

**TOMÁS RODRÍGUEZ GARCÍA**Email: [t.rodriguez@upm.es](mailto:t.rodriguez@upm.es)

Civil Engineering Department: Transport Infrastructure. ETSIC

Technical University of Madrid

C/ Alfonso XII, 3 y 5, 28014 Madrid, Spain

**NICOLETTA GONZÁLEZ CANCELAS**Email: [nicoleta.gcancelas@upm.es](mailto:nicoleta.gcancelas@upm.es)**FRANCISCO SOLER-FLORES**Email: [f.soler@upm.es](mailto:f.soler@upm.es)

Civil Engineering Department: Transport. ETSICCP

Technical University of Madrid

C/ Profesor Aranguren, S/N, 28040 Madrid, Spain

Transport Logistics

Preliminary Communication

Submitted: Nov. 23, 2014

Approved: Oct. 6, 2015

# SETTING THE PORT PLANNING PARAMETERS IN CONTAINER TERMINALS THROUGH BAYESIAN NETWORKS

## ABSTRACT

*The correct prediction in the transport logistics has vital importance in the adequate means and resource planning and in their optimisation. Up to this date, port planning studies were based mainly on empirical, analytical or simulation models. This paper deals with the possible use of Bayesian networks in port planning. The methodology indicates the work scenario and how the network was built. The network was afterwards used in container terminals planning, with the support provided by the tools of the Elvira code. The main variables were defined and virtual scenarios inferences were realised in order to carry out the analysis of the container terminals scenarios through probabilistic graphical models. Having performed the data analysis on the different terminals and on the considered variables (berth, area, TEU, crane number), the results show the possible relationships between them. Finally, the conclusions show the obtained values on each considered scenario.*

## KEY WORDS

*containerised traffic (trade); Bayesian networks; planning; forecast; port capacity;*

## 1. INTRODUCTION

As taken up in paper [29], one of the main issues in port logistics or other related freight transport engineering fields is the general forecast of those parameters related to space, means, and resources requirements, as well as their optimisation. Physical and equipment parameters related to a container terminal (i.e. stocking surface, necessary berthing length, dock cranes number,...) represent a high investment and are characterised by important social, economic or environmental impacts.

Therefore, a correct forecast of these parameters and of the actual surface requirements (least possible geographical impact, thus its least modification), leads the performed research to provide a highly useful tool to any planning agent, so it can anticipate and/or forecast its space and means needs, way before strategic, marketing or planning decision-making.

Up to this date, port planning has been rather based on empirical, analytical or simulation models. Empirical methods are based on productivity average indicators issued by planning agents. These indicators set a relationship between the main activities of a subsystem and the total annual production. These methods are thus very useful when dealing with new terminals planning or master plans development. The reference indicators have been constantly studied and updated by different authors over the years ([4, 10, 11, 16, 27, 33 and 34] among others). Analytical methods use mathematical concepts and formulas, based on the queuing theory and requiring large databases. These methods have been studied by several authors [27, 41 and 1]. Dragovic emphasized it in his paper "Port and container terminals modelling" [12]. The paper mentions several studies ([22, 23 and 26] among others), based on different aspects of the berthing system planning, as the occupation ratio, port congestion percentage, minimum waiting time, total port system costs, optimal number of berthing points and dock cranes, the optimal ratios berthing points/terminal or dock cranes/berthing points, etc. As indicated by the United Nations Conference on Trade and Development [40] simulation techniques use models to represent complex processes, whose mathematical description is not performable due to random behaviour and non-linear characteristics of the process. A detailed description of the method and the results of its appli-

cation to the Casablanca Port are included in a paper published by UNCTAD [41].

The USA [20], published a paper that performs a revision on the literature related to the capacity factors, focused on port planning. Another paper has been issued in Singapore [15], dealing with strategic planning issues.

Spanish bibliographical references started back in 1977 [32] with a paper stating the basics of port planning. Rafael Soler [34] would publish later on a comparison between exploitation conditions in several Spanish ports, using empirical methods. More recently, paper [6] presents the parameters and processes to be considered in container terminal planning. In 2007, in his Ph.D. thesis MN [17], González, M. González Cancelas determines the characteristic parameters and ratios of the port operation, obtaining their values for each container port terminal. Other papers on logistic planning could also be mentioned [27].

During the last decades, several data analysis and modelling techniques have been developed in the statistics and artificial intelligence fields [13, 2]). Data Mining (DM) is a modern and interdisciplinary area that joins together those techniques automatically operating (requiring minimum human involvement) and, besides that, it is highly efficient in processing huge quantities of information, like those available in several practical processes of data bases. The application of these disciplines is extended to a great number of commercial and research environments, when dealing with prediction, classification or diagnosis processes [27, 29, 30, 31, 9, 43, 44, 42 and 8] among others). Data mining uses different techniques, as probabilistic networks or Bayesian networks, allowing to model jointly all the relevant information for a given problem using probabilistic interference mechanisms to obtain conclusions based on the available evidences [18, 25, 35 and 7].

Bayesian networks are a compact representation of a multi-variant probability distribution. Formally, a Bayesian network is an acyclically conducted graph, whose nodes represent a random variable; the relationships between variables are coded by the graph's structure following the d-separation criterion. Each node has an associated probability distribution, conditioned by its origins, in such a way that the overall distribution can be expressed as the product of all conditioned distributions associated to the network's nodes. Thus, for  $n$  variables network  $X_1, X_2, \dots, X_n$  (Equation 1):

$$p(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i | x_{pa(i)}) \quad (1)$$

where,

$p$  is the probability distribution of node variables,

$p(x_i | x_{pa(i)})$  is a conditional probability matrix,

$x_i$  is the variable value in each case.

This technique's study allows a good global perspective of the statistic learning process and data mining, as well as a better understanding of other alternative techniques [17-13]. Bayesian networks are used preferably in the transportation systems to develop highway models, as by Sun, Zhang and Yu [35] who showed the use of Bayesian networks to perform traffic predictions, or as shown in [37]. The use of Bayesian networks can also be found in [3] where an integrated management of water resources has been developed, or in [5] dealing with planning improvement of natural resources.

The strength of Bayesian networks can be explained by the fact that, once the network structure is clearly defined, they allow any inference given the available information. Thus, predictive inferences can be performed (given the transport terminal surface  $X$ , you can know what is the probability of having  $Z$  cranes in the storage area), as well as abductive (if the terminal has less than  $X$  cranes in the storage area, one can know the probability of knowing its storage surface). In this way, each node can be at the same time a source of information or a prediction subject. Inferences are performed by applying probabilities propagation algorithms, specially developed to this purpose. The use of Bayesian networks requires identification of the variables and their relationships, and to quantify these relationships by assigning a-priori and conditioned probabilities.

## 2. METHODOLOGY

In order to determine the Bayesian network characterising the operation of the main international container ports, the following methodology has been developed, divided into two main tasks: one that defines the work environment and the second that develops the artificial intelligence model.

### 2.1 Work environment definition

It consists in reviewing the state of the art in order to identify the operation measurement variables' array of the containers maritime terminals, by means of specialized browsers and application managers. It is developed in two steps:

- Terminal variables definition and selection: all possible variables are studied.
- Assign a value to each variable for each studied terminal: different information sources are used to obtain the values to be assigned to each variable.

The variables (features) taken into account to calculate the a-priori and conditioned probabilities are the port variables. The number of variables (features) forms a table of  $2^n$  combinations, where  $n$  is the natural number that could be considerably high; this is a difficulty that can be solved by reducing the space of

initial representation, in such a way that, if there are superfluous variables, their use can be analysed according to their importance in the methodology.

The study has been carried out for the main international port terminals.

The initial data are the traffic volume. The main traffic parameter is given in yearly TEUs operated at the terminal. The research used the following variables: *berth, area, crane number, and TEU*.

These data were obtained from selected ports and international bodies such as UNCTAD, Drewry, JOC, IAPH and others.

## 2.2 Artificial intelligence model construction

Building a Bayesian network from data represents a learning process in two steps: the structural learning and parametrical learning [25]. The first step allows obtaining the Bayesian network structure, i.e. the dependence and independence relationships between the involved variables. In the second step, the required a-priori and conditioned probabilities are obtained for a given structure. The following chapters describe the variables discretisation, the model construction, the inference and the classification.

### 2.2.1 Variables discretisation

Having selected the variables to be studied, they have to be subsequently discretised to allow the model construction. Bayesian networks usually use discrete or nominal variables; in case they are not, they have to be discretised before constructing the model. Although there are Bayesian network models with continuous variables, these are limited to Gaussian variables and linear relationships. There are two main types of discretisation methods: (i) not supervised and (ii) supervised; thus, different discretisation types are studied and this will be an option of the developed software.

In order to apply the Bayesian networks to this study, the variables obtained during the work environment definition will be used. These variables are discrete, so the continuous ones have been discretised according to intervals determined by 25, 50 and 75 percentiles, reproduced in the following table:

Table 1 – Variables and percentiles used by the network

Variable	P25	P50	P75
Berth(m)	1,480	3,000	5,100
Area (ha)	31.8	80.6	208
Crane number	11	26.5	62
TEU number	524,791	1,892,231	6,709,818

### 2.2.2 Model construction

During this step of the process, the structural learning is equivalent to finding the relationships between variables, so that the Bayesian network topology or structure can be determined. According to the structure type, different structural learning methods can be applied: tree learning, poly-tree learning, multi-connected networks learning, measurements and search-based learning, dependence relationships based methods.

Bayesian networks are increasingly included in the supervised classification tasks, but not so in the planning activity related to ports. Based on the ideas expressed in [21, 24], and developed in [25], showing that probability expressions represented by Bayesian networks can be used to carry out classifications considering one special variable – the variable to be classified, predicted by a group of variables – in such a way that the obtained network structure can be used to predict its value by assigning values to the predictor variables, and then propagating the evidence introduced in the network, by calculating the a-posteriori probability in the node associated to the special variable, given the values of the rest of them.

To build the network, the Elvira software has been used, specially developed for Bayesian networks [14].

The computer has been used to study works in Windows 7 operating system, CPU Intel Core i7 3.4 GHz processor and 8 MB of RAM. However, this program does not require such a powerful computer, but if you have JAVA V.5 or later installed software.

The code Elvira uses its own format to encode the models, a reader-interpreter module for codified models, a graphical interface for network construction – with specific options for the case of canonical models, exact and approximate (stochastic) reasoning algorithms for both discrete and continuous variables, reasoning explanation methods, decision-making algorithms, model learning based on databases, networks fusion, etc.

### 2.2.3 Inference and classification

After constructing the models, the inference capacity is studied. A Bayesian network offers an inference system, in which – when new evidences on the state of certain nodes are encountered – their probability tables are modified and subsequently the new probabilities are diffused to the rest of nodes. The probability diffusion, or probabilistic inference, is the probability for some variables to be calculated, given the evidences on other variables. The probabilities before the introduction of evidences are known as a-priori probabilities; after introducing the evidences, the new diffused evidences are called a-posteriori probabilities.

In this phase the Bayesian networks characteristics in learning tasks are used. Each observed example will

modify the probability that the formulated hypothesis is correct (increasing or decreasing it). Thus, a hypothesis that would not fit within an example array is not completely discarded, but its associated probability is decreased.

These methods are robust to the possible noise in the training examples and to the possibility of having incomplete or possibly wrong data among these training examples.

Bayesian methods allow taking into account – in the hypothesis prediction – the a-priori knowledge or domain knowledge as probabilities.

The network built using the K2 algorithm is showed below (Figure 1). This algorithm was developed by Cooper and Herskovits (1992). It is a search algorithm that optimizes the probability of the given network database. Actually, what makes this algorithm is to find the most probable set of origins, using Bayesian metrics that accurately measures the probability given the data structure. This heuristic algorithm is based on a topological order that must be specified by the user. This will greatly reduce the search space, because the system makes a node that then orders another one not to be the origin.

The network obtained using inference allows to calculate the probabilities to be adopted by the different variables discretisation, and to obtain efficient scenarios for the container terminals; if the efficiency [39, 38], is understood as the capacity to reach the programmed objectives with a minimum mobilization of resources, thus optimizing them, different scenarios are obtained (Figure 2) corresponding to different crane number.

The different scenarios (Figure 3, 4, 5 and 6) correspond to discretisations of the variable crane number, S1 corresponds to allocate less than 11 cranes, S2 to allocate between 11 and 26 cranes, S3 between 26 and 62, and S4 to allocate more than 64 cranes.

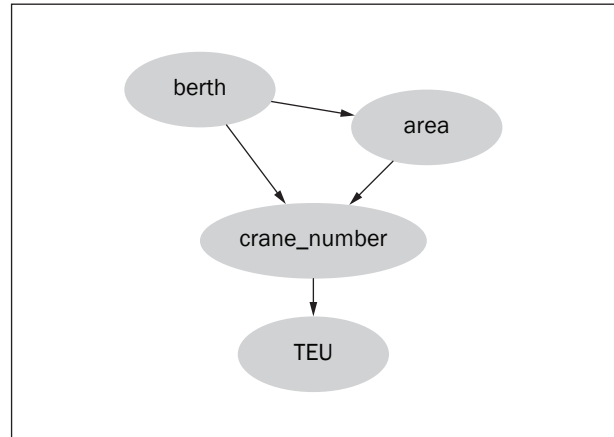


Figure 1 – Bayesian network. Algorithm K2.

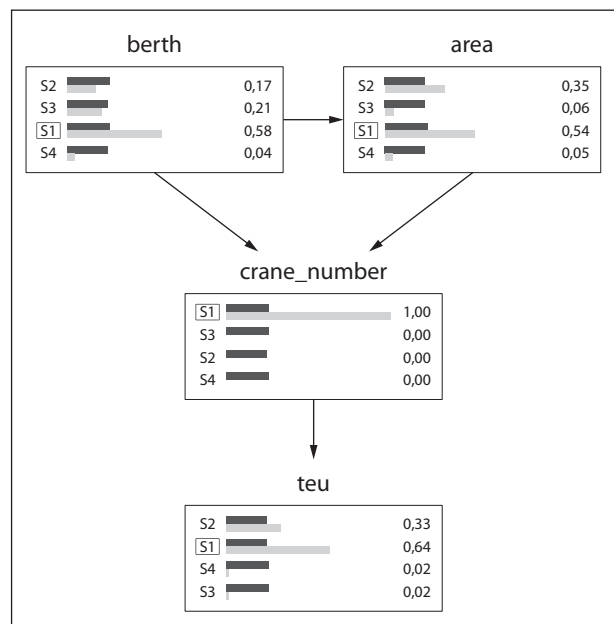


Figure 2 – Bayesian networks. Algorithm K2. Inference: crane number discretisation layer S1

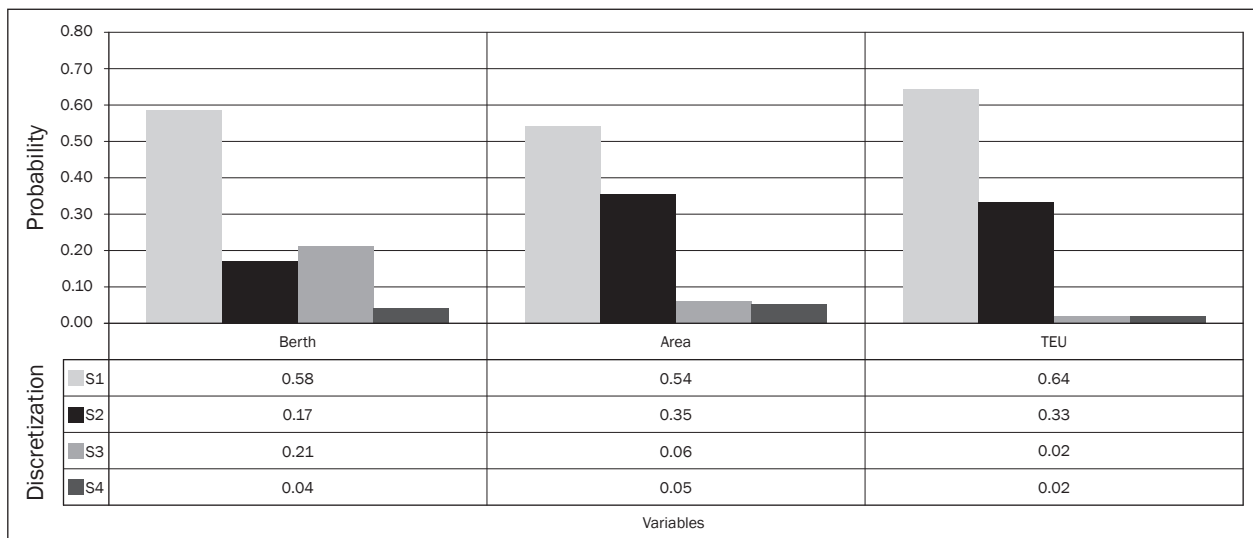


Figure 3 – Bayesian networks. Algorithm K2. Scenario 1 crane number discretisation layer S1

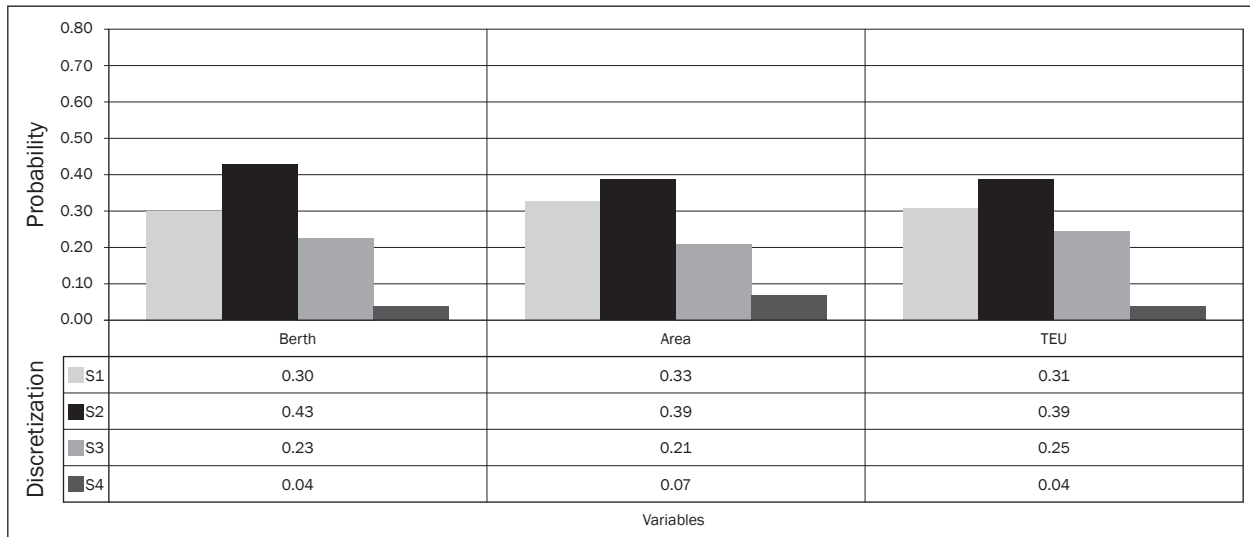


Figure 4 - Bayesian networks. Algorithm K2. Scenario 2 crane number discretisation layer S2

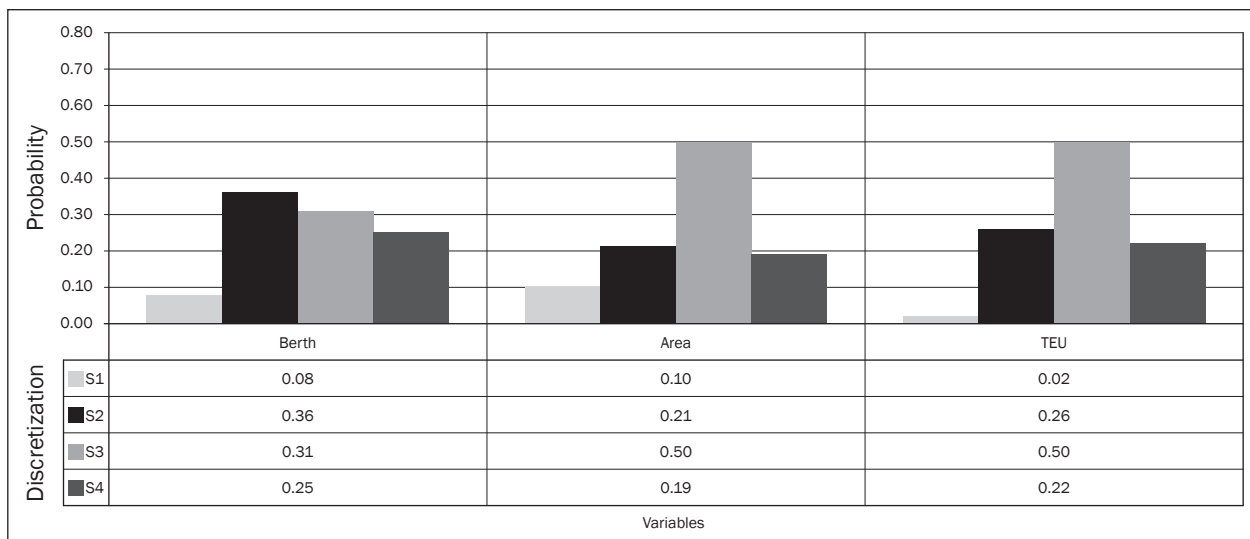


Figure 5 - Bayesian networks. Algorithm K2. Scenario 3 crane number discretisation layer S3

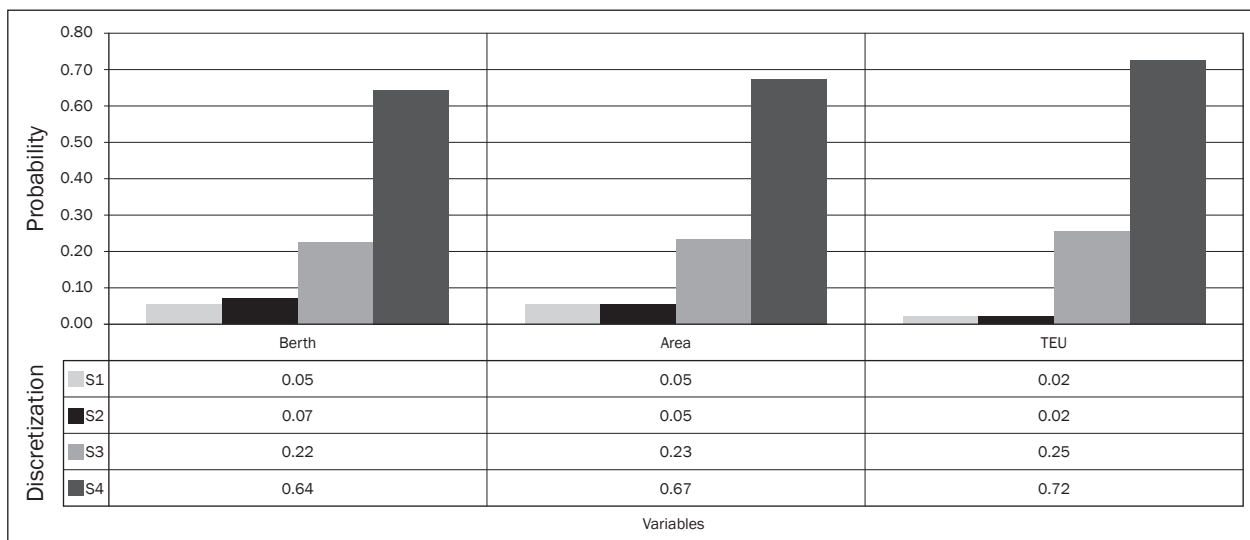


Figure 6 - Bayesian networks. Algorithm K2. Scenario 4 crane number discretisation layer S4

Each scenario with different probabilities adopted by each variable are described in the following Figures (Figure 3 for layer S1, Figure 4 for layer S2, Figure 5 for layer S3 and Figure 5 for layer S4). The discretisations have been carried out on intervals determined by 25, 50 and 75 percentiles, whose values have been described above.

Analyzing the results of Figures 3, 4, 5 and 6, one can see a good network performance, such as better behavior occurs on setting 4 layer S4 for a number greater than 62 cranes, where there is a probability above 60% to meet the discretization of layer S4 (berth > 5,100m, area >210 m, TEU >6.7 M.) .

### 3. RESULTS

Network topology or structure provides information on the probabilistic dependence between variables and their conditioned dependencies, given other variable(s), which is the purpose of this paper. This dependence simplifies the knowledge representation (fewer parameters) and the reasoning (diffusion of probabilities). Thus, the Bayesian Network provides a compact and modular way of representing the common distribution of several random variables. Bayesian networks are a compact representation of a multivariable probability distribution by a directed acyclic graph where each node represents a random variable and dependencies between variables are encoded in the structure of the graph at the discretion of separation.

A Bayesian Network comprises a qualitative section that describes the relationships between different variables and a quantitative section that describes the strength of these relationships by means of conditioned probabilities. Using these models allows obtaining the relationships between the variables associated to the dimensioning of a containers terminal and the terminal itself. These qualities of the Bayesian networks allow the relationships to be studied in order to identify the links between variables (berth, area, crane number), as shown subsequently.

#### 3.1 Relationship 1

The variable *berth* shall be the main decision variable in the planning process. It shows in the network as a variable: as a node generating all arches, by configuring a divergent connection, an origin that is linked towards several nodes. In other words, the outward arrows are divergent towards the nodes. When the state of origin variable is known, there is dependence between variables; an unknown state of the origin will produce independent variables and information diffusion is impossible when adding evidences on the other nodes. In this paper, the variables *crane number* and *area* are independent, given the variable *berth* (Figure 7).

#### 3.2 Relationship 2

In the convergent connections (also known as head to head) several variables target their arches towards a convergence variable. In this kind of connections, the origin variables are independent. Nevertheless, if there is an evidence of a generated variable, the origins will turn dependent. This study considers that *berth* and *area* are conditionally dependent, given the *crane number*. The important characteristic of this connection type when the information is diffused is that when an evidence is available on the convergence variable, the origin nodes will turn dependent and one evidence of the state of one of them will diffuse to the others (Figure 7).

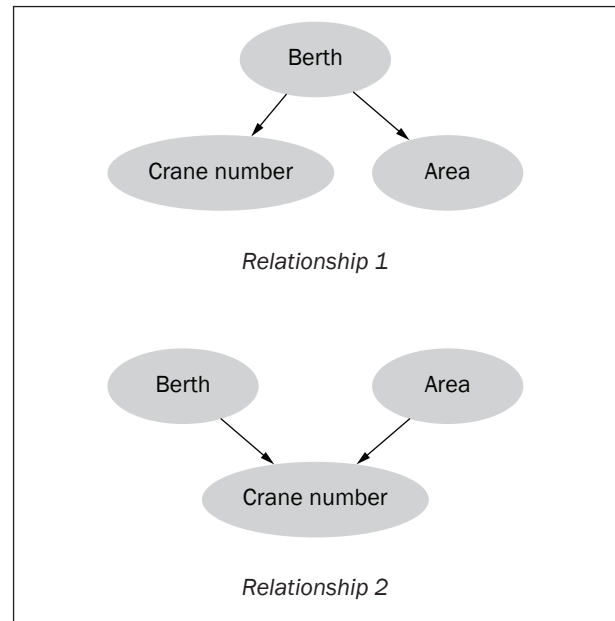


Figure 7 – Bayesian networks. Algorithm K2. Relationship 1 and 2 (berth, crane number, area)

#### 3.3 Relationship 3

The variable *berth* and the variable *area* are independent, given the variable *crane number*; it can be highlighted that the variable *AREA* depends on the variable *crane number*, and *crane number* depends on *berth*: this is equivalent to state that *berth* is the cause of *crane number* and that *crane number* is the cause of *area*. In this case, given the dependence between variables, if the information on the variable *berth* is known, the certainty of *crane number* state can be modified; also, when information is available on *crane number* state, the knowledge probability on *berth*'s state is modified. Nevertheless, if the *crane number* state is known, knowing some information about the variables *berth* and *area* will not modify the certainty of their state. It can be stated that information diffusion is blocked, and that the variables *berth* and *area* turn conditionally independent given the *crane number*; this is known as serial connection (Figure 8).

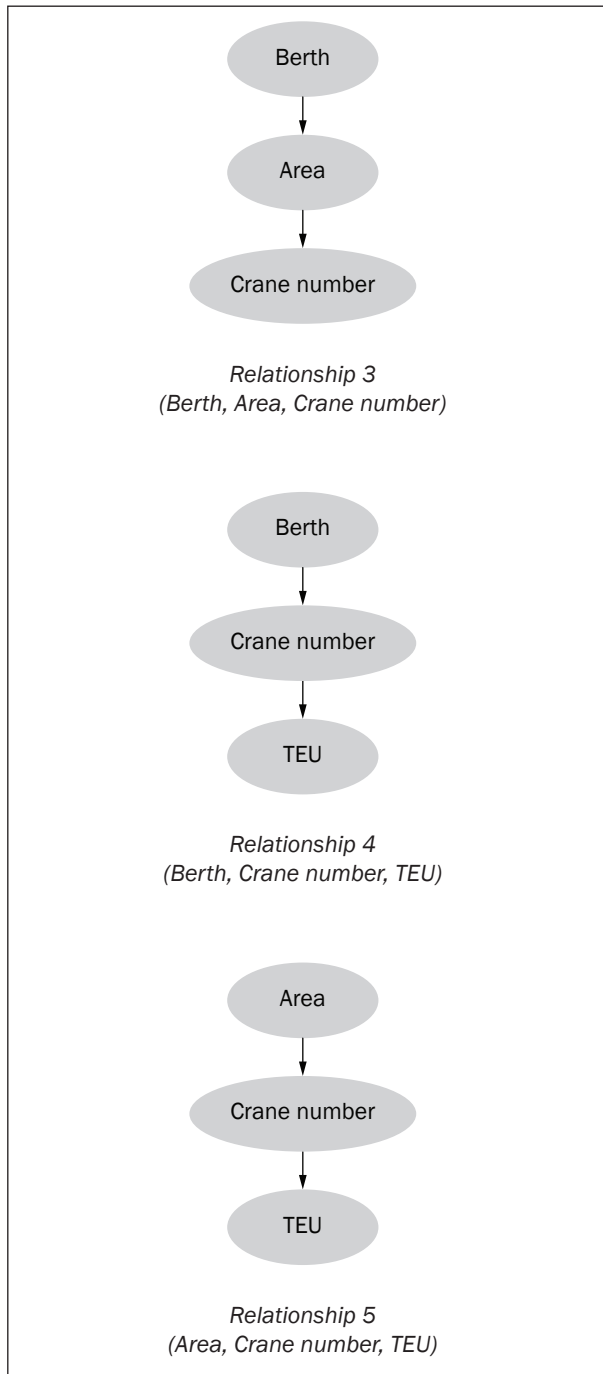


Figure 8 – Bayesian networks. Algorithm K2.  
Relationships 3, 4 and 5

### 3.4 Relationship 4

Variables *berth* and *TEU* are independent given the *crane number* (Figure 8), so the variable *TEU* depends on the variable *crane number*, and *crane number* depends on *berth*: this is equivalent to stating that *berth* is the cause of *crane number* and *crane number* is the cause of *TEU*. Knowing information on *berth* allows modifying the certainty on *crane number* state; also, if information is available on *crane number* state, the certainty on *berth* state is modified. Nevertheless, giv-

en the state of *crane number*, the information on the variables *berth* and *TEU* will not modify the certainty on their state. It can be stated that information diffusion is blocked, and that the variables *berth* and *TEU* turn conditionally independent given the *crane number*; this is known as serial connection.

### 3.5 Relationship 5

Variables *area* and *TEU* are independent given the variable *crane number*, so the variable *TEU* depends on the variable *crane number*, and *crane number* depends on *area*: this is equivalent to stating that *area* is the cause of *crane number* and *crane number* is the cause of *TEU*. Knowing information on *area* allows modifying the certainty on *crane number* state; and also, if information is available on *crane number*, the certainty on *area* is modified. Nevertheless, given the state of *crane number*, the information on the variables *area* and *TEU* will not modify the certainty on their state. It can be stated that information diffusion is blocked, and that the variables *area* and *TEU* turn conditionally independent given the *crane number*; this is known as serial connection.

## 4. CONCLUSIONS

The origin variable in international terminals planning is *Berth*, because its connections are totally outward.

Studying the scenarios according to the variable *crane number* leads to the conclusion that those terminals working with a few cranes – fewer than 11 – have a short berth length (about 1,500 meters), reduced areas generally of less than 30 hectares, and that they operate less than 500,000 TEUs, with a probability of about 60% for these combinations, situation that corresponds to Scenario 1 (Figure 3). The opposite situation corresponds to Scenario 4, (Figure 6), with more operating cranes (more than 60) and bigger berth length (more than 5,000), areas of more than 200 hectares and operating more than 6.7 million TEUs per year, with a probability between 65% and 72%.

Scenario 2 (Figure 4) and Scenario 3 (Figure 5) have lower probabilities. Scenario 2 represents operation of 11 to 26 cranes and has the same probability for different combinations, divided between the discretisations S1, S2 and S3. The probability that a terminal with this crane number may have a berth length of more than 5,000 m, nor an area of more than 200 hectares; nor operate more than 6.7 million TEUs is not considered.

Scenario 3 has homogeneous probability combinations divided between discretisation layers S2, S3 and S4; layer S1 was not considered because there is higher probability for a terminal operating 26 to 62 cranes to have an area between 80 and 200 hectares and a volume of about 1.9 to 6.7 million TEUs.

Among further research work in the field includes the following:

- Use other tools to analyse the models used in this research.
- Use other models and/or tools to analyse the results.
- Expand variables in the study and analyse new results that may be obtained in these same models.
- Analyse models with other port traffic (bulk, passengers,...).
- Expand data in the future years studied and ports that have been modified (upgrades, improvement works,...) to analyse the models and their behaviour in specific cases.

## ACKNOWLEDGEMENTS

I am sincerely thankful to Teodor Comeaga, for the help given so selflessly.

Finally, I would like to thank the Transportation Department of Civil Engineering Division in the Civil Engineering School of the Madrid Polytechnic University for their support.

### TOMÁS RODRÍGUEZ GARCÍA

E-mail: t.rodriguez@upm.es  
 Departamento de Ingeniería Civil: Infraestructura del Transporte. ETSIC  
 Universidad Politécnica de Madrid  
 C/ Alfonso XII, 3 y 5, 28014 Madrid, España

### NICOLETTA GONZÁLEZ CANCELAS

E-mail: nicoleta.gcancelas@upm.es

### FRANCISCO SOLER-FLORES

E-mail: f.soler@upm.es  
 Departamento de Ingeniería Civil: Transportes. ETSICCP  
 Universidad Politécnica de Madrid  
 C/ Profesor Aranguren, S/N, 28040 Madrid, España

## RESUMEN

*La correcta predicción en el ámbito de la logística de transportes, es de vital importancia para una adecuada planificación de medios y recursos, así como de su optimización. Hasta la fecha los estudios sobre planificación portuaria, se basan principalmente en modelos empíricos, analíticos o de simulación. El estudio refleja el alcance del posible uso de las redes bayesianas a la planificación portuaria. En la metodología se indica el escenario de trabajo y la construcción de la red considerada en el estudio para su aplicación en la planificación de terminales portuarias de contenedores, apoyado en las herramientas que proporciona el programa Elvira. Para el análisis de los escenarios de las terminales de contenedores y mediante el empleo de modelos gráficos probabilísticos, se han definido las principales variables y se ha realizado inferencia en escenarios virtuales. Una vez analizados los datos de las distintas terminales del estudio, así como las variables consideradas (berth, area, TEU, crane number), se muestran las posibles relaciones entre las citadas variables. Por último en las conclusiones, se muestran valores obtenidos con la red, para cada uno de los escenarios considerados.*

## PALABRAS CLAVE

*Tráfico de contenedores; Redes bayesianas; planificación; predicción; capacidad de puertos.*

## REFERENCES

- [1] Agerschou H. Facilities Requirements In Planning and Design of Ports and Marine Terminals. 2nd ed. London: Thomas Telford Ltd., 2004; p. 5-20.
- [2] Bishop CM. Pattern recognition and machine learning. New York: Springer-Verlag; 2006.
- [3] Bromley J, Jackson NA, Clymer OJ, Giacomello AM, Jensen FV. The use of hugin® to develop bayesian networks as an aid to integrated water resource planning. Environmental Modeling & Software. 2005;20(2):231-242.
- [4] Bureau of Transport and Regional Economics. Waterline. Issue no. 41. Australian Government, Department of Transport and Regional Services; 2006.
- [5] Cain J. Planning improvements in natural resource management. Guidelines for using bayesian networks to support the planning and management of development programmes in the water sector and beyond. Wallingford, Oxon: CEH Wallingford; 2001.
- [6] Camarero A, González MN, Pery P. Optimización y estudio de la capacidad de las terminales portuarias mediante modelos de simulación y explotación. Determinación de los niveles de servicio. UPM, UPV, Cenix, Valencia Port, Cedex, España. Código PT-2066-004-14IAPM; 2009.
- [7] Castillo E, Gutiérrez JM, Hadi AS. Expert systems and probabilistic network models. New York: Springer-Verlag; 1997.
- [8] Cios KJ, Kacprzyk J. Medical data mining and knowledge discovery. Heidelberg: Physica-Verlag; 2001.
- [9] Cios KJ, Pedrycz W, Swiniarski RW, Kurgan LA. Data mining: A knowledge discovery approach. New York: Springer; 2007.
- [10] Drewry, editor. World Container Terminals. Drewry Shipping Consultants; 1997.
- [11] Drewry, editor. Global Port Congestion: No Quick Fix. Drewry Shipping Consultants; 2005.
- [12] Dragovic B, Zrnic D, Radmilovic Z. Ports & Container Terminals Modeling. Research Monograph. Faculty of Transport and Traffic Engineering, University of Belgrade; 2006.
- [13] Duda RO, Hart PE, Stork DG. Pattern classification. New York Wiley; 2000.
- [14] Elvira Consortium. Elvira: An environment for creating and using probabilistic graphical models. In: Gámez JA, Salomón A, editors. Proceedings of the 1st European Workshop on Probabilistic Graphical Models (PGM'02). 2002; p. 222-230.
- [15] Fan HSL, Cao JM. Sea space capacity and operation strategy analysis system. Transportation Planning and Technology. 2000;24(1):49-63. DOI: 10.1080/03081060008717660.
- [16] Fourgeaud P. Measuring Port Performance. World Bank TWUTD, mimeo; 2000.
- [17] González, MN. Metodología para la determinación de parámetros de diseño de terminales portuarias de contenedores a partir de datos de tráfico marítimo [PhD thesis]. Madrid: Universidad Politécnica de Madrid; 2007.



- [18] González MN, Soler Flores F, Camarero Orive A. Modelo de eficiencia de las terminales de contenedores del sistema portuario español. *Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA*. Rect@. 2013;14:49-67.
- [19] Hastie T, Tibshirani R, Friedman JH. *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer-Verlag; 2009.
- [20] Maloni M, Jackson EC. North American Container Port Capacity. An Exploratory Analysis. *Transportation Journal*. 2005;44(2):16-36.
- [21] Michalski RS. A theory and methodology of inductive learning. *Artificial Intelligence*. 1983;20(2):111-161.
- [22] Nicolaou SN. Berth Planning by Evaluation of Congestion and Cost. *Journal of the Waterways and Harbors Division*. Proceedings of the American Society of Civil Engineers. 1967;93.
- [23] Nicolaou SN. Berth Planning by Evaluation of Congestion and Cost – Closure. *Journal of the Waterways and Harbors Division*. Proceedings of the American Society of Civil Engineers. 1969;95 (WW3):419-425.
- [24] Orallo JH, Quintana, MJR, Ramírez CF. *Introducción a la minería de datos*. Madrid: Editorial Pearson Educación SA; 2004.
- [25] Pearl J. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann; 1988.
- [26] Plumlee CH. Optimum Size Seaport. *Journal of the Waterways and Harbors Division*. Proceedings of the American Society of Civil Engineers. 1966;92:1-24.
- [27] Quijada-Alarcon J, González Cancelas MN, Camarero Orive A, Soler Flores F. Road network analysis using decision trees algorithm: A case of study of Panama. The 1st Virtual International Conference ARSA; 2012. Available from: <http://www.arsa-conf.com/archive/?vid=1&aid=2&kid=60101-242>
- [28] Rodríguez García T, González Cancelas MN, Soler Flores F. Forecast of container terminal capacity in a crisis scenario using Neural Network. *Virtual Conference*; 2013. Available from: <http://scieconf.com/archive/?vid=1&aid=3&kid=90101-38&q=f1>.
- [29] Rodríguez García T, González Cancelas MN, Soler Flores F. Setting the port planning parameters in container terminals through artificial neural networks. *Global Virtual Conference*; 2013. Available from: <http://www.gv-conference.com/archive/?vid=1&aid=2&kid=30101-9>.
- [30] Rodríguez García T, González Cancelas MN, Soler Flores F. Forecasting models in ports transport systems. 2nd Electronic international Interdisciplinary Conference; 2013. Available from: <http://www.eiic.cz/archive/?vid=1&aid=3&kid=20201-29&q=f1>.
- [31] Rodríguez García T, González Cancelas MN, Soler Flores F. The Artificial Neural Networks to obtain port planning parameters. XVIII Congreso Panamericano de Ingeniería de Tránsito, Transporte y Logística (PAN-AM 2014). *Procedia – Social and Behavioral Sciences*; 2014.
- [32] Rodríguez Pérez F. *Dirección y explotación de puertos*. Bilbao; 1977.
- [33] Schreuder M. Application of Approximate Performance Indicators for Master Planning of Large Ports. *Port Technology International*. 2005;26:19-22.
- [34] Soler R. Índices Portuarios Españoles. *Revista de Obras Públicas*. 1979;3166: 91-104.
- [35] Soler Flores F, González Cancelas MN, Camarero Orive A, Almazán Gárate JL, Palomino Monzón MC. Diseño de un modelo de planificación de zonas de actividades logísticas mediante el empleo de redes bayesianas. *Revista Ingeniería Industrial*. 2013;1:7-26.
- [36] Sun S, Zhang C, Yu G. A bayesian network approach to traffic flow forecasting. *Intelligent Transportation Systems, IEEE Transactions*. 2006;7(1):124-132.
- [37] Tebaldi C, West M. Bayesian inference on network traffic using link count data. *Journal of the American Statistical Association*. 1998;93(442):557-573.
- [38] Tongzon J, Heng W. Port privatization, efficiency and competitiveness: Some empirical evidence from container ports (terminals). *Transportation Research Part A: Policy and Practice*. 2005;39(5):405-424.
- [39] Tongzon JL. Determinants of port performance and efficiency. *Transportation Research Part A: Policy and Practice*. 1995;29(3):245-252.
- [40] UNCTAD, editor. *Desarrollo de los Puertos. Mejora de las Operaciones Portuarias e Instalaciones Conexas*. Nueva York: Naciones Unidas; 1969.
- [41] UNCTAD, editor. *Desarrollo Portuario. Manual de Planificación Para Países en Desarrollo*. Nueva York: Naciones Unidas; 1984.
- [42] Vityaev E, Kovalerchuk B. Empirical theories discovery based on the measurement theory. *Mind and Machine*. 2004;14(4):551-573.
- [43] Witten IH, Frank E. *Data mining: Practical machine learning tools and techniques*. Massachusetts: Morgan Kaufmann Pub; 2005.
- [44] Wong RCW, Li J, Fu AWC, Wang K. ( $\alpha, k$ )-anonymity: An enhanced k-anonymity model for privacy preserving data publishing. Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Philadelphia, USA: Association for Computing Machinery; 2006.