

Modelización y predicción del tráfico marítimo en los puertos españoles

Coto-Millán, Pablo cotop@unican.es

Departamento de Economía

Universidad de Cantabria

Inglada-Pérez, Lucía linglada@eco.uc3m.es)

Departamento de Economía

Universidad Carlos III de Madrid

Casares, Pedro casaresp@unican.es

Departamento de Economía

Universidad de Cantabria

Inglada, Vicente vinglada@cee.uned.es

Departamento de Economía Aplicada y Estadística

UNED

RESUMEN

Tradicionalmente el transporte marítimo ha desempeñado un papel relevante en la economía mundial a través de su estrecha conexión con el comercio mundial ya que una gran parte de los intercambios internacionales de mercancías se realizan utilizando dicho modo de transporte. Esta relevancia se ha visto acrecentada, si cabe, en un mundo profundamente globalizado e interdependiente como el actual, donde la competitividad comercial de un país depende significativamente de poder disponer de un acceso adecuado a los servicios de transporte marítimo internacional y de un sistema portuario eficiente. En el caso español el sistema portuario desempeña un papel estratégico muy relevante en la actividad económica y en el sistema económico y social ya que cerca del setenta por ciento del comercio exterior se lleva a cabo mediante el transporte marítimo y una parte importante del territorio tiene un carácter insular. La modelización y predicción de la demanda de transporte marítimo es particularmente relevante para una eficiente gestión del sistema portuario, así como en la planificación eficiente de las inversiones, especialmente teniendo en cuenta el elevado coste de las inversiones portuarias, así como su carácter de largo plazo. En general su interés se extiende a todos los agentes intervinientes en el transporte marítimo: armadores, fletadores, etc. y a las actividades

económicas desarrolladas en su hinterland. El objetivo general de este trabajo es la modelización y predicción del tráfico marítimo en los puertos españoles. Para ello se utiliza una amplia batería de modelos lineales y no lineales como SARIMA, Suavizado Exponencial y redes neuronales. Para comparar la precisión de las predicciones obtenidas con los modelos ajustados se emplean varios criterios de evaluación como es el error absoluto medio o el error cuadrático medio.

ABSTRACT

Traditionally, maritime transport has played an important role in the world economy through its close connection with world trade and a large part of the international exchanges of products used in this regard. This is the reason why it is considered enhanced, if possible, in a completely globalized and interdependent world like the current one, where the commercial competitiveness of a country depends on a place of adequate access to an international maritime transport service and an efficient port system. In the Spanish case, the port system plays a very important strategic role in economic activity and in the economic and social system that is close to the foreign system, foreign trade is carried out with maritime transport and a significant part of the territory has an insular character. The modeling and prediction of maritime transport demand is also relevant for efficient system management, as well as the efficient future of investments, taking into account the cost of investments, as well as their long-term nature. . In general, their interest extends to all agents involved in maritime transport: shipowners, charterers, etc. and the activities developed inside it. The general objective of this work is the modeling and prediction of maritime traffic in Spanish ports. For this, a wide range of linear and non-linear models is used, such as SARIMA, Exponential Smoothing and neural networks. The accuracy of the predictions with the fitted models is also used to evaluate the mean absolute error or the mean square error.

Palabras claves:

Modelos ARIMA; transporte marítimo; predicción; puertos españoles

Área temática: A3 Métodos cuantitativos en un entorno de incertidumbre

1. INTRODUCCIÓN

Secularmente el transporte marítimo ha desempeñado un papel relevante en la economía mundial a través de su estrecha conexión con el comercio mundial ya que una gran parte de los intercambios internacionales de mercancías, se realizan utilizando dicho modo de transporte. Esta relevancia se ha visto acrecentada, si cabe, en un mundo profundamente globalizado e interdependiente como el actual, donde la competitividad comercial de un país depende significativamente de poder disponer de un acceso adecuado a los servicios de transporte marítimo internacional y de un sistema portuario eficiente (Coto-Millán et al., 2018). En el caso español el sistema portuario desempeña un papel estratégico muy relevante en la actividad económica y en el sistema económico y social ya que cerca del setenta por ciento del comercio exterior se lleva a cabo mediante el transporte marítimo y una parte importante del territorio tiene un carácter insular (Inglada-Pérez, 2015).

Por ello, la modelización y predicción de la demanda de transporte marítimo es particularmente relevante para una eficiente gestión del sistema portuario, así como para la planificación de inversiones, y otras decisiones estratégicas especialmente teniendo en cuenta el elevado coste de las inversiones portuarias así como su carácter de largo plazo. En general su interés se extiende a todos los agentes intervinientes en el transporte marítimo: armadores, fletadores, etc. y a las actividades económicas que se desarrollan al calor de los puertos, en su correspondiente hinterland.

El objetivo principal de este trabajo es la modelización y predicción de la serie de tráfico marítimo en los puertos españoles en el periodo de enero de 1997 a diciembre de 2017. Para ello se utiliza una amplia gama de modelos univariantes lineales y no lineales como son: SARIMA, suavizado exponencial y redes neuronales. Para comparar la adecuación de cada modelo y la precisión de las predicciones obtenidas con los modelos ajustados se emplean varios criterios de evaluación: error porcentual absoluto medio (MAPE), error absoluto medio (MAE) y la raíz cuadrada de los errores cuadráticos medios (RMSE) que son frecuentemente utilizados en la literatura internacional para evaluar los resultados obtenidos en la modelización y predicción de series temporales

Dada su relevancia en la economía mundial y española, no es extraño que la modelización y predicción de las series temporales de tráfico marítimo, hayan constituido el

objetivo de numerosos trabajos de investigación. Sin embargo, en nuestro conocimiento, este trabajo es el primero que estudia el tráfico de forma desagregada según su presentación - graneles líquidos, graneles sólidos, contenedores y mercancía convencional- y utilizando cuatro tipos de modelos, así como tres criterios de evaluación de la precisión de las predicciones del tráfico.

El trabajo se organiza de la manera siguiente: en el apartado segundo se lleva a cabo una revisión de la literatura internacional sobre modelización y predicción de la demanda de transporte marítimo relacionada con esta investigación. En la sección tercera se describen las fuentes de datos utilizadas y se analizan con detalle los datos empleados. A continuación, se describe el proceso metodológico adoptado en esta investigación tanto en relación con los tipos de modelos empleados como con los criterios para la evaluación de las predicciones. En el capítulo quinto se exponen y comentan los principales resultados obtenidos para cada tipo de transporte marítimo. Finalmente, en el último capítulo se extraen una serie de conclusiones sobre la investigación realizada y se sugieren las posibles líneas de investigación futuras.

2. REVISIÓN DE LA LITERATURA

A partir del análisis de la literatura internacional¹, se deduce que la mayoría de los trabajos sobre modelización y predicción de la demanda de transporte, y de transporte marítimo, en particular, se realizan utilizando modelos lineales como son el análisis de regresión o los modelos ARIMA basados en la metodología Box-Jenkins (Box y Jenkins, 1976 y Box et al., 1994). Sin embargo, durante las últimas décadas ha crecido el interés por la aplicación de modelos no lineales en el ámbito de la predicción del transporte marítimo. En particular, han proliferado las investigaciones que utilizan modelos de redes neuronales para la predicción de la evolución futura de las series temporales analizadas. A continuación, se describen las investigaciones que se consideran más relevantes en el contexto de esta investigación.

Peng y Chu (2009) comparan seis modelos univariantes para predecir el tráfico marítimo en tres puertos de Taiwan. Obtienen que el método clásico de descomposición y el modelo SARIMA ofrecen las mejores predicciones basadas en los criterios de precisión

¹ En Washington et al, (2003) se realiza una revisión completa de los modelos econométricos utilizados para el análisis de las series de datos de transporte. Asimismo, Glen (2006) lleva a cabo una revisión de los modelos para el tráfico marítimo de graneles sólidos y líquidos. En Coto-Millán et al. (2010) se estudian los modelos de regresión de la demanda de transporte marítimo de diferentes tipos de mercancías.

adoptados. Entre las investigaciones que utilizan la metodología de redes neuronales, en Lam et al. (2004) se llevan a cabo predicciones sobre la evolución futura de varias series de tráfico de mercancías en el puerto de Hong Kong, utilizando el modelo de redes neuronales, comparando los resultados obtenidos con los del modelo de regresión lineal. Concluyen que los modelos de redes neuronales se muestran significativamente superiores a los de regresión lineal, en relación con su capacidad de predicción. Mostafa (2004) utiliza también el modelo de Redes Neuronales para predecir el tráfico mensual en el canal de Suez. Al comparar los resultados obtenidos con los correspondientes a otros métodos, como los modelos ARIMA, concluye que los modelos de Redes Neuronales superen a dichos modelos en su capacidad de predicción del tráfico futuro. Finalmente, Chen y Chen (2010) comparan para la serie de tráfico marítimo de contenedores, la capacidad de predicción del modelo de programación genética con otros modelos como es el SARIMA. Utilizan como medida de la precisión en las predicciones, al valor absoluto del error porcentual medio, y obtienen que el método de programación genética se muestra significativamente como el mejor modelo de predicción de las series analizadas. El contenido de los epígrafes se escribe aquí.

3. DATOS

Las cinco series de tráfico marítimo en los puertos españoles utilizadas en esta investigación, para el periodo que transcurre entre enero de 1997 y diciembre de 2016 (240 observaciones), corresponden, respectivamente, al volumen de mercancías embarcadas y desembarcadas en los puertos españoles en forma de graneles líquidos, graneles sólidos, contenedores y mercancía convencional, así como al total de mercancías.

Tabla 1. Estadísticos descriptivos de las variables de tráfico marítimo. (Miles de toneladas)

	Total	Graneles líquidos	Graneles sólidos	Mercancía en contenedores	Mercancía convencional
Media	33701,1	11907,3	7745,3	9729,4	4319,1
Mediana	34723,7	11942,4	7620,9	10499,3	4351,2
Máximo	44624,4	15329,6	10940,4	17004,8	6248,9

Mínimo	20356,6	8813,3	5238,2	3412,5	2318,7
Coficiente de variación	0,171	0,122	0,157	0,344	0,199
Observaciones	240	240	240	240	240

Fuente: elaboración propia.

Todos los datos han sido tomados del Ministerio de Fomento (<http://www.fomento.es/>). En la Tabla 1 se muestran los principales estadísticos descriptivos de las cinco series investigadas. Se observa en dicha tabla que el mayor coeficiente de variación corresponde al tráfico marítimo de contenedores mientras que el tráfico marítimo de graneles líquidos es el que presenta un menor coeficiente de variación.

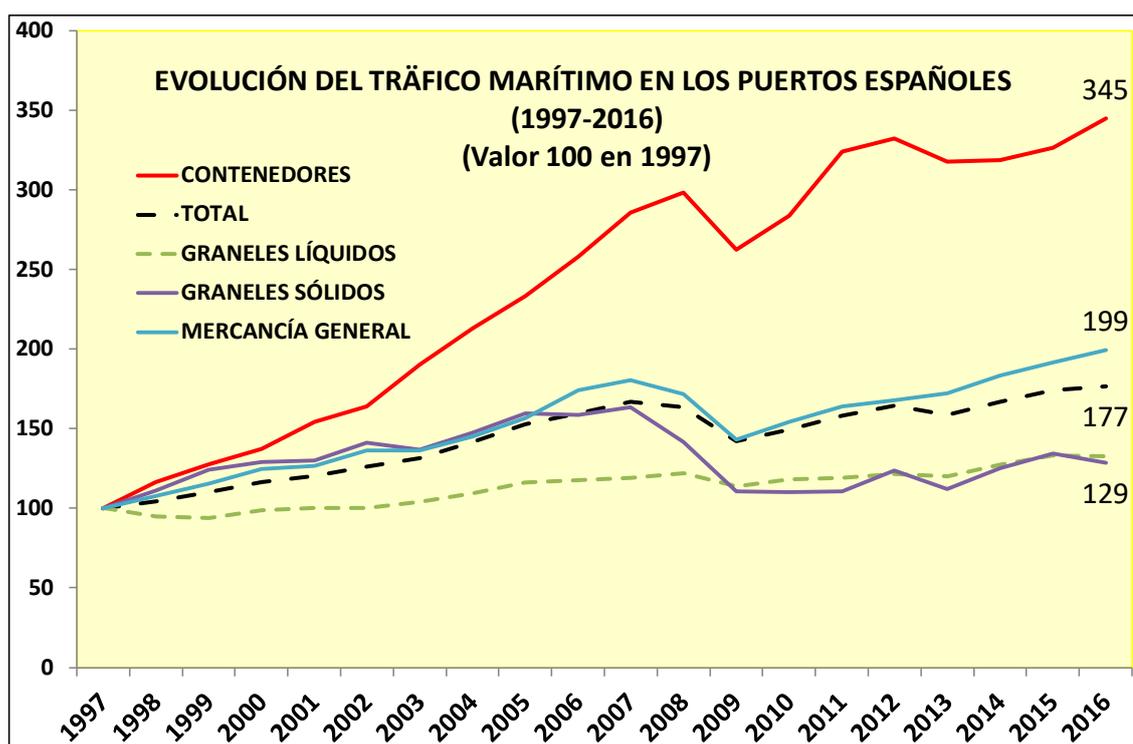


Gráfico 1. Evolución del tráfico en los puertos españoles (2007-2016).

(Valor 100 en 1997)

Fuente: Elaboración propia

En el Gráfico 1 se muestra la evolución del tráfico marítimo de mercancías durante el periodo 1997-2016 en los puertos españoles (total, graneles líquidos, graneles sólidos, contenedores y mercancía convencional). Se observa que el mayor crecimiento en este periodo corresponde al tráfico de contenedores con un crecimiento del 245% mientras que el menor corresponde al tráfico de graneles sólidos con un crecimiento del 29%.

Asimismo, en la tabla 2 se muestra el efecto de la crisis económica sobre cada tipo de tráfico marítimo. Para medirlo se considera la tasa de variación en 2009 respecto a 2008. Se observa que el impacto sobre el tráfico total ha sido del 13%. Desglosando por tipo de tráfico, el mayor impacto se produce en el tráfico de graneles sólidos (21,9%), seguido por el de mercancía convencional (16,6%). Por el contrario, el menor impacto de la crisis económica corresponde al tráfico de graneles líquidos (6,8%).

Tabla 2. Efecto de la crisis económica sobre cada tipo de tráfico marítimo. (Tasa de variación en 2009 respecto 2008).

Tipo de mercancía				
Total	Granel es Líquidos	Granel es Sólidos	Conte nedores	Conve ncional
-	-	-	-	-
13,03%	6,80%	21,92%	12,02%	16,60%

Fuente: elaboración propia.

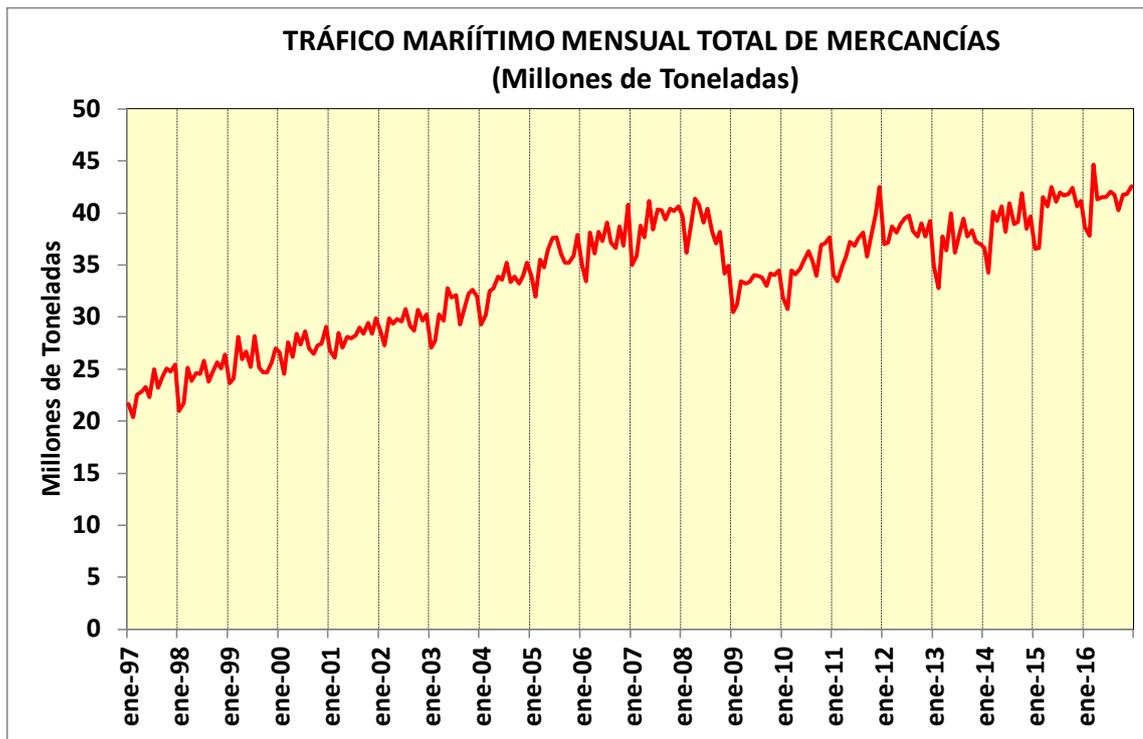


Figura 2. Tráfico marítimo total de mercancías

Fuente: Elaboración propia.

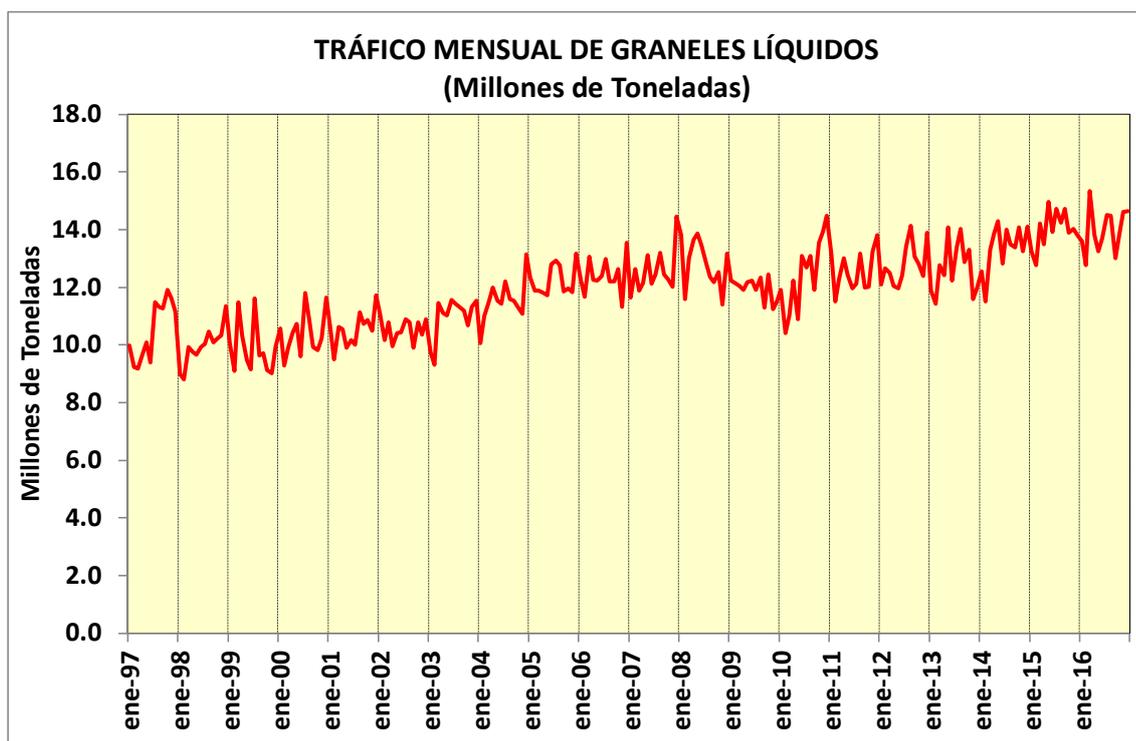


Figura 3. Tráfico marítimo de graneles líquidos

Fuente: Elaboración propia.

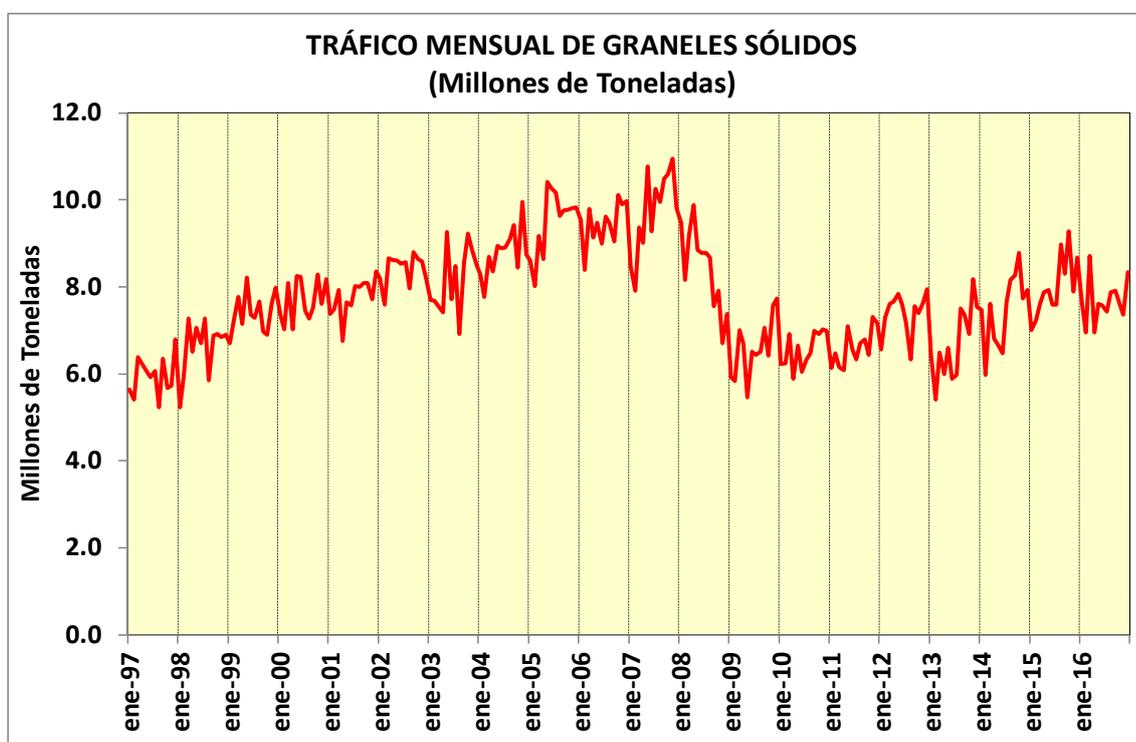


Figura 4. Tráfico marítimo de graneles sólidos

Fuente: Elaboración propia.

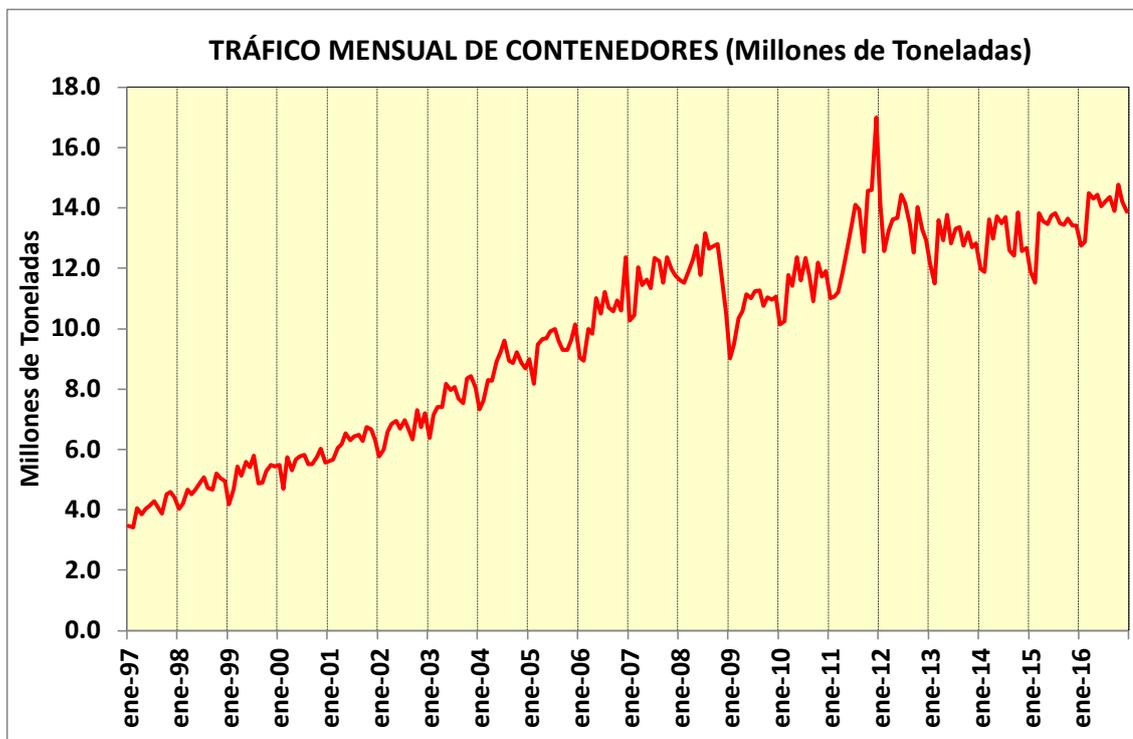


Figura 5. Tráfico marítimo de contenedores.

Fuente: Elaboración propia.

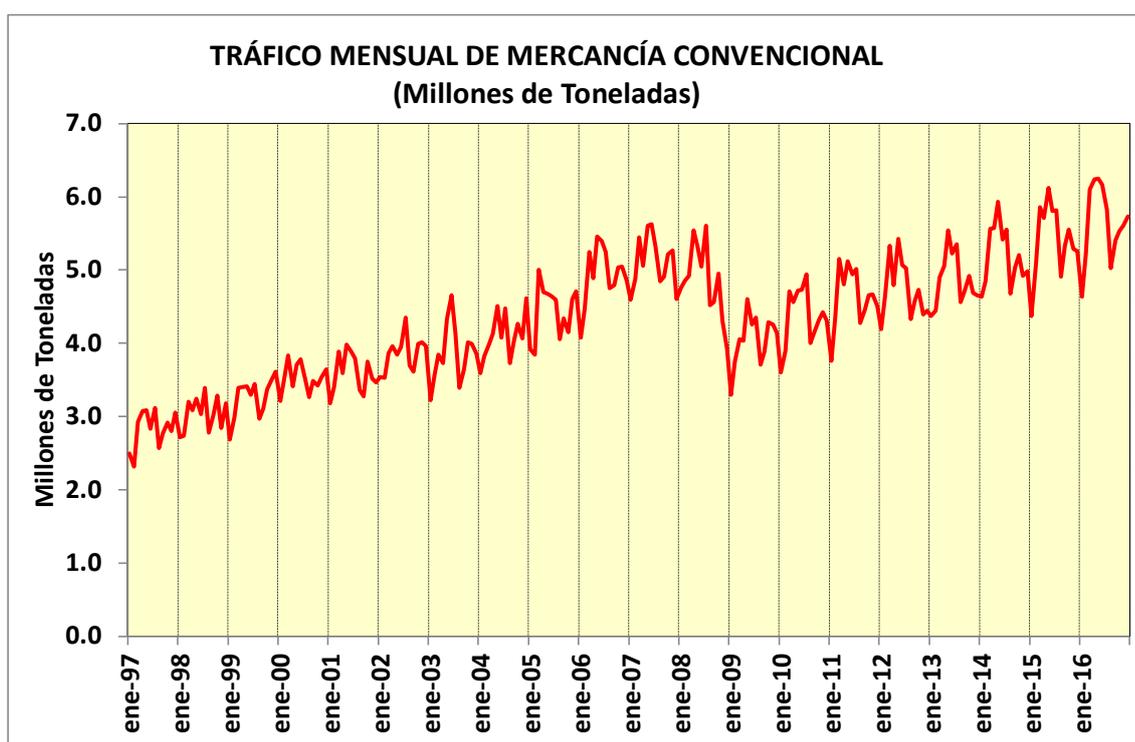


Figura 6. Tráfico marítimo de mercancía convencional.

Fuente: Elaboración propia.

En los gráficos 2, 3, 4, 5 y 6 se muestran, respectivamente, la evolución del tráfico marítimo entre enero de 1997 y diciembre de 2016, de todas las mercancías; de graneles líquidos; de graneles sólidos; de contenedores; y de mercancía convencional. Aparte de las pautas relacionadas con la tendencia y el impacto de la crisis económica que se analizaron anteriormente, las series mensuales del tráfico marítimo nos permiten esclarecer el grado de estacionalidad de las series. En general se observa la existencia de posible estacionalidad en todas las series, aunque este patrón parece menos marcado en la serie de tráfico de graneles líquidos. Esta característica deberá ser considerada en la modelización de las series mediante los modelos SARIMA, utilizando el operador diferencia de orden 12 para eliminar dicho ciclo estacional. El contenido de los epígrafes se escribe aquí.

4. METODOLOGÍA

4.1. Modelos: de predicción

Se han considerado los siguientes tipos de modelos para llevar a cabo la modelización y predicción².

1) SARIMA: Modelo Arima estacional (SARIMA): $(p,d,q) \times (P,D,Q)_{12}$.

Dentro del marco metodológico de este trabajo se ha utilizado como modelo lineal para explicar la dinámica subyacente en el tráfico marítimo a los modelos multiplicativos estacionales SARIMA (Box y Jenkins, 1970 y Box et al., 1994). Esta metodología se compone básicamente de los pasos siguientes: a) Se aplica el operador logarítmico con el fin de estabilizar la varianza. 2) A continuación se aplican los filtros adecuados a la serie original, en forma de diferencias regulares y estacionales, para obtener las correspondientes series estacionarias y sin ciclos estacionales. La estacionaridad se contrasta mediante el test de Dickey-Fuller. 3) A continuación se procede al proceso de modelización de las series estacionarias obtenidas, para el periodo que transcurre entre enero de 1997 y diciembre de 2016, utilizando modelos SARIMA. Se elige el modelo que presente un menor valor de los criterios de información Bayesiano (BIC) o de Schwarz (SIC).

2) Alisado exponencial en el espacio de los estados (ETS). En este tipo de modelos las predicciones están basadas en las observaciones anteriores ponderadas con sus correspondientes pesos (Hyndman et. al., 2002; y Hyndman et. al., 2008).

3) Extreme Learning Machine (ELM). Una extensa descripción de este tipo de modelos puede encontrarse en Huang et al. (2006.),

4) Redes neuronales (NNETAR). En este simple modelo de redes neuronales se toman como inputs a los valores retardados de la serie y una sola capa. Es un modelo del tipo feed-forward. Como todo modelo de redes neuronales permite recoger el efecto de una dinámica en la serie que no sea lineal.

² Se ha utilizado el lenguaje de programación R para la estimación de los modelos y las correspondientes predicciones.

4.2. Criterios de evaluación

Para la predicción con los diferentes modelos se ha considerado un periodo de predicción de 12 meses. Los tres criterios adoptados para evaluar y comparar la bondad del ajuste de los modelos considerados y la precisión de sus predicciones son los siguientes³:

1) Error absoluto medio (MAE; Mean Absolute Error): Es el promedio de los valores absolutos de los errores.

Se calcula mediante la expresión siguiente donde y_t es la serie considerada e \hat{y}_t las predicciones.

$$\frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

2) Error Porcentual Absoluto Medio (MAPE; Mean Absolute Percentage Error): Tiene la ventaja de no ser dependiente de la escala al medir el tamaño del error en términos porcentuales. Se calcula mediante la expresión siguiente:

$$\frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

3): Raíz cuadrada del error cuadrático medio: (RMSE; Root Mean Square Error): Se calcula mediante la expresión siguiente:

$$\sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

5. ANÁLISIS DE LOS RESULTADOS

Tabla 3. Resultados de la evaluación de los modelos.

Tipo de mercancía	Modelos (puesto en precisión)	Criterios de evaluación		
		MAE	MAPE	RMSE

³ En Hyndman y Koehler (2006), y Hyndman y Athanasopoulos (2018,) se describen las principales características de estos criterios.

total	Tráfico	SARIMA (2)	918669	2,8446	1663278
		ETS (1)	873059	2,7391	1580615
		ELM (4)	1288365	3,8664	2111130
		NNETAR (3)	1099928	3,3634	1851218
líquidos	Graneles	SARIMA (2)	538814	5.1061	786136
		ETS (2)	537163	5,8880	712565
		ELM (4)	590042	4,9990	782254
		NNETAR (1)	536718	4.5276	678443
sólidos	Graneles	SARIMA (1)	431360	5,6644	571937
		ETS (1)	430055	5,7461	555897
		ELM (3)	446335	5,8302	581545
		NNETAR (4)	465826	6,1439	591478
ores	Contenedores	SARIMA (2)	388806	4,6005	799155
		ETS (1)	358889	4,3888	758766
		ELM (4)	560699	5,9944	990516
		NNETAR (3)	462827	5,3465	890798
Mercancía convencional		SARIMA (2)	166288	4,0982	269577
		ETS (1)	164957	4,0933	264011
		ELM (4)	244820	5,7832	362121
		NNETAR (3)	198925	4,8107	306353

Fuente: Elaboración propia.

En la tabla 3 se muestran los resultados obtenidos en la evaluación de la predicción de los modelos adoptados, para cada criterio de evaluación y tipo de mercancía transportada. Se observa que el modelo más preciso en sus predicciones depende del tipo de mercancía transportada. El modelo ETS presenta las mejores predicciones en el tráfico total de mercancías, así como en los de graneles sólidos (junto al SARIMA), contenedores y mercancía

convencional. El modelo SARIMA es el que presenta mejores resultados en relación con la predicción del tráfico de graneles sólidos (junto al ETS). Finalmente, el modelo NNETAR es el que presenta predicciones más precisas en el tráfico de graneles líquidos. En resumen, cabe concluir que en general, la precisión de las predicciones de los modelos univariantes lineales es mayor que la correspondiente a los modelos univariantes no lineales. Tan solo en el caso del tráfico de graneles líquidos la precisión del modelo adoptado de redes neuronales es superior a la de los modelos lineales. Estos resultados podrían estar relacionados con el grado de linealidad de la serie correspondiente ya que los modelos no lineales como los de redes neuronales son más adecuados para el caso de que exista una dinámica no lineal en la serie analizada.

6. CONCLUSIONES

En este trabajo se lleva a cabo un análisis comparativo entre varios tipos de modelos univariantes, lineales y no lineales, en relación con su capacidad para modelizar y predecir el tráfico marítimo de mercancías en los puertos españoles. El estudio se lleva a cabo no solo para el tráfico marítimo total sino también específicamente para cada tipo de mercancía transportada: graneles líquidos, graneles sólidos, contenedores y mercancía convencional. Para la evaluación de la precisión de las predicciones correspondientes a cada modelo se utilizan tres criterios que son habitualmente empleados en la literatura internacional.

A partir de los resultados obtenidos cabe concluir que el modelo más preciso en sus predicciones depende del tipo de mercancía transportada. En general, la precisión de las predicciones de los modelos univariantes lineales (ETS y SARIMA) es mayor que la correspondiente a los modelos univariantes no lineales. Tan solo en el caso del tráfico de graneles líquidos la precisión del modelo adoptado de redes neuronales es superior a la de los modelos lineales.

Los resultados obtenidos sobre la dinámica existente en las series temporales de tráfico de mercancías en los puertos españoles son particularmente interesantes para los gestores de las autoridades portuarias, y de los diversos agentes que intervienen en la operación del transporte marítimo. Asimismo, los resultados obtenidos presentan interés para la planificación de las futuras inversiones en el sistema portuario, ya que para realizar un plan de actuaciones en infraestructuras portuarias es necesario disponer de predicciones precisas sobre la evolución futura del tráfico marítimo.

Las futuras investigaciones que den continuidad a este trabajo consisten en la comparación de las predicciones del tráfico portuario obtenidas con otros métodos alternativos: por ejemplo, modelos GARCH, en el caso de haberse detectado comportamiento no lineal estocástico, así como la consideración de otros modelos no lineales como Support Vector Machines o la ampliación del modelo de redes neuronales utilizado en esta investigación. Otros posibles trabajos consisten en la aplicación de la metodología empleada en esta investigación a otras series de tráfico marítimo como son el tráfico de pasajeros o el tráfico desagregado de cada puerto, así como a las series de datos de tráfico marítimo de otros países. El contenido de los epígrafes se escribe aquí.

3. REFERENCIAS BIBLIOGRÁFICAS

- BOX, G. y JENKINS, G. (1970). “Time Series Analysis: Forecasting and Control”. Holden-Day, San Francisco.
- BOX, G.E.P., JENKINS, G.M. y REINSEL, G.C. (1994). “Time Series Analysis; Forecasting and Control”. 3rd Edition, Prentice Hall, Englewood Cliff, New Jersey.
- CHEN, S.H. Y CHEN, J.N., (2010). “Forecasting container throughputs at ports using genetic programming”. *Expert Systems with Applications*, 37(3), pp. 2054–2058.
- COTO-MILLÁN, P., INGLADA-PÉREZ, L., CASARES, P. y INGLADA LÓPEZ DE SABANDO, V. (2018). “Modelización del transporte marítimo de contenedores”, *Estudios de Economía Aplicada*, pp. 675-690.
- COTO-MILLAN, P., BANOS-PINO J., PESQUERA, M.A., CASTANEDO, J.A. y INGLADA-PEREZ, L. (2010,). “The Conditioned Demands of “General Merchandise”, “Dry Bulk” and “Liquid Bulk” Sea Transport”. En *Essays on Port Economics*. Springer Verlag.
- GLEN, D. R. (2006). “The modelling of dry bulk and tanker markets: a survey”. *Maritime Policy and Management*, 33 (1), pp. 431-455.
- HUANG, G.-B., ZHU, Q.-Y. y SIEW, C.-K. (2006). “Extreme learning machine: Theory and applications”. *Neurocomputing*, 70, pp.489–501.
- HYNDMAN, R. J., KOEHLER, A. B., SNYDER, R. D., y GROSE, S. (2002). “A state space framework for automatic forecasting using exponential smoothing methods”. *International Journal of Forecasting*, 18(3), pp. 439–454.

- HYNDMAN, R.J. y KOEHLER, A.B. (2006). "Another look at measures of forecast accuracy". *International Journal of Forecasting*, 22(4), pp.679-688.
- HYNDMAN, R. J., KOEHLER, A. B., ORD, J. K., y SNYDER, R. D. (2008). "Forecasting with Exponential Smoothing: The State Space Approach". Berlin: Springer-Verlag.
- HYNDMAN, R.J. y ATHANASOPOULOS, G. (2018). "Forecasting: principles and practice," 2nd ed., OTexts, Melbourne, Australia.
- INGLADA-PEREZ L. (2015). "Evaluación de la no linealidad y del comportamiento caótico en el transporte marítimo". *Revista de Evaluación de Programas y Políticas Públicas (Journal of Public Programs and Policy Evaluation)*, 4, pp. 40-58.
- LAM, W. H., NG, P. L., SEABROOKE, W., y HUI, E. C. (2004). "Forecasts and reliability analysis of port cargo throughput in Hong Kong." *Journal of Urban Planning and Development*, 130(3), pp. 133–144.
- MOSTAFA, M. M. (2004). "Forecasting the Suez Canal traffic: A neural network analysis". *Maritime Policy and Management*, 31 (2), pp. 139-156.
- PENG, W.Y. y CHU, C.W., (2009). "A comparison of univariate methods for forecasting container throughput volumes". *Mathematical and Computer Modelling*, 50, pp. 1045–1057.
- WASHINGTON, S., KARLAFTIS, M. y MANNERING, F. (2003). "Statistical and Econometric Methods for Transportation Data Analysis", Chapman and Hall/CRC Press, Boca Raton, FL.
- CEA, F. y MORA, J.G. (1992). "Análisis socioeconómico de la demanda de estudios superiores". *Estadística Española*, 34, 129, pp. 61-92.