

Hacia la sostenibilidad portuaria mediante modelos probabilísticos: redes bayesianas

Towards port sustainability through probabilistic models: Bayesian networks

B. Molina (*), N. Gonzalez (*), F. Soler (*)

RESUMEN

En la explotación y gestión es necesario que el gestor de la infraestructura conozca las relaciones entre las variables en juego, lo cual es factible con el uso de redes bayesianas que permiten clasificar, predecir y diagnosticar las variables, pudiendo incluso estimar la probabilidad posterior de las no conocidas en base a las conocidas. En la metodología propuesta se ha generado una base de datos con variables portuarias, clasificadas en económicas, sociales, ambientales e institucionales, tal como se abordan los estudios de smart ports, en el Sistema Portuario Español, desarrollando una red mediante un grafo dirigido acíclico para conocer las relaciones en términos de padres e hijos. En términos probabilísticos, se observa que las variables más decisoras para la sostenibilidad portuaria son las institucionales.

Se concluye que las redes bayesianas permiten modelar la incertidumbre de forma probabilística incluso cuando el número de variables es elevado como en la planificación y explotación portuaria.

Palabras clave: variables portuarias; sostenibilidad; puerto; gestión portuaria; redes Bayesianas.

ABSTRACT

It is necessary that a manager of an infrastructure knows relations between variables. Using Bayesian networks, variables can be classified, predicted and diagnosed, being able to estimate posterior probability of the unknown ones based on known ones. The proposed methodology has generated a database with port variables, which have been classified as economic, social, environmental and institutional, as addressed in of smart ports studies made in all Spanish Port System. Network has been developed using an acyclic directed graph, which have let us know relationships in terms of parents and sons. In probabilistic terms, it can be concluded from the constructed network that the most decisive variables for port sustainability are those that are part of the institutional dimension.

It has been concluded that Bayesian networks allow modeling uncertainty probabilistically even when the number of variables is high as it occurs in port planning and exploitation.

Keywords: port variables; sustainability; port; port management; Bayesian networks.

(*) Universidad Politécnica de Madrid.

Persona de contacto/Corresponding author: beatriz.molinas@alumnos.upm.es (B. Molina)

ORCID: <http://orcid.org/0000-0002-7832-9573> (B. Molina); <http://orcid.org/0000-0001-7167-1563> (N. Gonzalez); <http://orcid.org/0000-0002-1636-834X> (F. Soler)

Cómo citar este artículo/Citation: Molina, B., Gonzalez, N., Soler, F. (2018). Hacia la sostenibilidad portuaria mediante modelos probabilísticos: redes bayesianas. *Informes de la Construcción*, 70(549): e244. <https://doi.org/10.3989/id.54678>

Copyright: © 2018 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

1. INTRODUCCIÓN

El concepto de desarrollo sostenible partió de la búsqueda de compatibilidad entre el desarrollo económico y la protección y uso adecuado de los recursos naturales. Por primera vez se incluyó en la Conferencia de Estocolmo (1972), si bien no fue asumido por la sociedad hasta 1987 con el Informe Brundtland donde el concepto de desarrollo sostenible pasó a definirse como «el desarrollo que satisface las necesidades de la generación presente, sin comprometer la capacidad de las generaciones futuras de satisfacer sus propias necesidades» [1]. Este hecho supuso un cambio social, ambiental y económico importante al incluir cuestiones medioambientales que nunca antes habían sido debatidas. En los años siguientes, esta visión fue ampliándose a otras dimensiones, de forma que actualmente, el concepto de desarrollo sostenible posee un carácter integral, multidimensional e interactivo, que involucra más áreas del conocimiento [2].

En la actualidad, el concepto de sostenibilidad está siendo aplicado en forma emergente por autoridades del sector transporte y muchos otros campos de actividad e industrias a nivel mundial, fuertemente impulsada por iniciativas que incorporan la variable ambiental y la responsabilidad social empresarial en la gestión estratégica de las empresas [3]. En el caso de los puertos, la sostenibilidad portuaria tiene sus raíces en las propuestas del GRI [4] de las cuáles conserva, entre otras cuestiones, los cuatro ejes o dimensiones que conforman un enfoque de desarrollo sostenible, es decir, el institucional, el económico, el ambiental y el social.

De esta forma se considera que la gestión sostenible de una empresa u organismo tiene como meta el mantenimiento equilibrado en el largo plazo de su función y actividad, de modo que se tengan en cuenta el efecto recíproco de su actividad con el entorno económico, social y ambiental con el que se relaciona y el efecto de su actividad con su propia estructura al tiempo que se el efecto de sus funciones y actividades en sus aspectos o dimensiones económica, social, medioambiental e institucional, buscando un desarrollo equilibrado de estas cuatro dimensiones [5].

En este contexto, merece una atención particular el transporte marítimo, al transportar alrededor del 80% del volumen del comercio internacional (en toneladas-kilómetros) a nivel mundial [6]. Así, en el sector portuario, se debe entender la gestión sostenible como «aquella que permite que crezca el volumen de tráfico de contenedores, graneles sólidos y líquidos, mercancía general y número de pasajeros, disminuyendo a su vez el consumo de energía y recursos naturales, el volumen de residuos generados y los impactos negativos a los sistemas sociales y ecosistemas en las áreas de influencia del puertos» [7].

Sin embargo, uno de los principales desafíos que plantea la introducción de criterios de sostenibilidad en el modelo de gestión y desarrollo portuario es romper la inercia existente en relación a la consideración «única» del factor económico como variable de desarrollo, y lograr que las variables ambiental y social tomen la importancia que deben a fin de que el modelo de gestión y desarrollo portuario realmente tienda hacia la sostenibilidad [8].

El Texto Refundido de la Ley de Puertos del Estado y de la Marina Mercante (BOE núm. 253, de 20/10/2011), incor-

pora la sostenibilidad como uno de los principios que deben regir el modelo de planificación y de gestión de los puertos. Para ello, en el artículo 55.4 prevé que el proyecto de Plan de Empresa de cada Autoridad Portuaria debe acompañarse de una Memoria de Sostenibilidad, la cual constituye una herramienta de análisis y diagnóstico. Dicha memoria se lleva a cabo a través de una metodología a la cual se incorpora la constitución de indicadores específicos, si bien no determina el comportamiento de la Autoridad Portuaria, sino que describe, por medio de indicadores de desempeño, los resultados derivados de la adopción y aplicación de códigos, políticas y sistemas de gestión [9].

Estos indicadores de sostenibilidad permiten evaluar el desempeño de la gestión de desarrollo sostenible, más allá de los reportes de sostenibilidad, los indicadores permiten controlar de manera objetiva la marcha de la gestión de la Autoridad Portuaria en esta materia. La aplicación de los indicadores es útil para que las Autoridades Portuarias controlen su gestión sostenible, evalúen el impacto de los programas aplicados y los modifiquen cuando sea necesario. Los indicadores permiten realizar Benchmarking de gestión sostenible entre puertos para determinar las mejores prácticas y comparar el desempeño de una Autoridad Portuaria frente a la industria similar, y en el marketing puede ser un elemento diferenciador y de competitividad en el mercado. La aplicación generalizada a un sistema portuario permitiría realizar un benchmarking preciso en materia de sostenibilidad entre puertos de una misma región o país [10].

De la aplicación de dichas herramientas surgen los objetivos (económicos, medio ambientales, sociales e institucionales) que una autoridad o empresa portuaria debe alcanzar para asegurar el desarrollo sostenible y crecimiento de su puerto [11]. Los objetivos económicos pueden ser: el incremento del volumen de negocio, aumentar los ingresos por concesiones, reducir el endeudamiento con el fin de asegurar la sostenibilidad financiera del puerto y optimizar y rentabilizar las inversiones de los activos portuarios. En el caso de la dimensión medio ambiental, los objetivos pueden ser tales como accionar con respeto al medio ambiente, minimizar los impactos ambientales derivados de la actividad portuaria, minimizar los accidentes ambientales y mejorar la gestión ambiental en el recinto portuario. Los objetivos sociales que pueden ser internos y externos, deben estar enmarcados en ámbitos tales como desarrollar y modernizar sistemas de gestión de los recursos humanos, desarrollar un equipo humano motivado y comprometido y lograr un respaldo sostenido y activo de la comunidad del entorno. Para los objetivos institucionales se puede estar buscando impulsar ciertos cambios legales y normativos para modernizar la forma de desarrollo y operación del puerto, reorganizar el mercado portuario incorporando competencia, gestión e inversión privada para mejorar su eficiencia y capacidad de expansión, modernizar el régimen laboral para mejorar su competitividad, desarrollar la comunidad portuaria para incrementar la eficiencia operacional y calidad de los procesos, institucionalizar y optimizar la relación ciudad puerto, expandir la gestión operativa del puerto a la cadena logística para agregar valor e integrar a la comunidad logística local al desarrollo del puerto (Figura 1).

Por tanto, el objetivo que se persigue es que través de estos cuatro ejes de la sostenibilidad, los puertos se confor-

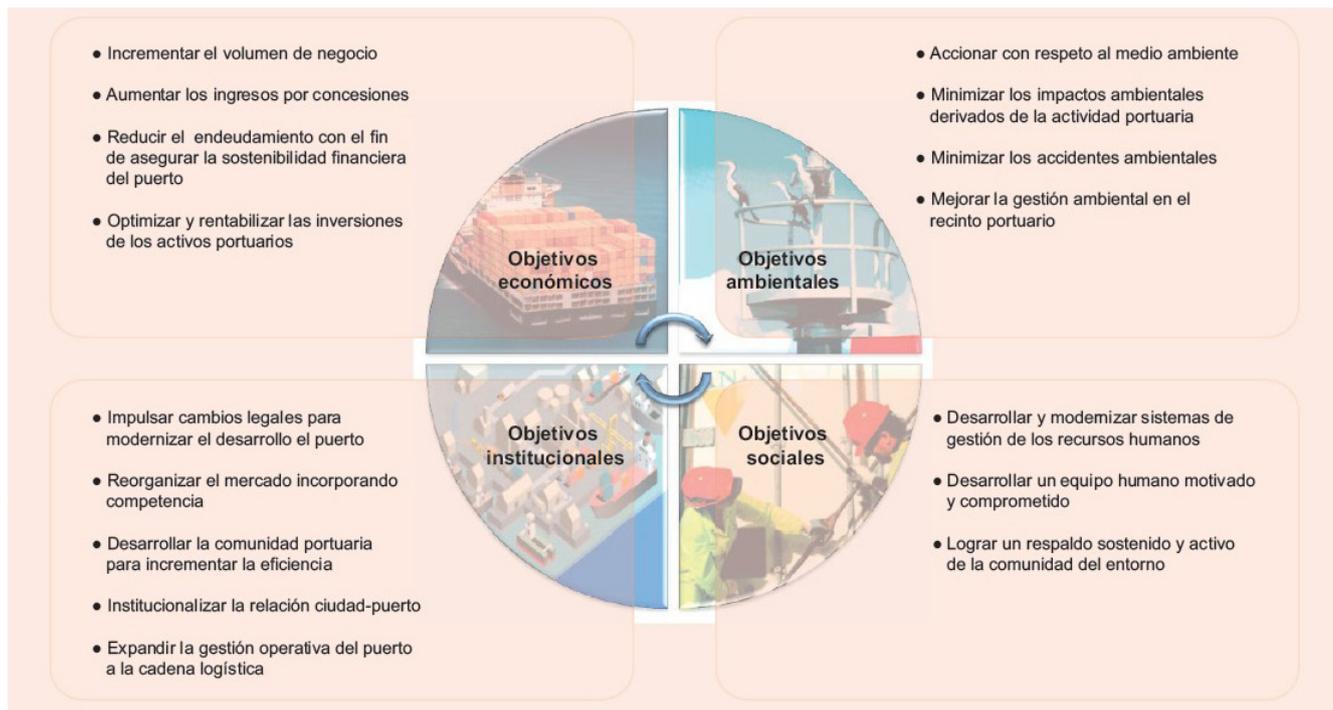


Figura 1. Objetivos estratégicos para las cuatro dimensiones de la sostenibilidad. Fuente: Doerr, 2011 [3]

men como un sistema y no sean vistos como entes aislados y sujetos a una coyuntura comercial concreta, sino como elementos que se interrelacionan con un entorno físico, social y ambiental, en el que han de integrarse de forma efectiva, esto es, siendo capaces de adaptarse a una coyuntura cambiante y a la vez, apuntando a una renovación que contribuya a alcanzar el mejor de los escenarios futuros posibles [12].

Sin embargo, una fuerte limitación metodológica no ha permitido una más extensa aplicación del concepto. Aún es una cuestión crítica y no resuelta para la gestión del desarrollo sostenible, la disponibilidad de metodologías que permitan evaluar el impacto del accionar de las instituciones y empresas en cada una de las dimensiones de la sostenibilidad, determinando el valor y las variables que cuantifiquen la verdadera contribución o aporte de esa gestión al desarrollo sostenible.

Una de las metodologías a emplear son las Redes Bayesianas, a partir de las cuales se puede obtener de una forma gráfica las relaciones entre las variables consideradas de cada una de las cuatro dimensiones, con objeto de poder determinar a posteriori los valores que cuantifiquen su contribución a la sostenibilidad.

En la tabla siguiente (Tabla 1), se incluyen los principales trabajos sobre sistemas de transportes que han sido desarrollados usando Redes Bayesianas:

2. METODOLOGÍA Y RESULTADOS OBTENIDOS

Para la consecución del objetivo de caracterizar los parámetros físicos de las terminales de contenedores del sistema portuario español, mediante redes bayesianas, se ha desarrollado la siguiente metodología (Figura 2). Esta se divide en dos tareas: una para determinar el escenario de trabajo

y la segunda para desarrollar el modelo de inteligencia artificial.

Las redes bayesianas están diseñadas para hallar las relaciones de dependencia e independencia entre todas las variables que conforman el dominio de estudio. Basado en ello, se utilizan métodos de razonamiento probabilístico que permiten realizar predicciones sobre el valor de cualquier variable desconocida basados en los valores de las conocidas.

Las redes bayesianas proveen una forma compacta de representar el conocimiento y métodos flexibles de razonamiento –basados en las teorías probabilísticas– capaces de predecir el valor de variables no observadas y explicar las observadas. Entre las características que poseen las redes bayesianas se puede destacar que permiten aprender sobre relaciones de dependencia y causalidad, permiten combinar conocimiento con datos y pueden manejar bases de datos incompletas

Las redes bayesianas pueden realizar la tarea de clasificación –caso particular de predicción– que se caracteriza por tener una sola de las variables de la base de datos (clasificador) que se desea predecir, mientras que todas las otras son los datos propios del caso que se desea clasificar. Pueden existir una gran cantidad de variables en la base de datos, algunas de las cuales estén directamente relacionadas con la variable clasificadora pero también otras variables que tienen una influencia directa sobre dicha clase.

El obtener una red bayesiana a partir de datos es un proceso de aprendizaje que se divide en dos etapas: el aprendizaje estructural y el aprendizaje paramétrico [34]. La primera de ellas consiste en obtener la estructura de la red bayesiana, es decir, las relaciones de dependencia e independencia entre las variables involucradas. La segunda etapa tiene como finalidad obtener las probabilidades a priori y condicionales requeridas a partir de una estructura dada.

Tabla 1. Principales trabajos sobre sistemas de transportes que han sido desarrollados usando Redes Bayesianas

Año	Título	Referencia
1996	Building classifiers using Bayesian networks	[13]
1997	Marginal costs and scale economies in Spanish ports	[14]
1998	Bayesian inference on network traffic using link count data	[15]
2001	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	[16]
2002	Using Bayesian networks to manage uncertainty in student modeling. User Modeling and User-Adapted Interaction	[17]
2005	The use of hugin® to develop Bayesian networks as an aid to integrated water resource planning	[18]
2006	A Bayesian network approach to traffic flow forecasting	[19]
2006	Short-term freeway traffic flow prediction: Bayesian combined neural network approach	[20]
2006	Integrating Bayesian networks and decision trees in a sequential rule-based transportation model	[21]
2008	Traffic estimation and optimal counting location without path enumeration using Bayesian networks	[22]
2008	Predicting traffic flow using Bayesian networks	[23]
2008	A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation	[24]
2009	A cross-disciplinary approach to minimising the risks of maritime transport in the Gulf of Finland.	[25]
2010	The complex network of global cargo ship movements.	[26]
2012	Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network	[27]
2013	Modelo de eficiencia de las terminales de contenedores del sistema portuario español	[28]
2013	Utilización de redes bayesianas como método de caracterización de parámetros físicos de las terminales de contenedores del sistema portuario español	[29]
2014	Diseño de un modelo de planificación de zonas de actividades logísticas mediante el empleo de redes bayesianas	[30]
2014	Bayesian network with quantitative input for maritime risk analysis	[31]
2015	Setting the Port Planning Parameters in Container Terminals through Bayesian Networks	[32]
2016	Aplicaciones tecnológicas en la logística de transportes portuarios. Las terminales de contenedores	[33]

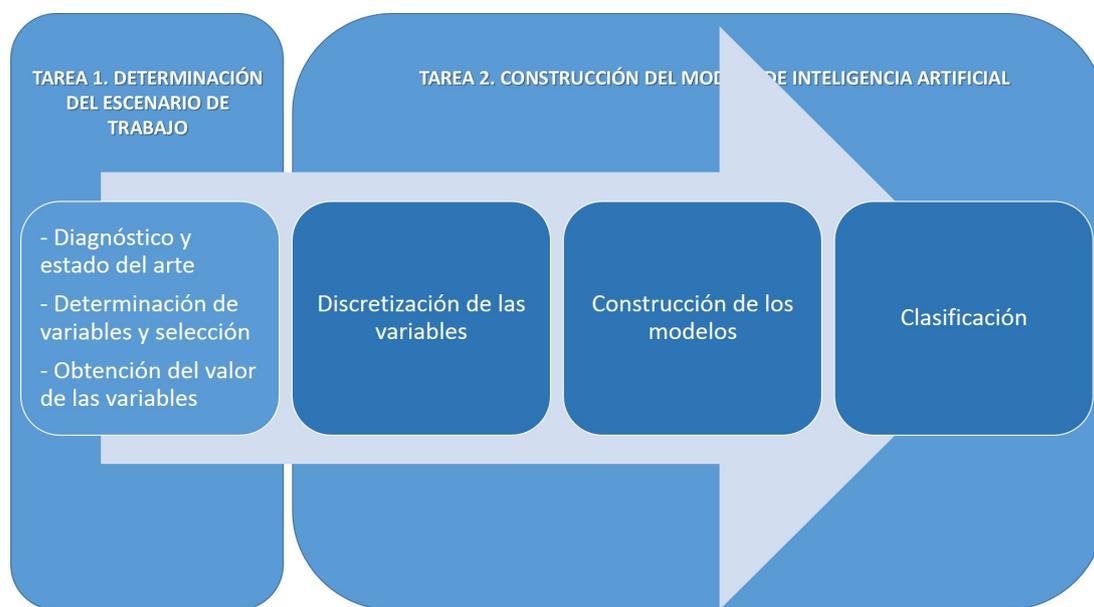


Figura 2. Esquema metodológico.

2.1. Tarea 1: determinación del escenario de trabajo

2.1.1. Diagnóstico y estado del arte

Consiste en la revisión del estado del arte para identificar el conjunto de variables de medida de la sostenibilidad, mediante el empleo de buscadores especializados y gestores de aplicaciones.

2.1.2. Determinación de las variables y selección

Se realiza un estudio las variables de sostenibilidad susceptibles de investigación para los puertos marítimos. Dada la variabilidad de las terminales y puertos componen el sistema portuario español se debe recurrir a las memorias de sostenibilidad de las autoridades portuarias del sistema portuario español.

2.1.3. Obtención del valor de las variables

El Sistema Portuario español de titularidad estatal está integrado por 46 puertos de interés general, gestionados por 28 Autoridades Portuarias, cuya coordinación y control de eficiencia corresponde al Organismo Público Puertos del Estado, órgano dependiente del Ministerio de Fomento y que tiene atribuida la ejecución de la política portuaria del Gobierno. (www.puertos.es/es-es). Los valores de las variables con las que se trabaja para la construcción del modelo de la red bayesiana corresponden a los datos de las Memorias de Sostenibilidad publicadas anualmente por las Autoridades Portuarias de los puertos españoles y completados con información suministrada por el Organismo Público Puertos del Estado, corresponden al registro histórico desde el año 2010, contando con casi 3000 registros.

Las variables seleccionadas en el estudio se incluyen en la Tabla 2.

Sin embargo, los modelos directos tienen una noción más compleja de independencia que los modelos indirectos, pero tienen, en cambio, varias ventajas. La ventaja principal de estos modelos es que se puede crear un arco desde A hasta B para indicar que A es la «causa» de B, lo cual se puede usar para crear la estructura de grafo. Es por ello que los modelos directos pueden codificar relaciones determinísticas, siendo más fáciles de aprender (encajar los datos).

Asimismo, es necesario especificar los parámetros del modelo para así poder definir la estructura del grafo. En el caso de modelos directos de Distribución Condicional de la Probabilidad (CPD), dicha definición se debe especificar para cada uno de los nodos. En el caso de que las variables sean discretas, se deben representar como en la tabla (CPT), es decir, definiendo la probabilidad que un nodo hijo tiene para cada uno de los diferentes valores para cada una de las combinaciones de valores de sus nodos padres.

2.2. Tarea 2: construcción del modelo de inteligencia artificial

Los modelos gráficos de probabilidad son grafos en los cuales los nodos representan variables aleatorias y en los que la existencia o la falta de arcos representan supuestos de independencia condicional. Por lo tanto, proporcionan una representación completa de la articulación de las distribuciones de la probabilidad. Los modelos gráficos indirectos, también llamados Campos aleatorios de Markov (MRFs) o Redes de Markov, cuenta con una definición simple de independencia: dos nodos o conjunto de nodos, A y B, son condicionalmente independientes considerando un tercer conjunto (C), si todos los caminos existentes entre los nodos A y B están separados por un nodo en C. Por el contrario, los modelos gráficos directos, también llamados Redes Bayesianas o Redes de Creencias (BNs), cuentan con una noción más compleja de independencia, la cual tiene en cuenta la direccionalidad de los arcos, tal y como se explica a continuación [35].

Los modelos gráficos de tipo indirectos son más populares en el caso de comunidades físicas y de visión, mientras que los modelos directos lo son en comunidades AI y estadísticas. No obstante, es posible tener un modelo en el cual se incluyen arcos directos e indirectos, denominándose a este modelo como Grafo encadenado. Para realizar un estudio más exhaustivo

de las relaciones entre los modelos gráficos directos e indirectos, revisar las siguientes referencias: [36], [37] y [38].

Con objeto de especificar la Red Bayesiana y así representar totalmente las uniones en la distribución de la probabilidad, se necesita definir la distribución de la probabilidad condicionada de X hacia sus padres en cada uno de los nodos X. La distribución condicional de X hacia sus padres puede adoptar cualquier forma. Es muy común el uso de distribuciones discretas o Gaussianas cuando esto simplifica los cálculos. Otras veces, sólo se conocen las restricciones de la distribución, en cuyo caso, se puede usar el máximo principio de la entropía para definir una distribución sencilla: tomar el que tiene una mayor entropía de las restricciones. De forma análoga, en el contexto específico de una Red Bayesiana dinámica, es habitual que se especifique la distribución condicional del estado oculto de la evolución temporal para maximizar el ratio de proceso estocástico implícito.

A menudo, estas distribuciones condicionales incluyen parámetros que son desconocidos y deben ser estimados a partir de los datos. Algunas veces, los parámetros se estiman usando el enfoque de máxima verosimilitud. La maximización de la verosimilitud (o probabilidad posterior) de forma directa resulta generalmente complejo cuando existen variables no observables. Un enfoque típico de este problema es el algoritmo de expectación-maximización. Este algoritmo alterna el cálculo de los valores esperados de las variables no observadas, que se determinan a partir de los datos observados, y la maximización de la verosimilitud completa (o posterior), considerando que los valores esperados calculados previamente son correctos. Bajo condiciones de regularidad suave, este proceso converge en valores de parámetro de máxima verosimilitud (o máximo posterior).

Un enfoque más profundo de los parámetros bayesianos es tratar a los parámetros como variables no observadas adicionales y computarlas en una distribución completa posterior con todos los nodos condicionales sobre los datos observados, con objeto de sacar los parámetros. Este enfoque puede ser costoso y conducir a modelos de gran dimensión, por lo que en la práctica es más común el uso de métodos clásicos de parametrización.

Dado que una Red Bayesiana es un grafo acíclico dirigido, se puede interpretar de dos formas:

- Distribución de probabilidad: Representa la distribución de la probabilidad conjunta de las variables representadas en la red.
- Base de reglas: cada arco representa un conjunto de reglas (cuantificadas por las probabilidades respectivas) que asocian a las variables involucradas.

La estructura de una Red Bayesiana se puede determinar de la siguiente manera:

- 1) Se asigna un vértice o nodo a cada variable (X_j) y se indica de qué otros vértices es una causa directa; a ese conjunto de vértices «causa del nodo (X_j)» se lo denota como el conjunto πX_j y se lo llamará «padres de X_j ».
- 2) Se une cada padre con sus hijos con flechas que parten de los padres y llegan a los hijos.
- 3) A cada variable X_j se le asigna una matriz $P(X_j|\pi X_j)$ que estima la probabilidad condicional de un evento $X_j = x_j$

Tabla 2. Variables de trabajo.

Dimensión	ID	Marca temporal	Descripción
Dimensión Institucional (ID: dimins): Conjunto de objetivos e indicadores de sostenibilidad que reflejan prioridades estratégicas en el ámbito de la sostenibilidad para el conjunto del sistema portuario de interés general.	herramgestion_dimins	Herramientas de apoyo a la gestión	Sistemas de gestión de apoyo a la toma de decisiones: sistemas de gestión de la calidad, cuadros de mando integral, campañas de caracterización de mercados, etc.
	geninfraortuaria_dimins	Generación de infraestructura portuaria	Papel de la Autoridad Portuaria como proveedor de la infraestructura
	mercservidos_dimins	Mercados servidos	Estructura y evolución de los principales tráfico de mercancías
	dinamact_dimins	Papel del sector portuario como dinamizador de la actividad productiva	Principales sectores o actividades relevantes en el desarrollo económico local que se apoyan en el puerto para su desarrollo
	serviciosconcauto_dimins	Servicios y concesiones/ autorizaciones	Tipos, marco de prestación y regulación
	inicprivada_dimins	Presencia de la iniciativa privada	Número de empresas que operan en el puerto, superficie terrestre ocupada, caracterizada como uso comercial concesionado, etc.
	transconcu_dimins	Transparencia y libre concurrencia	Iniciativas dirigidas a garantizar que todo operador que desee prestar servicios en el puerto u optar a una concesión pueda conocer de modo transparente las condiciones para operar en el puerto y los mecanismos administrativos que regulan dicho proceso
	calidserv_dimins	Calidad en la prestación de los servicios	Iniciativas promovidas por la Autoridad Portuaria dirigidas a mejorar la eficiencia, la calidad del servicio y el rendimiento de los servicios prestados a la mercancía
	intetrans_dimins	Integración de los puertos en el sistema de transporte	Eficiencias con la que son coordinados los diferentes modos de transporte que confluyen en el puerto
Dimensión económica (ID: dimecon): Situación económica financiera del sistema portuario español	sitecofin_dimecon	Situación económica financiera	Entre otros indicadores: rentabilidad sobre activos, EBIDTA/tonelada, servicio de la deuda, relación gastos de explotación e ingresos de explotación, etc.
	inv_dimecon	Nivel y estructura de las inversiones	Entre otros indicadores: inversión pública en relación con el cash-flow, inversión ajena frente a la inversión pública, renovación de activos
	negserv_dimecon	Negocio y servicios	Entre otros indicadores: ingresos por tasas de ocupación y actividad, uso comercial de la superficie, uso de los muelles, etc.
	vgenprod_dimecon	Valor generado y productividad	Entre otros indicadores: productividad del trabajo según ingresos, productividad del trabajo según EBIDTA, etc.
Dimensión social (ID: dimsoci): Capital humano y empleo	caphum_dimsoci	Capital humano de la actividad portuaria	Empleo, comunicación interna y participación, formación, estructura de plantilla y equidad, seguridad y salud en el trabajo, etc.
	empl_dimsoci	Empleo y seguridad laboral en la comunidad portuaria	Empleo en la comunidad portuaria, seguridad laboral y formación en servicios y concesiones portuarios, etc.
Dimensión medioambiental (ID: dimma): Calidad y gestión ambiental	gestamb_dimma	Gestión ambiental	Grados de implantación de los sistemas de gestión ambiental (EMAS, ISO 14001 y PERLS) y recursos económicos invertidos gastos, así como inversiones en su caso, asociados a la implantación, certificación y mantenimiento de un sistema de gestión ambiental
	calaire_dimma	Calidad del aire	Principales focos de emisión del puerto que suponen emisiones significativas, evolución del número de quejas o denuncias registradas por la Autoridad Portuarias procedentes de grupos de interés relativas a emisiones de polvo o a la calidad del aire en general, medidas implantadas por la Autoridad Portuaria para controlar las emisiones ligadas a la actividad del conjunto del puerto
	calagua_dimma	Calidad del agua	Principales focos de vertido situados en el puerto que tienen un impacto significativo en la calidad del agua y sedimentos de las dársenas del puerto, medidas implantadas por la Autoridad Portuaria para controlar las emisiones ligadas a la actividad del conjunto del puerto, superficie de la zona de servicio que cuenta con recogida y tratamiento de aguas residuales
	calacust_dimma	Calidad acústica	Principales focos de emisión (puntuales y difusos) del puerto que suponen emisiones acústicas significativas, evolución del número de quejas o denuncias registradas por la Autoridad Portuaria procedentes de grupos de interés, elaboración de mapa de ruido y plan de acción acústica
	residuos_dimma	Gestión de residuos	Residuos generados por la Autoridad Portuaria que son segregados y valorizados, actividades o fuentes de generación de residuos dentro del puerto, iniciativas promovidas por la Autoridad Portuaria para la mejora de la gestión de residuos de la comunidad portuaria
	ecoef_dimma	Ecoeficiencia	Eficiencia en el uso del suelo, consumo de agua y energía eléctrica por la Autoridad Portuaria
	comport_dimma	Comunidad portuaria	Condiciones o exigencias sobre aspectos ambientales en los pliegos de prescripciones técnicas particulares de los servicios portuarios, en términos de otorgamiento y en títulos de concesión o autorización

dada una combinación de valores de los πX_j . Las Redes Bayesianas permiten definir modelos y utilizarlos tanto para hacer razonamiento de diagnóstico (pues obtienen las causas más probables dado un conjunto de síntomas), como para hacer razonamiento predictivo (obteniendo la probabilidad de presentar un cierto síntoma suponiendo que existe una causa conocida). Una de las características de las Redes Bayesianas es que un mismo nodo puede ser fuente de información u objeto de predicción dependiendo de cuál sea la evidencia disponible. A continuación se muestran cuáles son las características de estos dos tipos de inferencia utilizando una Red Bayesiana.

2.2.1. Discretización de variables

A menudo conviene representar un fenómeno continuo en la naturaleza usando variables discretas. Para ello, las medidas continuas tienen que ser discretizadas. Esto puede hacerse proyectando la escala de valores continua en un conjunto finito de intervalos. Los valores que caigan en el mismo rango se considerarán como un mismo estado.

Una vez seleccionadas las variables de estudio en las tareas anteriores es necesario, para el proceso de construcción de los modelos, la discretización de las variables. Normalmente las redes bayesianas consideran variables discretas o nominales, por lo que si no lo son, hay que discretizarlas antes de construir el modelo. Aunque existen modelos de redes bayesianas con variables continuas, éstos están limitados a variables gaussianas y relaciones lineales.

2.2.2. Construcción de los modelos

Después de definir las variables, el siguiente paso en la construcción de un modelo es definir su estructura. Esto se hace conectando variables con arcos (también llamados enlaces). En las redes bayesianas los arcos son dirigidos. Cambiar la dirección de un arco cambia su significado. La ausencia de un arco entre dos variables indica que no existen relaciones de dependencia directa entre ellas, sino a lo sumo a través de otras variables. La presencia de un arco indica una relación de influencia causal entre dos variables.

En esta parte, el aprendizaje estructural consiste en encontrar las relaciones de dependencia entre las variables, para determinar la topología o estructura de la red bayesiana. De acuerdo al tipo de estructura, se aplican diferentes métodos de aprendizaje estructural: aprendizaje de árboles, aprendizaje de poli-árboles, aprendizaje de redes multiconectadas, métodos basados en medidas y búsqueda, métodos basados en relaciones de dependencia.

Uno de los principales problemas encontrados en este tipo de métodos es la cantidad tan grande de posibles estructuras. Es por ello que se han desarrollado diferentes heurísticas para no explorar todo el espacio de búsqueda, encontrándose entre ellos los algoritmos genéticos (Larragaña) y los algoritmos «greedy» como es el K2 [39] que se ha empleado en el estudio.

El mejor modelo es aquél que es más probable. Para ello resulta muy útil la teoría bayesiana pues proporciona herramientas para calcular la probabilidad de un modelo y posteriormente comparar las probabilidades de varios modelos, detectando cuál de ellos es el más probable.

No obstante, método más sofisticado es el denominado algoritmo K2, que se describe a continuación. Dicho algoritmo está basado en la optimización de una medida. Esa medida se usa para explorar, mediante el algoritmo, el espacio de búsqueda formado por todas las redes que contienen las variables de la base de datos. Así, partiendo de una red inicial, ésta se va modificando (añadiendo arcos, eliminándolos o cambiándolos de dirección) obteniendo una nueva red con mejor medida [39].

2.2.3. Clasificación

El último paso en el proceso de modelado es especificar parámetros, basta con proporcionar las probabilidades a priori de los nodos raíz y las probabilidades condicionales del resto de los nodos

En esta fase se aprovechan las características que poseen los métodos bayesianos en tareas de aprendizaje. Cada ejemplo observado va a modificar la probabilidad de que la hipótesis formulada sea correcta (aumentándola o disminuyéndola). Es decir, una hipótesis que no concuerda con un conjunto de ejemplos significativos no es desechada por completo sino que se disminuirá la probabilidad estimada para la hipótesis. Los métodos bayesianos permitirán tener en cuenta en la predicción de la hipótesis el conocimiento a priori o conocimiento del dominio en forma de probabilidades.

2.3. Resultados

Es necesario especificar dos aspectos para describir una Red Bayesiana: topología de grafo (estructura) y los parámetros de cada Distribución Condicional de Probabilidad (CPD). Esto hace posible generar ambos a partir de los datos. Sin embargo, definir la estructura es más difícil que definir los parámetros. Además, cuando algunos de los nodos están ocultos o faltan datos, definirlo es más difícil que definirlo cuando se observa todo.

La Red Bayesiana es desconocida en muchos de los sucesos prácticos y, por tanto, se necesita definir a partir de los datos. Este problema se conoce como problema de definición de una Red Bayesiana. Describiéndolo de una forma informal, sería: obtención de los datos base y priorización de la información (por ejemplo, conocimiento del experto, relación causal); estimación de la topología del grafo (estructura de la red) y distribución de los parámetros JPD (parámetros de probabilidad conjunta) en la Red Bayesiana.

El desarrollo de la estructura de una Red Bayesiana se considera un problema más complejo que la definición de los parámetros de dicha red. Asimismo, existe otro handicap en el caso de situaciones de observación parcial cuando los nodos están ocultos o cuando directamente faltan.

El caso simple es una Red Bayesiana definida y detallada por un experto. Por lo que se usa para desarrollar inferencias. En otras aplicaciones, la tarea de definir la red es demasiado compleja, siendo necesario que la estructura de la red y la distribución local de los parámetros sea desarrollada a partir de datos.

Automáticamente, el desarrollo de la estructura gráfica de una Red Bayesiana es un reto que se persigue dentro del aprendizaje de la máquina.

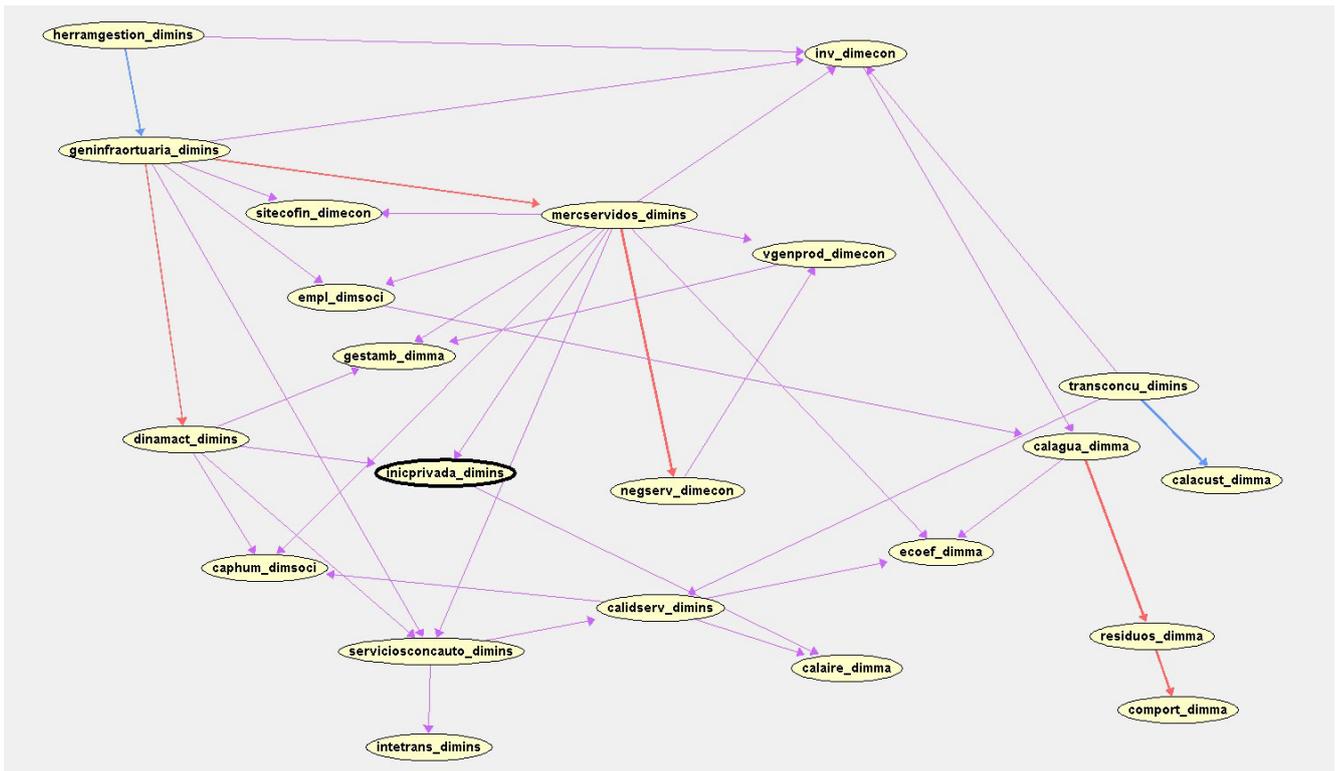


Figura 3. Red Bayesiana. Algoritmo K2

La red construida es la que se muestra en la Figura 3. Como se aprecia ninguna variable ha quedado descolgada.

Los pasos que son necesarios seguir son los siguientes:

- Identificación de los factores relevantes
- Determinación de cómo esos factores están relacionados causalmente entre ellos
- El arco causa-efecto significa que la causa es un factor involucrado en provocar ese efecto

A modo de ejemplo, las relaciones más significativas se muestran en el caso siguiente. En él, en el caso de las variables padres, *inv_dimecon* y *geninfraortuaria_dimins*, es el nodo de *herramgestion_dimins*. Cuando se da *herramgestion_dimins*, entonces *inv_dimecon* y *geninfraortuaria_dimins*, son condicionalmente independientes (Figura 4).

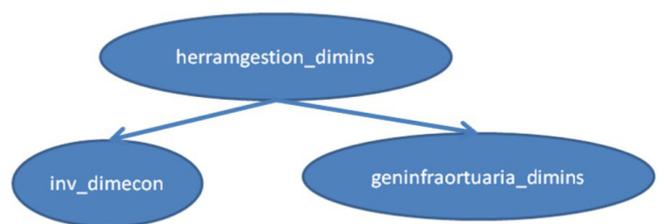


Figura 4. Relación 1.

La variable *herramgestion_dimins* es una variable decisora, aparece en la red como una variable «nodo» de la que salen arcos, por lo que se genera una conexión divergente, se tiene un nodo padre que proyecta todos sus arcos hacia varios hijos. O lo que es lo mismo las flechas salen de él y divergen hacia sus hijos (Figura 4).

Recordemos que Herramientas de apoyo a la gestión (ID: *herramgestion_dimins*) son los sistemas de gestión de apoyo a la toma de decisiones, es decir, sistemas de gestión de la calidad, cuadros de mando integral, campañas de caracterización de mercados, etc.

Cuando se conoce el estado de la variable padre existe una dependencia entre las variables, sin embargo, cuando el estado de la variable padre no se conoce, las variables hijo se tornan

independientes y la información no se propaga por la red si se añaden evidencias sobre los nodos hijos (Figura 5).

Un efecto que tiene dos o más arcos de entrada desde otros vértices se considera como un efecto común de esas causas. Una causa que tiene dos o más arcos de salida a otros vértices se considera como una causa común (factor) de esos efectos. Los efectos de una causa común son, normalmente, observables.

Siguiendo la suposición de independencia de la Red Bayesiana, en este caso, se observan varias declaraciones de independencia respecto de cada uno de los factores.

Cuando *mercervidos_dimins* se introduce, *sitecofin_dimecon*, *serviciosconcauto_dimins*, *inicprivada_dimins*, *ecoef_dimma*, *vgenprod_dimecon*, *empl-dimsoci*, *inv_dimecon*, *negserv_dimecon*, *gestamb_dimma*, and *caphum_dimensoci* son independientes condicionalmente o su antecesor *geninfraortuaria_dimins* (Figura 5).

Grafo casual: la variable *mercervidos_dimins* tiene diez efectos comunes: *sitecofin_dimecon*, *serviciosconcauto_di-*

mins, inicprivada_dimins, ecoef_dimma, vgenprod_dimecon, empl-dimsoci, inv_dimecon, negserv_dimecon, gestamb_dimma, and caphum_dimensoci (Figura 5).

3. CONCLUSIONES

De la red obtenida mediante al empleo del algoritmo K2 se aprecia que la categoría más decisora en la sostenibilidad: categoría institucional, luego económica y social a la misma altura, y por último la ambiental.

Los sistemas de gestión de apoyo a la toma de decisiones: sistemas de gestión de la calidad, cuadros de mando integral, campañas de caracterización de mercados, etc, representados por *herramgestion_dimins*. Se representa como una variable padre de la red (solo salen flechas de ese nodo). Lo mismo ocurre con *transconcu_dimins* que representa a las iniciativas dirigidas a garantizar que todo operador que desee prestar servicios en el puerto u optar a una concesión pueda conocer de modo transparente las condiciones para operar en el puerto y los mecanismos administrativos que regulan dicho proceso, que es un padre de la red. La variable *transcocu_dimins* es una variable decisora, aparece en la red como una variable «nodo» de la que salen arcos, por lo que se genera una conexión divergente, se tiene un nodo padre que proyecta todos sus arcos hacia varios hijos. O lo que es lo mismo las flechas salen de él y divergen hacia sus hijos.

Otra variable fundamental en la estructura de la red es *merc-servidos_dimins*, de ella salen 10 flechas a 10 nodos diferentes, son los efectos de la estructura y evolución de los principales tráficos de mercancías, estos efectos son tanto sociales, económicos, institucionales y medioambientales; es decir los mercados servidos tienen efectos sobre el tipos, marco de prestación

y regulación de los servicios del puerto, sobre el número de empresas que operan en el puerto, dentro de la categoría institucional; en la categoría económica sobre el EBIDTA, EBIDTA/tonelada, inversión pública en relación con el cash-flow y : ingresos por tasas de ocupación y actividad entre otros; sobre la categoría social ejerce un efecto sobre sus dos variables que representan el empleo en la comunidad portuaria, seguridad laboral y formación en servicios y concesiones portuarios, empleo, comunicación interna y participación, formación, estructura de plantilla y equidad, seguridad y salud en el trabajo, entre otros. Para la categoría medioambiental los mercados servidos son la causa de los grados de implantación de los sistemas de gestión ambiental (EMAS, ISO 14001 y PERLS) y recursos económicos invertidos gastos, así como inversiones en su caso, asociados a la implantación, certificación y mantenimiento de un sistema de gestión ambiental y de la eficiencia en el uso del suelo, consumo de agua y energía eléctrica por la Autoridad Portuaria. Por lo tanto los mercados servidos suponen una variable muy importante dentro de la planificación desde una perspectiva sostenible.

Las variables de la categoría dimensión institucional se encuentran muy relacionadas entre sí.

Las económicas fundamentales en términos de causa-efecto, son efecto de la variables de los mercados servidos que es de dimensión institucional. El valor generado y la productividad dependiente de negocio y servicios (ingresos por tasas de ocupación y actividad, uso comercial de la superficie, uso de los muelles, etc.).

Las variables de dimensión sociales son efecto de de variables de dimensión institucional pero no tienen una relación directa entre variables de la misma dimensión, la social.

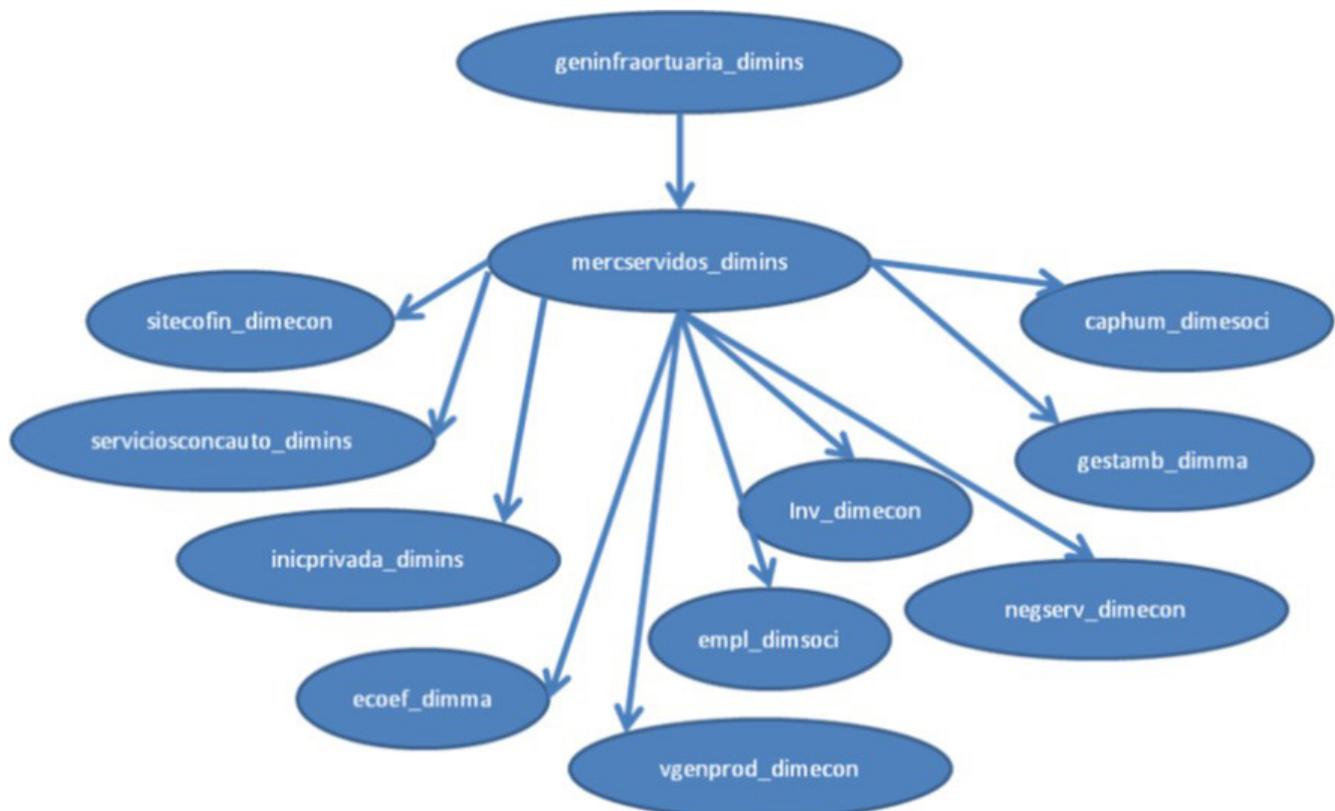


Figura 5. Relación 3.

Las variables de categoría ambiental se encuentran muy relacionadas entre ellas en la red bayesiana y son efectos de variables de categoría institucional principalmente. Por ello, aunque las Autoridades Portuarias no tienen competencias ambientales, desempeñan un papel clave en la adecuada gestión ambiental del puerto (ya que actúan como gestores de infraestructura, reguladores, coordinadores de los servicios prestados y, en especial como líderes de la comunidad). La actividad del puerto genera impactos tanto en el medio acuático como en el terrestre y aéreo, y este capítulo evalúa los impactos y las medidas implantadas para reducirlos, dándose información también sobre los sistemas de recepción y gestión de quejas ambientales disponibles en los puertos.

Las variables de las dimensiones económica, social y ambiental son efecto de variables de dimensión institucional. En consecuencia, un aspecto clave es que las Autoridades Por-

tuarias vayan incorporando a los instrumentos de que disponen para la regulación de los servicios portuarios y la gestión del dominio público, los elementos exigibles en materia de sostenibilidad que la ley de Puertos contempla.

Por lo tanto, se ha visto que las redes bayesianas permiten modelar la incertidumbre de forma probabilística incluso cuando el número de variables es elevado. A partir de la red construida, el paso siguiente para continuar con la investigación la inferencia, es decir conocer la distribución de unas variables cuando se observan otras, teniendo en cuenta que los métodos de razonamiento aproximado, entre los que se encuentran los métodos bayesianos, aportan modelos teóricos que simulan la capacidad de razonamiento en condiciones de incertidumbre, cuando no se conoce con absoluta certeza la verdad o falsedad de un enunciado o hipótesis, e imprecisión, enunciados en los que se admite un rango de variación.

REFERENCIAS

- (1) World Commission on Environment and Development (WCED) (1987). *Our Common Future* (Brundtland Report), United Nations
- (2) Madero Gómez, S.M.; Zarate Solís, I.A. (2017). La sostenibilidad desde una perspectiva de las áreas de negocios. *Cuadernos de Administración*, vol. 32, no 56, p. 7-19
- (3) Doerr, O. (2011). Políticas portuarias sostenibles. *Boletín FAL. CEPAL*. Edición nº 299, número 7 de 2011
- (4) Global Reporting Initiative (2000). *Guía para la elaboración de Memorias de Sostenibilidad*. Versión 3.1. GRI
- (5) Serrano, O. (2015). Operativa portuaria y sostenibilidad. *CONAMA LOCAL 2015*, 7 octubre de 2015. Málaga
- (6) Sánchez, R.J.; Jauimurzono, A.; Willmsmeier, G.; Pérez Salas, G.; Doerr, O.; Pinto, F. (2015). Transporte marítimo y Puertos. Desafío y oportunidades en busca de un desarrollo sostenible de América Latina y El Caribe. CEPAL. Serie Recursos Naturales e Infraestructura.
- (7) Crespo Soler, C.; Giner Fillol, A.; Morales Baraza, J.A.; Ponte Tubal, N; Ripoll Feliu; V.M. (2007). La información de sostenibilidad en el marco de las cuentas anuales: análisis aplicado al caso de la Autoridad Portuaria de Valencia. *Revista de Contabilidad de Maestrado em Ciências Contábeis da UERJ, Rio Janeiro*, v-12, n.3, p-11 set./dez, 2007
- (8) Grupo De Trabajo 23 (2004). La sostenibilidad en los puertos. *CONAMA VII Cumbre del desarrollo sostenible*, 24 de noviembre de 2004
- (9) Crespo Soler, C.; Ripoll Feliu; V.M., Crespo Trujillo, A.M.; Giner Fillol, A. (2005). La sostenibilidad ambiental en el sistema portuario de titularidad estatal. XIII Congreso AECA. Armonización y Gobierno de la Diversidad. 22-24 Septiembre 2005
- (10) González, F.; Guerra, A.; Martín, F.; Nóvoa, J.J.; Otero, C. Y Penela, J. (2010). 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
- (11) Autoridad Portuaria De A Coruña, Autoridad Portuaria De Valencia, Organismo Público Puertos Del Estado (2008). *Guía para la elaboración de memorias de sostenibilidad en el sistema portuario español*. FEPORTS
- (12) Puertos Del Estado (2011). *Memoria de sostenibilidad del sistema portuario de interés general*.
- (13) Friedman, N., & Goldszmidt, M. (1996). Building classifiers using bayesian networks. *Proceedings of the National Conference on Artificial Intelligence*, Menlo Park, Ca: AAAI Press
- (14) Jara-Díaz, S., Martínez-Budría, E., Cortes, C., & Vargas, A. (1997). Marginal costs and scale economies in Spanish ports. 25th European Transport Forum, Proceedings Seminar L, PTRC, London, pp. 137-147.
- (15) Tebaldi, C., & West, M. (1998). Bayesian inference on network traffic using link count data. *Journal of the American Statistical Association*, 93(442), 557-573.
- (16) Cain, J. (2001). 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.
- (17) Conati, C., Gertner, A. Y Vanlehn, K. (2002). Using bayesian networks to manage uncertainty in student modeling. *User Modeling and User-Adapted Interaction*, 12, 371-417. doi: 10.1023/A:1021258506583
- (18) Bromley, J., Jackson, N. A., Clymer, O. J., Giacomello, A. M., & Jensen, F. V. (2005). The use of hugin® to develop bayesian networks as an aid to integrated water resource planning* 1. *Environmental Modelling & Software*, 20(2), 231-242. <https://doi.org/10.1016/j.envsoft.2003.12.021>
- (19) Sun, S., Zhang, C., & Yu, G. (2006). A bayesian network approach to traffic flow forecasting. *IEEE Transactions on Intelligent Transportation Systems*, 7(1), 124-132. doi: 10.1109/TITS.2006.869623
- (20) Zheng, W., Lee, D. H., & Shi, Q. (2006). Short-term freeway traffic flow prediction: Bayesian combined neural network approach. *Journal of transportation engineering*, 132(2), 114-121. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2006\)132:2\(114\)](https://doi.org/10.1061/(ASCE)0733-947X(2006)132:2(114))

- (21) Janssens, D., Wets, G., Brijs, T., Vanhoof, K., Arentze, T., & Timmermans, H. (2006). Integrating Bayesian networks and decision trees in a sequential rule-based transportation model. *European Journal of operational research*, 175(1), 16-34. <https://doi.org/10.1016/j.ejor.2005.03.022>
- (22) Castillo, E., Menéndez, J. M., & Sánchez-Cambronero, S. (2008). Traffic estimation and optimal counting location without path enumeration using Bayesian networks. *Computer-Aided Civil and Infrastructure Engineering*, 23(3), 189-207. doi: 10.1111/j.1467-8667.2008.00526.x
- (23) Castillo, E., Menéndez, J. M., & Sánchez-Cambronero, S. (2008). Predicting traffic flow using Bayesian networks. *Transportation Research Part B: Methodological*, Vol. 42, No. 5; 482-509. <https://doi.org/10.1016/j.trb.2007.10.003>
- (24) Trucco, P., Cagno, E., Ruggeri, F., & Grande, O. (2008). A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation. *Reliability Engineering & System Safety*, 93(6), 845-856. <https://doi.org/10.1016/j.res.2007.03.035>
- (25) Klemola, E., Kuronen, J., Kalli, J., Arola, T., Hanninen, M., Lehtikoinen, A. & Tapaninen, U. (2009). A cross-disciplinary approach to minimising the risks of maritime transport in the Gulf of Finland. *World Review of Intermodal Transportation Research*, Vol. 2, No. 4, 343-363. <https://doi.org/10.1504/WRITR.2009.026212>
- (26) Kaluza, P., Kölzsch, A., Gastner, M. T., & Blasius, B. (2010). The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, rsif20090495. doi: 10.1098/rsif.2009.0495
- (27) Hofleitner, A., Herring, R., Abbeel, P., & Bayen, A. (2012). Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network. *IEEE Transactions on Intelligent Transportation Systems*, 13(4), 1679-1693. doi: 10.1109/TITS.2012.2200474
- (28) Cancelas, N. G., Flores, F. S., & Orive, A. C. (2013). Modelo de eficiencia de las terminales de contenedores del sistema portuario español. *Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA. Rect@*, Vol. 14; 49-67.
- (29) Camarero, A., González-Cancelas, N., Soler, F., & López, I. (2013). Utilización de redes bayesianas como método de caracterización de parámetros físicos de las terminales de contenedores del sistema portuario español. *Revista de Ingeniería*, (39), 31-38.
- (30) Flores, F. S., Cancelas, N. G., Orive, A. C., Gárate, J. L. A., & Monzón, M. D. C. P. (2014). Diseño de un modelo de planificación de zonas de actividades logísticas mediante el empleo de redes bayesianas. *Revista Ingeniería Industrial*, 12(1).
- (31) Li, K. X., Yin, J., Bang, H. S., Yang, Z., & Wang, J. (2014). Bayesian network with quantitative input for maritime risk analysis. *Transportmetrica A: Transport Science*, 10(2), 89-118. <http://dx.doi.org/10.1080/18128602.2012.675527>
- (32) Rodríguez García, T., González Cancelas, N., & Soler-Flores, F. (2015). Setting the Port Planning Parameters In Container Terminals through Bayesian Networks. *PROMET-Traffic&Transportation*, 27(5), 395-403. <http://dx.doi.org/10.7307/ptt.v27i5.1689>
- (33) García, T. R. (2016). Aplicaciones tecnológicas en la logística de transportes portuarios. Las terminales de contenedores. *Revista Transporte y Territorio*, (14), 5-26. <http://revistascientificas.filo.uba.ar/index.php/rtt/article/view/2426>
- (34) Pearl J. (1988) Probabilistic reasoning in intelligent systems: Networks of plausible inference. Morgan Kaufmann.
- (35) Almazán-Gárate, J.L.; Palomino-Monzón, M.C.; González-Cancelas, N.; Soler-Flores, F. (2014). Relationship between air pollution and natural gas with respect to maritime transport. Methodology based on Bayesian Networks. Global Virtual Conference. 7-11 April 2014. Transport and Logistics Section.
- (36) Castillo, E.; Gutiérrez, J.M. and Hadi, A.S. (1997). *Expert Systems and Probabilistic Network Models*. Springer Verlag, New York
- (37) Duda, R.O.; Hart, P. E. and Stork, D.G. (2001). *Pattern Classification*. Wiley, New York
- (38) Pearl, J. (1982). The Solution for the Branching Factor of the Alpha-Beta Pruning Algorithm and its Optimality. *Communications of the ACM*. 1982. Vol 25, no.8. doi: 10.1145/358589.358616
- (39) Cooper, G. and E. Herskovits (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning* 9, 309-348

* * *