A Decision Support System for Vessel Speed Decision in Maritime Logistics using Weather Archive Big Data

Lee, H, Ayden, N, Choi, Y, Irani, Z & Lekhavat, S

Original citation:
Lee, H, Ayden, N, Choi, Y, Irani, Z & Lekhavat, S 2017, 'A Decision Support System for Vessel Speed Decision in Maritime Logistics using Weather Archive Big Data' Computers & Operations Research, vol (in press), pp. (in press)
https://dx/.doi.org/10.1016/j.cor.2017.06.005

DOI 10.1016/j.cor.2017.06.005
ISSN 0305-0548
ESSN 1873-765X

Publisher: Elsevier

Computers and Operations Research 000 (2017) 1–13

Contents lists available at ScienceDirect

Computers and Operations Research

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

A decision support system for vessel speed decision in maritime logistics using weather archive big data

Habin Lee[a,*], Nursen Aydin[b], Youngseok Choi[c], Saowanit Lekhavat[a], Zahir Irani[d]

[a] Brunel Business School, Brunel University London, UK
[b] Warwick Business School, University of Warwick, UK
[c] Coventry Business School, Coventry University, UK
[d] Bradford Business School, University of Bradford, UK

ARTICLE INFO

Article history:
Received 14 November 2016
Revised 8 May 2017
Accepted 4 June 2017
Available online xxx

Keywords:
Speed optimization
Sustainable maritime logistics
Weather archive data
Liner shipping
Particle swarm optimization

ABSTRACT

Speed optimization of liner vessels has significant economic and environmental impact for reducing fuel cost and Green House Gas (GHG) emission as the shipping over maritime logistics takes more than 70% of world transportation. While slow steaming is widely used as best practices for liner shipping companies, they are also under the pressure to maintain service level agreement (SLA) with their cargo clients. Thus, deciding optimal speed that minimizes fuel consumption while maintaining SLA is managerial decision problem. Studies in the literature use theoretical fuel consumption functions in their speed optimization models but these functions have limitations due to weather conditions in voyages. This paper uses weather archive data to estimate the real fuel consumption function for speed optimization problems. In particular, Copernicus data set is used as the source of big data and data mining technique is applied to identify the impact of weather conditions based on a given voyage route. Particle swarm optimization, a metaheuristic optimization method, is applied to find Pareto optimal solutions that minimize fuel consumption and maximize SLA. The usefulness of the proposed approach is verified through the real data obtained from a liner company and real world implications are discussed.

1. Introduction

Speed optimization in liner shipping has significant economic and environmental impact for reducing fuel cost and Green House Gas (GHG) emission as the shipping over maritime logistics takes more than 70% of world transportation (UNCTAD, 2010; Psaraftis and Kontovas, 2013). While slow steaming is widely used as best practices for liner shipping companies, they are also under the pressure to maintain service level agreement (SLA) with their cargo clients (Lee et al., 2015; Parthibaraj et al., 2016). Thus, deciding optimal sailing speed which minimizes fuel consumption while maintaining SLA is an important managerial decision problem for liner companies.

Sailing speed decision mainly depends on the vessel schedule and it is a challenging problem due to the uncertainties imposed in maritime logistics such as stochastic port times and weather conditions. Port time uncertainty significantly affects the time that vessels spend at ports in anchorage, berthing, unberthing or drifting status. Increased port congestion and delays can negatively affect service level of shipping lines to their customers and put pressure on schedule reliability and might incur logistics costs to the customer (Notteboom, 2006). On the other hand, weather conditions including current and wind affect journey times and the routing decisions (Kontovas, 2014).

The majority the literature work on the speed optimization problem based on a theoretical fuel consumption function. For example, Fagerholt et al., (2010) and Yao et al., (2012) propose a fuel consumption function which is based on the empirical data from a shipping company. However, these functions do not reflect the actual fuel consumption of vessels that are affected by weather conditions. In reality, certain routes may encounter harsher weather conditions than others and speed optimization needs to consider such different voyage environments.

In Fig. 1, we compare the theoretical fuel consumption based on the empirical model proposed by Yao et al., (2012) with the historical fuel consumption data obtained from a liner shipping company. The data belongs to a Turkish liner service with 10 ports-of-call operated in the Mediterranean region. 15 voyages performed by the same vessel of this service in 2013 are analyzed. Fig. 1 illustrates the change in total fuel consumption with respect to time in sea in terms of day. Although fuel consumption mainly depends

* Corresponding author.
E-mail address: Habin.Lee@brunel.ac.uk (H. Lee).

http://dx.doi.org/10.1016/j.cor.2017.06.005

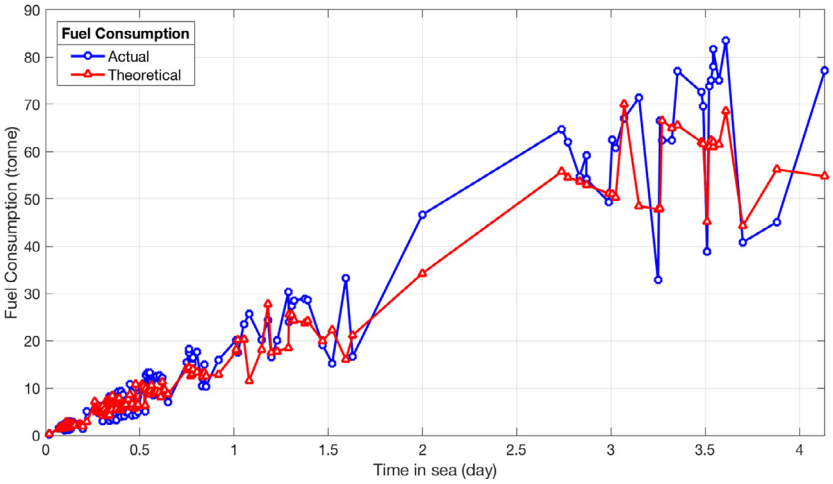

Fig. 1. The actual and theoretical fuel consumption levels with respect to time in sea.

on the vessel sailing speed, there are other affecting factors such as the weather conditions (winds, currents, etc.). The differences between the estimated consumption and the actual one illustrate the effect of these factors. In particular, fuel consumption difference becomes larger when the time in sea is longer. In this study, we focus on the speed optimization problem by considering the effect of weather conditions on fuel consumption.

Different vessel routes have different weather conditions hence, it is difficult to have unified weather adjustment functions to correct the differences between actual and theoretical fuel consumption. The impact of current and winds on fuel consumption varies depending on the routes due to the geographical characteristics. Thus, it is more realistic to identify different impacts of weather conditions in different routes based on historical voyage data and weather data. The analysis of weather archive big data which is publicly available on the Internet in comparison with actual fuel consumption data from liner companies provides an opportunity to measure different impacts of weather conditions on fuel consumption of vessels.

Despite the opportunity, using weather archive big data in vessel speed optimization requires overcoming following challenges. Firstly, weather archive data provides an opportunity to apply big data analytics to estimate the degree of the impacts of weather conditions on fuel consumption of vessels in different routes based on its huge volume of historical data. However, most of such archive data is not easy to use due to the format, volume, and velocity of data. Secondly, the relationship between weather conditions and fuel consumption is different for different routes and difficult to model as a single mathematical formula. In this study, we apply a data mining technique to explore such non-linear relationships based on historical weather and voyage big-data from a liner company.

This paper proposes a decision support system (DSS) that uses weather archive big data in vessel speed optimization overcoming above challenges. To the best of our knowledge, the impact of weather conditions on fuel consumption in liner shipping has not been explicitly considered in the literature. This paper aims to fill this research gap. In particular, we focus on speed optimization problem in liner shipping by considering the weather impact. The speed decision affects the transit time between ports, and in turn, affects the service level. Hence, we also study the trade-off between minimizing fuel cost and maximizing service level. A particle swarm optimization (PSO) technique based solver is proposed to solve this multi-objective problem. Based on a real shipping data, we analyze the impact of weather conditions on the fuel consumption.

The remainder of the paper is organized as follows. Section 2 reviews related studies with regard to speed optimization in maritime logistics. Section 3 then formulates the target problem as a multi-objective optimization problem. The details of the decision support system are given in Section 4. In Section 5, experiment results based on data obtained from a real liner shipping company are provided to verify the usefulness of the proposed decision support system. Finally, Section 6 concludes the paper.

2. Literature review

Optimization techniques have been widely applied to maritime operations including ship routing and scheduling, fleet management, disruption handling, and bunkering. Christiansen et al., (2013) provide a survey of studies on ship routing and scheduling. The literature on bunker optimization methods in maritime shipping has been summarized by Wang et al., (2013). Tran and Haasis (2015) review the literature on container liner shipping with respect to container routing, fleet management and network design. Recently, Mansouri et al., (2015) have reviewed existing studies in maritime operations from sustainability and decision support perspective.

Speed optimization is one of the important problems for sustainable maritime operations as the CO_2 emission is directly affected by the fuel consumption which is determined by vessel speeds. Early studies on the speed optimization problem assume deterministic port times and strict time windows (Fagerholt et al., 2010; Hvattum et al., 2013; Norstad et al., 2011; Andersson et al., 2015). The proposed models restrict vessels to arrive at the contracted time windows to meet 100% service level agreement. However, in reality such assumption is too strong and it is reported that only 55% to 89% vessels arrive on time at ports (Drewry, 2016). Port and travel times can be highly variable due to congestion, handling and weather conditions (Notteboom, 2006). Thus, recent studies in this field extend the speed optimization problem by considering uncertainties at ports and voyage routes (Qi and Song, 2012; Aydin et al., 2017). Qi and Song (2012) propose a vessel scheduling model to minimize the total fuel cost by considering uncertain port times and frequency requirements. In their formulation, they relax the port time window constraint and allow vessels to arrive at any time. On the other hand, Aydin et al., (2017) extend the problem by considering the time windows and bunkering decisions.

The speed optimization models generally assume that fuel consumption solely depends on the vessel speed (Psaraftis and Kontovas, 2013). Yao et al., (2012) propose optimal bunker manage-

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005

ment strategy by solving an integrated mathematical model that includes decision variables with regard to bunkering port selection, bunkering amount decision, and vessel speeds between ports. They discuss that different fuel consumption functions need to be considered for different vessel size based on empirical data obtained from Asia-Europe and Asia-Pacific services. Wang and Meng (2012) work on deterministic speed optimization problem for container routing problem. By using historical data, they analyze the relation between sailing speed and fuel consumption. The authors discuss that the fuel consumption depends on voyage legs as weather conditions can be different at different legs. In this study, we focus on the speed optimization problem by considering the effect of weather conditions.

While above studies are aiming at developing optimization models, the application of decision support systems in maritime logistics are rare compared to other industries in the literature and this is attributed to unique culture of maritime industry (Mansouri et al., 2015). In practice, commercial software solutions (for example, SPOS[1] and NETPAS[2]) are being adopted by liner services. The supporting functionalities of these systems are limited to provide weather and voyage data management rather than automatically finding the optimal sailing speeds by considering the environmental variables including weather and port conditions. A DSS proposed by Besikci et al., (2016) is one of few efforts to support vessel speed optimization problem using various factors including weather condition, trim, cargo quantity, and vessel speeds. Artificial Neural Network (ANN) is applied to learn the impacts of those factors on fuel consumption based on historical data obtained via Noon Data reports which are recorded by the crews of vessels. However, the rules identified by the DSS is applicable to only one vessel and they cannot be applied to other vessels that have different specifications and providing services in different routes. As highlighted by Mansouri et al., (2015), the majority of the previous studies in sustainable ship scheduling problem pay attention to only mathematical modelling and algorithms to solve the problem. Existing literature on DSS for vessel scheduling is relatively scarce and therefore this paper seeks to fill this gap in the literature.

Kim and Lee (1997) propose one of the pioneer studies for the use of optimization-based DSS for scheduling vessels. The proposed DSS assigns bulk cargoes to a schedule in tramp shipping. LINDO optimizer is used as a tool in scheduling process in order to maximize the profit obtained from the transportation of cargoes. Another similar bulk cargo scheduling problem in tramp shipping is proposed by Bausch et al., (1998). The authors aim to assign cargoes into the vessel schedules so that all loads are transported at a minimum cost and satisfy all constraints such as time windows and compatibility between ports and vessels. The output of this optimization process is presented as a schedule on a spreadsheet for the users. Since the study by Bausch et al., (1998), there has been a lack of literature related to the use of DSS in vessel scheduling problem. Later, Fagerholt (2004) argues that one of the main reasons why managers in marine shipping are not willing to use DSS is because of its limitations to consider all of the constraints in the scheduling process. To address this problem in the industry, a DSS called TurboRouter was introduced for vessel fleet scheduling. Fagerholt and Lindstad (2007) extend TurboRouter to meet all the requirements for vessel scheduling problem in industrial and tramp shipping. Time windows, vessel capacities, compatibility between port and vessel, bunker consumption rate, bunkering port calls are taken into account for planning the vessels to arrive at port within specific time period and with the maximum profit. As a result, the decision maker can easily see the schedule through user interface. TurboRouter also receives satellite positions from ships in real time and computes the estimated arrival times to given ports. Apart from industrial and tramp shipping, Lam (2010) focused on designing DSS for scheduling liner shipping problem. The proposed integrated approach first selects the ports of call and then schedules vessels with respect to given time windows and finally analyzes the financial factors. In scheduling process, a planner can edit the service route manually and then the system updates the optimal schedule automatically.

Due to the recent environmental concerns in maritime shipping, later studies on DSS for vessel scheduling have focused on minimizing CO_2 emissions. Ballou et al., (2008) presented a DSS called Voyage and Vessel Optimization Solutions (VVOS) in order to schedule vessels to reach ports of call with minimum CO_2 emissions within a given time window. The system makes ship scheduling decisions based on the wind, wave and current data. VVOS is considered to be user friendly as it is flexible for the user to choose whether they would like to use optimization module. Similarly, Windeck and Stadtler (2011) also focused on developing DSS for network design problem to minimize cost and CO_2 emission by considering weather factors.

While studies on big data are common in computer science and information systems (Agarwal and Dhar, 2014 for example), the application of big data analytics are gaining popularity in operations research field recently. Choi et al., (2017) proposes a novel method to integrate a qualitative decision model with open big data available on the Internet to support public procurement processes. Fang et al., (2016) applies random forecast regression to big data obtained from insurance companies to forecast the profitability of insurance customers. Song and Wang (2016) find that enterprises that are participating to global value chain tend to have the higher green technology level via regression analysis on difference-in-difference panel data on Chinese enterprises. Psaraftis et al., (2016) review the literature on dynamic vehicle routing problem. They discuss the importance of using big data in vehicle routing problems to enhance decision making. They also point out that the literature should focus on how to make use of big data. While these studies are processing large amount of data, the nature and size of the data used in this paper is more complex and huge. Weather archive big data in this paper contains vast amount of observation data on weather in different points of Sea. In addition, the format of the archive data is usually not directly accessible by general purpose programming tools therefore pre-processing is required. This paper shows a systematic method to process the archive data to build weather information for chosen vessel routes from the vast amount of archive.

3. Problem formulation

The objective of the problem is to minimize fuel consumption for a vessel that travels through a predefined route while maximizing the total service level. Since these two objectives are conflicting, we have a multi-objective optimization problem. Decision makers are interested in learning the trade-off relationship between vessel operation cost and service level for a given liner route.

We use the problem structure defined by Aydin et al., (2017). They use a single objective function to minimize the total operation cost by synthesizing the fuel cost with penalty cost incurred from missing required service level. However, in reality normalizing the scale of penalty cost with the fuel cost is very difficult. Therefore, finding Pareto optimal solutions that show the trade-off relationships between two components can make more sense for decision makers. Thus, we define a bi-criteria model to solve the multi-objective optimization problem.

[1] Ship performance optimization system, http://www.meteogroup.com/
[2] http://www.netpas.net/

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005

We consider a vessel providing a liner shipping service over a route that is a predefined sequence of ports of call denoted by set $N = \{0, 1, 2, \ldots, n\}$. Port 0 denote the starting node of the route. Leg i represents a trip from port $(i-1)$ to port i. We assume that the vessel has a contracted time window for each port and port service can start within the specified time window. If the vessel arrives earlier than the contracted time, then it needs to wait until the starting point of the time window. We assume that the vessel consumes a fixed amount of fuel per hour during waiting and service time at each port. In particular, vessels usually use more expensive fuel when they are waiting at ports, therefore we distinguish the waiting cost from sailing cost. We also assume that a vessel has a minimum and maximum speed limit and operates within the capacity. We will use the following notation in the paper to explain the structure of the optimization model.

N: set of ports
μ_i: average service time in hour at port i
t_i^a: arrival time of vessel at port i
t_i^d: departure time of vessel at port i
$[\alpha_i,\beta_i]$: contracted time window at port i
d_i: length of leg i in nautical mile
ϕ: fuel cost per hour during waiting and service time at a port
r_s: price of fuel per ton consumed during sailing
r_p: price of fuel per ton consumed at ports
v_i: average speed at leg i (nautical mile per hour), which is limited by $[v_{min}, v_{max}]$

The vessel operational cost consists of two major components: sailing cost and port cost. Sailing cost corresponds to the fuel cost incurred during sailing. Yao et al., (2012) present an empirical model to reflect the relation between fuel consumption rate and the sailing speed by considering the size of the vessels. The estimated fuel consumption rate is given by $k_1 v_i^3 + k_2$, where k_1 and k_1 are constants and their values depend on the size of the vessel. Multiplication of fuel consumption rate by the transit time between ports yields the total fuel consumption. We extend Yao et al., (2012)'s fuel consumption model by considering the weather factor at each leg. The fuel consumption function for leg i is represented by $f_i(g(v_i), w_i)$ where $g(v_i) = (d_i/24v_i)k_1 v_i^3 + k_2$ and w_i denote the weather factor at leg i. The fuel consumption function is convex and increasing with v_i and adjusted by the weather factor (w_i) at let i.

Port cost also corresponds to a fuel cost which is incurred while a vessel waits for berthing or receives a service from a port. We assume that port cost is proportional to the entire time spent at the port including waiting time and service time. If we let κ be the average amount of fuel (tons) consumed per hour, then the fuel cost per hour (ϕ) at a port is given by $\phi = r_p\kappa$. Finally, the total vessel fuel cost is defined as in Eq. (1).

$$C_f = \sum_{i=1}^{n} \left(r_s f_i(g(v_i), w_i) + \varphi\left(t_i^d - t_i^a\right)\right) + \varphi\mu_0 \tag{1}$$

Given the vessel speed v_i and average service time τ_i at port i, the arrival and departure times at each port are defined by the following system dynamics equations:

$$\begin{aligned} t_i^a &= t_{i-1}^d + d_i/v_i \\ t_i^d &= \max\left\{t_i^a, \alpha_i\right\} + \mu_i,\ i = 1, \ldots, n. \end{aligned} \tag{2}$$

Since port 0 denote the starting node of the route, we assume that $t_0^a = 0$ and $t_0^d = \mu_0$.

Our second objective is to maximize service level. When a vessel arrives at the port before or within the time window, such port is satisfied 100%. However, the service level starts to decrease if the vessel arrives later than the contracted time window. On time delivery of the containers is very important for liner shipping companies since delayed cargo may result in high cost by customers. Stepwise function is suitable for representing the increasing margin of delay effect, where ports may tolerate a small delay but a large delay will result in deviation from the planned schedule and will have a large negative impact on the service level. The service level at port i is computed as follows;

$$s_i\left(t_i^a\right) = \begin{cases} 1 & \text{if } t_i^a \le x_{i1} \\ h_i(x_{i2}) & \text{if } x_{i1} < t_i^a \le x_{i2} \\ \cdots & \\ h_i\left(x_{im_i}\right) & \text{if } x_{im_i-1} < t_i^a \le x_{im_i} \end{cases} \tag{3}$$

where $x_{i1} = \beta_i$ and $x_{i1} < x_{i2} < \cdots < x_{im_i}$. Function $h_i(x_{i\_})$ returns a service level value according to the time points $x_{i1}, x_{i2}, \cdots, x_{im_i}$. The first time point corresponds to the latest start time of the service, β_i. Therefore, if the vessel arrives before the first time point, the port is satisfied 100%. Through conversation with a major liner company, it was realized that missing contracted time windows at busy ports results in higher delay than the idle ports due to the difficulty of finding alternative service time slots. Therefore, we assume that the function $h_i(x_{i\_})$ can take different form for different ports. For example, a function for a busy port may return lower service level compared to idle ports for the same amount of delay. The multi-objective speed optimization problem is formulated as:

$$minimize \sum_{i=1}^{n} \left(r_s f_i(g(v_i), w_i) + \varphi\left(t_i^d - t_i^a\right)\right) + \varphi\mu_0 \tag{4}$$

$$maximize \sum_{i=0}^{n} \frac{1}{n+1} s_i\left(t_i^a\right) \tag{5}$$

$$subject\ to\ t_i^a = t_{i-1}^d + d_i/v_i,\ i = 1, \ldots, n, \tag{6}$$

$$t_i^d = \max\left\{t_i^a, \alpha_i\right\} + \mu_i,\ i = 1, \ldots, n, \tag{7}$$

$$v_{min} \le v_i \le v_{max},\ i = 1, \ldots, n, \tag{8}$$

where $t_0^a = 0$ and $t_0^d = \mu_0$. Constraints (6) and (7) correspond to the system dynamics equations for arrival and departure time. Constraints (8) ensure that the vessel sailing speed is within the lower and upper limits in all legs. While objective function (4) minimizes the total fuel cost incurred during sailing and service at ports, objective function (5) maximizes the average service level at all ports. These objectives conflict with each other, i.e., increasing one objective deteriorates the other.

4. Decision support system for big data based speed optimization

The overall architecture of the decision support system is shown in Fig. 2.

The DSS consists of four major components: user interface, weather archive data parser, weather impact miner, and PSO solver. User interface is a web-based system for effective and platform independent interaction with end users. Weather archive data parser has interface with weather archive data source and it converts original archive data into data format that can be interpreted by other components of the DSS. Weather impact minor aims at finding rules with respect to weather impact, w_i, on fuel consumption function $f_i(g(v_i),w_i)$ for each leg. PSO solver use the weather impact data to generate Pareto optimal solutions that show trade-off relationships between fuel consumption and service level for speed solutions.

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005

Table 1
Data types in the NetCDF file.

sowavenu	Wave number (number of waves that exist for a specified distance)
somestdy	Meridional-stokes drift velocity at surface (m/s) (in the north – south direction)
vomecrty	Meridional velocity (current in the north – south direction)
sozastdx	Zonal-stokes drift velocity at surface (m/s) (in the west – east direction)
vozocrtx	Zonal velocity (current in the west – east direction)
wind_speed	Wind speed (m/s)
wind_to_dir	Wind direction (degree)

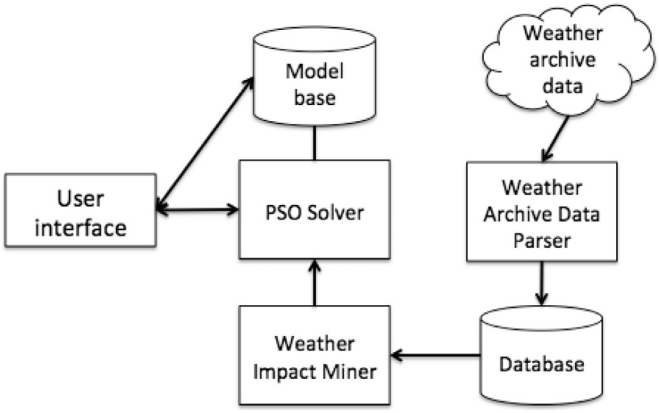

Fig. 2. The architecture of the decision support system for multi-objective speed optimization.

4.1. Weather archive data parser

In this section, our aim is to estimate the effect of the sea state on the vessel creating either drag or forward push depending on the direction. We use real time marine condition data provided by Copernicus Maritime Environment Monitoring Service (Copernicus, 2016). We analyze the marine data for 3 years (obtained from 2012 to 2014) for the Mediterranean Sea. We follow several steps before applying data analytics techniques. Next, we explain our data processing steps.

The data is stored as a segmented (i.e. quarterly data packages) NetCDF (network common data form) file. NetCDF file is a set of software libraries that support a machine-independent format to represent scientific data (Rew and Davis, 1990). We access the data by using Matlab 2016a programming language. This file includes several data types including temperature, salinity, drift velocity, current and wind speed. Fig. 3 illustrates the content of a quarterly data package and Table 1 presents the explanation of the terms in this file.

Each data point for a given latitude and longitude presents 24-h mean value of the corresponding data type. Meridional and zonal directions correspond to the north-south and west-east orientations, respectively. In this section, we define how to extract current data along the vessel route. Other data types (e.g. wind speed and wind direction) in NetCDF file can also be extracted in the same way. Fig. 4 presents the average current for a given day. The colour differentiates the direction and magnitude of the current in the north-south and west-east directions.

To compute the net effect of the current on the vessel, both the vector (direction and magnitude) of travel and the vector of the current should be considered. By making use of the NetCDF file, the vector of the current can be computed along the vessel route. The coordinates of the vessel route are provided by the liner shipping company. The traversed grids on each day can be identified by the route coordinates, vessel sailing speed and the port service time. The distance between two geographical coordinates is computed by using Haversine formula (Sinnott, 1984).

The coordinates of the vessel route can include either the route diversion points or ports. At the diversion points the vessel changes the direction intravenously whereas at ports it waits for berthing and service. Considering this information, the travelled route for each day is computed as illustrated in Fig. 5. Interim points marking the end of the day on the route are also captured and computed. The traversed grids along the path are then determined by using the Bresenham's line algorithm (Bresenham, 1965). In this analysis, the sailing time of the vessel is computed by only considering the vessel speed. For more realistic approximation, the effect of the current can be recalculated iteratively. As it is seen from the marked green grids travelled on the first segment by the vessel (Fig. 5), each grid is traversed in different durations. In order the calculate the average net effect of current, the weighted average of the resultant current vector is computed as follows;

$$\sum_{j=1}^{N} \frac{y_j}{D_i}\left(c_j\hat{u} + c_j\hat{v}\right) \tag{9}$$

where D_i is the distance travelled at segment i (between two coordinates), y_j is the distance travelled in grid j, N is the number of grids travelled in segment i and $(c_j\hat{u},\ c_j\hat{v})$ denote the vector components of the current in zonal and meridional directions, respectively.

Fig. 6 illustrates the variation in the magnitude of current along the vessel route for different days. The colour map in each graph corresponds the resultant magnitude of the current velocity in m/s. Traveled route for different days is presented by coloured lines on the map. As it is seen in Fig. 6, significant changes in current are observed at each different day.

4.2. Weather impact miner

The role of weather impact miner is to identify important factors that can affect the fuel consumption of vessels. We use the data passed from Weather Archive Data Parser, which provides the weather data for given routes including current vectors and wind speed/direction information. By combining weather data extracted from Copernicus dataset and service history of the liner shipping company, we can identify the important weather factor. We analyze the fuel consumption of the same route for specific dates by using the extracted weather data. The service data or voyage abstract covers the information including arrival and departure ports, running distance between two ports, average speed, arrival and departure time, fuel consumption at sea, route coordinates, etc. Based on this data, weather impact miner considers the average fuel consumption (total fuel consumption between ports divided by the distance) as a dependent variable to find important factors that may affect to fuel consumption.

Weather impact miner mainly investigates the impact of the wind and current data. The direction and magnitude of wind can affect the fuel efficiency with respect to the date and time. To identify the impact of wind on the fuel consumption, weather impact miner extracts the rule prioritizing the direction and magnitude combination. Since the wind magnitude may not show a linear impact on the fuel consumption due to its direction, weather

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005

Str	DimID	AttArray	Type	VSize	Begin	Data
'time'	0	1x8 struct	4	476	3752	119x1 int32
'lon '	2	1x7 struct	5	2708	4228	677x1 single
'sowavenu'	[0;1;2]	1x7 struct	5	81529756	6936	119x253x67...
'somestdy'	[0;1;2]	1x7 struct	5	81529756	81536692	119x253x67...
'vomecrty'	[0;3;1;2]	1x7 struct	5	81529756	163066448	4-D single
'sozostdx'	[0;1;2]	1x7 struct	5	81529756	244596204	119x253x67...
'depth '	3	1x9 struct	5	4	326125960	1.4721
'lat '	1	1x7 struct	5	1012	326125964	253x1 single
'vozocrtx'	[0;3;1;2]	1x7 struct	5	81529756	326126976	4-D single

(a) Current data

Str	DimID	AttArray	Type	VSize	Begin	Data
'lon '	3	1x9 struct	4	1364	7020	341x1 int32
'measurement_time'	[0;1;2;3]	1x7 struct	4	175956	8384	4-D int32
'eastward_wind '	[1;2;3]	1x9 struct	3	87980	272320	1x129x341 int16
'wind_speed '	[0;1;2;3]	1x9 struct	3	87980	448280	4-D int16
'wind_to_dir '	[0;1;2;3]	1x8 struct	3	87980	536260	4-D int16
'northward_wind '	[1;2;3]	1x9 struct	3	87980	1064148	1x129x341 int16
'lat '	2	1x9 struct	4	516	1152128	129x1 int32

(b) Wind data

Fig. 3. Data types captured from the NetCDF data in Matlab.

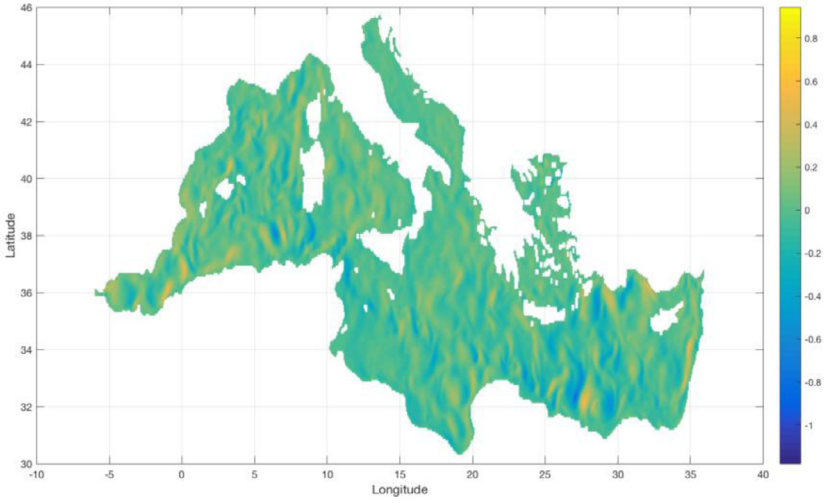

Fig. 4. Illustration of current in Mediterranean for a given day.

impact miner collects all data for given routes and then compares multiple voyage histories by controlling the wind directions. Then, the prioritized wind direction can be obtained and weather impact miner calculates the relative weight of each wind direction and magnitude combinations and returns the rule accordingly using statistical rule mining. Regarding the wind data, we have wind direction and wind force variables, which are both categorical. The direction and magnitude have 5 and 7 levels, respectively. The wind force is coded as ordinal scale (1 to 7) where the highest scale corresponds to the strongest wind force. Wind direction code indicates relative direction of wind with respect to the sailing direction as shown in Fig. 7.

In case we do not have enough voyage data for a given route, bootstrapping method can be applied to derive meaningful statistical rule mining results. To develop the preference rule for reducing fuel consumption, based on the fuel consumption record, we define the average fuel consumption rate function ($N^2 \rightarrow R$) between ports A and B with respect to wind force i and wind direction j

$$AFR_{A-B}(i, j),\ 1 \le i \le 7,\ 1 \le j \le 5.$$

This function returns the mean value of fuel consumption between ports A and B for given period. Using this function, weather impact miner conducts pairwise comparison to derive the preference rule, which can be defined as

$$(i, j) \rightarrow (k, l) \text{ if } AFR_{A-B}(i, j) < AFR_{A-B}(k, l)$$

where $(i, j) \rightarrow (k, l)$ means (i, j) is preferred over (k, l). This pairwise comparison returns all preferences between 35 possible wind direction and force combination. Using this result and the overall mean of fuel consumption between two ports (AFR_{A-B}), the relative weights can be calculated and assigned to all combination. If the $AFR_{A-B}(i, j)$ do not show the statistically significant difference comparing to overall mean of fuel consumption, the (i, j) combination between ports A and B will have no weight. Otherwise, we will use Ratio of Mean values (RoM) for the weight of (i, j) combination, by dividing $AFR_{A-B}(i, j)$ with AFR_{A-B}.

The impact of current can vary from the geographical locations of the routes. To estimate the accurate impact of current data, weather impact miner requests the current vector data for all the

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005

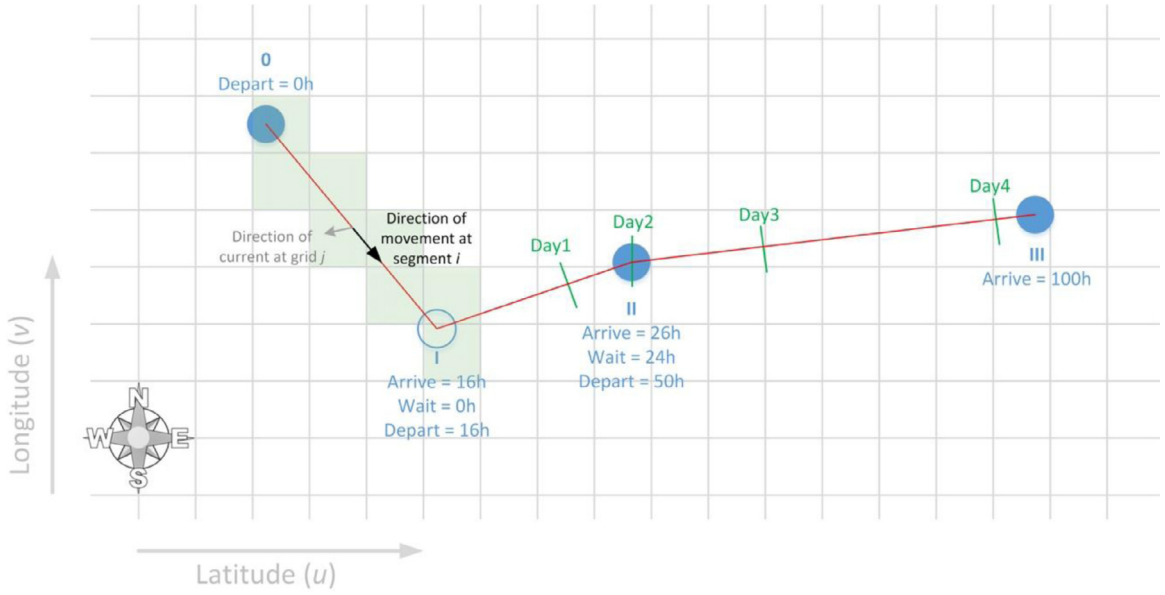

Fig. 5. Illustration of vessel route and daily traversed grids.

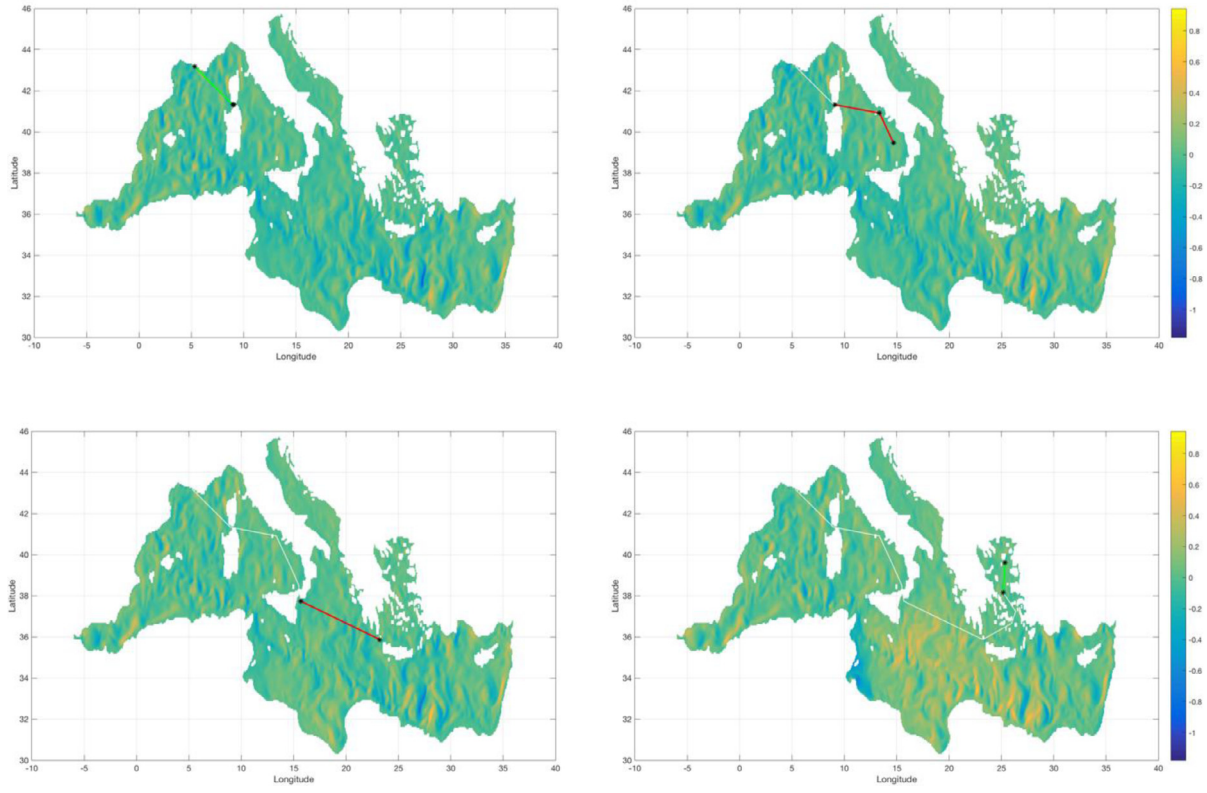

Fig. 6. Change in the current along the vessel route.

past voyage data to weather big data parser. Using the date and route information such as trajectory coordinates and route schedule (i.e., time and date on each coordinates), weather impact miner can match the current information of each route in voyages and then, create historical records regarding to current. By controlling wind effect, weather impact miner estimates the net impact of current on the fuel consumption using regression analysis and identifies which voyages are easily affected by the current magnitude in terms of fuel consumption. The standardized coefficient with statistical significance will be the current weight for given route on the sailing period and 1 will be assigned to the route which was not affected by current.

4.3. Particle swarm optimization solver

Particle swarm optimization (PSO) is one of the successful metaheuristic algorithms which has been applied to many real-world applications (Ai and Kachitvichyanukul, 2009a; Ai and Kachitvichyanukul, 2009b). It is a population-based search method developed by Kennedy and Eberhart (1995). It mimics the social behaviour of a group of birds or a school of fish in foraging their

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005

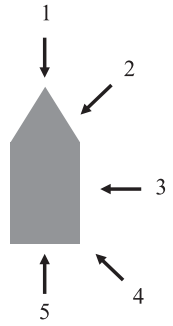

Fig. 7. Wind direction coding scheme.

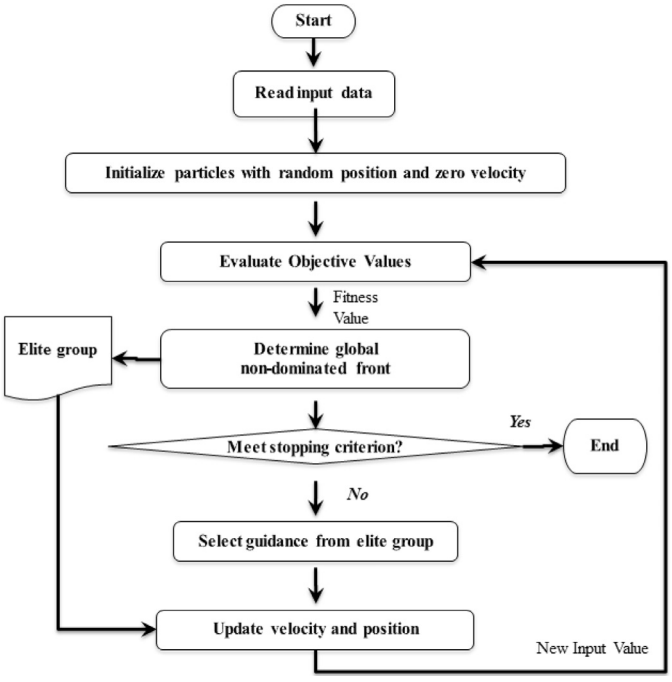

Fig. 8. MOPSO framework.

food. The searching algorithm is motivated by the movements of the individuals or particles in the swarm. There are L particles in a swarm and each particle is characterized by its current position, velocity, personal best position and fitness value. While the current position represents the current location of each particle, the velocity specifies the direction that the particle moves. At each PSO iteration, every particle move to new position according to its velocity. The position of each particle represents a solution of the problem. The personal best position can be considered as a metaphor of cognitive learning of each particle. It keeps the best location of a particle which gives the best objective function value compared to its previous positions. In addition, each particle can also learn from other particles in the swarm. Thus, the best location in the swarm can be found as well. This location is called as the global best location.

The PSO proposed by Kennedy and Eberhart (1995) considers only single objective and hence, it cannot be directly applied to the multi-objective problems. In this study, we utilize multi-objective PSO (MOPSO) framework presented in Nguyen and Kachitvichyanukul (2010). We adapt one of the proposed movement strategy and conduct experiment to fine-tune the PSO parameters for suitability to our model. The MOPSO framework is illustrated in Fig. 8 (Nguyen and Kachitvichyanukul, 2010).

It should be noted that we apply the direct encoding scheme to represent the decision variables (average vessel speed v_i at leg i) in the position of the particle in the swarm. The position of particle l is represented by a H-dimensional vector $\theta_{lh}(\tau)$ where $l=1, \ldots, L$, h is the dimension of the vector and τ denote the iteration number. The corresponding velocity is given by $\omega_{lh}(\tau)$ and the personal best position of particle l is represented by $\psi_{lh}(\tau)$. The steps of MOPSO framework are given as follows.

Step 1: Initialize the particles in the swarm and specify the maximum number of iterations T for the stopping criterion. Positions of two particles in the swarm are set by the minimum and the maximum sailing speeds in all legs in order to guarantee the lower and upper bound solutions in the initial Pareto frontier. These two positions are denoted by θ^{min} and θ^{max}, respectively. For the remaining particles, the position $\theta_{lh}(1)$ at iteration 1 is randomly generated by setting the average speed v_i at leg i between the range of minimum and maximum sailing speed $[v_{min}, v_{max}]$. Velocity $\omega_{lh}(1)$ is set to 0 for every particle in the swarm.

Step 2: Calculate the fuel consumption and the service level by using the Eqs. (4)–(5) presented in Section 3 for each particle in the swarm.

Step 3: Evaluate both objective functions computed in step 2 for the non-dominated front (Pareto front). The non-dominated front is stored in the external archive called elite group. To choose particle l for the elite group, it should satisfy one of the following three criteria;

(1) Both objective function values of particle l should be better than the objective function values of the compared particles in the swarm.
(2) While one objective function value of particle l is better than the one of the other particles, its other objective function value is equal to the objective function value of any particle.
(3) While one objective function of particle l is better than the one of the other particles, its other objective function value is worse than the objective function of any particle.

Step 4: Check whether the stopping criterion is met or not. If it is not satisfied ($\tau < T$), go to step 5; otherwise, the process stops and the final non-dominated front is obtained.

Step 5: Select some particles from the elite group to guide the direction of movement for all particles by following the movement strategy proposed by Nguyen and Kachitvichyanukul (2010). The aim of this strategy is to identify the gaps in the elite group and move particles to the space that has a high gap in the elite group. The advantage is that it helps to improve the distribution of the elite group. For the details of particle selection criteria, we refer to the second movement strategy proposed by Nguyen and Kachitvichyanukul (2010). Basically, this movement strategy checks the gap between particles in the elite group. If the gap is higher than the predefined percentage, then the corresponding particles are added to the unexplored position set as a pair. Then, the movement is performed by randomly choosing a pair of particles (P_{1h}, P_{2h}) from the unexplored position set to be a global guide in the search.

Step 6: Update velocity and position for particles to move to the next position. Velocity of a particle at iteration $\tau+1$ is updated by considering three main components which are velocity at iteration τ, its personal best position and the global best position. The velocity of particle l at iteration $\tau+1$ is computed as follows:

$$\omega_{lh}(\tau+1) = \sigma(\tau)\omega_{lh}(\tau) + a_p U(\psi_{lh}(\tau) - \theta_{lh}(\tau)) + a_g[U(P_{1h} - Q_{lh}(\tau)) + U(P_{1h} - P_{2h})] \quad (10)$$

where $\sigma(\tau)$ is the inertia weight at iteration τ, a_p and a_g denote the acceleration constants for its personal best and the global best position and U is a uniform random variable in the interval [0, 1].

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005

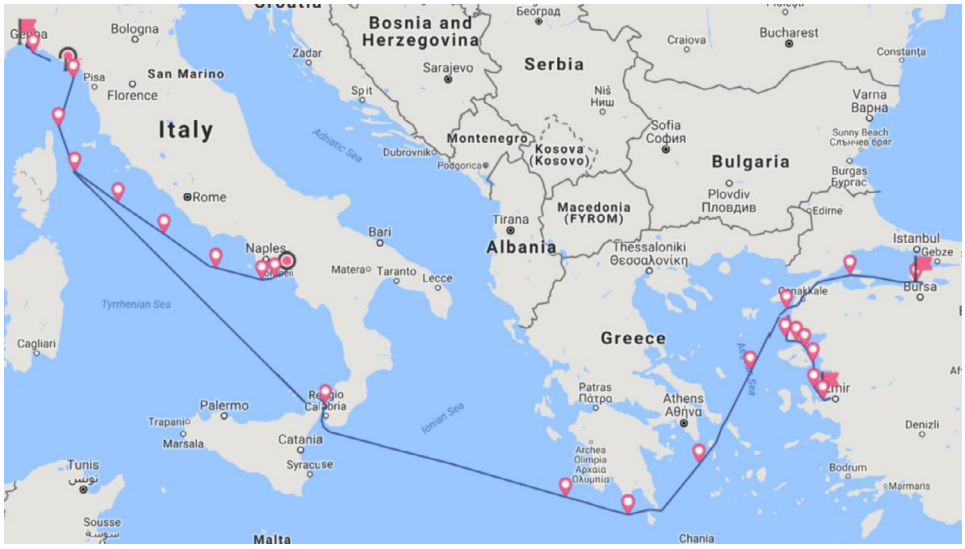

Fig. 9. Service route.

The inertia weight at iteration τ is computed as follows:

$$\sigma(\tau) = \sigma(T) + \frac{\tau - T}{1 - T}[\sigma(1) - \sigma(T)]. \quad (11)$$

After updating the velocity, the position of a particle is computed based on its new velocity and previous position as demonstrated in Eq. (12).

$$\theta_{lh}(\tau + 1) = \theta_{lh}(\tau) + \omega_{lh}(\tau + 1) \quad (12)$$

However, the new position can correspond to an infeasible solution where the vessel speed at each leg does not satisfy the sailing speed constraints. Therefore, we introduce the following conditions to force the position of a particle to be in minimum (θ^{min}) and maximum (θ^{max}) position value.

(i) If $\theta_{lh}(\tau+1) > \theta^{max}$, then $\theta_{lh}(\tau+1)=\theta^{max}$ and $\omega_{lh}(\tau+1)=0$.
(ii) If $\theta_{lh}(\tau+1) < \theta^{min}$, then $\theta_{lh}(\tau+1)=\theta^{min}$ and $\omega_{lh}(\tau+1)=0$.

The iterative process repeats from Step 2 until it reaches the termination criterion.

5. Computational study

In this section, we test the usability of the proposed DSS against the data obtained from a liner shipping company that provides services in Mediterranean and Black Sea regions. The operations team of the company makes the decision on vessel speeds for their services. This team is largely responsible for the scheduling of the vessels and the planning cargo loading on the vessels. Through private conversation with the case liner company, it was realized that speed decision is influenced by several factors and the most significant factor is port situation reports. These reports are usually dispatched through daily emails. The port authority publishes the port status data and forwards it to liner companies through the subscribed agents. These reports are one of the major data sources for the speed decision as they provide information on preferred arrival time at each port.

For the experiment, we choose one of the services operated by the liner company in the Mediterranean region. This service starts from port Salerno in Italy and visits ports La Spezia and Genoa in Italy and ports Gemlik, Yilport, Marport and Izmir in Turkey. After completing the route, the vessels return to the port Salerno for the next voyage. This service covers 2790 nautical miles on average by staying 7.9 days in the sea. The service route is depicted in Fig. 9.

We first show how accurately the fuel consumption function adjusted by weather impact miner. We collect the actual voyage abstract data between 2012 and 2014 from the company. Table 2 presents a part of the sample abstract data for the selected service. The abstract data shows the time stamp on each port arrival and departure with general operation statistics such as average sailing speed, sea days, and fuel consumption for between ports.

As we described in Section 4, we combine the service abstract data with weather information parsed from Copernicus Maritime Environment Monitoring Service. The size of the extracted weather data for this experiment is 43 GB covering three years of nautical data for the Mediterranean and the Black Sea region.

5.1. Weather impact on fuel consumption

We compare our fuel consumption estimation with the theoretical estimation obtained by the empirical model in Yao et al., (2012) and the actual fuel consumption for the given service route. We set the constants in the empirical model as $k_1 = 0.004595$ and $k_2 = 16.42$. The selected service for the experiment had been operated 43 times between 2012 and 2014. The service is divided into 11 legs that correspond to the sea legs between ports and/or straits. Table 3 presents the list of legs and the detail information of each leg including distance and average sailing speed. The table also presents the estimation error in percentage in the right two columns. The estimation error indicates the root mean squared error (RMSE) calculated based on past 43 voyages. In Table 3, the legs are sorted by distance. Fig. 10 illustrates the estimation error for each leg. We refer our weather dependent fuel consumption function and the empirical model proposed by Yao et al., (2012) as WFC and EM, respectively.

As depicted in Fig. 10 and discussed in Fig. 1, estimation error of the empirical model (EM) tends to increase dramatically for the legs with longer distances while the fuel consumption estimations for short legs are relatively accurate. The results show that our proposed fuel consumption model with weather weights can decrease the estimation error for the voyage legs with long distances. For instance, for the longest leg (leg no 11, between two straits), our model (WFC) gives 7.5% error whereas EM has an error of 9.3%.

As depicted in Fig. 9, legs 10 and 11 cross the Mediterranean Sea and hence, these legs are exposed to stronger current compared to the other legs, which lie in the Tyrrhenian Sea and the Aegean Sea. In addition, the impact of weather is more significant

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005

Table 2
Sample voyage abstract data for the Mediterranean service.

General information								
Voyage No.	.....							
Vessel name	.....							
Call sign	TCLA							
Port rotation								
No	Port name	Distance	Avg. speed	Engine load	Avg. fuel C.	Sea (day)	EOSP	COSP
1	Salerno						02/01/2013 15:30	02/01/2013 21:40
2	Genoa	376.0	14.4	50	30.3	1.09	03/01/2013 23:50	04/01/2013 15:00
3	La Spezia	49.0	15.5	55	4.1	0.13	04/01/2013 18:10	11/01/2013 16:50
4	Genoa	42.0	15.8	55	3.8	0.11	11/01/2013 19:30	13/01/2013 00:15
5	Strait of Messina	475.0	14.3	55	43.0	1.39	14/01/2013 09:30	14/01/2013 10:15
6	Dardanelles Strait	648.0	14.3	55	58.4	1.89	16/01/2013 07:30	16/01/2013 10:30
7	Gemlik	106.0	15.0	55	8.8	0.30	16/01/2013 17:35	17/01/2013 10:30
8	Yilport	43.0	15.2	55	3.5	0.12	17/01/2013 13:20	18/01/2013 00:35
9	Marport	31.0	12.8	55	3.0	0.10	18/01/2013 03:00	19/01/2013 07:50
10	Dardanelles Strait	94.0	14.5	55	8.1	0.27	19/01/2013 14:20	19/01/2013 17:45
11	Izmir	131.0	14.4	68	11.4	0.38	20/01/2013 02:50	20/01/2013 23:40
12	Strait of Messina	666.0	15.3	82	64.2	1.82	22/01/2013 19:20	22/01/2013 20:10
13	Salerno	133.0	15.5	82	15.0	0.36	23/01/2013 04:45	
Total		2794.0				7.94		

Table 3
Fuel consumption estimation summary.

Leg no.	Port/Strait		Distance	Avg. speed	Estimation error % (RMSE)	
	From	To			EM	WFC
1	Yilport	Marport	32.33	16.02	0.73	0.73
2	Gemlik	Yilport	42.18	15.71	0.78	0.74
3	La Spezia	Genoa	46.25	15.93	0.93	0.94
4	Marport	Dardanelles Strait	90.08	16.32	1.52	1.46
5	Dardanelles Strait	Gemlik	106.24	15.92	1.47	1.27
6	Dardanelles Strait	Izmir	131.24	16.74	1.50	1.27
7	Strait of Messina	Salerno	149.48	15.68	1.42	1.36
8	Salerno	La Spezia	330.01	15.58	1.92	1.86
9	Genoa	Strait of Messina	473.40	15.77	5.93	5.61
10	Izmir	Strait of Messina	625.64	16.13	6.31	5.35
11	Strait of Messina	Dardanelles Strait	647.26	15.71	9.32	7.48

Table 4
MOPSO parameter setting.

Inertia weight	$\sigma(0)=0.4,\ \sigma(T)=0.9$
Acceleration constants	$a_p=0.05,\ a_g=0.05$
Percent of unexplored gap	5%
Swarm size	500
Number of maximum iterations	1000
Number of elites	100

in the long voyage legs compared to the short legs. Since our fuel consumption function considers the impact of weather conditions on the fuel efficiency, it performs better than the empirical model (EM) especially for long sea legs. Considering intercontinental long voyages where weather will be more severe and highly variable than the exemplified closed seas, the proposed fuel consumption function is anticipated to provide better estimates.

5.2. Numerical results on multi-objective speed optimization problem

In this section, we test the performance of our multi-objective speed optimization model given by Eqs. (4)–(8). MOPSO is used to find the optimal sailing speed at each leg which minimizes the fuel consumption and maximizes the average service level. The parameters used in the PSO solver are shown in Table 4. In our experiments, we used a computer with 1.80 GHz Intel (R) Core (TM) and 8.00 GB of RAM. The solution algorithm is implemented in Visual C# running under Windows 8.1 operating system.

In this experiment, we investigate three voyages of the same liner service operated by the sane vessel between 2013 and 2014 and discuss the potential fuel savings by optimizing the vessel sailing speed. In particular, we compare the fuel consumption obtained by our multi-objective model with the actual fuel consumption of the liner service. We also test the performance of our fuel consumption function against the empirical model proposed by Yao et al., (2012).

The analyzed voyages include seven ports and two strait passages in the sequence following SAL (Salerno), LAS (La Spezia), GEN (Genoa), ST1 (Strait of Messina), ST2 (Dardanelles Strait), GEM (Gemlik), YIL (Yilport), MAR (Marport), ST2 (Dardanelles Strait), IZM (Izmir), ST1 (Strait of Messina), and back to Salerno. The distance between ports are given in Table 3. The start and end dates of these three voyages are 12.08.2013 and 26.08.2013 for the first voyage (V1), 6.03.2014 and 23.03.2014 for the second voyage (V2), and 09.05.2014 and 24.05.2014 for the third voyage (V3).

According to the data obtained from the liner company, the vessel has always arrived before the end of the contracted time window in these voyages. Therefore, we compare the results for the target service level of 100%. As we discussed in Section 3, we assume that the service level degradation for busy and idle ports are different and it is given in Eqs. (13) and (14), respectively.

$$s_i\left(t_i^a\right)=\begin{cases}1 & \text{if } t_i^a \le \beta_i \\ 0.5 & \text{if } \beta_i < t_i^a \le \beta_i+1 \\ 0 & \text{if } \beta_i+1 < t_i^a\end{cases} \tag{13}$$

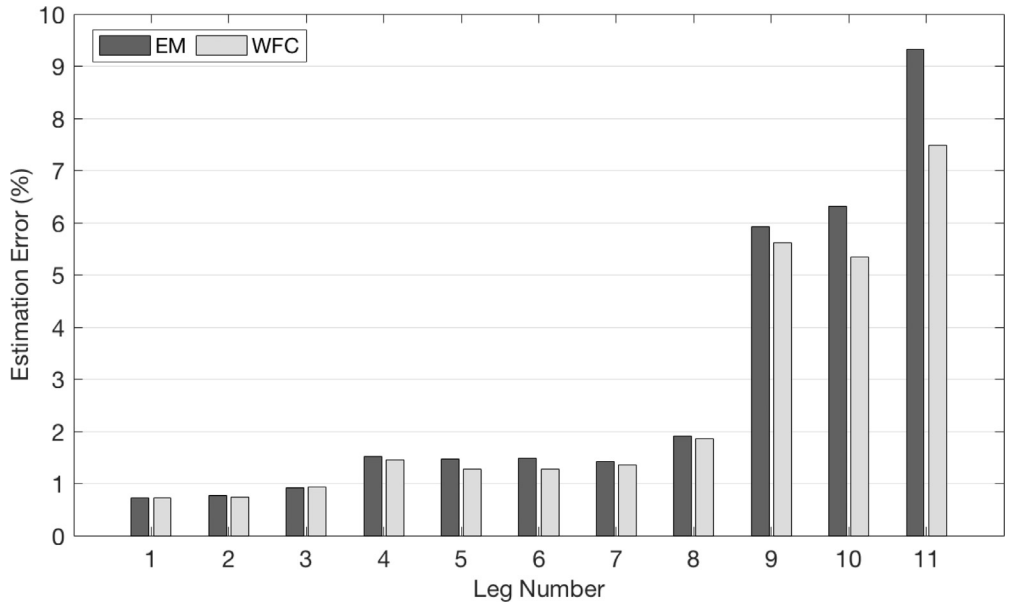

Fig. 10. Percentage error in fuel consumption estimation (RMSE).

Table 5
Optimal sailing speeds and fuel consumption.

			SAL	LAS	GEN	ST1	ST2	GEM	YIL	MAR	ST2	IZM	ST1	SAL	Total
	w_i		–	1.07	1.00	0.97	0.97	1.00	1.20	1.02	1.00	1.00	1.10	1.00	
		WS	–	17.18	16.23	17.21	17.09	17.33	17.40	15.12	15.36	16.02	13.22	14.93	
	v_i	BS	–	17.66	13.03	17.35	16.74	18.07	16.38	13.91	18.42	15.65	13.20	13.24	
V1		AD	–	18.05	18.43	17.48	16.59	17.42	16.26	16.80	14.87	15.17	13.50	14.69	
		WS	3.81	40.65	9.04	44.46	61.03	13.49	7.64	11.91	8.72	21.53	60.14	12.57	295.0
	F_i	BS	3.81	39.10	8.72	46.13	61.99	13.84	6.65	11.75	9.81	20.12	54.69	12.10	288.7
		AD	3.93	39.40	9.60	51.90	76.50	14.70	9.30	12.80	11.80	18.20	55.00	10.90	313.9
	w_i		–	1.00	1.04	1.21	1.00	1.04	1.00	0.97	0.97	0.97	1.00	0.97	
		WS	–	13.24	14.55	16.12	16.03	15.61	16.56	14.38	15.44	14.77	13.54	13.84	
	v_i	BS	–	13.34	13.67	15.66	16.25	15.88	16.77	16.35	14.68	14.89	13.51	13.98	
V2		AD	–	13.46	12.84	16.66	15.96	16.46	13.12	12.77	13.61	18.49	17.62	18.06	
		WS	2.18	36.52	12.98	53.70	60.73	13.60	13.02	9.06	8.80	19.74	55.30	11.94	297.6
	F_i	BS	2.18	36.57	12.33	43.67	61.23	13.30	13.05	9.34	8.86	20.13	55.27	12.40	288.3
		AD	5.10	37.30	12.32	55.56	78.63	15.57	17.25	10.24	15.73	20.83	72.84	15.19	356.6
	w_i		–	0.97	0.99	1.00	0.98	0.97	0.99	0.97	0.98	1.02	1.00	1.00	
		WS	–	15.78	16.04	15.87	16.04	16.28	16.26	15.92	15.94	14.02	13.87	14.01	
	v_i	BS	–	16.73	15.70	15.68	16.04	15.31	16.27	15.52	14.45	14.48	12.76	13.64	
V3		AD	–	17.47	12.92	16.14	15.96	16.00	13.30	14.22	15.48	16.16	14.78	13.22	
		WS	3.93	36.81	10.72	43.91	59.06	13.90	6.22	7.92	8.67	17.56	54.90	12.18	275.8
	F_i	BS	3.93	38.95	10.71	43.60	60.30	13.90	6.30	8.00	8.50	17.50	54.00	12.10	277.6
		AD	5.91	43.60	10.60	52.80	81.40	17.10	9.20	9.40	13.80	21.50	72.40	13.10	350.8

$$s_i\left(t_i^a\right) = \begin{cases} 1 & \text{if } t_i^a \le \beta_i \\ 0.8 & \text{if } \beta_i < t_i^a \le \beta_i + 1 \\ 0.6 & \text{if } \beta_i + 1 < t_i^a \le \beta_i + 2 \\ 0.4 & \text{if } \beta_i + 2 < t_i^a \le \beta_i + 3 \\ 0.2 & \text{if } \beta_i + 3 < t_i^a \le \beta_i + 4 \\ 0 & \text{if } \beta_i + 4 < t_i^a \end{cases} \tag{14}$$

Lastly, we assume that the sailing speed ranges between 12.5 and 19.5 knots.

Table 5 presents the optimal sailing speeds and the corresponding fuel consumption at sea legs. We use the emprical fuel consumption function proposed by Yao et al., (2012) in our multi-objective speed optimization model for benchmarking. Both models are solved by MOPSO with the same parameter set. In the sequel, we refer to our weather dependent speed optimization model and the benchmarking model as WS and BS, respectively. The actual liner data (sailing speed and fuel consumption data) is denoted as AD. The first column of Table 5 denotes the voyages. The next columns present the sailing speeds (v_i) and corresponding fuel consumption (F_i) at each leg. Since our speed optimization model considers weather impact, we also provide the corresponding weather weights (w_i) for each leg in Table 5. These weather weights are obtained through the weather impact miner.

Comparing the sailing speeds, we observe that the empirical fuel consumption model provides slightly lower sailing speeds and hence, the estimated fuel consumption is less than the one obtained by WS. This difference can be attributed to the impact of weather conditions on the sailing speed and fuel consumption. The fuel consumption function in BS does not consider the weather effect. As seen from Table 5, the differences are more significant when the variability in the weather weights is high. We can deduce from this result that in long voyages where weather-sea conditions can be highly variable and severe, vessel speed should be determined by considering the impact on fuel consumption. On the other hand, the gap between the total fuel consumption obtained by WS and the actual fuel data shows the maximum potential fuel savings that can be achieved through optimizing sailing speeds by considering service level.

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005

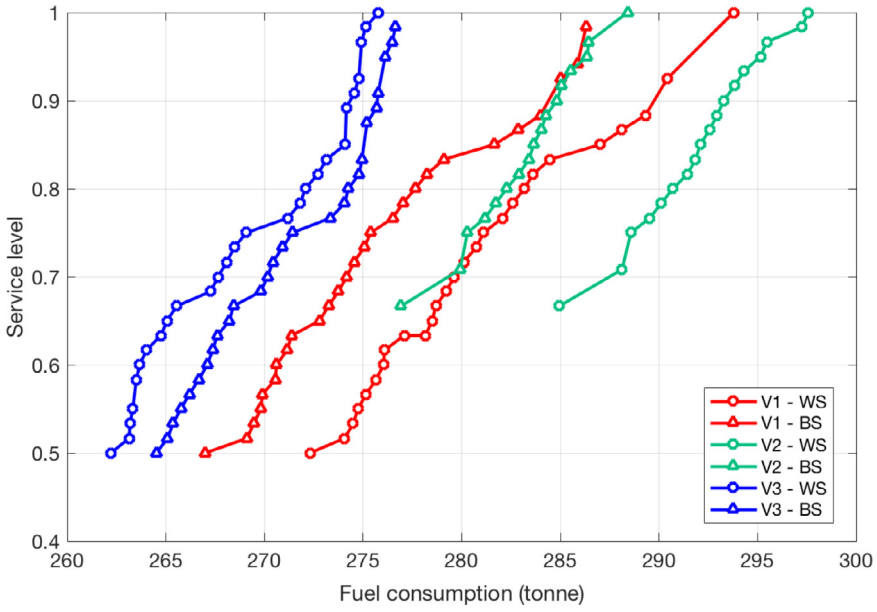

Fig. 11. Pareto front lines for the three voyages.

Fig. 11 illustrates the Pareto front-lines of the three voyages for the Pareto optimal solutions of WS and BS models. The Pareto front lines show the trade-off relationship between fuel consumption and service level. As seen in the figure, achieving high service level requires more fuel consumption. Comparing the Pareto front lines of WS and BS, we observe that the estimated fuel consumption for a given service level is generally higher when we use weather dependent fuel consumption function. In voyage 3, variability in weather conditions is low and hence, Pareto front-lines of WS and BS are closer. This observation is in line with the results in Table 3. The slopes of the Pareto front-lines show how much more fuel is required to achieve higher service level. The managers can use the front-lines to decide required service level and fuel consumption depending on different priorities coming from their clients and operations teams.

6. Conclusion

This paper contributes to vessel speed optimization literature by proposing a way to explore weather archive big-data. In particular, a novel method to parse weather archive data and apply data mining techniques to learn the impact of weather condition on fuel consumption was proposed. Revised fuel consumption function considers the impact of wind and current on fuel consumption of vessels. We focus on speed optimization problem in liner shipping by considering the trade-off between minimizing fuel cost and maximizing service level. PSO technique based solver is used to solve this multi-objective problem. We conduct a computational study by using real-life cases from a liner shipping company. The numerical experiments demonstrate that the revised fuel consumption function provides better fuel consumption estimates compared to the benchmark method which ignores the weather impact. The improvement on fuel estimation is more significant in long voyage legs. Therefore, considering intercontinental long voyages where weather-sea conditions are highly variable than the exemplified closed seas, the proposed DSS can bring significant cost improvements. Moreover, the PSO solver of the DSS generates Pareto optimal solutions that show trade-off analysis between fuel consumption and port service level. Liner operators can decide sailing speeds of vessels for each leg considering the customer requirements.

In spite of its merits, this study has limitations which provide future research directions. Firstly, the source of the weather archive data of the DSS is currently fixed to Copernicus Maritime Environment Monitoring Service and the weather archive data parser can be applied only to this data source. As different archive data sources have different data format and contents, the parser needs to be extended to be able to parse other data sources. Secondly, though our method considers the variabilities in weather conditions when computing fuel consumption, it does not address uncertainties generated from ports. In reality, port side uncertainties are common and can affect the actual service times at ports. A promising research direction would be to include port side uncertainties in the mathematical model and in the PSO solver.

Acknowledgements

This study was partially supported by Korea National Research Foundation through Global Research Network Program (Project no. 2016S1A2A2912265) and an EU Marie Skłodowska-Curie action funded project, MINI-CHIP, under grant number 611693.

References

Agarwal, R., Dhar, V., 2014. Big data, data science, and analytics: the opportunity and challenges for IS research. Inf. Syst. Res. 25, 443–448.
Ai, T.J., Kachitvichyanukul, V., 2009a. A particle swarm optimization for the vehicle routing problem with simultaneous pickup and delivery. Comput. Operat. Res. 36, 1693–1702.
Ai, T.J., Kachitvichyanukul, V., 2009b. A particle swarm optimization for vehicle routing problem with time windows. Int. J. Operat. Res. 6, 519–537.
Andersson, H., Fagerholt, K., Hobbesland, K., 2015. Integrated maritime fleet deployment and speed optimization: cast study from RoRo shipping. Comput. Operat. Res. 55, 233–240.
Aydin, N., Lee, H., Mansouri, A., 2017. Speed optimization and bunkering in liner shipping in the presence of uncertain service times and time windows at ports. Eur. J. Operat. Res. 259, 143–154.
Ballou, P., Chen, H., Horner, JD., 2008. Advanced methods of optimizing ship operations to reduce emissions detrimental to climate change. In: Proceedings of OCEANS 2008, Quebec Convention Centre, Quebec City, Canada, Sep 2008 Print ISSN: 0197-7385.
Bausch, D.O., Brown, G.G., Ronen, D., 1998. Scheduling short-term marine transport of bulk products. Maritime Policy Manage. 25, 335–348.
Besikci, E.B., Arslan, O., Turan, O., Olcer, A.I., 2016. An artificial neural network based decision support system for energy efficient ship operations. Comput. Operat. Res. 66, 393–401.
Bresenham, J.E., 1965. Algorithm for computer control of a digital plotter. IBM Syst. J 4, 25–30.

Choi, Y., Lee, H., Irani, Z., 2017. A big-data drive fuzzy cognitive maps for public policy modelling and impact analysis. Ann. Operat. Res. doi:10.1007/s10479-016-2281-6, forthcoming.
Christiansen, M., Fagerholt, K., Nygreen, B., Ronen, D., 2013. Ship routing and scheduling in the new millennium. Eur. J. Operat. Res. 228, 467–483.
Copernicus Maritime Environment Monitoring Service (Last accessed: 5 October 2016).
Drewry Schedule reliability insight (Last accessed: 1 August 2016).
Fagerholt, K., 2004. A computer-based decision support system for vessel fleet scheduling—experience and future research. Dec. Support Syst. 37, 35–47.
Fagerholt, K., Laporte, G., Norstad, I., 2010. Reducing fuel emissions by optimizing speed on shipping routes. J. Oper. Res. Soc. 61, 523–529.
Fagerholt, K., Lindstad, H., 2007. TurboRouter: an interactive optimisation-based decision support system for ship routing and scheduling. Maritime Econ. Logist. 9, 214–233.
Fang, K., Jiang, Y., Song, M., 2016. Customer profitability forecasting using big data analytics: a case study of the insurance industry. Comput. Ind. Eng. 101, 554–564.
Hvattum, L.M., Norstad, I., Fagerholt, K., Laporte, G., 2013. Analysis of an exact algorithm for the vessel speed optimization problem. Networks 62, 132–135.
Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks, 4, pp. 1942–1948.
Kim, S., Lee, K., 1997. An optimization-based decision support system for ship scheduling. Comput. Ind. Eng. 33, 689–692.
Kontovas, C.A., 2014. The Green Ship Routing and Scheduling Problem (GSRSP): a conceptual approach. Transport. Res. Part D 31, 61–69.
Lam, J.S.L., 2010. An integrated approach for port selection, ship scheduling and financial analysis. NETNOMICS 11, 33–46.
Lee, C.Y., Lee, H.L., Zhang, J., 2015. The impact of slow ocean steaming on delivery reliability and fuel consumption. Transport. Res. Part E 76, 176–190.
Mansouri, A., Lee, H., Aluko, O., 2015. Multi-objective decision support to enhance sustainability in maritime shipping. Transport. Res. Part E 78, 3–18.
Nguyen, S., Kachitvichyanukul, V., 2010. Movement strategies for multi-objective particle swarm optimization. Int. J. Appl. Metaheurist. Comput. 1, 59–79.
Norstad, I., Fagerholt, K., Laporte, G., 2011. Tramp ship routing and scheduling with speed optimization. Transport. Res. Part C 19, 853–865.
Notteboom, T., 2006. The time factor in liner shipping services. Maritime Econ. Logist. 8, 19–39.
Parthibaraj, C.S., Subramanian, N., Palaniappan, P.L.K., Lai, K., 2016. Sustainable decision model for liner shipping industry. Comput. Operat. Res. forthcoming.
Psaraftis, H., Kontovas, C., 2013. Speed models for energy-efficient maritime transportation: a taxonomy and survey. Transport. Res. Part C 26, 331–351.
Psaraftis, N.H., Wen, M., Kontovas, C.A., 2016. Dynamic vehicle routing problems: three decades and counting. Networks 67, 3–31.
Qi, X., Song, D.-P., 2012. Minimizing fuel emissions by optimizing vessel schedules in liner shipping with uncertain port times. Transport. Res. Part E 48, 863–880.
Rew, R., Davis, G., 1990. NetCDF: an interface for scientific data access. IEEE Comput. Graph. Appl. 10, 76–82.
Sinnott, R.W., 1984. Virtues of the haversine. Sky and Telescope 68, 159.
Song, M., Wang, S., 2016. Participation in global value chain and green technology progress: evidence from big data of Chinese enterprises. Environ. Sci. Pollut. Res. doi:10.1007/s11356-016-7925-1.
Tran, N.K., Haasis, H., 2015. Literature survey of network optimization in container liner shipping. Flexible Serv. Manuf. J. 27, 139–179.
UNCTAD (2010). United nations conference on trade and development, Review of Maritime Transport 2010, Geneva.
Wang, S., Meng, Q., 2012. Sailing speed optimization for containerships in a liner shipping network. Transport. Res. Part E 48, 701–714.
Wang, S., Meng, Q., Liu, Z., 2013. Bunker consumption optimization methods in shipping: a critical review and extensions. Transport. Res. Part E 53, 49–62.
Windeck, V., Stadtler, H., 2011. A liner shipping network design–routing and scheduling impacted by environmental influences. In: Network Optimization. Springer, pp. 574–576.
Yao, Z., Ng, S.H., Lee, L.H., 2012. A study on bunker fuel management for the shipping liner services. Comput. Operat. Res. 39, 1160–1172.

Please cite this article as: H. Lee et al., A decision support system for vessel speed decision in maritime logistics using weather archive big data, Computers and Operations Research (2017), http://dx.doi.org/10.1016/j.cor.2017.06.005