Application of Bayesian networks in analysing tanker shipping bankruptcy risks

Grace W.Y. Wang
Maritime Administration, Texas A&M University at Galveston, Galveston, Texas, USA

Zhisen Yang
Liverpool Logistics, Offshore and Marine Research Institute, Liverpool John Moores University, Liverpool, UK

Di Zhang
Intelligent Transport Systems Research Centre, National Engineering Research Centre for Water Transport Safety, Wuhan University of Technology, Wuhan, China

Anqiang Huang
School of Economics and Management, Beijing Jiaotong University, Beijing, China, and

Zaili Yang
Liverpool Logistics, Offshore and Marine Research Institute, Liverpool John Moores University, Liverpool, UK

Abstract

Purpose – This study aims to develop an assessment methodology using a Bayesian network (BN) to predict the failure probability of oil tanker shipping firms.

Design/methodology/approach – This paper proposes a bankruptcy prediction model by applying the hybrid of logistic regression and Bayesian probabilistic networks.

Findings – The proposed model shows its potential of contributing to a powerful tool to predict financial bankruptcy of shipping operators, and provides important insights to the maritime community as to what performance measures should be taken to ensure the shipping companies' financial soundness under dynamic environments.

Research limitations/implications – The model and its associated variables can be expanded to include more factors for an in-depth analysis in future when the detailed information at firm level becomes available.

Practical implications – The results of this study can be implemented to oil tanker shipping firms as a prediction tool for bankruptcy rate.

Originality/value – Incorporating quantitative statistical measurement, the application of BN in financial risk management provides advantages to develop a powerful early warning system in shipping, which has unique characteristics such as capital intensive and mobile assets, possibly leading to catastrophic consequences.

Keywords Bankruptcy, Logistic regression, Bayesian network, Maritime risk, Oil tanker shipping firm

Paper type Research paper

1. Introduction

With severe market illiquidity in global financial markets and possible overcapacity in shipping, evaluation of shipping financial performance becomes an important issue for the companies to ensure their sustainability. The shipping cycle inevitably experienced a down term in tanker and bulk shipping industries. As a result, among 27 tanker shipping firms included in the Bloomberg tanker index, seven of them have filed bankruptcy in recent years. Given that shipping markets are highly competitive in nature, and shipping business features risks and challenges in a constantly changing environment, it will be necessary and beneficial to investigate financial risk management in the shipping segment.

It is however well-noted that although analysing financial performance and predicting bankruptcy risk have a long tradition in economics, studies focusing on tanker shipping failure prediction are rarely to be found in the literature. This study aims to develop an assessment methodology using a Bayesian network (BN) to predict the failure probability of oil tanker shipping firms. To provide a foundation for building the BN model, we focus on the development of a qualitative BN using the casual factors which are identified using correlation analysis in this study. Of all, 27 publicly traded tanker shipping companies in the Bloomberg Tanker Shipping Index are investigated for the period from 2000 to 2010. Shipping companies' performance and the likelihood of filing bankruptcy are influenced by many factors, including the global business cycle, demand for crude oil transport, supply of oil tankers fleet, shift of the trade routes and cost of tanker transport service at a macro level and sales performance, ownership structure, risk management, bank financing and merger and acquisition (M&A) at a micro/firm level. The dependency among the key factors and the relationship between factors and financial performance can be simulated using qualitative diagram in BN, while the quantitative configuration of such dependency (i.e. conditional probabilities) can be obtained using statistical regression analysis based on historical data and through a correlation analysis.

It proposes a new framework capable of incorporating quantitative statistical measurement with BN in shipping financial risk estimation and prediction. A well-structured assessment of shipping performance can provide an early warning system (EWS) in reducing possible financial disturbances and enhancing long-term sustainability in the industry. In addition, by analysing international shipping firms, a better understanding of the effects of financial risk management across national boundaries and shipping segments can be obtained. The results of the paper will also provide important insights to the maritime community as to what performance measures should be taken to ensure the shipping companies' financial soundness under dynamic environments.

The remainder of this paper is organised as follows. Section 2 reviews the current literature relating to bankruptcy prediction. Section 3 describes the proposed methodology and its associated supporting techniques. It is followed by an empirical study to

demonstrate the feasibility of the methodology in Section 4. Section 5 concludes this paper with references to its contributions and limits.

Analysing tanker shipping bankruptcy risks

2. Literature review

2.1 Background of understanding of bankruptcy law

The federal bankruptcy court provides the basic information about how the bankruptcy laws are designed in assisting the troubled businesses to have access to restructure their debts to pay their creditors using possible liquidation:

> Bankruptcy laws are normally designed to protect firms experiencing short term financial difficulties that can be overcome by restructuring. A written disclosure statement and a plan of reorganization must be filed with the court. The disclosure statement is a document that must contain information concerning the assets, liabilities, and business affairs of the debtor sufficient to enable a creditor to make an informed judgment about the debtor's plan of reorganization (www.pacer.gov/psc/eresources.html; www.uscourts.gov).

For example, to keep business survived and pay creditors over time, illiquid and/or insolvent tanker shipping firms that filed bankruptcy – Chapters 11 or 15 – in the USA are able to save costs in ways non-bankrupt firms cannot. The purpose of disclosure is to provide sufficient information to creditors to make fair judgement regarding the feasibility of restructuring process and possible liquidation. Thus, other than the standard information such as debtor's name, tax identification, location, etc., the debtor also needs to disclose the important information about financial and operating performance such as the schedules of assets and liabilities, current income and expenditures, contracts and unexpired leases and financial affairs in a voluntary petition.

2.2 Studies on methodology of predicting bankruptcy

The earliest bankruptcy prediction was in 1932 when FitzPatrick compared 13 financial ratios to distinguish failed and successful firms (FitzPatrick, 1932). Statistical methods have been applied in bankruptcy prediction since 1960s. Knowing that accounting ratios might have linear relationship, Altman (1968) built a multivariate discriminant analysis model. Then, "Z-score" measuring liquidity, profitability, earnings, solvency and management was introduced to the bankruptcy study. Since then, research on the bankruptcy prediction has dramatically increased, and logit analysis, probit analysis and neural networks have become the main stream.

Sun and Shenoy (2007) described the development of bankruptcy prediction models from the simple univariate analysis (Beaver, 1966) and multiple discriminant analysis (Altman, 1968) in the 1960s, to logit and probit models in the 1980s (Ohlson, 1980, Zmijewski, 1984), to neural network models (NN) (Tam and Kiang, 1992), rough set theory (McKee, 1998), discrete hazard models (Shumway, 2001), BN models (Sarkar and Sriram, 2001) and genetic programming (McKee and Lensberg, 2002). Among these techniques, BN models have many attractive features, including easiness to interpret, incorporation of both objective and subjective data and capability of conducting bi-directional analysis (i.e. forward risk prediction and backward risk diagnosis) (Yang *et al.*, 2009). Use of BNs in financial risk prediction in general and bankruptcy in specific has been seen in the literature. According to Thomson Reuters Web of Science, 29 journal papers were found (Yan and Suo, 2013; Andersen *et al.*, 2012) using the keywords BN and financial risk, and 22 results (Sun and Shenoy, 2007) using the keywords BN and bankruptcy. Such works have created an invaluable platform for further research in BN aiming to study financial risk analysis.

Among those recent studies on BN prediction and classification, most of the work is related to applications in finance and banking. Focusing bankruptcy predictions in a probabilistic way, Pena *et al.* (2011) compared traditional statistical methods such as discriminant analysis and logistic regressions to the machine-learning techniques such as Gaussian process and Bayesian Fisher discriminant in an application of computational finance. Obtaining knowledge of causal relations and significant conceptual patterns of attributes from the credit scoring models, Wu (2011) found that classification accuracy can be improved using a hybrid model of different machine learning techniques. While using Gaussian marginal densities from Bayesian model to select new variables, Feki *et al.* (2012) concluded a significant improvement of prediction performance in an application of bank risk prediction. Shao *et al.* (2011) applied BN classification to assess and forecast enterprise risk of Chinese listed companies. To identify causes and influencing factors contributing to financial crisis, Andersen *et al.* (2012) applied BN to four types of financial organisations and found that the industry widely failed to manage operational risks. Using accounting information, Salehi *et al.* (2016) compared data mining techniques such as artificial neural networks and naive Bayesian classifier in predicting corporate financial distress. Some earlier studies can be found in Gestel *et al.* (2006), Yin and Peng (2006) and Zhang *et al.* (1999) while discussing BN in risk prediction in comparison with the use of liner statistical models of discriminant analysis and logistic regression. They applied BN with posterior class probabilities of bankruptcy to analyse corporate clients, and concluded that the technique outperformed liner models with cross-validation applied.

Applications of bankruptcy and failure prediction using quantitative statistical tools on financial institutions can also be found (Gorton and Souleles, 2007; Brunnermeier and Pedersen, 2009; Shleifer and Vishny, 2010; Acharya and Yorulmazer, 2007; Cox and Wang, 2013). However, studies focusing on the failure prediction of shipping in general and tanker shipping in specific are rarely to be found in the literature, revealing the necessity of this work. Few statistical applications in risk prediction on the shipping and maritime segments are as follows. Given that macroeconomic factors impact heavily the performance of maritime industry, Drobetz *et al.* (2013) found that leverage is counter-cyclical, meaning slower adjustment back to the equilibrium during recession. Leverage ratios provided an implication of liquidity and solvency status. Similarly, Yeo (2016) identified the trade-off between liquidity and leverage of 130 shipping companies. A higher leverage ratio impacted cash holdings and the access to debt markets.

3. Methodology
This section is composed of three parts. First, the supporting techniques such as BN and logistic regression models are described to provide the foundation of developing the proposed BN on bankruptcy prediction. Second, independent variables are identified based on a thorough literature investigation and refined by correlation analysis together with experts' domain knowledge. Third, the structure of the network is determined based on experts' domain knowledge and is verified by the d-separation test. Then, the prior probability values in the conditional probability table (CPT) are obtained by applying a series of logistic regression models.

3.1 Bayesian network
A BN (also called belief network or probabilistic network) is a graphical presentation of probability combined with mathematical inference calculation. It is used to represent dependencies between random variables. To form a BN, each variable or a node is

connected by directed links of arrows or arcs with CPT that is filled with values of probability assigned to the variables. The nodes in a BN are called chance nodes. Chance nodes represent uncertain events or variables. Each of them can be a continuous or discrete random variable or a set of events. A deterministic node is a special case of chance nodes, which operates deterministically on other nodes. The arrows are the directed links between nodes, and this direction represents the conditional dependent relationship of these nodes.

The graphical representation makes BNs a flexible tool for modelling causal impacts between events, in particular when the causal impact has a random nature. Also, specification of probabilities can be calculated to a very detailed set of the variables and the parent nodes. Having constructed the model, it can be used to compute effects of information as well as interventions from deterministic nodes. When the states of some variables are fixed, the posterior probability distributions for the remaining variables can be computed. Algorithms based on the Bayes' rule and Chain rule (Jensen, 2001) are developed for probability updating, and they perform efficiently on a large variety of models. This makes BNs well suited to forecasting and diagnosing risks (Yang *et al.*, 2009).

Bayes' theorem has so far been proven to be a coherent method of mathematically expressing a decrease in uncertainty gained by (or proportional to) an increase in knowledge. As an imperative phase of probability analysis, this is achieved by combining probability distributions or functions of different parameters (such as specific outcomes or events) and revising their probabilities when new information/data are obtained.

The Bayes' theory can be expressed by a parameter θ and given observed data x, as follows:

$$p(\theta|x) = \frac{p(x|\theta) \times p(\theta)}{p(x)} \tag{1}$$

where $p(\theta|x)$ represents the posterior probability of θ conditional on x, $p(\theta)$ denotes the prior probability of θ, $p(x)$ denotes the marginal (total) probability of x and is effectively constant because the obtained data are in hand, and $p(x|\theta)$ refers to the conditional probability of x conditional on θ.

Generally Bayes' rule can be considered for the problem of estimating values of k parameters (causes), $\theta = (\theta_1, \ldots, \theta_k)$, using n observations (effects), $x = (x_1, \ldots, x_n)$. In the rule, then, given the observations $x = (x_1, \ldots, x_n)$, the posterior probability distribution of θ can be computed as:

$$p(\theta|x_1, \ldots, x_n) = \frac{p(x_1, \ldots, x_n|\theta) \times p(\theta)}{p(x_1, \ldots, x_n)} \tag{2}$$

Alternative forms of Bayes' rule can also be applied to occurrence probabilities of events. In other words, the probability that an event A occurs given the condition that another event B occurs results in:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \tag{3}$$

It then follows that given the situation where event B is represented by $(B_1, \ldots, B_m)$, the posterior probability of event A can be computed using Bayes' rule as:

$$P(A|B_1, \ldots, B_m) = \frac{P(B_1, \ldots, B_m|A) \times P(A)}{P(B_1, \ldots, B_m)} \tag{4}$$

3.2 Logistic regression
In this research, logistic regression is used to provide the CPT for a discrete BN. There are two types of logistic regression: binary logistic regression and multinomial regression.

In a multinomial regression, the dependent variable y is multinomial and is modelled with different range of values for different status. Usually, the discrete-dependent variable is specified in the form of unobserved but continuous variable y^*, where $y^* \in (-\infty, +\infty)$.

Consider an independent variable set $X = (x_1, x_2, \ldots, x_n)$ leading to the dependent variable y, where each independent variable has several status (j). Defining the unobserved variable y^* as a function of X:

$$y^* = \sum_{i=1}^{n} \sum_{j=1}^{m-1} \beta_{ij} x_{ij} + \varepsilon \tag{5}$$

where β_{ij} represents the contribution of x_i in status j, ε is an unobserved stochastic component and the value of x_{ij} is defined as $x_{ij} = 1$ when status j of x_i occurs; otherwise, $x_{ij} = 0$.

Therefore, we can get the conditional probability of y under a configuration of independent variable set X^0 through multinomial regression (Rijmen and Vomlel, 2008):

$$P(y = y_j) = \frac{e^{\beta_j X^0 + \varepsilon}}{1 + \sum_{i=1}^{m-1} e^{\beta_i X^0 + \varepsilon}} \tag{6}$$

where y_i represents the j_{th} status of y, and m is the number of status of y.

Figure 1 illustrates a simple example as to how we calculate CPTs in a multi-level BN:

First, we calculate the CPT for node C based on logistics regression, then the CPT for node C is used to calculate the CPT for node E. In terms of nodes A, B and D, the percentage
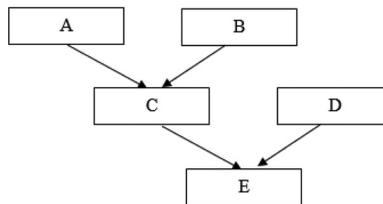
Figure 1.
Hierarchy for calculation of CPTs of a multi-level BN

of each state is used as the conditional probability. In other words, we calculate the CPTs layer by layer, until the CPT of the dependent variable E is obtained.

The binary logistic regression is a special case of multinomial regression. Hence, the calculation process used in multinomial regression also fit in a binary case.

3.3 Shipping bankruptcy prediction using Bayesian network

The process of developing a data-based model consists of four phases: data acquisition, BN structure learning, BN monitoring and analysis and model validation. With respect to the four steps, a new conceptual shipping financial risk analysis methodology is developed, including the following six steps.

3.3.1 Data acquisition. To explore if some tanker operators were more susceptible than others to adverse shocks, a list of shipping companies that have filed bankruptcy recently is first investigated. In the absence of a publicly available official compilation, we rely on keyword search on Lloyd's List Bankruptcy News. Then, we narrow the sample to the companies that are listed in Bloomberg Tanker Shipping Index and/or publicly traded in the stock markets.

3.3.2 Variable identification. Possible factors influencing shipping company's performance and the probability of bankruptcy are first analysed through a careful literature search. They are as follows:

- macroeconomic conditions, including shipping business cycle, global demand for crude oil transportation and the supply of oil tankers;
- microeconomic factors such as operating and sale performance, ownership structure and risk management; and
- other factors such as accessibility to capital markets (i.e. public debt and equity market initial public offerings) and traditional bank financing.

Next, the dependency among the key factors and the quantitative configuration of such dependency (i.e. conditional probabilities) can be obtained using statistical regression analysis based on historical data (i.e. Step 5 later in the text). During this process, the factors of insignificant effect on shipping financial performance will be eliminated to ensure that only major factors are taken into account in financial risk analysis and predictions. We focus on the balance of the global supply and demand for tanker transport, the interactions between movements of oil and petroleum products, capacity of the world tanker fleets, pricing of tanker spot and charter freight rates and the managerial decisions such as financing, investment, merger, risk assessment and asset and liability allocation.

3.3.3 Variable interdependency modelling in a hierarchy. Historical firm-level financial information (from Step 1) combined with macro-level data from Step 2 shows tanker market fluctuations. The dependency of the variables will be established through step-wise correlation analysis at 10, 5 and 1 per cent significant levels. Bivariate correlations between various parameters measure the strength of liner dependence between two variables. Logit regression could also identify critical factors in predicting failures of the shipping companies when a dependent variable is binary. In this case, a categorical variable that has two values such as failed or survived is included in the logit regression. After statistical analysis, significant variables are selected based on both correlation analysis and expert judgements.

3.3.4 Bayesian network-based risk model establishment. After identifying all risk variables in the hierarchy in Step 3, one can start to confirm the relationship among

them and construct a qualitative BN to represent their interactive dependencies. The knowledge about the financial risk problem and intuitive understanding of the various dependencies are then used to construct the causal structure. Here, the graphical representation becomes very handy and permits the risk analysts to express the fundamental relationships of direct or indirect influence between risk variables. The influence relationships expressed in BNs have a feature with causality. To ensure that the BN models correspond with a real-world situation, it is required to check their d-separation properties (Jensen, 2001; Yang *et al.*, 2009). D-separation is a very important concept in the interaction of BNs.

3.3.5 Prior probability distribution. When the qualitative BNs have been built and verified (via d-separation), the prior probabilities to all nodes of the networks require to be distributed to model the uncertainties of risk variables. They can be calculated through the statistical regression analysis in Steps 2 and 3.

3.3.6 Financial risk prediction and diagnosis. Once the prior probabilities are distributed along the established BN structure, the next task is to analyse the networks to obtain the posterior probabilities of the financial performance, given observations of the relevant variables from a realistic situation. The objective of using BNs in a risk model is to predict and infer the unobservable situations (uncertainties) related to the shipping financial performance through the posterior probabilities when observable evidence (e.g. the factors' values at both macro and micro levels) is provided. Such posterior probabilities can be obtained by using the Bayes' rule and Chain rule (Jensen, 2001) theoretically and calculated on the basis of computing software from a practical viewpoint (Yang *et al.*, 2009).

4. Shipping financial risk modelling
4.1 Data
Table I lists sound tanker shipping companies and the suspicious firms that failed, ceased trading in the stock exchange or were close to file bankruptcy in the study period from 2000 to 2010.

Macro-level variables such as annual gross domestic product (GDP) growth rate and industrial production are obtained from the World Bank. Oil prices and US stocks of crude oil and gasoline are collected from the US Energy Information Administration (EIA). Supply of oil tankers and world existing tanker fleets is obtained from the Review of Maritime Transport from UNCTAD. Tanker shipping indices such as the Baltic dirty/clean tanker indices and the Bloomberg tanker shipping index are from Bloomberg. Firm-level financial data and the decision of mergers and acquisition (M&A) strategy are available from the Thomson One. A panel data set containing cross-sectional and time-series data is applied. The sequence of the data is summarised by a number of shipping firms over the study period. If multiple financial filings are available in the same year for a company because of further adjustment, we adopt updated reclassified or restated financial data rather than original filings. We also drop the observations of a particular year out of the sample if a firm had missing data for that year.

4.2 Risk variables
This research studies why certain publicly traded shipping companies were underperformed and chose to file bankruptcy in the recent years. Given the highly unpredictable dynamics with continuous fluctuations in tanker operation, we identify factors that affect tanker performance and possible determinants of filing Chapters 11 or 15 as follows.

Symbol	Company	Exchange	Country
Normal/survived			
TK:US	TEEKAY CORP	New York	USA
SNI:NO	STOLT-NIELSEN LTD	Oslo	NO
SFL:US	SHIP FINANCE INTL LTD	New York	USA
GESCO:IN	GREAT EASTERN SHIPPING CO	Natl India	IN
NAT:US	NORDIC AMERICAN TANKERS LTD	New York	USA
ODF:NO	ODFJELL SE-A SHS	Oslo	NO
CPLP:US	CAPITAL PRODUCT PARTNERS LP	NASDAQ GS	USA
FSJ:LN	FISHER (JAMES) & SONS PLC	London	GB
VLCCF:US	KNIGHTSBRIDGE TANKERS LTD	NASDAQ GS	USA
9119:JP	IINO KAIUN KAISHA LTD	Tokyo	JP
EURN:BB	EURONAV SA	EN Brussels	BE
TNP:US	TSAKOS ENERGY NAVIGATION LTD	New York	USA
PR:IM	PREMUDA SPA	Brsa Italiana	IT
DIS:IM	D'AMICO INTERNATIONAL SHIPPI	Brsa Italiana	IT
CCORB:SS	CONCORDIA MARITIME AB-B SHS	Stockholm	SE
DHT:US	DHT HOLDINGS INC	New York	USA
9130:JP	Kyoei Tanker Co Ltd	Tokyo	JP
TOPS:US	TOP SHIPS INC	NASDAQ GS	USA
ECHEM:NO	EITZEN CHEMICAL ASA	Oslo	NO
A1MN:GR	NEWLEAD HOLDINGS LTD	Berlin	DE
Suspicious/failed			
OSG:US	OVERSEAS SHIPHOLDING GROUP	New York	USA
FRO:NO	FRONTLINE LTD	Oslo	NO
BLTA:IJ	BERLIAN LAJU TANKER TBK PT	Indonesia	ID
OCT:CY	OCEAN TANKERS HOLDINGS PUBLI	Nicosia	CY
SRAB:SS	SRAB HOLDING AB-B SHS	Stockholm	SE
QGTS:QD	QATAR GAS TRANSPORT(NAKILAT)	Qatar	QA
GMR:US	GENERAL MARITIME CORP	New York	USA

Sources: Compiled by authors from various sources; List of companies is from the Lloyd's list market data of Bloomberg Shipping Company Index. Suspicious companies include companies currently ceased trading in the stock exchange (OSG, OCT, SRAB, QGTS and GMR) and companies were close to the failure or have filed Chapter 11 (FRO and BLTA) based on the news search on Lloyd's list

Table I.
Tanker shipping
companies listed in
Bloomberg Tanker
Shipping Index

4.2.1 Global business cycle and growth of gross domestic product. Macroeconomic factors such as industrial production, economic growth and systemic risks are important in a shipping firm's operating performance. Current economic fundamentals are crucial in predicting future shipping cycles. Reversed business cycle affects aggregated demand negatively for both public and private sectors. The information of GDP growth and industrial activities is available from the World Bank.

4.2.2 Price of oil and petroleum products. The interaction of total demand of oil and petroleum products, crude oil production and oil stock building affects the spot price of crude oil and gasoline prices (EIA; www.eia.gov). Figure 2 shows the crude oil spot price.

4.2.3 Demand for tanker transport services. Tankers have different sizes and capacities to transport crude oil from limited source of origin to innumerable pre-determined oil refineries and to customers. Demand for oil transportation measured in ton-miles shows the movements of oil and petroleum products.

Figure 2.
Crude oil spot prices
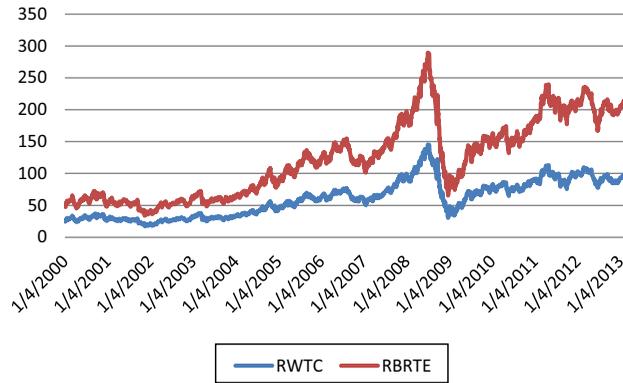
Notes: RWTC: Cushing, OK WTI Daily Crude Oil Spot Price
FOB (dollars per barrel); RBRTE Europe Brent Daily Crude Oil
Spot Price FOB (dollars per barrel)
Source: www.eia.gov and summarised by authors

4.2.4 Shifting of global demand for crude oil transportation services. There has been a
growing trend of replacing import oil and petroleum products from foreign suppliers with
domestic production of crude oil in the USA since the global recession in 2008. The proxy to
capture the shift of demand is the US consumption of oil produced by the OPEC countries
and the stocks of crude oil and gasoline available in the USA from the EIA. The US ending
stocks of crude oil are presented in Figure 3.

Figure 3.
US ending stocks of
crude oil
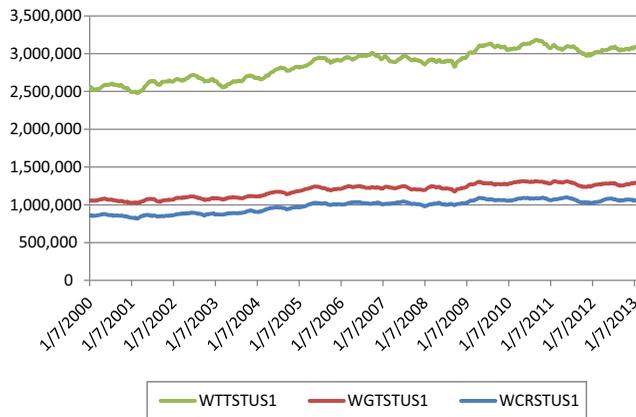
Notes: WTTSTUS1: Weekly US ending stocks of crude oil and
petroleum products (thousand barrels); WGTSTUS1: Weekly US
ending stocks of total gasoline (thousand barrels); WCRSTUS1:
Weekly US ending stocks of crude oil (thousand barrels)
Source: www.eia.gov and summarised by authors

4.2.5 Supply of oil tankers (total supply of tanker fleet and capacity). Supply of oil tankers, including very large crude carriers (VLCC) and smaller size tankers affect charter rate for crude oil transport and pricing of oil cargos. Vessel sales, available number of tankers, total capacity of the active tankers, the existing tanker fleet and tonnage of shipping fleet in million dwt also influence the time charter and spot freight rates and the pricing of crude oil transportation services (Shipping Intelligence Weekly, Clarkson Research Service Ltd.; Review of Maritime Transport, UNCTAD). Other indicators such as newbuilding prices, second-hand prices and scrap prices will affect the number of world tanker fleet and the available capacity in the tanker market (Lloyds Shipping Economist; Drewry's Shipping Statistics).

4.2.6 Spot and time charter tanker freight rates (Cost of tanker transport services). Oil shipping costs and freight rates vary on different types of charter agreements. In the previous tanker cycle, a shortage of shipbuilding capacity led to a rapid growth in the market, and unanticipated heavy overbook and mismatch of new ship delivery and scrapping explained why these cycles are so irregular. The unpredictable volatility of tanker charter rates and the time charter equivalent provide an unstable stream of voyage revenue. Spot rates (Worldscale), time charter rates ($/day) and average freight earnings ($/day) such as the Baltic Dirty Tanker Index, Baltic Clean Tanker Index and Bloomberg Tanker Company Index are examples.

4.2.7 Mergers and acquisition strategy. Shipping finance that supports the acquisition decision is through the combination of issuing new common equity offerings and the offer of shipping loans. A successful fund raising through the common stock offerings and the availability of bank loans determine the likelihood of illiquidity and the possibility of filing a bankruptcy. The financial performance of a shipping firm depends on the accessibility to capital markets such as public debt and the initial public offerings in the equity market and traditional bank financing. Kavussanos and Marcoulis (2004) studied stock market performance of a publicly traded shipping firm and found it is highly related to the success of M&A strategies.

4.2.8 Company's balance sheet, operating performance and financial risk-taking behaviour. Internal liquidity, operating income, EBITDA, etc., are the measures of financial health of a company. Level of capital and the involvement of private equity funds or hedge funds also ensure whether a smooth refinancing transaction is feasible.

From the micro aspect, financial risks in the shipping industry are broadly categorised by Kavussanos and Visvikis (2006) and Wang *et al.* (2014) as follows:

- how easily a firm can turn its assets into cash indicates liquidity risk;

- possibility that a firm will not be able to make payments to honour its obligations on time presents default risk;

- financial risk refers to how dependent a firm is on borrowing and financial leverage;

- credit risk captures if the counterparties of a business will fulfil their agreement in a timely manner; and

- market risk emphasizes the correction of a firm's performance to shocks in the stock and financial markets.

Shipping companies can manage their risk-taking behaviour through risk assessment tools and financing/investment decision involved.

Through the supervision rating system, the CAMELS, referring to the components of the bank supervision rating system in the USA include capital adequacy, assets quality, management quality, earnings ability, liquidity and sensitivity to the market conditions, the EWS of above risks can be revealed. Other than the above influencing factors, Z-score also reflects the degrees of systemic risk. Indicators that are provided to distinguish surviving shipping companies from their failed peers serve as the early warning signals that predict

shipping companies' failures. We need to identify dependence among risk management/ assessment factors that influence company's decision of filing bankruptcy. Also, we need to filter surviving firms as comparison based on market segment, employee or total assets.

The first-level variables are summarised in Table II. Firm-level data from 1999Q2 to 2013Q2 is also analysed, and the result is presented in Table III. The coefficients are means from the sample shipping companies, and the standard deviations are in the parenthesis. Except S, DBEV, DEBT, referring to the ratios of growth, assets or current liabilities, all variables in the Columns 2 and 6 are in terms of millions of US dollars for the failed companies, and the variables in Columns 3 and 7 are for the survived firms. We also present the difference in mean and the t-statistics in the Columns 4 and 8 which test the mean difference of both sample firms.

The above indicators not only reflect a firm's operating performance and the managerial decision of filing bankruptcy but also influence a shipping company's stock market performance. The dependency among the key factors and the relationship between factors can be simulated using qualitative description and diagram in BN, while the quantitative configuration of such dependency can be obtained using statistical mean difference test and regression analysis based on historical data.

The results of the correlation analysis to identify the interdependency among the variables are presented in Table IV, in which all variables in the second column are 1 per cent correlated with their individual own variable listed in the first column. To simplify the conceptual BN model for ship financial risk prediction, the variables are shown in Table V.

Variable	Definition	Variable	Definition
S	Growth of sales	QUIK	Quick ratio
NSAL	Net sales or revenue	CURR	Current ratio
AST	Total assets	ROE	Return on equity
OP	Operating profit per employee	ROA	Return on assets
CE	Capital expenditures	EBITDA	Earnings before income tax
DBEV	Debt to enterprise value	NICM	Net income
DEBT	Long-term debt to asset ratio		

Table II. Definition of variables

Variable	Failure(SD)	Survive(SD)	Diff(t-test)−
S	0.563 (2.922)	0.034 (0.314)	−0.529 (−4.149)***
NSAL	148.289 (131.791)	147.987 (183.521)	−0.302 (−0.024)
AST	2557.252 (2271.324)	1690.828 (2071.205)	−866.424 (−5.30)***
OP	461.105 (1134.359)	9.254 (30.938)	−451.841 (−8.30)***
CE	−192.691 (344.165)	−111.970 (146.111)	80.720 (4.429)***
DBEV	0.618 (0.220)	0.603 (0.256)	−0.015 (−0.759)
DEBT	0.514 (0.196)	0.420 (0.183)	−0.094 (−6.485)***
QUIK	3.264 (14.067)	2.66 (5.222)	−0.189 (−0.93)
CURR	3.171 (12.591)	2.981(5.657)	0.175 (0.736)
ROE	−0.475 (6.817)	−0.436 (1.907)	0.432 (1.416)
ROA	0.063 (0.100)	0.054 (0.089)	−0.009 (−1.275)
EBITDA	0.472 (8.362)	−1.003 (21.441)	−1.475 (−0.874)
NICM	0.981 (10.157)	0.319 (11.100)	−0.662 (−0.793)

Table III. Descriptive status of variables

Note: ***Significant at 1% level in the two-way test

Parameters	Correlation (significant at 1%)
Failure (F)	FL, GOP, USOI, S, CE, OP, DA, GDP
Fleet	
Fleet (FL)	D, PNB, PSH, FS, FTC, CAE, BDI, GOP, USOI, USOS, COSP, DA, GDP
Demolition (D)	FL, PNB, PSH, FS, FTC, CAE, BDI, GOP, USOI, USOS, COSP, CE
Price of newbuilding (PNB)	FL, D, PSH, FS, FTC, CAE, BDI, GOP, USOI, USOS, COSP, CE, OP
Price of 10-year second-hand (PSH)	FL, D, PNB, FS, FTC, CAE, BDI, USOI, USOS, COSP, CE, OP
Oil trade	
Crude oil spot price (COSP)	FL, D, PNB, PSH, FS, FTC, CAE, GOP, USOI, USOS, DA
Global oil production (GOP)	F, FL, D, PNB, FS, FTC, CAE, BDI, USOI, USOS, COSP, DA
US crude oil import (USOI)	F, FL, D, PNB, PSH, FS, FTC, CAE, BDI, GOP, USOI, USOS, COSP, CE, OP, DA, GDP
US oil stock (USOS)	FL, D, PNB, PSH, FS, FTC, CAE, BDI, GOP, USOI, COSP, DA, GDP
Economic growth	
GDP growth (GDP)	F, FL, FS, FTC, CAE, BDI, USOI, USOS, S, CE, OP, DA
Firm financial performance	
Growth in sale (S)	F, FS, CE, OP, DA, GDP
Capital expenditure (CE)	F, D, PNB, PSH, FTC, CAE, BDI, USOI, S, CE, OP, DA, GDP
Operating profit per employee (OP)	F, PNB, PSH, CAE, USOI, S, OP, DA, GDP
Long-term debt to asset ratio (DA)	F, FL, CAE, BDI, GOP, USOI, USOS, COSP, S, CE, GDP
Freight rate	
Freight spot price (FS)	FL, D, PNB, PSH, FTC, CAE, BDI, GOP, USOI, USOS, COSP, S, GDP
Freight Time charter (FTC)	FL, D, PNB, PSH, FS, CAE, BDI, GOP, USOI, USOS, COSP, CE, GDP
Clarksons tanker average earnings (CAE)	FL, D, PNB, PSH, FS, FTC, CAE, BDI, GOP, USOI, USOS, COSP, CE, OP, DA, GDP
Baltic dirty index (BDI)	FL, D, PNB, PSH, FS, FTC, CAE, GOP, USOI, USOS, CE, DA, GDP

Table IV.
Results of correlation
analysis

Fleet	*Likelihood of failure*	*Firm financial performance*
Price scrap ($/ldt)		Current ratio (liquidity)
Fleet demolition (#; DWT)		Sales (growth rate)
Price second-hand ($mill)		EBITDA (profitability)
Price newbuilding ($mill)		ROA/ROE (management)
Fleet (#, $mill DWT)		
		Economic growth
Freight rate		GDP (annual growth)
Spot (route, WS)		Industrial value added
Time charter (1,3,5, $/day)		(annual, % of GDP)
Freight Baltic Tanker earning ($/day)		
BDIY/BCIY (index)		
	Oil trade	
	Stocks of crude oil in US	
	US crude oil Import/export (MBPD)	
	Spot crude oil price ($/barrel)	
	Oil production (MBPD)	
	OPEC: crude oil	

Table V.
The conceptual
dependency among
factors influencing
failure

Given a long list of variables that can share the same/similar influence, only the variables that can sufficiently represent the influence as well as receive effective historical data to support the quantification of its influence are selected. For instance, if we want to capture the impact of fleet, five suitable variables are price scrap ($/ldt), fleet demolition (#; DWT), price second hand ($mill), price newbuilding ($mill) and fleet (#, $mill DWT).

4.3 A proposed model
The prototype of a EWS model for analysing and predicting the failure rate of an oil tanker firm is developed by taking the risk variables as well as their interdependent relationships into consideration. The model is shown in Figure 4. In this process, domain experts are interviewed to rationalise the grades used to define each variable.

Four root nodes, five intermediate nodes and one leaf node are presented in the model. The states of each node are shown as follows:

- supply: high ($\geq$150), low ($<$150);
- demand: high ($\geq$80), low ($<$80);
- GDP growth: high ($\geq$3), medium (0-3), low ($<$0);
- freight rate: high ($\geq$150), medium (70-150), low ($<$70);
- sales growth: high ($\geq$0.5), medium (0-0.5), low ($<$0);
- EBITDA: high ($\geq$1), medium (0-1), low ($<$0);
- ROA: high ($\geq$0), low ($<$0);
- debt ratio: high ($\geq$0.6), medium (0.4-0.6), low ($<$0.4);
- current ratio: high ($\geq$1), medium (0-1), low ($<$0); and
- failure: yes (1), no (0).
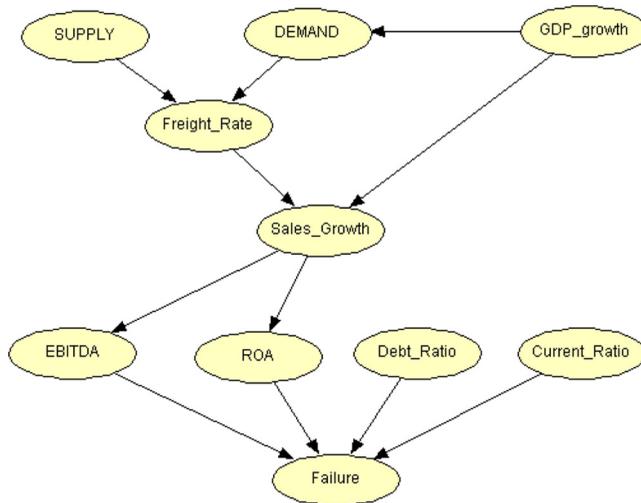
Figure 4.
Simplified BN model
for tanker firm

Dependent variable	Independent variable status	Coefficient
Failure	Current Ratio (CR) (high)	1.552
(yes)	CR(medium)	0
	CR(low)	0.621
	EBITDA(high)	0.034
	EBITDA(medium)	0
	EBITDA(low)	−0.074
	ROA(high)	0.078
	ROA(low)	0
	CR(high)	0.462
	CR(medium)	0
	CR(low)	−1.252
	Constant	−2.281
Sales(high)	GDP(high)	0.642
	GDP(medium)	0
	GDP(low)	−0.803
	FR(high)	0.407
	FR(medium)	0
	FR(low)	−0.878
	Constant	−1.18
Sales(low)	GDP(high)	0.273
	GDP(medium)	0
	GDP(low)	−0.675
	FR(high)	−0.252
	FR(medium)	0
	FR(low)	0.556
	Constant	−0.036
ROA(low)	Sales(high)	−19.588
	Sales(medium)	0
	Sales(low)	0.185
	Constant	−1.615
EBITDA	Sales(high)	0.547
(high)	Sales(medium)	0
	Sales(low)	1.553
	Constant	−4.554
EBITDA	Sales(high)	−1.325
(low)	Sales(medium)	0
	Sales(low)	0.212
	Constant	−2.682
Demand	GDP(high)	−0.123
(low)	GDP(medium)	0
	GDP(low)	−0.389
	Constant	−1.349
FR(high)	Supply(high)	−17.548
	Supply(low)	0
	Demand(high)	18.449
	Demand(low)	0
	Constant	−19.933
FR(low)	Supply(high)	4.216
	Supply(low)	0
	Demand(high)	−1.723
	Demand(low)	0
	Constant	−0.969

Table VI.
Coefficients
summary

4.4 The conditional probability table and prior probabilities for each node

Once the BN model is developed, the next step is to establish the CPT for each node. When executing the model, the CPT of each node will be calculated through the logistics regression method.

Table VI presents the logistic regression estimates of the contribution, β_{ij} in Equation (5), of different variables.

When putting these coefficients β_{ij} into Equation (6), it is possible to calculate the conditional probabilities of different nodes in this network. With regard to the root nodes, the proportion of each defined state is used as their conditional probabilities. The relevant conditional probabilities of the other (i.e. leaf) nodes are listed in Appendices 1-6.

4.5 Results

With the established CPT of each node in the BN model, the marginal probability of each child node can be obtained using Bayes' rules and the associated computing software packages (e.g. Netica). Figure 5 indicates that the estimated failure rate (bankruptcy rate) of an oil tanker shipping firm is approximately 30.2 per cent given the input data covering the period of 2001-2010. Such a result is consistent with the tanker company failure ratio of 26 per cent (seven failures out of 27 companies) to a large extent. It therefore partially verified the model.
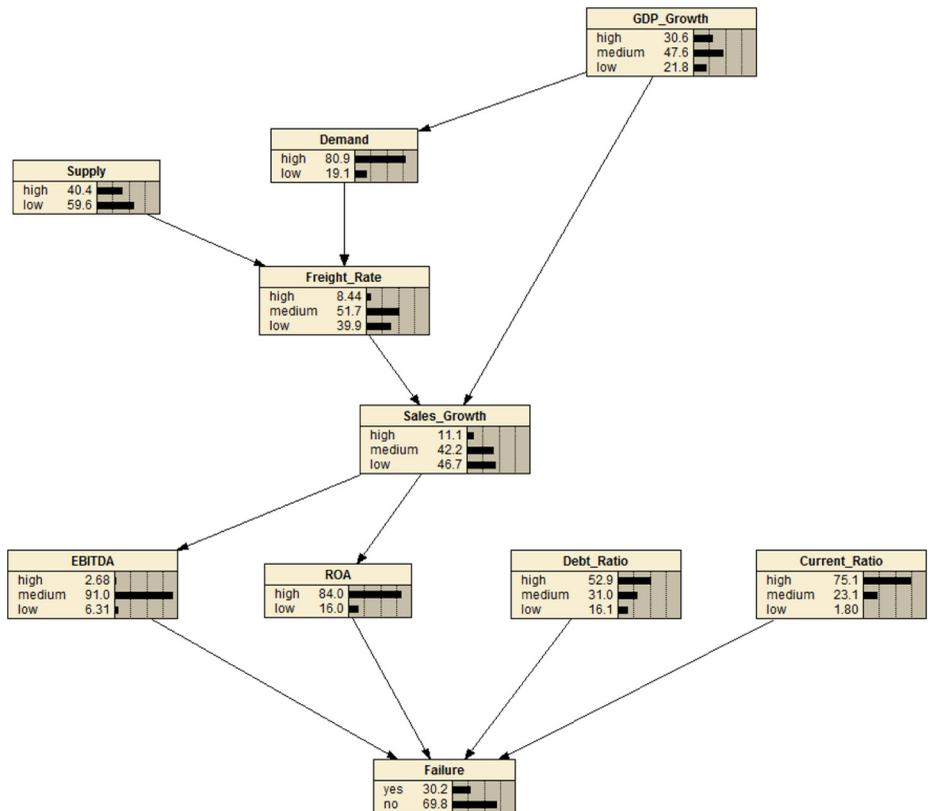

Figure 5.
The result of the BN model

5. Discussions and conclusions

After the sub-prime mortgage crises in 2007-2009 in the USA, the maritime sector has experienced a dramatic downturn in the shipping cycle. While how shipping companies respond to supply and demand imbalance remains unclear, the concern of oversupply and the high leverage ratios increases the uncertainties and further exposes shipping companies to unforeseen risks. We focus on the tanker shipping sector because we observed more companies tend to choose Chapters 11, 13 and 15 bankruptcy laws to have access to restructure their debts to pay their creditors.

This study develops a conceptual assessment methodology using a BN to predict failure or underperformance of oil tanker shipping firms. We provide a comprehensive review of the current literature of statistical and non-statistical risk assessment methods. We demonstrate step by step on how to build a qualitative BN model. Further, we identify critical factors in a hierarchical interdependent relationship using correlation analysis and provide foundation to establish a simplified discrete BN. It can be used to provide useful insights for shipping companies to prevent themselves from financial risks.

The model and its associated variables can be expanded to include more factors for an in-depth analysis in future. To do so, it requires the establishment of a comprehensive database capable of detailing the data to a level where the interdependent relationship among all the variables is modelled in a timeframe containing several financial cycles. Another possible development of this work is to use the model to investigate the risk regarding the merge of shipping companies given its rising concerns in the maritime industry, particularly in the container shipping sector.

References

Acharya, V. and Yorulmazer, T. (2007), "Too many to fail: an analysis of time-inconsistency in bank closure policies", *Journal of Financial Intermediation*, Vol. 16 No. 1, pp. 1-31.

Andersen, L.B., Hager, D., Maberg, S., Naess, M.B. and Tungland, M. (2012), "The financial crisis in an operational risk management context-a review of causes and influencing factors", *Reliability Engineering and System Safety*, Vol. 105, pp. 3-12.

Altman, E.I. (1968), "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *Journal of Finance*, Vol. 23 No. 4, pp. 589-609.

Beaver, W. (1966), "Financial ratios as predictors of failures", *Journal of Accounting Research*, Vol. 4, pp. 71-111.

Brunnermeier, M.K. and Pedersen, L.H. (2009), "Market liquidity and funding liquidity", *The Review of Financial Studies*, Vol. 4 No. 6, pp. 2201-2238.

Cox, A.M.G. and Wang, J. (2013), "Root's barrier: construction, optimality and applications to variance options", *The Annals of Applied Probability*, Vol. 23 No. 3, pp. 859-894.

Drobetz, W., Gounopoulos, D., Merikas, A. and Schroder, H. (2013), "Capital structure decisions of globally-listed shipping companies", *Transportation Research Part E*, Vol. 52, pp. 49-76.

Feki, A., Ben Ishak, A. and Feki, S. (2012), "Feature selection using Bayesian and multiclass support vector machines approaches: application to bank risk prediction", *Expert Systems with Applications*, Vol. 39 No. 3, pp. 3087-3099.

FitzPatrick, P.J. (1932), "A comparison of the ratios of successful industrial enterprises with those of failed companies", *Journal of Accounting Research*, The Certified Public Accountant Beaver. 1968. (In three issues: October, 1932, 598-605; November, 1932, 656-662; December, 1932, pp. 727-733.)

Gestel, T.V., Baesens, B., Suykens, J.A.K., Van den Poel, D., Baestaens, D.M. and Willekens, M. (2006), "Bayesian kernel based classification for financial distress detection", *European Journal of Operational Research*, Vol. 172 No. 3, pp. 979-1003.

Gorton, G.B. and Souleles, N.S. (2007), "Special purpose vehicles and securitization", *Therisks of Financial Institutions*, University of Chicago Press.

Jensen, F.V. (2001), *Bayesian Network and Decision Graphs*, Springer-Verlag, New York, NY.

Kavussanos, M.G. and Marcoulis, S.N. (2004), "Cross-industry comparisons of the behaviour of stock returns in shipping transportation and other industries", *Research in Transportation Economics*, Vol. 12 No. 1, pp. 107-142.

Kavussanos, M.G. and Visvikis, L.D. (2006), "Shipping freight derivatives: a survey of recent evidence", *Maritime Policy and Management*, Vol. 33 No. 3, pp. 233-255.

Mckee, T. (1998), "A mathematically derived rough set model for bankruptcy prediction", *Collected Papers of the Seventh Annual Research Workshop on Artificial Intelligence and Emerging Technologies in Accounting, Auditing and Tax, Artificial Intelligence/Emerging Technologies Section of the American Accounting Association*.

Mckee, T. and Lensberg, T. (2002), "Genetic programming and rough sets: a hybrid approach to bankruptcy classification", *European Journal of Operational Research*, Vol. 138 No. 2, pp. 436-451.

Ohlson, J.A. (1980), "Financial ratios and the probabilistic prediction of bankruptcy", *Journal of Accounting Research*, Vol. 18 No. 1, pp. 109-131.

Pena, T., Martinez, S. and Abudu, B. (2011), "Bankruptcy prediction: a comparison of some statistical and machine learning techniques", *Computational Methods in Economic Dynamics*, Vol. 13, pp. 109-131.

Rijmen, F. and Vomlel, J. (2008), "Assessing the performance of variational methods for mixed logistics regression models", *Journal of Statistical Computation and Simulation*, Vol. 78 No. 8, pp. 7765-7779.

Salehi, M., Shiri, M.M. and Pasikhani, M.B. (2016), "Predicting corporate financial distress using data mining techniques an application in Tehran Stock Exchange", *International Journal of Law and Management*, Vol. 58 No. 2, pp. 216-230.

Sarkar, S. and Sriram, R. (2001), "Bayesian models for early warning of bank failures", *Management Science*, pp. 1457-1475.

Shao, J., Wang, S. and Liu, Y. (2011), "Enterprise risk assessment and forecast: based on Chinese listed companies in 2009-2010", Wu, Dash (Ed.), *Quantitative Financial Risk Management*, Springer, Berling, Heidelberg.

Shleifer, A. and Vishny, R.W. (2010), "Unstable banking", *Journal of Financial Economics*, Vol. 97 No. 3, pp. 306-318.

Shumway, T. (2001), "Forecasting bankruptcy more accurately: a simple hazard model", *The Journal of Business*, Vol. 74 No. 1, pp. 101-124.

Sun, L.L. and Shenoy, P.P. (2007), "Using Bayesian networks for bankruptcy prediction: some methodological issues", *European Journal of Operational Research*, Vol. 180 No. 2, pp. 738-753.

Tam, K. and Kiang, M. (1992), "Managerial application of neural networks: the case of bank failure predictions", *Management Science*, Vol. 38 No. 7, pp. 926-947.

Wang, G.W., Woo, S.-H. and Mileski, J. (2014), "The relative efficiency and financial risk assessment of shipping companies", *Maritime Policy and Management*, Vol. 41 No. 7, pp. 651-666.

Wang, G.W.Y., Yang, Z., Zhang, D., Huang, A. and Yang, Z. (2016), "Application of Bayesian networks in analysing tanker shipping bankruptcy risks", *International Association of Maritime Economists Conference*, Germany.

Wu, W.W. (2011), "Improving classifications accuracy and causal knowledge for better credit decisions", *International Journal of Neural System*, Vol. 21 No. 4, pp. 297-309.

Yan, Y. and Suo, B. (2013), "Risks analysis of logistics financial business based on evidential Bayesian network", *Mathematical Problems in Engineering*, Vol. 2013 No. 2013, pp. 1-8.

Yang, Z.L., Bonsall, S. and Wang, J. (2009), "Use of hybrid multiple uncertain attribute decision making techniques in safety management", *Expert Systems with Applications*, Vol. 36 No. 2, pp. 1569-1586.

Yeo, H. (2016), "Solvency and liquidity in shipping companies", *The Asian Journal of Shipping and Logistics*, Vol. 32 No. 4, pp. 235-241.

Yin, H. and Peng, Y. (2006), "Bayesian-based credit risk early warning", International conference on Artificial Intelligence, *Beijing, China*.

Zhang, G., Hu, M. and Patuwo, B. (1999), "Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis", *European Journal of Operational Research*, Vol. 116 No. 1, pp. 16-32.

Zmijewski, M. (1984), "Methodological issues related to the estimation of financial distress prediction models", *Journal of Accounting Research*, Vol. 22, pp. 59-82.

Corresponding author
Zaili Yang can be contacted at: z.yang@ljmu.ac.uk

Appendix 1

DR			CR		
high	52.86			high	75.10
medium	31.02			medium	23.06
low	16.12			low	1.84
GDP					
high	30.60		Supply	high	40.40
medium	47.60			low	59.60
low	21.80				

Appendix 2

GDP	High	Medium	Low
Demand high	81	79.1	84.8
Demand low	19	20.9	15.2

Appendix 3

Supply	High		Low	
DEMAND	high	low	high	low
FR high	0.1	0.1	17.5	0.1
FR medium	17.8	3.6	77.3	72.4
FR low	82.1	96.3	5.2	27.5

Appendix 4

GDP		High			Medium			Low		
FR	high	medium	low	high	medium	low	high	medium	low
Sales high	30.6	20.5	7.0	14.6	13.5	4.5	13.0	8.5	3.0
Sales medium	35.0	35.0	29.0	51.5	44.0	35.6	63.0	61.3	52.2
Sales low	34.4	44.5	64.0	33.9	42.5	59.9	24.0	30.2	44.8

Table AIV.
CPT of sales growth
(%)

Appendix 5

Sales	High	Medium	Low	Sales	High	Medium	Low
EBI high	1.8	1.0	4.4	ROA high	1	83.4	80.7
EBI medium	96.4	92.7	88.2	ROA low	0	16.6	19.3
EBI low	1.8	6.3	7.4				

Table AV.
CPT of ROA and
EBITDA (%)

Appendix 6

EBITDA	high						medium						low					
ROA	high			low			high			low			high			low		
DR	high	medium	low	high	medium	low	high	medium	low	high	medium	low	high	medium	low	high	medium	low
CR *High*																		
Yes	46.1	35.0	13.4	44.2	33.3	12.5	45.3	34.3	13.0	43.4	32.5	12.1	43.5	32.6	12.2	41.6	30.9	11.4
No	53.9	65.0	86.6	55.8	66.7	87.5	54.7	65.7	87.0	56.6	67.5	87.9	56.5	67.4	87.8	58.4	69.1	88.6
CR *Medium*																		
Yes	15.4	10.3	3.2	14.4	9.6	2.9	14.9	9.9	3.1	14.0	9.3	2.8	14.0	9.3	2.8	13.1	8.7	2.6
No	84.6	89.7	96.8	85.6	90.4	97.1	85.1	90.1	96.9	86.0	90.7	97.2	86.0	90.7	97.2	86.9	91.3	97.4
CR *Low*																		
Yes	25.2	17.5	5.7	23.8	16.4	5.3	24.6	17.1	5.6	23.2	16.0	5.2	23.3	16.0	5.2	21.9	15.0	4.8
No	74.8	82.5	94.3	76.2	83.6	94.7	75.4	82.9	94.4	76.8	84.0	94.8	76.7	84.0	94.8	78.1	85.0	95.2

Table AVI.
CPT of failure (%)