On reconciling macro and micro energy transport forecasts for strategic decision making in the tanker industry

Wessam Abouarghoub[a], Nikos Nomikos[b], Fotios Petropoulos[c,*]

[a] *Cardiff Business School, Cardiff University, UK*
[b] *Cass Business School, City University of London, UK*
[c] *School of Management, University of Bath, UK*

Abstract

We propose the use of hierarchical structures for forecasting freight earnings. Hierarchical time series approaches are applied and contrasted in the dry-bulk and tanker markets in order to identify the most suitable strategy for forecasting freight rates. We argue that decision making for shipping practitioners should be based on forecasts of freight earnings at different hierarchical levels. In other words, different strategic shipping decisions such as operations management, choice of freight charter contract and type of investment should be matched with the appropriate level of forecasts of freight earnings that are aggregated at different macro-levels: operating route, vessel size and type of trade, and for different forecasting horizons: short-, medium- and long-run.

Keywords: forecasting, freight revenues, shipping energy, freight earnings, hierarchical aggregation

1. Introduction

It is well known that the dynamics of freight revenues differ across the various segments of the shipping market, depending on the market sector, the size of the vessel as well as the underlying trading route at which each vessel operates. Hitherto, empirical work in terms of modeling and forecasting freight rates has focused on a specific level of the market thus ignoring whether the structure and hierarchy within the market have an impact in forecasting and modeling freight rates.

In many applications, one may find time series that are hierarchically organized and can be aggregated in groups based on their features. In shipping markets, we can find many instances of similar "hierarchical time series" that can be aggregated at different levels according to market sector (dry bulk or tanker), market segment across the same

*Correspondance: F Petropoulos, East Building, School of Management, University of Bath, Claverton Down, Bath, BA2 7AY, UK.

Email addresses: abouarghoubw@cardiff.ac.uk (Wessam Abouarghoub), n.nomikos@city.ac.uk (Nikos Nomikos), f.petropoulos@bath.ac.uk (Fotios Petropoulos)

sector (e.g. Very Large Crude Carriers (VLCC) of 260,000 metric tonnes dead-weight (mt dwt) versus Suezmax tanker vessels of 125,000 mt dwt) as well as in terms of geographical criteria according to the underlying trade routes (e.g. East-bound or West-bound VLCC routes from Middle-East Gulf). Typically, the approach that the market has adopted is to model and forecast the series either by considering each of the time series in isolation or by examining its interactions with other variables at the same level of hierarchy.

Hierarchical time series are commonly analyzed using either a "top-down" or a "bottom-up" method, or a combination of the two. The top-down method entails forecasting the completely aggregated series, and then disaggregating the forecasts based on historical proportions. The bottom-up method involves forecasting each of the disaggregated series at the lowest level of the hierarchy, and then using simple aggregation to obtain forecasts at higher levels of the hierarchy. In practice, many businesses combine these methods, in what is sometimes called the "middle-out" method, where forecasts are obtained for each series at an intermediate level of the hierarchy, and then aggregation is used to obtain forecasts at higher levels and disaggregation is used to obtain forecasts at lower levels.

Similar approaches should also be helpful for stakeholders in the shipping industry (shipowners, charterers, brokers, investors) who are faced with challenging strategic decisions that require valuing and assessing shipping revenues at different levels. On the one hand, decisions at the micro-operational level refer mainly to the choice of trading routes, type of contract to use (i.e. period time-charter contract or spot voyage contract), bunker arrangements and steaming speeds. Typically, such decisions are captured by the lower level of a hierarchical structure. On the other hand, decisions at the macro-investment level refer to which segment and sector of the market to operate and are typically more macro-based. Ideally, market participants should generate forecasts across all possible freight routes at which each vessel operates, but also across each vessel type and even market sector in order to decide how to deploy their fleet but also how to invest their capital.

Hitherto, the maritime economic literature that investigates freight dynamics has mainly focused on comparing forecasting methodologies for freight revenues and not on discussing forecasting strategies, as if, the generated forecasts for freight revenues at any hierarchy level are appropriate for making investment and operational decisions alike. The focus of the current literature is the accuracy of freight revenues forecasts across different methodologies and not at different aggregation levels. For instance, most studies consider the series across each tier, either on an individual basis using univariate models as in Cullinane (1992) or multivariate models as in Kavussanos and Nomikos (2003). However, none of these studies considers the inherent correlation structure of the hierarchy in freight revenues. This research explores, for the first time, how reconciliation of revenues forecasts produced at different levels can improve forecast accuracy and hence decision making in the shipping transportation sector. This may benefit shipping practitioners in making more informed decisions, as a direct result of reconciled forecasts that consider information at various levels of the hierarchy. This is important since the various components of the hierarchy can interact in varying and complex ways. A change in one series at one level, can have a consequential impact on other series at the same level, as well as series at higher and lower levels. By modeling the entire hierarchy of a time series simultaneously, we can obtain better forecasts of the component series simply because such complex interactions can be accounted for.

The remainder of the paper is organized as follows. Section 2 provides a background review of the literature on shipping freight rate dynamics and forecasting. Section 3 describes the data and presents the empirical methodology used in the paper. Section 4 presents the empirical results and some practical examples on how the proposed methodology can be used in practice. Finally, Section 5 concludes.

2. Literature Review

The main focus of the existing freight modeling literature is on investigating the dynamics of spot freight rates Cullinane (1992), or between spot and forward freight rates (see studies by Kavussanos and Nomikos (1999) and (2003), Batchelor et al. (2007) and Zhang et al. (2014)). The consensus from these studies is that univariate models are useful tools for forecasting individual freight rate series while, when one combines spot and forward rates a Vector Autoregressive (VAR) or a Vector Error Correction Model (VECM) may be preferred. In addition, the majority of studies (see Table 1) that investigate freight dynamics using spot, Forward Freight Agreements (FFAs) and Time-Charter time series, actually model freight rates across the same level of hierarchy. For example, studies by Cullinane (1992), Kavussanos and Nomikos (2003), Batchelor et al. (2007), and Zhang et al. (2014) focus on one particular level of hierarchy, while studies by Kavussanos and Dimitrakopoulos (2011) and Abouarghoub et al. (2014) focus on more than one level but these studies are rather limited in their scope and do not consider all possible levels of hierarchy. In addition, none of these studies use forecasting at different levels of hierarchy for strategic decisions.

Table 1: Key references to hierarchical mapping in maritime literature.

Focus of the Study	**Freight Market**			**Hierarchy-Level**		
	Spot	Forward	TC	Trade	Size	Route
Cullinane (1992)	✓				✓	
Kavussanos and Nomikos (2003)	✓	✓				✓
Batchelor et al. (2007)	✓	✓				✓
Kavussanos and Dimitrakopoulos (2011)	✓			✓		✓
Zhang et al. (2014)	✓	✓				✓
Abouarghoub et al. (2014)	✓		✓	✓		✓
This Study	✓		✓	✓	✓	✓

It seems therefore that the focus in the literature has been in comparing different forecasting methodologies to improve accuracy rather than looking at forecasting strategies for different hierarchy levels. The motivation for the use of hierarchical models is as follows: Shipping transport is a non-storable service that is provided by a capital intensive industry where short and long-term revenues are not linked through an arbitrage relationship (Kavussanos and Nomikos (1999)). Freight revenues are highly volatile (Alizadeh and Nomikos (2009)), seasonal (Kavussanos and Alizadeh (2001)), sensitive to energy prices and market sentiment (Papapostolou et al. (2014)) and are considered to be mean-reverting in the long-run and subject to demand and supply imbalances in the short-run. For instance, Adland and Cullinane (2006) show that freight rates are non-stationary over short periods of time, yet are mean reverting over longer periods, as implied by maritime economic theory. This finding is consistent with the view that associates the dependency of conditional freight revenues to different regimes in the market (Abouarghoub et al. (2014)). In addition, shipping freight revenues for different vessel sizes and on different trading routes are highly correlated and move together in the long-run, but behave differently in the short-run. The co-movement in the long-run is due to the nature of the demand function for sea transport that is derived from macroeconomic factors such as global economic growth and demand for seaborne trade. In the short-run, the dynamics of freight revenues are shaped by factors such as the availability of vessels and cargoes in each specific market segment. Thus, it is reasonable to suggest that forecasting freight revenues in the short-term may require a different approach compared to forecasting in the long-term. In that respect, hierarchical approaches are particularly useful, especially for longer term horizons when the complex interactions across the hierarchies may play a more important role.

The main argument of this study can be summarized as follows: strategic decision making for shipping practitioners should be based on forecasts of freight earnings at different hierarchical levels. In other words, different strategic shipping decisions such as operations management, choice of freight charter contract and type of investment should be matched with the appropriate level of forecasts of freight earnings that are aggregated at different macro-levels: operating route, vessel size and type of trade, and for different forecasting horizons: short-, medium- and long-run. This argument is presented graphically in Table 2.

Table 2: Hierarchical Shipping Decision Matrix.

Level	**Description**	**Level of Earnings**	**Decision Type**
Top level	Trade (Macro)	Sector Earnings (e.g. Tankers)	Financing and Investment
Middle level	Size (Middle)	Segment Earnings (e.g. VLCC)	Fleet Diversification
Bottom level	Route (Micro)	Route-Specific Earnings	Fleet Repositioning

First, forecasts of shipping earnings generated at the bottom level for different tanker routes can be used to make short-term operational decisions, such as fleet repositioning as owners relocate their fleet according to short- and medium-term profitability in different routes. Second, forecasts of shipping earnings generated at the middle level, enable decisions on fleet diversification across segments (e.g. by chartering-in different size vessels in the same sector) and chartering policy; for instance, if future spot voyage earnings are forecasted to be below the corresponding period rate then this may be taken as a signal to fix the vessel in the period market, as we demonstrate in section 4.2. Finally, forecasts of shipping earnings generated at the top level may be used to make medium to long-term investment decisions such as buying or selling second-hand ships or ordering new building vessels.

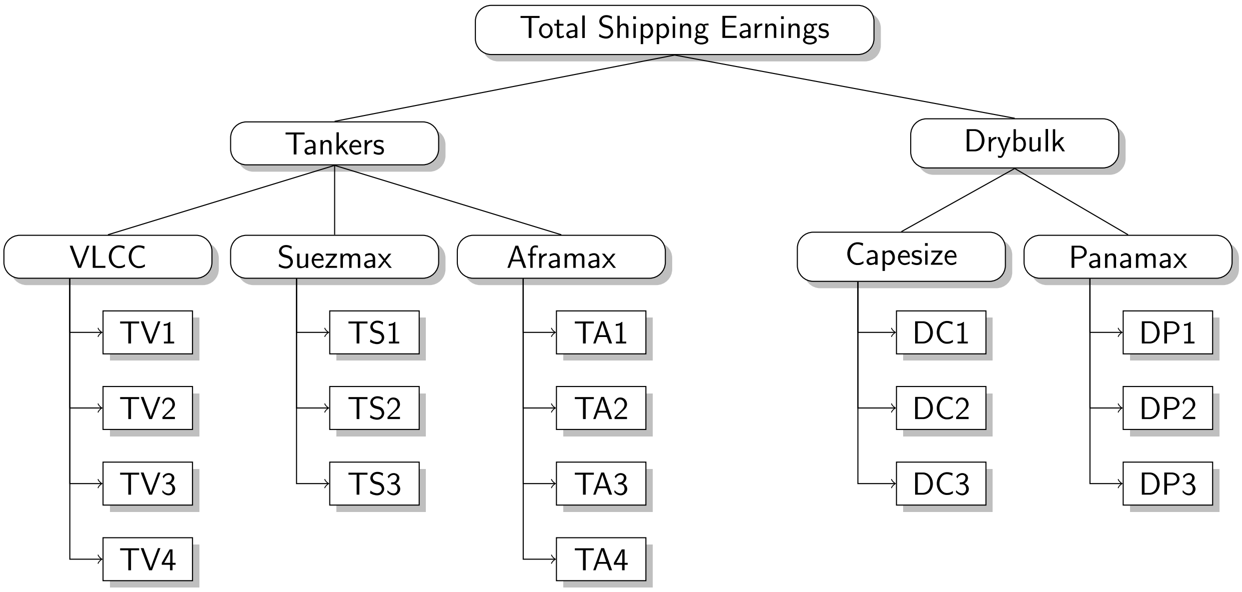

Figure 1: Hierarchy structure of shipping earnings categorized by trade, size and route.

3. Design and data description

Shipping markets are considered as one entity with a common market sentiment (Randers and Gluke (2007)) which results in a highly correlated freight market. Thus, our proposed hierarchical structure reflects the strong degree of substitutability of cargoes across different routes and shipping freight revenues and their forecasts are aggregated at four levels. First, the lower-level of the hierarchy represents shipping routes; forecasts generated at this level are more relevant for short-term operational decisions, such as fleet repositioning. Second, the middle-level hierarchy represents market segment (size of vessel) and forecasts generated at this level are more relevant for long- to medium-term operational and investment decisions, such as fleet diversification. The third hierarchy represents the type of trade (tanker or dry bulk) and forecasts generated at this level are more relevant for macro-level planning. Finally, the top level reflects aggregate shipping earnings. The different hierarchy levels for trade, size and route are presented in Figure 1.

The primary data used in this study for the lower level of hierarchy are freight earnings for large tanker and dry bulk carriers that specialize in transporting wet (crude oil) and dry (bulk) commodities. Tanker vessels are categorized by cargo capacity into three sizes: VLCC of 260,000 mt dwt, Suezmax (125,000 mt dwt) and Aframax (75,000 mt dwt). Dry bulk vessels are categorized into two sizes: Capesize (170,000 mt dwt) and Panamax (74,000 mt dwt). Table 3 describes the shipping routes under investigation which also refer to the most disaggregated level of the hierarchy depicted in Figure 1. The first and second columns in Table 3 are the number and code of the nodes depicted in Figure 1, the third and fourth columns describe the trade and shipping route, respectively and the final column refers to the size of the vessels. For a complete list of shipping routes and historical developments of tanker trades see Alizadeh and Nomikos (2009). The underlying data are monthly freight earnings, measured in US dollars per day, also known as time-charter-equivalent (TCE) rates. These are calculated by taking the total revenue for a specific voyage and then deducting current bunker costs based on prices at representative regional bunker ports, estimated port costs and total commission and then dividing the result by the number of voyage days. In addition, we also use one-year time charter rates to test the reliability and usefulness of the results for ship-owners by running an operational exercise, which is presented in section 4.2. Data are provided by Clarkson Shipping Intelligence Network [1].

3.1. ***Experimental framework***

We apply an ARIMA model for modeling freight earnings at different hierarchical levels. ARIMA models are selected in this case due to their versatility and proven effectiveness in modeling freight rates (see e.g. Cullinane (1992)). We note that the proposed methodology is flexible and is model-independent; as such, ARIMA could be replaced by any other forecasting model. One could even select different models for each level, for example ARIMA for the bottom-level series, VAR for the middle level and judgment for the top level. However, we do not consider other forecasting models in this study as the purpose here is not to compare different models, but to present a framework that lies one conceptual level higher than the models.

Monthly freight earnings are the base for our forecasts at the lowest hierarchy level, the sums of those earnings for each vessel size are the base for our forecasts at the middle level, the total sums of size freight earnings are the base for our forecasts at the trade level and finally, the sums of those are the base for our forecasts at the top level. In other words, our forecasts at lower, lower middle, upper middle and top levels are average earnings for specific routes, aggregate earnings for each size category, aggregate earnings for tankers and drybulk trades and aggregate shipping earnings, respectively.

Statistical forecasts based on historical data can be produced at every of the three levels of each sub-hierarchy considered in Figure 1, namely Tankers and Drybulk, but also for each of the four levels of the complete hierarchy (Total Shipping Earnings). However, forecasts created at the lower levels do not exactly sum up to the forecasts calculated directly at the upper levels. Equally, forecasts produced at the top level can be disaggregated to the lower level forecasts, however there will be some deviations from the estimates directly produced at the lower levels.

[1]An anonymous reviewer suggested to take into account the effect of fleet size, geographical distribution and duration of different routes by considering weightings in the TCE data. Although we believe that the estimated TCE accounts for those dynamics, as it is calculated by discounting bunker consumption and other voyage costs from the voyage revenue, we considered two additional approaches for calculating shipping freight earnings as further robustness tests. The first approach, accounts for fleet size development (DWT) on each route using weights calculated from number of fixtures on each route as follows: $Earnings_{r,t} = (TCE_{r,t}) \times (FleetSize_{s,t}) \times (Fixtures_{r,t})/(TotalFixtures_{s,t})$, where r stands for route and s stands for vessel size. The second approach accounts for average haul of each route to capture demand for seaborne trade using the following formula: $Earnings_{r,t} = (TCE_{r,t}) \times (VesselSize) \times (Distance) \times (Fixtures_{r,t})/(TotalFixtures_{s,t})$. Results using those series are qualitatively similar to the results based on TCE, reported in the next section, and are available upon request from the authors.

Table 3: Hierarchical at the bottom level and shipping routes.

No	Node	Trade	Shipping Route	Vessel Size
1	TV1	Arab Gulf-East	Ras Tanura (SA) - Singapore (SG)	VLCC
2	TV2	WAF-T/A	Bonny Offshore (NG) - LOOP (US)	VLCC
3	TV3	Arab Gulf-West	Ras Tanura (SA) - Rotterdam (NL)	VLCC
4	TV4	Arab Gulf-Japan	Ras Tanura (SA) - Chiba (JP)	VLCC
5	TS1	Other Trades	Rest of the world	Suezmax
6	TS2	WAF-T/A	Bonny Offshore (NG) - Philadelphia (US)	Suezmax
7	TS3	Med-Med	Sidi Kerir (EG) - Fos (FR)	Suezmax
8	TA1	Other Trades	Rest of the world	Aframax
9	TA2	Med-Med	Sidi Kerir (EG) - Trieste (IT)	Aframax
10	TA3	Baltic-Continent		Aframax
11	TA4	Arab Gulf-SG	Ras Tanura (SA) - Singapore (SG)	Aframax
12	DC1	T/A RV	BCI Route C8_03	Capesize
13	DC2	Trip Far-East	BCI Route C9_03 & C11_03	Capesize
14	DC3	NOPAC RV	BCI Route C10_03	Capesize
15	DP1	T/A RV	BPI Route P1A_03	Panamax
16	DP2	Trip Far-East	BPI Route P2A_03 & P4_03	Panamax
17	DP3	NOPAC RV	BPI Route P3A_03	Panamax

In this table we describe shipping routes for the most disaggregated level in our hierarchy, the columns from left to right are route number, node codes, loading and discharging ports (country) and vessel size.
The letters T, V, S, A, D, C and P in the node column stands for Tanker, VLCC, Suezmax, Aframax, Drybulk, Capesize and Panamax, respectively.
WAF stands for West Africa. T/A stands for Trans (across) the Atlantic.
LOOP stands for Louisiana Offshore Oil Port.
T/A RV stands for Trans-Atlantic Round Voyage.
NOPAC RV stands for North Pacific Round Voyage.
BCI stands for Baltic Capesize Index.
BPI stands for Baltic Panamax Index.
Country codes are NL: Netherlands; SG: Singapore; US: United States; JP: Japan.
FR: France; IT: Italy; SA: Saudi Arabia; EG: Egypt; NG: Nigeria;

To tackle this problem, several hierarchical reconciliation approaches have been considered in the literature, with bottom-up and top-down being the most popular ones.

Bottom-up approach is arguably the most widely applied approach in hierarchical forecasting. (see for example Dangerfield and Morris (1992); Zellner and Tobias (2000); Athanasopoulos et al. (2009)). In the bottom-up approach, forecasts are calculated only at the lowest hierarchical level; these forecasts are subsequently summed up to create forecasts of higher levels.

Top-down approach produces forecasts only at the highest hierarchical level; these forecasts are then split down to other levels using appropriate proportions. The proportions are usually calculated as the historical percentages that each lower level node contributes to the grand total (see for example Lütkepohl (1984); Fliedner (1999); Gross and Sohl (1990)). More recently, Athanasopoulos et al. (2009) proposed the disaggregation of the top level forecasts using forecasted proportions. By doing so, several dynamics that appear at the lower aggregation levels (but are smoothed at the higher levels due to aggregation) can be reintroduced. They showed that top-down with forecasted proportions can offer improvements when such dynamics exist.

Middle-out is yet another approach that is a mix of the bottom-up and top-down approaches. Assuming that at least three hierarchical levels exist, forecasts are produced at a middle level. Forecasts for levels lower to that level are derived by disaggregation (in this study we assume historical proportions); forecasts for higher levels are produced by summing up the forecasts of the respective nodes. If more than one middle levels exist, then one should first decide on the middle level to be used for forecasting. In this paper, we have only one middle level for each of the two sub-hierarchies considered, but two middle levels for the total hierarchy. In the latter case, two middle-out approaches are investigated (one for each middle level).

More recently, combination approaches have been considered in the literature. Specifically, Hyndman et al. (2011) and Athanasopoulos et al. (2009) proposed the "optimal reconciliation" approach. This approach, statistically and optimally combines the forecasts produced at all levels. The calculation of the optimal weights is given by $w = S(S'S)^{-1}S'$, where S is a Boolean matrix that directly corresponds to the hierarchical structure of the data. For the total hierarchy (Total Shipping Earnings) used in this paper,

$$S = \begin{bmatrix} 1_{17} \\ 1_{11}\ 0_6 \\ 0_{11}\ 1_6 \\ 1_4\ 0_3\ 0_4\ 0_3\ 0_3 \\ 0_4\ 1_3\ 0_4\ 0_3\ 0_3 \\ 0_4\ 0_3\ 1_4\ 0_3\ 0_3 \\ 0_4\ 0_3\ 0_4\ 1_3\ 0_3 \\ 0_4\ 0_3\ 0_4\ 0_3\ 1_3 \\ diag(1_{17}) \end{bmatrix}, \quad (1)$$

where x_y is an horizontal vector of values x and order y while $diag(\cdot)$ refers to the diagonal matrix. The full matrix is presented in equation A.1 of the appendix. Each row corresponds to one node of the hierarchy (25 for the total hierarchy) and each column to one node at the very bottom level (17 for the total hierarchy). Each row suggests (0 or 1) which bottom-level series should be summed to form the series for the respective level. For instance, the 17 ones of the first row suggest that the grand total is the sum of all bottom-level series. At the same time, the second row suggests that Tanker Earnings can be calculated as the sum of the first 11 series. The reconciled forecasts, $\tilde{y}$, can be calculated as $\tilde{y} = w\hat{y}$, where $\hat{y}$ are the original forecasts calculated directly at each level.

The corresponding S matrix for the first sub-hierarchy (Tankers Earnings) is:

$$S = \begin{bmatrix} 1_{11} \\ 1_4 \; 0_3 \; 0_4 \\ 0_4 \; 1_3 \; 0_4 \\ 0_4 \; 0_3 \; 1_4 \\ diag(1_{11}) \end{bmatrix} \tag{2}$$

and for the second (Drybulk):

$$S = \begin{bmatrix} 1_6 \\ 1_3 \; 0_3 \\ 0_3 \; 1_3 \\ diag(1_6) \end{bmatrix}. \tag{3}$$

A disadvantage of the optimal combination approach is that it gives more emphasis to the higher level forecasts. Furthermore, it assumes that errors are equivariant and uncorrelated. Towards both ends, Athanasopoulos et al. (2017), focusing on the temporal hierarchies, propose the use of a scaling matrix W so that $w = S(S'W^{-1}S)^{-1}S'W^{-1}$. One way to estimate W is directly from the structure of the hierarchy and the number of bottom level nodes contributing to each node. For the total hierarchy of this study, $W = diag(17, 11, 6, 4, 3, 4, 3, 3, 1_{17})$, with $W = diag(11, 4, 3, 4, 1_{11})$ and $W = diag(6, 3, 3, 1_6)$ for the first and second sub-hierarchies respectively.

We additionally propose and evaluate two simple combination approaches. The first, simply combines the outputs of the bottom-up and top-down approaches, while the second additionally considers the outputs of the middle-out approach. Once forecasts have been calculated through these approaches, then a simple (equally-weighted) combination is applied. Effectively, we combine information (forecasts) with equal weights across top and bottom or across all levels. As such, this new approach is considered to lie in-between the bottom-up and optimal combination in terms of computational complexity, however it is simpler than the optimal combination approach with regards to calculating the final forecasts. In any case, we do not suggest that the application of optimal combination is not practical, as recent research has proposed fast computations of the reconciled forecasts even for large hierarchies (Hyndman et al., 2016).

In brief, we apply and compare the following hierarchical approaches:

- Bottom-up approach, where forecasts are calculated only on the most granular level and forecasts of other levels are directly derived as appropriate sums of the bottom level forecasts. We denote this approach as BU.
- Middle-out, where forecasts are produced at a middle level. Forecasts at other levels are calculated by appropriate aggregations and disaggregations. We denote this approach as MO, followed by the name of the level if more than one middle levels exist.
- Top-down approach, where only the forecast for the grand total is produced and subsequently disaggregated to the lower levels. We consider both disaggregations via historical and forecasted proportions, denoted by TDHP and TDFP, respectively.
- Simple combination approaches, where the forecasts derived from BU, TD (and MO) are combined using equal weights. We denote these as COMB BU+TD and COMB BU+MO+TD.
- Optimal combination approach, based on the scaled weights as shown above. We denote this approach as COMB Optimal.

Forecasts at all levels are produced using the automatic ARIMA approach (Hyndman and Khandakar, 2008, `auto.arima()` function) implemented in the `forecast` package for the R statistical software. In more detail, this approach identifies and estimates the most appropriate ARIMA model for each time series individually. The data cover the period from January 1996 for the tanker earnings sub-hierarchy and from March 1999 for the dry-bulk earnings sub-hierarchy and the total hierarchy.

Forecasting performance is evaluated over the last three years of the available data (from 01/2013 till 12/2015) in a rolling origin manner. First, the in-sample period consists of data up to December 2012 and monthly forecasts for the next year (01/2013 till 12/2013) are produced. Then, the in-sample increases by one month and forecasts are produced for the period 02/2013 till 1/2014. The procedure is repeated till the point that the forecast origin is the end of 2014 (12/2014 is the last available observation in the in-sample) and monthly forecasts are produced for 2015. By following this process, we produce forecasts from 25 origins. In all cases, the forecasting horizon is equal to 12 months.

Note that as we roll through the sample, different optimal models (and coefficients) may be estimated for a time series. Table 4 shows the percentage of times an order (autoregressive or moving average, seasonal or not) is selected as optimal at the various aggregation levels. We can see that, while at the very top level the aggregated series appears to be white noise, different models are identified as optimal at lower aggregation levels. Also, it should be noted that at Size and Route levels, seasonality might exist, which is not the case for Trade or Total levels.

4. Results

4.1. Empirical evaluation

We measure the forecasting performance by the Mean Absolute Error (MAE), an error measure that is suitable for measuring forecast accuracy. MAE refers to the average (across origins and horizons) of the absolute errors. The forecast error is defined as the actual minus the forecast, or $y - \hat{y}$. The MAE is calculated for each node of the hierarchy and for each

Table 4: Frequencies that each autoregressive and moving average order is identified as optimal in the fitted ARIMA models.

Level	Non-seasonal part						Seasonal part					
	AR			MA			AR			MA		
	0	1	2	0	1	2	0	1	2	0	1	2
Total	100.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	0.0
Trade	98.0	2.0	0.0	50.0	2.0	48.0	100.0	0.0	0.0	100.0	0.0	0.0
Size	88.8	11.2	0.0	69.6	29.6	0.8	86.4	12.8	0.8	73.6	0.0	26.4
Route	97.6	2.4	0.0	87.1	12.7	0.2	91.5	7.8	0.7	84.5	0.5	15.1

of the hierarchical approaches. The percentage differences in forecasting performance are captured by the percentage decrease in the value of MAE relatively to the value of a MAE for the bottom-up approach, MAE_{BU}, or

$$\%Improvement = 100\left(1 - \frac{MAE}{MAE_{BU}}\right), \tag{4}$$

A positive difference indicates increase (improvement) in the forecasting performance (accuracy). For example, the improvements of optimal combination over bottom-up can be captured by replacing the numerator of equation 4 with MAE_{OC}. These differences are subsequently averaged across all nodes of the same hierarchical level.

Note that the ratio we consider here is equivalent to the Relative MAE ($RelMAE$) discussed in Hyndman and Koehler (2006). They suggest that interpretability is an advantage of this measure, however one should be careful when applying this to measure the performance across series of different scales. Thus, we calculate the improvement in performance not only on the hierarchy as a whole but also at each level separately. Additionally, we considered the percentage improvement based on the Relative Mean Squared Error ($RelMSE$). The results provide the same insights as the MAE and as such are not presented here.

Table 5 reports the percentage improvements of the various hierarchical approaches over the bottom-up (BU) approach for the first sub-hierarchy (Tankers Earnings). The table presents the performance for the three levels of the hierarchy (trade, size and route) as well as the improvements across the hierarchy as a whole, by calculating the mean improvement across all relative improvements. Columns report the performance improvements for different forecasting horizons (in months).

We observe that, on average, all approaches outperform the bottom-up approach. This is a direct result of the high volatility observed at the very granular (route) level, which is smoothed-out at higher aggregation levels and thus, improves forecastability. Performance improvements are also positively correlated with the forecasting horizon, with larger improvements recorded for the longer horizons. In most cases, improvements seem to be consistent across the various aggregation levels within each approach. The only exception is the top-down based on forecasted proportions (TDFP) approach, which for the shortest horizons (1 to 3 months ahead) has a decrease in the performance of the middle and bottom levels.

Top-down with historical proportions (TDHP) outperforms all other methods on average, followed by the middle-out approach (which is based, as well, on historical proportions when disaggregating the middle level forecasts to the bottom-level ones). Combination approaches, which consider combinations of forecasts across all levels, offer much lower improvement compared to the top-down and middle-out approaches. Interestingly, the optimal combination approach (COMB Optimal) is outperformed by its simpler counterpart, the simple combination based on all levels (COMB BU+MO+TD); this is especially true for the bottom-level. Note that by construction, the top-level forecasts coincide for approaches TDHP and TDFP as well as for COMB BU+MO+TD and COMB Optimal.

Table 5: Percentage improvements over the bottom-up (BU) approach for the first sub-hierarchy (Tankers Earnings).

Approach	**Level**	**Description**	**Forecasting Horizon**				
			1-3	4-6	7-9	10-12	1-12
MO	Top level	Trade	12.72	33.05	59.68	81.61	67.87
	Middle level	Size	9.60	34.26	58.77	78.80	64.91
	Bottom level	Route	11.18	38.46	61.22	79.45	66.23
	Overall		11.07	35.46	59.93	79.92	66.29
TDHP	Top level	Trade	6.12	31.53	64.83	88.77	72.94
	Middle level	Size	16.39	38.26	68.12	88.30	73.62
	Bottom level	Route	17.46	42.36	69.83	88.56	74.47
	Overall		13.86	37.81	67.78	88.54	73.71
TDFP	Top level	Trade	6.12	31.53	64.83	88.77	72.94
	Middle level	Size	-16.08	17.55	56.93	81.77	62.83
	Bottom level	Route	-48.71	2.75	48.52	76.54	53.78
	Overall		-21.69	16.16	56.17	82.18	62.72
COMB BU+TD	Top level	Trade	5.24	24.30	40.52	50.30	42.95
	Middle level	Size	11.93	28.05	41.42	48.09	42.10
	Bottom level	Route	13.48	28.13	39.46	46.91	40.94
	Overall		10.59	26.98	40.44	48.38	41.95
COMB BU+MO+TD	Top level	Trade	7.83	28.18	47.70	62.24	52.46
	Middle level	Size	11.82	31.33	49.27	59.77	51.22
	Bottom level	Route	13.62	33.25	48.35	58.79	50.62
	Overall		11.34	31.12	48.47	60.21	51.39
COMB Optimal	Top level	Trade	7.83	28.18	47.70	62.24	52.46
	Middle level	Size	5.34	25.45	43.69	54.02	45.45
	Bottom level	Route	2.17	4.61	4.46	0.68	2.14
	Overall		4.89	18.55	30.52	38.08	32.21

Similarly to table 5, tables 6 and 7 report the percentage improvements of the various hierarchical approaches over the BU approach for the Drybulk Earnings sub-hierarchy as well as the total hierarchy, which integrates Tankers and Drybulk Earnings. Note that for table 7 we report the performance of two middle levels, referring to trade and size respectivelly.

Also, the percentage improvements are presented for each of the four levels (bottom, top and two middle levels) for each approach.

Tables 6 and 7 provide similar insights to that of table 5. TDHP is overall the best approach for Drybulk, followed by TDFP. TDHP and MO on Trade are the top performers for the total hierarchy. Improvements in the case of Drybulk are lower compared to Tankers and the whole hierarchy, but still significant. At the same time, COMB Optimal provides the least improvements (over BU) across all approaches. In fact, COMB Optimal is the only approach that on average (across all horizons) decreases in performance compared to BU for a single aggregation level (as is the case for the total hierarchy, route level).

Table 6: Percentage improvements over the bottom-up (BU) approach for the second sub-hierarchy (Drybulk Earnings).

Approach	**Level**	**Description**	**Forecasting Horizon**				
			1-3	4-6	7-9	10-12	1-12
MO	Top level	Trade	0.55	10.25	19.61	33.07	19.73
	Middle level	Size	-0.09	8.60	17.99	30.07	17.46
	Bottom level	Route	1.22	10.03	20.25	31.15	18.94
	Overall		0.57	9.62	19.30	31.41	18.70
TDHP	Top level	Trade	5.51	23.39	34.87	52.13	34.53
	Middle level	Size	5.78	23.13	34.70	53.01	34.31
	Bottom level	Route	7.31	24.45	36.24	53.13	35.21
	Overall		6.25	23.67	35.29	52.77	34.69
TDFP	Top level	Trade	5.51	23.39	34.87	52.13	34.53
	Middle level	Size	6.10	21.07	31.46	50.13	31.95
	Bottom level	Route	3.20	19.69	31.19	48.06	30.35
	Overall		4.88	21.34	32.48	50.06	32.22
COMB BU+TD	Top level	Trade	5.10	13.00	17.81	27.49	18.46
	Middle level	Size	9.55	14.08	18.25	28.28	19.52
	Bottom level	Route	10.49	15.05	19.92	28.37	20.34
	Overall		8.57	14.07	18.68	28.06	19.47
COMB BU+MO+TD	Top level	Trade	4.43	12.18	18.41	29.52	19.07
	Middle level	Size	8.34	12.93	18.57	29.47	19.58
	Bottom level	Route	9.19	14.04	20.59	29.97	20.67
	Overall		7.48	13.07	19.22	29.66	19.80
COMB Optimal	Top level	Trade	4.43	12.18	18.41	29.52	19.07
	Middle level	Size	3.93	11.00	17.34	26.85	17.28
	Bottom level	Route	3.44	10.01	15.75	23.93	15.48
	Overall		3.90	11.04	17.14	26.71	17.23

To evaluate the significance in the differences of the performance across the various hierarchical approaches, we perform Diebold and Mariano (1995) (DM) tests. The predictive accuracy of any two hierarchical approaches is tested as follows. We consider the 25 absolute errors (one from each forecast origin) corresponding to the same node (time series) and forecast horizon. We test if an approach is significantly better (or worse) than another and

Table 7: Percentage improvements over the bottom-up (BU) approach for the total hierarchy (Total Shipping Earnings).

Approach	**Level**	**Description**	**Forecasting Horizon**				
			1-3	4-6	7-9	10-12	1-12
MO Trade	Top level	Total	11.61	43.98	74.84	92.71	82.12
	Middle level	Trade	10.87	40.52	71.86	91.57	79.82
	Middle level	Size	19.20	44.43	73.29	91.34	79.81
	Bottom level	Route	20.39	47.90	74.25	91.37	80.12
	Overall		16.11	44.40	73.57	91.74	80.44
MO Size	Top level	Total	21.21	48.29	72.78	89.04	79.90
	Middle level	Trade	16.68	44.70	70.51	88.12	77.83
	Middle level	Size	13.66	42.49	68.71	86.02	74.95
	Bottom level	Route	15.07	45.90	69.98	86.21	75.45
	Overall		16.28	45.24	70.43	87.33	76.96
TDHP	Top level	Total	14.90	45.91	76.10	93.52	83.20
	Middle level	Trade	2.63	36.44	70.36	90.82	78.32
	Middle level	Size	12.40	40.52	71.88	90.60	78.37
	Bottom level	Route	13.88	44.03	72.81	90.62	78.65
	Overall		11.10	41.73	72.75	91.38	79.58
TDFP	Top level	Total	14.90	45.91	76.10	93.52	83.20
	Middle level	Trade	-3.60	40.16	72.79	92.45	80.00
	Middle level	Size	-31.57	21.16	64.42	87.22	70.79
	Bottom level	Route	-59.56	7.72	57.18	83.65	63.89
	Overall		-23.96	27.21	67.15	89.14	74.16
COMB BU+TD	Top level	Total	9.37	31.59	47.62	50.60	46.99
	Middle level	Trade	7.85	28.91	46.19	50.43	46.02
	Middle level	Size	12.71	29.54	42.95	48.80	44.13
	Bottom level	Route	14.29	29.46	41.30	47.72	43.00
	Overall		11.38	29.82	44.36	49.37	44.97
COMB BU+MO+TD	Top level	Total	-23.50	13.70	44.72	58.27	48.92
	Middle level	Trade	-21.11	13.91	43.95	58.19	48.25
	Middle level	Size	-8.33	18.04	43.75	57.53	47.91
	Bottom level	Route	-5.56	22.49	44.60	57.45	48.34
	Overall		-13.50	17.39	44.25	57.85	48.34
COMB Optimal	Top level	Total	13.34	41.01	63.34	72.52	65.94
	Middle level	Trade	11.48	36.88	58.74	67.82	61.02
	Middle level	Size	8.52	31.14	48.40	54.37	48.87
	Bottom level	Route	4.73	7.20	4.47	-4.39	-0.86
	Overall		9.07	27.84	42.38	47.07	42.77

repeat for every combination of node and horizon. Tables 8, 9 and 10 present the results for the two sub-hierarchies and the total hierarchy, respectively. Each entry in these tables reports the percentage of cases where an approach presented in rows is significantly better than an approach presented in columns. For example, table 8 shows that TDHP significantly outperforms BU in 41.67% of the cases, while BU significantly outperforms TDHP in just 5% of the cases.

Table 8: Diebold-Mariano (DM) tests for comparing the predictive accuracy of the various hierarchical approaches for the first sub-hierarchy (Tankers Earnings).

	BU	MO	TDHP	TDFP	COMB BU+TD	COMB BU+MO+TD	COMB Optimal
BU		13.89	5.00	21.67	1.11	2.78	13.89
MO	41.11		0.56	25.00	26.67	23.33	45.56
TDHP	41.67	20.56		38.89	30.00	27.22	50.00
TDFP	26.11	0.00	0.00		15.56	12.22	29.44
COMB BU+TD	48.89	13.33	10.56	25.56		8.89	40.00
COMB BU+MO+TD	46.11	17.22	5.00	23.33	28.89		48.33
COMB Optimal	34.44	1.11	1.11	13.89	8.33	0.00	

Table 9: Diebold-Mariano (DM) tests for comparing the predictive accuracy of the various hierarchical approaches for the second sub-hierarchy (Drybulk Earnings).

	BU	MO	TDHP	TDFP	COMB BU+TD	COMB BU+MO+TD	COMB Optimal
BU		3.70	2.78	0.93	0.00	1.85	0.00
MO	36.11		3.70	1.85	0.93	0.93	11.11
TDHP	51.85	25.93		24.07	39.81	39.81	39.81
TDFP	54.63	15.74	12.96		37.04	35.19	45.37
COMB BU+TD	56.48	12.04	6.48	1.85		6.48	5.56
COMB BU+MO+TD	59.26	12.96	4.63	3.70	1.85		24.07
COMB Optimal	59.26	12.96	8.33	0.93	6.48	8.33	

The results in table 8 corroborate with these in table 5: TDHP is the top-performing approach for the Tankers Earnings. On average, it significantly outperforms the other approaches in almost 35% of the cases, while the opposite is true in less than 4% of the cases. TDHP also performs very well for the Drybulk Earnings (table 9), followed by the TDFP which outperforms TDHP in almost 13% of the cases. It should be noted that for the Tankers and Drybulk Earning sub-hierarchies, the simpler combination approach involving all levels (COMB BU+MO+TD) is significantly better compared to the optimal one in a much greater percentage of cases than the other way around. When the total hierarchy is considered (table 10), MO on Trade significantly outperforms all other approaches in a much greater percentage of cases than TDHP, despite that the two approaches show similar levels of improvement according to table 7.

Table 10: Diebold-Mariano (DM) tests for comparing the predictive accuracy of the various hierarchical approaches for the total hierarchy (Total Shipping Earnings).

	BU	MO Trade	MO Size	TDHP	TDFP	COMB BU+TD	COMB BU+MO+TD	COMB Optimal
BU		5.67	8.33	19.33	19.33	10.67	30.67	15.33
MO Trade	48.33		22.33	34.67	26.33	44.67	63.33	52.67
MO Size	44.00	3.33		24.00	22.00	41.00	63.00	44.67
TDHP	30.33	2.00	14.00		34.33	20.33	27.33	41.00
TDFP	31.00	0.33	4.00	14.00		26.33	31.00	38.00
COMB BU+TD	37.33	10.33	11.00	24.00	23.67		36.33	33.00
COMB BU+MO+TD	24.67	0.67	3.33	1.00	11.33	10.00		20.00
COMB Optimal	32.00	1.67	3.00	15.67	17.33	16.33	27.67	

In balance, we provide some insights on the forecasting performance of the different approaches:

- The bottom-up approach scores the worst performance compared to other hierarchical approaches, due to high volatility in the route-level data.
- Producing forecasts at an aggregate level, which are then disaggregated using historical proportions, significantly increases forecast accuracy. Performance is maximized when forecasts are produced at the trade-level data (top level for the two sub-hierarchies, second to top-level for the total hierarchy).
- A simple combination of the forecasts across all levels performs better than optimal combination.

4.2. Economic Evaluation

The results so far indicate that utilizing the proposed hierarchy structure can significantly improve forecasting performance. In this section, we consider the economic significance of our results and demonstrate their use in practice. Specifically, we consider the case of a shipowner who uses the forecasts generated in 4.1 to optimize his chartering strategy. Our hypothetical shipowner owns a VLCC and has three operational choices: 1) fix the vessel in the spot market and earn the spot voyage rate for the next n months. 2) fix the vessel using a n-period time-charter (TC) contract. 3) switch between the previous two options, depending on the following signal from the generated forecasts: If at time t, the average of the forecasts for the next n months is greater than the corresponding n-period TC rate, the shipowner will use the spot market for the next n months (and the period market otherwise). This strategy is evaluated empirically in the out-of-sample period, from 2013:01 to 2015:12 and for a combination of period contracts of different durations: 3-, 6-, 9- and 12-months. The strategy can be expressed more formally as follows:

$$F_{n,t}^{h_l} \begin{cases} + & fix\ spot\ for\ the\ next\ n\ months\ at\ period\ (t) \\ - & fix\ TC\ for\ the\ next\ n\ months\ at\ period\ (t) \end{cases} \tag{5}$$

where $F_{n,t}^{h_l} = \sum_{i=1}^{n} \hat{y}_{i,t+1}^{h_l}/n - TC_t^n$ is the difference between the average forecasts for the next

n months and the n-period TC rate at t, h_l refers to the hierarchy level $l = \{1, 2, 3, ..., 25\}$ and $n = \{3, 6, 9, 12\}$ is the forecast horizon. The benchmark position in this case refers to operating in the spot market and the gains from this strategy are estimated as the difference between operating in the spot versus operating in the period market, as follows:

$$G^{h_l}_{n,t} = \sum_{i=1}^{n} Spot_i/n - TC^n_t \tag{6}$$

where $G^{h_l}_{n,t}$ is the average gain or loss at time t, for hierarchy level h_l and forecast horizon n and $Spot_t$ are the spot earnings at time t [2].

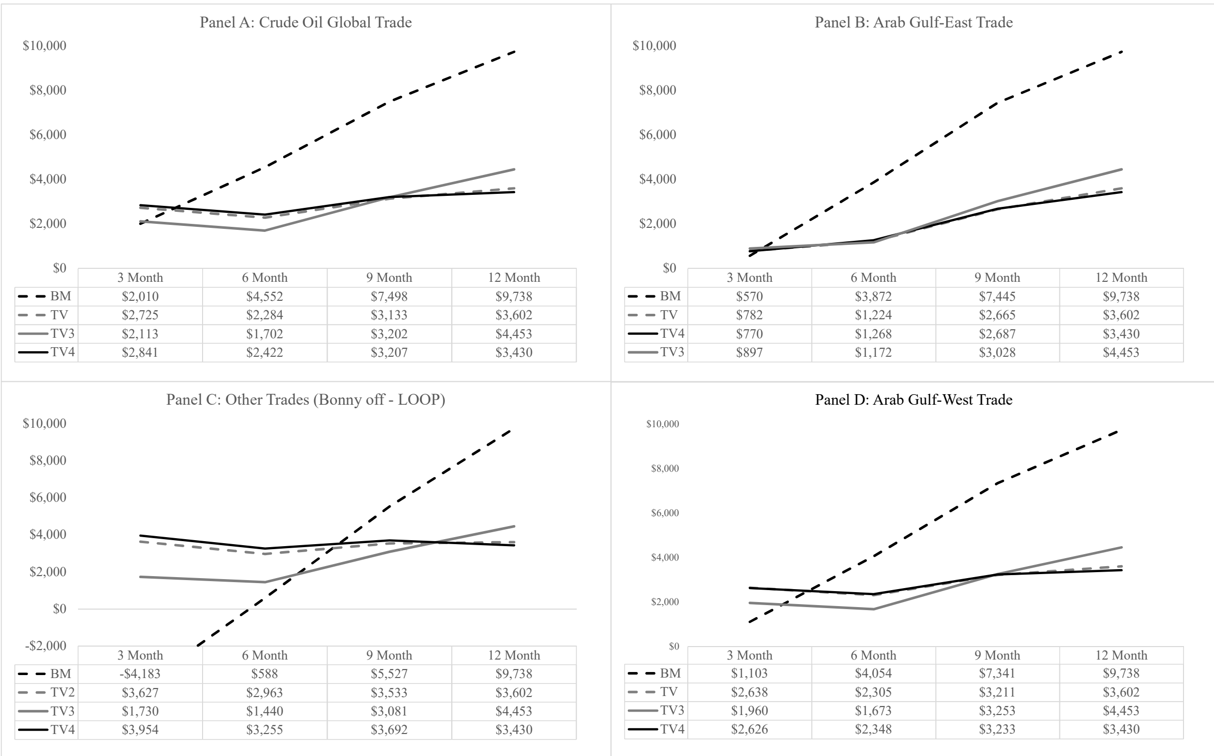

Figure 2: Average daily gains from Active Strategy vs Benchmark

Figure 2 shows average daily gains from applying the above strategy in comparison to the benchmark. In each case, we consider forecasts that are generated at the aggregate

[2] In this exercise, we assume that the owner can use period TC contracts with durations of 3-, 6-, 9- and 12-months. In practice, only rates for 1-year and 3-year contracts are available from Clarkson's for the tanker market. Thus, to derive the required contract maturities, we use linear interpolation between the spot and 1 year TC rates.

level (TV) and the individual route level (TV2 - TV4), as shown in Figure 1. There are four panels that show and report the gain and loss over four different horizons: 3-months, 6-months, 9-months and 12-months. Panel A reports average daily gains for VLCCs trading globally; panel B reports average gains for VLCCs trading in the Arab Gulf - East route; panel C reports average daily gains for VLCCs trading in the Bonny offshore - LOOP route and panel D reports average daily gains for VLCCs trading in the Arab Gulf - West trade. The results show that our proposed strategy generates positive gains for shipowners and outperforms the (spot-only) benchmark over short horizons of three months (six-months in the case of the Bonny - LOOP route), suggesting that a rolling 3-months forecasting strategy is superior for strategic operational planning.

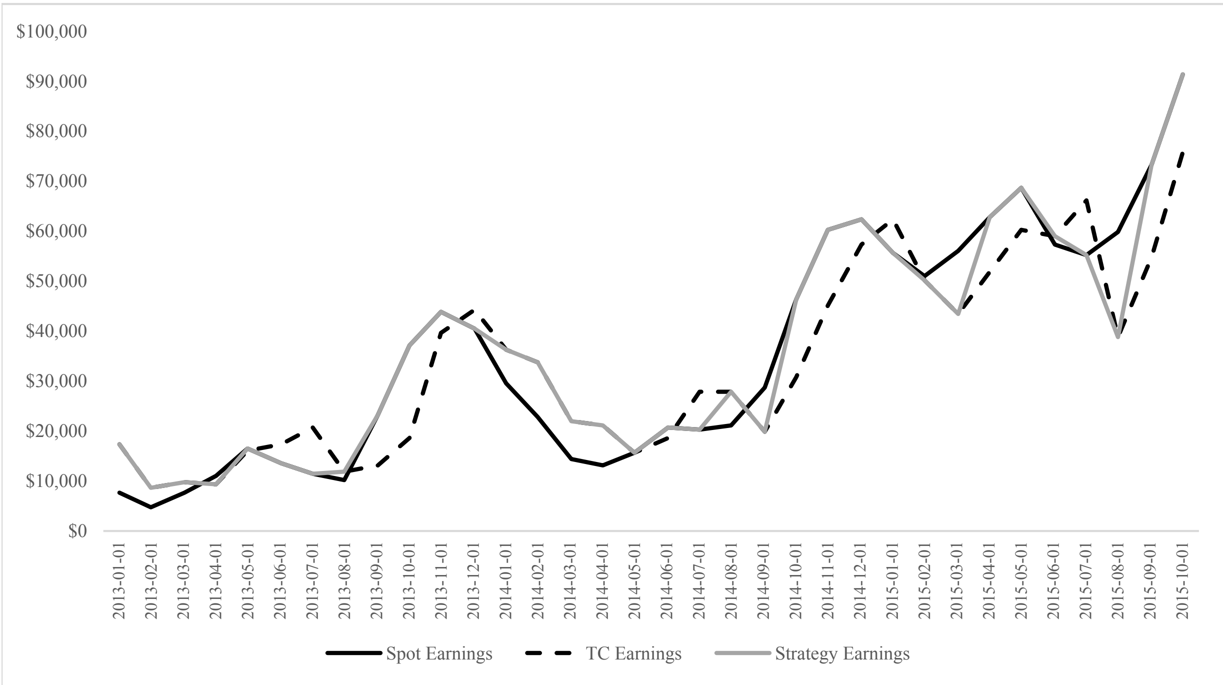

Figure 3: Cumulative Earnings from Different VLCC Chartering Strategies

Figure 3 shows average cumulative gains from the proposed active strategy (spot vs 3-month TC contracts), compared to a spot-only strategy and a strategy where the vessel is chartered continuously using 3-month TC contracts. The average earnings for the three strategies are \$35,700/day, \$33,690/day and \$36,115/day for the spot, 3-month TC and active strategies, respectively. When the earnings mentioned above are scaled to match the days that a VLCC operates, these differences become more noticeable; for example, assuming that a VLCC follows each strategy 350 days in a year, the cumulative performance for each strategy is \$37,484,511, \$35,374,004 and \$37,920,604, respectively. This means that, for the period from 2013 to 2015, a switching strategy between spot and TC operations that is based on a rolling 3 month forecast had the potential to increases earnings for a shipowner operating only spot by more than \$436,093 and by more than \$2,500,000 for a shipowner operating only in the TC market.

5. Conclusions

The forecasting literature in maritime economics has hitherto focused on comparing different forecasting methodologies to improve forecasting accuracy and directional predictability. In this paper, we distinguish between forecasts of freight earnings that are generated at different levels of hierarchy and present a framework that is model independent and lies one conceptual level higher than the econometric models used for forecasting. The advantage of the proposed methodology is that it provides a link between strategic decision making in shipping and the hierarchical structure of freight earnings and matches the appropriate type of strategic decision with their corresponding forecasts aggregated at the appropriate level. As a result, forecasts of freight earnings generated at any hierarchy level are suitable for both macro and micro decision making in the shipping industry.

This paper argues that strategic decision making in shipping can be improved by generating forecasts at different hierarchy levels that better match the type of decision needed. In doing so, this paper contributes to the literature in two ways. First, it associates type of strategic decision making in shipping to the appropriate aggregated level of forecasting. Second, it provides evidence that combination of forecasts from different levels improves strategic decision making in shipping markets. In addition to evidence of accuracy improvements, the approach provides reconciled forecasts and as a result reconciled strategic decision across all hierarchical levels.

One path for future research would be the application of the recently proposed concept of temporal hierarchies (Athanasopoulos et al., 2017) in the shipping industry, where temporal aggregation can capture and model dynamics in the time series by not assuming that monthly is the only frequency where forecasts should be built. Additionally, further research could focus on investigating the link between strategic decision making in different shipping markets and the forecasting hierarchical structure by matching operational, chartering and investment decisions with the appropriate level of freight earnings forecasts.

Appendix A. The S matrix of the hierarchical data

$$S = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix} \tag{A.1}$$

Appendix B. Correlation Matrix

Table B.11: Correlation Matrix for freight earning level prices

	TV1	TV2	TV3	TV4	TS1	TS2	TS3	TA1	TA2	TA3	TA4	DC1	DC2	DC3	DP1	DP2	DP3
TV1	**100%**	98%	97%	**100%**	88%	88%	85%	80%	75%	75%	81%	56%	57%	56%	53%	53%	52%
TV2	98%	**100%**	98%	99%	93%	93%	89%	86%	80%	80%	86%	56%	56%	56%	52%	53%	52%
TV3	97%	98%	**100%**	97%	89%	89%	85%	82%	77%	76%	84%	50%	50%	50%	46%	47%	46%
TV4	**100%**	99%	97%	**100%**	89%	88%	85%	81%	76%	75%	82%	55%	55%	55%	51%	52%	51%
TS1	88%	93%	89%	89%	**100%**	99%	99%	95%	90%	90%	88%	54%	54%	53%	49%	50%	50%
TS2	88%	93%	89%	88%	99%	**100%**	98%	93%	89%	88%	86%	53%	53%	53%	49%	50%	49%
TS3	85%	89%	85%	85%	99%	98%	**100%**	93%	91%	90%	83%	54%	54%	54%	50%	51%	50%
TA1	80%	86%	82%	81%	95%	93%	93%	**100%**	95%	95%	90%	45%	46%	46%	40%	42%	42%
TA2	75%	80%	77%	76%	90%	89%	91%	95%	**100%**	92%	80%	45%	45%	45%	39%	40%	41%
TA3	75%	80%	76%	75%	90%	88%	90%	95%	92%	**100%**	79%	47%	47%	47%	42%	44%	44%
TA4	81%	86%	84%	82%	88%	86%	83%	90%	80%	79%	**100%**	40%	40%	40%	35%	37%	37%
DC1	56%	56%	50%	55%	54%	53%	54%	45%	45%	47%	40%	**100%**	99%	99%	97%	97%	96%
DC2	57%	56%	50%	55%	54%	53%	54%	46%	45%	47%	40%	99%	**100%**	**100%**	97%	97%	97%
DC3	56%	56%	50%	55%	53%	53%	54%	46%	45%	47%	40%	99%	100%	**100%**	96%	97%	96%
DP1	53%	52%	46%	51%	49%	49%	50%	40%	39%	42%	35%	97%	97%	96%	**100%**	99%	97%
DP2	53%	53%	47%	52%	50%	50%	51%	42%	40%	44%	37%	97%	97%	97%	99%	**100%**	99%
DP3	52%	52%	46%	51%	50%	49%	50%	42%	41%	44%	37%	96%	97%	96%	97%	99%	**100%**

Table B.12: Correlation Matrix for freight earning returns

	TV1	TV2	TV3	TV4	TS1	TS2	TS3	TA1	TA2	TA3	TA4	DC1	DC2	DC3	DP1	DP2	DP3
TV1	**100%**	95%	76%	84%	32%	34%	23%	24%	15%	19%	32%	3%	6%	6%	3%	5%	0.2%
TV1	**100%**	95%	76%	84%	32%	34%	23%	24%	15%	19%	32%	3%	6%	6%	3%	5%	0.2%
TV2	95%	**100%**	83%	84%	45%	45%	34%	34%	23%	26%	40%	8%	10%	10%	7%	6%	2%
TV3	76%	83%	**100%**	75%	32%	32%	23%	22%	13%	18%	31%	13.0%	9%	9%	18%	14%	13%
TV4	84%	84%	75%	**100%**	24%	23%	16%	21%	17%	12%	26%	1.0%	-0.3%	0.3%	-1.9%	-1.8%	-3.2%
TS1	32%	45%	32%	24%	**100%**	96%	97%	76%	66%	65%	43%	16%	11%	14%	19%	16%	14%
TS2	34%	45%	32%	23%	96%	**100%**	91%	72%	63%	64%	40%	13%	9%	12%	18%	13%	9%
TS3	23%	34%	23%	16%	97%	91%	**100%**	73%	67%	62%	35%	13%	8%	12%	21%	18%	17%
TA1	24.1%	34.3%	22.2%	20.9%	76.1%	72.1%	73.0%	**100%**	85%	88%	52%	17%	11%	18%	19%	15%	12%
TA2	15%	23%	13%	17%	66%	63%	67%	85%	**100%**	72%	27%	17%	13%	19%	17%	13%	13%
TA3	19%	26%	18%	12%	65%	64%	62%	88%	72%	**100%**	34%	14%	8%	15%	20%	17%	13%
TA4	32%	40%	31%	26%	43%	40%	35%	52%	27%	34%	**100%**	7%	6%	6%	11%	8%	2%
DC1	3%	8%	13%	**100%**	16%	13%	13%	17%	17%	14%	7%	**100%**	92%	88%	48%	55%	49%
DC2	6%	10%	9%	-0.3%	11%	9%	8%	11%	13%	8%	6%	92%	**100%**	90%	44%	54%	50%
DC3	6%	10%	9%	0.3%	14%	12%	12%	18%	19%	15%	6%	88%	90%	**100%**	45%	56%	53%
DP1	3%	7%	18%	-2%	19.%	18%	21%	19%	17%	20%	11%	48%	44%	45%	**100%**	85%	68%
DP2	5%	6%	14%	-2%	16%	13%	18%	15%	13%	17%	8%	55%	54.4%	56.2%	85.2%	**100%**	87.9%
DP3	0.2%	2.4%	13.%	-3%	14%	9%	17%	12%	13%	13%	2%	49%	50%	53%	68%	88%	**100%**

References

Abouarghoub, W., Mariscal, I., Howells, P., 2014. A two-state markov-switching distinctive conditional variance application for tanker freight returns. International Journal of Financial Engineering and Risk Management 1 (3), 239–263.

Adland, R., Cullinane, K., 2006. The non-linear dynamics of spot freight rates in tanker markets. Transportation Research Part E 42, 211–224.

Alizadeh, A., Nomikos, N., 2009. Shipping derivatives and risk management. Palgrave, Macmillan, New York.

Athanasopoulos, G., Ahmed, R. A., Hyndman, R. J., 2009. Hierarchical forecasts for australian domestic tourism. International Journal of Forecasting 25, 146–166.

Athanasopoulos, G., Hyndman, R. J., Kourenztes, N., Petropoulos, F., 2017. Forecasting with temporal hierarchies. European Journal of Operational Research 262 (1), 60–74.

Batchelor, R., Alizadeh, A., Visvikis, I., 2007. Forecasting spot and forward prices in the international freight market. International Journal of Forecasting 23, 101–114.

Cullinane, K., 1992. A short-term adaptive forecasting model for biffex speculation: a box-jenkins approach. Maritime Policy and Management 19, 91–114.

Dangerfield, B. J., Morris, J. S., 1 Oct. 1992. Top-down or bottom-up: Aggregate versus disaggregate extrapolations. International journal of forecasting 8 (2), 233–241.

Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253–263.

Fliedner, G., 1 Sep. 1999. An investigation of aggregate variable time series forecast strategies with specific subaggregate time series statistical correlation. Computers & operations research 26 (10), 1133–1149.

Gross, C. W., Sohl, J. E., 1 May 1990. Disaggregation methods to expedite product line forecasting. Journal of forecasting 9 (3), 233–254.

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., Shang, H. L., 2011. Optimal combination forecasts for hierarchical time series. Computational Statistics & Data Analysis 55 (9), 2579–2589.

Hyndman, R. J., Khandakar, Y., 7 2008. Automatic time series forecasting: The forecast package for R. Journal of Statistical Software 27 (3), 1–22.

Hyndman, R. J., Koehler, A. B., 2006. Another look at measures of forecast accuracy. International Journal of Forecasting 22, 679–688.

Hyndman, R. J., Lee, A. J., Wang, E., 1 May 2016. Fast computation of reconciled forecasts for hierarchical and grouped time series. Computational statistics & data analysis 97, 16–32.

Kavussanos, M. G., Alizadeh, A. H., 2001. Seasonality patterns in dry bulk shipping spot and time charter freight rates. Transport Research Part E 37, 443–467.

Kavussanos, M. G., Dimitrakopoulos, D. N., 2011. Market risk model selection and medium-term risk with limited data: Application to ocean tanker freight markets. International Review of Financial Analysis 20, 258–268.

Kavussanos, M. G., Nomikos, N. K., 1999. The forward pricing function of the shipping freight futures market. Journal of Futures Markets 19 (3), 353–376.

Kavussanos, M. G., Nomikos, N. K., 2003. Price discovery, causality and forecasting in the freight futures market. Review of Derivatives Research 6 (3), 203–230.

Lütkepohl, H., 1984. Forecasting contemporaneously aggregated vector ARMA processes. Journal of business & economic statistics: a publication of the American Statistical Association 2 (3), 201–214.

Papapostolou, N., Nomikos, N. K., Pouliasis, P., Kyriakou, I., 2014. Investor sentiment for real assets: The case of dry bulk shipping market. Review of Finance 18, 1507–1539.

Randers, J., Gluke, U., 2007. Forecasting turning points in shipping freight rates: lessons from 30 years of practical effort. System Dynamics Review 23 (2-3), 253–284.

Zellner, A., Tobias, J., 1 Sep. 2000. A note on aggregation, disaggregation and forecasting performance. Journal of forecasting 19 (5), 457–465.

Zhang, J., Zeng, Q., Zhao, X., 2014. Forecasting spot freight rates based on forward freight agreement and time charter contracts. Applied Economics 46 (29), 747–782.