Transport Reviews: A Transnational Transdisciplinary Journal
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ttrv20

Classifying Ports for Efficiency Benchmarking: A Review and a Frontier-based Clustering Approach

Beatriz Tovar[a] & Héctor Rodríguez-Déniz[b]
[a] Department of Applied Economics, Infrastructure and Transport Research Group, University of Las Palmas de Gran Canaria, Campus de Tafira, Modulo D, Despacho 2.20, 35017 Las Palmas de Gran Canaria, Spain
[b] Statistics Division, Department of Computer and Information Science, Linköping University, 58183 Linköping, Sweden
Published online: 10 Mar 2015.

Transport Reviews, 2015
http://dx.doi.org/10.1080/01441647.2015.1016473

Routledge
Taylor & Francis Group

Classifying Ports for Efficiency Benchmarking: A Review and a Frontier-based Clustering Approach

BEATRIZ TOVAR*[§] AND HÉCTOR RODRÍGUEZ-DÉNIZ**

**Department of Applied Economics, Infrastructure and Transport Research Group, University of Las Palmas de Gran Canaria, Campus de Tafira, Modulo D, Despacho 2.20, 35017 Las Palmas de Gran Canaria, Spain; **Statistics Division, Department of Computer and Information Science, Linköping University, 58183 Linköping, Sweden*

(*Received 9 September 2014; revised 25 January 2015; accepted 2 February 2015*)

ABSTRACT *Port efficiency and port clustering are two aspects that have received different degrees of attention in the existing literature. While the actual estimation of port efficiency has been extensively studied, the existing literature has paid little attention to developing robust methodologies for port classification. In this paper, we review the literature on classification methods for port efficiency, and present an approach that combines stochastic frontier analysis, clustering and self-organized maps (SOM). Cluster methodologies that build on the estimated cost function parameters could group ports into performance metrics' categories. This helps when setting improvement targets for ports as a function of their specific cluster. The methodology is applied to a database of Spanish port authorities. The dendrogram features three clusters and five outlier Spanish Port Authorities. SOM are employed to track the temporal evolution of Spanish Port Authorities that are of special interest for some reasons (i.e. outliers). Results show that use of a combination of cost frontier and cluster methods to define robust port typology and SOMs, jointly or in isolation, offers useful information to the decision-makers.*

1. Introduction

Interest in productivity and efficiency analyses of companies has significantly grown in recent decades. They are especially relevant in sectors where adjustment is normal: for example, infrastructure industries and public utilities. Indeed, the possibility of comparing the performance of regulated companies does contribute to easing the problem of information asymmetries, by increasing efficiency in the regulatory agencies. The problem of distinguishing between heterogeneity and inefficiency is widely acknowledged in benchmarking and it is aggravated when international data sets are used.

The study of efficiency and productivity in the port sector using frontier techniques has also received a great deal of attention in recent years (Pallis, Vitsounis, De Langen, & Notteboom, 2011). In that regard, like-to-like comparisons are required, in order to set realistic targets for improvement (Jessop, 2012).

[§]Corresponding author. Email: btovar@daea.ulpgc.es

However, in most cases, there are differences across firms that are not reflected in the database. This information which is not picked up by the variables included in the data is named as unobserved heterogeneity.

Port efficiency and port clustering are two aspects that have received different degrees of attention in the existing literature. While the actual estimation of port efficiency has been extensively studied, the existing literature has paid little attention to developing robust methodologies for port classification. Cullinane and Song (2006), referring to this, stated the desirability of "developing appropriate clusters of ports in the sample that can be benchmarked against one another, in order to identify sources of inefficiency and measures for its amelioration". Despite this necessity, only a few papers have attempted to produce a robust typology of ports by using specific techniques, such as cluster analysis, to avoid classifications based on ad hoc criteria.

Port efficiency benchmarking could benefit from combined efficiency estimates with cluster analysis, especially if the sample is large and heterogeneous. In this paper, we review previous contributions that combine both frontier-based productivity and/or efficiency estimation with port classification. About 70% (10 out of 14) of the surveyed papers are based on international data. Although the best practice frontier may be estimated by using either non-parametric (e.g. Data Envelopment Analysis — DEA) or parametric techniques (e.g. Stochastic Frontier Analysis –SFA), almost all reviewed papers have used DEA after splitting the sample.

Environment heterogeneity is one characteristic manifestation of firm heterogeneity. Due to ports being characterized by their geographical and operational settings, DEA could not be a proper approach, unless we split the sample into homogeneous groups before the frontier estimation; this is to avoid any producer-specific heterogeneity being considered as inefficiency. The latter approach is not without problems, as the researcher has to decide the criteria to split the sample; secondly, some groups could be so small that it would be illogical to estimate their frontier. Conversely, in an SFA context, a wide range of models that incorporate unobserved heterogeneity have been proposed, as they let us distinguish between unobserved individual heterogeneity and inefficiency. Whether a SFA model that accounts for unobserved heterogeneity is used the need for splitting the sample can be avoided and a reliable classification could be generated. Moreover, other additional advantages could be achieved as Rodríguez-Déniz and Voltes-Dorta (2014) have recently shown.

Their approach combines all the available information from the efficiency estimation to produce clusters based on the relevant multi-dimensional criteria. Regarding the comparative analysis, the hierarchical method allows to precisely identify the efficiency benchmarks within each cluster. In this paper, we present an empirical application of the latter methodology on the Spanish port authorities, using the cost frontier parameters and efficiency estimates by Rodríguez-Álvarez and Tovar (2012); these develop a stochastic frontier model that controls unobserved heterogeneity. We will then be able to define the port authority categories that mirror the performance indicator by using hierarchical clustering, thus filling a gap in the port benchmarking literature.

The article is organized as follows. Section 2 addresses the previous literature measuring productivity and/or port efficiency through frontier techniques, which also classify ports in groups. Section 3 introduces the methodological framework, that is, variable weighting, hierarchical clustering and Self-Organized-

Maps (SOMs). Section 4 describes the resulting port clusters and efficiencies, and the temporal analysis of selected ports. Section 5 concludes the paper.

2. Literature Review

Productivity and efficiency are related but different concepts although both are equally utilized in ports' benchmarking. Productivity is the ratio between the products obtained and the factors used in its production. On the other hand, efficiency is, broadly speaking, the ability to do something or produce something without wasting materials, time, or energy.[1] With the previous definitions, we can easily deduce that efficiency is only one of the factors that determine productivity.

Woo, Pettit, Beresford, and Kwak (2012) show that a topic relatively widely researched throughout the 1980s and 1990s was port performance. Port performance studies evolved in the 2000s due to the development of new measures and approaches. The frontier approach became popular not only to measure port productivity but also as a measure of port competitiveness. A recent brief review of papers measuring total factor productivity (TFP) in ports shows that the frontier approach is the most popular method (see Chang & Tovar, 2014). Probably the reason behind this popularity is that a frontier approach must be used in order to take into account the contribution of efficiency change upon productivity change. The traditional approach does not take into account companies' inefficiency; this means that it is not capable of distinguishing which part of the productivity changes is due to efficiency changes.

Broadly speaking, productivity change sources include efficiency change and technical change. These main components could be disaggregated as well. The first one, efficiency change, could be decomposed into the pure technical efficiency change (PECH) and the Scale Efficiency Change (SECH). This distinction enables to contemplate those situations where a productive unit can be technically efficient, as the production volume uses the least quantity of factors; however, it is not situated in the optimum production scale, because it is not adequately sized (it is either too small or too big). Therefore, the changes in productivity that are strictly related to technical efficiency appear in PECH, while these related to the productive unit size appear in SECH (Wilmsmeier, Tovar, & Sanchez, 2013). On the other hand, the second component of productivity, technical change, could also be decomposed into three components: pure technical change (PTCH), associated with a parallel shift of the efficient frontier; non-neutral change (NNTCH) and scaling augmenting (SATCH) that are associated with whether technological change is biased towards any inputs or any outputs, respectively (Chang & Tovar, 2014).

The literature on the measurement of port productivity and/or port efficiency using frontier models can be grouped into two main categories, in terms of the method used to estimate the frontier. The first sets of studies use non-parametric models, or DEA; a review of the literature on port efficiency using DEA can be found in Panayides, Maxoulis, Wang, and Koi Yu (2009) and Simoes and Marques (2010). The second set is formed by studies using parametric techniques, namely SFA, and there are far fewer papers compared to the DEA group.[2] A recent brief review of papers measuring efficiency in ports using SFA can be found in Tovar and Wall (2015).

DEA and SFA represent two alternative methods to measure efficiency (e.g. Wanke, Barbastefano, & Hijar, 2011). Both techniques allow derivation of relative efficiency ratios within a group of analysed units, so the efficiency of the units is compared through an efficient envelopment. Both methods have advantages and drawbacks. DEA does not impose any functional form to the frontier nor does it assume a distributional form for the inefficiency error terms, but it could be influenced by noise, and traditional hypothesis tests are not possible except by using bootstrapping techniques (Simar & Wilson, 2000). On the other hand, SFA involves the cost of imposing a particular functional form and making particular distributional assumptions for the one-side error term associated with technical efficiency, but it is capable of managing random shocks and/or measurement error. Moreover, traditional hypothesis tests could be used and, finally, environmental variables and firm heterogeneity are easier to deal with.

Table 1 shows a summary of previous papers measuring productivity and/or port efficiency through frontier techniques that also classify ports into groups. It can be seen that these vary widely in scope. Most of the studies use international data sets (e.g. Notteboom, Coeck, & Van Den Broeck, 2000; Sharma & Yu, 2009). Those that analyse efficiency for a single country include Martinez-Budria, Diaz-Armas, Navarro-Ibanez, and Ravelo-Mesa (1999), Kaisar, Pathomsiri, and Haghani (2006) and Medal-Bartual and Sala-Garrido (2011). The studies also differ in terms of the frontier approach used to measure productivity and/or efficiency, although DEA is the most commonly used approach (13 out of 14 papers). However, there are some differences in the models used, and these could be categorized into three groups. First are studies measuring technical efficiency using standard DEA (Bichou, 2013; Kaisar et al., 2006; Koster, Balk, & van Nus, 2009; Martinez-Budria et al., 1999; Medal-Bartual & Sala-Garrido, 2011; Sharma & Yu, 2009; Wu & Goh, 2010). Second are those studies measuring technical efficiency using more sophisticated DEA models (Cheon, 2009; Cullinane & Wang, 2010; Hung, Lu, & Wang, 2010; Quaresma-Dias, Garrido Azevedo, & Ferreira, 2009). In the third category are the studies measuring TFP (Cheon, Dowall, & Song, 2010; Guironnet, Peypoch, & Solonandrasana, 2009). Finally, we only found one study based on SFA, Notteboom et al. (2000), which employs a Bayesian Stochastic Frontier Modelling; this present study counts as the second to use such methodology.

The goal of clustering is to find meaningful groups in data (that is, clusters). This is normally achieved through the use of distance measures, so that the similarity among the elements in the same cluster is greater than the similarity among the elements belonging to different clusters. When it comes to efficiency benchmarking, ports are classified into a number of groups or categories to facilitate a better comparative analysis (see Table 1). However, as Table 1 corroborates, there has been a clear trend in the recent literature towards ad hoc classification. Most of these studies have lacked a methodological framework, which gives theoretical support to their clustering. Thus, they rely on classification schemes based on ad hoc criteria selected by authors; for example, traffic mix (Bichou, 2013), geographical situation (Hung et al., 2010; Kaisar et al., 2006), or their functional role (Notteboom et al., 2000).

Among the papers using ad hoc classification one, Martinez-Budria et al. (1999) is of special interest to us, since it studied the relative efficiency of all the Spanish port authorities during the 1993–1997 period, by using DEA. Recognizing explicitly the heterogeneity of Spanish port authorities, they stated: “In the first place we have needed to group all the ports into homogeneous sets, in order to reach conclusive results from the application of the DEA technique”. Their classification

Table 1. Classification/clustering in port efficiency studies

Author (year)	Data	Approach Efficiency	Clustering Variables	Clustering Method	Number of Clusters	Cluster Criteria
Martinez-Budria et al. (1999)	26 Spanish port authorities 1993–1997	TE, DEA (IO)	Level of complexity of ports	Ad hoc	3 groups: namely, high-, medium- and low-complexity ports	Ad hoc
Notteboom et al. (2000)	36 European container terminals, supplemented with 4 Asian container ports, 1994	TE, Bayesian Stochastic Frontier modelling	Grouping of pooled efficiencies according to (1) Port range (2) Functional role	Ad hoc	(1) 4 groups: Hamburg-Le Havre, Mediterranean, Atlantic port ranges and UK (2) 2 groups: hub and feeder	Ad hoc
Kaisar et al. (2006)	25 US container ports 1998–2003	TE, DEA (OO)	Regional grouping	Ad hoc	3 group: Gulf coast, West Coast and East Coast	Ad hoc
Sharma and Yu (2009)	70 WWide container terminals	TE, DEA (OO)	Quay length, Terminal area, Quay cranes, Transfer cranes, Straddle carriers, Reach stackers	SOM	4 terminal clusters obtained from SOM	SOM
Quaresma Dias et al. (2009)	10 Iberian Peninsula container terminals 2008	TE, Tiered DEA (OO)	Number of cranes, number of employees, terminal area, number of trailers, yard equipments, terminal length	SOM	4 terminal clusters obtained from SOM	SOM
Koster et al. (2009)	38 WWide Terminals 2006	TE, DEA (OO and IO)	Geography, operations, handling system, size, owner	Ad hoc	21 groups: geography (6),operations (2) handling system (2), size (4), owner (7)	*T*-tests and ANOVA
Guironnet et al. (2009)	24 Italian and 13 French ports 2003–2004	DEA, TFP, LUM	Classify ports depending on the EFFCH and TECH	Ad hoc	2 groups	Group 1: improvement in both Group 2: improvement in TECH coexist with deterioration of TE

(*Continued*)

Table 1. Continued

Author (year)	Data	Approach Efficiency	Clustering Variables	Clustering Method	Number of Clusters	Cluster Criteria
Cheon (2009)	110 World container ports 2004	TE, Recursive DEA (OO)	Recursive DEA Classifies ports into rank-ordered peer groups	Recursive DEA	7 groups according	Technical efficiency = 1
Cullinane and Wang (2010)	25 World Ports 1992–1999	TE, DEA, Window An.	Local competition	Ad hoc	2 groups — local competition	Ad hoc
Wu and Goh (2010)	21 WWide container ports 2005	TE, DEA (OO)	CCR and CCR Cross-efficiency levels	Ad hoc	3 groups according self-assessment using CCR model and the relative-assessment CCR cross-efficiency	Assessment values are divided into 3 groups: less favourable, favourable and more favourable
Hung et al. (2010)	31 container ports in the Asia-Pacific region 2003	TE, DEA, MPSS Bootstrap	Geographical location	Ad hoc	3 region groups according to geographical situation	Geographical situation
Cheon et al. (2010)	98 World Ports 1991 and 2004	DEA,MPI (OO)	Port ownership	Ad hoc	4 groups: ownership change (2) and corporate change structure (2)	2 groups — ownership change
			Corporate structure			2 groups corporate change
Medal-Bartual and Sala-Garrido (2011)	28 Spanish ports authorities 1994–2008	TE, DEA (IO)	Fixed assets regarding total cargo Gross return	HC	4 groups after dendrogram observation	Dendrogram cut-off
Bichou (2013)	60 World Container terminals 2004–2010	TE, DEA (IO)	Container mix (size, type and op. status): High % of inbounds, outbounds, T/S, FEUs, empties, laden containers (FCLs, LCL); Low % inbounds, FEUs	Ad hoc	7 groups according to container mix	50% cut-off proportion
Present study	26 Spanish port authorities 1993,1997,2003,2007	CE, SFA	4 outputs: bulk solids, bulk liquids, general cargo and passengers; 2 input prices: materials & labour, and a quasi-fixed input	HC	3 Clusters	Dendrogram cut-off

Note: TE, technical efficiency; CE, cost efficiency; TFP, total factor productivity; EFFCH, efficiency change; TECH, technical change; DEA, data envelopment analysis; LUM, Luenberger productivity indicator; MPI, Malmquist productivity index; IO, input oriented; OO, output oriented; HC, hierarchical clustering; SOM, Self-Organizing Map; CCR, Charnes, Cooper and Rhode model.

divided the Spanish port authorities into three groups using "a complexity criterium given by port size and the composition of the output vector". Their results show a different evolution for every group, in terms of efficiency; and they concluded that, by having gone closer to the frontier over time, highly complex ports show higher comparative efficiency levels.

Only four papers do not follow this ad hoc classification trend. Two use neural networks (SOM) to cluster[3] their data (Quaresma-Dias et al., 2009; Sharma & Yu, 2009). A third uses recursive DEA, in order to group ports with similar efficiency levels (Cheon, 2009). The fourth performs a hierarchical clustering (Medal-Bartual & Sala-Garrido, 2011); this is of special interest to us, because it is another study on the relative efficiency of all the Spanish port authorities during the 1994–2008 period. Sharma and Yu (2009) developed a decision support framework that comprises homogeneous groups and a stepwise improvement path for each container terminal. They use SOM to cluster the container terminals according to their input characteristics, and recursive DEA to obtain stratified efficiency levels. Similarly, Quaresma-Dias et al. (2009) carried out a benchmarking analysis of the main Iberian seaports, focusing on the efficiency of their container terminals in 2007, by using a number of input/output performance indicators. Cheon (2009) attempted to estimate whether the participation of global terminal operators increases port efficiency; this was done by using recursive DEA to generate a set of seven port efficiency peer clusters, from the most to the least efficient. Finally, Medal-Bartual and Sala-Garrido (2011) classified Spanish port authorities using agglomerative hierarchical clustering and an Euclidean distance, based on two variables: fixed assets and gross returns. The clustering was performed on the average values of the sample period, 1994–2008, and four clusters were identified. Subsequent DEA analysis reveals that these highly specialized ports are likely to achieve better gross returns and also higher activity levels (in terms of traffic).

We actually borrow many details from these studies, including the preference for multi-dimensional clustering, and the use of self-organized maps (SOM). In addition, we contribute to the literature by classifying ports using a methodology that combines SFA and cluster analysis, as opposed to Cheon (2009), who does not apply traditional clustering techniques. Moreover, the use of hierarchical clustering allows for a more detailed analysis and identification of the efficiency benchmarks; this is in contrast to other methods such as *K*-Means or SOMs that produce flat clusters (e.g. Quaresma-Dias et al., 2009; Sharma & Yu, 2009). The weighting of the relevant clustering variables has not been considered in any of the previous papers, not even by Medal-Bartual and Sala-Garrido (2011), though they use hierarchical clustering as well. Finally, the present paper goes further when using clustering variables from a model. It explicitly takes into account the existence of unobserved heterogeneity, to avoid a bias in relative efficiency scores.

Therefore, to the best of our knowledge, the present paper is the first to combine efficiency estimates obtained through SFA with hierarchical clustering analysis. Environmental variables and firm heterogeneity are easier to deal with using SFA. The combination of both methodologies has the advantage of building clusters, based on the relevant multi-dimensional criteria, which also take into account unobserved heterogeneity. In this way, we ensure that we are making like-to-like comparisons by getting clusters that are based on a two-step procedure involving efficiency/productivity estimation and port clustering based on the same explanatory set (cost elasticities and factor shares).

3. Methodological Framework

3.1. *Variable Selection, SFA Model and Weighting*

We take advantage of the database, coefficients, and efficiency estimates from Rodríguez-Álvarez and Tovar (2012) whose model is especially suitable for our purpose because it distinguishes individual unobserved heterogeneity, inefficiency and stochastic errors. Another reason to choose this study is the availability of a sample of Spanish port authorities, which overlaps both in terms of ports and time periods with the paper by Medal-Bartual and Sala-Garrido (2011). This offers us the opportunity to compare the results obtained by both approaches, and to analyse how these results change when the individual unobserved heterogeneity is taken into account; this is the case for both in the efficiency estimation and in the clustering method followed.[4]

Rodríguez-Álvarez and Tovar (2012) estimated a stochastic cost frontier using a panel of Spanish port authorities during the 1993–2007 period, which provides all that we require to apply the proposed methodology to ports.[5] The sample is made up of 390 observations that correspond to a total of 26 port authorities between 1993 and 2007. The port authorities analysed are: A Coruña, Alicante, Avilés, Bahía de Algeciras, Bahía de Cádiz, Baleares, Barcelona, Bilbao, Cartagena, Castellón, Ceuta, Huelva, Las Palmas, Málaga, Marín y Ría de Pontevedra, Melilla, Motril, Pasajes, Sta. Cruz de Tenerife, Santander, Sevilla, Tarragona, Valencia, Vigo and Vilagarcía.

The port authorities included in the sample vary widely in terms of size, specialization, logistics connectivity, geographical situation, maritime connectivity, and so on. Some port authorities manage cargo and passenger traffic, whereas others run ports whose passenger transport activity is virtually non-existent. Even for cargo-orientated ports, there is a variety of classifications, depending on the type of merchandise; furthermore, their sizes and roles as distribution centres and ports of final destination are important. Rodríguez-Álvarez and Tovar (2012) considered the possibility of there being unobservable heterogeneity among Spanish port authorities and formulated a model to test it. Theirs is a so called "true" model[6] because it includes two terms for unobserved heterogeneity. One term controls time-variant factors, the other takes into account the producer-specific invariant characteristics. The basic assumption is the existence of producer-specific and time-invariant factors that cannot be captured by explanatory inefficiency variables, due to the variation of the latter over time and/or omitted variables.

The set of explanatory variables used in efficiency estimation and port clustering are those used in the estimation of our reference cost frontier which features four outputs: bulk solids (y1), bulk liquids (y2), general merchandise (y3), and passengers (y4); a quasi-fixed input (K) plus two input prices: labour (wl) and materials (wi). In addition, the time variable (t) was added to the specification in order to account for technical change and a Spanish port authorities dummy variables (D_i) to capture unobservable heterogeneity. Table 2 presents the descriptive statistics for the variables. The maximum and minimum values in the sample show a high degree of heterogeneity, in terms of the size and specialization of each port.

Finally, following Rodríguez-Déniz and Voltes-Dorta (2014), we will use the output cost elasticities (η) and input cost shares (s), derived from Rodríguez-Álvarez and Tovar (2012) estimation,[7] as optimal variable weights (see Table 3).

Table 2. Descriptive analysis of the data

Variable	Unit	Description	Average	Standard deviation	Minimum	Maximum
CT	€ deflated	Total cost	1.88E+07	1.38E+07	2.47E+06	8.04E+07
CV	€ deflated	Variable cost	1.14E+07	8.68E+06	1.86E+06	5.38E+07
y1	Tm	Bulk solids	3.24E+06	3.39E+06	36128	1.97E+07
y2	Tm	Bulk liquids	4.93E+06	6.13E+06	377	2.28E+07
y3	Tm	General merchandise	4.44E+06	7.71E+06	79606	4.71E+07
y4	Number	Passengers	691040	1.31E+06	0	5.93E+06
wL	€ deflated /Workers	Labour price	30610.9	6538.69	16789.8	53016.8
wI	€/m^2	Intermediate inputs price	2.60082	1.79541	0.198284	16.1609
K	m^2	Quasi-fixed input	2.47E+06	2.84E+06	194314	1.72E+07
T	Year	Trend	8	4.32604	1	15

Source: Rodríguez-Álvarez and Tovar (2012)

Table 3. Summary of cluster analysis properties and methodological decisions

Clustering method	Hierarchical clustering results in a more informative structure than the flat clusters obtained from divisive methods (e.g. K-Means) or neural network-based clustering (e.g. SOM). In addition, HC does not require us to predefine the number of clusters
Performance indicator	Logged total operating costs (ln C)
Unidimensional vs. Multi-dimensional classification	Multi-dimensional scaling is preferred because neither, for example, container traffic nor geographic location fully captures the cost structure of a port. Consistency requires that clustering be based upon the same variables that explain cost efficiency
Clustering variables	Logged outputs (bulk solids, bulk liquids, general merchandise, and passengers) input prices (labour and materials) and a quasi-fixed input (K)
Variable weighting and standardization	Output cost elasticities (η) and input cost shares (s) derived from the industry's translog cost frontier estimated in Rodríguez-Álvarez and Tovar (2012) $\eta_{ik} = \left.\frac{\partial \ln C}{\partial \ln w_h}\right\rvert_{(y,w)i} \quad s_{ih} = \left.\frac{\partial \ln C}{\partial \ln w_h}\right\rvert_{(y,w)i}$
Distance measure	Euclidean distance is the most widely used metric for interval data
Clustering algorithm	Average linkage (0.96) was chosen over single linkage (0.93), complete linkage (0.91), centroid linkage (0.95), median linkage (0.95) and Ward's method (0.81) according to the cophenetic correlation coefficient c (Sokal and Rohlf, 1962). $c = \frac{\Sigma_{i<j}(\mathrm{d}(i,j) - \bar{\mathrm{d}})(t(i,j) - \bar{t})}{\sqrt[2]{\Sigma_{i<j}(\mathrm{d}(i,j) - \bar{\mathrm{d}})^2 \Sigma_{i<j}(t(i,j) - \bar{t})}},$ where d,t represent Euclidean and cophenetic (tree-) distances, respectively
Cluster determination	The dendrogram was truncated at a dissimilarity level of 0.42
Clustering software	MATLAB. Statistics toolbox
Dendrogram viewer	Interactive tree of life (Letunic and Bork, 2007)

Source: Adapted from Rodríguez-Déniz and Voltes-Dorta (2014)

These optimal variables have the additional advantage that the optimal variable weights, which are to be used for building port authorities clusters, come from a model that took into account unobserved heterogeneity. As we have indicated, it is very convenient to introduce these individual effects, in order to avoid any diversity being termed as inefficiency; this is because we are dealing with units, that is, Spanish port authorities, with widely differing characteristics.

3.2. *Hierarchical Clustering*

Cluster analysis, or simply clustering, is one of the most popular data analysis techniques,[8] with a plethora of applications in fields ranging from astronomy to medical imaging. There is a wide variety of clustering techniques, which can be roughly classified into hierarchical and divisive methods (Abonyi & Feil, 2007). Agglomerative hierarchical clustering (*bottom-up*) starts with each object assigned to an individual cluster. This is followed by an iterative process of merging, which is performed until all objects are assigned to a single cluster. As a result, a tree-like diagram, a dendrogram, is usually generated to illustrate this process and to analyse the underlying structure of the data. On the other hand, divisive methods (e.g. k-means) provide a flat partition of the input data set into a fixed number of clusters, instead of a group hierarchy; they are particularly useful when a specific number of output categories is required, or the number of objects complicates the dendrogram inspection.

We used the frontier-based hierarchical clustering approach proposed by Rodríguez-Déniz and Voltes-Dorta (2014)[9] and present the results in a dendrogram, which allows natural groups to be identified graphically. It is also possible, as is detailed in Section 4.1, to include the efficiency estimates in the dendrogram, and analyse them at both cluster and individual levels. Regarding the partitioning method, we chose average linkage (using an Euclidean distance as distance measure) instead of other approaches like single linkage (nearest neighbour) and complete linkage (farthest neighbour) which is sensitive to outliers. A methodological summary of the cluster analysis performed is presented in Table 3.

The clustering application has to be done over a cross-section of the original data set in Rodríguez-Álvarez and Tovar (2012). This article evaluates the impact of the legislative reforms that have taken place in the Spanish port sector in the last 15 years, within an appropriate short run total cost model. In their analysis, three legislative periods have been defined. The first starts with the 27/1992 Law, and covers the period from 1993 to 1997. The second begins with the passing of the 62/1997 Law and runs from 1998 to 2003. The third corresponds to the 42/2003 Law and covers the period from 2004 to 2007.

Four clustering processes were performed, and they correspond to the selected years in the sample period that mark the three periods defined by the legislative reforms; that is, 1993, 1997, 2003, and 2007. Dendrograms were generated for each year. The level of dissimilarity between two clusters is given by the height[10] of the point at which their branches merge. Note that hierarchical clustering does not require us to define the number of clusters a priori. This can instead be identified by direct examination or by using a variety of "tree-cutting" techniques; see Milligan and Cooper (1985). A very intuitive way to determine the number of port authorities clusters was chosen, that is, cutting the dendrogram at a level of dissimilarity of 0.42. This produces eight clusters, three actual clusters and five outlier ports, which will be described in Section 4.1.

As we mentioned, dendrograms, although illustrative, cannot be readily interpreted in terms of the original variables. This is especially true when dealing with multi-dimensional data, as this present study does. In addition, the different dendrograms are hardly comparable, which restricts the analysis to a cross-sectional study. In order to overcome these limitations, we have complimented our cluster analysis using SOM, which will enable us to track the temporal evolution of some interesting ports, at each variable level; this provides a better understanding of the dynamics of these ports.

3.3. *Kohonen's SOMs*

A SOM, developed by Kohonen (1982), is a two-layered neural network model that projects a k-dimensional input space onto a low-dimensional space by using unsupervised learning.[11] The output space is normally arranged in a rectangular lattice of i x j neurons, as shown in Figure 1. Samples from the original input space are fed into the network through the input layer, which consists of n neurons that are fully connected to the output map. These connections have weights associated (codebook vectors), W_{ij}, that determine the strength of the response of the (i, j)th output neuron to an input pattern x.

Weights are computed following a competitive learning process. When an input sample is presented to the network, each output unit (neuron) calculates the similarity (e.g. Euclidean distance) between its codebook vector (weights) and the input vector. The neuron whose weight vector is closest to the input vector is designed as the Best Matching Unit (BMU). Then, the codebook vector of the BMU is adjusted towards the input vector, thus increasing the sensitivity of the BMU to such an input pattern. This process is carried out until weight changes are negligible. However, the originality of Kohonen's approach relies on the fact that neurons near the BMU will have their weights updated as well, according to some kind of neighbourhood function; for example, Von Neumann neighbourhood.[12] The result is a SOM that preserves the topology of the input space. It does this in such a way that similar input patterns are mapped to points that are neighbouring in the output space. Still, Kohonen's algorithm does not produce actual clusters, that is, disjoint groups of similar objects. There are a number of

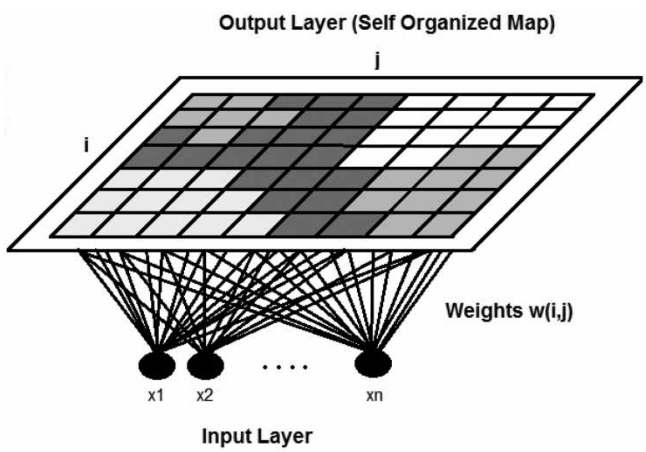

Figure 1. Architecture of a two-dimensional SOM.

methods, such as the U-Matrix by Ultsch and Siemon (1990), aimed at identifying and delimiting clusters in SOM; these allow the automatic classification of input vectors, depending on which region within the map their corresponding BMUs lie.

There are an extensive number of issues that can be addressed using SOM, such as clustering, feature extraction and vector quantization. A particularly interesting application of SOMs is the temporal analysis, as long as the available data is time-ordered, see for example, Martín-del-Brío and Serrano-Cinca (1993) and Sarlin (2013). Once the network is trained and the map generated, one can easily create time series of BMUs (i.e. trajectories) in order to track the progression of the samples. This strategy has been used in the literature to study cluster membership and temporal evolution. In this last regard, this paper presents the first application of SOMs to analyse the temporal response of port authorities.

4. Results and Discussion

4.1. *Port Authorities Clusters and Efficiency*

Figure 2 shows the dendrogram for 2003, which features both the resulting partition and the efficiency estimates for each port authority in that year.[13] As we mention before, we cut the dendrogram at a level of dissimilarity of 0.42, which results in eight clusters, three actual clusters and five outlier port authorities. Figure 2 also presents the Economic Efficiency (EE) estimates obtained from our reference study. Note that these values take into account the distance to port authorities cost frontier. The industry average EE is 0.9533 for the year 2003.

The port authorities dendrogram allows for a more precise identification of efficiency benchmarks, which are now characterized by a vector of cophenetic distances to their "peers". This information effectively adds another dimension to the comparative analysis, and helps in setting sharper targets for improvement. However, the use of multi-dimensional scaling also leads to a more complex characterization of clusters, which may complicate the generalizability of results for out-of-sample port authorities. In our case, the resulting three clusters are characterized by seven different dimensions, and hence the interpretation of results is not so straightforward. Table 4 presents the minimum, maximum, and average values of each dimension per cluster.[14] These values serve as reference

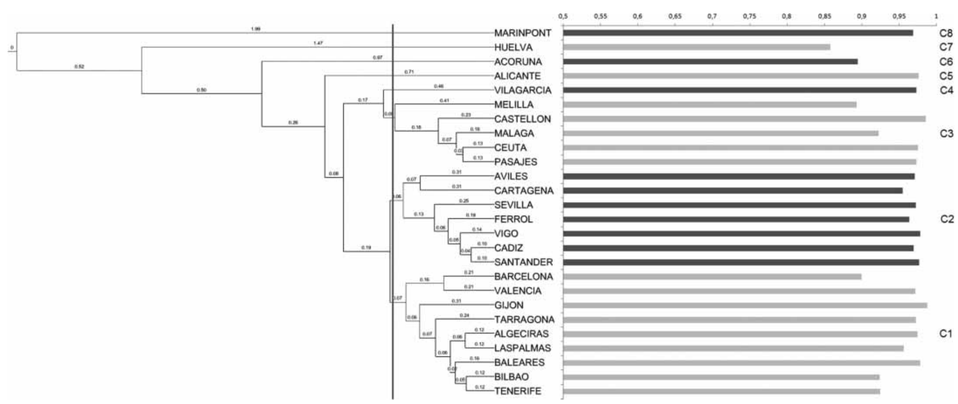

Figure 2. Dendrogram (truncated at 0.42), port clusters, and efficiency benchmarks for 2003.

Table 4. Spanish port authorities clusters

		y1	y2	y3	y4	wl	wi	k	EE
Cluster 1	**Max**	16,895,000	21,570,000	32,370,000	5,011,416	48,908	2.24	8,288,910	0.9873
	Average	**5,313,667**	**9,344,556**	**13,160,444**	**1,708,168**	**36,559**	**2.86**	**3,544,658**	**0.9540**
	Min	1,558,000	1,439,000	467,000	1,044	31,016	3.21	1,743,306	0.8992
Cluster 2	**Max**	7,596,000	16,543,000	3,172,000	1,71,264	35,801	0.93	4,802,939	0,9779
	Average	**3,315,167**	**3,118,333**	**1,611,667**	**43,240**	**32,855**	**1.16**	**2,741,709**	**0.9679**
	Min	774,000	57,000	3,76,000	0	31,215	1.64	1,285,709	0.9543
Cluster 3	**Max**	3,447,000	6,706,000	2,200,000	2,194,744	41,780	3.33	957,191	0.9854
	Average	**1,535,000**	**1,556,400**	**9,77,600**	**5,98,888**	**35,923**	**3.43**	**715,997**	**0.9496**
	Min	52,000	74,000	3,68,000	0	30,654	4.74	274,646	0.8928

indicators for out-of-sample ports, in order to easily determine cluster membership.

At a glance, the dendrogram (see Figure 2) features three broad clusters (C1, C2, and C3) and five outlier port authorities. Note that the main clusters have been clearly defined according to the scale of production, and thus comprise a variety of port authorities profiles, in terms of traffic mix. Therefore, it seems that a port's complexity, in the sense defined by Martinez-Budria et al. (1999), plays a role in clustering the Spanish port authorities. This partition at the top of the hierarchy indicates that grouping, by complexity, size and traffic mix, may be adequate when it comes to benchmarking the cost-efficiency of Spanish port authorities. This contrasts with other studies, such as Medal-Bartual and Sala-Garrido (2011), which obtain clusters that are basically characterized by the type of traffic. Note that each port authority in the dendrogram should be analysed carefully, since their classification might not be explained by a single variable, but by a combination of factors.

Cluster 1 (C1) includes the following nine port authorities: Algeciras, Baleares, Barcelona, Bilbao, Gijón, Las Palmas, Tenerife, Tarragona, and Valencia, which are characterized by large-scale production. These are the main commercial ports, with total through outputs higher than 10 million tonnes in 2003. There is also a predominance of general cargo in this cluster, although the ports have all types of cargo including passenger traffic. The average EE for this cluster in 2003 was 0.95. Upon closer examination, we found that Barcelona and Valencia, two of the three biggest general cargo ports, are merged into the same sub-cluster. Note that Algeciras has been classified with Las Palmas, and that both ports are also among the biggest general cargo ports. Although they represent a significant share of liquid bulk, as they are among the top bunkering ports in Spain, they are also important in transhipment traffic. In this case, the liquid bulk traffic acted as a discriminator between both subgroups, Algeciras-Las Palmas and Barcelona-Valencia. Furthermore, Gijón and Tarragona, the top solid bulk ports, also appear next to each other. Nonetheless, they are clearly separated from the rest of the elements of the lower sub-cluster of C1, due to the absence of passenger traffic in these two particular port authorities. Moreover, Gijón, the port authority with the highest level of solid bulk traffic, appears as the top C1 performer, (0.99).

The second cluster (C2) includes seven port authorities: Avilés, Cádiz, Cartagena, Ferrol, Santander, Sevilla, and Vigo which are characterized by medium-scale production. They have bulk-based traffic, both solids, and liquids, although they also have general cargo and, to lesser extent, passenger traffic as well. The cluster has been also partitioned into two subgroups, according to the quasi-fixed input levels. Avilés and Cartagena, with very different traffic characteristics, were separated from the rest of the elements due to their lower quasi-fixed input level. The five remaining port authorities include Cadiz, Ferrol, Santander, Sevilla and Vigo, Vigo is the top performer and the two elements closest to each other are Cadiz and Santander, which are the only port authorities with significant passenger traffic. The average EE for C2 in 2003 was 0.97. Cluster 3 groups ports are characterized by low-scale production, and are mainly gateways for their natural hinterland. They present a similar pattern to that shown by C2; that is, main subgroups are defined by the quasi-fixed: Castellón, Ceuta, Málaga, Melilla and Pasajes. Melilla is isolated, given its very low quasi-fixed level. Similarly, a port with a very significant level of liquid bulk traffic, Castellón, was also grouped apart from the remaining four. Ceuta and Pasajes appear to be the most similar port authorities, although this case cannot be clearly explained by any single component. Finally, the average EE for this cluster was 0.95 in 2003.

After truncating the dendrogram at the selected level (0.42), the remaining port authorities (C4−C8) can be considered outliers, maybe with the exception of Vilagarcía (C4), which would fit in C3 at the truncation level 0.5. Their average EE is 0.93, with Alicante as top performer (0.98). Marín-Pontevedra is probably the most conspicuous outlier in 2003, as it is the only port authority in the sample with virtually zero bulk liquid traffic. Conversely, the position of A Coruña and Alicante cannot be so readily explained. They would perfectly fit in C1 and C3, respectively, and it is difficult to justify their positions as outliers, given the variables under study. On the other hand, the Huelva port authority, which is a consistent outlier throughout the four years analysed, features abnormal levels of both intermediate input prices and quasi-fixed input.

Finally, Figure 3 shows, grouping by cluster, the evolution in the EE for the entire sample of Spanish port authorities during the period of analysis, 1993, 1997, 2003, and 2007. From the results, it can be noted that port authorities as Algeciras, Vigo, Cartagena, and Málaga are leaders in their group, and reach the highest average levels of EE during the time horizon; these results are Algeciras (0.950), Vigo (0.937), Cartagena (0.938), and Málaga (0.933).

The port classification described above presents a number of desirable characteristics, especially when compared with similar articles published. The closest reference to our paper, both in scope and in methodology, is Medal-Bartual and Sala-Garrido (2011). Despite the apparent methodological similarities, there are actually some noticeable differences. The most relevant is probably the approach to and the way in which both papers address the efficiency estimation. Whereas they calculate their DEA efficiency frontier after splitting port authorities into clusters to take into account the “observable” heterogeneity, we have estimated only one frontier for the whole sample using an SFA model. This has let us take into account not only the “observable” but also any producer-specific unobserved heterogeneity.[15] Afterwards, we used the estimated parameters to cluster port authorities, which has let us to take into account all the relevant cost information.

Regarding the actual clustering methodology, Medal-Bartual and Sala-Garrido (2011) grouped their sample according to two predefined variables, while we perform a technology-based multi-dimensional clustering founded on both input and output factors, which ultimately serve as optimal variable weights.[16] Their resulting dendrogram features four high-level clusters that clearly

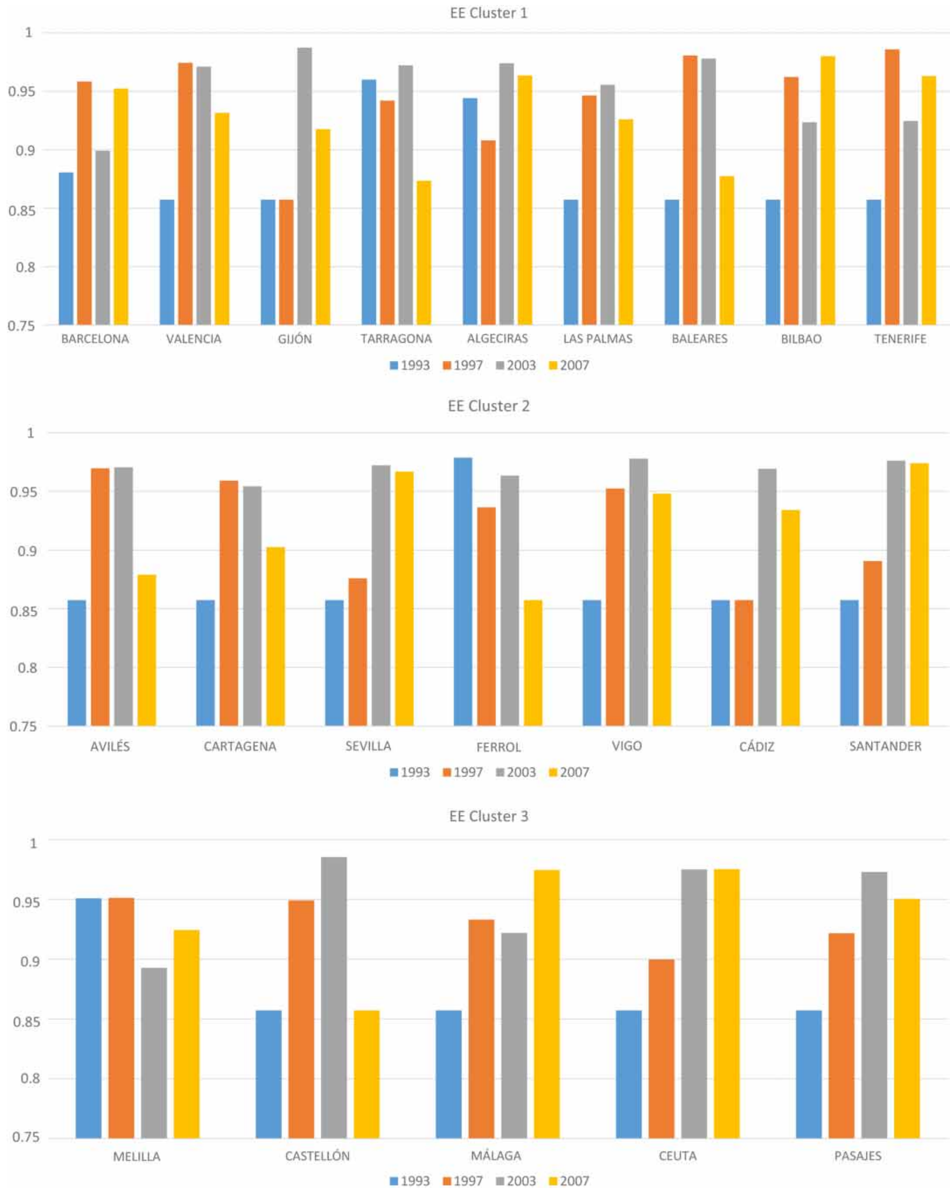

Figure 3. EE evolution by cluster.

correspond to specific traffic profiles, for example, solid bulk ports. Despite the difference in scale of operations among many Spanish port authorities, the size of ports cannot be readily detected even at lower levels of the hierarchy. In contrast, the complexity of Spanish port authorities is a key criterion in our cluster definition. While our classification tends to group port authorities according to their scale at a high level, similar port authorities[17] tend to be neighbouring inside each cluster. This results in a multi-levelled characterization of port authorities. This result, which better reflects the Spanish port authorities' actual situation, is a consequence of our hierarchical clustering; this in turn is heavily based on all the available information from an efficiency estimation, which takes unobservable heterogeneity into account.

Another work worthy of mention at this point, for comparative purposes, is Martinez-Budria et al. (1999). In common with the previous reference, they share a DEA approach and also the same way of conducting the efficiency estimation; however, they differ in the criteria followed to split the sample (see Section 2). They categorized their ports into three levels of complexity; however, their criteria are not clearly detailed, and may be regarded as an ad hoc approach.[18] This is in contrast to our efficiency-frontier-based clustering alternative. Nevertheless, their resulting classification (low-, medium-, and high-complexity ports) is very similar to that of our dendrogram, in the sense that port authorities with similar scale are grouped together. It also shows the existing relationship between size and complexity of port authorities. Their clusters resemble our own results in many cases; for example, their first cluster (high-complexity ports) almost mirrors our C1, port authority for port authority. However, the absence of the hierarchical component in their paper hinders a more detailed characterization of port authorities and limits the subsequent analysis.

In summary, Spanish port authorities have been clustered according to the relevant outputs and input prices, with the cost elasticities and factor shares serving as optimal variable weights. In this way, we have been able to generate dendrograms that define the port authorities categories that mirror the performance indicator. This information effectively adds another dimension to the comparative analysis, and helps in setting sharper targets for improvement specific to each cluster; this is even the case within each cluster, since we can take advantage of the possibility of the sub-cluster identification offered by the hierarchical approach. Nevertheless, there are outliers that cannot be easily or obviously classified into a specific cluster. In these particular cases, SOM are used to analyse their temporal response, in order to shed light on the reasons why they are outliers.

4.2. *Temporal Pattern Analysis*

The use of hierarchical methods makes temporal analysis difficult, and limits the clustering to a cross-section study; this is because different dendrograms cannot be straightforwardly compared. Some authors tackle this problem by clustering the average values of each port authority for the entire period; for example, Medal-Bartual and Sala-Garrido (2011). This is questionable and could be very misleading, as important changes that affect Spanish port authorities could be masked by the average figures. We propose the use of SOMs, in order to track the evolution of Spanish port authorities. Training a map will allow us to visually follow the history of ports that for whatever reason are of special interest; for example, being an outlier. Additionally, dendrograms are intended to provide a broad overview of the underlying structure of the data, but they are not directly interpretable in terms of the clustering variables. This problem can be easily addressed by projecting each variable, or component, on the SOM using Kohonen networks.

Our sample features a variety of port authority profiles, in terms of temporal evolution. Some port authorities do not seem to be affected by time changes, and thus show consistent behaviour. What is more, there are a number of outlier port authorities, which can be readily identified by inspecting the resulting dendrogram, not only for 2003 but for the whole 1993–2007 period. These observations are usually characterized by drastically reduced levels of cargo, either temporary or permanent, or by unusual traffic mixes. Moreover, port authorities may face difficulties, such as the recent economic downturn, which has temporarily affected their traffic volumes. Hence, we will focus on four port authorities to illustrate the utility of SOMs as a tool to follow their temporal evolution. They are Algeciras which is a relevant and consistent port, Huelva which is a consistent outlier, and Port Pasajes and Marín-Pontevedra which have suffered important variations in their traffic.

Figure 4 shows the U-matrix (Figure 4(a)) and component planes (Figure 4(b)–(h)) of Algeciras, Huelva, Marín-Pontevedra, and Pasajes port authorities. The network was trained[19] using the complete sample; that is, 1993, 1997, 2003, and 2007. Once the SOM has been generated, the individual response, the BMU, for each port authority at a given period is stored. The result is shown as a four-step trajectory. Each port authority trajectory has been assigned a different colour to facilitate visual inspection; see the legend in the bottom right of Figure 4.

The U-matrix is intended to reveal the clustering structure of the data. Map units with high values, the brighter colours in the figure, indicate important differences among nearby units, while homogeneity is represented by darker tones. We

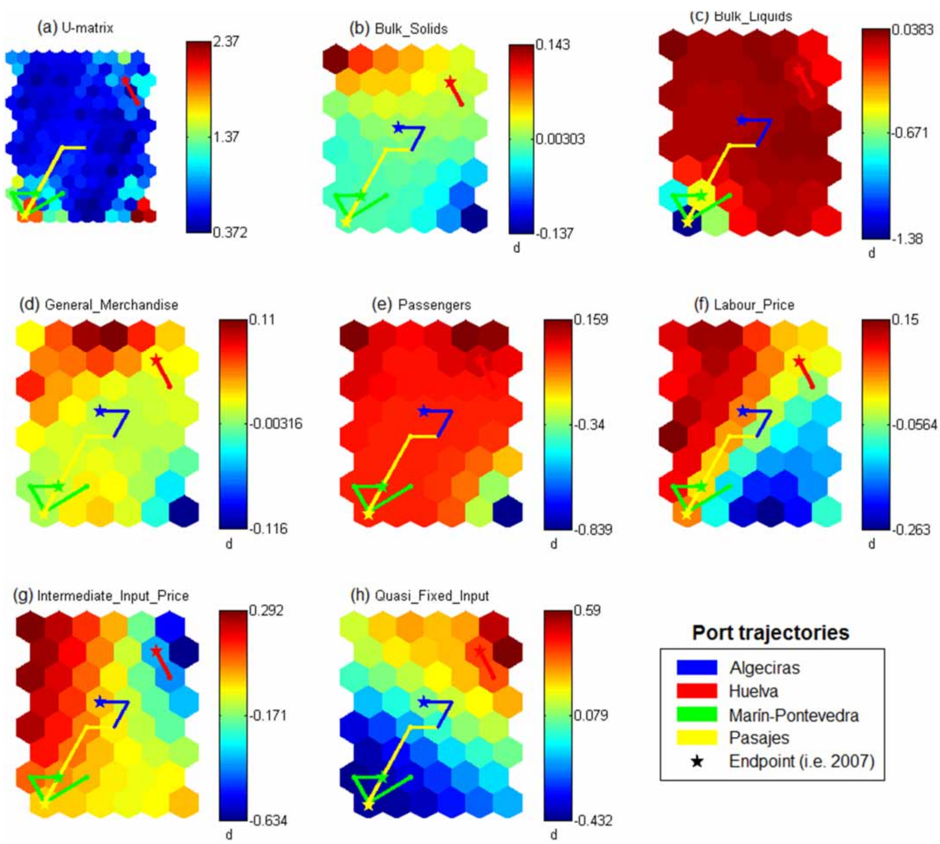

Figure 4. U-matrix, component planes and trajectories for selected ports (1993–2007).

have noticed three well-differentiated areas in our U-matrix. The first is a large, homogeneous dark-blue area of the map, which includes the majority of port authorities of the sample, for example, Algeciras and Las Palmas. The second is a relatively small piece of the map, which is in the bottom left corner, and whose feature is very low levels of liquid bulks; for example, Marín-Pontevedra port authority. Third is another small portion of the map, in the bottom right corner, which cannot be readily related to a single component.

We will focus, however, on the temporal evolution of the selected port authorities, which can be assessed through the component planes at an attribute level. Each component plane represents the value of a given variable on each map unit; in Figure 4(b)–(h), there are colour bars indicating the levels of such variables. For each port authority, its consecutive BMUs are connected by a line, or trajectory, and are superimposed onto the component maps; thus, their progress in time is represented, and this allows for a ready visual analysis. As a rule of thumb, the pronounced colour variations of BMUs within the same trajectory are normally associated with large variations in the levels of components. The Pasajes port authority perfectly illustrates the case of a sudden decline in traffic, and its effects on classification; in 2005, after years of a slow but consistent decline, its bulk liquid traffic reduced to zero. In this case, what was a consistent port from 1993 to 2003, subsequently in 2007 became an outlier, joining Marín-Pontevedra port authority in the uppermost side of the corresponding dendrogram.

In the case of Pasajes, this declining trend can be immediately detected by inspecting its trajectory for the liquid bulk component; see the yellow line in Figure 4(c). This corroborates its evolution from the “safe” region of the map to the area associated with extreme, low values. Figure 5 presents a three-dimensional representation of this particular trajectory and component. Moreover, a very similar pattern was observed for Marín-Pontevedra port authority in 1997, when their liquid bulk traffic stopped. This occurred following its strategy to focus on general merchandise[20] and solid bulks, and because of competition from A Coruña and Ferrol. All these ports are located in Galicia.

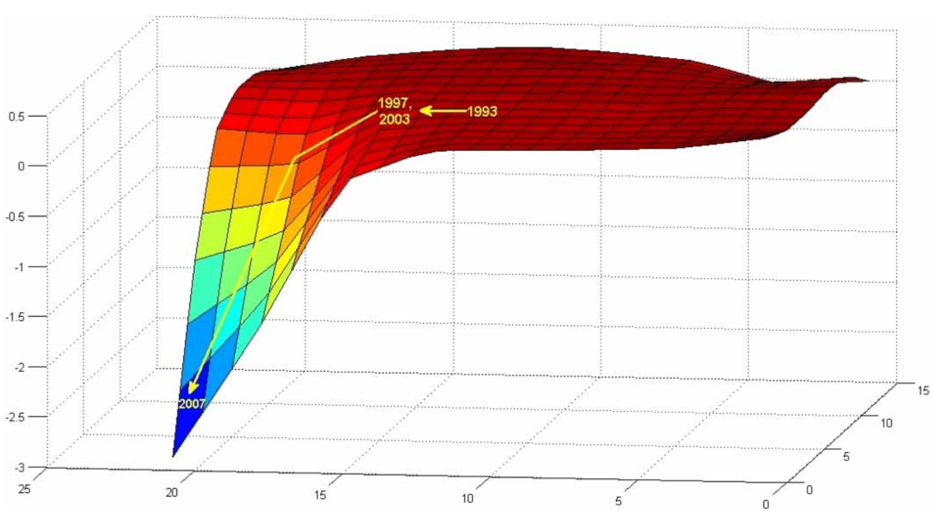

Figure 5. Evolution of Pasajes port between 1993 and 2007. Liquid bulk plane.

The case of the port of Huelva is particularly interesting. This port authority is a consistent outlier in each one of the periods considered, with apparently low levels of intermediate prices and high quasi-fixed input. Note that the outlier status of Huelva was not clearly represented by the U-matrix; hence, this shows the importance of analysing the component planes for each port.

5. Conclusions

In this paper, we review previous studies that combine both frontier-based productivity and/or efficiency estimation with port classification. The problem of distinguishing between heterogeneity and inefficiency is widely acknowledged in benchmarking, and aggravated when international data sets are used. Due to ports being characterized by their geographical and operational settings, this environment heterogeneity — which is one characteristic manifestation of firm heterogeneity — should be taken into account to avoid considering as "inefficiency" what are really "differences". Our survey has shown that authors using DEA (the vast majority) have tried to solve this problem by splitting the sample into homogeneous groups before the frontier estimation. As the latter approach is not without disadvantages, we propose using a frontier-based clustering to classify ports following the methodological framework from Rodríguez-Déniz and Voltes-Dorta (2014).

In this paper, we have also contributed to the SFA literature on port efficiency measurement by presenting an empirical application of the above-mentioned methodology on the Spanish port authorities. They were classified by means of hierarchical clustering, with the cost elasticities and factor shares serving as optimal variable weights. We have used the cost frontier parameters and efficiency estimates from Rodríguez-Álvarez and Tovar (2012), who developed a stochastic frontier model that controls unobserved heterogeneity. We defined port authorities' categories that mirror the performance indicator for the years 1993, 1997, 2003, and 2007 by using hierarchical clustering. We also proposed the use of Kohonen's SOMs, in order to track the temporal evolution of special port authorities, for example, outliers.

From the analysis of the structure and efficiency of Spanish port authorities, we can conclude that there is no homogeneous set of individuals, but a number of well-defined groups with similar properties; these depend on the level of complexity, scale and mix of production, in the corresponding Spanish port authorities. Specifically, the cluster analysis applied suggests the existence of three clusters of port authorities. The first one is compounded of the main commercial ports, which have greater complexity. These are characterized by large scales of production. There is a predominance of general cargo, although their port authorities have all types of cargo, including passenger traffic. The second cluster groups medium-complexity port authorities, with medium scales of production, whose traffic is based on both solid and liquid bulk traffic; nonetheless, they also have general cargo and, to lesser extent, passenger traffic. Finally, the last port cluster is characterized by a low scale of production, and ports are mainly gateways for their natural hinterland. The classification also allows for fine-grained comparisons inside each cluster. Although the main divisions of the dendrogram are defined according to the scale of the port authorities, similar elements tend to form sub-clusters at lower levels, in accordance with their traffic mixes or quasi-fixed input; this is the case for the top solid bulk ports Gijón and Tarragona in C1, and for Avilés and Cartagena in C2. Since the defined categories of port authorities mirror the performance indicator, another dimension is effectively added to the comparative analysis; and this helps in setting sharper targets for improvement that are specific to each cluster. We also analysed the relative efficiency of each Spanish port authority, with respect to the others belonging to the same cluster. From the results, it is noted that ports such as Algeciras, Vigo, Cartagena, and Málaga are leaders in their group, and reach the highest average levels of EE during the time horizon, 1993–2007.

Given that different dendrograms for each period are not straightforward to compare, we completed our cluster analysis using SOMs, and this represents one original contribution of this article. In this paper, we have proposed the use of SOMs to track the temporal evolution of Spanish port authorities. The resulting maps have allowed us to visually follow the history of port authorities, which are of special interest for several reasons, and shed light on the origin of such patterns. We have selected four port authorities' profiles to be tracked. Algeciras represents those port authorities that behave in a consistent way during the whole period. A second, Huelva, shows the case of a port authority which behaves as an outlier all the time. Finally, we use Pasajes and Marin-Pontevedra to illustrate the case of a port authority changing from being consistent to becoming outlier and vice versa. The SOMs were particularly helpful in visualizing the temporal patterns at each variable level, allowing, us to identify the cause of some outliers.

In summary, we have reviewed the previous literature on classification for port benchmarking and also advanced the literature by classifying port authorities using a methodology that combines SFA and cluster analysis. We go further than ad hoc approaches and also avoid making misleading comparisons, by taking into account unobserved heterogeneity in the SFA estimation. The use of hierarchical clustering allows for a more detailed analysis and identification of the efficiency benchmarks. This is in contrast to other existing alternatives that produce flat clusters. Moreover, the weighting of the relevant variables has not been considered in any previous study. In addition, the weights applied to the clustering variables come from a model that explicitly takes into account the existence of unobserved heterogeneity. We conclude that combining cost frontier and cluster methods into cost-frontier-based clustering methods and the use of SOMs, either in isolation or jointly, offer particularly useful information for administrators, regulators, and policy-makers. Cost-frontier-based clustering may be used in when defining a robust port typology, and SOM for analysing the temporal response of particular cases. For this reason, we believe that these methods could be useful for future studies on efficiency and/or productivity in ports.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. Formally, technical efficiency is the capacity of obtaining the maximum amount of output from certain inputs (output orientation). Alternatively, as the capacity of obtaining a given output level using the minimum amount of inputs (input orientation). Also, a company presents efficiencies of scale, if it reaches the maximum productivity with the current technology.

2. One of the advantages of DEA is that the researcher can generate results with relatively small data sets. As a result, DEA has been used extensively in the economic analysis of port efficiency.
3. Note that in our paper, the SOMs are employed to analyse the temporal evolution of ports, and not to produce actual clusters, as Sharma and Yu (2009) and Quaresma-Dias et al. (2009) did. The aim of their application of SOMs was to generate an input-based clustering, which was subsequently combined with a set of efficiency tiers obtained from the DEA.
4. The optimal variable weights which are going to be used for building port clusters come from an SFA model which took unobserved heterogeneity into account.
5. For the sake of brevity, we only summarize the details needed to understand what are we doing here and refer the interested readers to the original paper.
6. They began by following the model proposed by Battese and Coelli (1995), which allows them to specify economic inefficiency in terms of a set of explicative variables that may change with time. This model did not need to recur to second-stage analysis, thus avoiding inconsistency problems; see Wang and Schmidt (2002). However, they proposed the re-estimation of the model within a Fixed Effect Model (FEM) framework, by introducing port authorities dummy variables into the frontier equation; this was done in order to capture possible systematic differences between ports (unobservable heterogeneity), If this heterogeneity exists and it is not explicitly picked up in the model, then a problem of omitted variables exists; consequently, the estimated coefficients of the included variables will be biased. In this way, the FEM model nests the previous pooled model and, on the basis of likelihood ratio tests, the restricted model was rejected; thus, the FEM was found to be a better representation of the technology for the sample. The immediate implication is that a model, which does not account for individual effects, would be misspecified, and therefore provides biased parameter estimates and misleading inference.
7. The estimated cost function fulfils the properties required by the theory; the regularity conditions are satisfied as the outputs are increasing, and the input prices are non-decreasing and quasi-concave.
8. Everitt, Landau, Leese, and Stahl (2011) is a general reference to cluster analysis. Xu and Wunsch (2005) present a comprehensive survey of algorithms for data clustering.
9. Here, we only present their methodology briefly, and we refer the interested readers to the original paper.
10. By using iTOL (Letunic & Bork, 2007), we generated the dendrograms with their branches labelled according to their height.
11. In unsupervised learning, the goal is to describe the associations and patterns among a set of unlabelled samples (Duda, Hart, & Stork, 2000; Hastie, Tibshirani, & Friedman, 2009).
12. On a two-dimensional squared lattice, the Von Neumann neighbourhood comprises the four nodes orthogonally surrounding a central node (Schiff, 2008).
13. To show the advantages of the methodology of combining SFA and hierarchical clustering, we only need to analyse a single year from within the sample. Since we get very similar port cluster for the selected years, we only report results for 2003; moreover, in that year, the Spanish port authorities showed the highest average economic efficiency for the 1993–2007 period.
14. The port authorities clusters mirror the performance indicator because the optimal variable weights used to build them come from a stochastic frontier cost efficiency model.
15. Unobserved heterogeneity is not reflected in measured variables, but it is expressed in the form of effects (Greene, 1993).
16. Although Medal-Bartual & Sala-Garrido classified ports according to multiple characteristics, no variable weights were assigned.
17. For example, Barcelona and Valencia, two out of the three biggest general cargo ports, were merged into the same C1 sub-cluster.
18. Martinez-Budria et al. state that they use the port classification provided by the former General Management of Ports, but the only explanation they offer is "is based on a complexity criterium given by port size and the composition of the output vector". They offer no reference on where more information about this port classification may be found.
19. Self-Organized Maps were generated using the SOM Toolbox for Matlab: http://www.cis.hut.fi/projects/somtoolbox/.
20. In 2011, Marín-Pontevedra overtook A Coruña as general merchandise port.

References

Abonyi, J., & Feil, B. (2007). *Cluster analysis for data mining and system identification*. Basel: Birkhäuser.
Battese, G. E., & Coelli, T. J., (1995). Model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, *20*, 325–332.
Bichou, K. (2013). An empirical study of the impacts of operating and market conditions on container-port efficiency and benchmarking. *Research in Transportation Economics*, *42*, 28–37.
Chang, V., & Tovar, B. (2014). Efficiency and productivity changes for Peruvian and Chilean ports terminals: A parametric distance functions approach. *Transport Policy*, *31*, 83–94.
Cheon, S. (2009). Impact of global terminal operators on port efficiency: A tiered data envelopment analysis approach. *International Journal of Logistics: Research and applications*, *12*(2), 85–101.
Cheon, S., Dowall, D., & Song, D-W. (2010). Evaluating impacts of institutional reforms on port efficiency changes: Ownership, corporate structure, and total factor productivity changes of world container ports. *Transportation Research Part E*, *46*, 546–561.
Cullinane, K., & Wang, T. (2010). The efficiency analysis of container port production using DEA panel data approaches. *OR Spectrum*, *32*(3), 717–738.
Cullinane, K. P. B., & Song, D. W. (2006). Estimating the relative efficiency of European container ports: A stochastic frontier analysis. In K. P. B. Cullinane, & W. K. Talley (Eds.), *Port economics. Research in transport economics* (pp. 85–115). Amsterdam: Elsevier.
Duda, R. O., Hart, P. E., & Stork, D. G. (2000). *Pattern classification*. New York, NY: Wiley-Interscience.
Everitt, B., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Chichester: Wiley.
Greene, W. H., (1993). The econometric approach to efficiency analysis. In Hal O. Fried, C. A. Knox Lovell, & Shelton S. Schmidt (Eds.), *The measurement of productive efficiency* (pp. 68–119). New York: Oxford University Press.
Guironnet, J-P., Peypoch, N., & Solonandrasana, B. (2009). A note on productivity change in French and Italian seaports. *International Journal Shipping and Transport Logistics*, *1*(3), 216–226.
Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. New York, NY: Springer-Verlag.
Hung, S.-W., Lu, W.-M., & Wang, T.-P. (2010). Benchmarking the operating efficiency of Asia container ports. *European Journal of Operational Research*, *203*, 706–713.
Jessop, A. (2012). A decision aid for finding performance groups. *Benchmarking: An international journal*, *19*(3), 325–339.
Kaisar, E., Pathomsiri, S., & Haghani, A. (2006). *Efficiency measurement of US ports using data envelopment analysis*. National urban freight conference, Long Beach, CA, 1–16.
Kohonen, T. (1982). Self-Organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59–69.
Koster, M. B. M., Balk, B. M., & van Nus, W. T. I. (2009). On using DEA for benchmarking container terminals. *International Journal of Operations & Production Management*, *29*(11), 1140–1155.
Letunic, I., & Bork, P., 2007. Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics*, *23*(1), 127–128. Retrieved from http://itol.embl.de/
Martín-del-Brío, B., & Serrano-Cinca, C. (1993). Self-organizing neural networks for the analysis and representation of data: Some financial cases. *Neural Computing & Applications*, *1*(2), 193–206.
Martinez-Budria, E., Diaz-Armas, R., Navarro-Ibanez, M., & Ravelo-Mesa, T. (1999). A study of the efficiency of Spanish port authorities using data envelopment analysis. *International Journal of Transport Economics*, *XXVI*(2), 237–253.
Medal-Bartual, M. A., & Sala-Garrido, R. (2011). Análisis de la Eficiencia del Sistema Portuario Español: estructura, evolución y perspectivas. Edita: Fundacion VALENCIAPORT. Serie Planificción y gestión portuaria. España.
Milligan, G., & Cooper, M. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrica*, *50*(2), 159–179.
Notteboom, T., Coeck, C., & Van Den Broeck, J. (2000). Measuring and explaining the relative efficiency of container terminals by means of Bayesian stochastic frontier models. *Maritime Economics & Logistics*, *2*, 83–106.
Pallis, A., Vitsounis, T., De Langen, P., & Notteboom, T. (2011). Port economics, policy and management: Content classification and survey. *Transport Reviews*, *31*, 445–471.
Panayides, P., Maxoulis, C., Wang, T-F., & Koi Yu, A. (2009). A critical analysis of DEA application to seaport economic efficiency measurement. *Transport Reviews*, *29*(2), 183–206.
Quaresma Dias, J. C., Garrido Azevedo, S., & Ferreira, J. (2009). A comparative benchmarking analysis of main Iberian container terminals: A DEA approach. *International Journal of Shipping and Transport Logistics*, *1*(3), 260–275.

Rodríguez-Álvarez, A., & Tovar, B. (2012). Have Spanish port reforms during the last two decades been successful? A cost frontier approach. *Transport Policy*, *24*, 73–82.

Rodríguez-Déniz, H., & Voltes-Dorta, A. (2014). A frontier-based hierarchical clustering for airport efficiency benchmarking. *Benchmarking: An International Journal*, *21*(4), 486–508.

Sarlin, P. (2013). Self-organizing time map: An abstraction of temporal multivariate patterns. *Neurocomputing*, *99*, 496–508.

Schiff, J. L. (2008). *Cellular automata: A discrete view of the world*. Hoboken, NJ: Wiley.

Sharma, J., & Yu, S. (2009). Performance based stratification and clustering for benchmarking of container terminals. *Expert Systems with Applications*, *36*, 5016–5022.

Simar, L., & Wilson, P. (2000). A general methodology for bootstrapping in nonparametric frontier models. *Journal of applied Statistics*, *27*(6), 779–802.

Simoes, P., & Marques, R. C. (2010). Influence of congestion efficiency on the European seaports performance: Does it matter? *Transport Reviews*, *30*, 517–539.

Sokal, R., & Rohlf, F. (1962). The comparison of dendrograms by objective methods. *Taxon*, *11*, 33–40.

Tovar, B., & Wall, A. (2015). Can ports increase traffic while reducing inputs? Technical efficiency of Spanish port authorities using a directional distance function approach. *Transport Research Part A*, *71*, 128–144.

Ultsch, A., & Siemon, H. P. (1990, July 1990). Kohonen's self organizing feature maps for exploratory data analysis. In B. Widrow, & B. Angeniol (Eds.), *Proceedings of the international neural network conference (INNC-90). Paris, France 1* (pp. 305–308). Dordrecht: Kluwer.

Wang, H., & Schmidt, P. (2002). One-Step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis*, *18*(2), 129–144.

Wanke, P., Barbastefano, R., & Hijar, M. (2011). Determinants of efficiency at major Brazilian port terminals. *Transport Reviews*, *31*, 653–677.

Wilmsmeier, G., Tovar, B., & Sanchez, R. J. (2013). The evolution of container terminal productivity and efficiency under changing economic environments. *Research in Transportation Business & Management*, *8*, 50–66.

Woo, S., Pettit, S., Beresford, A., & Kwak, D. (2012). Seaport research: A decadal analysis of trends and themes since the 1980s. *Transport Reviews*, *32*, 351–377.

Wu, Y-C., & Goh, M., (2010). Container port efficiency in emerging and more advance markets. *Transportation Research Part E*, *46*, 1030–1042.

Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, *16*(3), 645–678.