INTRODUCTION

Shipping is a very dynamic sector of global transportation as over 90% of international trade is carried by the sea due to its cost-effectiveness. However, the occurrence of some important maritime accidents such as the grounding of the Exxon Valdez, the capsizing of the Herald of Free Enterprise and the foundering of Estonia passenger ferry attracted the attention of the world on maritime safety. The common causes of these accidents include, insufficient monitoring of the ship’s technical condition, inadequate training of the crew and deficiencies in the safety management on board. The Port state control (PSC) regimes have been developed in response to this situation, as ”second line of defense” allowing countries to inspect foreign-registered ships in port other than those of the flag state and take action against ships that are not in compliance with international rules.


With the introduction of the NIR, the initially defined 25% quota for inspections to be performed by each individual Member State was abandoned. As an alternative, a 'fair share' scheme was developed. Moreover, under this New Inspection Regime the old Target Factor system is replaced by a risk-based one, the Ship Risk Profile. The new approach to inspections has been designed to reward quality shipping with a reduced inspection burden, whereas ships considered to be high risk will be subject to more frequent in-depth inspections. Several factors are used in order to calculate Ship Risk Profile such as the type of ship, age of ship, flag, company performance and others. In this system ships are described as High Risk Ships (HRS), Low Risk Ships (LRS) or Standard Risk Ships (SRS). The ship’s priority for inspection, the interval between inspections and the scope of the inspection are determined according to Ship Risk Profile.

Due to the importance of maritime transportation and the variety of hazards there is an extensive research effort on maritime safety. Goerlandt and Montewka (2015) presented a review of risk definitions, perspectives and scientific approaches to risk analysis found in the maritime application area. The most common definition of risk is as the expected value of the probability of an event occurrence and the utility of the consequences. Among the early research studies on Maritime Risk Analysis, qualitative analysis was more common. Moreover, most studies were based on the accident statistics.

ABSTRACT: The Port State Control (PSC) regimes have been developed after several important maritime accidents allowing countries to inspect foreign-registered ships in port other than those of the flag state and take action against ships that are not in compliance with international rules. In the present research, the risk influencing factors adopted in the definition of the Ship Risk Profile used for selecting ships for inspections under the New Inspection Regime of the Paris Memorandum of Understanding (MoU) on Port the State Control are characterized. Moreover, the PSC inspections at two ports, Thessaloniki and Liverpool ports, are analyzed in terms of type and age of ships and other factors influencing the ships’ risk. In addition, the results of inspections in terms of detentions and number and type of deficiencies found at the two ports are analyzed. Finally, four Bayesian network models are developed using the data from Port State Control inspections. Two models are used to analyze the Thessaloniki port and two models the Liverpool port. The first two Bayesian network models, one for each port, are used to assess how risk factors such as the flag, the age, the recognized organization among others, influence the number of deficiencies and the detention of the ship. The other two models focus on the categories of deficiencies and how the risk factors influence specific deficiencies.
Guedes Soares and Teixeira (2001) have reviewed some applications of quantified risk assessment within the maritime transportation sector and identified early studies on the probability of ship loss by foundering and capsizing. However, many frameworks have been developed for assessing risks in maritime industry with both quantitative and qualitative approaches. Montewka et al. (2014) introduced a systematic, transferable and proactive framework estimating the risk for maritime transportation systems. Their framework is developed with the use of Bayesian Belief Networks and combines discrete and continuous variables. Moreover, Chai et al. (2017) developed a Quantitative Risk Assessment (QRA) model in order to evaluate the risk of ship collisions. Since PSC is the main strategy to eliminate substandard ships, more and more research works have addressed this subject. The literature on Port State Control is broad and covers many aspects such as elements of law, effectiveness, targeting, and discrepancies in implementation among others. Sage (2005) proposed a criterion for the identification of ‘High Risk Vessels’ (HRV) in European waters that would allow coastal States to monitor the movements of ships posing significant risk. The idea was to develop a risk index, which could be calculated and attached to ships individually. Cariou et al. (2009) investigated how the different target factors contribute to deficiencies and detentions of the vessel using data from 26,515 PSC inspections that took place within the Indian Ocean MoU region from 2002 to 2006. Cariou and Wolff (2015) proposed a methodology to complement existing practices focusing on detention with the objective to improve the selection process. They concluded that current practices should still rely on detention statistics but could be improved through a combination of quintile regressions for count data and Probit models. Graziano et al. (2017) conducted a study to determine the factors leading to differences in treatment during the Port State Control inspection process among EU Member States using data from 14 elite interviews. According to them, Paris MOU is the most effective of the regional agreements. However, a single training policy for Port State Control Officers will reduce the differences from country to country during the inspection. Moreover, Graziano et al. (2018) analyzed 25 inspection reports prepared by EMSA to monitor the level of implementation and enforcement of the Directive 2009/16/EC on Port State Control. Chen et al. (2019) analyzed the factors behind the detention of ships using ship detention 2008–2017 data from port states in the Asia-Pacific Region collected by Tokyo MOU. The factors considered are not ship-related (type, age, flag etc) but the types of deficiencies obtained from Tokyo MOU annual reports.

Bayesian Networks (BNs) is a common tool used in several research works to develop frameworks for risk assessment. The capability of representing rather complex, not necessarily causal but uncertain relationships makes Bayesian Networks a powerful tool for risk modelling. Bayesian Network models can be developed directly from data via expert elicitation or a combination of both. The data-driven approach reduces the dependence on human experts and in some cases increases the accuracy of the model. Hänninen (2014) used Bayesian networks for maritime safety modelling. They showed that the most important advantages of BNs are the easy combination of data with expert knowledge and the capability of updating the model with new information and of propagating evidences related to specific scenarios. Wang et al. (2012) presented an innovative approach towards integrating logistic regression and a BN. In their model they used a binary logistic regression method of providing input for a BN, making use of different maritime accident data resources. Hänninen and Kujala (2014) explored the dependencies of deficiencies and detentions of the ship and ship’s involvement in accidents and incidents. They used Bayesian networks with the method of learning from the inspection, accident and incident data. Sotiralis et al. (2016) presented an approach that more adequately incorporates human factors considerations into quantitative risk analysis of ship operation. The approach is based on the development of a Bayesian Network model that take into account the human performance and calculates the collision accidental probability due to human error. Yang et al. (2018a) proposed a data-driven Bayesian Network (BN) based approach to analyze risk factors influencing PSC inspections, and predict the probability of vessel detention. Data used are related to bulker carriers of seven major countries in Europe at the ‘Pre-NIR’ time, from 2005 to 2008 in Paris MoU. Yang et al. (2018b) presented a Bayesian Network model to determine vessel detention rates, emphasizing on company performance as a risk factor. The network was developed based on inspection data of bulk carriers in 2015–2017 involving nine major countries from the Paris MoU. The remainder of the work is organized as follows. Section 2 describes the methodology used. Section 3 addresses the Bayesian Network models developed and the results obtained are presented in Section 4. Finally, Section 5 presents the conclusions and suggestions of improvements in future works.

2 METHODOLOGY

2.1 Paris MoU on Port State Control

To select the ships for inspection, Port State Control Officers (PSCOs) use an information system, known as ‘THETIS’, that supports the new Port State Control inspection regime (NIR). THETIS indicates which ships have priority for inspection, depending on Ship Risk Profile and allows the results of inspections to be
recorded. Via THETIS these reports are made available to all Port State Control authorities in the Paris MOU. Data on ships particulars and reports of previous inspections carried out within the Paris MoU region are provided by the information system as well. Every ship in the information system will be attributed a ship risk profile (SRP), according to Annex 7 of Paris MOU (Paris MoU on Port State Control 2014). This SRP will determine the ships priority for inspection. Ships can be “high risk”, “standard risk” or “low risk”. The profile is calculated using generic and historic factors. A ship’s risk profile is recalculated daily taking into account changes in more dynamic parameters such as age, the 36 month inspection history and company performance. Recalculation also occurs after every inspection and when the applicable performance tables for flag and ROs are changed. Annex 7 of Paris MOU defines the factors and their weighting on the overall risk of the ship. High Risk Ships (HRS) are ships with 5 or more points in the risk profile point system. Low Risk Ships (LRS) are ships which meet all the criteria of the Low Risk Parameters and have had at least one inspection in the previous 36 months. Standard Risk Ships (SRS) are ships that are neither HRS nor LRS.

Risk factors used for defining Ship Risk Profile are generic and historical. The generic factors are:

- **Type of ship**
- **Age of ship**
- **Flag**
- **Company Performance**
- **Recognized Organization Performance**

If a ship is a Chemical tanker, a Gas Carrier, an Oil tanker, a Bulk carrier or Passenger ship, it has 2 weighting points in the ship risk profile while all other types do not have points.

Every year Paris MOU organization publishes a new list with White, Grey and Black flags. The flag ranking is based on the total number of inspections and detentions over a 3-year rolling period for flags with at least 30 inspections in the period. Flags that are in the Black flag list in the categories of Very High Risk (VHR), High Risk (HR) and Medium to High Risk (MtoHR) take 2 weighting points in the system. Black flags in the category of Medium Risk (MR) take 1 point.

**2.2 Theoretical background on Bayesian Networks**

Probabilistic networks are graphical models of casual interaction among a set of variables, where the variables are represented as nodes of a graph and the interactions as directed edges between the notes. Any pair of unconnected nodes of such a graph indicates conditional independence between the variables represented by these nodes. A Bayesian network can be described briefly as an acyclic directed graph (DAG), which defines a more compact factorization of the joint probability distribution over the variables that are represented by the nodes of the DAG, where the factorization is given by the directed links of the DAG. Generally, a Bayesian network can be described in terms of a qualitative component, consisting of a DAG, and a quantitative component, consisting of a joint probability distribution that factorizes into a set of conditional probability distributions governed by the structure of the DAG. The fact that the structure of a Bayesian network can be characterized as a DAG derives from basic axioms of probability calculus leading to recursive factorization of a joint probability distribution into a product of lower-dimensional conditional probability distributions.

First, any joint probability can be decomposed (or factorized) into a product of conditional distributions of different dimensionality, where the dimensionality of the largest distribution is identical to the dimensionality of the joint distribution.
Second, statements of local conditional independences manifest themselves as reductions of dimensionalities of some of the conditional probability distributions. Most often, these independence statements give rise to dramatic reductions of complexity of the DAG, such that the resulting DAG appears to be quite sparse.

A discrete Bayesian network, over variables $X$ consists of an acyclic, directed graph and a set of conditional probability distributions $P$. Each node in graph corresponds one-to-one with a discreet variable $X_n \in X$ with a finite set of mutually exclusive states. The directed links of graph specify assumptions of conditional dependence and independence between the random variables.

There is a conditional probability distribution, $P(X_n|\text{pa}(n))$ for each variable $X_n$. The set of variables represented by the parents, $\text{pa}(n)$, are sometimes called the conditioning variables of $X_n$ – the conditioned variable.

The chain rule from probability using the information provided by the BN structure is given by:

$$P(X) = \prod_{v \in V} P(X_v|\text{pa}(v))$$  \hspace{1cm} (1)

Figure 1 represents the topology of a simple Bayesian network with 3 nodes. $X_1$ is the parent of $X_2$ and $X_3$. For this network the equation of chain rule is the following:

$$P(x_1, x_2, x_3) = P(x_1)P(x_2|x_1)P(x_3|x_1)$$

Figure 1 Simple Bayesian network

The construction of a Bayesian network thus runs in two phases. First, given the problem at hand, one identifies the relevant variables and the casual relations among them. The resulting graph, called topology, specifies a set of dependence and independence assumptions that will be enforced on the joint probability distribution, which is next to be specified in terms of a set of conditional probability distributions, $P(X_v|\text{pa}(v))$, one for each “family”, $\{v\} \cup \text{pa}(v)$ of the topology.

A Bayesian network can be constructed manually, semi-automatically from data or through a combination of a manual and a data-driven process, where partial knowledge about structure as well as parameters (conditional probabilities) blend with statistical information extracted from databases.

Data-driven modeling is the task of identifying a Bayesian network model from a source of data. Structure learning from data is the task of inducing the structure i.e., the graph, of a Bayesian network from a source data. There exists different classes of algorithms for learning the structure of a Bayesian network such as search-and-score algorithms and constraint-based algorithms as well as combinations of the two. The main part of structure learning is to identify a graph structure that best encodes a set of conditional dependence and independence relations (CIDR’s) between the variables.

The PC algorithm is a constraint-based algorithm for learning the structure of a Bayesian network (Kjaerulff and Madsen 2008). The main steps of the PC algorithm are

1. Test for conditional independence between each pair of variables represented in order to derive the sets of conditional dependence and independence relations (CIDR’s)
2. Identify the skeleton of the graph
3. Identify colliders
4. Identify derives directions

Parameter estimation in a Bayesian Network is the task of estimating the values of the parameters corresponding to a specific topology and specific distributions $P$ from a database.


Let $N$ be a Bayesian network, with a specific topology, for which we would like to estimate the parameters $T$ of $P$ from a database of cases $D$ (datasets). The estimation of the parameters $T$ from $D$ proceeds, as mentioned above, by iterating the E-step and the M-step. Given an initial assignment to the parameters $T$, the E-step is to compute the expected sufficient statistics under $T$, while the subsequent M-step is to maximize the log-likelihood of the parameters under the expected sufficient statistics. These two steps are alternated iteratively until a stopping criterion is satisfied (Kjaerulff and Madsen 2008).

3 BAYESIAN NETWORK MODELS

Four Bayesian network models are developed using the data collected as described above. Two models are used to analyze the Thessaloniki port and two models the Liverpool port.

The first two Bayesian network models, one for each port, will be used to analyze how risk factors such as the flag, the age, the Recognized Organization etc, influence the number of deficiencies and the detention.
The other two models focus on categories of deficiencies and how the risk factors influencing specific deficiencies.

Bayesian Network is a powerful tool used to create models for risk assessment. An advantage over the simple Bayesian theory is that BNs provide a better graphical representation while comparing to other techniques, BN has a stronger mathematical background. Moreover, taking advantage of causal inference, BN can be used to analyze the importance of different factors influencing ship risk profile and the relationships between them.

3.1 Data

The data used for developing the Bayesian Network (BN) models are collected from Paris MOU Inspection Database and consist of the inspections carried out at two different ports during 2018. The Thessaloniki and the Liverpool ports are selected as case study. The port of Thessaloniki is located on the southern part of Paris MOU community. The port of Thessaloniki is the most important port in Macedonia, a region of Greece, and one of the most important ports in Southeast Europe. During 2018, a total number of 1.929 ships visited the Thessaloniki port. Most of them, 752 ships, arrived at conventional terminal while 492 ships arrived at container terminal. During the same year, 424.500 containers loaded in TEU’s while 3.844.522 tons of conventional cargo went in and out. Concerning ship’s flag, 659 ships were under Greek flag while 402 were under Malta’s flag. Thessaloniki Port Authority S.A (2018). On the other hand, the port of Liverpool is located on the northern part. The port of Liverpool is located on both banks of the River Mersey within the North West of the United Kingdom. Commodities at Liverpool port include mainly automotive, containers, dry bulks. Liverpool port is the fourth biggest port in UK, with million tons of cargo loaded and unloaded every year. Moreover, it is the fifth bigger port in UK handling RO-RO main freight Department for Transport (2018) Moreover, the number of inspections carried out in these two ports are comparable as in Thessaloniki 75 vessels were inspected during 2018 while in Liverpool 95 inspections were carried out during the same year.

As mentioned above, Paris MOU organization provides access to inspection data through its website. With a custom search for a specific year and location all inspections at Thessaloniki and Liverpool ports during 2018 can be listed. However, all the data required to calculate Ship Risk Profile according to Annex 7 of Paris MOU are not displayed directly in the search result screen. Data such as company performance, vessel’s history and deficiencies categories obtained and processed manually.

3.2 Risk Variables

All models use the same variables except from those that focus on categories of deficiencies that have one extra variable. The risk variables along with their states are explained below.

1. Flag

This variable has three states: white, grey and black. Using the data collected as described above and the “White, Grey and Black (WGB) list” which is published by Paris MOU organization all flags was characterized as white, grey or black.

2. Recognized Organization (RO)

The four states (High, Medium, Low and Very Low) of this variable was created by classifying the ROs using the performance list that Paris MOU organization publishes.

3. Age

This variable has nine states as vessel age was categorized in groups of five years each in order to be manageable for the model. So state “$0\to5$” refers to vessels being under 5 years old while state “$5\to10$” refers to vessels which age is $5 < x \leq 10$.

4. Type

This variable describes the different types of ships inspected and has different states for each type (Container, bulk carrier, chemical tanker, etc.). The name for every state is the same as Paris MOU organization uses to describe ship types.

5. Type of inspection

There are three types of inspections that can be carried out by a Port State Control officer: initial inspection, more detailed inspection, expanded inspection. Therefore, this variable has the three states mentioned above depending on how much in depth the vessel was examined.

6. Gross Tonnage

This variable has three states: small, medium and big. Vessels with gross tonnage under 5.000 are characterized as small while vessels with tonnage over 30000 as big. The remaining are classified as medium.

7. Company Performance

This variable has three states: high, medium and low. Company performance was calculated manually using data collected as described above and the methodology as described in Annex 7 of Paris MOU.
8. Deficiencies

This variable represents the number of deficiencies found in the particular inspection. The number of deficiencies was also categorized in groups in order to be manageable by the network. Therefore, this variable has four states. “Zero” state for vessels with no deficiencies, “1to3” for vessels found with 1, 2 or 3 deficiencies, “4to9” and “over10”.

9. Detention

This variable has two states, “Detention” or “No_Detention” depending on the result of this particular inspection.

10. Risk Profile

This variable represents the ship risk calculated according to Annex 7 of Paris MOU. Therefore, it has 3 states: “Low_Risk_Ship” for vessels meeting low risks criteria, “High_Risk_Ship” for vessels considered dangerous and “Standard_Risk_Ship” for the others.

3.3 Topology

Genie software can automatically define the topology of a Bayesian Network model using a data file through a learning algorithm.

A data file consisting of 10 columns, one for each risk variable, was created using “Microsoft Excel” software and given as an input to Genie in order to apply structural learning. The PC algorithm was chosen as the most suitable among the available algorithms. The result was a complex graph with many links between variables. Therefore, some of those links considered meaningless or even incorrect were deleted in order to simplify the model.

Figure 2 presents the Bayesian Network model topology. Risk variables described above are represented as nodes. The following topology is the same for both ports models.

- Nodes representing the risk variables “Type”, “Flag” and “Age” have no parents. All these variables are describing vessel characteristics and are not influenced by other variables. They are on the top of the Bayesian Network model.
- The variable that represents the Recognized Organization (RO) is influenced by the type and the flag due to the trend for specific types and flags to choose specific ROs.
- Gross Tonnage is influenced by the type of the vessel as long as different types lead to different tonnage categories.
- “Risk Profile” is influenced by all factors contributing to its definition according to Paris MOU text as explained above.
- Nodes representing the risk variables “Deficiencies”, “Risk Profile” and “Detention” are not influencing other variables. All these variables represent the results of the inspection and have no influence on the risk factors. They are on the bottom of the Bayesian Network model.

As mentioned above, two models are also developed in order to analyze the influence of the risk factors on the deficiency categories at the two ports. A different Microsoft Excel file was created in which every column describes a variable. However, many vessels have more than one deficiency, the record of each ship is repeated as many times as the number of deficiencies found. For this reason, those models will be used only to analyze categories of deficiencies.

Figure 3 presents the Bayesian Network model topology of the model developed to analyze the influence of risk factors on the deficiency categories.

Figure 3 Topology of the Bayesian Network model developed to analyze categories of deficiencies

- As explained above, nodes representing the variables about vessel characteristics such as “Age”, “Type” and “Flag” are not influenced by other nodes.
- Nodes representing inspection’s results such as “Detention” and “Category of Deficiency”
are not influencing risk factors. “Risk Profile” also does not influence any node.

- “Category of Deficiency” is on the bottom of the Bayesian Network model as it does not influence other nodes. It is influenced by vessel characteristics such as “Age”, “Type” and “Flag”. Moreover, nodes such as “Company Performance” and “RO” influence the “Category of Deficiency” due to the fact that the performance of the company and the RO can lead to specific types of deficiencies.

- The node “Deficiencies”, representing the number of deficiencies, is not connected with “Category of Deficiency” as the number of deficiencies found in a Port State Control inspection is not directly associated with the category they belong to.

### 3.4 Parameters

The network parameters of all models are learned using Genie software and the EM learning algorithm Dempster et al. (1977). Using as input the same data files used for topology definition, Genie software generates the probability tables of all model nodes. Variables that do not have parents, such as the “Age”, have marginal probability tables. Table 1 shows the probability table of the variable “Age” from Thessaloniki’s model as learned by Genie software using EM algorithm.

<table>
<thead>
<tr>
<th>Age</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0to5</td>
<td>0.0694</td>
</tr>
<tr>
<td>5to10</td>
<td>0.2549</td>
</tr>
<tr>
<td>10to15</td>
<td>0.1483</td>
</tr>
<tr>
<td>15to20</td>
<td>0.1224</td>
</tr>
<tr>
<td>20to25</td>
<td>0.0541</td>
</tr>
<tr>
<td>25to30</td>
<td>0.1119</td>
</tr>
<tr>
<td>30to35</td>
<td>0.1161</td>
</tr>
<tr>
<td>35to40</td>
<td>0.0683</td>
</tr>
<tr>
<td>Over40</td>
<td>0.0541</td>
</tr>
</tbody>
</table>

Table 1 Probability table of the variable “Age”

However, the probability tables of the variables that are influenced by others, i.e. have parents, are not simple marginal probability distributions but Conditional Probability Tables (CPTs), for each state of the parent variable.

Table 2 presents the Conditional Probability Table (CPT) of the variable “Company Performance”, obtained from Thessaloniki’s model, which is influenced by the variable Flag.

<table>
<thead>
<tr>
<th>Flag</th>
<th>White</th>
<th>Grey</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.1836</td>
<td>0.1666</td>
<td>0.85</td>
</tr>
<tr>
<td>Medium</td>
<td>0.7551</td>
<td>0.8333</td>
<td>0.15</td>
</tr>
<tr>
<td>High</td>
<td>0.0612</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2 Probability table of the variable “Company Performance”

For variables that have more than one parent, CPTs are very complex and are not illustrated in this dissertation. However, the CPTs of the variables have been considered below in results.

### 4 RESULTS

#### 4.1 Prior probability distributions of model variables

Based on the topology of the Bayesian Network model and on the inspection data the Genie software calculates the prior probability distribution of each node. Figure 4 shows the results of the model of the Thessaloniki port.

Figure 4 Bayesian Network model results of the Thessaloniki port

- The detention rate, according to the model, for vessels inspected at Thessaloniki port is 15%. This number is higher than the actual detention rate calculated from inspection data (9.33%).

- Concerning Ship Risk Profile, the probability of a vessel to be characterized as High Risk Ship (HRS) is very close to the actual one (29% to 26%). However, according to model results, 25% of the vessels are characterized as Low Risk Ships (LRS) while the number calculated from data is much lower, 4%.

- The probability of finding zero deficiencies after a Port State Control inspection at Thessaloniki port is 37% according to the model while calculated from data as 43%. On the other hand, the Bayesian Network model calculates the probability of finding 10 or more deficiencies during the inspection as 12% while the actual probability obtained from data is 9%.

Based on the dependent and independent relations between the model variables along with the one year inspection data at the port of Liverpool, the Bayesian Network model calculates the prior probability distri-
The detention rate is 3% as calculated from the model while the actual detention rate calculated from the inspections records is 1.05%. The two values are close, a good evidence about the reliability of the model developed.

Concerning Ship Risk Profile, the model calculates the probability of ship classified as “Low Risk Ship” as 25% while the actual probability is 11%. Moreover, according to the results obtained from the model, the probability of a vessel categorized as “Standard Risk Ship” is 75% which is much lower than the value calculated using Port State Control inspection data (89%). Finally, both model and data eliminate the probability of having a vessel classified as “High Risk Ship” inspected at Liverpool’s port.

Bayesian Network model calculates the probability of finding zero deficiencies as 35%, value that is very close to the actual probability calculated using the data (36%). Moreover, according to model result, the probability of finding 1 to 3 deficiencies is 45% while calculated as 49%. Finally, the model calculates the probabilities of finding 4 to 9 and over 10 as 18% and 3% respectively. The actual values calculated above are 13% and 1% respectively.

### 4.2 Posterior probability distributions of model variables

Given the results presented above, a further analysis for the Port State Control inspections can be done through the Bayesian Network model. Genie software allows calculating posterior probability distributions of each variable given evidences on some particular states of model nodes. Such analysis helps having a better picture about how each risk factor influences Port State Control inspection results in terms of number of deficiencies and detentions. Conclusions can help vessel owners take better decisions and avoid many deficiencies and detentions

- Evidence: Company Performance= “Low” at Thessaloniki port

Company Performance is a risk factor introduced with the implementation of the NIR and is calculated as described in Chapter 3. Given the fact that the company performance of vessel inspected is classified as “Low”, the posterior probability distribution of the number of deficiencies can be calculated.

Figure 6 presents the posterior probability distribution of the risk variable “Number of deficiencies”, given the fact that the company performance is “Low”, in comparison with the prior probability distribution.

The information that the company performance of the vessel inspected is classified as “Low”, according to Paris MOU criteria, decreases the probability of finding zero or 1 to 3 deficiencies during a Port State Control inspection. On the other hand, a significant increase in the probability of the vessel being found with more than 10 deficiencies is observed. Moreover, the probability of finding 4 to 9 deficiencies is increased too. Including to the model the evidence that the Company Performance is low changes the detention rate as well. Before adding the information, the detention rate was 15% while the new information updates the detention rate calculation to 29%.

- Evidence: RO “Medium” at Liverpool port

Figure 7 presents the posterior probability distribution, given the information that the Recognized Organization (RO) of the vessel inspected is classified as “Medium” performance.
The information that the RO performance of the ship inspected is categorized as “Medium” decreases the probability of finding zero or 1 to 3 deficiencies. On the other hand, the probability of having 4 to 9 deficiencies is significantly increased as well as the probability of finding more than 10 deficiencies. The detention rate, after including the evidence, is calculated as 10% while without the evidence was calculated as 3%.

4.3 Probabilistic analysis of deficiencies categories

As mentioned above, two separate models are developed to analyze how risk factors influence the different categories of deficiencies found in a Port State Control Inspection. The Bayesian Network models developed before calculated the probability of finding deficiencies of any category during a Port State Control Inspection at Thessaloniki and Liverpool ports. In an attempt to analyze how risk factors such as age, flag and RO influence the categories of deficiencies found in a Port State Control some evidences are included in the second set of models developed with detailed information on the type of deficiencies found in the two ports.

- Evidence: Old vessel with black flag and poor company performance at Thessaloniki port

Including evidences to the model in order to simulate a scenario in which an old vessel (age 30 to 35) with a flag classified as “Black” and a company performance categorized as “Low”, a new probability distribution of the variable “Category of deficiency” is obtained.

Figure 8 presents the posterior probability distribution of the categories of deficiencies, given the information described above, in comparison with the prior probability distribution (with no evidences). It is easily observed that the information about the age, the flag and the company performance of the inspected vessel leads to a significant decrease in the probability of having no deficiencies. The categories of deficiencies with the most important increase are “Labour Conditions” and “Safety of navigation” followed by “Certificate and Documentation”, “Lifesaving appliances” and “Emergency Systems”.

- Evidence: Bulk carrier with “white” flag, “high” RO and “high” company performance at Liverpool port

In the following scenario a bulk carrier with a flag classified as “white”, a RO categorized as “high” and a company performance classified as “high” is arriving for inspection at Liverpool port. Including the evidences mentioned above to the model, a new probability distribution of the categories of deficiencies is calculated. The ship, due to its low risk level, is not expected to have many deficiencies. However, using the BN model, the inspector can notice the categories that are more likely to be defective items.

Figure 9 presents the posterior probability distribution, given the evidences mentioned above, of the variable “Category of Deficiency” in comparison with the prior probability distribution.
Even though the classification of vessel’s flag, RO and company performance is good, there are specific categories of deficiencies with high probability. “Labour Conditions” is the most likely deficiency category followed by “Safety of Navigation”. Port State Control officers can use such models in order to predict which categories are more likely to be defective and give special attention to them during the inspection.

5 CONCLUSIONS AND FUTURE WORK

The Port State Control regimes have been developed to eliminate substandard ships that threatened the safety of life, property and the environment. Paris Memorandum of Understanding on Port State Control (Paris MoU) was established in 1982 by fourteen European countries. On the 1 January 2011 the Paris MoU introduced a New Inspection Regime (NIR). Under this New Inspection Regime the old Target Factor system is replaced by a risk-based one, the Ship Risk Profile.

The risk influencing factors that are used to define Ship Risk Profile along with the way it is calculated is analyzed above.

It has been shown that there is a growing interest in using Bayesian networks for probabilistic modelling of Port State Control inspections. The capability of representing rather complex, not necessarily causal but uncertain relationships makes Bayesian Networks a powerful tool for risk modelling. Especially, the data-driven approach on developing Bayesian Network models reduces the dependence on human experts and in some cases increases the accuracy of the model. In the present dissertation, four Bayesian Network models are developed using Port State Control inspection data from inspections carried out at Thessaloniki and Liverpool port during 2018. The main objective is to characterize probabilistically the influence of the risk influencing factors on the Ship Risk Profile and on the deficiencies identified in the inspection at the two ports. The process of including evidences helps to have a better picture of how individual risk factors influence number of deficiencies found and detentions on a Port State Control inspection. In addition to the results presented in Chapter 4, some useful conclusions can be drawn:

- Respectively young ships, under 10 years old, seem to have less deficiencies and not get detained. However, the age of the ship is not a major factor that can lead to significant change on the number of deficiencies found.
- Factors such as the “Company Performance” and the “RO performance” seems to have a major impact on the number of deficiencies found and detention rate among the other factors. Poor company or RO performance lead to more deficiencies found on a Port State Control inspection.

Two separate models are developed to analyze how risk factors influence the different categories of deficiencies found in a Port State Control Inspection. Such Bayesian Network models can be served as the prediction tool for port authorities under different conditions. When a vessel arrives at a port, the port authority can use the proposed prediction tool to predict which categories are more likely to be defective. This can lead to a more focused and effective inspection process.

Such scenarios are simulated by including evidences in the Bayesian Network models and some useful conclusion can be drawn:

- For a vessel considered to be low risk, a bulk carrier with a flag classified as “white”, a RO categorized as “high” and a company performance classified as “high”, inspected at Liverpool port, “Labour Conditions” and “Safety of navigation” are the two deficiency categories that are more possible to be defective.
- For a vessel considered to be high risk, a General cargo/multipurpose vessel with a RO classified as “Very Low”, a company performance categorized as “Low” and a flag classified as “Black”, inspected at Thessaloniki port, “Labour Conditions” and “Water/Watertight conditions” are the two deficiency categories that are more possible to be defective.

The present dissertation has shown the merits of Bayesian Network models for analyzing the influence of risk factors on the inspection results in terms of number and type of deficiencies. Moreover, Bayesian Network models can be used by port authority as prediction tools to predict which categories are more likely to be defective. However, the Port State Control
inspection data used for the development of the Bayesian Network models in this dissertation refer to one year and to two ports only, which limits the potential of the models. Based on the present dissertation, some improvements can be done. Further work should focus on the following aspects:

- Data used for the Bayesian Network model development should be referred to more than one year and to more than two ports of the Paris MOU region. A larger sample of Port State Control inspection records can lead to a better model and to more accurate inferences.
- To develop a ship risk index that extends the concept of the Ship Risk Profile, which is heavily dependent on generic static parameters and on the inspection history of the flag State of the ships, by incorporation additional risk influencing factors.

REFERENCES


Kjaerulf, Uffe B., and Anders L. Madsen. 2008. Bayesian Networks and Influence Diagrams. Springer US.


