ARTICLE DRAFT

Modelling and optimization of ship's fuel consumption using Random Forest Regression (RFR)

Muhammad Fakhruriza Pradana^{a,b}, Hibatul Wafi^b, Bernd Noche^b

^aDepartment of Civil Engineering, University of Sultan Ageng Tirtayasa, Cilegon, Indonesia ; ^bInstitute of Transport Systems and Logistics, University of Duisburg-Essen, Duisburg, Germany

ARTICLE HISTORY

Compiled August 14, 2023

ABSTRACT

Efforts to model energy-efficient operation of shipping operations using machine-learning methods have emerged due to volatile bunker fuel prices and stringent environmental regulations. It is widely regarded that ship speed is one of the most influential factors impacting ships' fuel oil consumption and as such, accurate modelling of ship speed is paramount to ensure the accuracy of subsequent FOC prediction.

This study proposes an intuitive data-driven modelling approach, integrating Automatic Identification System and weather data for modelling of ship states and environmental conditions' impact on FOC. Grey Box Modelling approach divides the speed and FOC prediction into stages, the first stage involves the prediction of speed over ground using Random Forest Regressor. Consequently, the FOC prediction based on predicted speed employs the empirical formula by Holtrop-Mennen, maintaining adherence with established vessel knowledge.

In the presented case study, optimised SOG prediction achieves 3.94% mean absolute percentage error (MAPE) and 93.41% R^2 score. Subsequent FOC prediction from estimated speed yields 86.57% R^2 and 12.06% MAPE. The results affirm the proposed approach's viability in predicting energy-efficient ship operations.

KEYWORDS

Energy-efficient operation; Random Forest Regression; Ship speed prediction; Fuel consumption prediction; Grey Box Model; AIS

1. Introduction

The marine industry is actively researching efficient ship operation due to rising fuel prices and stricter environmental rules. Fuel costs, known as "bunkers," comprise over 50% of voyage expenses and up to 75% of total operating costs, impacting profitability (Bialystocki and Konovessis 2016). Energy-efficient practices reduce costs and greenhouse gas emissions, crucial with shipping contributing 2.51% of global emissions (IMO 2020). This mutual motivation aligns economic benefits with environmental compliance. Stakeholders seek solutions to energy-efficient operation by considering technical and operational approaches. Technical solutions require costly

structural and power system alterations (Yan, Wang, and Psaraftis 2021; Li et al. 2022), prompting interest in the cost-effective, optimisation of operational measures.

Significant emphasis is given in this study on optimisation of ship speed due to its substantial impact on fuel consumption which is caused by a third-order non-linear correlation between fuel consumption and ship speed (Wang and Meng 2012; Du et al. 2019). However, the process of optimising the speed prediction model is intricate, appropriate features must be considered as the ship speed is influenced by factors like vessel performance and weather conditions.

Fuel consumption models based on historical data and ship parameters lack robustness and sensitivity to noise. To address this, recent research employs data-driven techniques, like machine learning (ML), for ship speed and fuel consumption prediction. ML models showcase strong generalisation capabilities and low prediction errors, although some experts are reluctant to accept the generated models by the machine learning approach due to their complexity, unintuitiveness, and potential violation of vessel physics. The success of data-driven models is also highly dependent on data quality and quantity (Yan, Wang, and Psaraftis 2021; Gkerekos, Lazakis, and Theotokatos 2019). Given volatile fuel prices, developing an accurate Fuel Oil Consumption (FOC) prediction model is valuable for maritime stakeholders. This aids in timely economic decisions without violating environmental regulations.

2. Literature Review

2.1. Modelling Approach for Ship Operation

Haranen et al. (2016) and Coraddu et al. (2017) categorised fuel consumption prediction models into three strategies:

White Box Models (WBM): Built on prior mechanistic knowledge and physical principles of a vessel's system, including its structure, design parameters, and propulsion configuration.

Black Box Models (BBM): Data-driven and developed using data from different sailing journeys and historical observations. The Machine Learning (ML) modelling approach focuses on the prediction of bunker consumption at different points in time.

Grey Box Models (GBM): A fusion of WBM and BBM, resulting in a single model that considers both *a priori* knowledge of the vessel and historical sailing data, This method aims to complement the performance of WBM and BBM.

Each strategy has strengths and weaknesses. WBM is transparent and comprehensible, rooted in physics and hydrodynamics, but lacks adaptability and generalisation due to its deterministic nature and dependence on prior knowledge. BBM excels in fitting and predicting data but lacks vessel-specific knowledge and can be complex. To achieve good prediction, it requires an abundance of data quantity and good data quality (Halevy, Norvig, and Pereira 2009). GBM mitigates these limitations by combining mechanistic understanding with predictive capabilities.

The modelling of FOC using GBM requires both components of WBM and BBM. For the BBM modelling part using ML approach, For black-box modelling using ML techniques, it is crucial to have sufficient high-quality data for accurate training (Halevy, Norvig, and Pereira 2009). Yan, Wang, and Psaraftis (2021) categorise the data sources for FOC modelling as follows:

Besides its intended role as a collision avoidance system, Automatic Identification System (AIS) data finds potential in ship behaviour analysis and environmental assessment. The International Maritime Organization (IMO) utilized AIS data to study Greenhouse Gas (GHG) emissions, estimating global shipping emissions (IMO 2020; Smith et al. 2015). Rakke (2016) introduced ECAIS as a methodology to compute ship emissions from AIS-derived fuel consumption data using Holtrop-Mennen and literature-based approximations. The study by Kim et al. (2020) used AIS data, ship information, and environmental data for estimating Energy Efficiency Operational Indicator (EEOI). The use of AIS data in research aims for data independence, reducing reliance on commercial databases.

2.2. Predictive performance of tree based models

Tree-based model is a supervised, highly interpretable BBM modelling approach using machine learning approach which is adept in classification and regression tasks. The model is inherently resistant to multicollinearity problems (Yan, Wang, and Psaraftis 2021). Several literature studies reveal its advantages and performance superiority. Soner, Akyuz, and Celik (2018) employed ferry data to predict FOC using tree-based models including bagging, random forest (RF), and bootstrap. RF achieved 43.5 L/h RMSE for fuel consumption, outperforming Artificial Neural Network (ANN) model employed by Petersen, Jacobsen, and Winther (2012).

Yan, Wang, and Du (2020) predicted FOC for a dry bulk ship's voyage using RF. The model incorporated sailing speed, cargo weight, and meteorological conditions, it is able to attain mean absolute percentage error (MAPE) of 7.91% and the RF model outperformed decision tree, ANN, LASSO, and SVR. Gkerekos, Lazakis, and Theotokatos (2019) compared ML models to predict daily FOC, RF model achieved 89% and 96% R^2 scores with noon data and ADLM system data, respectively. Li et al. (2022) fused meteorological, voyage, and AIS data to explore the effect of data on ML models for FOC prediction. Tree-based models (bagging and boosting ensembles) including ETR, RFR, AB, GB, XG, and LB were recommended for energy-efficient operation modelling, with RFR particularly displaying the best robustness among the presented ML models in the study. Abebe et al. (2020) predicted ship speed over ground (SOG) using AIS and weather data. The RF model achieved 98% R^2 score and 0.25 knots RMSE.

WBMs for predicting FOC utilize physics and hydrodynamic laws to compute the vessel's resistance, encompassing calm water resistance and additional effects like wind and waves. Then the engine power can be subsequently estimated at a specific speed, facilitating FOC calculation (Haranen et al. 2016). Holtrop-Mennen power estimation method Holtrop (1984), is applicable in a wide range (Rakke 2016; Kim

et al. 2020). Rakke (2016) utilized AIS data and mechanical information to estimate engine power using Holtrop-Mennen, achieving about 5% model testing error for FOC and GHG emissions estimation. Similarly, Kim et al. (2020) estimated Energy Efficiency Operational Index (EEOI) through Holtrop-Mennen-based engine power estimation, enabled by AIS data and weather information.

3. Methodology

This chapter covers the methodology used to construct the grey box model (GBM). The grey box approach employed in this study is categorised as sequential GBM, which entails a two-stage development process. The initial stage focuses on machine learning modelling using tree-based models. The modelling is carried out using Python in conjunction with Scikit-Learn (Pedregosa et al. 2011).

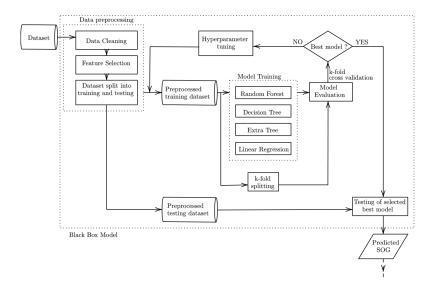


Figure 1. Scheme of proposed BBM methodology

The second stage of the modelling process revolves around the power estimation method (Holtrop 1984). This involves an initial conversion of SOG to STW for estimation of encountered resistance during the voyage, this then facilitates the estimation of the required power i.e. the energy required to propel the ship.

3.1. Data Acquisition

The data is collected from a ferry serving between the ports of Køge, Rønne, Ystad, and Sassnitz. The trip duration between Køge and Rønne is approximately 5 hours and 30 minutes, while the voyage between Rønne and Sassnitz takes around 3 hours and 20 minutes. The Danish Maritime Authority's (DMA) T-AIS system tracks the journey. Weather data along the ferry's route is sourced from ECMWF, providing information on wind, waves, and seawater temperature. This data has a temporal resolution of 1 hour and a spatial granularity of 0.25° (longitude) x 0.25° (latitude).

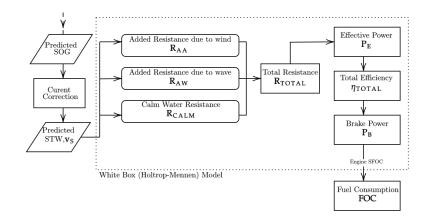


Figure 2. Scheme of proposed BBM methodology

Current information, obtained from CMEMS, is available at a temporal resolution of 3 hours and a spatial granularity of 0.25° (longitude) x 0.25° (latitude).

IMO	9812107
Type & Service	Passenger ferry
L_{OA}	158.00 m
L_{WL}	144.80 m
B (moulded)	24.5 m
T_{DESIGN}	5.70 m
T_{MAX}	5.85 m
Gross Tonnage (GT)	18,009
Deadweight (dwt)	4,830 t
Main Engines	Wärtsillä 8V31 2 x 4,880 kW
SFOC	169.4 g/kWh
Service Speed	17.7 knots
Bow Thrusters	$2 \ge 1500 \text{ kW}$

Køge Ystod
Rønne
Sassnitz

Figure 3. Particular of M/S Hammershus

Figure 4. Journey of the ferry

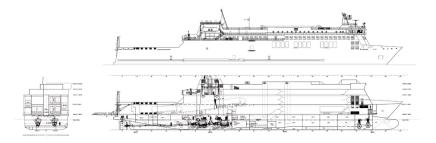


Figure 5. Schematics of M/S Hammershus

3.2. Modelling methodologies

3.2.1. Decision Tree (DT) Regressor

The Decision Tree operates by employing nested if-then statements based on predefined rules, resulting in a partitioned data space. This process can also be visualized as a binary tree, enhancing interpretability by representing diverse input

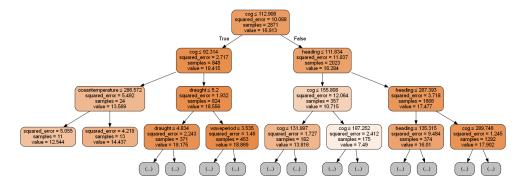


Figure 6. Structure of a Decision Tree (DT) Regressor

responses within a single tree Kuhn and Johnson (2013); Hastie, Tibshirani, and Friedman (2009).

A Decision Tree encompasses distinct nodes: the **Root node** represents the top-level node; **Leaf nodes** (or terminal nodes) yield final prediction outcomes; and **Internal nodes** lie between the root and leaf nodes. The procedure of dividing a node into subsequent nodes is termed **splitting**, where the original node is the **parent node** and the resultant nodes are **child nodes**. In regression tasks, tree growth is often controlled by Mean Square Error (MSE), guided by the Classification and Regression Tree (CART) algorithm, The notation $J(k, t_k)$ represents the cost function that needs to be minimised.

$$MSE_{s_i} = \frac{1}{n_{s_i}} SSE_{s_i}$$
 where $i = (1, 2)$ (1)

$$J(\mathbf{k}, \mathbf{t}_k) = \frac{1}{n_{s_1}} SSE_{s_1} + \frac{1}{n_{s_2}} SSE_{s_2} \begin{cases} SSE_{s_i} = \sum_{i \in s_i} (\hat{y}_{s_i} - y_{s_i})^2 \\ \hat{y}_{s_i} = \frac{1}{n_{s_i}} \sum_{i \in s_i} y \end{cases}$$
(2)

The process of tree growth stops until either the number of samples for splitting reaches a predefined threshold or when no further split can be found which reduces the MSE. The decisions from optimal splits are visualised through a binary tree representation, enhancing the interpretability and ease of implementation. The inherent logic structure of decision trees enables them to handle diverse data types without extensive preprocessing, including sparse, skewed, continuous, and categorical data. Decision trees also inherently perform feature selection, which is a valuable aspect in modelling (Kuhn and Johnson 2013).

However, an unconstrained single decision tree is prone to overfitting due to its tendency to closely match the training data. This model's instability can lead to substantial changes in its structure when the data is altered, resulting in a completely different interpretation of splits (Hastie, Tibshirani, and Friedman 2009; Kuhn and Johnson 2013). To mitigate overfitting, it becomes essential to regularise the decision tree's growth during training. The following parameters control the growth of a single decision tree:

- max_depth: This hyperparameter is defined as the count of nodes along a path from the root node to its parent node. The default parameter allows full unpruned growth of the tree.
- min_samples_leaf: This hyperparameter controls the number of samples required to be at the leaf node, where the split point will be considered if the leaf contains at least min_samples_leaf=n training samples in each left and right branch.
- min_samples_split: This hyperparameter controls the minimum number of samples i.e. data points required to split a node.

3.2.2. Random Forest (RF) Regressor

Ensemble learning offers a solution to enhance the performance of Decision Tree (DT) regressors. This concept involves combining the strengths of multiple simpler base models (Hastie, Tibshirani, and Friedman 2009). One prominent ensemble method is the **Random Forest**, introduced by Breiman (2001), which involves creating bootstrap samples, randomly selecting splitting features, and aggregating predictions. This approach combines various learning algorithms, referred to as weak learners, with each corresponding to an individual decision tree in the Random Forest. Random forest uses the Bagging (bootstrap aggregating) strategy, where it trains each tree using bootstrap samples, where instances from the training set are randomly selected with replacement.

To further improve bagging, reducing the correlation between trees is applied. This involves introducing randomness during tree construction. Random split selection, as introduced by Dietterich (2000), involves selecting a feature from a random subset for each split. This, coupled with the inherent instability of a single decision tree, addresses overfitting and the lack of robustness of DTR. The Random Forest methodology addresses these issues by creating an ensemble of independent, strong learners, resulting in reduced variance and robustness against noisy data (Breiman 2001). While losing some interpretability compared to basic tree-based models, the impact of each feature in the ensemble can still be quantified (Kuhn and Johnson 2013). Random Forest performs better with larger sample sizes, and extensive parameter tuning is often unnecessary for good prediction results (Kuhn and Johnson 2013; Hastie, Tibshirani, and Friedman 2009).

In addition to the hyperparameters used to fine-tune the decision tree, the RF model provides additional hyperparameters to control the growth of the tree:

- max_features: This hyperparameter controls the number of features to be considered when looking for the best split. The default parameter considers all features during training.
- n_estimators: This hyperparameter controls the number of trees i.e. predictors in a forest.

4. Using the interact class file

For convenience, simply copy the interact.cls file into the same directory as your manuscript files (you do not need to install it in your TEX distribution). In order to

use the interact document class, replace the command \documentclass{article} at the beginning of your document with the command \documentclass{interact}.

The following document-class options should *not* be used with the interact class file:

- 10pt, 11pt, 12pt unavailable;
- oneside, twoside not necessary, oneside is the default;
- leqno, titlepage should not be used;
- twocolumn should not be used (see Subsection ??);
- onecolumn not necessary as it is the default style.

To prepare a manuscript for a journal that is printed in A4 (two column) format, use the largeformat document-class option provided by interact.cls; otherwise the class file produces pages sized for B5 (single column) format by default. The geometry package should not be used to make any further adjustments to the page dimensions.

5. Additional features of the interact class file

5.1. Title, authors' names and affiliations, abstracts and article types

The title should be generated at the beginning of your article using the \maketitle command. In the final version the author name(s) and affiliation(s) must be followed immediately by \maketitle as shown below in order for them to be displayed in your PDF document. To prepare an anonymous version for double-blind peer review, you can put the \maketitle between the \title and the \author in order to hide the author name(s) and affiliation(s) temporarily. Next you should include the abstract if your article has one, enclosed within an abstract environment. The \articletype command is also provided as an optional element which should only be included if your article actually needs it. For example, the titles for this document begin as follows:

\articletype{ARTICLE TEMPLATE}

\title{Taylor \& Francis \LaTeX\ template for authors (\textsf{Interact})
layout + Chicago author-date reference style)}

\author{

\name{A.~N. Authora\thanks{CONTACT A.~N. Author.
Email: latex.helpdesk@tandf.co.uk} and John Smithb}
\affil{aTaylor \& Francis, 4 Park Square, Milton
Park, Abingdon, UK; bInstitut f\"{u}r Informatik,
Albert-Ludwigs-Universit\"{a}t, Freiburg, Germany} }

\maketitle

\begin{abstract}

This template is for authors who are preparing a manuscript for a Taylor \& Francis journal using the \LaTeX\ document preparation system and the \texttt{interact} class file, which is available via selected journals' home pages on the Taylor \& Francis website. \end{abstract}

An additional abstract in another language (preceded by a translation of the article title) may be included within the abstract environment if required.

A graphical abstract may also be included if required. Within the abstract environment you can include the code

\\\resizebox{25pc}{!}{\includegraphics{abstract.eps}}

where the graphical abstract is to appear, where abstract.eps is the name of the file containing the graphic (note that 25pc is the recommended maximum width, expressed in pica, for the graphical abstract in your manuscript).

5.2. Abbreviations

A list of abbreviations may be included if required, enclosed within an abbreviations environment, i.e. \begin{abbreviations}...\end{abbreviations}, immediately following the abstract environment.

5.3. Keywords

A list of keywords may be included if required, enclosed within a keywords environment, i.e. \begin{keywords}...\end{keywords}. Additional keywords in other languages (preceded by a translation of the word 'keywords') may also be included within the keywords environment if required.

5.4. Subject classification codes

AMS, JEL or PACS classification codes may be included if required. The interact class file provides an amscode environment, i.e. \begin{amscode}...\end{amscode}, a jelcode environment, i.e. \begin{jelcode}...\end{jelcode}, and a pacscode environment, i.e. \begin{pacscode}...\end{pacscode} to assist with this.

5.5. Additional footnotes to the title or authors' names

The \thanks command may be used to create additional footnotes to the title or authors' names if required. Footnote symbols for this purpose should be used in the order * (coded as * ast\$), † (\dagger), ‡ (\dagger), ‡ (\dagger), § (\ship), ¶ (\hip), ¶ (\hip).

Note that any **footnotes** to the main text will automatically be assigned the superscript symbols 1, 2, 3, etc. by the class file.¹

6. Some guidelines for using the standard features of LATEX

6.1. Sections

The Interact layout style allows for five levels of section heading, all of which are provided in the interact class file using the standard LATEX commands \section,

¹If preferred, the **endnotes** package may be used to set the notes at the end of your text, before the bibliography. The symbols will be changed to match the style of the journal if necessary by the typesetter.

\subsection, \subsubsection, \paragraph and \subparagraph. Numbering will be automatically generated for all these headings by default.

6.2. Lists

Numbered lists are produced using the enumerate environment, which will number each list item with arabic numerals by default. For example,

- (1) first item
- (2) second item
- (3) third item

was produced by

```
\begin{enumerate}
  \item first item
  \item second item
  \item third item
\end{enumerate}
```

Alternative numbering styles can be achieved by inserting an optional argument in square brackets to each item, e.g. \item[(i)] first item to create a list numbered with roman numerals at level one.

Bulleted lists are produced using the itemize environment. For example,

- First bulleted item
- Second bulleted item
- Third bulleted item

was produced by

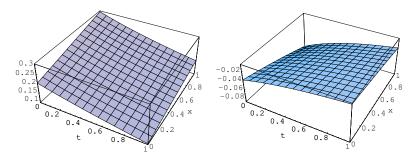
```
\begin{itemize}
  \item First bulleted item
  \item Second bulleted item
  \item Third bulleted item
\end{itemize}
```

6.3. Figures

The interact class file will deal with positioning your figures in the same way as standard LATEX. It should not normally be necessary to use the optional [htb] location specifiers of the figure environment in your manuscript; you may, however, find the [p] placement option or the endfloat package useful if a journal insists on the need to separate figures from the text.

Figure captions appear below the figures themselves, therefore the \caption command should appear after the body of the figure. For example, Figure 7 with caption and sub-captions is produced using the following commands:

```
\begin{figure}
\centering
\subfigure[An example of an individual figure sub-caption.]{%
\resizebox*{5cm}{!}{\includegraphics{graph1.eps}}}\hspace{5pt}
\subfigure[A slightly shorter sub-caption.]{%
\resizebox*{5cm}{!}{\includegraphics{graph2.eps}}}
```



- (a) An example of an individual figure sub-caption.
- (b) A slightly shorter sub-caption.

Figure 7. Example of a two-part figure with individual sub-captions showing that captions are flush left and justified if greater than one line of text.

\caption{Example of a two-part figure with individual sub-captions
showing that captions are flush left and justified if greater
than one line of text.} \label{sample-figure}
\end{figure}

To ensure that figures are correctly numbered automatically, the \label command should be included just after the \caption command, or in its argument.

The \subfigure command requires subfigure.sty, which is called in the preamble of the interacttfssample.tex file (to allow your choice of an alternative package if preferred) and included in the Interact IATEX bundle for convenience. Please supply any additional figure macros used with your article in the preamble of your .tex file.

The source files of any figures will be required when the final, revised version of a manuscript is submitted. Authors should ensure that these are suitable (in terms of lettering size, etc.) for the reductions they envisage.

The epstopdf package can be used to incorporate encapsulated PostScript (.eps) illustrations when using PDFIATEX, etc. Please provide the original .eps source files rather than the generated PDF images of those illustrations for production purposes.

6.4. Tables

The interact class file will deal with positioning your tables in the same way as standard LATEX. It should not normally be necessary to use the optional [htb] location specifiers of the table environment in your manuscript; you may, however, find the [p] placement option or the endfloat package useful if a journal insists on the need to separate tables from the text.

The tabular environment can be used as shown to create tables with single horizontal rules at the head, foot and elsewhere as appropriate. The captions appear above the tables in the Interact style, therefore the \tbl command should be used before the body of the table. For example, Table 1 is produced using the following commands:

```
\begin{table}
\tbl{Example of a table showing that its caption is as wide as
  the table itself and justified.}
{\begin{tabular}{lccccc} \toprule
  & \multicolumn{2}{1}{Type} \\ \cmidrule{2-7}
```

Class & One & Two & Three & Four & Five & Six \\ \midrule

Table 1. Example of a table showing that its caption is as wide as the table itself and justified.

	Туре							
Class	One	Two	Three	Four	Five	Six		
Alpha ^a	A1	A2	A3	A4	A5	A6		
Beta	B2	B2	B3	B4	B5	$_{\rm B6}$		
Gamma	C2	C2	C3	C4	C5	C6		

^aThis footnote shows how to include footnotes to a table if required.

```
Alpha\textsuperscript{a} & A1 & A2 & A3 & A4 & A5 & A6 \\
Beta & B2 & B2 & B3 & B4 & B5 & B6 \\
Gamma & C2 & C2 & C3 & C4 & C5 & C6 \\ bottomrule
\end{tabular}}
\tabnote{\textsuperscript{a}This footnote shows how to include footnotes to a table if required.}
\label{sample-table}
\end{table}
```

To ensure that tables are correctly numbered automatically, the \label command should be included just before \end{table}.

The \toprule, \midrule, \bottomrule and \cmidrule commands are those used by booktabs.sty, which is called by the interact class file and included in the Interact LATEX bundle for convenience. Tables produced using the standard commands of the tabular environment are also compatible with the interact class file.

6.5. Landscape pages

If a figure or table is too wide to fit the page it will need to be rotated, along with its caption, through 90° anticlockwise. Landscape figures and tables can be produced using the rotating package, which is called by the interact class file. The following commands (for example) can be used to produce such pages.

```
\setcounter{figure}{1}
\begin{sidewaysfigure}
\centerline{\epsfbox{figname.eps}}
\caption{Example landscape figure caption.}
\label{landfig}
\end{sidewaysfigure}
\setcounter{table}{1}
\begin{sidewaystable}
\tbl{Example landscape table caption.}
{\begin{tabular}{@{}llllcll}
.
.
.
.
\end{tabular}\label{landtab}
\end{sidewaystable}
```

Before any such float environment, use the \setcounter command as above to fix

the numbering of the caption (the value of the counter being the number given to the preceding figure or table). Subsequent captions will then be automatically renumbered accordingly. The \epsfbox command requires epsfig.sty, which is called by the interact class file and is also included in the Interact LATEX bundle for convenience.

Please note that if the endfloat package is used, one or both of the commands

\DeclareDelayedFloatFlavor{sidewaysfigure}{figure}
\DeclareDelayedFloatFlavor{sidewaystable}{table}

will need to be included in the preamble of your .tex file, after the endfloat package is loaded, in order to process any landscape figures and/or tables correctly.

6.6. Theorem-like structures

A predefined proof environment is provided by the amsthm package (which is called by the interact class file), as follows:

Proof. More recent algorithms for solving the semidefinite programming relaxation are particularly efficient, because they explore the structure of the MAX-CUT problem.

This was produced by simply typing:

\begin{proof}

More recent algorithms for solving the semidefinite programming relaxation are particularly efficient, because they explore the structure of the MAX-CUT problem.

\end{proof}

Other theorem-like environments (theorem, definition, remark, etc.) need to be defined as required, e.g. using \newtheorem{theorem}{Theorem} in the preamble of your .tex file (see the preamble of interactcadsample.tex for more examples). You can define the numbering scheme for these structures however suits your article best. Please note that the format of the text in these environments may be changed if necessary to match the style of individual journals by the typesetter during preparation of the proofs.

6.7. Mathematics

6.7.1. Displayed mathematics

The interact class file will set displayed mathematical formulas centred on the page without equation numbers if you use the displaymath environment or the equivalent \[...\] construction. For example, the equation

$$\hat{\theta}_{w_i} = \hat{\theta}(s(t, \mathcal{U}_{w_i}))$$

was typeset using the commands

```
\[
\hat{\theta}_{w_i} = \hat{\theta}(s(t,\mathbb{U}_{w_i}))
\]
```

For those of your equations that you wish to be automatically numbered sequentially throughout the text for future reference, use the equation environment, e.g.

$$\hat{\theta}_{w_i} = \hat{\theta}(s(t, \mathcal{U}_{w_i})) \tag{3}$$

was typeset using the commands

```
\begin{equation}
\hat{\theta}_{w_i} = \hat{\theta}(s(t,\mathbb{U}_{w_i}))
\end{equation}
```

Part numbers for sets of equations may be generated using the **subequations** environment, e.g.

$$\varepsilon \rho w_{tt}(s,t) = N[w_s(s,t), w_{st}(s,t)]_s, \tag{4a}$$

$$w_{tt}(1,t) + N[w_s(1,t), w_{st}(1,t)] = 0, (4b)$$

which was typeset using the commands

This is made possible by the amsmath package, which is called by the class file. If you put a \label just after the \begin{subequations} command, references can be made to the collection of equations, i.e. '(4)' in the example above. Or, as the example also shows, you can label and refer to each equation individually – i.e. '(4a)' and '(4b)'.

Displayed mathematics should be given end-of-line punctuation appropriate to the running text sentence of which it forms a part, if required.

6.7.2. Math fonts

6.7.2.1. Superscripts and subscripts. Superscripts and subscripts will automatically come out in the correct size in a math environment (i.e. enclosed within $\(\ldots)$) or ldots commands in running text, or within $\[\ldots]$ or the equation environment for displayed equations). Sub/superscripts that are physical variables should be italic, whereas those that are labels should be roman (e.g. C_p , T_{eff}). If the subscripts or superscripts need to be other than italic, they must be coded individually.

6.7.2.2. Upright Greek characters and the upright partial derivative sign. Upright lowercase Greek characters can be obtained by inserting the letter 'u' in the control code for the character, e.g. \umu and \upi produce μ (used, for example, in the symbol for the unit microns $-\mu m$) and π (the ratio of the circumference of a circle to its diameter). Similarly, the control code for the upright partial derivative ϑ

is \upartial. Bold lowercase as well as uppercase Greek characters can be obtained by {\bm \gamma}, for example, which gives γ , and {\bm \Gamma}, which gives Γ .

Acknowledgement(s)

An unnumbered section, e.g. \section*{Acknowledgements}, may be used for thanks, etc. if required and included in the non-anonymous version before any Notes or References.

Disclosure statement

An unnumbered section, e.g. \section*{Disclosure statement}, may be used to declare any potential conflict of interest and included in the non-anonymous version before any Notes or References, after any Acknowledgements and before any Funding information.

Funding

An unnumbered section, e.g. \section*{Funding}, may be used for grant details, etc. if required and included in the non-anonymous version before any Notes or References.

Notes on contributor(s)

An unnumbered section, e.g. \section*{Notes on contributors}, may be included in the non-anonymous version if required. A photograph may be added if requested.

Nomenclature/Notation

An unnumbered section, e.g. \section*{Nomenclature} (or \section*{Notation}), may be included if required, before any Notes or References.

Notes

An unnumbered 'Notes' section may be included before the References (if using the endnotes package, use the command \theendnotes where the notes are to appear, instead of creating a \section*).

7. References

7.1. References cited in the text

References should be cited in Chicago author-date style, e.g. '(Albiston 2005; Greenberg 2008; Schuman and Scott 1987)' or '... see Smith (1985, 75)'. If there are three

authors, list them all in every citation, e.g. '(Jacobs, Thomas, and Lang 1997)'. For more than three authors, cite the first author's name followed by et al. For two or more sources by the same author(s) in the same year, use lower-case letters (a, b, c, ...) with the year to order the entries in the References list and use these letters with the year in the in-text citations, e.g. '(Fogel 2004a,b)'. If two or more authors have the same surname, use their initials with the surnames, e.g. '(C. Doershuk 2010; J. Doershuk 2009)'. If the first author's names and the years of publication are identical for several references, include enough co-author names to eliminate ambiguity, e.g. '(Schonen, Baker, et al. 2009; Schonen, Brooks, et al. 2009)'. For further details on this reference style, see the Instructions for Authors on the Taylor & Francis website.

Each bibliographic entry has a key, which is assigned by the author and is used to refer to that entry in the text. In this document, the key Fow89 in the citation form \citep{Fow89} produces '(Fowler 1989)', and the keys {Bro86,Bro02,Roh08} in the citation form \citep{Bro86,Bro02,Roh08} produce '(Brooks and Wiley 1986; Brooks and McLennan 2002; Rohde, Levy, and Kehler 2008)'. The appropriate citation style for different situations can be obtained, for example, by \citet{Sam06} for 'Samples (2006)', \citealt{Lev05} for 'Levitt and Dubner 2005', or \citealp{Mor08} for 'Morasse, Guderley, and Dodson 2008'. Citation of the year alone may be produced by \citeyear{Cho08}, i.e. '2008', or \citeyearpar{ChoGul08}, i.e. '(2008)', or of the author(s) alone by \citeauthor{Tep05}, i.e. 'Teplin et al.'. Optional notes may be included at the beginning and/or end of a citation by the use of square brackets, e.g. \citep[see] [275]{El168} produces '(see Ellet 1968, 275)'; \citep[e.g.] []{Wau50} produces '(e.g. Wauchope 1950)'; \citet[chap.~2]{Str00} produces 'Strunk and White (2000, chap. 2)'. A 'plain' \cite command will produce the same result as a \citet, i.e. \cite{Wei02} will produce 'Weigel and Glazebrook (2002)'.

7.2. The list of references

References should be listed at the end of the main text in alphabetical order by authors' surnames, then chronologically (earliest first). If references have the same author(s), editor(s), etc., arrange by year of publication, with undated works at the end. A single-author entry precedes a multi-author entry that begins with the same name. If the reference list contains two or more items by the same author(s) in the same year, add a, b, etc. and list them alphabetically by title. Successive entries by two or more authors when only the first author is the same are alphabetized by co-authors' surnames. If a reference has more than ten named authors, list only the first seven, followed by 'et al.'. If a reference has no author or editor, order by title; if a date of publication is impossible to find, use 'n.d.' in its place.

The following list shows some sample references prepared in the Taylor & Francis Chicago author-date style.

References

Abebe, Misganaw, Yongwoo Shin, Yoojeong Noh, Sangbong Lee, and Inwon Lee. 2020. "Machine Learning Approaches for Ship Speed Prediction towards Energy Efficient Shipping." *Applied Sciences* 10 (7): 2325.

Bialystocki, Nicolas, and Dimitris Konovessis. 2016. "On the estimation of ship's fuel consumption and speed curve: A statistical approach." *Journal of Ocean Engineering and Science* 1 (2): 157–166.

- Breiman, Leo. 2001. "Random Forests." Machine Learning 45 (1): 5-32.
- Coraddu, Andrea, Luca Oneto, Francesco Baldi, and Davide Anguita. 2017. "Vessels fuel consumption forecast and trim optimisation: A data analytics perspective." *Ocean Engineering* 130: 351–370.
- Dietterich, Thomas G. 2000. "An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization." *Machine Learning* 40 (2): 139–157.
- Du, Yuquan, Qiang Meng, Shuaian Wang, and Haibo Kuang. 2019. "Two-phase optimal solutions for ship speed and trim optimization over a voyage using voyage report data." *Transportation Research Part B: Methodological* 122: 88–114.
- Gkerekos, Christos, Iraklis Lazakis, and Gerasimos Theotokatos. 2019. "Machine learning models for predicting ship main engine Fuel Oil Consumption: A comparative study." *Ocean Engineering* 188: 106282.
- Halevy, Alon, Peter Norvig, and Fernando Pereira. 2009. "The Unreasonable Effectiveness of Data." *IEEE Intelligent Systems* 24 (2): 8–12.
- Haranen, Michael, Pekka Pakkanen, Risto Kariranta, and Jouni Salo. 2016. "White, grey and black-box modelling in ship performance evaluation." In 1st Hull performance & insight conference (HullPIC), 115–127.
- Hastie, Trevor, Robert Tibshirani, and J. H. Friedman. 2009. The elements of statistical learning: Data mining, inference, and prediction / Trevor Hastie, Robert Tibshirani, Jerome Friedman. 2nd ed., Springer series in statistics. New York: Springer.
- Holtrop, J. 1984. "A statistical re-analysis of resistance and propulsion data." Published in International Shipbuilding Progress, ISP, Volume 31, Number 363.
- IMO. 2020. "Fourth IMO GHG Study 2020." International Maritime Organization London, UK.
- Kim, Seong-Hoon, Myung-Il Roh, Min-Jae Oh, Sung-Woo Park, and In-Il Kim. 2020. "Estimation of ship operational efficiency from AIS data using big data technology." *International Journal of Naval Architecture and Ocean Engineering* 12: 440–454.
- Kuhn, Max, and Kjell Johnson. 2013. Applied predictive modeling. New York: Springer.
- Li, Xiaohe, Yuquan Du, Yanyu Chen, Son Nguyen, Wei Zhang, Alessandro Schönborn, and Zhuo Sun. 2022. "Data fusion and machine learning for ship fuel efficiency modeling: Part I Voyage report data and meteorological data." *Communications in Transportation Research* 2: 100074.
- Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2011. "Scikit-learn: Machine learning in Python." the Journal of machine Learning research 12: 2825–2830.
- Petersen, Jóan Petur, Daniel J. Jacobsen, and Ole Winther. 2012. "Statistical modelling for ship propulsion efficiency." *Journal of Marine Science and Technology* 17 (1): 30–39.
- Rakke, Stian Glomvik. 2016. "Ship emissions calculation from AIS." .
- Smith, T.W.P., J.P. Jalkanen, B.A. Anderson, J. J. Corbett, J. Faber, S. Hanayama, E. O'Keeffe, et al. 2015. "Third IMO Greenhouse Gas Study 2014.".
- Soner, Omer, Emre Akyuz, and Metin Celik. 2018. "Use of tree based methods in ship performance monitoring under operating conditions." Ocean Engineering 166: 302–310.
- Wang, Shuaian, and Qiang Meng. 2012. "Sailing speed optimization for container ships in a liner shipping network." *Transportation Research Part E: Logistics and Transportation Review* 48 (3): 701–714.
- Yan, Ran, Shuaian Wang, and Yuquan Du. 2020. "Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship." Transportation Research Part E: Logistics and Transportation Review 138: 101930.
- Yan, Ran, Shuaian Wang, and Harilaos N. Psaraftis. 2021. "Data analytics for fuel consumption management in maritime transportation: Status and perspectives." Transportation Research Part E: Logistics and Transportation Review 155: 102489.

References

- Albiston, Catherine R. 2005. "Bargaining in the Shadow of Social Institutions: Competing Discourses and Social Change in the Workplace Mobilization of Civil Rights." Law and Society Review 39 (1): 11–47.
- Brooks, Daniel R., and Deborah A. McLennan. 2002. The Nature of Diversity: An Evolutionary Voyage of Discovery. Chicago: University of Chicago Press.
- Brooks, Daniel R., and E. O. Wiley. 1986. *Evolution as Entropy*. 2nd ed. Chicago: University of Chicago Press.
- Choi, Mihwa. 2008. "Contesting *Imaginaires* in Death Rituals during the Northern Song Dynasty." PhD diss., University of Chicago. ProQuest (A AT 3300426).
- Choi, Stephen J., and G. Mitu Gulati. 2008. "Bias in Judicial Citations: A Window into the Behavior of Judges?" *Journal of Legal Studies* 37 (January): 87–129.
- Ellet, Elizabeth F. L. 1968. "By Rail and Stage to Galena." In *Prairie State: Impressions of Illinois*, 1673–1967, by Travelers and Other Observers, edited by Paul M. Angle, 271–79. Chicago: University of Chicago Press.
- Fogel, Robert William. 2004a. The Escape from Hunger and Premature Death, 1700–2100: Europe, America, and the Third World. New York: Cambridge University Press.
- Fogel, Robert William. 2004b. "Technophysio Evolution and the Measurement of Economic Growth." *Journal of Evolutionary Economics* 14 (2): 217–21.
- Fowler, Melvin L. 1989. *The Cahokia Atlas: A Historical Atlas of Cahokia Archaeology*. Studies in Illinois Archaeology 6. Springfield: Illinois Historic Preservation Agency.
- Greenberg, Joel, ed. 2008. Of Prairie, Woods, and Water: Two Centuries of Chicago Nature Writing. Chicago: University of Chicago Press.
- Jacobs, Sue-Ellen, Wesley Thomas, and Sabine Lang, eds. 1997. Two-Spirit People: Native American Gender Identity, Sexuality, and Spirituality. Urbana: University of Illinois Press.
- Levitt, Steven D., and Stephen J. Dubner. 2005. Freakonomics: A Rogue Economist Explores the Hidden Side of Everything. New York: William Morrow.
- Morasse, Sébastien, Helga Guderley, and Julian J. Dodson. 2008. "Paternal Reproductive Strategy Influences Metabolic Capacities and Muscle Development of Atlantic Salmon (Salmo salar L.) Embryos." Physiological and Biochemical Zoology 81 (4): 402–13.
- Rohde, Hannah, Roger Levy, and Andrew Kehler. 2008. "Implicit Causality Biases Influence Relative Clause Attachment." Poster presented at the 21st CUNY Conference on Human Sentence Processing, Chapel Hill, NC, March.
- Samples, John. 2006. "The Origins of Modern Campaign Finance Law." In *The Fallacy of Campaign Finance Reform*, Chap. 7. Chicago: University of Chicago Press.
- Schuman, Howard, and Jacqueline Scott. 1987. "Problems in the Use of Survey Questions to Measure Public Opinion." *Science* 236: 957–59.
- Strunk, William, Jr., and E. B. White. 2000. The Elements of Style. 4th ed. New York: Allyn and Bacon.
- Teplin, Linda A., Gary M. McClelland, Karen M. Abram, and Jason J. Washburn. 2005. "Early Violent Death in Delinquent Youth: A Prospective Longitudinal Study." Paper presented at the Annual Meeting of the American Psychology-Law Society, La Jolla, CA, March.
- Wauchope, Robert. 1950. A Tentative Sequence of Pre-Classic Ceramics in Middle America. Vol. 1 of Middle American Research Records. New Orleans, LA: Tulane University.
- Weigel, Detlef, and Jane Glazebrook. 2002. Arabidopsis: A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

This was produced by typing:

\begin{thebibliography}{}

\bibitem[Albiston(2005)]{Alb05} Albiston, Catherine~R. 2005. 'Bargaining in the Shadow of Social Institutions: Competing Discourses and Social Change in the Workplace Mobilization of Civil Rights.', \emph{Law and Society Review} 39 (1): 11--47.

\bibitem[Brooks and McLennan(2002)]{Bro02}

Brooks, Daniel~R., and Deborah~A. McLennan. 2002. \emph{The Nature of Diversity: An Evolutionary Voyage of Discovery}. Chicago: University of Chicago Press.

\bibitem[Brooks and Wiley(1986)]{Bro86}

Brooks, Daniel~R., and E.~O. Wiley. 1986. \emph{Evolution as Entropy}. 2nd ed. Chicago: University of Chicago Press.

\bibitem[M. Choi(2008)]{Cho08}

Choi, Mihwa. 2008. ''Contesting \emph{Imaginaires} in Death Rituals during the Northern Song Dynasty.'' PhD diss., University of Chicago. ProQuest (A\,AT\,3300426).

\bibitem[S.~J. Choi and Gulati(2008)]{ChoGul08}

Choi, Stephen J., and G. Mitu Gulati. 2008. "Bias in Judicial Citations: A Window into the Behavior of Judges?" \emph{Journal of Legal Studies} 37 (January): 87--129.

\bibitem[Ellet(1968)]{Ell68}

Ellet, Elizabeth F.~L. 1968. ''By Rail and Stage to Galena.'' In \emph{Prairie State: Impressions of Illinois, 1673--1967, by Travelers and Other Observers}, edited by Paul~M. Angle, 271--79. Chicago: University of Chicago Press.

\bibitem[Fogel(2004a)]{FogEHPD04}

Fogel, Robert~William. 2004a. \emph{The Escape from Hunger and Premature Death, 1700--2100: Europe, America, and the Third World}. New York: Cambridge University Press.

\bibitem[Fogel(2004b)]{FogJEE04}

Fogel, Robert~William. 2004b. "Technophysic Evolution and the Measurement of Economic Growth." \emph{Journal of Evolutionary Economics} 14 (2): 217--21.

\bibitem[Fowler(1989)]{Fow89}

Fowler, Melvin~L. 1989. \emph{The Cahokia Atlas: A Historical Atlas of Cahokia Archaeology}. Studies in Illinois Archaeology~6. Springfield: Illinois Historic Preservation Agency.

\bibitem[Greenberg(2008)]{Gre08}

Greenberg, Joel, ed. 2008. \emph{Of Prairie, Woods, and Water: Two Centuries of Chicago Nature Writing}. Chicago: University of Chicago Press.

\bibitem[Jacobs, Thomas, and Lang(1997)]{JTL97}

Jacobs, Sue-Ellen, Wesley Thomas, and Sabine Lang, eds. 1997. \emph{Two-Spirit People: Native American Gender Identity, Sexuality, and Spirituality}. Urbana: University of Illinois Press.

\bibitem[Levitt and Dubner(2005)]{Lev05} Levitt, Steven~D., and Stephen~J. Dubner. 2005. \emph{Freakonomics: A Rogue Economist Explores the Hidden Side of Everything}. New York: William Morrow.

\bibitem[Morasse, Guderley, and Dodson(2008)]{Mor08} Morasse, S{\'e}bastien, Helga Guderley, and Julian~J. Dodson. 2008. ''Paternal Reproductive Strategy Influences Metabolic Capacities and Muscle Development of Atlantic Salmon (\emph{Salmo salar} L.) Embryos.'' \emph{Physiological and Biochemical Zoology} 81 (4): 402--13.

\bibitem[Rohde, Levy, and Kehler(2008)]{Roh08}
Rohde, Hannah, Roger Levy, and Andrew Kehler. 2008. ''Implicit
Causality Biases Influence Relative Clause Attachment.'' Poster
presented at the 21st CUNY Conference on Human Sentence Processing,
Chapel Hill, NC, March.

\bibitem[Samples(2006)]{Sam06}

Samples, John. 2006. "The Origins of Modern Campaign Finance Law." In \emph{The Fallacy of Campaign Finance Reform}, Chap. 7. Chicago: University of Chicago Press.

\bibitem[Schuman and Scott(1987)]{Sch87}
Schuman, Howard, and Jacqueline Scott. 1987. 'Problems in the Use of
Survey Questions to Measure Public Opinion.' \emph{Science} 236: 957--59.

\bibitem[Strunk and White(2000)]{Str2000}
Strunk, William, Jr., and E.~B. White. 2000. \emph{The Elements of Style}. 4th ed. New York: Allyn and Bacon.

\bibitem[Teplin et~al.(2005)]{Tep05}

Teplin, Linda~A., Gary~M. McClelland, Karen~M. Abram, and Jason~J. Washburn. 2005. ''Early Violent Death in Delinquent Youth: A Prospective Longitudinal Study.'' Paper presented at the Annual Meeting of the American Psychology-Law Society, La Jolla, CA, March. \bibitem[Wauchope(1950)]{Wau50}

Wauchope, Robert. 1950. \emph{A Tentative Sequence of Pre-Classic Ceramics in Middle America}. Vol.~1 of \emph{Middle American Research Records}. New Orleans, LA: Tulane University.

\bibitem[Weigel and Glazebrook(2002)]{Wei02}
Weigel, Detlef, and Jane Glazebrook. 2002. \emph{\emph{Arabidopsis}:
 A Laboratory Manual}. Cold Spring Harbor, NY: Cold Spring Harbor
 Laboratory Press.

\end{thebibliography}

Each entry takes the form:

```
\bibitem[authors' names(date of publication)]{key}
Bibliography entry
```

where 'authors' names' is the list of names to appear where the bibitem is cited in the text, and 'key' is the tag that is to be used as an argument for the \cite{} commands in the text of the article. 'Bibliography entry' is the material that is to appear in the list of references, suitably formatted. The commands

```
\usepackage{natbib}
\bibpunct[, ]{(){)}{;}{a}{},}
\renewcommand\bibfont{\fontsize{10}{12}\selectfont}
```

need to be included in the preamble of your .tex file in order to generate the citations and bibliography as described above.

Instead of typing the bibliography by hand, you may prefer to create the list of references using a BibTeX database. The tfcad.bst file needs to be in your working folder or an appropriate directory, and the lines

```
\bibliographystyle{tfcad}
\bibliography{interactcadsample}
```

included where the list of references is to appear, where tfcad.bst is the name of the BIBTEX bibliography style file for Taylor & Francis' Chicago author-date reference style and interactcadsample.bib is the bibliographic database included with the Interact-CAD IATEX bundle (to be replaced with the name of your own .bib file). IATEX/BIBTEX will extract from your .bib file only those references that are cited in your .tex file and list them in the References section.

Please include a copy of your .bib file and/or the final generated .bbl file among your source files if your .tex file does not contain a reference list in a thebibliography environment.

8. Appendices

Any appendices should be placed after the list of references, beginning with the command \appendix followed by the command \section for each appendix title, e.g.

```
\appendix
\section{This is the title of the first appendix}
\section{This is the title of the second appendix}
produces:
```

Appendix A. This is the title of the first appendix

Appendix B. This is the title of the second appendix

Subsections, equations, figures, tables, etc. within appendices will then be automatically numbered as appropriate. Some theorem-like environments may need to have their counters reset manually (e.g. if they are not numbered within sections in the main text). You can achieve this by using \numberwithin{remark}{section} (for example) just after the \appendix command.

Please note that if the endfloat package is used on a document containing appendices, the \processdelayedfloats command must be included immediately before

the \appendix command in order to ensure that the floats in the main body of the text are numbered as such.

Appendix A. Troubleshooting

Authors may occasionally encounter problems with the preparation of a manuscript using LATEX. The appropriate action to take will depend on the nature of the problem:

- (i) If the problem is with LATEX itself, rather than with the actual macros, please consult an appropriate LATEX 2ε manual for initial advice. If the solution cannot be found, or if you suspect that the problem does lie with the macros, then please contact Taylor & Francis for assistance (latex.helpdesk@tandf.co.uk).
- (ii) Problems with page make-up (e.g. occasional overlong lines of text; figures or tables appearing out of order): please do not try to fix these using 'hard' page make-up commands the typesetter will deal with such problems. (You may, if you wish, draw attention to particular problems when submitting the final version of your manuscript.)
- (iii) If a required font is not available on your system, allow TEX to substitute the font and specify which font is required in a covering letter accompanying your files.

Appendix B. Obtaining the template and class file

B.1. Via the Taylor & Francis website

This article template and the interact class file may be obtained via the 'Instructions for Authors' pages of selected Taylor & Francis journals.

Please note that the class file calls up the open-source LATEX packages booktabs.sty, epsfig.sty and rotating.sty, which will, for convenience, unpack with the downloaded template and class file. The template calls for natbib.sty and subfigure.sty, which are also supplied for convenience.

B.2. Via e-mail

This article template, the interact class file and the associated open-source LATEX packages are also available via e-mail. Requests should be addressed to latex.helpdesk@tandf.co.uk, clearly stating for which journal you require the template and class file.