



Article

Use of Bayesian Networks to Analyze Port Variables in Order to Make Sustainable Planning and Management Decision

Beatriz Molina Serrano ^{1,*}, Nicoleta González-Cancelas ¹, Francisco Soler-Flores ², Samir Awad-Nuñez ³ and Alberto Camarero Orive ¹

¹ Departamento de Ingeniería Civil, Transportes, Universidad Politécnica de Madrid, 28040 Madrid, Spain; nicoleta.gcancelas@upm.es (N.G.-C.); alberto.camarero@upm.es (A.C.O.)

² Departamento Tessella-Altran World Class Center for Analytics, Altran Innovación, 28022 Madrid, Spain; francisco.solerflores@altran.com

³ Departamento de Ingeniería Civil, Universidad Europea de Madrid, Madrid 28670, Spain; samir.awad@universidadeuropea.es

* Correspondence: beatriz.molinas@alumnos.upm.es; Tel.: +34-913-366-428

Received: 26 October 2017; Accepted: 13 December 2017; Published: 11 January 2018

Abstract: In the current economic, social and political environment, society demands a greater variety of outcomes from the public logistics sector, such as efficiency, efficiency of managed resources, greater transparency and business performance. All of them are an indispensable counterpart for its recognition and support. In case of port planning and management, many variables are included. Use of Bayesian Networks allows to classify, predict and diagnose these variables and even to estimate the subsequent probability of unknown variables, basing on the known ones. Research includes a data base with more than 40 variables, which have been classified as smart port studies in Spain. Then a network was generated using a non-cyclic conducted grafo, which shows port variable relationships. As conclusion, economic variables are cause of the rest of categories and they represent a parent role in the most of cases. Furthermore, if environmental variables are known, subsequent probability of social variables can be estimated.

Keywords: Bayesian Networks; graph theory; sustainability; port management; artificial networks

1. Introduction

Sustainable development is being applied emergently by transport authorities as it has been made in other activity sectors and industries all over the World. So, transport must include environmental variables and social responsibility in the strategic management of companies [1,2] because it is considered as a ‘blood system’ of global economies. In case of ports, port sustainability is rooted in the proposals of the Global Reporting Initiative (GRI) [3]. It preserves the four main ideas or dimensions that knock into shape sustainable development. Consequently, it is necessary looking for an equilibrate development of economic, social, environmental, and institutional dimensions [4]. However, global warming increase and climate change during the past decade, has created a greater emphasis on the environmental management issue [5].

In this context, maritime transport requires an especial attention, because shipping operations contribute to the international trade activities growth [6,7]. It transports over 80% foreign trades (TON-Km) [8,9]. So, sustainable management must be understood as “management which allows containers traffic, solid and liquid granel, general goods and number of passenger grow up while energy and natural resources purchase, rubbish volume and negative impacts over social systems and ecosystems decrease in port influence areas” [10].

However, one of the principal challenges to integrate sustainable criteria in management model and port development is broken out present inertia. This inertia considers economic factor as unique development variable. It is necessary that environmental and social variables gain importance to make port development travel to sustainability [11]. Ports are infrastructures for which the definition of performance is not easy because port operation are carried out thorough complex systems [12,13]. Each port is integrated by several interrelated subsystems that provide services to the ships, and to the companies that send/receive cargoes through maritime transport. Terminal planning is carried out in the medium and long term in an efficient exploitation and it must include a systemic study [7]. In none of the port terminal subsystems should be a bottlenecks that hinder the terminal operation.

Each operative subsystem capacity is conditioned by diverse variables. Some of them are endogenous (own competence) and others exogenous. From the broad group of variables that affect capacity, the following may be highlighted as primary: exogenous variables (regularity of work and rationality of work) and endogenous variables (terminal dimension and productivity). Remember that sustainability concept is a highly theoretical concept and it is very sensitive to the management and operation of each port. So the capacity obtained from each terminal must be analyzed, studying at the same time the exploitation that takes place on it.

Nowadays, ports have ceased to be simply an infrastructure where operations are carried out to constitute important logistics operations, which have a significant economic relevance in the environment in which they are located. Port planning aims to be a management tool that guarantees, through decision-making, a certain success degree for the studied port. This way, environmental conditions are considered in order to do so with the development of this application to provide a tool to planners. Before planners had limited their options to a mere indicator study.

Spanish Port Law includes sustainability as one of the principles that must regulate planning models and port management. Article 55.4 foresees company planning of each Port Authority must be gone with a sustainability report. It is considered as an analysis and diagnostic tool which is based on GRI's one. However, it includes specific quality indicators and must be approved by Spanish State Ports and Port Authorities. Redaction of these reports is an important effort because it integrates information about Port Authority operation and information about its environmental, economic and social performance. However, this report only describes results using performance quality indicators [14].

These indicators allow an evaluation of sustainable development management. So, they allow to control Port Authority management in an objective way. Consequently, these indicators must cover four dimensions of sustainability. Implementation of these quality indicators is useful for Port Authorities to control their sustainable management and, even, 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 [15].

Application of these objectives must be achieved by a Port Authority or port company to assure its sustainable development and its port growth [16]. Therefore, the pursued objective is that ports conform as a system and not be seen as isolated entities through these four main sustainability ideas. They must be considered as elements that interact with an environment physical, social, and environmental [17]. Main objectives of each port sustainability dimension are included in Figure 1.



Figure 1. Sustainability dimensions: main objectives. Source: [18].

The literature about port industry efficiency is relatively new, because the first investigations appeared in the mid-90 s. It is also poor, especially when compared to studies carried out in other public services (including the transport sector) where most of the publications correspond to publications related to the railway and air sectors [19].

In the last ten years, there has been an important advance in the work that analyzes the efficiency and productivity of the port sector. The processes of technological innovation that occurred in the maritime and port industries and the changes in the organization and port management have conditioned a change in the nature of the operations. It favored a greater factor specialization. These events have had a great impact on the productivity and operation efficiency. Consequently, Law 33/2010, of August 5, on economic regime and provision of services in ports of general interest, established that: the competitiveness of our productive system is conditioned by the efficiency and efficiency of the ports. It also adds the requirement to adopt measures in Spain that improve the management of our ports and their efficiency, boosting their competitiveness in a situation of strong competition international.

So that, it has been investigated using probabilistic graphical models (Bayesian Networks). The objective is to provide a tool to managers and planners through sustainability variables classification. The general trend of exploitation analysis of a port terminal is to compare ratios and parameters of the international literature on port exploitation and planning. The most used bibliography deals with planning and management, but it does not include sustainability. This kind of analysis establishes the reference criteria which terminals should be exploited. Moreover, it determines what a terminal moves away from or approaches the reference farm. This reference value is represented by the values of these international parameters.

As it has been said, one of the methodologies used are Bayesian Networks. It allow to get in a graphical form relationships between variables of each dimension in order to determine a posteriori values that quantify their contribution to sustainability. There are several studies that address the use of Bayesian Networks in transport area, in which data mining has a great advance [20–25].

In case of maritime transport, probabilistic models have been collected in various investigations [26–28]. However, the first work applying Bayesian Networks for exploitation and port planning was in 2013 by Camarero, González-Cancelas, Soler, López [29]. A similar study is written by Cancelas, Flores, Orive [30], in which naive bayes is used with port traffic as the principal variable. In more recent studies [31], Bayesian Networks are used to find the logistic potential of a country, and, in others [32], main variables are defined and virtual scenarios inferences are determined in order to analyse container terminals scenarios using probabilistic graphical models. Furthermore, a K2

algorithm was created to determine relationships between all of the variables that are involved in decision, using a complete cartography by ArcGIS to obtain scores of each variable in [33].

Li, Yin, Bang, Yang, Wang [34] presents an innovative approach towards integrating logistic regression and Bayesian Networks into maritime risk assessment. In other investigations, it was been able to establish a model of planning zones of logistical activities by using Bayesian Networks [35].

2. Materials and Methods

Nodes represent random variables in probabilistic graphical models. In this case, arcs or a lack of them represent conditional independence assumptions, which provide a representation of probability distributions.

Markov Random Fields and Markov Networks are direct graphical models that have a simple independence definition. They consider that two nodes of set of them, A and B, are conditionally independent from a third node, C, if all the arcs that joined A and B are separated by node C. However, these kinds of networks can represent significant attributes impacts using independent logical formulas weighted by their relationships [36]. On the other hand, Bayesian Networks are directed graphical models. They have a more complex independence notion because they consider arc directionality, as it is explained below [37]. Undirected graphical models are usually used in physics and optical communities and directed models in AI and statistic ones. These kinds of models are called chain graphs. Furthermore, if a more careful study of relationship between directed and undirected graphical models is expected, it is necessary to review references from [38–40].

In order to construct a Bayesian network, methodology represented in Figure 2 has been developed. It is divided into two tasks: processing and construction of an artificial intelligence model.

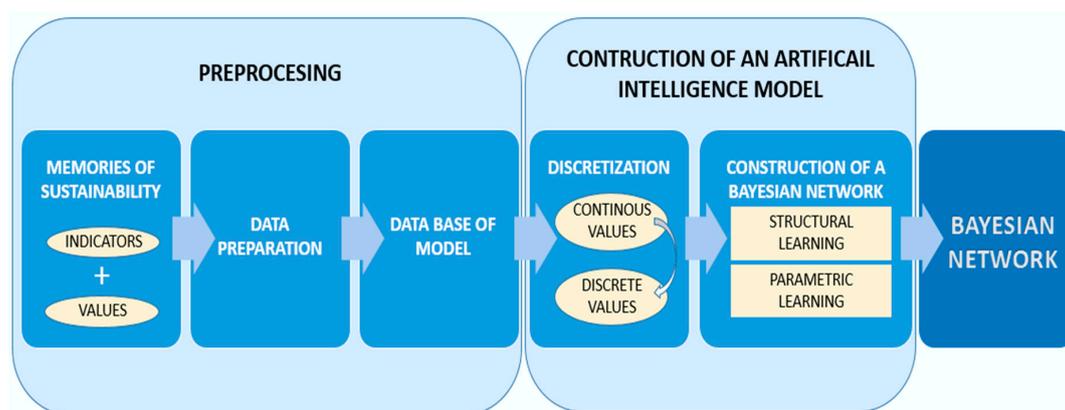


Figure 2. Process to construct a Bayesian network.

Construction of a Bayesian network from data is a learning process that is developed in two steps: structural learning and parameter learning [41]. The first one provides network structure because it establishes dependence/independence relationship between considered variables. The second one, will allow to obtain probabilities (not only a priori ones, even conditioned ones too), if previous structure is considered. Following, variables discretization and model construction is explained.

Model construction requires discretizing selected variables that are not discrete or nominal ones, because Bayesian Networks usually use this kind of variables. However, it is possible to create a Bayesian network using continuous variables. However, continuous variables are limited to Gaussian and ones with linear relationships.

There are two kinds of discretization methods: supervised and not supervised. They depend on software that will be use in the process. Variables obtained in environment work definition are considered in this research. They are discrete, so the continuous ones have been discretized according to intervals determined by 25, 50, and 75. So that, values in the same range will be considered in the same state.

3. Calculation

The variables used in this research correspond to the indicators used by Puertos del Estado in its tasks. Selected variables in the study are included in Table 1. They have been classified in the four sustainability port dimensions: environmental, economic, institutional and social dimension.

Table 1. Selected factors that have been classified in the four sustainability port dimensions.

Dimension	ID	Description	Measure Units
Environmental	<i>MSED</i>	Improving Initiatives. Environmental Management Systems	Millions of euros
Environmental	<i>AQED</i>	The main source of emissions at port can be punctual or diffused	Total number
Environmental	<i>WQED</i>	Main source discharges located at port have a significant impact on water quality and sediment from docks, measures implemented by Port Authority to control emissions linked to entire port activity, surface area service that offers collection and treatment of wastewater	Number
Environmental	<i>SQED</i>	Major emission sources (point and diffuse) port involving significant noise pollution, changes in number of complaints or complaints registered by Port Authority from stakeholders, preparation of noise maps and noise action plan	Number
Environmental	<i>RMED</i>	Waste classification. Management of dregrad material. Waste from vessels (Marpol waste)	Number
Environmental	<i>WEED</i>	Water, electrical power, fuels	Kwh/m ²
Environmental	<i>EIED</i>	Conditions or requirements on environmental issues in specifications of particular technical requirements for port services, in terms of grant and concession titles or authorization	Number
Economic	<i>FSED</i>	Return on assets. EBITDA created per ton moved. Debt servicing. Operating expenses over operating revenue	%
Economic	<i>SIED</i>	Public investment in relation to cash-flow. Foreign investment opposite to public investment. Renewal of assets	Investment/Cash flow
Economic	<i>BSED</i>	Income from occupancy fees over the Net Turnover. Income from activity fees over the Net Turnover. Performance of surface eligible for concession. Performance of active docks	Tm/m ²
Economic	<i>VPED</i>	Creation of Net Turnover per employee. Creation of EBITDA per employee	Net amount of turnover/person
Institutional	<i>MSID</i>	Balanced Scorecards for the implementation of strategies and goals definition: Quality Management Systems according to norm ISO 9001; Occupational Risks Management Systems according to standard OSHAS 18001; Environmental Management Systems according to standard ISO 14001 o EMAS; Management Ex	Number
Institutional	<i>RIID</i>	Investment in basic and complementary infrastructure Industrial or logistic promotion	Number
Institutional	<i>STID</i>	Evolution of traffic in the port system. Distribution of traffic in the Port Authorities	Millions of tonnes
Institutional	<i>EDID</i>	Ports have far key role in the foreign business of Spain. Competitiveness and dynamism of the country's economy receives a direct impact from: Availability of adequate infrastructures; management of such infrastructures with competitive operational costs	% turnover
Institutional	<i>TRID</i>	In the ports several services are provided that, according to their nature, are developed directly by the Port Authority or by private companies accredited by the Port Authority to provide them	% commodity

Table 1. Cont.

Dimension	ID	Description	Measure Units
Institutional	<i>PIID</i>	Progression in the use of land eligible for concession. Diversification of clients, users and port land	Number
Institutional	<i>PSID</i>	Port's Law establishes different mechanisms to guarantee that companies operating the port public domain provide their services in a regimen of free competitiveness and free concurrency	Number
Institutional	<i>IEID</i>	A service management is quality oriented when it seeks client or user's maximum satisfaction	Number
Institutional	<i>TSID</i>	In the ports meet the sea, road and railroad transport, enabling the development of integrated chains of sea-land transport. The financial and environmental efficiency such chains of transport are conditioned by the efficiency in which such ways of transport are coordinated	% railway
Social	<i>ETSD</i>	Employment, internal communication and participation, training, staffing structure and equity, health and safety at work, etc.	% employees within agreement
Social	<i>PCSD</i>	Employment in port community, job security and training services, port concessions, etc.	Number

The main objective is to develop a tool that both, Ports of the State and Port Authorities, can use when making decisions. Besides, they could know relationships between the different variables. In addition, it is expected that the created database can evolve over the years so, that more activity records are available.

National Spanish Port System is composed of 46 ports of general interest, which are managed by 28 Port Authorities. Their coordination and efficiency control corresponds to the Spanish public entity that is responsible for ports, called "Puertos del Estado". It depends on Spanish Fomento Ministry, and it has attributed execution of Government Port policy (www.puertos.es). Variable values uses to construct Bayesian network model correspond to data from the Sustainability Reports that are published annually by Spanish Port Authorities. They have been supplemented by information that is provided by the Public Agency State Ports. They correspond to the historical record since the year 2010—counting on almost 3000 records.

Using these sustainability reports, Puertos del Estado and Port Authorities materialize their commitment to transparency in their management because these documents allow to know the current situation and evolution suffered at an institutional, financial, social, and environmental level. They include a serie of common indicators that allow to standardize the methodology. Current report collects the aggregated information of the port system.

Institutional dimension describes the main challenges and achievements of sustainability in relation with aspects such as: infrastructures, objective market, financial feasibility, institutional communication, operational efficiency, or service quality.

In case of financial dimension, every indicator is gathered regarding the financial situation of the port authority, as well as the level and a structure of investments and some indicators regarding productivity.

The social dimension is based mainly in the human resources policy, including the training actions developed under the scheme of the competency-based management (which has as a goal obtaining an optimal efficiency of the human resources of the company, by means of developing the individual and collective competencies), the quality plan and the efforts made regarding safety and health.

Last, and with regard to the environmental dimension, although the port authorities do not have environmental competencies, they develop a key role in the adequate environmental port management (due to the fact that operate as administrators of infrastructure, regulators, coordinators of the service provided, and, especially, as leaders of the community).

Port activity causes an impact not only in the aquatic environment, but also in the land and air too. Such chapter assesses the impact on the measures that are carried out to reduce them. Although directed models have a more complicated independence notion than undirected ones, but they have several advantages. The main advantage is that everyone can regard an arc from A to B to indicate A “causes” B. This can be used as a guide to build graph structure. Moreover, directed models can encode deterministic relationships. Even, they are easier to learn (fit to data). In addition, it is necessary to specify parameters of the model to define graph structure. In a directed model Conditional Probability Distribution (CPD) must be specify for each node. If variables are discrete, it can be represented as a table (CPT), in which there is listed probability of a child node takes on all of its different values for each combination of its parent’s values.

As it has been said, constructing a Bayesian network from data is a learning process that is divided into two stages: structural learning and parametric learning [40]. The first one consists on obtaining Bayesian network structure, which is relation of dependence and independence between all of its variables involved. The second stage has the purpose of obtaining a priori probabilities and conditional ones that are required from a given structure.

The following describes variables discretization, model construction and classification. In this case, obtained network is displayed by K2 algorithm. It can be observed that variable has been pulled away (Figure 3).

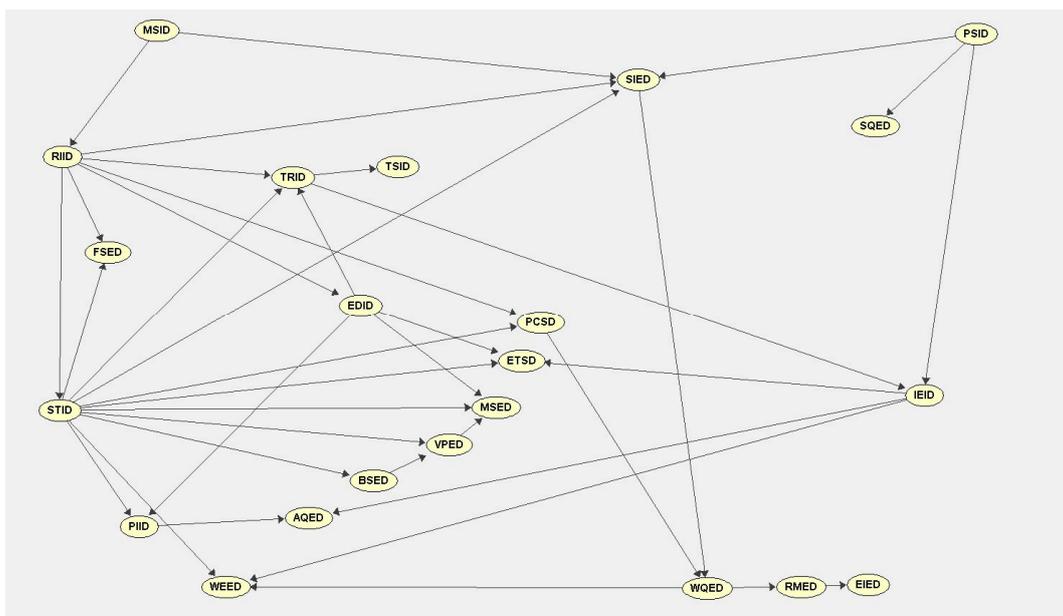


Figure 3. Bayesian Network.

3.1. Variable Discretization

When variables in study are selected, it is necessary to make a variable discretization in model construction process. Bayesian Networks used to consider discrete variables, so called nominal variables. So, if they are not, they must be discretized before model construction.

Discretization consists on dividing continuous variables into a finite number of intervals. In case of Bayesian Networks, it is necessary to select the number of intervals to discretize attributes. Intervals are created when considering frequency, that is to say, the number of instances in each interval or using the same distance. Obviously, some ‘information’ is lost when discretizing. In present study, it has been considered to use discrete variables, so it has been necessary to discretize continuous variables. This task is performed according to expert criteria for the strata selection.

Continuous attributes are transformed into intervals that can be used as discrete tags. Number of intervals is decided by domain expert, and they reject or correct instances with outliers, and so on.

When pre-processed data is finished, learning process begins (data mining process). In the case of Bayesian Networks, it consists on learning network structure and generating probability tables for each node. In this step, domain experts usually will also be necessary in some tasks: all or part of the network structure definition, direction of the arcs decision, algorithms, and so on. When the network is obtained, it may need to be modified by adding, deleting, or reversing the arcs.

3.2. Model Construction

In this stage, structural learning consists of finding dependence relationship between variables. So topology or structure of Bayesian Networks could be determined. Depending on structure type, different methods of structural learning are applied: trees learning, poli-trees learning, multiconnected networks learning, methods based on measures and search, and methods based on dependence relationship.

To construct the model, K2 algorithm has been used in this investigation. K2 algorithm is based on measure optimization, which is the main target in planning to optimize exploitation ratios. This measure is used to exploit search space, which includes all of the networks that contain variables from the database. It uses an algorithm of colines ascension. It begins from an initial network and it is modifying (including arcs, erasing them or changing their direction). The result is a new network, which has a better measure. Specifically, K2 measure [42] for a G network and a D database is (1):

$$f(G : D) = \log P(G) + \sum_{i=1}^n \left[\sum_{k=1}^{s_j} \left[\frac{\log(\Gamma(\eta_{ik}))}{\Gamma(\eta_{IK} + \eta_{IK})} + \sum_{j=1}^{r_i} \frac{\log(\Gamma(\eta_{N_{ijk}} + \eta_{ijk}))}{\Gamma(\eta_{ijk})} \right] \right] \quad (1)$$

where is N_{ikj} the frequency of the configurations found in D database of the variables x_i , where n is the number of variables, taking its j value and its parents in G taking its k configuration, where if it is the number of possible configurations of set of parents, and r_i is the number of values that can take the variable x_i . Furthermore, $N_{ik} = \sum_{j=1}^{r_i} N_{ijk}$ and Gamma function is represented by Γ .

K2 algorithm is different from algorithm of stochastic search with variable neighborhood because:

- the search algorithm is not a simple coline ascension, but the points of the search space are grouped by neighborhoods and explored locally in those neighborhoods; and,
- the quality measure of the networks is not necessarily the K2 one.

Algorithm of stochastic search with variable neighborhood is defined, as follows (2):

$$f(G : D) = \sum_{i=1}^n \sum_{j=1}^{r_i} \sum_{k=1}^{s_i} N_{ijk} \cdot \log \frac{N_{ijk}}{N_{IK}} - \frac{1}{2} C(G) \log N \quad (2)$$

where N is the register number of data base and $C(G)$ is a complexity measure of G network. It is defined by (3):

$$C(G) = \sum_{i=1}^n (r_i - 1) s_i \quad (3)$$

Bayesian Networks have long been incorporated into supervised classification tasks, but not into port planning considering ideas presented by Acid & de Campos [43], Friedman & Goldszmidt [44], and extended by Sierra & Larranaga [45]. Probability factorizations that are represented by this kind of networks are useful to perform classification of some variables. These variables are considered as special because they are predicted by the rest of variables. So, network structure is able to predict values of the type of variables when considering values of the predictor ones. That is, it is possible to calculate a posteriori probability of a specific node if the values of the rest of variables are known.

Elvira software was used in the network construction. It is a specific issue in Bayesian Networks [46] and it is intended for the editing and evaluation of probabilistic graphic models, specifically Bayesian Networks and influence diagrams. Elvira has its own format for coding the models, a reader-interpreter for the coded models, and a graphical interface for the construction of networks. It has specific options for canonical models (OR, AND, MAX, etc.), algorithms exact and approximate (stochastic) reasoning for both discrete and continuous variables, reasoning explanation methods, decision making algorithms, model learning from databases, network fusion, etc. Being compiled in JAVA, its most important advantage is that it allows working on different platforms and operating systems.

4. Results and Discussion

Bayesian models are used to solve problems from a descriptive and predictive perspective. As a descriptive method, these models focus on discovering dependency/independence relationships. From this perspective, it can be affirmed that they sometimes complement and/or even exceed the association rules. Regarding the predictive function, it is circumscribed to Bayesian techniques as classification methods.

Figure 3 shows network obtained by the algorithm K2. It shows that *MSID* and *PSID* are the father of the network because they only leave arrows. Connections of type $A \rightarrow B$ indicate the dependence or direct relevance between these variables. In this case, it is indicated that *B* depends on *A*, or that *A* is the cause of *B*, so *B* is the effect of *A*. It is also said that *A* is the father and *B* is the son. The absence of arcs between the nodes is providing us with valuable information, since in this case, the graph informs us of conditional independence.

Topology needs to identify factors that are relevant to determine how those factors are causally related to each other. The arc cause-effect means that cause is a factor that is involved in causing effect. In this case, parent of *SQED* and *SIED* variables is the node *PSID*. Then, when *PSID* is known, *SQED* and *SIED* are conditionally independent (Figure 4).

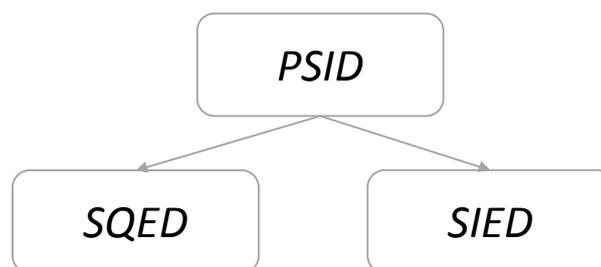


Figure 4. Divergent relationship

Then, *PSID* is a resolution variable that appears in network as a “node”. Some arcs started on it, so this variable generates a divergent connection. This way, *PSID* is a parent node that projects arcs to several sons, that is to say, some arrows start in this variable and diverge to its sons (Figure 4). As it can be remembered: *PSID*, timestamp transparency and free competition. So, its initiatives go 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, some variables are taken in an independent way and information will not spread along network if some evidences are included over son nodes.

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 *PSID* is a parent of *SQED* and *SIED*. So, *SQED* and *SIED* are children of *PSID*, so, *PSID* influences or causes *SQED* and *SIED*. Consequently, *SQED* and *SIED* depends on *PSID*.

As it is shown in Figure 5, *RIID* has an influence on *EDID*, and at the same time, on *ETSD*. So, a *RIID* evidence will influence on *EDID* certainty, and then, it will influence on *ETSD* certainty too. Analogously, *ETSD* evidence will influence a *RIID* certainty through *EDID* variable. However, if *EDID* state is given, the channel will be blocked and *RIID* and *ETSD* variables become independent. Then, it can be said that *RIID* and *ETSD* are d-separated given *EDID*. When a variable state is given, it can be said that this variable is instantiated. It shows that evidence could be transmitted through a serial connection unless a variable state in the connection is given.

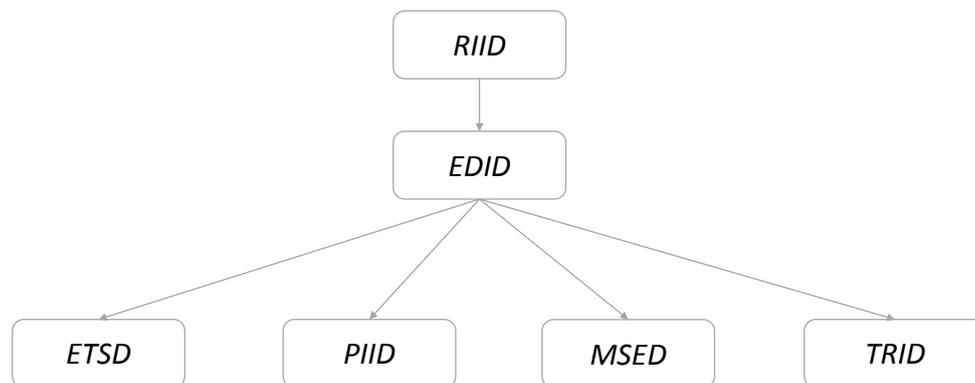


Figure 5. Serial relationship.

A causal model for the relations between *RIID*, *EDID* and *TRID*, or the relations between *RIID*, *EDID* and *PIID* is represented in Figure 5. On one hand, if *EDID* is not observed, *PIID* will increase believing that *EDID* is high. It shows something about *RIID*. On the other hand, when considering the same line of reasoning, if *EDID* is already known, *PIID* will show anything new about *RIID*.

Conditioning the Bayesian Network independence assumption, several independence statements can be observed in this case. So, if *EDID* is known, *ETSD*, *PIID*, *MSED* and *TRID* are conditionally independent of its ancestor's *RIID*. Casual graph: The variable *EDID* has four common effects *ETSD*, *PIID*, *MSED* and *TRID* (Figure 5).

Management systems supporting decision-making includes quality management systems, scorecards, market characterization campaigns, etc. All of them are represented by *MSID*. *MSID* is considered as a parent variable in network, so arrows only start on it. The same goes for *PSID*, 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. *PSID* is a father node in network. Furthermore, it 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 *STID* (Figure 3). Ten arrows start on it and go to ten different nodes. These are structure effects 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, it has effects on variables representing port community employment, job security, and training services and health work, among others (social category).

Finally, in the environmental category, served markets causes different grades of environmental management systems implementation (EMAS, ISO 14001 y PERLS), economic resource investment, and investments that are associated to implementation, certification, and maintenance of environmental

management system. Therefore, served markets are a very important variable in planning considering a sustainable perspective.

Evidence may only be transmitted through the converging connection if either *AQED* or one of its descendants has received evidence (see Figure 6). In Figure 6 it can be observed that if nothing is known about *AQED* except what may be inferred from knowledge of its parents *PIID* and *IEID*, then the parents are independent. Consequently, evidence about them cannot influence the certainties of the rest through *AQED*. Moreover, knowledge of one of the causes does not inform anything about the rest of them. However, if any consequence is known, information about one possible cause could tell something about the other causes.

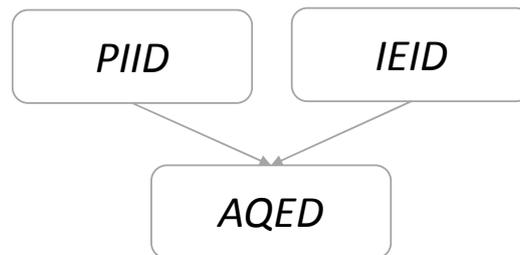


Figure 6. Convergent relationship.

First, if there is no available evidence about *AQED* state, information about *PIID* state will not provide any information about *IEID* state. So, *PIID* is not an *IEID* indicator and vice versa. Thus, a converging connection will not transmit information if there is no available evidence for the middle variable, as it is shown in Figure 5. Convergent connection has no evidence on *AQED* or on any of its sons. So, lack of information about *PIID* will affect the belief about *IEID* state, and vice versa.

In second term, if there is available *AQED* evidence, some information about *PIID* state will provide an explanation about evidence received about *AQED* state. On one hand, this issue confirms or dismisses *IEID* as cause of the evidence received for *AQED*. Obviously, the opposite situation holds true. On the other hand, convergent connections will allow information transmission if middle variable evidence is available.

As it can be seen in Figure 6, if anything is known about a common effect of two (or more) causes, they will be considered as independent. So, information about one of them that is received will have no effect on other(s) belief. However, if some evidence is available on a common effect, causes do not turn on independent.

PIID and *IEID* are conditionally dependent if *AQED* is observed. So if it already knows that air quality went off, knowing the number of companies that operate in a port will increase general belief in that initiatives that are aimed at improving efficiency promoted by the Port Authority. A deep knowledge allows to increase number of initiative promoted by Port Authority to improve efficiency, quality of service, and performance of services that are provided to the merchandise. Thus, increases the belief in growth of the number of companies operating in the port, land occupied.

In Figure 7, it can be observed that only *PCSD* indirectly through information about *RMED*. So, knowing the state of *RMED*, something can be shown about the state of *WQED*, which in turn, shows something about *PCSD*.

Moreover, if a variable is instantiated, then it is called evidence. Otherwise, it is soft. In Figure 7, hard evidence about the variable *RMED* provides soft evidence about the variable *PCSD* because serial and diverging connections require a hard evidence.

If nothing is known about *WQED* or *PCSD*, information about whether *RIID* will not show anything about *STID*. However, if something is noticed about *WQED*, information about *RIID* will make to believe something about *STID*.

If neither *PCSD* nor any of its descendants are observed, *RIID* and *STID* will be independent.

Information cannot be transmitted through *PCSD* among parents of *PCSD*. It leaks down *PCSD* and its descendants. Besides, if *PCSD* or any of its descendants is observed, *RIID* and *STID* will be dependent. Information can be transmitted through *PCSD* among parents of *PCSD* if *PCSD* or any of its descendants are observed. Observation of *PCSD* or its descendants opens the information path.

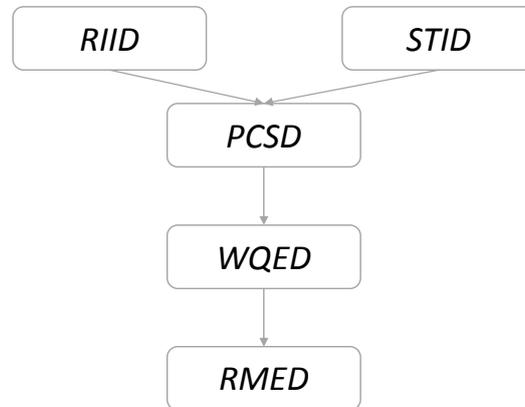


Figure 7. Convergent + serial relationship.

5. Conclusions

Through Bayesian network construction, the most decision-making category is institutional category on the bottom. It is followed by economic and social at the same level, and finally, by the environmental category.

Institutional variables are interconnected. Economics ones are important as cause-effect and they are effects of served markets, which belong to institutional dimension. Generated value and productivity depend on kind of business and service. Moreover, social variables are the effects of institutional, but they have not a direct relationship with their same dimension. Finally, environmental variables are closely interconnected and they are effects of institutional category.

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

Sustainability should be understood as a value contributor element for entities and society as a whole. It is a transversal and multidimensional concept that must be integrated in the strategy policies and actions to answer social, environmental and business needs. It will simultaneously allow increasing value creation capacity in organizations and their condition of long-term success factor.

In current port system structure, institutional variables take greater weight in the achievement of sustainability objectives. A dependent relationship between these kind of variables has been obtained in the Bayesian network developed. For this reason, the port system will be able to act on sustainability variables if it acts on environmental dimension variables. To study the possible effects, it is necessary to know about in which parameters can act Port Authority in an institutional point of view. The basic functions of the Port Authority are the planning, projection, construction, conservation and operations of port works and services, collaboration with official bodies and port private companies coordination and management of the port domain.

Relationships between private initiative and public body are decisive in achievement of sustainability objectives, considering institutional dimension. Regarding to the different types of investments (Port-City, safety, environment, or business promotion), its right identification and analysis present 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. As for institutional transparency, it is relevant to mention that the Ports law establishes different mechanisms to guarantee that companies operating in

the port public domain deliver their service in a regimen of free competitiveness and free concurrency. Regarding 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 emphasizing the wide group of social, financial, and administrative nature collectives that are affected by the Port Authority activity. Even they affect the development and performance of such Port Authority activities. Considering 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, they Management Boards without more limits than those regulated by Labour and Budget Regulations. Port Authorities personnel are linked to these by a relation subject to the Labour or private Law rules that are applicable. Recruitment is carried out according to systems that are based on the principles of merit and capacity, and with the exception of directive or trusted staff, through public call. The regimen of remuneration and non-compatibility adjusts to what is in general established regarding staff of Public Law Entities which article 6 of the amendment Budget General Law, refers to. So, in case of social dimension, it seems less probable to influence in a direct way.

Therefore, if a Bayesian network is built in a port environment which is based on the four pillars of sustainability, a tool that allows to actuate on global port system sustainability 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 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. It is important to keep in mind that the Port Authority has not environmental competencies; neither does it has the ultimate responsibility to enforce the environmental legislation in the port. In general, this competency relies on the Autonomous Communities, who are provided with a sanctioning regimen that allows for them to act against possible violations. However, Port Authorities develop a key role in the adequate environmental port management due to the fact that they act as administrators of infrastructures, regulators, coordinators of services provided, and, specially, as leaders of the port community. So, it can be observed that Bayesian Networks become very optimal models. They can incorporate information from experts in the study area and, further, they optimize the correct answer percentage.

Bayesian Networks allow that information that is provided by one or more observed variables that are observed (evidence) is propagated through the network and updates the belief about the non-observed variables. This process is called inference. It is possible to learn the conditional probabilities that describe the relationships between the variables from the data. It is even possible to learn the complete structure of the network from complete data or with some of its unknown values. Besides, Bayesian Networks can be used to make optimal decisions by introducing possible actions and the usefulness of their results. The next step to develop would be inference, so that the main objective is to find the probability distribution of certain variables of interest given the values of other observed variables.

Probabilistic inference allows to give certain known variables (evidence). To carry out classifications, Bayesian Networks have been intended to use in the next research, considering for this the existence of a special variable: the variable to be classified. This variable is predicted by a group of variables (the rest). So, obtained network structure can be used to predict the value of the kind of this variable to be classified. This is obtained assigning values to the predictors (rest of variables) and the subsequent propagation of the evidence that is introduced in the network. That is, if a posteriori probability of the node is calculated, it will be associated to the special variable, given the values of the rest.

Author Contributions: Nicoleta González-Cancelas, Beatriz Molina Serrano and Francisco Soler-Flores conceived and designed the experiments; Francisco Soler-Flores performed the experiments; Nicoleta González-Cancelas

and Beatriz Molina Serrano analyzed the data; Samir Awad-Nuñez and Alberto Camarero Orive contributed analysis tools; Nicoleta González-Cancelas and Beatriz Molina Serrano wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Doerr, O. *Políticas Portuarias Sostenibles*; Boletín FAL; Edición n° 299, Número 7 de 2011; CEPAL: Santiago, Chile, 2011.
2. Robèrt, K.-H.; Borén, S.; Ny, H.; Broman, G.A. Strategic approach to sustainable transport system development—Part 1: Attempting a generic community planning process model. *J. Clean. Prod.* **2016**, *140*, 53–61. [[CrossRef](#)]
3. Global Reporting Initiative. *Guía Para la Elaboración de Memorias de Sostenibilidad*; Versión 3.1; GRI: Amsterdam, The Netherlands, 2000.
4. Serrano, O. *Operativa Portuaria y Sostenibilidad*; Conama Local: Málaga, Spain, 2015.
5. Lirn, T.C.; Lin, H.W.; Shang, K.C. Green shipping management capability and firm performance in the container shipping industry. *Marit. Policy Manag.* **2014**, *41*, 159–175. [[CrossRef](#)]
6. Lun, Y.H.V.; Lai, K.-H.; Wong, C.W.Y.; Cheng, T.C.E. Green shipping practices and firm performance. *Marit. Policy Manag.* **2014**, *41*, 134–148. [[CrossRef](#)]
7. Thi Thuy Hong, N.; Thi Lich, H.; Thanh Nga, B.T. The Social Investment Capital and the Cargo Volume Transported by Sea: A VAR Approach for Vietnam. *Logistics* **2017**, *1*, 8. [[CrossRef](#)]
8. Sánchez, R.J.; Jauimurzono, A.; Willmsmeier, G.; Pérez Salas, G.; Doerr, O.; Pinto, F. *Transporte Marítimo y Puertos. Desafío y Oportunidades en Busca de un Desarrollo Sostenible de América Latina y El Caribe*; Serie Recursos Naturales e Infraestructura; CEPAL: Santiago, Chile, 2015.
9. Duru, O. The Origin and Consistency of the Ton–Mile Metric in the Shipping Economics. *Logistics* **2017**, *1*, 3. [[CrossRef](#)]
10. Crespo Soler, C.; Ripoll Feliu, V.; Crespo, A.; Giner, A. La Sostenibilidad Ambiental en el Sistema Portuario de Titularidad Estatal. In Proceedings of the XIII Congreso AECA, Oviedo, Spain, 22–24 September 2005. Armonización y Gobierno de la Diversidad.
11. Doménech, J.L.; Matias, A.; Muñoz-Calero, R. Ecoeficiencia y sostenibilidad en Puertos: Aplicaciones en el puerto de Gijón. In *Comunicaciones Técnicas del VIII Congreso Nacional de Medio Ambiente (CONAMA 8)*; Fundación CONAMA: Madrid, Spain, 2004.
12. Monie, G.D. *Measuring and Evaluating Port Performance and Productivity*; United Nations: New York, NY, USA, 1987.
13. Monie, G.D. Medición y evaluación del rendimiento y de la productividad de los puertos. In *Monografías de la UNCTAD Sobre Gestión de Puertos*; Naciones Unidas National Park: New York, NY, USA, 1989.
14. Fillol, A.G.; Bataza, J.A.M.; Ubal, N.P.; Feliu, V.M.R.; Soler, C.C. La Información de Sostenibilidad en el Marco de las Cuentas Anuales: Análisis Aplicado al Caso de la Autoridad Portuaria de Valencia. *Revista de Contabilidade do Mestrado em Ciências Contábeis da UERJ* **2007**, *12*, 3.
15. González, F.; Guerra, A.; Martín, F.; Nóvoa, J.J.; Otero, C.Y.; Penela, J. Medición de la sostenibilidad en el sistema portuario Español: Propuesta metodológica a través de indicadores sintéticos de desarrollo sostenible. In Proceedings of the XII Reunión de Economía Mundial, Santiago de Compostela, Spain, 26–28 May 2010.
16. Del Estado, P. *Guía Para la Elaboración de Memorias de Sostenibilidad en el Sistema Portuario Español*; Working Paper; Puertos del Estado: Madrid, Spain, 2008.
17. Del Estado, P. *Memoria de Sostenibilidad del Sistema Portuario de Interés General*; Puertos del Estado: Madrid, Spain, 2011.
18. Serrano, B.M.; González-Cancelas, N.; Soler-Flores, F.; Camarero-Orive, A. Classification and prediction of port variables using Bayesian Networks. *Transp. Policy* **2017**. [[CrossRef](#)]
19. Serrano, M.M.G.; Trujillo, L. La Medición de la Eficiencia en el Sector Portuario: Revisión de la Evidencia Empírica. DOCUMENTO DE TRABAJO 2005-06. Available online: https://www.researchgate.net/profile/Lourdes_Trujillo2/publication/4999078_La_medicion_de_la_eficiencia_en_el_sector_portuario_revision_de_la_evidencia_empirica/links/570257fb08ae61e0c46cfcb4/La-medicion-de-la-eficiencia-en-el-sector-portuario-revision-de-la-evidencia-empirica.pdf (accessed on 9 January 2018).

20. Hofleitner, A.; Herring, R.; Abbeel, P.; Bayen, A. Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network. *Intell. Trans. Syst. IEEE Transp.* **2012**, *13*, 1679–1693. [CrossRef]
21. Janssens, D.; Wets, G.; Brijs, T.; Vanhoof, K.; Arentze, T.; Timmermans, H. Integrating Bayesian Networks and decision trees in a sequential rule-based transportation model. *Eur. J. Oper. Res.* **2006**, *175*, 16–34. [CrossRef]
22. Castillo, E.; Menéndez, J.M.; Sánchez-Cambronero, S. Predicting traffic flow using Bayesian Networks. *Transp. Res. Part B Methodol.* **2008**, *42*, 482–509. [CrossRef]
23. Sun, S.; Zhang, C.; Yu, G. A bayesian network approach to traffic flow forecasting. *Intelligent Transportation Systems. IEEE Trans.* **2006**, *7*, 124–132.
24. Zheng, W.; Lee, D.H.; Shi, Q. Short-term freeway traffic flow prediction: Bayesian combined neural network approach. *J. Transp. Eng.* **2006**, *132*, 114–121. [CrossRef]
25. Tebaldi, C.; West, M. Bayesian inference on network traffic using link count data. *J. Am. Stat. Assoc.* **1998**, *93*, 557–573. [CrossRef]
26. Trucco, P.; Cagno, E.; Ruggeri, F.; Grande, O. A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation. *Reliab. Eng. Syst. Saf.* **2008**, *93*, 845–856. [CrossRef]
27. Klemola, E.; Kuronen, J.; Kalli, J.; Arola, T.; Hanninen, M.; Lehtikoinen, A.; Tapaninen, U. A cross-disciplinary approach to minimising the risks of maritime transport in the Gulf of Finland. *World Rev. Intermodal Transp. Res.* **2009**, *2*, 343–363. [CrossRef]
28. Kaluza, P.; Kölsch, A.; Gastner, M.T.; Blasius, B. The complex network of global cargo ship movements. *J. R. Soc. Interface* **2010**, *7*, 1093–1103. [CrossRef] [PubMed]
29. Camarero, A.; González-Cancelas, N.; Soler, F.; López, I. Utilización de redes bayesianas como método de caracterización de parámetros físicos de las terminales de contenedores del sistema portuario español. *Rev. Ing.* **2013**, *39*, 31–38.
30. Cancelas, N.G.; Flores, F.S.; Orive, A.C. Modelo de eficiencia de las terminales de contenedores del sistema portuario español. *Rev. Electrón. Comun. Trab. Asepuma.* **2013**, *14*, 49–67.
31. Quijada-Alarcón, J.; González-Cancelas, N.; Camarero Orive, A.; Soler Flores, F. Logistic groups in the Republic of Panama. *J. Maps* **2015**, *12*, 1248–1252. [CrossRef]
32. Rodríguez García, T.; González-Cancelas, N.; Soler-Flores, F. Setting the Port Planning Parameters in Container Terminals through Bayesian Networks. *Promet Traffic Transp.* **2015**, *27*. [CrossRef]
33. Awad-Núñez, S.; González-Cancelas, N.; Soler Flores, F.; Camarero Orive, A. How should the sustainability of the location of dry ports be measured? A proposed methodology using Bayesian Networks and multi-criteria decision analysis. *Transport* **2015**, *30*, 312–319.
34. Li, K.X.; Yin, J.; Bang, H.S.; Yang, Z.; Wang, J. Bayesian network with quantitative input for maritime risk analysis. *Transp. A Transp. Sci.* **2014**, *10*, 89–118. [CrossRef]
35. Flores, F.S.; Cancelas, N.G.; Orive, A.C.; Gárate, J.L.A.; Monzón, M.D.C.P. Diseño de un modelo de planificación de zonas de actividades logísticas mediante el empleo de redes bayesiana. *Rev. Ing. Ind.* **2014**, *12*, 1.
36. Li, T.; Guan, H.; Ma, J.; Zhang, G.; Liang, K. Modeling travel mode choice behavior with bounded rationality using Markov Logic Networks. *Transp. Lett.* **2017**, 1–8. [CrossRef]
37. Almazán-Gárate, J.L.; Palomino-Monzón, M.C.; González-Cancelas, N.; Soler-Flores, F. Relationship between air pollution and natural gas with respect to maritime transport. Methodology based on Bayesian Networks. In Proceedings of the Global Virtual Conference; Available online: <http://www.gv-conference.com/archive/?vid=1&aid=3&kid=30201-114&q=f6> (accessed on 11 January 2018). Transport and Logistics Section.
38. Castillo, E.; Gutiérrez, J.M.; Hadi, A.S. *Expert Systems and Probabilistic Network Models*; Springer: New York, NY, USA, 1997.
39. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern Classification*; Wiley: New York, NY, USA, 2001.
40. Pearl, J. The Solution for the Branching Factor of the Alpha-Beta Pruning Algorithm and its Optimality. *Commun. ACM* **1982**, *25*, 559–564. [CrossRef]
41. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*; Morgan Kaufmann: Burlington, MA, USA, 1988.
42. Cooper, G.F.; Herskovits, E. A Bayesian Method for the Induction of Probabilistic Networks from Data. *Mach. Learn.* **1992**, *9*, 309–347. [CrossRef]
43. Acid, S.; De Campos, L. Approximations of Causal Networks by Polytrees: An Empirical Study. In Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Paris, France, 4–8 July 1994; pp. 149–158.

44. Friedman, N.; Goldszmidt, M. Building Classifiers Using Bayesian Networks. In Proceedings of the National Conference on Artificial Intelligence, Portland, Oregon, 4–8 August 1996; AAAI Press: Menlo Park, CA, USA, 1996.
45. Sierra, B.; Larranaga, P. Predicting Survival in Malignant Skin Melanoma using Bayesian Networks Automatically Induced Genetic Algorithms: An Empirical Comparison between Different Approaches. *Artif. Intell. Med.* **1998**, *14*, 215–230. [[CrossRef](#)]
46. Consortium, E. Elvira: An environment for creating and using probabilistic graphical models. In Proceedings of the First European Workshop on Probabilistic Graphical Models, Cuenca, Spain, 6–8 November 2002; pp. 222–230.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).