Looking for an efficient port planning: analysis of Spanish Port System through artificial intelligence

Beatriz Molina Serrano\textsuperscript{1,a}, Nicoleta González-Cancelas\textsuperscript{1,b}, Francisco Soler-Flores\textsuperscript{2,c}, Alberto Camarero-Orive\textsuperscript{1,d}

\textsuperscript{1}Civil Engineering Department, Transport, Universidad Politécnica de Madrid, Avda. Profesor Aranguren s/n 28040, Madrid, Spain

\textsuperscript{2}Mathematics and Computers Applied to Civil and Naval Engineering Department, Universidad Politécnica de Madrid, Avda. Profesor Aranguren s/n 28040, Madrid, Spain

\textsuperscript{a-d}E-mail address: beatriz.molinas@alumnos.upm.es, nicoleta.gcancelas@upm.es, f.soler@upm.es, alberto.camarero@upm.es

ABSTRACT

Nowadays, society claim for efficiency and efficiency of managed resources, so, making-decision people look for transparency and good planning practices. In particular case of ports, many variables are involved on it, so Bayesian networks are a useful tool to make an efficient planning because this kind of networks allow to know relationships between variables, even when they are a great number of them. In order to facilitate this work, the following study is carried out in which, through the construction of a Bayesian network, the Agent can know interesting information about how port variables are connected and can insert possible actions based on utility of their results, even when number of variables is high as occurs in planning and management of port terminals. There are a low number of studies related to this, so, our research includes more than 40 port variables, belonging to the four dimension of port sustainability. Main obtained conclusion shows that economic variables are the network parents in most of cases, so, their consideration is very important in planning decision-making. Their knowledge even allows to estimate probability of the rest of variables including in the network.

Keywords: port planning, decision-making, Spanish Port System, artificial intelligence, Bayesian networks
1. INTRODUCTION

Sustainable development is applying emergently by transport authorities and in other activity sectors and industries all over the World. It is proposed by initiatives which include environmental variables and social responsibility in strategic management of companies [1]. This is port case. Port sustainability is rooted in the proposals of the GRI [2] and it preserves the four main ideas or dimensions which knock into shape sustainable development – institutional, economic, environmental and social dimensions- So, it is necessary looking for an equilibrate development of these dimensions [3].

This concept is applying emergently by transport authorities and in other activity sectors and industries all over the World. It is proposed by initiatives which include environmental variables and social responsibility in strategic management of companies [1]. This is port case. Port sustainability is rooted in the proposals of the GRI [2] and it preserves the four main ideas or dimensions which knock into shape sustainable development – institutional, economic, environmental and social dimensions- (Figure 1).

![Figure 1. Four dimension of port sustainability. Source: Own production from reference [4]](image-url)

It is considered that sustainable management of a company or organism has as main target equilibrate keeping of its function and activity for a long time. So, it is necessary looking for an equilibrate development of economic, social, environmental and institutional dimensions [3].

In this context, maritime transport requires an especial attention, because it transports over 80% foreign trade (tonelates-kilometres) [5]. Two of the advantages of maritime transport
are the maintenance of transport and individual cost degressivity. They allow to say that this transport is cost-effective during long haul [6]. So, sustainable management must be understood in port sector as “management which allows containers traffic, solid and liquid granels, general goods and number of passenger grow up at the same time that energy and natural resources purchase, rubbish volume and negative impacts over social systems and ecosystems decreases in port influence areas” [7].

However, it is necessary that environmental and social variables gain importance to make port development travels to sustainability [8].

Spanish State Ports and Merchand Navy Law includes sustainability as one of the principles which must regulate planning model and port management. Each Port Authority must generate a sustainability report, which is considered as an analysis and diagnostic tool. However, this report does not determine Port Authority behaviour, it only describes results using performance quality indicators [9].

Quality indicator allow benchmarking sustainable management between different ports to define the best practices and to compare Port Authority performance and similar industry performance. Widespread application to a port system is useful to perform accurate benchmarking in sustainability between ports in the same region or country [10].

Application of these tools emerge objectives (economics, environmental, socials and institutionals) whose must be achieved by a port authority or port company to assure its sustainable development and its port growth [11].

Therefore, pursued objective is that through these four main sustainability ideas, ports conform as a system and not be seen as isolated entities and subject to a specific business situation, but as elements that interact with an environment physical, social and environmental, which are to be integrated effectively, that is, being able to adapt to a changing situation and in turn, pointing to a renewal that will help achieve the best possible future scenarios [12].

However, at operational level, the application of the concept is a critical issue because there is a strong methodological limitation. But, Bayesian networks form one of the possible methodologies.

3. ARTIFICIAL INTELLIGENCE IN PLANNING

Artificial neural networks (RNA) are new programming technologies based on computer systems for the analysis and study of learning and automatic processing sustained on the animal neural systems. One of the methodologies to be used are Bayesian Networks, from which you can get in a graphical form relationships between variables of each of four dimensions, in order to determine a posteriori values that quantify their contribution sustainability.

During the last decades, several data analysis and modelling techniques have been developed in the statistics and artificial intelligence fields [13,14]. Data Mining (DM) is a modern and interdisciplinary area that joins together those techniques automatically operating (requiring minimum human involvement) and, besides that, 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.
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 [24-27].

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.

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 [28,29]. Bayesian networks are used preferably in the transportation systems to develop highway models, as by Sun, Zhang and Yu [30] who showed the use of Bayesian networks to perform traffic predictions, or as shown in [31]. The use of Bayesian networks can also be found in [32] where an integrated management of water resources has been developed, or in [33] dealing with planning improvement of natural resources.

There are several studies that address the study of land transport using Bayesian networks in different environments, a transport area in which data mining has advanced the most [34-40]. In maritime transport, probabilistic models have been collected in various investigations [41-43]. Of the first works in applying Bayesian networks for exploitation and port planning is [44]. A similar study is [45] in which naive bayes is used with port traffic as the main variable. In more recent studies [46] Bayesian networks are used to find the logistic potential of a country, in this case the Republic of Panama. In the case of [47], main variables were defined and virtual scenarios inferences were realized in order to carry out the analysis of the container terminals scenarios through probabilistic graphical models. In [48], in order to determine the relationships between all the variables involved in the decision, giving us the importance of each factor and variable, we built a K2 BN algorithm. To obtain the scores of each variable, they used a complete cartography analyzed by ArcGIS. Continuing in a port environment [49] presents an innovative approach towards integrating logistic regression and Bayesian networks (BNs) into maritime risk assessment. In the investigation of [50] the authors are able to establish a model of planning zones of logistical activities by using Bayesian networks.

4. RESULTS

To build the network, the Elvira software has been used, specially developed for Bayesian networks [51]. 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.

Selected variables in the study are listed in Table 1, which is included below:
Table 1. Port factors which have been included in the research classified in sustainability dimensions

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Notion</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental dimension</td>
<td>Environmental quality and management</td>
<td>Environmental</td>
</tr>
<tr>
<td>gestamb_dimma</td>
<td>Environmental management</td>
<td>Environmental</td>
</tr>
<tr>
<td>calaire_dimma</td>
<td>Air quality</td>
<td>Environmental</td>
</tr>
<tr>
<td>calagua_dimma</td>
<td>Water quality</td>
<td>Environmental</td>
</tr>
<tr>
<td>calacust_dimma</td>
<td>Sound quality</td>
<td>Environmental</td>
</tr>
<tr>
<td>residuos_dimma</td>
<td>Rubbish management</td>
<td>Environmental</td>
</tr>
<tr>
<td>ecoef_dimma</td>
<td>Eco-efficiency</td>
<td>Environmental</td>
</tr>
<tr>
<td>comport_dimma</td>
<td>Port community</td>
<td>Environmental</td>
</tr>
<tr>
<td>sitecofin_dimecon</td>
<td>Economic and financial situation</td>
<td>Economic</td>
</tr>
<tr>
<td>inv_dimecon</td>
<td>Level and structure of investments</td>
<td>Economic</td>
</tr>
<tr>
<td>negserv_dimecon</td>
<td>Business and service</td>
<td>Economic</td>
</tr>
<tr>
<td>vgenprod_dimecon</td>
<td>Generated value and productivity</td>
<td>Economic</td>
</tr>
<tr>
<td>herramgestion_dimins</td>
<td>Tool management support</td>
<td>Institutional</td>
</tr>
<tr>
<td>geninfraortuaria_dimins</td>
<td>Port infrastructure generation</td>
<td>Institutional</td>
</tr>
<tr>
<td>mercservidos_dimins</td>
<td>Served Markets</td>
<td>Institutional</td>
</tr>
<tr>
<td>dinamact_dimins</td>
<td>Port sector role as revitalizing of port activity</td>
<td>Institutional</td>
</tr>
<tr>
<td>serviciosconcauto_dimins</td>
<td>Services and licenses of concessions/authorizations</td>
<td>Institutional</td>
</tr>
<tr>
<td>inicprivada_dimins</td>
<td>Private initiative presence</td>
<td>Institutional</td>
</tr>
<tr>
<td>transconcu_dimins</td>
<td>Transparency and free competition</td>
<td>Institutional</td>
</tr>
<tr>
<td>calidserv_dimins</td>
<td>Service quality</td>
<td>Institutional</td>
</tr>
<tr>
<td>intettrans_dimins</td>
<td>Port integration in transport system</td>
<td>Institutional</td>
</tr>
<tr>
<td>caphum_dimsoeci</td>
<td>Human capital of port activity</td>
<td>Social</td>
</tr>
<tr>
<td>empl_dimsoeci</td>
<td>Employment and job security in the port community</td>
<td>Social</td>
</tr>
</tbody>
</table>
Variable values uses to construct Bayesian network model correspond to data from the Sustainability Reports published annually by Spanish Port Authorities and they have been supplemented by information provided by the Public Agency State Ports, which corresponds to the historical record since the year 2010 - counting on almost 3000 records. Using K2 algorithm, obtained Bayesian networks can be seen in Figure 2.

![Bayesian network diagram](image)

**Figure 2.** Obtained Bayesian network

Topology needs to identify factors that are relevant, to determine how those factors are causally related to each other. The arc cause-effect does mean that cause is a factor involved in causing effect. In this case, for example, parent of the variables calacust_dimma and inv_dimecon is the node tranocu_dimins. Then, when tranocu_dimins is known, calacust_dimma and inv_dimecon, are conditionally independent (Figure 3).
Then, trancocu_dimins is a resolution variable which appears in network as a “node”. Some arcs started on it, so this variable generates a divergent connection. This way, trancocu_dimins is a parent node which projects arcs to several sons, that is to say, arrows start in this variable and diverges to its sons (Figure 3). As it can be remembered: trancocu_dimins, timestamp transparency and free competition, it’s Initiatives to ensure that any operator wishing to provide services in port or to qualify for a concession to hear a transparent manner the conditions to operate port and administrative mechanisms governing this process.

When parent variable state is known, there is a dependence relationship between variables. However, when a parent state is unknown, son variables are taken in an independent way and information will not spread along network if some evidences are included over son nodes (Figure 3). An effect that has two or more ingoing arcs from other vertices is a common effect of those causes. A cause that has two or more outgoing arcs to other vertices is a common cause (factor) of those effects. The effects of a common cause are usually observables.

Therefore, transparency and free competition in initiatives promoted directly affect: main emission sources of port that involve significant noise emissions; and level and structure of investments. It would be said that trancocu_dimins is a parent of calacust_dimma and inv_dimecon, calacust_dimma and inv_dimecon are childs of trancocu_dimins, that trancocu_dimins influences, or causes calacust_dimma and inv_dimecon, calacust_dimma and inv_dimecon depends on trancocu_dimins (Figure 4).

Considering the situation in Figure 4, geninfraortuaria_dimins has an influence on dinamact_dimins, which in turn has an influence on caphum_dimsoci. Obviously, evidence about geninfraortuaria_dimins will influence the certainty of dinamact_dimins, which then influences the certainty of caphum_dimsoci. Similarly, evidence about caphum_dimsoci will influence the certainty of geninfraortuaria_dimins through dinamact_dimins. On the other hand, if the state of dinamact_dimins is known, then the channel is blocked, and geninfraortuaria_dimins and caphum_dimsoci become independent; we say that geninfraortuaria_dimins and caphum_dimsoci are d-separated given dinamact_dimins. When the state of a variable is known, we say that the variable is instantiated. We conclude that evidence may be transmitted through a serial connection unless the state of the variable in the connection is known.
Figure 4 shows a causal model for the relations between geninfraortuaria_dimins, dinamact_dimins and serviciosconcauto_dimins, or the relations between geninfraortuaria_dimins, dinamact_dimins and inicprivada_dimins. For example, if I have not observed the dinamact_dimins, then knowing that there has been inicprivada_dimins will increase my belief that the dinamact_dimins is high, which in turn will tell me something about the geninfraortuaria_dimins. The same line of reasoning holds in the other direction. On the other hand, if I already know the dinamact_dimins, then knowing that there has been inicprivada_dimins will not tell me anything new about geninfraortuaria_dimins.

Following the Bayesian Network independence assumption, several independence statements can be observed in this case, in respect to each of the factors. When dinamact_dimins is known, caphum_dimsoci, inicprivada_dimins, gestamb_dimma and serviciosconcauto_dimins are conditionally independent of its ancestor’s geoinfraortuaria_dimins. Casual graph: The variable dinamact_dimins has four common effects caphum_dimsoci, inicprivada_dimins, gestamb_dimma and serviciosconcauto_dimins (Figure 4).

Management systems supporting decision-making includes quality management systems, scorecards, market characterization campaigns, etc., and they are represented by herramgestion_dimins. It is considered as a parent variable in network, so arrows only start on it. The same goes for transconcu_dimins which represents initiatives to ensure that any operator could provide services in port or qualify for a concession because operator can know, in a transparent way, conditions to operate and administrative mechanisms governing this process. It is a father node in network. Furthermore, transcocu_dimins is decision-making variable, so it appears in network as a “node” variable and bows only start on it. So, a divergent connection is created and this father node throws its bows toward several of its sons, that is to say, arrows start on it and go to its sons.

Other essential variable in network structure is mercservidos_dimins (Figure 5). 10 arrows star on it and go to 10 different nodes. These are effects of structure and main good traffic evolution, so they are social, economic, institutional and environmental effects. That is to say, served markets have effects on rates, delivery framework and regulation of port.
services the number of companies operating in the port (institutional category). It has effects on EBITDA, EBITDA/tonne, public investment relative to cash flow and income from employment and activity rates among others too (economic category). Even, about social status, it has effects on variables representing port community employment, job security and training services and health work, among others. Finally, in environmental category, served markets causes different grades of environmental management systems implementation (EMAS, ISO 14001 y PERLS), economic resource investment and investments associated to implementation, certification and maintenance of environmental management system. Therefore, served markets are a very important variable in planning from a sustainable perspective.

**Figure 5.** List of mercservidos_dimins sons
Evidence may only be transmitted through the converging connection. If either calaire_dimma or one of its descendants has received evidence (Figure 6). A description of the situation in Figure 3 requires a little more care. If nothing is known about calaire_dimma except what may be inferred from knowledge of its parents inicprivada_dimins and calidserv_dimins, then the parents are independent: evidence about one of them cannot influence the certainties of the others through calaire_dimma. Knowledge of one possible cause of an event does not tell us anything about the other possible causes. However, if anything is known about the consequences, then information on one possible cause may tell us something about the other causes.

**Figure 6.** Partial views of the network. Relationship 3

First, if no evidence is available about the state of calaire_dimma, then information about the state of inicprivada_dimins will not provide any derived information about the state of calidserv_dimins. In other words, inicprivada_dimins is not an indicator of calidserv_dimins, and vice versa (again, of course, assuming correctness of the model). Thus, contrary to serial and diverging connections, a converging connection will not transmit information if no evidence is available for the middle variable. This fact is illustrated in Figure 6.

Converging connection with no evidence on calaire_dimma or any of its descendants. Information about inicprivada_dimins will not affect our belief about the state of calidserv_dimins and vice versa.

Second, if evidence is available on calaire_dimma, then information about the state of inicprivada_dimins will provide an explanation for the evidence that was received about the state of calaire_dimma, and thus either confirm or dismiss calidserv_dimins as the cause of the evidence received for calaire_dimma. The opposite, of course, also holds true. Again, contrary to serial and diverging connections, converging connections allow transmission of information whenever evidence about the middle variable is available.

The rule illustrated in Figure 6 tells us that if nothing is known about a common effect of two (or more) causes, then the causes are independent; i.e., receiving information about one of them will have no impact on the belief about the other(s). However, as soon as some evidence is available on a common effect the causes become dependent. inicprivada_dimins and calidserv_dimins are conditionally dependent if calaire_dimma is observed.

If it already knows that air quality went off, knowing number of companies that operate in a port will increase general belief in that initiatives aimed at improving efficiency which are promoted by the Port Authority will be better. Further knowing let increase number of initiative promoted by Port Authority to improve efficiency, quality of service and performance of services provided to the merchandise. Thus, this increase belief in growth of
number of companies operating in the port, land occupied - characterized as commercial use concessioned -, etc.

In Figure 7, we observe only empl_dimsoci indirectly through information about residuos_dimma; knowing the state of residuos_dimma tells us something about the state of calagua_dimma, which in turn tells us something about empl_dimsoci. The conclusion is that evidence may be transmitted through a converging connection only if either the variable in the connection or one of its descendants has received evidence.

![Diagram](image)

**Figure 7.** Partial views of the network. Relationship 4

Evidence about a variable is a statement of the certainties of its states. If the variable is instantiated, we call it hard evidence; otherwise, it is called soft. In the example above, we can say that hard evidence about the variable residuos_dimma provides soft evidence about the variable empl_dimsoci. Blocking in the case of serial and diverging connections requires hard evidence, whereas opening in the case of converging connections holds for all kinds of evidence.

If we know nothing of calagua_dimma or empl_dimsoci, then the information on whether geoinfraortuaria_dimins will not tell us anything about mercservidos_dimins. However, if we have noticed something about calagua_dimma, then the information about geoinfraortuaria_dimins will make us to believe something about mercaservidos_dimins.

If neither empl_dimsoci nor any of its descendant are observed, geoinfraortuaria_dimins and mercaservidos_dimins are independent. Information cannot be transmitted through empl_dimsoci among parents of empl_dimsoci. It leaks down empl_dimsoci and its descendants. If empl_dimsoci or any of its descendant is observed, geoinfraortuaria_dimins and mercaservidos_dimins are dependent. Information can be transmitted through empl_dimsoci among parents of empl_dimsoci if empl_dimsoci or any of its descendants are observed. Observing empl_dimsoci or its descendants opens the information path.
3. CONCLUSIONS

Through the construction of a Bayesian network the relations between the different sustainability variables in a port environment are known. The most decision-making category, as network obtained by using the algorithm K2 shows, is institutional category, then economic and social at the same height, and finally environmental category.

In other hand, institutional variables are interconnected. Economics ones are important as cause-effect because they are effects of served markets which belong to institutional dimension. Generated value and productivity depend on kind of business and service (fee income occupation and activity, commercial use of the surface, use of docks, etc.). Moreover, social variables are effects of institutional variables but they have not a direct relationship with their same dimension, social one. Finally, environmental variables are closely interconnected in Bayesian Network and they are principally effects of institutional category.

Therefore, economic, social and environmental variables are effects of institutional ones. As a conclusion, key issue is that Port Authorities start to incorporate sustainable elements- included in Port Law- in their tools, used to regulate port services and public possession management.

Using relationships between variables that describe each of the four axes of sustainability, obtained using Bayesian networks, it has been determined that these axes cannot be considered as isolated entities and subjected to a concrete commercial conjuncture. They are interrelated with a physical, social and environmental environment, in which they must be effectively integrated to make ports create a system. This system lets ports are able to adapt to a changing conjuncture and at the same time and to point to a renewal that contributes to achieve the best of the possible future scenarios. It is only possible through knowledge of the relationships between different variables.

Considering current port system structure, institutional variables are those that take greater weight in achievement of sustainability objectives, because dependence relationship between these variables that has been obtained using a Bayesian network model developed. For this reason, port system will be able to act on sustainability variables if it acts on environmental dimension variables. To see possible effects, it is necessary to know about what parameters Port Authority can act from the institutional point of view. The basic functions of the port authority are the planning, projection, construction, conservation and operations of the works and services of the port, collaboration with official bodies, and coordination of the port private companies and management of the port domain.

Relationships between private initiative and public body are decisive in achievement of sustainability objectives, using institutional dimension. In regards to the different types of investments (Port-City, safety, environment or business promotion), its right identification and analysis presents a challenge. Looking into the future, it will become necessary to go deeper into defining its criteria, due to the large number of concepts that intervene in this type of actions and the confusion that this involves for its proper analysis and report. In regards to the quality of the services it is also worth mentioning that some authorities count with mechanisms to boost quality improvements and competitiveness of their services, and with mechanisms to assess their quality. It is worth to emphasize the wide group of social, financial and administrative nature collectives that are affected by the activity of the Port Authorities and that themselves affect the development and performance of such Port Authorities’ activities. Because of their institutional commitment, most of the ports identified...
their expectations and define communication or participation frameworks with each of those groups.

Staff competencies correspond to governmental bodies of the Port Authorities, in other words, to Management Boards, without more limits than those regulated by Labour and Budget Regulations. Social dimension is mainly based on the human resources policy, including structure of workforce, training actions developed under competency management scheme, equality plan and the efforts made in safety and health material.

Therefore, if a Bayesian network is built in a port environment which are based on the four pillars of sustainability, a tool that lets actuate on sustainability of global port system will be obtained because relationships between different variables are known. This tool is very necessary because the environmental management is clearly constrained by the exploitation scheme public-private. The port’s environmental efficiency does not exclusively rely on the Port Authority, but also on how rigorous the concessions, service providers and port users are regarding environmental management. However Port Authority does not have environmental competencies, it develops a key role in the adequate environmental management of the port (it acts as infrastructure administrator, regulator, etc.).

Despite great efforts made by Port Authorities by demanding these enterprises to be in possession of a certification, most of the Marpol, stowage or technical service delivery companies do not count yet with an EMS. Another challenge is extending the training regarding environmental matters, due to the fact that only 35 % of the personnel took it. In regards to air and water quality, Port Authorities actively contribute launching different initiatives to reduce dust and particle emissions and improve the quality of water and the acoustic quality.

References


Grupo De Trabajo 23. La sostenibilidad en los puertos. CONAMA VII Cumbre del desarrollo sostenible, 24 de noviembre de 2004

C. Crespo Soler, V.M. Ripoll Feliu; A.M. Crespo Trujillo, A. Giner Fillol. La sostenibilidad ambiental en el sistema portuario de titularidad estatal. XIII Congreso AECA. Armonización y Gobierno de la Diversidad. 22-24 Septiembre 2005

F. González, A. Guerra, F. Martín, J.J. Nóvoa, C. Otero, y J. Penela. Medición de la sostenibilidad en el sistema portuario Español: propuesta metodológica a través de indicadores sintéticos de desarrollo sostenible. XII Reunión de economía mundial, mayo de 2010. Santiago de Compostela


Puertos Del Estado. Memoria de sostenibilidad del sistema portuario de interés general, 2011


J.A. Moscoso, J.J. Ruiz, M.M. Cerbán, Técnicas de predicción en el tráfico Ro-Ro en el nodo logístico del Estrecho de Gibraltar, IAME 2011 conference


(Received 28 July 2017; accepted 14 August 2017)