine.ie

1.1 Abstract

Fuel costs are an important element in models used to analyse and predict fisher behaviour for application within the wider mixed fisheries and ecosystem approaches to management. This investigation explored the predictive capability of linear and generalised additive models (GAMs) in providing daily fuel consumption estimates for fishing vessels given knowledge of their length, engine power, fleet segment (annual dominant gear type) and fuel prices. Models were fitted to half of the Irish fishing vessel economic data collected between 2003 and 2011. The predictive capabilities of the seven best models were validated against the remaining, previously un-modelled, data.

The type of gear used by a fleet segment had an important influence on fuel consumption as did the price of fuel. The passive pot gear and Scottish seine gear segments indicated

consistently lower consumptions, while dredge and pelagic gears showed consistently higher fuel consumptions. Furthermore, increasing fuel price negatively affected fuel consumption, especially for more powerful, larger vessels.

Of the formulated models, the best fit to training data was a GAM with a gear main effect and two smooth functions; standardized vessel length, and engine power interacting with fuel price. For prediction, overall, this model showed the closest predictions with the least bias, followed by three linear models. However, all seven models compared for predictive capability performed well for the most sampled segments (demersal and pelagic trawlers).

1.2 Key words

Fuel consumption; fuel price; fuel cost predictions; fishing vessels; modelling; GAM; fishing gear

1.3 Introduction

Fishing, like any other business, aims to generate profits through achieving greater revenues than costs. Individual fishers hold a detailed understanding of the factors influencing their business, such as fishing location, gear configuration, and fuel costs. Scientists do not have such detailed information and must reconstruct, or predict, this knowledge from the information available.

Fuel represents one of the largest costs associated with individual fishing trips, while the actual proportion attributable to fuel varies greatly between fisheries (Sumaila *et al.*, 2008). Within Hong Kong’s commercial fisheries, fuel amounts to between 30% and 60% of total costs (Sumaila *et al.*, 2007). In South East Australia, trawlers report fuel costs of between 18% and 25%, while the proportions were lower (5%-10%) for Danish seiners

(FERM, 2004). Variation in fuel costs have also been reported among European fisheries: ranging from 15% to 38% of total costs for Irish demersal trawlers varying annually and with vessel length (BIM unpublished economic data); Cheilari *et al.* (2013) state average fuel costs represented 29% of total costs in 2008 across 54 fleet segments; Bastardie *et al.* (2013) detail variation in the fuel costs as a proportion of the value of landings for Danish fisheries between 2005 and 2010.

Since the Arab oil embargo of 1973 (Yergin, 1991), fuel "supply scares" have resulted in rapid fluctuations in fuel prices. The most recent event occurred between early 2007 and mid 2008 when fuel doubled in price. Such scares have prompted analyses of the energetic performance and economic vulnerability of a wide range of fisheries (see Tyedmers, 2001 and Tyedmers *et al.*, 2005 for examples). The most recent price fluctuations stimulated further investigations into fuel use within the fishing sector. From an economic perspective, Cheilari *et al.* (2013) evaluated the economic performance and energy efficiency of the EU fleet. Abernethy *et al.* (2010) examined the impact of fuel price on the structure, behaviour and vulnerability of the UK's southwest fishing fleet. Others have considered increased fuel prices from a more biological perspective. For example, Arnason (2007) conceptualises excessive fishing pressure could be reduced as a result of lower profitability (from higher fuel costs) further hypothesising that such reductions in pressure could aid fish stock recovery. However, Arnason highlights that this can be negated if governments increase fuel subsidies, such as the 38 cent per litre rebate described in Australian trawl fisheries by Chenhall and Magnet (2008). The points made by Arnason (2007) are further supported by Sumaila *et al.* (2008) who believe that positive reductions in fishing pressure due to increased fuel prices are reduced, if not completely negated by increasing fuel subsidies. Such variability illustrates the importance of fuel costs as a driver of fisher choices and behaviour.

The fuel consumption of a fishing vessel varies depending on a variety of factors and conditions, for example vessel size, age, and condition, engine power, vessel speed and gear configuration, sea state and weather conditions (Driscoll & Tyedmers, 2010; Schau *et al.*, 2009; Tyedmers, 2001).

Previous investigations have examined fuel consumption associated with fleets, fisheries, gears and specific species or stocks over time. Schau *et al.* (2009) developed fuel-use coefficients expressed as a value of fish per volume of fuel used. Tyedmers (2001) related the results of fish per fuel volume of various studies to their equivalents in terms of obtainable energy. This conversion into values of protein energy yield (Joules) and output (tonnes) allowed respective comparison between fisheries, and other protein producing sectors such as agriculture. Other studies generate fish per fuel volume values and convert these into their equivalent greenhouse gas emissions (e.g. CO_2 weight per volume of fuel) to address the implications of such emissions (e.g. Driscoll & Tyedmers, 2010).

Bio-economic models that examine the choices and responses of fishers to management impositions within mixed fisheries should include fuel as an important explanatory variable. Many such analyses utilise increasingly complex models to analyse and predict fisher behaviour. The importance of financial drivers to the decision making processes in fisheries is increasingly acknowledged and incorporated (Andersen *et al.*, 2010; Gourguet *et al.*, 2013; Marchal *et al.*, 2011; Ulrich *et al.*, 2007). However, these analyses often do not relate specifically to fuel consumption, but rather incorporate measures of fuel usage (e.g. total expenditure or price per quantity) as proxies for fishing cost. Disaggregating estimates of input variables within fishery simulations, e.g., by vessel length and engine size, should lead to models with increased accuracy and enhanced predictive capabilities.

This investigation utilised annual Irish vessels fuel cost data to estimate linear and generalized additive models to estimate, and subsequently predict, fuel consumption per day for different fleet segments (gears) by vessel length and engine power combinations whilst accounting for variation in fuel prices. Model outputs were designed for subsequent use in decision support tools to inform the development of mixed fisheries management plans by enumerating potential economic consequences and behavioural adaptations in response to management measures.

1.4 Materials and Methods

1.4.1 Data

Europe's implementation of fisheries Data Collection Regulations (DCF; EC, 2001) and Member State's subsequent commitment to the Data Collection Framework (DCF; EC, 2008) has increased the quantity and quality of economic data collected from the fishing sector. Individual Member States are required to collect a variety of detailed economic variables from a sample of the fleet considered representative of the overall fishing sector. More general economic data, such as total fuel costs are also collected.

Economic data on fishing vessels within Ireland are collected by Bord Iascaigh Mhara (BIM) as part of Ireland's commitment to the DCF. Information collected encompasses annual income and expenditure figures, from a sample of individual vessels. Questionnaires are sent to all ≥10m active vessels on an annual basis (~400 annually). Sampled vessels constitute those who completed and returned the questionnaire. These data were used in conjunction with the annual number of days-at-sea associated with the vessel (as the number of days absent from port) available from logbook entries, provided by the Department of Agriculture, Food, and the Marine. Annual average fuel prices per litre were provided by BIM between 2003 and 2011. This price was calculated by converting the Europe Brent Spot Price FOB[1] (Dollars per Barrel) to Euro and litres (159 litres/barrel), adding transport (32%) and handling (4.45%) costs and then averaged over the year (M. Keatinge, BIM, pers. coms.). There is temporal variation in the handling fee rates depending on fuel prices which tend to decrease when prices are high however 4.45% is the best fit (M. Keatinge, BIM, pers. coms.).

Annual estimates of fuel cost (in Euro) per vessel were divided by the vessel's days-at-sea effort within the same year. This resulted in a fuel cost per day-at-sea consumption (defining the "sample" used subsequently throughout the study) regardless of vessel activity (steaming or fishing). Within this investigation fuel consumption rates were assumed to be the same for fishing days and steaming days. Separate information on fuel consumption rates for steaming and fishing were not available for this investigation. Previous studies have shown that only a small proportion of time within fishing trips is spent steaming for the majority of the Irish fleet (Gerritsen & Lordan, 2011; unpublished data). Here fishing trips are recorded by fleet segment, increasing the homogeneity in travel distances between trips for different gear types, for example between those employing pots, demersal trawl gear, or pelagic gears. Furthermore, the intended purpose of this investigation was to provide fuel consumption volumes for fishing trips, internalising this assumption.

The two polyvalent gear classifications, PGO and PMP, contributing 1 and 2 samples respectively were removed from the dataset. Exploratory modelling had resulted in high by-gear leverage for these samples due to the small sample sizes. The resulting dataset contained 655 anonymous records spanning 312 individual vessels. This equates to around 72 vessels annually, covering around 18% of the ≥10m fleet. In terms of days at sea fishing effort, on average sampled vessels generate ~3.5% of total Irish effort. Data details:

[1] Available from: http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RBRTE&f=D

vessel length (rounded down to the nearest 0.5m); engine power (rounded down to the nearest 5kW); vessel annual fleet segment (Table A1; defined using DCF dominance criteria[2]; hereafter referred to as gear); fuel cost per day-at-sea (Euro); and a fuel consumption per day-at-sea (litres) value (here after referred to as fuel per day) derived from per day fuel cost divided by the average overall annual fuel price per litre (detailed in Table A2).

The dataset was split into two groups, a training set on which modelling was carried out (336 samples) and a testing set (319 samples) reserved for assessing the predictive capabilities of the fitted models. This was done using stratified random sampling by gear ensuring adequate samples of each gear type. Vessels employing different gears over time were allocated according to the gear with maximum number of samples, else when equal sample numbers existed the least sampled gear type of the split was assigned. Due to the low number of hook vessels, manual selection of one large and one small vessel was necessary to provide the training model with a vessel length range.

1.4.2 *Analysis*

Methods of data visualization are described first followed by a description of linear and additive models fit to the training per day fuel consumption data. Finally, the predictions of the best fitting set of linear and additive models are compared with the un-modelled test data to assess model predictive capability. All analyses were carried out within the R statistical environment version 2.15.2 (R core development team, 2012) and included use

[2] A vessel is allocated a gear annually based on the gear with the highest number of fishing days within the year (i.e. over 50% of fishing days), if no gear dominates the vessel is allocated to one of 3 polyvalent segments (all mobile gears, all passive gears, mixed mobile and passive gears), from http://datacollection.jrc.ec.europa.eu/dcf-fish/eco/dsgr visited 12/03/2013

of the following specific functions and packages: dredge (MuMIn; Barton, 2013) and gam (mgcv; Wood, 2006), normalmixEM (mixtools; Benaglia *et al.*, 2009).

1.4.3 Data visualisation

Fuel consumption per day values of sampled vessels 2003-2011 were visualised by gear for vessel length and engine power (Figure 1a and b; supplementary Figures S1 and S2 depict gears separately) to examine relationships between fuel usage and vessel characteristics. A clear relationship is observed between vessel length and engine power when plotted by gear (Figure 1c, depicted separately in supplementary Figure S3) highlighting a correlation which needs to be addressed within the modelling framework.

Fuel per day plots (not shown) by vessel length and engine power indicated power-curve relationships between consumption and vessel characteristics, with increasing variability with mean response in a log-normal fashion. This suggests a log-linear model with normally distributed errors on the log scale may be appropriate. Such a relationship will down-weight the influence of more extreme samples.

1.4.4 Modelling

Log-linear models

Based on the log-linear relationship identified during data visualisation the continuous variables of vessel length and engine power were converted to natural logarithmic values. This was also done for the response variable, fuel consumption per day. Preliminary linear model fits were carried out using all fuel consumption records primarily to determine the importance of year (as a categorical variable)in the presence of vessel length, engine power, and fuel price per litre. Using the Akaike Information Criterion (AIC, Akaike (1974)) to determine goodness of fit and comparison, these preliminary models identified inclusion of vessel length and engine power to be significantly better descriptors than either variable alone. Year was also determined to significantly improve model fit. However, the inclusion of fuel price per litre resulted in a significantly better model fit than the inclusion of a year affect, or both year and fuel price. Varying fuel prices were therefore assumed to be the dominant cause of the year effect. This is intuitive given vessels are unlikely to alter annually in length or in engine power beyond minor alterations in efficiency. Although it is possible for engines to deteriorate over time due to age and maintenance, if this had been the case within the available dataset, inclusion of both a year effect and fuel price may have resulted in a better fitting model.

The training 50% subset of sample data were used to develop a set of candidate models for predicting per day fuel consumption based on the main effects identified within the year effect analysis: vessel length and engine power characteristics for different gear types as well as the average price of fuel. The following initial log-linear model of fuel per day by length, power, gear, and fuel price was applied:

$$\ln(F_i) = a_{0,G[i]} + a_1 \ln(L_i) + a_2 \ln(E_i) + a_3 P_i + a_4 \ln(L_i) \ln(E_i) + \varepsilon_i \quad (1)$$

Where F_i is fuel consumption per day for the ith observation (i.e., average fuel per day for a given year and vessel), $a_{0,G[i]}$ is the intercept which varies by gear $G[i]$, L_i is vessel length, E_i is engine power, P_i is average fuel price per litre, G is gear (categorical variable with 8 levels: Table A1). A high level of correlation was identified between the variables vessel length and engine power (0.73; Figure 1), which was reflected in high correlation of the parameter estimates. The variables were standardized by subtracting the mean and dividing by the standard deviation. The model was expanded to include the full 4-way interaction, described as:

$$\ln(F_i) = a_{0,G[i]} + a_{1,G[i]} \ln(SL_i) + a_{2,G[i]} \ln(SE_i) + a_{3,G[i]} P_i + a_{4,G[i]} \ln(SL_i) \ln(SE_i) P_i + \varepsilon_i$$

(2)

Where SL_i and SE_i are the standardised vessel length and engine power, respectively. Within this model there were 167 possible sub-model combinations for the four variables. The two best fitting of these models (given by low AIC values) are presented within Table 1, while Table A3 lists the top 20. Examination of the scatter around the relationship between the predicted and observed fuel consumption values (Figure 2) suggest data may not conform to a strict linear relationship.

Generalized additive models

To investigate deviation from linearity assumptions, a series of three different types of Generalized Additive Model (GAM) (Hastie & Tibshirani, 1986, 1990; Wood, 2006) with integrated smoothness estimations were fitted, for example:

$$\ln(F_i) = a_{0,G[i]} + \beta_1 P_i + s_1(\ln(SL_i)) + s_2(\ln(SE_i)) + \varepsilon_i \quad (3)$$

Where s_1 and s_2 are smoother functions (thin plate regression splines). Examination of the residuals indicated that although the GAM model provided a good fit for the data, residual patterns were present. This indicated that not all patterns within the data were accounted for by the covariates using low basis dimension smoothing. This could be remedied by increasing the space over which the smoothers could operate (within R this is done by increasing the k value to a large number, such as 100). However, rather than increasing the degrees of freedom allowed for the smoothers, it was considered more appropriate to explore alternative formulations to account for these patterns. This increases the model complexity, whereby a series of GAMs were applied accounting for interactions between parameters within a smoother term. Including the interaction between, for example, vessel length and engine power modelled as:

$$\ln(F_i) = a_{0,G[i]} + \beta_1 P_i + s_1(\ln(SL_i), \ln(SE_i)) + \varepsilon_i \quad (4)$$

Where s_1 here is a 2-dimensional surface thin-plate spline (Wood, 2006). Preliminary investigations into smoother effect of fuel price indicated that a linear effect was adequate. The final fit trials included gear as part of the smoothing function of the GAM. This was trialled for both types of GAM models. Expanding the examples above, this gives a 2 smoother model (Equation 5) and a 2-dimensional smoother model (Equation 6):

$$\ln(F_i) = a_{0,G[i]} + \beta_1 P_i + s_{1,G[i]}(\ln(SL_i)) + s_{2,G[i]}(\ln(SE_i)) + \varepsilon_i \quad (5)$$

$$\ln(F_i) = a_{0,G[i]} + \beta_1 P_i + s_{1,G[i]}(\ln(SL_i), \ln(SE_i)) + \varepsilon_i \quad (6)$$

The two best fitting GAM models (according to AIC values) from those trialled are presented in Table 1 and a summary of the top 20 GAM models trialled is given in Table A4.

1.4.5 Model testing

To test the ability of formulated models to predict out-of-sample per day fuel consumption, the fuel prices per litre, vessel lengths, engine powers, and gears were taken from the 319 samples within the test dataset. Vessel length and engine power of the test samples were logged and standardised using the mean and standard deviation values obtained during standardisation of the training dataset to which the models were fitted. Predictions were made for all linear and GAM models. Proportional errors (predicted consumption - observed consumption)/observed consumption), absolute proportional errors and their means (MAPE; Abramowitz & Stegun, 1972) were calculated to compare the predictive capability.

1.5 Results

1.5.1 Modelling

Interestingly within the linear models of the training dataset, fuel price was not a strong explanatory variable. As can be seen from Table 1, the linear model LM.1 does not contain fuel price as an explanatory variable, whilst LM.2 with less than 1 AIC point difference, does. There is no significant difference in the explained variance between these two models (ANOVA: $p = 0.23$). A large number of GAM models (Table A4) achieved lower AIC values than the series of linear models fitted (Table A3).

GAM models which contained by gear considerations within a smoother (such as those in Equations 5 and 6) out performed those that considered gear only as a main effect. This included the two best AIC performing GAMs (GAM.1 and GAM.2 in Table 1), both of which contained a 2-dimensional smoother. GAM.1 containing two smoother functions showed the best fit to the test fuel consumption data. The first smoother was over log standardised engine power, the other smoothing over an interaction between log standardised vessel length and fuel price varying by gear.

These models result in very complex smoother terms with a large number of parameters using a large number of degrees of freedom. The large standard errors observed for a number of these smoothers are likely the result of the models being unable to determine underlying trends from the low sample numbers at such disaggregation levels. GAM.3 is the best fitting GAM without the inclusion of gear within the smoother, representing a much simpler model with a far fewer parameters (Table 2 lists gear coefficients for this model). The AIC of this model is still lower than that of the best performing linear models. The three contour plots in Figure 3 were used to visually compare the variability between the two more complex GAM models and the simpler GAM.3 for DST gear as the best sampled and mid fuel price range.

The diagnostic plots for this simpler GAM.3 model (Figure S4) indicate there is some deviation of residuals from the normal distribution, a similar pattern observed within many of the GAM model residuals. This was investigated post-hoc by applying a scale finite mixture model (McLachlan & Peel, 2000) to the residuals. The mixture model indicated two distributions were present within the residuals: ~80% of which were normally distributed with a small variance, with the remaining 20% constituting a more dispersed distribution (Figure 4).

1.5.2 Model testing

The sum of absolute proportional errors per model was used to determine the most accurate linear and GAM models. These two models were named LM.PE and GAM.PE (detailed in Table 1). The proportional errors around predictions were visually compared for each of the, now seven, candidate models in Table 1 (Figure 5). Predictions for each of these seven models on average, overestimate fuel consumption per day with the exception of Scottish seine gear (SSC) which consistently exhibited slight underestimates.

Demersal (DTS) and pelagic (TM) trawl gears show the most accurate and consistent fuel consumption predictions over the majority of comparisons (Figure 5). These two gear types had the greatest sample levels within the training and testing datasets. There is greater variability between model consumption estimates, in particular GAM.1, GAM.2 and GAM.PE, for gillnet (DFN), dredge (DRB), and to a lesser extent pots (FPO) gear. GAM.2 had a greater tendency toward under predicting fuel consumption for gillnet and pots gear. GAM.3 shows continuity with linear models. With the exception of pelagic gear

(TM) LM.PE model typically produced slightly lower mean consumptions than the other linear models.

A series of per day fuel consumption predictions were made over the range of observed fuel prices for each of the seven models detailed in Table 1. Three average Irish vessel sizes were chosen for predictions: large (24m with a 397kW engine); medium (16m with 177kW); and small (12m with 104kW). GAM models incorporating gear within smoother terms generated unusually high predictions of fuel consumption for at least one gear-vessel size combination. In these cases such high fuel consumption predictions indicate models were over-fitting to individual data points within the training dataset. For GAM.1 this occurred for beam trawl gears for each vessel length. For GAM.2 it occurred for the largest vessel size with gillnets, whilst for GAM.PE pots had extremely large predictions of consumption.

Two general trends were identified across model fuel consumption predictions. Firstly, with increasing vessel size greater volumes of fuel were used, as would be expected given increasing average engine size with vessel length. Secondly fuel consumption varied by gear type, the difference of which is more prominent at larger vessel sizes. Scottish seines (SSC) and pots (FPO) show relatively low consumption compared to pelagic trawls (TM) and dredging gears (DBR) where consumption is high.

GAM.3 performs more in line with linear model fuel consumption per day predictions (Figure 6) without spuriously high predictions. Predictions at the small vessel size indicate little variation between models and demonstrate minimal influence of varying fuel price on fuel consumption. For GAM.3 and LM.2, which reflect very similar trends, fuel price has a greater, negative, influence on fuel consumption across gear types as vessel size increases. These two models incorporate a fuel price effect. With increasing vessel size model LM.1, which does not contain this variable, diverges from these models maintaining a static trend across fuel prices. LM.PE generated some unlikely positive fuel consumption trends with increasing fuel price, particularly for dredges and beam trawls. These trends are likely to result from the interaction between gear and fuel price included within the model.

1.6 Discussion

1.6.1 Model

A general additive model incorporating two smoother terms, one encompassing vessel length, the other a 2-dimensional smoother of engine power and fuel price (GAM.3) was found to be the most appropriate fuel consumption descriptor of the linear and GAM models explored here.

A number of GAM models provided better AIC values, these models were more complex including a gear type effect on one or more smoother terms, greatly increasing the number of parameters. These models were considered to be over-fitted to the specific data points within the training data, in part due to the small subsample sizes these models subsequently fitted to (this included GAM.1 and GAM.2). Although a series of interactions may prove to be highly significant in a statistical stance, such as the above models, the scientific interpretation can be difficult, and rational explanation for higher-level interactions may not be evident (Maunder and Punt, 2004). Many of the fitted models, both linear and GAM, highlighted a more complex relationship than a simple scaling of fuel consumption between gear, fuel price, and vessel size characteristics. Such variation in fuel consumption between gear and fishing practice were observed in several other studies (Schau *et al.*, 2009; Tyedmers *et al.*, 2005; Winther *et al.*, 2009), as was the price of fuel (Poos *et al.*, 2013). Model GAM.3 is consistent with those studies reviewed by Tyedmers (2001), and more recently by Abernethy *et al.* (2010), who identified higher fuel consumption with towed gears and larger vessels. Bastardie *et al.* (2013) found seiners out-competed trawlers when targeting the same species. This finding is in line with the results of our model, in which Scottish seiners have lower fuel consumption than those of demersal trawlers. Fuel consumption variation between gears is understandable, given that while towing gear ~95% of fuel is used to tow the gear with the remainder propelling the vessel (BIM, 2009). This penalty is not incurred by vessels setting passive gears and leaving them for a period of time (e.g. pots). Longlining is an exception within the passive gear group. In the current analysis longlining was shown to have higher fuel consumption than other passive type gears. Tyedmers (2001) also found longlines to have a higher fuel consumption (litres per HP*sea day) than other gear types, including a combined trawl and dredge group. In relation to Irish longlining, increased fuel consumption in comparison to other passive gears may relate to the typically more offshore fishing grounds exploited by longliners requiring greater steaming distances. Furthermore, the nature of longline gear deployment and retrieval in conjunction to location of fishing grounds result in longline vessels tending to be larger than those employed in gillnetting or potting, giving rise to a greater energy (fuel) input. Longlining was found by Tyedmers (2001) to have higher energy intensity (litres/tonne) than other passive gears, due to the relatively high energy inputs (fuel) and low levels of fish landed (despite their sometimes high monetary value).

Dredge gears were the most fuel demanding fishing method identified within this investigation. Our investigation supports the anecdotal information that this gear is traditionally thought of as a high fuel consumer. This gear incurs the resistance of sea floor sediment during towing, resulting in a greater drag resistance on the boat moving forward, higher than if the gear were just passing over the surface of the seabed, and thus requiring greater fuel volumes. Consumption rates for pelagic gears were close to that of the dredges. The high fuel intensity indicated for pelagic gears may relate to high volume catches entering the mid-water net creating greater drag and additional effort maintaining position in the water column with the additional weight of catch. However, as our results are calculated on a per day at sea basis, a more plausible explanation would be that pelagic vessels exert substantially more effort searching for fish shoals, and also travel greater distances in often higher powered vessels (as also suggested by Schau *et al.*, 2009 and Winther *et al.*, 2009). Furthermore, greater cruising speeds to reach markets faster and thus provide a fresher, more valuable product (Reid *et al.*, 2011) would require greater fuel consumption, as would running the seawater refrigeration units which many pelagic vessels possess (Reid *et al.*, 2011).

Residual distributions from all the models evaluated indicate a level of violation to assumed normality. An investigation of the residual distribution, through the application of scale mixture models, indicated that two distributions were potentially present; one of normal distribution (mean of zero and standard deviation 0.392), the other containing broader tails (mean of zero and standard deviation 1.281). The mixture proportions were identified as ~80% normal with ~20% over dispersion contamination. This could result from omitted variables (e.g., engine age) or varying levels of accuracy in the returned estimates of annual fuel and effort usage. This reporting could result from a number of sources including submission of under- or over- estimated annual fuel cost, effort reported compared to actual effort, or deviation from the average fuel price. Cost data are presented by accountants and thus have a higher likelihood of being accurate representations of annual fuel costs. Fuel data may be distorted by the application of an average fuel price, if the prices paid by a fisher constantly varied from the average. The use of a "dominant" gear type may introduce some variability within the vessel’s fuel consumption, whereby a vessel may be classed as one dominant gear but have a percentage of time utilising an alternative gear. In addition, although the reporting variable for fuel excludes lubrication oil, some vessels may report it within the total. Furthermore, distortion may result from inaccuracies within the reporting of days-at-sea effort within the logbooks. The non-normal distribution could further be investigated through the application of mixed distribution models at the modelling stage. Whilst such approaches are at the forefront of CPUE modelling research (Thorson *et al.*, 2012), this type of modelling was beyond the scope of the present investigation. Particularly given the relatively low level of contamination, and would be more likely to affect uncertainty rather than the mean parameter values used for prediction here. Investigation of mixture models for this type of analyses may be a fruitful avenue for further research.

During the initial stages of investigation raw data indicated a power curve relationship between the vessel characteristics scaled by gear types. However the formal log linear relationship of fitted linear models (Table A3) appeared too restrictive to adequately describe the relationships within our data. The flexibility of the linear relationship within GAMs provided a more appropriate fit. The complexity of GAM models applied increased highlighting that better fitting models were possible (the top 18 of the 50 fitted GAM models here) but that these may have been over fitted to the specific variability within the training data rather than capturing persistent trends and effects. Whilst application of GAMs with gear type within the smoother term(s) resulted in better fits to the data, this was achieved at the expense of increased complexity and a reduced ability to predict unknown fuel consumptions, the overall goal of the investigation.

1.6.2 Prediction

Overall three models demonstrated better performance than other candidate models, both in predicting from previously un-modelled data and for an average set of vessels across fuel prices. Candidate models were tested on roughly half of available samples across all fleet segments increasing confidence in average predictions. This included linear models

LM.2 and LM.PE, although the later had the highest AIC of the seven prediction models. Furthermore, the underlying validity of applying a strict linear relationship was questionable, and highlighted by the better fit of GAM models to the training data.

The other was GAM.3, the simpler GAM without gear variation within the smoother terms. As in model selection, GAM.3 was chosen as the most appropriate for predicting fuel consumption. A large number of degrees of freedom can weaken the predictive power of models (Burnham & Anderson, 2002; Hastie et al., 2001), as was the case for GAM.1 and GAM.2, supporting the conclusion that the GAM models whose smoothers varied by gear were over-fitted to specific data points within the training data.

Models varied in their ability to predict fuel consumption between gears, most of which resulted in over-estimation bias (with the exception of Scottish seines). Using this knowledge, bias can be accounted for if using the models for predictive purposes. This would indicate that there are further explanatory elements not considered within our models, which explained between 58-74% of variation within the training data. However, these are the most widely available variables. The least variable predictions were for demersal (DTS) and pelagic (TM) trawl gears (MAPEs of 0.33 and 0.34 respectively). The quantity of the raw explanatory variables may also play a part in the over estimation bias. These groups contained the greatest sample numbers within the training data. These fleet segments also represent the largest capacity within Ireland, with the demersal trawl fleet receiving the greatest research focus. Whilst several of the candidate models applied to the testing data showed poor capability to predict gillnet (DFN) and dredge (DRB) gear fuel consumptions. These are a comparatively lesser sampled gear types from a segment with a diverse range of vessel sizes, both containing just one vessel length sample at the upper end of the range in the training set, while predicting for several. Such restriction in the training set may result in high level uncertainty in the upper vessel length range. This highlights the importance of sample size and that of collecting data across the whole fleet, not just from vessels of primary interest.

1.6.3 Perspectives

Previous fisheries energy consumption and emission studies have often focussed on small sample numbers of interviewed fishers and have been conducted for specific purposes. The results from such studies are presented as fuel usage per fish weight landed (e.g. Schau *et al.*, 2009), equivalent emissions per fuel usage (e.g. Driscoll & Tyedmers, 2010; Ziegler & Hansson, 2003), or expressed in terms of energy (e.g. Cheilari *et al.*, 2013; Tyedmers, 2001). Such estimations can be relative and changeable over time for numerous reasons including fluctuating species abundance and/or fuel prices (Schau *et al.*, 2009). Such alternative approaches may produce values with greater insight into the overall impact of fuel efficiency but they do not readily lend themselves for manipulation into input variables for alternative applications. This investigation however, took a more general modelling perspective to facilitate prediction of fuel consumption to the wider fleet, through usage of the more general data unit of fuel per day (litres/sea day) consumption rates, designed to generate fuel consumption figures per fishing trip. If desired, the consumption rates generated here can still be utilised as the basis for producing energy efficiency and emission estimates that form the focus of other studies.

Furthermore, our fuel consumption prediction outputs can also be used to generate fuel consumption figures at the vessel, or fleet level for integration as an economic variable within mixed fisheries bio-economic models on fisher choice and behaviour in which the economics of fuel use and price is becoming a more widely acknowledged driver. For example Suuronen et al. (2012) note that while fuel prices increase the fishing industry will suffer losses in profitability, with some conventional bottom trawl, beam trawl, and dredge fisheries becoming economically unviable, and forcing fishers to consider changes to their fishing practices. They argue that fuel consumption and costs of the fishing sector could be substantially lowered by adoption of low impact and fuel efficient technological improvements, and also behavioural adaptations. Behavioural adaptation to rising fuel costs was examined by Poos et al. (2013) within the Dutch beam trawl fleet. Through modelling the trade off between fuel savings and catch losses Poos et al. (2013) focussed on vessels adapting their speed to reduce fuel consumption. An Irish guide designed to advise the Irish fishing industry on energy efficiencies (BIM 2009) also refers to the determination of optimal speeds for highest fuel efficiency. The integration of fuel costs into fisher choice and decision models is thus needed, and demonstrated by Bastardie *et al.* (2013). The availability of a model capable of predicting fuel consumption will also enable fuel consumption and fuel price to be incorporated, and varied, independently within bio-economic models and response simulations. Such applications will likely increase the ability and utility of such models for predicting choices in fishing behaviour.

1.7 Conclusion

The GAM model, GAM.3, constructed within this analysis was capable of estimating fuel per day consumption for several different fleet segments (gears) utilising vessel length, engine power, and fuel price. This model included a gear type main effect, an independent smoother term for vessel length and a 2-dimensional smoother of engine size and fuel price. The interactions between gear type, vessel size, and fuel price have an important influence on fuel consumption. The greatest difference occurs between dredge and towed pelagic gears with high consumption compared to pot and Scottish seine gears with relatively low consumption. These daily fuel consumption predictions could be used for existing applications (such as translation into abundance calculations per litre, or emissions estimates) or used to estimate the fuel component of running costs within bio-economic models designed to examine drivers of fisher and fleet behaviour.

1.8 Acknowledgements

The authors would like to thank the Irish Department of Agriculture, Food and the Marine for providing logbook information and Bord Iascaigh Mhara for providing fuel cost data. Much of the data collection and analysis carried out here was supported by the European Commission through the Data Collection Framework. This study, in part, was supported by the EU European Regional Development Fund (ERDF) through Atlantic Area Programme funding of the GEPETO project. We are grateful to two anonymous reviewers who greatly improved the quality of our analyses and interpretation. Authorship sequence reflects levels of contribution through the “sequence-determines-credit” approach. Author contributions: data configuration: SD, EJ; data analyses: SD, CM, EJ; manuscript: SD, CM, RO, CL, EJ.

1.9 References

Abramowitz, M. and Stegun, I.A. 1972. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover, New York.

Abernethy, K. E., Trebilcock, P., Kebede, B., Allison, E. H., and Dulvy, N. K. 2010. Fuelling the decline in UK fishing communities? ICES Journal of Marine Science, 67: 1076-1085.

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19: 716-723.

Andersen, B. S., Vermard, Y., Ulrich, C., Hutton, T., and Poos, J. J. 2010. Challenges in integrating short-term behaviour in a mixed-fishery Management Strategies Evaluation frame: A case study of the North Sea flatfish fishery. Fisheries Research, 102: 26-40.

Arnason, R. 2007. The Economics of Rising Fuel Costs and European Fisheries. Euro Choices, 6: 22-29.

Barton, K. 2013. MuMIn: Multi-model inference. R package version 1.9.5. http://CRAN.R-project.org/package=MuMIn

Bastardie, F., Nielsen, J. R., Andersen, B. S., and Eigaard, O. R. 2013. Integrating individual trip planning in energy efficiency – Building decision tree models for Danish fisheries. Fisheries Research, 143: 119-130.

Benaglia, T., Chauveau, D., Hunter, D.R., and Young, D. 2009. mixtools: An R Package for Analyzing Finite Mixture Models. Journal of Statistical Software, 32: 1-29. http://www.jstatsoft.org/v32/i06/.

BIM, 2009. Improving fuel efficiency on fishing vessels. BIM user friendly guides, 7. 24 pp. ISSN 1649-9395. ISBN 978-1-903-412-36-7.

Burnham, K.P., and Anderson, D.R. 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd edition. Springer, New York, 496 pp..

Chenhall, S. and Magnet, E. 2008. Fuel for thought. AFMA update, 5(20). ISSN 1449-7697 ABN 8109-849-7517 http://www.afma.gov.au/afma_update/docs/update_0520/update_0520.htm (last accessed 17 May 2013).

Cheilari, A., Guillen, J., Damalas, D., and Barbas, T. 2013. Effects of the fuel price crisis on the energy efficiency and the economic performance of the European Union fishing fleets. Marine Policy, 40: 18-24.

Driscoll, J. and Tyedmers, P. 2010. Fuel use and greenhouse gas emission implications of fisheries management: the case of the New England Atlantic herring fishery. Marine Policy, 34: 353-359.

EC, 2001. Commission Regulation (EC) No. 1639/2001 of July 2001 establishing the minimum and extended Community programmes for the collection of data in the fisheries sector and laying down detailed rules for the application of Council Regulation (EC) No 1543/2000. Official Journal of the European Communities, L222: 53-115.

EC, 2008. Commission Regulation (EC) No 665/2008 of 14 July 2008 laying down detailed rules for the application of Council Regulation (EC) No 199/2008 concerning the establishment of a Community framework for the collection, management and use of data in the fisheries sector and support for scientific advice regarding the Common Fisheries Policy. Official Journal of the European Communities, L186: 3-5.

FERM, 2004. ITQs, ageing boats, and the price of fish. Profitability and autonomous adjustment in the South East Trawl Fishery. Fisheries economics research and management pty. Ltd.; 2004.

Gerritsen, H., and Lordan, C. 2011. Integrating vessel monitoring systems (VMS) data with daily catch data from logbooks to explore the spatial distribution of catch and effort at high resolution. ICES Journal of Marine Science, 68: 245–252.

Gourguet, S., Macher, C., Doyen, L., Thébaud, O., Bertignac, M., and Guyader, O. 2013. Managing mixed fisheries for bio-economic viability. Fisheries Research, 140: 46-62.

Hastie, T. and Tibshirani, R. 1986. Generalized Additive Models. Statistical Science, 1: 297-310.

Hastie, T. and Tibshirani, R. 1990. Generalized additive models (Vol. 43). Chapman & Hall/CRC, Boca Raton, Florida.

Hastie, T.J., Tibshirani, R.J., and Friedman, J. 2001. Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer, New York, 533 pp..

Marchal, P., Little, L. R., and Thébaud, O. 2011. Quota allocation in mixed fisheries: a bioeconomic modelling approach applied to the Channel flatfish fisheries. ICES Journal of Marine Science, 68: 1580-1591.

Maunder, M.N., and Punt, A.E. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research, 70: 141–159.

McLachlan, G.J. and Peel, D. 2000. Finite mixture models. John Wiley, New York.

Poos, J. J., Turenhout, M. N. J., van Oostenbrugge, H., and Rijnsdorp, A. D. 2013. Adaptive response of beam trawl fishers to rising fuel cost. ICES Journal of Marine Science, 70: 675-684.

R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Reid, D.G., Graham, N., Rihan, D.J., Kelly, E., Gatt, I.R., Griffin, F., Garritsen, H.D. and Kynock, R.J. 2011. Do big boats tow big nets? ICES Journal of Marine Science, 68: 1663-1669.

Schau, E.M., Ellingsen, H., Endal, A., and Aanondsen, S.A. 2009. Energy consumption in the Norwegian fisheries. Journal of Cleaner Production, 17: 325-334.

Sumaila, U. R., Cheung, W.W. L., and Teh, L. 2007. Rebuilding Hong Kong’s marine fisheries: an evaluation of management option. Fisheries Centre Research Reports, 15(3). Fisheries Centre, University of British Columbia, Vancouver, Canada. 112 pp.

Sumaila, U. R., Teh, L., Watson, R., Tyedmers, P., and Pauly, D. 2008. Fuel price increase, subsidies, overcapacity, and resource sustainability. ICES Journal of Marine Science, 65: 832-840.

Suuronen, P., Chopin, F., Glass, C., Løkkeborg, S., Matsushita, Y., Queirolo, D., and Rihan, D. 2012. Low impact and fuel efficient fishing – looking beyond the horizon. Fisheries Research, 119: 135–146.

Thorson, J. T., Stewart, I. J., and Punt, A. E. 2012. Development and application of an agent-based model to evaluate methods for estimating relative abundance indices for shoaling fish such as Pacific rockfish (*Sebastes* spp.). ICES Journal of Marine Science, doi:10.1093/icesjms/fss003.

Tyedmers P. 2001. Energy consumed by North Atlantic fisheries. Fisheries Centre Research Reports. *In* Fisheries impacts on North Atlantic ecosystems: catch, effort and national/regional datasets, 9(3). pp. 12-34. Ed. by D. Zeller, R. Watson, and D. Pauly. Fisheries Centre, University of British Columbia, Vancouver.

Tyedmers, P., H., Watson, R., and Pauly, D. 2005. Fuelling global fishing fleets. Ambio, 34: 635-638.

Ulrich, C., Andersen, B. S., Sparre, P. J., and Nielsen, J. R. 2007. TEMAS: fleet-based bio-economic simulation software to evaluate management strategies accounting for fleet behaviour. ICES Journal of Marine Science, 64: 647-651.

Winther, U., Ziegler, F., Skontorp Hognes, E., Emanuelsson, A., Sund, V., and Ellingsen, H. 2009. Carbon footprint and energy use of Norwegian seafood products. SINTEF Report Nr. SHF80 A096068, 91 pp. (www.sintef.no).

Wood, S. N. 2006. Generalized additive models: an introduction with R (Vol. 66). Chapman & Hall.

Yergin, D. 1991. The Prize: The Epic Quest for Oil, Money, and Power. Simon & Schuster, New York.

Ziegler, F. and Hansson, P. A. 2003. Emissions from fuel combustion in Swedish cod fishery. Journal of Cleaner Production, 11: 303-314.

1.10 Figures

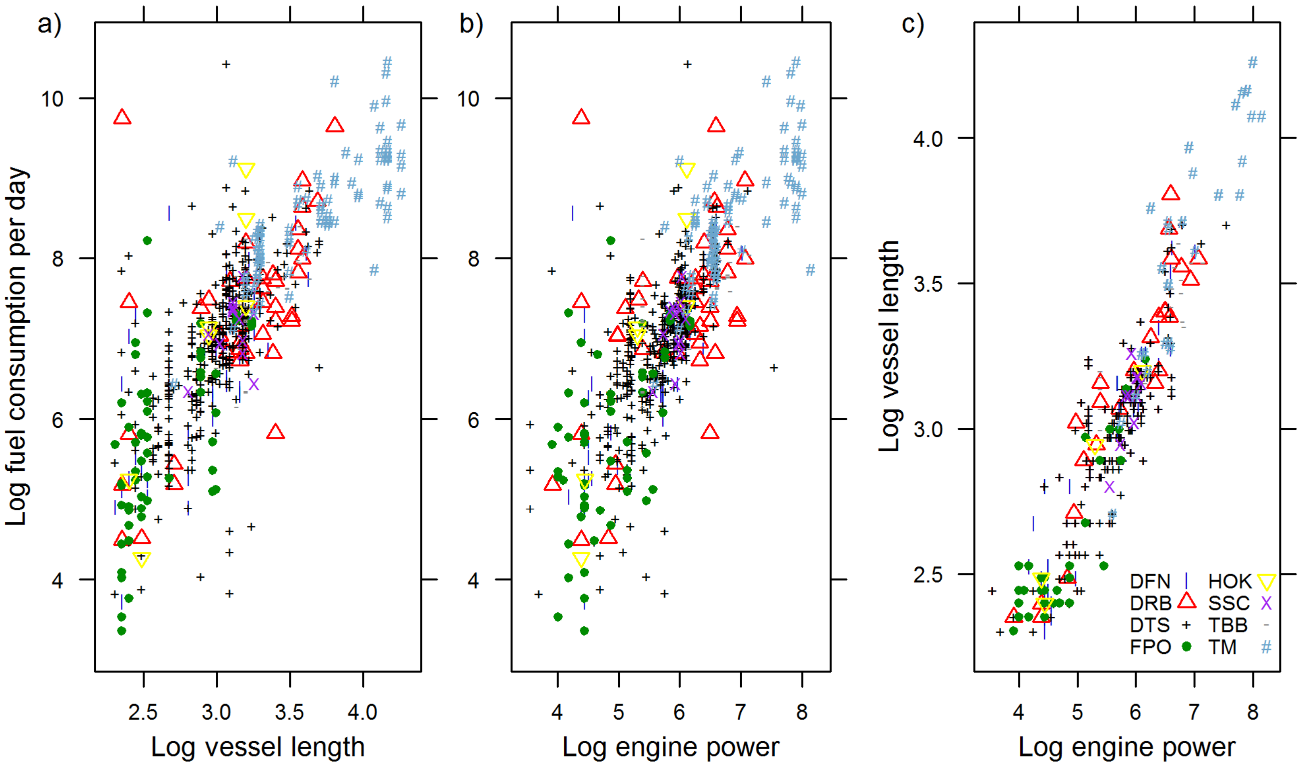

Figure 1. Sampled vessels 2003-2011 fuel per day consumption for different gears by a) vessel length and b) engine power in addition to c) the relationship between vessel length and engine power. Depicted on the natural logarithmic scale.

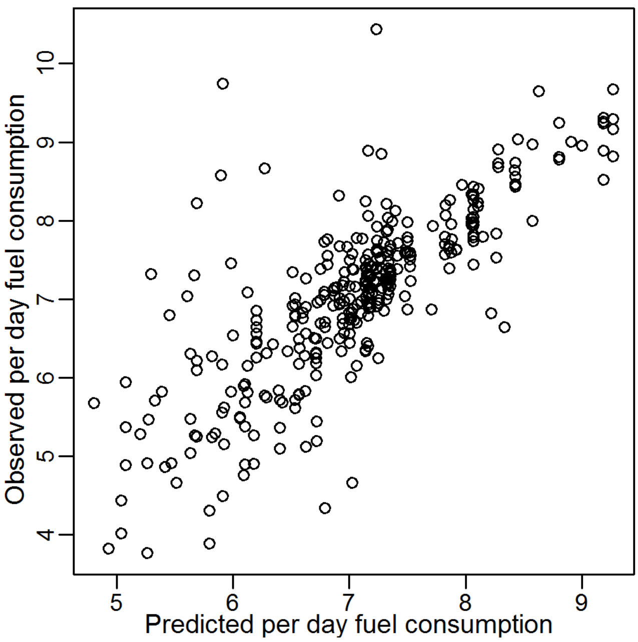

Figure 2. Relationship between observed and predicted values of natural log standardised per day fuel consumption for LM.1.

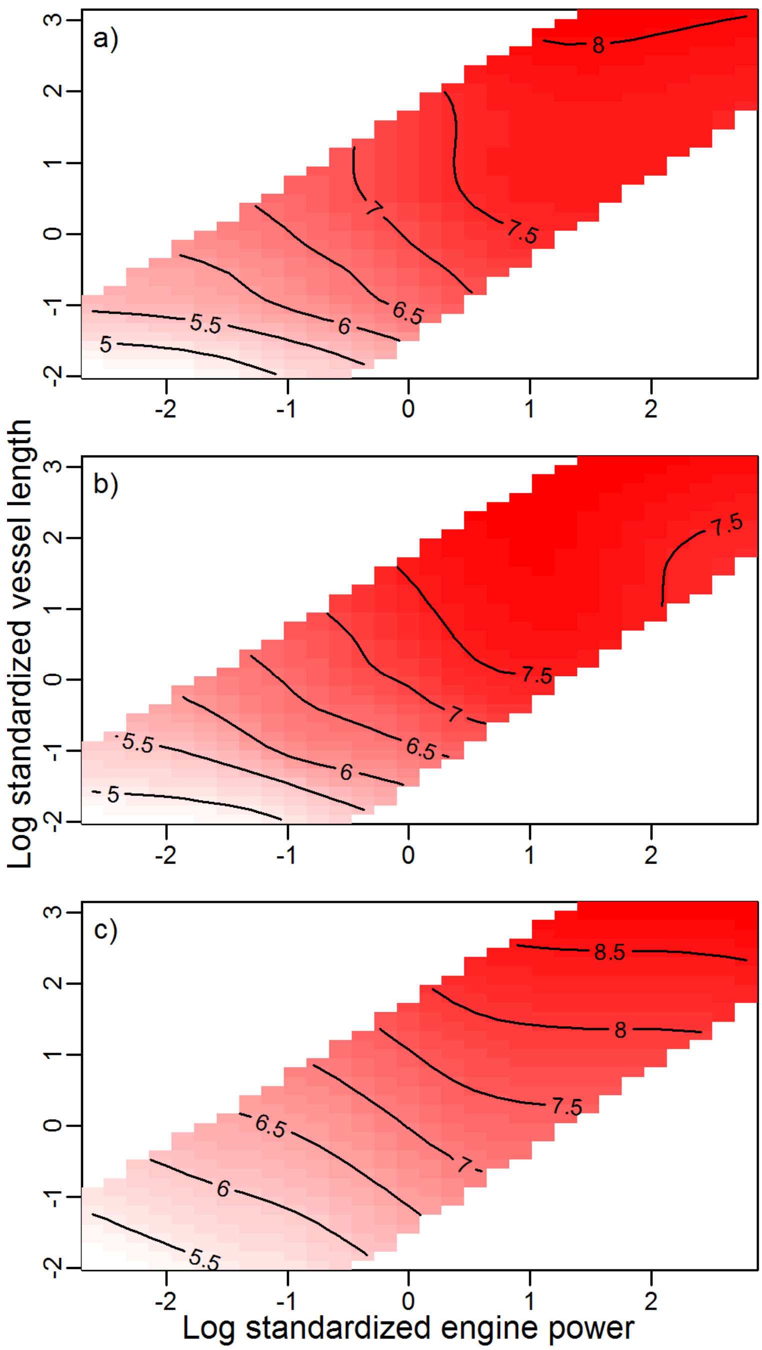

Figure 3. Contour plots representing variation in log (ln) fuel consumption per day between GAM model fits applying DTS gear at a fuel cost of €0.45 per litre across log standardised vessel lengths and engine powers, a) GAM.1, b) GAM.2, c) GAM.3.

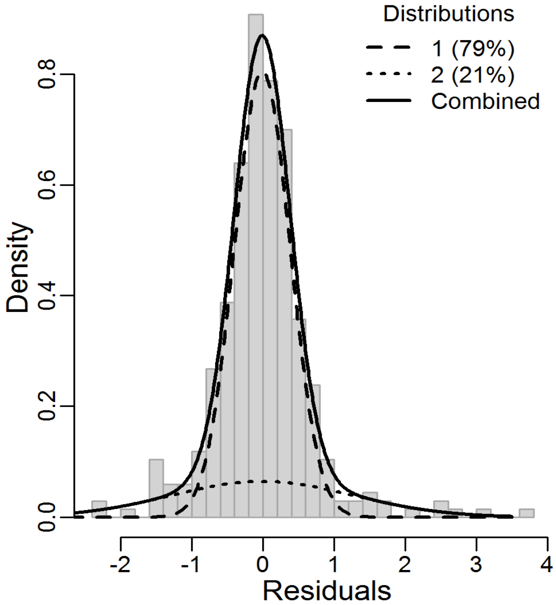

Figure 4. Residuals histogram of GAM.3 fit displaying results of scale mixture model broken down by identified mixed distributions.

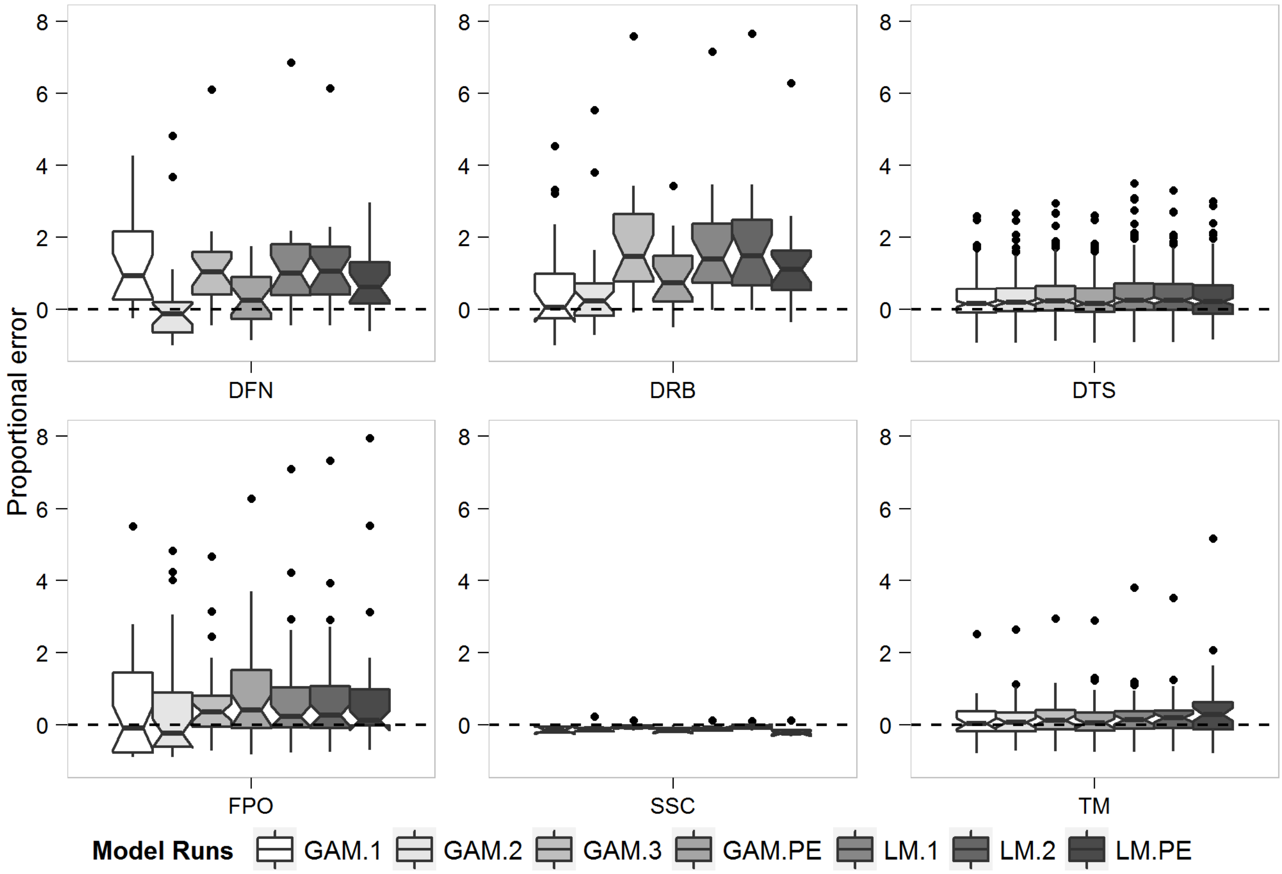

Figure 5. Proportional errors from predicted fuel consumptions of the seven models (Table 1) by gear types. Median depicted with the upper and lower quartiles corresponding to the 25th and 75th percentiles.

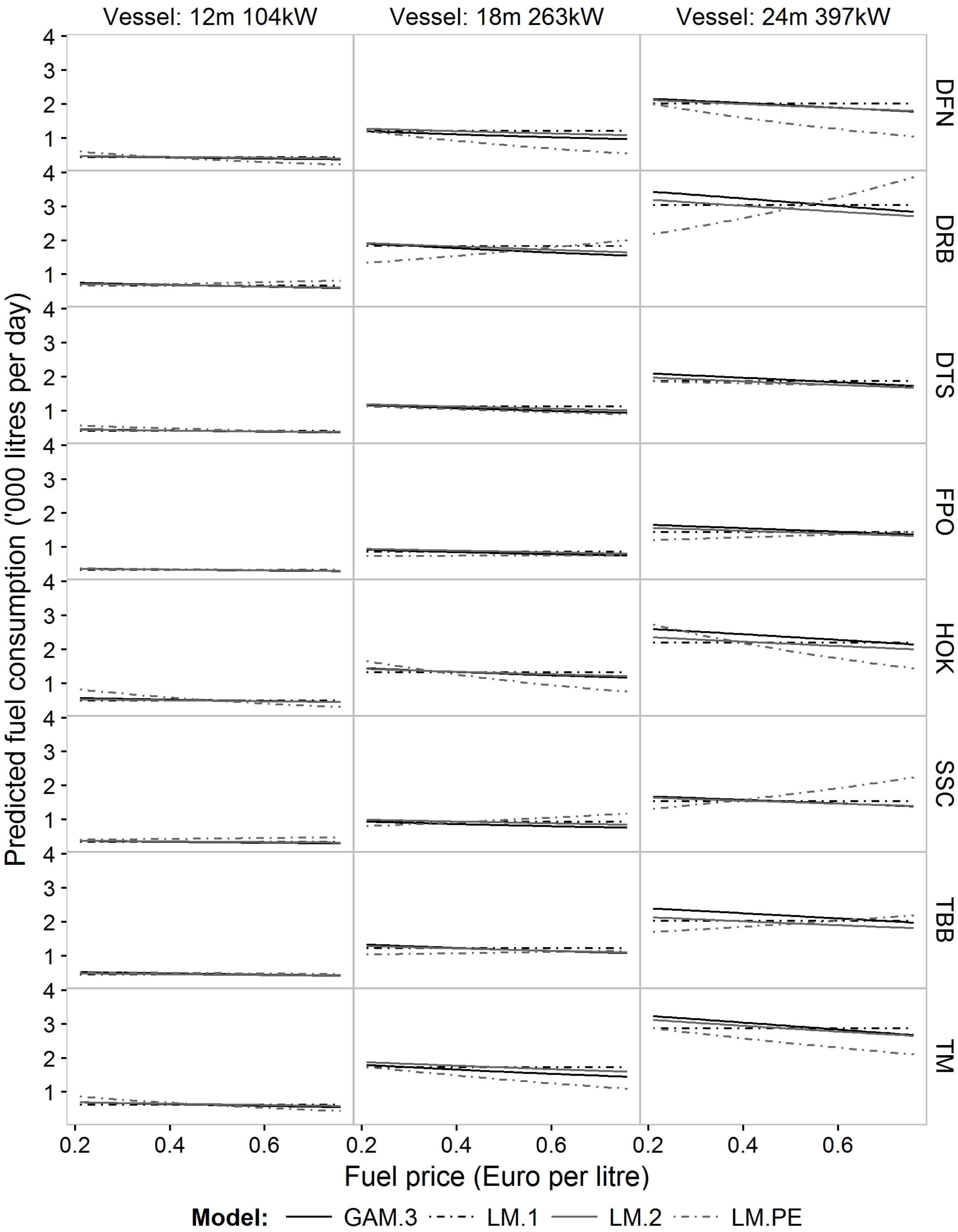

Figure 6. Fuel consumption predictions (litres per day) for models GAM.3, LM.1, LM.2, and LM.PE across the range of observed per litre fuel prices for small (12m), medium (18m), and large (24m) vessel lengths with corresponding average engine power.

1.11 Tables

Table 1. Summary of chosen models, detailing degrees of freedom (df), log likelihoods, AIC values, difference in AIC to best fitting model, and sum of absolute proportional errors from predictions (SumAPE). With *P* as fuel price per litre, *SE* as standardised engine, *SL* as standardised length, and *G* as gear

ID	Model	df	logLik	AIC	ΔAIC	SumAPE
GAM.1	$a_{0,G[i]} + s_1(\ln(SE_i)) + s_{2,G[i]}(\ln(SL_i), P_i) + \varepsilon_i$	62	-254.9	634.4	0	311.8
GAM.2	$a_{0,G[i]} + \beta_1 P_i + s_{1,G[i]}(\ln(SE_i), \ln(SL_i)) + \varepsilon_i$	47	-287.9	670.0	35.6	285.7
GAM.3	$a_{0,G[i]} + s_1(\ln(SL_i)) + s_2(\ln(SE_i), P_i) + \varepsilon_i$	15	-347.4	724.9	90.5	309.1
GAM.PE	$a_{0,G[i]} + s_1(\ln(SL_i)) + s_{2,G[i]}(\ln(SE_i), P_i) + \varepsilon_i$	42	-298.2	679.8	45.4	273.8
LM.1	$a_{0,G[i]} + a_{1,G[i]} \ln(SE_i) + a_{2,G[i]} \ln(SL_i) + a_{3,G[i]} \ln(SE_i) \ln(SL_i) + \varepsilon_i$	12	-353.1	731.1	96.6	313.4
LM.2	$a_{0,G[i]} + a_{1,G[i]} \ln(SE_i) + a_{2,G[i]} \ln(SL_i) + a_{3,G[i]} P_i + a_{4,G[i]} \ln(SE_i) \ln(SL_i) + \varepsilon_i$	13	-352.3	731.8	97.4	310.8
LM.PE	$a_{0,G[i]} + a_{1,G[i]} \ln(SL_i) + a_{2,G[i]} P_i G_i + a_{3,G[i]} \ln(SL_i) P_i + \varepsilon_i$	19	-362.0	764.3	129.9	288.9

Table 2. Coefficients resulting from selected model, GAM.3. The intercept encompasses DFN gear.

Parameter		Estimate	Std. Error	t value	Pr(>|t|)
Intercept		7.005	0.185	37.944	<2e-16
Gear	DRB	0.467	0.247	1.887	0.06
	DTS	-0.031	0.193	-0.161	0.87
	FPO	-0.271	0.232	-1.169	0.24
	HOK	0.184	0.443	0.417	0.68
	SSC	-0.257	0.304	-0.847	0.40
	TBB	0.101	0.256	0.395	0.69
	TM	0.402	0.237	1.701	0.09

1.12 Appendix A

Table A1.Details of gear codes (DCF fleet segments), and the gear types to which they refer. * Demersal seine (SSC) gear under the DCF is included within the demersal trawl (DTS) category. For the purposes of this investigation, demersal seiners were examined as a separate group owing to likely differences in fuel consumption.

Gear code	Gear description
DFN	Drift and/or fixed netters
DRB	Dredgers
DTS	Demersal trawlers
SSC*	Demersal seiners
FPO	Pots and/or traps
HOK	Hooks
PGP	Polyvalent passive gears only
PMP	Mixed active and passive gears
TM	Pelagic trawlers
TBB	Beam trawlers

Table A2. Average annual fuel price per litre 2003-2011 applied in analyses, provided by BIM.

Year	Fuel Price (€/liter)
2003	0.22
2004	0.26
2005	0.38
2006	0.45
2007	0.45
2008	0.56
2009	0.41
2010	0.52
2011	0.75
2012	0.75
2013	0.73

Table A3. Summary of top 20 linear model fits, detailing the model degrees of freedom (df), log-likelihood, AIC value and difference in AIC to the best fitting model. Models ordered by increasing AIC value. With *F* as fuel per day, *P* as fuel price per litre, *SE* as standardised engine, *SL* as standardised length, and *G* as gear. N.B. Degrees of freedom in models containing a three way interaction between vessel length, engine power and gear type are less than expected due to limited number of HOK samples with differing length:power combinations.

Model	df	logLik	AIC	ΔAIC
$a_{0,G[i]} + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	12	-353.1	731.1	0
$a_{0,G[i]} + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	13	-352.3	731.8	0.73
$a_{0,G[i]} + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SL_i) + a_5 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	14	-351.3	731.8	0.73
$a_{0,G[i]} + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SE_i) + a_5 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	14	-351.6	732.4	1.34
$a_{0,G[i]} + a_1 \ln(SE_i) + a_{2,G[i]} \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SL_i) + a_{5,G[i]} \ln(SE_i) \ln(SL_i) + \varepsilon_i$	21	-344.2	733.3	2.23
$a_{0,G[i]} + a_{1,G[i]} \ln(SE_i) + a_{2,G[i]} \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SE_i) + a_{5,G[i]} \ln(SE_i) \ln(SL_i) + \varepsilon_i$	33	-330.1	733.6	2.56
$a_{0,G[i]} + a_{1,G[i]} \ln(SE_i) + a_{2,G[i]} \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SL_i) + a_{5,G[i]} \ln(SE_i) \ln(SL_i) + \varepsilon_i$	33	-330.2	733.8	2.69
$a_{0,G[i]} + a_{1,G[i]} \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SL_i) + \varepsilon_i$	20	-345.6	733.9	2.80
$a_{0,G[i]} + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SE_i) + a_5 P_i \ln(SL_i) + a_6 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	15	-351.2	734.0	2.88
$a_{0,G[i]} + a_{1,G[i]} \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SE_i) + \varepsilon_i$	20	-345.7	734.0	2.88
$a_{0,G[i]} + a_1 \ln(SE_i) + a_{2,G[i]} \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SE_i) + a_6 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	21	-344.6	734.1	3.03
$a_0 + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SL_i) + a_4 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	7	-360.1	734.5	3.46
$a_{0,G[i]} + a_1 \ln(SE_i) + a_2 \ln(SL_i) + \varepsilon_i$	11	-356.0	734.9	3.77
$a_0 + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SL_i) + \varepsilon_i$	6	-361.3	734.9	3.80
$a_{0,G[i]} + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + \varepsilon_i$	12	-355.1	735.1	4.00
$a_0 + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	5	-362.5	735.1	4.03
$a_0 + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SE_i) + a_5 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	7	-360.5	735.3	4.22
$a_{0,G[i]} + a_1 \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SL_i) + \varepsilon_i$	13	-354.1	735.4	4.26
$a_{0,G[i]} + a_{1,G[i]} \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SL_i) + a_5 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	21	-345.2	735.4	4.26
$a_{0,G[i]} + a_{1,G[i]} \ln(SE_i) + a_2 \ln(SL_i) + a_3 P_i + a_4 P_i \ln(SE_i) + a_5 \ln(SE_i) \ln(SL_i) + \varepsilon_i$	21	-345.3	735.5	4.43

Table A4. Summary of top 20 general additive model (GAM) fits, detailing the model degrees of freedom (df), log-likelihood, AIC value and difference in AIC to the best fitting model. Models are ordered by increasing AIC value. With *F* as fuel per day, *P* as fuel price per litre, *SE* as standardised engine, *SL* as standardised length, and *G* as gear.

Model	df	logLik	AIC	ΔAIC
$a_{0,G[i]} + s_1(\ln(SE_i)) + s_{2,G[i]}(\ln(SL_i), P_i) + \varepsilon_i$	62	-254.9	634.4	0
$a_{0,G[i]} + \beta_1 P_i + s_{1,G[i]}(\ln(SE_i), log(SL_i)) + \varepsilon_i$	47	-287.9	670.0	35.58
$a_{0,G[i]} + \beta_{1,G[i]} P_i + s_{1,G[i]}(\ln(SE_i)) + s_{2,G[i]}(\ln(SL_i)) + \varepsilon_i$	55	-282.1	674.9	40.47
$a_{0,G[i]} + s_{1,G[i]}(\ln(SE_i), P_i) + s_2(\ln(SL_i)) + \varepsilon_i$	42	-298.2	679.8	45.37
$a_{0,G[i]} + \beta_1 P_i + s_1(\ln(SE_i)) + s_{2,G[i]}(\ln(SL_i)) + \varepsilon_i$	33	-309.1	685.1	50.65
$a_{0,G[i]} + s_1(\ln(SE_i), P_i) + s_{2,G[i]}(\ln(SL_i)) + \varepsilon_i$	33	-310.9	687.5	53.04
$a_{0,G[i]} + \beta_{1,G[i]} P_i + s_1(\ln(SE_i)) + s_{2,G[i]}(\ln(SL_i)) + \varepsilon_i$	41	-302.8	687.9	53.45
$a_{0,G[i]} + s_1(\ln(SE_i)) + s_{2,G[i]}(\ln(SL_i)) + \varepsilon_i$	32	-311.5	687.9	53.50
$a_{0,G[i]} + \beta_{1,G[i]} P_i + s_1(\ln(SE_i), P_i) + s_{2,G[i]}(\ln(SL_i)) + \varepsilon_i$	42	-302.6	688.5	54.03
$a_{0,G[i]} + s_{1,G[i]}(\ln(SE_i)) + s_2(\ln(SL_i), P_i) + \varepsilon_i$	49	-297.6	693.1	58.69
$a_{0,G[i]} + s_{1,G[i]}(\ln(SE_i), \ln(SL_i) + \varepsilon_i$	36	-311.6	694.8	60.31
$a_{0,G[i]} + \beta_1 P_i + s_{1,G[i]}(\ln(SE_i)) + s_{2,G[i]}(\ln(SL_i)) + \varepsilon_i$	39	-309.0	695.5	61.06
$a_{0,G[i]} + s_{1,G[i]}(\ln(SE_i)) + s_{2,G[i]}(\ln(SL_i)) + \varepsilon_i$	35	-315.3	699.9	65.48
$a_{0,G[i]} + \beta_{1,G[i]} P_i + s_{1,G[i]}(\ln(SE_i)) + s_2(\ln(SL_i), P_i) + \varepsilon_i$	51	-299.9	702.2	67.75
$a_{0,G[i]} + \beta_{1,G[i]} P_i + s_{1,G[i]}(\ln(SE_i)) + s_2(\ln(SL_i)) + \varepsilon_i$	39	-316.1	710.1	75.64
$a_{0,G[i]} + \beta_1 P_i + s_{1,G[i]}(\ln(SE_i)) + s_2(\ln(SL_i)) + \varepsilon_i$	25	-335.4	720.8	86.35
$a_{0,G[i]} + s_{1,G[i]}(\ln(SE_i)) + s_2(\ln(SL_i)) + \varepsilon_i$	23	-339.2	723.7	89.29
$a_{0,G[i]} + s_1(\ln(SE_i), P_i) + s_2(\ln(SL_i)) + \varepsilon_i$	15	-347.4	724.9	90.47
$a_{0,G[i]} + s_1(\ln(SE_i)) + s_2(\ln(SL_i), P_i) + \varepsilon_i$	22	-340.4	725.3	90.81
$a_{0,G[i]} + \beta_1 P_i + s_1(\ln(SE_i)) + s_2(\ln(SL_i)) + \varepsilon_i$	14	-348.3	725.4	90.95

1.13 Supplementary

The following supplementary material is available at ICESJMS online.

The first three figures below are by gear breakdowns of those within Figure 1 of the paper providing greater clarity of detail. Fuel per day consumption is depicted individually by the gears of sampled vessels, 2003-2011, by: vessel length, engine power, and the relationship between vessel length and engine power. Plotted on the natural logarithmic scale. Gear codification is detailed within Table A2 of the appendix.

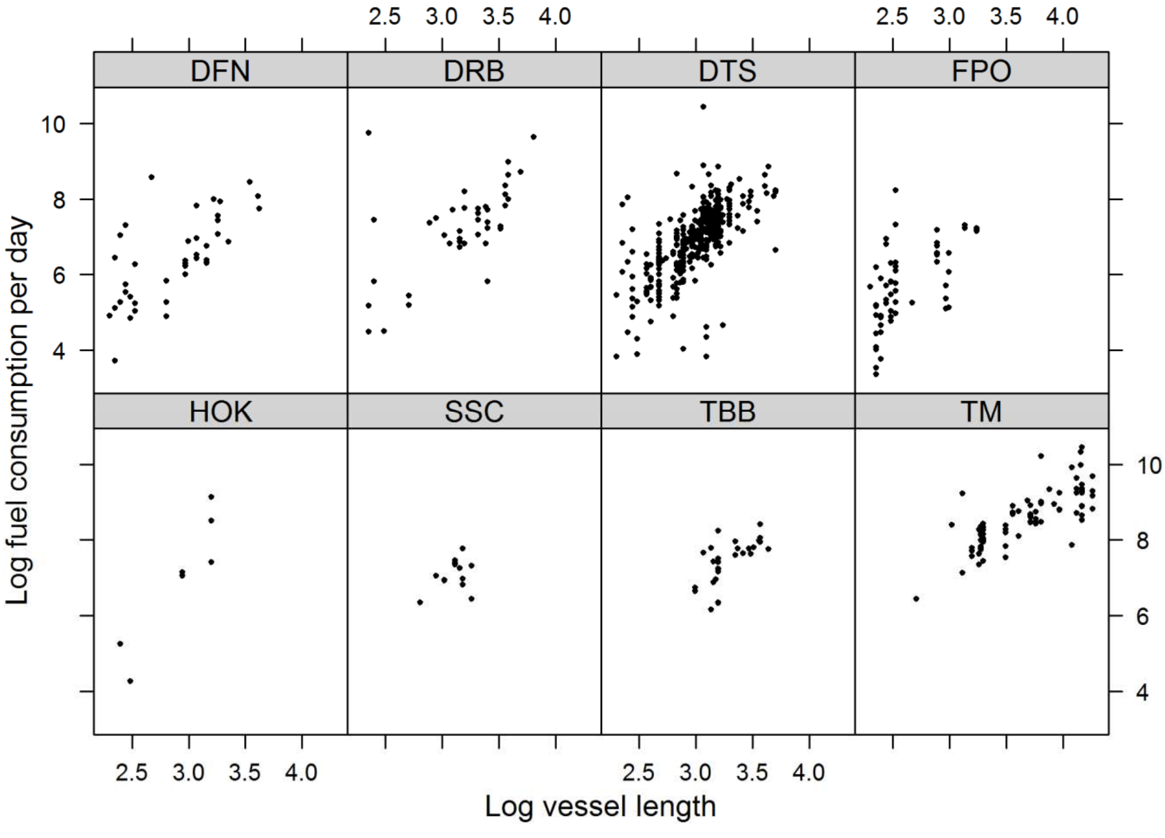

Figure S1. Fuel per day consumption by vessel length.

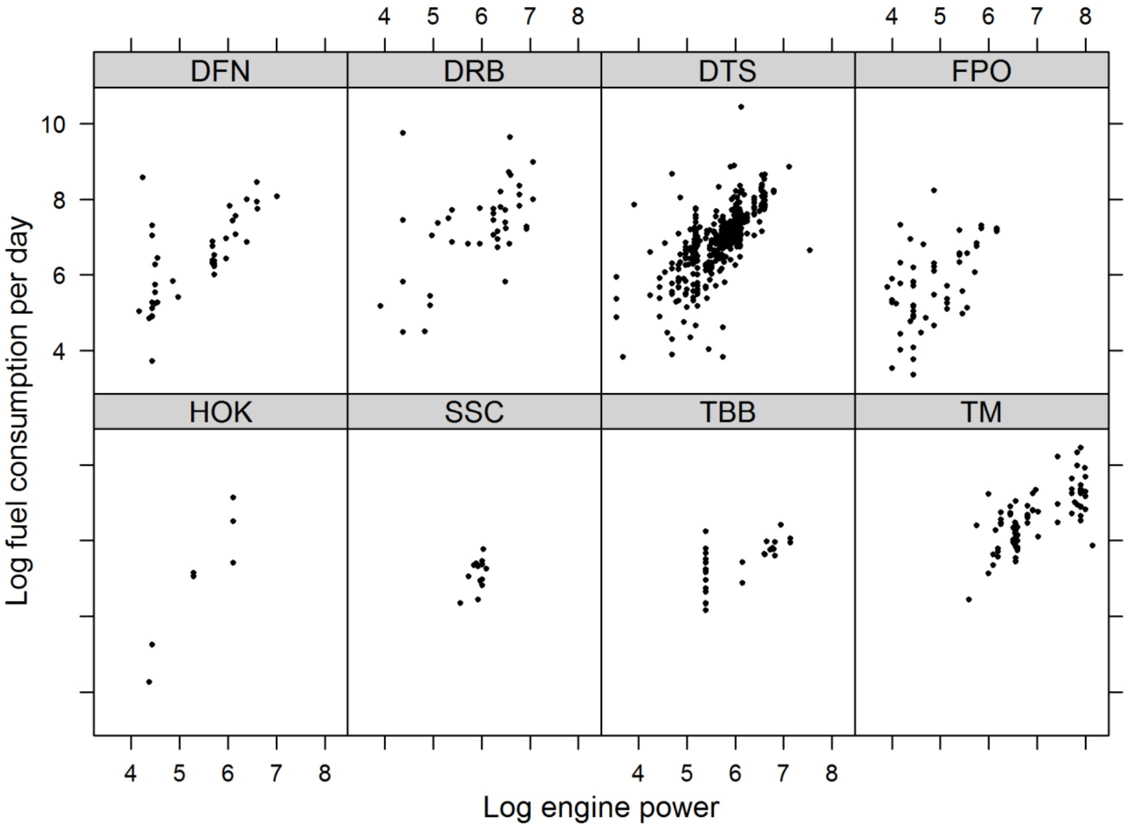

Figure S2. Fuel per day consumption by engine power.

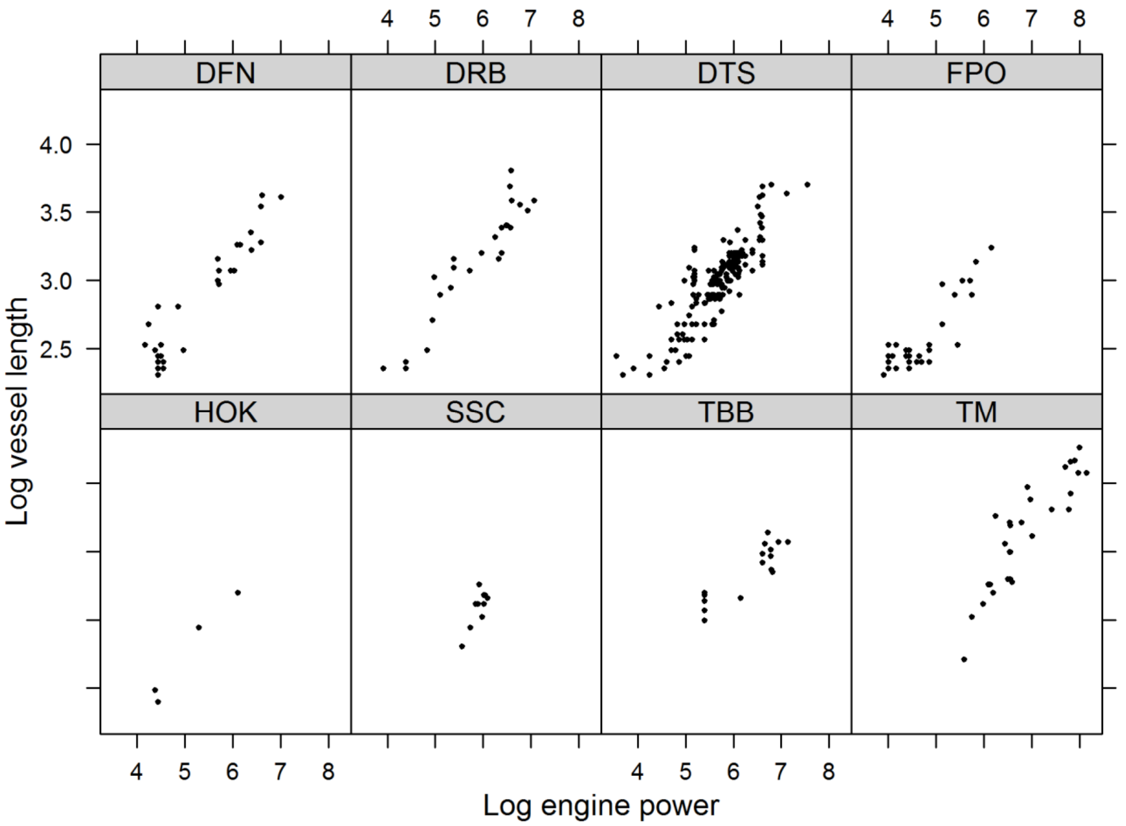

Figure S3. The relationship between sampled vessel length and engine power.

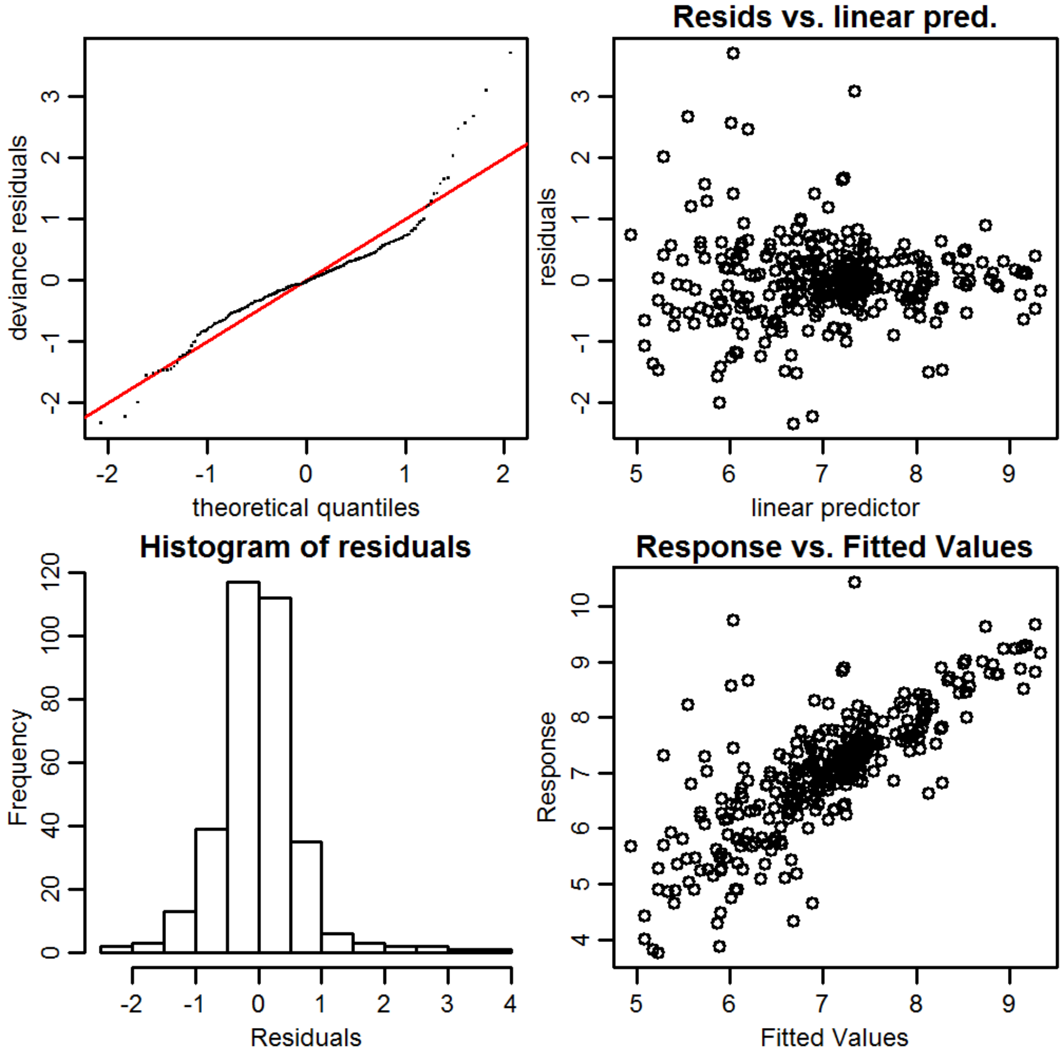

Figure S4. Diagnostic plots from GAM.3. Detailing four plots (top left to bottom right): Q-Q, residuals vs. linear predictor, histogram of residuals, and response vs. fitted values.