

Aplicación de inteligencia artificial para analizar el uso del gas natural licuado en el Short Sea Shipping en Europa: análisis específico de las componentes social y poblacional



Beatriz Molina Serrano

Departamento de Ingeniería Civil, Transportes, Universidad Politécnica de Madrid, España

Nicoletta González-Cancelas

Departamento de Ingeniería Civil, Transportes, Universidad Politécnica de Madrid, España

Francisco Soler-Flores

Facultad de Educación, Universidad Internacional de La Rioja, España

Alberto Camarero Orive

Departamento de Ingeniería Civil, Transportes, Universidad Politécnica de Madrid, España

Recibido: 22 de agosto de 2017. Aceptado: 31 de octubre de 2017

Resumen

Los problemas de congestión en los principales corredores europeos como consecuencia de la globalización de la economía provocaron la necesidad de acometer medidas y desarrollar políticas para evitar las externalidades generadas por tal situación. Así, la Unión Europea promovió el Short Sea Shipping como una alternativa, si bien se observó que esta medida ocasionó problemas ambientales en las zonas costeras, mostrando la necesidad de metodologías que ayuden a mitigar estos efectos. Así, el gas natural licuado se muestra como una alternativa a la hora de buscar combustibles más amigables. Es por ello que se ha llevado a cabo una investigación desarrollando una herramienta de inteligencia artificial que permite a los gestores tomar decisiones para minimizar la congestión del transporte en Europa basado en una metodología con redes bayesianas que permite conocer las relaciones entre las variables y clasificarlas. El principal resultado obtenido establece que la capacidad de las terminales de regasificación de gas natural licuado en construcción y la distribución modal de transporte de carga por aguas continentales son las dos variables más importantes en la decisión de implementar el gas natural licuado como combustible. Sin embargo, las variables de las categorías de medioambiente y energía y población y condición social son las más dependientes, si bien estas últimas se relacionan mucho entre sí, aunque no inciden directamente sobre variables relacionadas con el gas natural licuado y el Short Sea Shipping, pero sí sobre el producto interior bruto de cada país.

Palabras clave

Gas natural licuado
Short sea shipping
Inteligencia artificial
Población y condición social

Artificial intelligence application to analyze liquid natural gas use in European short sea shipping: Specific analysis of social and population components

Abstract

The problems of congestion in the main European corridors as a consequence of the globalization of the economy made it necessary to undertake measures and develop policies to avoid the externalities generated by this situation. Thus, the European Union promoted the Short Sea Shipping as an alternative, although it was observed that this measure caused environmental problems in the coastal areas with the environment in the Short Sea Shipping and methodologies are needed to help mitigate these effects. Thus, liquefied natural gas is shown as an alternative in the need to look for more friendly fuels. That is why research has been carried out developing an artificial intelligence tool that allows managers to make decisions to minimize transport congestion in Europe based on a Bayesian network methodology that allows to know the relationships between variables and classify them. The main result obtained establishes that the capacity of liquefied natural gas regasification terminals under construction and the modal distribution of cargo transport by inland water are the two most important variables in the decision to implement liquefied natural gas as fuel. However, the variables of the categories of environment and energy and population and social condition are the most dependent, although the latter are closely related to each other and do not directly affect variables related to liquefied natural gas and Short Sea Shipping, but yes on the gross domestic product of each country.

Keywords

*Liquefied natural gas
Short sea shipping
Artificial intelligence
Population and social status*

Palavras-chave

*Gás natural liquefeito
Transporte marítimo de curta
distância
Inteligência artificial
População e condições sociais*

Introducción

El transporte marítimo en Europa tiene una gran importancia para la cohesión y desarrollo económico de la Unión Europea, pero a su vez es una fuente de contaminación, sobre todo en las zonas próximas a la costa que es normalmente donde vive la mayor parte de la población (Orive, Orive, Cancelas & Guijarro, 2012).

El crecimiento del comercio internacional en los años 90 del siglo pasado llevó aparejado un aumento en la congestión de los principales corredores europeos, generando importantes externalidades (principalmente contaminación y accidentalidad). Como respuesta, se planteó la implantación de cadenas intermodales en las que cada modo de transporte se emplea en su tramo más eficiente (González, 2016). Así, el objetivo era llevar la carga la mayor parte del recorrido por vía marítima, por su eficiencia energética, por producir poco daño medioambiental, aumentar seguridad y reducir la saturación de la infraestructura (Cánovas, 2015). Por esta razón, la Comisión Europea promueve el Short Sea Shipping como una alternativa, puesto que complementa el transporte por carretera, es más eficiente energéticamente, contamina menos y resta congestión a las infraestructuras viales (Comisión Europea, 2001).

En los proyectos RECORDIT (RECORDIT, 2003) y REALISE (REALISE, 2002) y (REALISE, 2005) se demuestra que el transporte intermodal reduce los costos internos y externos totales, lo que hace que este se convierta en una mejor alternativa económica cuando se compara con el transporte unimodal (Lloyd, 2003) y (Vassallo, et al., 2004). Dentro de la estrategia también se buscaba alcanzar un desarrollo territorial sostenible, el cual pretende “alcanzar un equilibrio a largo plazo entre el desarrollo económico, la protección del medio ambiente, el uso eficiente de los recursos y la equidad social” (Awad, González y Camarero,

2015). Para la planificación portuaria, una metodología que se considera innovadora y novedosa sería la empleada por Macías, González-Cancelas y Soler-Flores en su estudio, desarrollada mediante la utilización de herramientas como el Análisis DAFO (también llamado análisis FODA o DOFA), DELPHI y la explotación estadística descriptiva del programa Labbtex (Macías, González-Cancelas & Soler-Flores, 2016). Por otro lado, dado que las terminales son uno de los principales nodos logísticos, la mercancía que estas mueven cada año representa unos volúmenes de inversión millonarios, y consecuentemente los usuarios exigen una seguridad elevada y garantías para ella (Romero, 2016). Existen diversas metodologías para realizar una identificación, valoración y análisis de los riesgos sobre activos en cualquier unidad de negocio, logístico o de operación, tales como el planteado por (Sáenz-Löbsack, Camarero y González, 2016). Dicho método defiende un sistema de presentación muy estructurado, sintético y, sobre todo, gráfico, que permite una rápida interpretación de los resultados y facilita claramente la validación, análisis y toma de decisiones derivados del análisis. Así, define los riesgos posicionándolos en una matriz dimensional y los califica mediante un conjunto de símbolos.

Con este impulso del Short Sea Shipping por parte de la Comisión Europea se esperaba optimizar los modos de transporte, descongestionar las carreteras, reducir las emisiones de dióxido de carbono (CO₂) y óxidos de azufre (SO_x) y ofrecer un transporte seguro y de calidad, garantizando la conexión entre las vías navegables y el ferrocarril y la creación de las autopistas del mar (Cánovas, 2015), teniendo en cuenta que las instalaciones deben producir el menor impacto en los ciudadanos, causándoles el menor trastorno posible, y minimizando los cambios en sus costumbres, siempre buscando el bienestar de la comunidad (Palomino, Almazán, González y Soler, 2012).

Sin embargo, la puesta en marcha de esta política incrementó la contaminación, pues el sector de la navegación ha sido una de las fuentes más importantes de contaminación. Así surgió la necesidad de buscar combustibles alternativos para reducir los niveles de contaminación en zonas costeras densamente pobladas y con presencia de alto tráfico de embarcaciones (Comisión Europea, 2006).

Un combustible alternativo es el gas natural licuado que cuenta con emisiones de dióxido de carbono (CO₂) inferiores al tradicional y con ausencia de emisiones de dióxido de azufre (SO₂) y óxidos de nitrógeno (NO_x) (Bengtsson, Andersson y Fridell, 2011). Su uso respeta los reglamentos de la política de transportes en Europa y permitirá, en el futuro, una diversificación de las fuentes de abastecimiento energético, ahorrará costos, generará un impacto positivo en el medio ambiente, contribuyendo a evitar el cambio climático (Orive, de la Fuente, López, Domínguez y Cancelas, 2012) al tiempo que fortalecerá las autopistas del mar (Acedo y Almazán, 2015).

Dada la importancia de este nuevo combustible, se ha puesto en marcha un proyecto, el proyecto CORE LNGas hive, para el impulso del gas natural licuado como combustible en el transporte. Dicho proyecto está liderado por Puertos del Estado y coordinado por Enagás, y que cuenta con un total de 42 socios de España y Portugal: 8 socios institucionales; 13 autoridades portuarias, y 21 socios industriales -operadores de GNL, constructoras navales, regasificadoras y otras empresas-. Contempla la realización de 25 estudios, impulsados por los socios, para la adaptación de las infraestructuras y el desarrollo logístico-comercial, que permitan la prestación de servicios small scale (abastecimiento a pequeña escala) y bunkering (suministro de GNL como combustible para barcos) (Enagas, 2017).

No obstante, en la actualidad, el uso del gas natural licuado en el transporte marítimo no es significativo, siendo también escasos los estudios que analizan el cambio de combustible. Por tanto, el objetivo del presente artículo es aplicar una metodología para determinar la relación entre las variables que definen el uso del gas natural

licuado en el Short Sea Shipping en Europa, mediante técnicas de minería de datos que permita establecer las variables más relevantes en la toma de decisiones y ayudar en la planificación y toma de decisiones en la política de transporte europea. En base a los resultados obtenidos, se podría analizar la extrapolación de dicha metodología y sus resultados a la problemática Argentina y regional, así como a otros ámbitos de estudio.

Las redes bayesianas abordan un determinado problema mediante un conjunto de variables y las relaciones de dependencia entre ellas. Estos modelos bayesianos poseen diferentes aplicaciones para diagnóstico, clasificación y decisión que brinde información importante en cuanto a cómo se relacionan las variables, las cuales pueden ser interpretadas como relaciones de causa-efecto. La elección de esta metodología se explica debido a que permite relacionar conjuntos de probabilidades, por lo que permite tener una afirmación o valor acertado, debido a su base de probabilidades. Además, las redes bayesianas permiten reducir el sobre ajuste de los datos y posibilitan el descubrimiento de la estructura causal subyacente en un conjunto de datos. Por lo que permite al planificador conocer la causa-efecto de sus acciones.

Metodología

Ante la necesidad de manejar una gran cantidad de datos surgen nuevas técnicas basadas en la minería de datos con objeto de descubrir patrones (Rodríguez, 2016), corresponden al conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto. Los datos son la materia prima bruta, de forma que, en el momento que un usuario les atribuye algún significado especial pasan a convertirse en información. Cuando los especialistas elaboran o encuentran un modelo, haciendo que la interpretación que surge entre la información y ese modelo represente un valor agregado, entonces hay que referirse al conocimiento. En esta investigación dentro de la minería de datos se han empleado redes bayesianas, para ello es necesario contar con información o antecedentes del suceso. Para su construcción, la metodología seguida se divide en tres fases, tal y como se muestra en la Figura 1.

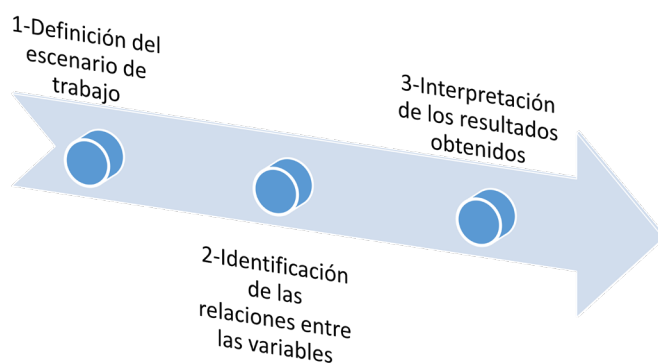


Figura 1. Esquema metodológico.
Fuente: Elaboración propia.

Definición del escenario de trabajo

En la primera fase se toma como base la revisión del estado del arte elaborada con anterioridad para establecer las categorías de las variables que conforman la base de datos y las fuentes de donde se consultaron.

La delimitación geográfica en el estudio es Europa, contando con los datos de 30 países: los 28 de la Unión Europea más Noruega e Islandia que se incluyeron por ser países marítimos que resultaban relevantes en la investigación.

Todo proceso de minería de datos parte de la generación de la base de datos, a partir de un panel de expertos se determinaron las variables de estudio, y se consultaron diferentes bases de datos, para poder generar la base de datos final del estudio. El criterio de selección de las variables se realiza de acuerdo a la relación con el Short Sea Shipping y gas natural licuado. Finalmente se han empleado para la investigación 35, todas ellas consensuadas en el panel de expertos asociado a la investigación, para obtener los valores de las variables, los cuales se extrajeron de Eurostat (28 variables), del documento The European LNG terminal infrastructure 2015: Status and Outlook de la Asociación de infraestructura del gas en Europa (GIE) (3 variables) y los 4 restantes fueron de elaboración propia a partir de otras bases de datos para completar la base de datos (Cuadro 1) Respecto a la tarea de agrupación de las variables, dichas variables se clasificaron en 5 tipos o grupos de acuerdo a su naturaleza, siendo estas categorías las siguientes:

- » Transporte y comercio internacional
- » Economía y finanzas
- » Medioambiente y energía
- » Institucional y política
- » Población y condición social

Cuadro 1. Variables seleccionadas para confeccionar la base de datos. Fuente: Elaboración propia

Categoría	Variable	Descripción	Unidades	Fuente
Transporte y comercio internacional	MTG	Transporte Marítimo	Miles de toneladas	Eurostat
	MSFT RAIL	Distribución modal de transporte de carga por ferrocarril	Porcentaje	Eurostat
	MSFT ROAD	Distribución modal de transporte de carga por carretera	Porcentaje	Eurostat
	MSFT INW	Distribución modal de transporte de carga por aguas continentales	Porcentaje	Eurostat
	SSS VTEU	TEUs (término en inglés Twenty-foot Equivalent Unit, que significa Unidad Equivalente a Veinte Pies) en Short Sea Shipping	1000 TEU's	Eurostat
Economía y finanzas	GDP	Producto interno bruto	Millones de Euros	Eurostat
	EGS	Exportaciones de estados miembros a terceros países	Millones de Euros	Eurostat
	IGS	Importaciones de estados miembros a terceros países	Millones de Euros	Eurostat
	GOVDEF	Deuda, Excedente público	% GDP	Eurostat
	PROIND	Producción en la industria	Index, 2010=100%	Eurostat
	PROCON	Producción en la construcción	Index, 2010=100%	Eurostat
	SHAIMPT	Participación en las importaciones por miembro de la UE	% (EU=100%)	Eurostat
	SHAEXPO	Participación en las Exportaciones por miembro de la UE	% (EU=100%)	Eurostat

Categoría	Variable	Descripción	Unidades	Fuente
Medioambiente y energía	AETSO	Emisiones de óxidos de azufre (SOx) procedentes del transporte y comunicación	Toneladas	Eurostat
	GHGTR	Emisiones de gases de efecto invernadero procedentes del transporte	Miles de toneladas	Eurostat
	ETXTRANS	Impuestos ambientales sobre transporte	Millones de Euros	Eurostat
	GENWAS	Generación total de residuos	Toneladas	Eurostat
	GPICONS	Precios del gas natural de uso industrial	Euro / Kilowatt-hora	Eurostat
	WELEC	Residuos de aparatos eléctricos y electrónicos producidos por hogares	Kilogramos per cápita	Eurostat
	REPRO	Productividad de los recursos	Euro / Kilogramo	Eurostat
	GHGAGR	Emisiones de gases de efecto invernadero procedentes de la agricultura	Millones de toneladas	Eurostat
Institucional y política	SECA	Áreas de control de emisión de azufre	Número	Elaboración propia
	SFRONT	Ubicación geográficas de las costas	Número	Elaboración propia
	RCO	Capacidad de las terminales de regasificación de gas natural licuado en operación	billones m ³ por año	Asociación de
	RCU	Capacidad de las terminales de regasificación de gas natural licuado en construcción	billones m ³ por año	Asociación de
	RCP	Capacidad de las terminales de regasificación de gas natural licuado en planes	billones m ³ por año	Asociación de
	FDG	Forma de gobierno	Número	Elaboración propia
	EPG	Modelos de gestión portuaria en Europa	Número	Elaboración propia
Población y condición social	RDEXPN	Gastos totales en investigación y desarrollo	Euro / habitante	Eurostat
	POBL	Población	Habitantes	Eurostat
	LIFEEXP	Esperanza de vida	Edad	Eurostat
	RISKPOV	Población en riesgo de pobreza o exclusión social	Porcentaje población	Eurostat
	UNER	Tasa de desempleo	Porcentaje población activa	Eurostat
	STENTER	Estudiantes matriculados en educación terciaria	Población	Eurostat
	EXSPBE	Gastos en protección social	Euro / habitante	Eurostat

Identificación de las relaciones las variables

Los modelos gráficos directos, también llamados Redes Bayesianas o Redes de Creencias (BNs), cuentan con una noción compleja de independencia, la cual tiene en cuenta la direccionalidad de los arcos (Almazán-Gárate, Palomino-Monzón, González-Cancelas y Soler-Flores, 2014). Técnicamente, se trata de un grafo dirigido acíclico en donde cada uno de sus nodos hace referencia a una variable aleatoria y las relaciones entre las variables se encuentran codificadas en la propia estructura del grafo de acuerdo al

criterio de separación. A cada nodo de la red se le asocia una distribución de probabilidad condicionada a los padres de ese nodo, de tal forma que la distribución conjunta multiplica las distribuciones condicionadas asociadas a los nodos de la red; la ecuación (1) muestra una red con n variables X_1, X_2, \dots, X_n .

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

Rodríguez y Dolado (Rodríguez y Dolado, 2007) aseguran que una red bayesiana se construye a partir de la definición de la estructura o grafo y de los parámetros de la red, siendo necesarios seis pasos:

- » Obtención de información adicional de los datos disponibles por medio de un experimento en la materia o un proceso adicional (Moral, 2014).
- » Selección del conjunto de variables de manera cuidadosa, para no agregar complejidad innecesaria al estudio.
- » Tipificación de cada una de las variables en booleano, etiqueta o numérico, y definir su discretización.
- » Establecimiento de las relaciones entre las variables (preferibles relaciones causales).
- » Construcción de tablas de probabilidad para cada nodo.
- » Análisis de sensibilidad, donde se observa la forma en la que algunas variables afectan los resultados del resto.

Cuando el sistema contiene un número pequeño de variables y estados, la construcción de la red bayesiana puede realizarse de forma manual. No obstante, en la actualidad se ha hecho necesario automatizar el proceso de construcción de la red porque la mayoría de las bases de datos contienen un elevado número de variables que dificultan la construcción manual (Rodríguez y Dolado, 2007). Dicha automatización puede realizarse a través de los pasos de la técnica de minería de datos.

El uso de redes bayesianas para la representación gráfica de las relaciones entre variables trae consigo una gran cantidad de ventajas, que facilitan la comprensión de las relaciones causales a través del grafo (Puga, 2012). Algunas de estas son: permite representar de forma simultánea la dimensión cuantitativa (tablas de probabilidad) y cualitativa (grafo) de un problema, pueden ser trabajadas con datos incompletos (Puga, 2012) y (Heckerman, 2008), es posible realizar una inferencia bidireccional (Rodríguez y Dolado, 2007; Pearl, 2011 y Puga, 2012), es posible realizar un análisis de sensibilidad (Rodríguez y Dolado, 2007) y (Laskey, 2014) y pueden realizar predicciones con información incompleta, entre otras.

No obstante, éstas presentan algunas dificultades a la hora de su construcción, dentro de las que se encuentran el costo computacional de explorar una red previamente desconocida, y que la calidad y el alcance de los parámetros iniciales pueden distorsionar los resultados (Niedermayer, 2008).

En el estudio de ha empleado el algoritmo K2 para la construcción de la red bayesiana. Dicho algoritmo está basado en la optimización de una medida. Así, dicha medida es usada para explorar, mediante el algoritmo, el espacio de búsqueda formado por todas las redes que contienen las variables de la base de datos, de forma que, partiendo de una red inicial, dicha red se va modificando (añadiendo arcos, eliminándolos o cambiándolos de dirección) hasta obtener una nueva red con mejor medida (Cooper y Herskovits, 1992).

Interpretación de los resultados

En esta última fase se analiza la red, es decir, se analiza la forma en la que las distintas variables se encuentran relacionadas, identificando aquellas que no guardan ninguna relación con las demás.

Las redes bayesianas contemplan tres tipos de relaciones (Figura 2).

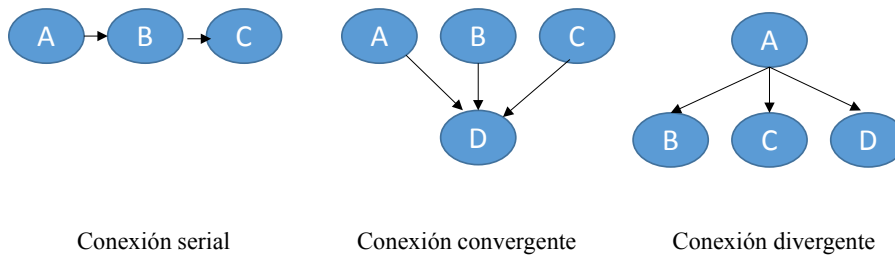


Figura 2. Tipos de conexiones en las redes Bayesianas. Fuente: Molina, González-Cancelas, Soler-Flores, Camarero (2017).

Generalmente, las redes bayesianas obtenidas tras un estudio suelen tener un aspecto bastante enmarañado, de manera que, a simple vista no se perciben las relaciones entre las variables, por lo que se hace necesario desenmarañar la red, de manera que un experto en la temática de estudio de la red vaya desplazando y agrupando los nodos de manera que estudie las relaciones del tipo, convergente, divergente y serial que se producen entre grupos de 3 variables.

Por otro lado, este tipo de redes permite 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.

Si se supone que es cierto un hecho del mundo real que está representado en la red como un nodo padre, la red puede deducir cuáles serán sus efectos. Para ello se debe introducir esta hipótesis en el nodo correspondiente y propagar esta información hacia el resto de los nodos. Este modo de razonamiento es de tipo predictivo y está regido por una inferencia “deductiva” donde el conocimiento se puede expresar de la forma “si A entonces B” y se cumple que el hecho conocido es “A” y el hecho deducido es “B”.

Por ejemplo las conexiones en serie (Figura 2) representan a un conjunto de variables asociadas linealmente que denota dependencia entre las variables (Puga et al., 2006).

Si B depende de A y C de B, en términos causales diríamos que A es causa de B y que B es causa de C. En este caso, dada la dependencia entre las variables, cuando se conoce información sobre A, podemos modificar nuestra certeza sobre el estado de C; y a la inversa, cuando sabemos algo sobre el estado de C la creencia sobre el estado de A se altera (Figura 3).

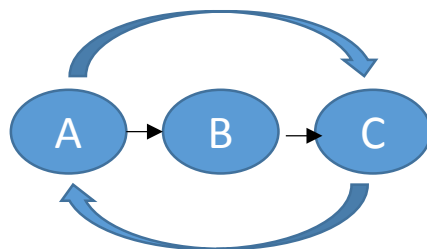


Figura 3. Flujo de información en una relación en serie. Fuente: propia.

Si el estado de B es conocido; conocer alguna información sobre A o C no modificará nuestra creencia sobre el estado de estas variables. Podríamos decir que la propagación de la información se bloquea y se dice que A y C se tornan condicionalmente independientes dado B (Figura 4).

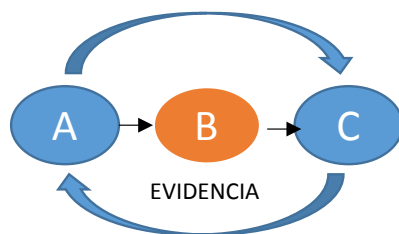


Figura 4. Evidencia en una conexión en serie. Fuente: propia.

Resultados obtenidos

De acuerdo con la metodología anteriormente propuesta, la red bayesiana obtenida es la de la Figura 5.

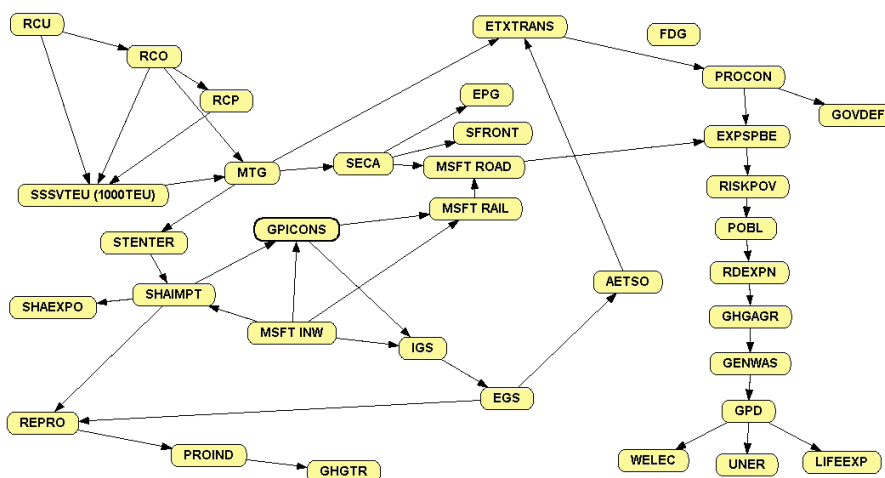


Figura 5. Red bayesiana construida para las variables de estudio con algoritmo K2. Fuente: Elaboración propia.

Las redes bayesianas (o redes de creencia) constituyen una manera práctica y compacta de representar el conocimiento incierto. Se trata a partir de las relaciones que se producen en toda red bayesiana: serial, de manera que se proporciona una información compacta y modular de representar la distribución conjunta de varias variables aleatorias. De manera que se dispone de una parte cualitativa que describe las relaciones entre las distintas variables y una parte cuantitativa que describe la fuerza de dichas relaciones mediante probabilidades condicionadas. En muchos problemas dado un conjunto de datos (evidencia), B tiene que seleccionar la hipótesis A más probable mediante $P(A|B)$. El modelo más extendido de representación de independencias lo constituye las redes bayesianas. En este modelo se representa de forma explícita la dependencia entre variables mediante un grafo. Los nodos del grafo se corresponden con variables y las dependencias se representan mediante arcos entre ellas. Intuitivamente, en una red bayesiana un arco entre X e Y significa una influencia directa de X sobre Y, y de esta manera se explica la red a continuación.

Se observa que la variable FDG (forma de gobierno) no posee relación alguna con las variables que conforman la red y por lo tanto no es ni causa ni efecto de ninguna. Por otro lado, RCU (capacidad de las terminales de regasificación de gas natural licuado en construcción) y MSFT INW (distribución modal de transporte de carga por aguas continentales) son los dos nodos raíz de la red, al ser las únicas que son padres y no hijos de otras variables. La primera se incluye dentro tipo comercio internacional y transporte y la segunda en tipo institucional y político.

Por tanto, el planificador debe prestar especial atención a estas variables de comercio internacional y de transporte o de tipo institucional o político, puesto que el arco existente entre X e Y significa una influencia directa de X sobre Y, por lo que ambos padres de la red poseen una influencia directa sobre el resto de variables (sus descendientes).

Dado que se pretende profundizar en el estudio en las variables de condición social y poblacional en la Figura 6 se han destacado dichas variables. A partir de la figura y, dado que las redes bayesianas permiten aprender sobre relaciones de dependencia y causalidad, permitiendo ver las relaciones entre estas variables, se observa que las variables de condición social y poblacional están muy relacionadas entre sí.

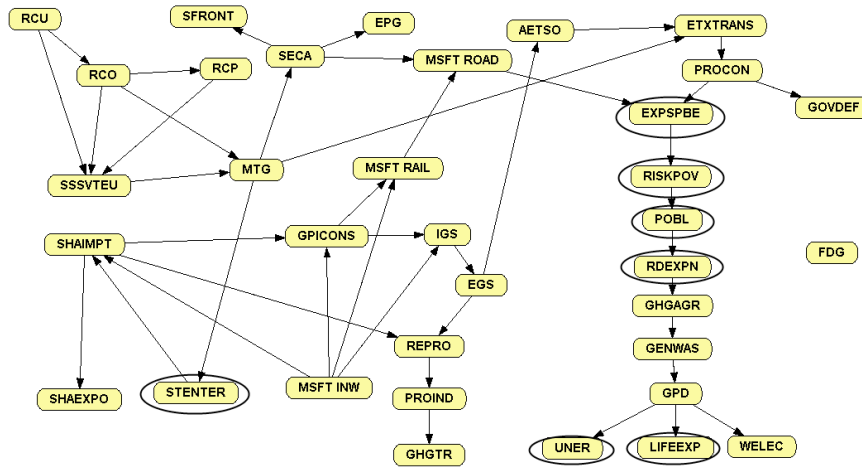


Figura 6. Red construida con algoritmo K2: variables de condición social y poblacional destacadas. Fuente: Elaboración propia.

En redes bayesianas se dice que X es un padre de Y si y sólo si existe un arco $X \rightarrow Y$. Se dice también que Y es hijo de X. Al conjunto de los padres de X se representa como $pa(X)$, y al de los hijos de X por $S(X)$. Si se realiza un estudio de padres e hijos, en el Cuadro 2 se puede observar para las variables de condición social y poblacional los padres e hijos de cada una de ellas. Así, se aprecia que las variables de condición social y poblacional se relacionan mucho entre sí, además son hijas de variables de transporte y comercio internacional como son el transporte marítimo y la distribución modal de transporte de carga por carretera y de variables de economía y finanzas como participación en las importaciones por miembro de la Unión Europea (UE), producto interno bruto y producción en la construcción.

Cuadro 2. Variables de condición social y poblacional: estudio de padres e hijos. Fuente: Elaboración propia

Variables sociales y poblacionales	Descripción	Variables Padres		Variables Hijos
RDEXPN	Gastos totales en investigación y desarrollo	POBL		GHGAGR
POBL	Población	RISKPOV		RDEXPN
LIFEEXP	Esperanza de vida	GDP		
RISKPOV	Población en riesgo de pobreza o exclusión social	EXPSPBE		POBL
UNER	Tasa de desempleo	GDP		
STENTER	Estudiantes matriculados en educación terciaria	MTG		SHAIMPT
EXPSPBE	Gastos en protección social	MSFT ROAD	PROCON	RISKPOV

Leyenda color de las categorías:

Población y condición social
Economía y finanzas
Medioambiente y energía
Transporte y comercio internacional

Si se analizan las relaciones de la variable STENTER (Estudiantes matriculados en educación terciaria), variable que permanece en un grupo aislado, respecto al resto de variables de condición social y poblacional (Figura 7) se aprecia que su padre es la variable transporte marítimo y su hijo es participación en las importaciones por miembro de la UE. Si se analizan las relaciones seriales A→B→C, las conexiones en serie representan a un conjunto de variables asociadas linealmente que denota dependencia entre las variables, en este caso, B depende de A y C de B. En términos causales diríamos que A es causa de B y que B es causa de C, así, en la red construida MTG depende de SSVTEU y de RCO, así como STENTER, ETXTRANS y SECA depende de MTG así RCO. Por lo que el número de estudiantes matriculados en educación terciaria dependen del número de TEUs (término en inglés Twenty-foot Equivalent Unit, que significa Unidad Equivalente a Veinte Pies) en Short Sea Shipping y de la capacidad de las terminales de regasificación de gas natural licuado en operación a través del volumen de Transporte Marítimo, es decir, que los países más evolucionados en términos de tráfico marítimo con capacidad de transporte de Short Sea Shipping elevada y con terminales de regasificación de gas natural licuado poseen mayor número de alumnos matriculados en ciclos superiores, así como mayores Áreas de control de emisión de azufre y mayores Impuestos ambientales sobre transporte.

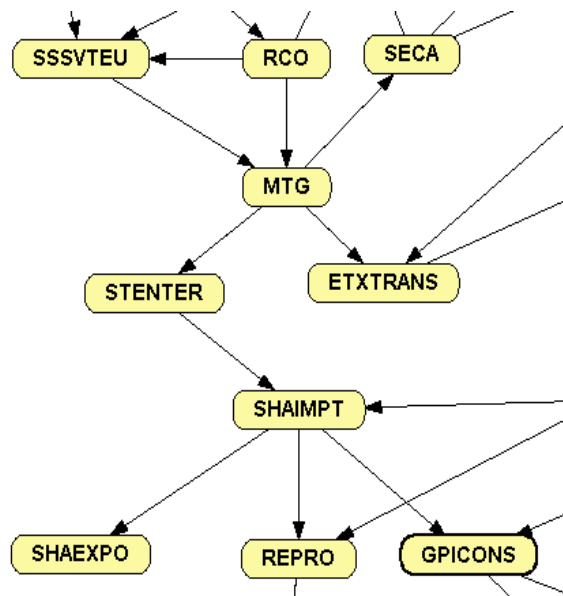


Figura 7. Relaciones entorno a la variable STENTER. Fuente: Elaboración propia.

De igual manera que en la descripción anterior, de la relación en serie MTG→STENTER→SHAIMPT se puede decir que STENTER es causa de SHAIMPT, así como MTG es causa de SHAIMPT a través de STENTER, es decir que el volumen de transporte marítimo de cada país es causa de la participación en las importaciones por miembro de la UE a través del número de estudiantes matriculados en educación terciaria.

Las relaciones en serie STENTER→SHAIMPT→SHAEXPO, STENTER→SHAIMPT→REPRO y STENTER→SHAIMPT→GPICONS indican que número de estudiantes matriculados en educación terciaria es causa de la participación en las importaciones por miembro de la UE y que la participación en las importaciones por miembro de la UE es causa de la Participación en las Exportaciones por miembro de la UE, de la productividad de los recursos y de los precios del gas natural de uso industrial. Llama la atención que una variable como el número de estudiantes matriculados en educación terciaria afecte a variables de medioambiente y energía y a las de economía y finanzas.

Si se realiza un análisis convergente, las conexiones convergentes varias variables apuntan con sus arcos hacia una variable de convergencia esto es, convergen en un mismo nodo hijo, se podría describir como $X \rightarrow Y \leftarrow Z$, X y Z convergen en Y. La propiedad importante de este tipo de conexión a la hora de propagar la información es que cuando se tiene evidencia sobre la variable de convergencia, los nodos padre se vuelven dependientes y la evidencia del estado de uno de ellos se propaga por los demás. Se diría, en nuestro ejemplo, SSSVTEU y RCO son dependientes dado MTG.

Quedaría un último análisis para estas relaciones resultantes, las conexiones divergentes, que puede notarse como $M \leftarrow N \rightarrow O$. En las conexiones divergentes tenemos un nodo padre que proyecta sus arcos hacia varios hijos; o lo que es lo mismo, las flechas que salen de él y divergen hacia sus hijos. Este tipo de conexión es el más apropiado para representar problemas de diagnóstico. Sin embargo, cuando el estado de la variable madre se conoce las variables hijo se tornan independientes y la información no se propaga si añadimos evidencias sobre los nodos hijo. En la red resultante SHAEXPO, RERO Y GPICONS son independientes dado SHAIMPT.

El resto de variables de condición social y poblacional tienen una relación más directa entre ellas como se aprecia en la Figura 8. La mayoría de las conexiones que se producen son en serie, como se ha visto antes se producen seriales $A \rightarrow B \rightarrow C$, de tal manera que desde la variable ETXTRANS a las variables UNER, LIFEEXP y WELEC se van sucediendo relaciones seriales de tal manera que B depende de A y C de B; en términos causales diríamos que A es causa de B y que B es causa de C, por lo que las relaciones entre las variables de condición social y poblacional: EXPSPBE, RISKPOV, POBL, RDEXPN, UNER, LIFEEXP se producen relaciones de causa efecto, de tal manera que los gastos en protección social son causa de Población en riesgo de pobreza o exclusión social en primera instancia, y de forma derivada de los gastos totales en investigación y desarrollo, la población y en mucha menor medida de la tasa de desempleo y la esperanza de vida. Otras relaciones seriales a destacar son $SECA \rightarrow MSFT ROAD \rightarrow EXPSPBE$ o $MSFT RAIL \rightarrow MSFT ROAD \rightarrow EXPSPBE$. De tal forma que los gastos en protección social son consecuencia directa de la distribución modal de transporte de carga por carretera y de forma indirecta a través de la distribución modal de transporte de carga por carretera, de las variables de distribución modal de transporte de carga por ferrocarril y de las áreas de control de emisión de azufre, por lo que los gastos en protección social son mayores en países con mayor reparto modal en la carretera y en aquellos donde las políticas medioambientales son más restrictivas a través del control de emisiones.

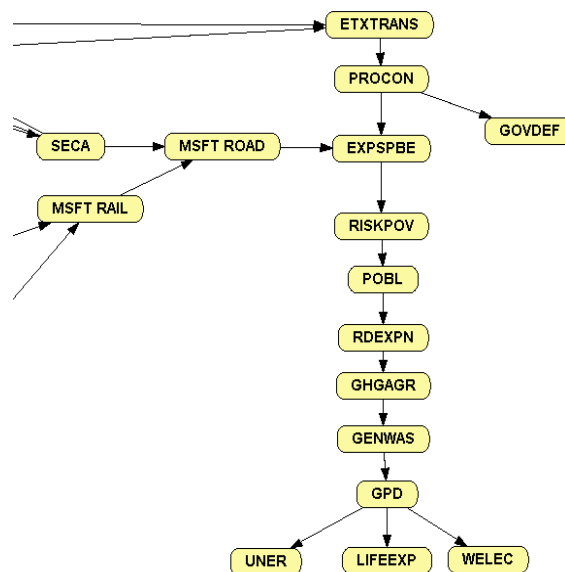


Figura 8. Relación entre el resto de variables de condición social y poblacional. Fuente: Elaboración propia.

Otra relación serial establece que EXTRANS→ PROCON→ EXPSBE de manera que los gastos en protección social son causa, entre otras variables de la producción en la construcción y de manera indirecta a través de Producción en la construcción de la variable impuestos ambientales sobre transporte.

Otras relaciones en serie a destacar son las que están entorno a GHGAGR, GENWAS y GDP que no son variables de condición social y poblacional pero que se encuentra directamente relacionadas con este tipo de variables. Así, las emisiones de gases de efecto invernadero procedentes de la agricultura son consecuencia directa de los gastos totales en investigación y desarrollo y a través de gastos totales en investigación y desarrollo son consecuencia de la población del país. De igual manera, la generación total de residuos son consecuencia de emisiones de gases de efecto invernadero procedentes de la agricultura y de forma indirecta a través de emisiones de gases de efecto invernadero procedentes de la agricultura de la variable de los gastos totales en investigación y desarrollo. Por lo tanto, variables de medioambiente y energía relacionadas con los residuos y las emisiones de gases de efecto invernadero son muy sensibles a variables sociales y de población relacionadas con gastos en investigación y desarrollo.

Al realizar un estudio aparte sobre la variable GPD, si se estudia la relación divergente que genera, el nodo padre GPD proyecta sus arcos hacia varios hijos, UNER, LIFEEX, WELEC; o lo que es lo mismo, las flechas salen de él y divergen hacia sus hijos. Este tipo de conexión es el más apropiado para representar problemas de diagnóstico. Sin embargo, cuando el estado de la variable padre se conoce, las variables hijo se tornan independientes y la información no se propaga si añadimos evidencias sobre los nodos hijo. En la red resultante UNER, LIFEEX, WELEC son independientes dado GPD, es decir conocido el producto interno bruto de un país, las variables tasa de desempleo, esperanza de vida y residuos de aparatos eléctricos y electrónicos producidos por hogares son independientes entre sí, pero no hay que olvidar que dichas variables (tasa de desempleo, esperanza de vida y residuos de aparatos eléctricos y electrónicos producidos por hogares) son efecto directo del el producto interno bruto. Así, a mayor producto interno bruto, menor tasa de desempleo, mayor esperanza de vida y mayor número de residuos de aparatos eléctricos y electrónicos producidos por hogares.

Así, también hay que destacar que los residuos de aparatos eléctricos y electrónicos producidos por hogares son consecuencia indirecta de la generación total de residuos a través del producto interno bruto de un país.

Conclusiones

La inteligencia artificial, y en especial las redes bayesianas, suponen una alternativa óptima para determinar las relaciones entre las variables que afectan al gas natural licuado como combustible para empleo en el Short Sea Shipping en Europa.

El gas natural licuado se muestra como una alternativa en la necesidad de buscar combustibles más amigables con el medioambiente en el Short Sea Shipping y es por ello que la herramienta desarrollada permite a los gestores tomar decisiones para minimizar la congestión del transporte en Europa. Dicha herramienta ha hecho posible establecer que las variables RCU (capacidad de las terminales de regasificación de gas natural licuado en construcción) y MSFT INW (distribución modal de transporte de carga por aguas continentales) son las variables más importantes en la decisión de implementar el gas natural licuado como combustible en el tramo marítimo de las autopistas del mar.

Asimismo, se establece que las variables de transporte y comercio internacional y economía y finanzas son las más relevantes en la toma de decisiones, mientras que las variables de las categorías de medioambiente y energía y población y condición social son las más dependientes y, por tanto, se ven afectadas en mayor proporción que el resto.

Respecto al estudio de las variables de condición social y poblacional se aprecia que dichas variables se relacionan mucho entre sí. Además, son hijas de variables de transporte y comercio internacional como son el transporte marítimo y la distribución modal de transporte de carga por carretera y de variables de economía y finanzas como participación en las importaciones por miembro de la UE, producto interno bruto y producción en la construcción. Destacar que el número de estudiantes matriculados en educación terciaria afecte a variables de medioambiente y energía y de economía y finanzas. Para el resto de variables de condición social y poblacional tienen una relación más directa entre ellas, la mayoría de las conexiones que se producen son en serie, es decir, de causa-efecto.

Variables de medioambiente y energía relacionadas con los residuos y las emisiones de gases de efecto invernadero son muy sensibles a variables sociales y de población relacionadas con gastos en investigación y desarrollo de los países de la UE.

Así, según la red construida, cuanto mayor es el producto interno bruto, menor es la tasa de desempleo, pero mayor es la esperanza de vida y el número de residuos de aparatos eléctricos y electrónicos producidos por los hogares.

Por último cabría destacar que las variables de población y condición social no afectan a las variables directamente implicadas en el gas natural licuado como combustible susceptible para el Short Sea Shipping como pueden ser el transporte marítimo, el número de TEUs en Short Sea Shipping, las terminales de regasificación de gas natural licuado, Áreas de control de emisión de azufre y, Ubicación geográficas de las costas, con la excepción de la variable que corresponde al número de estudiantes matriculados en educación terciaria que sí se relaciona con algunas de ellas.

Un avance al tema propuesto lo representa la inferencia bayesiana. La inferencia bayesiana es un tipo de inferencia estadística en la que las evidencias u observaciones se emplean para actualizar o inferir la probabilidad de que una hipótesis pueda ser cierta. El teorema de Bayes indica cómo actualizar las creencias sobre una variable a la luz de la evidencia (datos). Es un método general de inducción o de “aprendizaje de la experiencia”: a priori \rightarrow datos \rightarrow a posteriori. De manera que, conocida la información proporcionada por una o más variables que se observan (evidencia), ésta se propaga por la red y actualiza la creencia sobre las variables no observadas. Así, es posible aprender las probabilidades condicionales que describen las relaciones entre las variables a partir de los datos. Asimismo, es posible aprender la estructura completa de la red, no sólo a partir de datos completos, sino también en el caso en el que algunos de sus valores sean desconocidos. Por tanto, las redes bayesianas pueden utilizarse para tomar decisiones óptimas al introducir posibles acciones y la utilidad de sus resultados.

Bibliografía

- » Acedo, A., Almazán, J. L. (2015). Pasado, presente y futuro de las Autopistas del Mar en Europa. *Revista de Obras Públicas*, 3565, 31- 38.
- » 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.
- » Awad Núñez, S., González-Cancelas, N., Camarero Orive, A. (2015). Establecimiento de los factores a considerar para determinar la zona de ubicación de un puerto seco y de sus jerarquías a través de un panel DELPHI. *Revista Transporte y Territorio* 13, 100-121.
- » Bengtsson, S., Andersson, K., Fridell, E. (2011). A comparative life cycle assessment of marine fuels: liquefied natural gas and three other fossil fuels. Proceedings of the Institution of Mechanical Engineers, Part M: *Journal of Engineering for the Maritime Environment*, 225 (2), 97-110.
- » Cánovas, B. (2015). Short Sea Shipping, una estrategia Europea. *Documento de Opinión*, No. 70. Instituto Español de Estudios Estratégicos.
- » Comisión Europea (2001). *Libro Blanco, la política europea de transportes de cara al 2010: la hora de la verdad*. COM (2001) 370 final. Bruselas: Commission of the European Communities, 11- 53.
- » Comisión Europea (2006). *Libro Verde, estrategia europea para una energía sostenible, competitiva y segura*. COM (2006) 105 final. Bruselas: Commission of the European Communities, 3-19.
- » Cooper, G. F., Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning* 9, 309-348.
- » ENAGAS (2017). El proyecto CORE LNGas hive para el impulso del gas natural licuado como combustible en el transporte presenta en Barcelona sus últimos avances. Nota de prensa. Disponible en: <<http://www.enagas.es/stfls/ENAGAS/Notas%20de%20prensa/NP%20CORE%20LNGas%20hive%2022%20junio%202017.pdf>>.
- » González Cancelas, N., (2016). Nuevas cadenas de transporte de mercancías generadas por las infraestructuras logísticas de intercambio modal. *Revista Transporte y Territorio* 14.
- » Heckerman, David (2008). A tutorial on learning with Bayesian networks. En *Innovations in Bayesian networks*. Springer Berlin Heidelberg, 33-82.
- » Laskey, K. B. (2014). Sensitivity analysis for probability assessments in Bayesian networks. En *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence*. 136.
- » Lloyd, M. (2003). Regional action for logistical integration of shipping across Europe (realise). Work package 4. Task 4.1 multi-modal pricing and costing analyses inception report. Disponible en: <http://www.realise-sss.org/>
- » Macías Párraga, M., González-Cancelas, N., Soler-Flores, F. (2016). Estrategia de planificación del puerto de aguas profundas de la ciudad de Manta: inclusión

- de metodología DaDED (dafo-delphi-estadística descriptiva). *Revista Transporte y Territorio* 14.
- » Molina Serrano, B., González-Cancelas, N., Soler-Flores, F., Camarero Orive, A. (2017). Los recursos humanos y la huella ecológica portuaria. *Ingeniería Civil*, 188.
 - » MORAL, S. (2014). Una Introducción a las Redes Bayesianas. IMCOMPLETO
 - » Niedermayer, D. (2008). An introduction to Bayesian networks and their contemporary applications. *Innovations in Bayesian Networks*, FALTAN DATOS, 117-130.
 - » Palomino Monzón, M. C., Almazán Gárate, J. L., González Cancelas, N., Soler Flores, F. (2012). Planificación de la integración puerto ciudad: caso de Vigo. *Revista Transporte y Territorio*, 7, 150-165.
 - » Pearl, J. (2011). *Bayesian networks*. Department of Statistics, UCLA.
 - » Puga, J. L. (2012). Cómo construir y validar Redes bayesianas con netica. *REMA Revista electrónica de metodología aplicada*, 17 (1), 1-17.
 - » Orive, A., De La Fuente, P. M., López, Í. A., Domínguez, E. G.; Cancelas, N. (2012). Regeneración de las playas y rehabilitación de los canales de desagüe de Beira (Mozambique). *Revista de Obras Públicas*, 159 (3532), 7-22.
 - » Orive, A., Orive, A. C., Cancelas, M. N. G., Guijarro, E. D. (2012). Transporte marítimo: de la SECA a la ECA. *Revista de Obras Públicas: Organó profesional de los ingenieros de caminos, canales y puertos*, (3537), 65-76
 - » REALISE (2002). Regional Action for Logistical Integration of Shipping across Europe. *The Alliance of Maritime Regional Interests in Europe (AMRIE)*, GTC2-2000-33032. Disponible en: <http://www.realise-sss.org>
 - » REALISE (2005). The Alliance of Maritime Regional Interests in Europe, (AMRIE), *Final Report No. GTC2-2000-33032*. Disponible en: <http://www.realise-sss.org>
 - » RECORDIT (2003). Actions to Promote Intermodal Transport: *Final Report No. WP9*. Disponible en: <http://www.recordit.org/>
 - » Rodríguez, D., Dolado, J.(2007). Redes bayesianas en la ingeniería del software. *Cc. Uah. Es*, 1-21.
 - » Rodríguez García, T. (2016). Aplicaciones tecnológicas en la logística de transportes portuarios. Las terminales de contenedores. *Revista Transporte y Territorio*, 14.
 - » Romero Faz, D. (2016). Evaluación y organización de la seguridad en terminales portuarias. *Revista Transporte y Territorio*, 14.
 - » Sáenz-Löbsack, D., Camarero Orive, A., González Cancelas, N. (2016). Propuesta de una metodología simple para el análisis y representación gráfica de riesgos sobre activos. Aplicación al caso de una instalación portuaria. *Revista Transporte y Territorio*, 14.
 - » Vassallo, W., et al. (2004). Regional action for logistical integration of shipping across Europe (Realise). Disponible en: <http://www.realise-sss.org/>

Beatriz Molina Serrano / beatriz.molinas@alumnos.upm.es

Ingeniero de Caminos, Canales y Puertos por la Universidad Politécnica de Madrid. Doctorando de la Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos de la Universidad Politécnica de Madrid. Es especialista en transporte y seguridad vial.

Nicoletta González-Cancelas / nicoleta.gcancelas@upm.es

Doctor Ingeniero de Caminos, Canales y Puertos por la Universidad Politécnica de Madrid. Profesor Contratado Doctor del Departamento de Ingeniería Civil de la Universidad Politécnica de Madrid. Es especialista en Logística y transporte, profesora del Máster Universitario en Sistemas de Ingeniería Civil, del Master en Gestión y Planificación Portuaria e Intermodalidad y del Master del Campus Portuario de Puertos del Estado.

Francisco Soler-Flores / francisco.soler@unir.net

Licenciado en Matemáticas por la Universidad de Almería. Doctor en Tecnología y Sistemas de Información por la Universidad de Castilla-La Mancha. Es Profesor en la facultad de Educación de la Universidad Internacional de La Rioja. Es especialista en Analytics y Machine Learning.

Alberto Camarero-Orive / alberto.camarero@upm.es

Doctor Ingeniero de Caminos, Canales y Puertos por la Universidad Politécnica de Madrid y Licenciado en Economía y en Administración, y Dirección de Empresas por Universidad Carlos III de Madrid. Universidad Politécnica de Madrid. Profesor titular del Departamento de Ingeniería Civil y Transportes en la Unidad Docente de Explotación de Puertos. Además, es Profesor del Programa de Doctorado "Ingeniería e Infraestructura de los transportes".