Annals of Operations Research (2022) 318:713–741
https://doi.org/10.1007/s10479-022-04659-7

ORIGINAL - SURVEY OR EXPOSITION

The origins and development of statistical approaches in non-parametric frontier models: a survey of the first two decades of scholarly literature (1998–2020)

Amir Moradi-Motlagh[1] · **Ali Emrouznejad**[2]

Accepted: 7 March 2022 / Published online: 11 May 2022

Abstract
This paper surveys the increasing use of statistical approaches in non-parametric efficiency studies. Data Envelopment Analysis (DEA) and Free Disposable Hull (FDH) are recognized as standard non-parametric methods developed in the field of operations research. Kneip et al. (Econom Theory, 14:783–793, 1998) and Park et al. (Econom Theory, 16:855–877, 2000) develop statistical properties of the variable returns-to-scale (VRS) version of DEA estimators and FDH estimators, respectively. Simar & Wilson (Manag Sci 44, 49–61, 1998) show that conventional bootstrap methods cannot provide valid inference in the context of DEA or FDH estimators and introduce a smoothed bootstrap for use with DEA or FDH efficiency estimators. By doing so, they address the main drawback of non-parametric models as being deterministic and without a statistical interpretation. Since then, many articles have applied this innovative approach to examine efficiency and productivity in various fields while providing confidence interval estimates to gauge uncertainty. Despite this increasing research attention and significant theoretical and methodological developments in its first two decades, a specific and comprehensive bibliometric analysis of bootstrap DEA/FDH literature and subsequent statistical approaches is still missing. This paper thus, aims to provide an extensive overview of the key articles and their impact in the field. Specifically, in addition to some summary statistics such as citations, the most influential academic journals and authorship network analysis, we review the methodological developments as well as the pertinent software applications.

Keywords Data envelopment analysis · Bootstrap · Efficiency · Survey

✉ Ali Emrouznejad
a.emrouznejad@surrey.ac.uk
http://www.deazone.com

Amir Moradi-Motlagh
amoradi@swin.edu.au

[1] Swinburne University of Technology, Mail H23, PO Box 218, Hawthorn, VIC 3122, Australia

[2] Surrey Business School, University of Surrey, Guilford, UK

1 Introduction

Data Envelopment Analysis (DEA) is a linear programming technique measuring the relative efficiency of decision making units (DMUs), as introduced by Charnes et al. (1978). While Charnes and colleagues' article remains the most cited paper in the field of operations research (Laengle et al., 2017), there have been significant developments in both theory and applications of this non-parametric technique (Emrouznejad & Yang, 2018). Free Disposal Hull (FDH), introduced by Deprins et al. (2006), is the second most used non-parametric method which requires mixed integer programming formulation and relaxes the convexity assumption of DEA. The popularity of non-parametric techniques such as DEA and FDH is due to their advantage of not requiring any assumption on the functional form of the production frontier. However, conventional non-parametric methods are criticised as being deterministic, meaning that no noise is considered and all deviations from the frontier are assumed as inefficiency. Simar and Wilson (1998) propose a smooth bootstrap to deal with boundary issues arising from the "deterministic" nature of DEA estimators. Subsequently, it can be argued non-parametric approaches have a statistical basis and hence one of their main criticisms is no longer valid.

The last decade has seen exponential growth in the use of bootstrapping approach in studies of efficiency and productivity, with the number of citations to the bootstrap methods of Simar and Wilson (1998, 2000a, 2007) exceeding 7,200[1] (e.g., Agrell et al., 2020; Alberta Oliveira & Santos, 2005; Andersson et al., 2017; Boame, 2004; Brümmer, 2001; Davidova & Latruffe, 2007; De Witte & Marques, 2010; Eling & Luhnen, 2010; Essid et al., 2014; Fukuyama & Tan, 2021; Galariotis et al., 2021; Halkos & Tzeremes, 2013; Hawdon, 2003; Johnes, 2006; Kohl et al., 2019; Lothgren & Tambour, 1999; Merkert & Hensher, 2011; Moradi-Motlagh & Babacan, 2015; Rossi & Ruzzier, 2000; Salim et al., 2016; Tiemann & Schreyögg, 2009, 2012; Tortosa-Ausina, 2002; Worthington & Lee, 2008; Yang & Zhang, 2018). Additionally, more recent theoretical advancements in statistical inferences in the non-parametric context, necessitate survey research and analysis of bibliometric records, tracing the intellectual progression of the field and its applications. Bibliometric studies provide theoreticians and researchers with practical information about bibliometric indicators to enhance future methodological developments and applied research (Laengle et al., 2017; Liu et al., 2013). On the one hand, theoreticians can identify which models are employed and how frequently they are used. They can also learn from successful applications of methodological advancements of the techniques. Further, on the other hand, applied researchers will be informed about key articles, influential researchers, trend and research directions in the field (Liu et al., 2013).

Most previous DEA/FDH literature reviews are solely methodological or surveys on applications (e.g., Hatami-Marbini et al., 2011; Kao, 2014; Simar & Wilson, 2015; Witte & López-Torres, 2017). There are also several bibliometric analyses, as the following examples show. Lampe and Hilgers (2015) provide a bibliometric analysis of DEA and Stochastic Frontier Analysis (SFA) and discuss the methodological trends with each technique. Emrouznejad and Yang (2018) report a broad list of DEA related articles between 1978 and 2016 and provide some summary statistics including the most utilised journals, authorship, and keyword analysis.

Olesen and Petersen (2016) provide a review of stochastic DEA. Stochastic DEA is defined as "an efficiency analysis using non-parametric convex hull/convex cone reference technologies based on either statistical axioms or distributional assumptions that allow for a random

[1] The number of citations is based on Google Scholar, visited on 14th February 2022.

(estimator of the) reference technology" (Olesen & Petersen, 2016, p.3). They identify three main directions for stochastic DEA in the literature. First, methods which regard given input and output variables as a sample from a large population. Simar and Wilson (1998) approach is grouped in this category. Second, methods which are able to handle random noise like SFA. Banker and Maindiratta (1992) as the pioneer of this direction show how to interpret DEA residuals derived based on maximum likelihood function. Third, methods which are based on having information on random disturbances in the form of knowledge of distributions involved. Chance Constrained DEA is an example of such an approach (see Cooper et al., 1998; Olesen & Petersen, 1995). Among these three directions, our study only focuses on the first one as the most popular statistical approach in the literature.

Despite the recent prevalence of using statistical approaches in non-parametric efficiency studies, no previous study has focused on analysing a bibliography of this growing field of research. We address this gap by reviewing the statistical approaches introduced in non-parametric frontier analysis and identifying the most influential methodological papers over the past two decades. We then provide a bibliometric analysis for the most cited papers, including summary statistics of their citations, leading academic journals, keywords, and authorship network analysis.

The remainder of this paper is as follows. In the next section we briefly review the statistical approaches in the context of non-parametric models. Section 3 describes data and methodology used for bibliometric analysis. Section 4 identifies the most influential papers based on citations. Section 5 provides a bibliometric analysis of the leading methodological papers, Simar and Wilson (1998) and Simar and Wilson (2007). Section 6 reviews available software applications. Section 7 provides a conclusion.

2 Non-parametric frontier models and statistical approaches

The popularity of non-parametric techniques such as DEA and FDH is due to their advantage of not requiring parametric assumptions on the functional form of the production frontier as well as their ability to accommodate both multiple inputs and multiple outputs. However, conventional non-parametric methods were initially criticised as (i) being deterministic (meaning that no noise is considered, and all deviations from the frontier are assumed to result from inefficiency), and hence sensitive to outliers; and (ii) lacking any statistical interpretation. With regard to (i), a number of methods for detecting outliers that might distort DEA or FDH efficiency measurements have been developed, e.g., Wilson (1993, 1995), Simar (2003), and Porembski et al. (2005). Researchers can use these methods to detect outliers and then decide what to do. As discussed by Wilson (1993), an "outlier" is an atypical observation. When a researcher finds an atypical observation, more work is needed to determine whether the observation results from an error. Furthermore, Cazals et al. (2002) and Aragon et al. (2005) propose probabilistic approaches which provide robust estimators in the presence of outliers. With regard to (ii), Kneip et al. (1998) give a statistical model and use the assumptions of the model to establish statistical consistency and the rate of convergence of the variable returns-to-scale (VRS) version of DEA estimators. Meanwhile, Simar and Wilson (1998) show that conventional, "naive" bootstrap methods cannot provide valid inference when used with DEA or FDH estimators, and propose a smooth bootstrap to deal with boundary issues arising from the "deterministic" nature of DEA estimators.

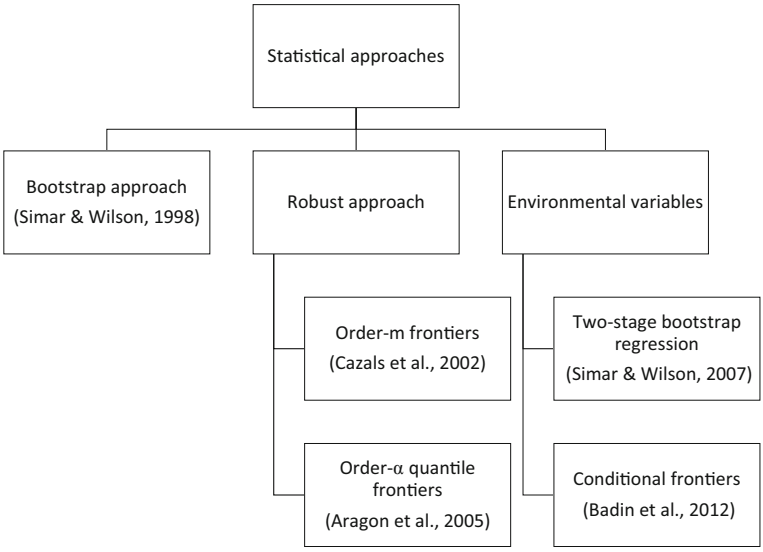

Fig. 1 Statistical approaches in non-parametric efficiency analysis

Another large strand of the efficiency literature involves estimation of efficiency, and then regression of the efficiency estimates on some additional, environmental variables in a second stage. As of 14 February 2022, a search on Google Scholar using the keywords "efficiency," "regression" and "second stage" found approximately 229,000 hits. These studies use a variety of parametric (and occasionally, nonparametric) models for the second-stage regressions. Figure 1 illustrates the main statistical approaches in examining efficiency using non-parametric models since 1998.

In this section, we focus on DEA as most of efficiency studies employ this non-parametric method. Radial[2] DEA efficiency scores can be obtained using input or output oriented approaches, as well as the VRS assumption. Let X be a set of p inputs and Y be a set of q outputs. The efficiency score of a DMU operating at level (x, y), in an input-oriented[3] framework with the assumption of variable returns to scale and can be estimated using

$$\widehat{\theta}(x, y) = \min\left\{\theta > 0 | y \leq \sum_{i=1}^{n} \lambda_i Y_i, \theta x \geq \sum_{i=1}^{n} \lambda_i X_i, \sum_{i=1}^{n} \lambda_i = 1,\ \lambda_i \geq 0\ \forall i = 1, \ldots, n\right\}, \tag{1}$$

where $\widehat{\theta}(x, y)$ is the estimation of the true unknown efficiency score of $\theta(x, y)$, λ refers to the vector of constants, and i is the number of DMUs. We use (1) and provide a summary of the methods illustrated in Fig. 1 in the following sections.

[2] In radial efficiency models, for any given DMU, it is assumed that inputs (outputs) can be proportionately decreased (increased) without altering the output (input) quantities to reach the frontier.

[3] To save space only an input-oriented approach is used for all models presented in this article.

2.1 Bootstrap DEA and FDH

It is a tautology that statistical inference is impossible without a statistical model and some knowledge of the distribution of an estimator of the feature about which one wishes to make inference. The limiting distribution of the FDH estimator was unknown until the work of Park et al. (2000), and the limiting distribution of the VRS-DEA estimator was unknown until it was established by Kneip et al. (2008). Simar (1996) emphasise on the importance of a valid bootstrap data generating process in statistical analysis of DEA models. Simar and Wilson (1998) proposed estimating the distributions of nonparametric, VRS-DEA efficiency estimators using a smooth bootstrap method, thereby allowing inference about inefficiency. The idea was extended in Simar and Wilson (2000a), where the distribution of efficiency was allowed to vary throughout the production set. Neither Simar and Wilson (1998) nor Simar and Wilson (2000a) offer proofs of validity of their suggested bootstrap methods since the limiting distribution of the VRS-DEA estimator remained unknown until eight years later. Nonetheless, simulation results provided in both papers seem to suggest that the proposed methods "work," at least within the context of the simulations of the papers.

Specifically, Simar and Wilson (1998) employ the bootstrap technique introduced by Efron (1992) to measure the sensitivity of efficiency scores of DEA/FDH models to the sampling variation. The homogeneous bootstrap procedure suggested by Simar and Wilson (1998), to estimate the bias corrected efficiency score of a given DMU, is presented as follows:

[1] Compute the technical efficiency using (1) for all DMUs to generate $\widehat{\theta}_1, \ldots, \widehat{\theta}_n$.

[2] Repeat the following five steps B times (B is a large number, say 2000) to provide a set of estimates $\left\{\widehat{\theta}^*_b(x, y), b = 1, \ldots, B\right\}$ for a given DMU operating at level (x, y).

[2-1] Draw with replacement from $\widehat{\theta}_1, \ldots, \widehat{\theta}_n$ to generate $\beta^*_{1,b}, \ldots, \beta^*_{n,b}$.

[2-2] Smooth the sampled values using:

$$\tilde{\theta}^*_{i,b} = \begin{cases} \beta^*_{i,b} + h\varepsilon^*_{i,b} & \text{if } \beta^*_{i,b} + h\varepsilon^*_{i,b} \leq 1 \\ 2 - \beta^*_{i,b} - h\varepsilon^*_{i,b} & \text{otherwise} \end{cases} \tag{2}$$

where h is the bandwidth and $\varepsilon^*_{i,b}$ is a random error drawn from the standard normal distribution. The bandwidth can be estimated using the likelihood cross validation method suggested by Daraio and Simar (2007a).

[2-3] Correct the variance of the generated bootstrap by

$$\theta^*_{i,b} = \overline{\beta}^*_b + \frac{\tilde{\theta}^*_{i,b} - \overline{\beta}^*_b}{\sqrt{1 + h^2/\widehat{\sigma}^2_\theta}}, \tag{3}$$

where $\overline{\beta}^*_b = \sum_{i=1}^{n} \beta^*_{i,b}/n$ and $\widehat{\sigma}^2_\theta$ refers to the sample variance of $\widehat{\theta}_1, \ldots, \widehat{\theta}_n$.

[2-4] Compute pseudo-data set $\left\{\left(X^*_{i,b}, Y^*_{i,b}\right), i = 1, \ldots, n\right\}$ given by $X^*_{i,b} = \frac{\widehat{\theta}_i}{\theta^*_{i,b}} \times X_i$ and $Y^*_{i,b} = Y_i$.

[2-5] Calculate the bootstrap estimate of $\widehat{\theta}^*_b(x, y)$ by solving

$$\widehat{\theta}^*_b(x, y) = \min\left\{\theta > 0 |\ y \leq \sum_{i=1}^{n} \lambda_i Y^*_{i,b},\ \theta x \geq \sum_{i=1}^{n} \lambda_i X^*_{i,b},\right.$$
$$\left.\sum_{i=1}^{n} \lambda_i = 1,\ \lambda_i \geq 0\ \ \forall i = 1, \ldots,\ n\right\} \tag{4}$$

[3] Calculate the bias corrected efficiency score using

$$\widehat{\widehat{\theta}}(\mathrm{x}, \mathrm{y}) = 2\widehat{\theta}(x, y) - \frac{1}{B}\sum\nolimits_{b=1}^{B} \widehat{\theta}_b^*(x, y)$$

Then as explained in detail by Simar and Wilson (2000b), bootstrap estimates are utilised to construct the confidence intervals. Simar and Wilson (1999) further extend the bootstrapping idea to examine whether Malmquist indices of productivity are significant in a statistical sense. In addition, Simar and Wilson (2000a) propose a general bootstrapping approach which allows for heterogeneity in the structure of efficiency. The double-smooth and subsampling bootstrap methods described by Kneip et al. (2008) were the first to be proved to provide asymptotically valid inference about inefficiency estimated by VRS-DEA estimators.

As discussed by Olesen and Petersen (2016), although the homogeneous bootstrap works well with relatively few observations in a moderate input output dimension, the assumption of identical random inefficiency distribution discredits the efficiency results in the eyes of DMUs evaluated. For example, the DMUs that are doing things differently (outliers) may not be presented in a fair way in such performance evaluations due to the focus on general tendency in statistical approaches (see Olesen and Petersen (2016) for detailed discussions).

Simar and Wilson (2011a) suggest a subsampling bootstrap approach in nonparametric models which does not require multivariate kernel smoothing. The subsampling approach is easier to implement and only requires drawing bootstrap pseudo-sample of size m out of n (the original sample size). Simar and Wilson (2011a) also provide a data-based algorithm for selecting m. The m out of n bootstrap method, to estimate the confidence interval for (x, y), with p inputs and q outputs is presented as follows:

[1] Draw m samples out of n ($m < n$).
[2] Compute the technical efficiency based on m samples using (1).
[3] Repat [1] and [2] B times to generate $\widehat{\theta}_1^*, \ldots, \widehat{\theta}_B^*$
[4] Construct a $(1-\alpha)$ confidence interval for θ using

$$\left[\widehat{\theta}\left(1 + n^{-\frac{2}{p+q+1}}\varphi_{\frac{\alpha}{2},m}\right), \widehat{\theta}\left(1 + n^{-\frac{2}{p+q+1}}\varphi_{1-\frac{\alpha}{2},m}\right)\right]$$

where $\varphi_{\alpha,m}$ is the α-quantile of the bootstrap distribution $m^{\frac{2}{p+q+1}}\left(\frac{\widehat{\theta}}{\widehat{\theta}^*} - 1\right)$.

The latent variable problem discussed by Simar and Wilson (2007) extends to other situations where researchers might want to make inference. For example, the true, underlying inefficiencies were observed, making inference about expected (or mean) inefficiency would be simple, relying on standard results such as the Lindeberg-Feller Central Limit Theorem (CLT). However, true inefficiency is not observed and must be estimated. Several applied studies have used inefficiency estimates obtained from DEA or FDH estimators to estimate confidence intervals for inefficiency using standard CLT results. However, Kneip et al. (2015) show that when using sample means of inefficiencies estimated by DEA or FDH estimators, the usual CLTs are valid only in special cases. In particular, under constant returns-to-scale (CRS) and using means of (CRS) DEA estimators, standard CLTs remain valid only if the number of dimensions (i.e., the number of inputs plus the number of outputs) is no larger than 3. Under VRS, standard CLT results hold for means of DEA estimates only if $\mathrm{d} \leq 2$, and when FDH estimators are used, standard CLT results never hold. Kneip et al. (2015) provide new CLTs for making inference about mean efficiency using either FDH, VRS-DEA or CRS-DEA estimators. Using these results, researchers can now estimate valid confidence intervals for mean inefficiency.

Kneip et al. (2016) use the new CLTs proved by Kneip et al. (2015) to develop tests of differences in mean inefficiency across groups of producers as well as a test of convexity

of the production set versus non-convexity and a test of CRS versus VRS. Given that FDH, VRS-DEA and CRS-DEA estimators require different assumptions regarding the shape of the frontier, the latter two tests are useful for deciding which estimator to use in a particular application. While Kneip et al. (2016) provide asymptotically normal test statistics for each test, bootstrapping is still needed for the tests of convexity and returns to scale. Both tests require randomly splitting the initial sample into two independent subsamples to implement valid tests. While the results of Kneip et al. (2016) are valid for any particular random split of the sample, different results are obtained from different splits of the same sample. Simar and Wilson (2020) provide a bootstrap method to combine information across multiple, random splits of a given sample, thereby removing the ambiguity resulting from a single random split of the sample.

2.2 Robust approach

As discussed, traditional non-parametric techniques such as DEA and FDH are sensitive to outliers. This drawback can be addressed by employing the order-m and order-α quantile frontiers introduced by Cazals et al. (2002) and Aragon et al. (2005), respectively. These two robust approaches are briefly reviewed in the following.

2.2.1 Order-m frontier

Cazals et al. (2002) initially introduce the order-m partial frontier based on a probabilistic formulation. Accordingly, the order-m frontier is defined as the expected minimum use of inputs among sample of m DMUs drawn from the population generating more than a given level of output. Due to its sampling effect, the order-m approach provides less-extreme benchmarks than the usual non-parametric estimators. Daraio and Simar (2005) develop the idea of Cazals et al. (2002) to a multivariate case and introduce external environmental factors in modelling the FDH estimators. Daraio and Simar (2007b) further extend the early works on the order-m partial frontier to a fully non-parametric methodology with the convexity assumption. The summary of the order-m algorithm to calculate the efficiency score of $\widehat{\theta}_m(x, y)$ is as follows:

[1] Repeat the next two steps for $b = 1, \ldots, B$ (B is a large number, say 1000).

[1-1] For a given y, generate $(X_{1,b}, \ldots, X_{m,b})$ by drawing a sample size m with replacement among those X_i such that $Y_i \geq y$.[4]

[1-2] Calculate $\overset{\sim b}{\theta}_m(x, y)$ using the following linear program

$$\tilde{\theta}_m^b(x, y) = min\left\{\theta > 0 | \theta x \geq \sum_{i=1}^{m} \lambda_i X_{i,b},\ \sum_{i=1}^{m} \lambda_i = 1,\ \lambda_i \geq 0\ \forall i = 1, \ldots, m\right\}. \tag{5}$$

[2] Calculate the order-m efficiency score using

$$\widehat{\theta}_m(x, y) \approx \frac{1}{B} \sum_{b=1}^{B} \tilde{\theta}_m^b(x, y).$$

[4] For multiple output scenarios, all elements of outputs should satisfy $Yi \geq y$. One of the limitations of this method is the fact that no or only few DMUs might satisfy $Yi \geq y$ especially where the sample size is small.

Daraio and Simar (2007b) also introduce a conditional measure of efficiency by including environmental factors in the probabilistic formulation of the production function. Accordingly, the joint distribution of (X, Y) conditional on environmental factors, denoted by Z, defines the production process. The summary of the conditional order-m algorithm to calculate the efficiency score of $\widehat{\theta}_m(x, y|z)$ is as follows:

[1] Repeat the next two steps for $b = 1, \ldots, B$ (B is a large number, say 1000).

[1-1] For a given y, generate the sample of $(X_{1,b}, \ldots, X_{m,b})$ by drawing a sample size m with replacement, and with a probability $K\left(\frac{z-z_i}{h}\right) / \sum_{j=1}^{n} K\left(\frac{z-z_j}{h}\right)$, among those X_i such that $Y_i \geq y$. h is the chosen bandwidth for kernel $K(.)$ with bounded support (see Bădin et al. (2010)).

[1-2] Calculate $\tilde{\theta}_m^{z,b}(x, y)$ using the following linear program

$$\tilde{\theta}_m^{z,b}(\mathrm{x}, \mathrm{y}) = min\left\{\theta > 0 | \theta x \geq \sum_{i=1}^{m} \lambda_i X_{i,b}, \sum_{i=1}^{m} \lambda_i = 1, \lambda_i \geq 0 \forall i = 1, \ldots, m\right\}. \tag{6}$$

[2] Calculate the conditional order-m efficiency score using

$$\widehat{\theta}_m(x, y|z) \approx \frac{1}{B} \sum_{b=1}^{B} \tilde{\theta}_m^{z,b}(\mathrm{x}, \mathrm{y}),$$

Simar and Vanhems (2012) adapt the order-m approach from the radial models to directional distances. For information about the difference between radial and directional distance models readers are referred to Chambers et al. (1998) and Färe and Grosskopf (2006).

2.2.2 Order- α quantile partial frontier

Aragon et al. (2005) introduce a robust α-quantile approach based on the quantiles of a univariate distribution function to deal with outliers in non-parametric efficiency estimation. The frontier of order- α quantile is defined as the input level not exceeded by $(1-\alpha) \times 100$ percent of DMUs from the population of peers generating more than a given level of output. In contrast to order-m, the order-α approach have better robustness properties as trimming is continuous ($\alpha \in (0,1]$) in terms of the order-α quantile. The idea is further extended to a full multivariate setup by Daouia and Simar (2007). The summary of the order-α algorithm with no convexity assumption to calculate the efficiency score of $\widehat{\theta}_\alpha(x, y)$ is as follows:

[1] Define

$$\mathfrak{x}_i = \max_{k=1,\ldots,p} \frac{x_i^k}{x^k}, i = 1, \ldots, n$$

[2] Denote $\mathfrak{x}_{(j)}^{y}$ as the jth order statistic of the observation $\mathfrak{x}_i$ such that $Y_i \geq y$[5] for $j = 1, \ldots, M_y$,

where $M_y = \sum_{i=1}^{n} 1(Y_i \geq y)$.

[5] See footnote 4

[3] Sort $\mathfrak{x}_j^y$ as:

$$\mathfrak{x}_{(1)}^y \leq \mathfrak{x}_{(2)}^y \leq \cdots \leq \mathfrak{x}_{(M_y)}^y$$

[4] Calculate the order-α efficiency score using

$$\hat{\theta}_\alpha(x,y) = \begin{cases} \mathfrak{x}^y_{\left((1-\alpha)M_y\right)} & if\ (1-\alpha)M_y\ is\ nonnegative\ intetger \\ \mathfrak{x}^y_{\left(\left[(1-\alpha)M_y\right]+1\right)} & otherwise \end{cases},$$

where $\left[(1-\alpha)M_y\right]$ is the integer part of $(1-\alpha)M_y$.

Recently, Daouia et al. (2017) propose an alternative way for the full multivariate environment based on the directional distance estimator of order-α suggested in Simar and Vanhems (2012).

2.3 Environmental variables

In addition to measuring the level of efficiency for DMUs, researchers and policy makers are interested to examine the impact of environmental factors on the production process. It is assumed that these environmental factors are exogenous and are not under control of management. Two recent approaches in the literature to deal with environmental factors are briefly discussed in the following.

2.3.1 Two-stage regression approach

A large part of the literature focuses on two-stage models where the level of efficiency is estimated in the first stage, and then estimated efficiency scores are regressed on a number of covariates in the second stage. However, as discussed by Simar and Wilson (2007), the majority of these studies provide invalid inferences due to the existence of unknown serial correlation among the estimated efficiency scores. Simar and Wilson (2007) examine this genre and provide important sets of results. They give the first (and as far as we know the only one) coherent statistical model encompassing both the first stage (where efficiency is estimated) and the second stage (where efficiency estimates from the first stage are regressed on environmental variables). Simar and Wilson (2007) show that the second-stage regression is a latent variable problem—one would like to regress inefficiency on the environmental variables, but inefficiency is unobserved and hence inefficiency estimates must be used. Moreover, the second-stage regression is shown to be a truncated regression, as opposed to censored (Tobit) regressions or linear regressions that are sometimes used in applied papers.

Specifically, Simar and Wilson (2007) propose bootstrap procedures to address aforementioned drawbacks using a bootstrap truncated regression given as

$$\widehat{\theta_i} = Z_i\beta + \varepsilon_i, \tag{7}$$

where $\widehat{\theta_i}$ is the estimated efficiency score, Z_i is a vector of environmental variables, β is the coefficient vector and ε_i is the stochastic error term. The summary of the double bootstrap procedure is as follows:

[1] Compute the efficiency scores using (1) for all DMUs to generate $\widehat{\theta_1}, \ldots, \widehat{\theta_n}$.
[2] Use the truncated regression (7) to obtain estimates $\widehat{\beta}$ of β and $\widehat{\sigma}_\varepsilon$ of σ_ε for $0 < \widehat{\theta_i} < 1$.

[3] Repeat the next four steps B times (B is large, say 1000) for all DMUs to obtain $\left\{\widehat{\theta}^{*}_{i,b},\ \mathrm{b}=1,\ldots,\mathrm{B}\right\}$.

[3-1] Draw $\varepsilon_{i,b}$ from a standard normal distribution with mean of zero and standard deviation of $\widehat{\sigma}_{\varepsilon}$ with left truncation at $\left(-Z_i\widehat{\beta}\right)$ and right truncation at $\left(1-Z_i\widehat{\beta}\right)$ for $i=1,\ldots,n$.

[3-2] Compute $\theta^{*}_{i,b}=Z_i\widehat{\beta}+\varepsilon_{i,b}$ for $i=1,\ldots,n$.

[3-3] Set $X^{*}_{i,b}=X_i\widehat{\theta}_i/\theta^{*}_{i,b}$, $Y^{*}_{i,b}=Y_i$ for $i=1,\ldots,n$.

[3-4] Calculate $\widehat{\theta}^{*}_{i,b}$ for $i=1,\ldots,n$ using (1) by replacing X_i and Y_i with $X^{*}_{i,b}$ and $Y^{*}_{i,b}$, respectively.

[4] Calculate the bias corrected efficiency scores for $i=1,\ldots,n$ using

$$\widehat{\widehat{\theta}}_i=2\widehat{\theta}_i-\frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}^{*}_{i,b}.$$

[5] Use the maximum likelihood method to estimate the truncated regression of $\widehat{\widehat{\theta}}_i$ on Z_i to obtain $\left(\widehat{\widehat{\beta}},\widehat{\widehat{\sigma}}\right)$.

[6] Loop over the next three steps B times (B is large, say 1000) to provide $\left(\widehat{\widehat{\beta}}^{*}_{b},\widehat{\widehat{\sigma}}^{*}_{b},b=1,\ldots,B\right)$.

[6-1] Draw $\varepsilon_{i,b}$ from a normal distribution with mean of zero and standard deviation of $\widehat{\widehat{\sigma}}$ with left truncation at $\left(-Z_i\widehat{\widehat{\beta}}\right)$ and right truncation $\left(1-Z_i\widehat{\widehat{\beta}}\right)$ for $i=1,\ldots,n$.

[6-2] Compute $\theta^{**}_{i,b}=Z_i\widehat{\widehat{\beta}}+\varepsilon_{i,b}$ for $i=1,\ldots,n$.

[6-3] Use the method of maximum likelihood to obtain estimates of $\widehat{\widehat{\beta}}^{*}_{b}$ and $\widehat{\widehat{\sigma}}^{*}_{b}$ in the truncated regression of $\theta^{**}_{i,b}$ on Z_i.

[7] Construct the confidence interval for β and σ_{ε} using the bootstrap results $\left(\widehat{\widehat{\beta}}^{*}_{b},\widehat{\widehat{\sigma}}^{*}_{b},b=1,\ldots,B\right)$ as detailed in Simar and Wilson (2007).

Despite the prevalence of using this method, it requires a restrictive assumption about the separability of the frontier production from the impact of Z. This assumption can be examined using a test procedure provided by Daraio et al. (2010).

Simar and Wilson (2007) show that a second-stage regression of efficiency estimates on some environmental variables can only be meaningful if the environmental variables have no relation to nor influence on the shape of the frontier. Simar and Wilson (2007) refer to this as a "separability" condition, meaning that the environmental variables are not included in the set of variables that influence the shape of the frontier, and note that this is a strong condition that should be tested. Testing the separability condition requires additional theoretical results, in particular those of Kneip et al. (2015). Daraio et al. (2018) provide a test of the separability condition, building on the work of Kneip et al. (2015, 2016).

For cases where the separability condition is satisfied, Simar and Wilson (2007) provide bootstrap methods for making inference in second-stage regressions. The latent variable problem in the second-stage regression complicates inference; in particular, Simar and Wilson (2007) show that conventional inference (e.g., involving inverting the negative Hessian of the log-likelihood for the second-stage regression) cannot provide valid, meaningful inference. As further explained by Simar and Wilson (2011b) and Kneip et al. (2015), the inefficiency

estimates used in the second stage are biased, creating problems beyond those described by Simar and Wilson (2007).

In contrast, some studies (e.g., Banker & Natarajan, 2008; Banker et al., 2019; McDonald, 2009) highlight the limitations of second stage bootstrap approach suggested by Simar and Wilson, and argue that a two-stage method followed by ordinary least square (OLS) regression provides consistent estimators of the impact of environmental variables. For example, Banker et al. (2019) argue that the effectiveness of bootstrap approach critically relies on the assumed data generating process and show that the bootstrap approach does not provide correct inferences in presence of stochastic noise. Using extensive simulations, they assert that the OLS second-stage model significantly outperforms the complex Simar–Wilson approach.

2.3.2 Conditional efficiency approach

Daraio and Simar (2005, 2007a, 2007b) describe how the comparison between conditional and unconditional efficiency measures can be used to examine the impact of environmental factors on the production process. They define the following ratio for such analysis:

$$\widehat{R}(x, y|z) = \frac{\widehat{\theta}(x, y|z)}{\widehat{\theta}(x, y)}, \tag{8}$$

where $\widehat{\theta}(x, y)$ is the unconditional efficiency score that can be simply obtained using (1) and $\widehat{\theta}(x, y|z)$ is the conditional efficiency score that can be estimated by solving

$$\widehat{\theta}(x, y|z) = \min\left\{\theta > 0|\ \theta x \geq \sum_{i|z-h\leq z_0\leq z+h} \lambda_i X_i,\quad y \leq \sum_{i|z-h\leq z_0\leq z+h} \lambda_i Y_i,\right.$$
$$\left.\sum_{i|z-h\leq z_0\leq z+h} \lambda_i = 1,\quad \lambda_i \geq 0\ \forall i = 1, \ldots, n\right\}, \tag{9}$$

where h is the local bandwidth and can be computed according to a data driven method suggested by Daraio and Simar (2005) or the least squares cross-validation procedure proposed in Bădin et al. (2010). Daraio and Simar (2005) also demonstrate how a smoothed non-parametric regression and a scatter diagram of $\widehat{R}(x, y|z)$ against a univariate Z can be used to describe the impact of environmental variables on efficiency.

Bădin et al. (2012) propose a location scale non-parametric regression to purify the conditional efficiency scores of $\widehat{\theta}(X, Y|Z = z)$ from the impact of Z as follows:

$$\widehat{\theta}(X, Y|Z = z) = \mu(z) + \sigma(z)\varepsilon, \tag{10}$$

where $\mu(z) = \mathbb{E}(\widehat{\theta}(X, Y|Z)|Z = z)$, $\sigma(z) = \mathbb{V}(\widehat{\theta}(X, Y|Z)|Z = z)$, $\mathbb{E}(\varepsilon|Z = z) = 0$, and $\mathbb{V}(\varepsilon|Z = z) = 1$. Analysing the residual of ε reveals the unexplained part of conditional efficiency score. While a large ε indicates poor managerial performance, a small ε specifies a good level of performance. Daraio and Simar (2014) show how this approach can be adapted in a directional distance setting.

Bădin et al. (2014) further develop the idea of Bădin et al. (2012) and suggest a bootstrap algorithm to provide confidence intervals for the local impact of the external factors on the ratio of conditional to unconditional efficiency scores as follows:

[1] Compute the n unconditional and conditional efficiency scores using (1) and (9), respectively.

[2] Calculate n ratios using (8).
[3] Select a fixed grid of values for Z, say $z_1, .., z_k$
[4] Use the following non-parametric regression to calculate $\tau_n^{z_j}$ for $j = 1, \ldots, k$

$$\tau_n^{z_j} = \sum_{i=1}^{n} W_n\left(Z_i, z_j, h_z\right) R(X_i, Y_i | Z_i), \tag{11}$$

where $W_n\left(Z_i, z_j, h_z\right)$ is the Nadaraya-Watson kernel weights and given by

$$W_n\left(Z_i, z_j, h_z\right) = \frac{K\left(\left(Z_i - z_j\right)/h_z\right.}{\sum_{i=1}^{n} K\left(\left(Z_i - z_j\right)/h_z\right.}.$$

h_z is the bandwidth selected by the least-squares cross validation method suggested in Bădin et al. (2010).

[5] Repeat the next three steps for b $= 1, \ldots,$ B (B is large, say B $= 1000$) for a given value of $m < n$. Note that the choice of m is based on the data driven method described in Simar and Wilson (2011a)

[5-1] Draw a random sample of $\left\{\left(X_i^{*,b}, Y_i^{*,b}, Z_i^{*,b}\right), i = 1, \ldots, m\right\}$ without replacement from $\left\{\left(X_i, Y_i, Z_i\right), i = 1, \ldots, n\right\}$.

[5-2] Use (8) to compute m ratios for the sample. Note that the bandwidth of $h_{m,i}^{*,b}$ for calculating conditional efficiency needs to be modified by rescaling the bandwidth of $h_{n,i}^{*,b}$ used to calculate the conditional efficiency in [1] using

$$h_{m,i}^{*,b} = \left(\frac{n}{m}\right)^{1/(r+4)} h_{n,i}^{*,b},$$

where r is the number of environmental variables.

[5-3] Estimate τ_m^{*,b,z_j} at the fix point z_j for $j = 1, \ldots, k$ using (11). Note that the bandwidth needs to be rescaled to the appropriate size using

$$h_{m,z} = \left(\frac{n}{m}\right)^{1/(r+4)} h_z,$$

where h_z is the bandwidth used in [4].

[6] For each $j = 1, \ldots, k$, compute the $\alpha/2$ and 1-$\alpha/2$ quantiles of the B bootstrapped values of τ_m^{*,b,z_j}-$\tau_n^{z_j}$.

Due to flexibility of the directorial distance approach, recent studies focus more on adapting the aforementioned algorithms in this section to directional models (e.g., Daraio & Simar, 2014; Simar & Vanhems, 2012; Simar et al., 2012). Recently, Daraio et al. (2020) provide a fast and efficient computation of directional distance estimators for full frontier, robust versions, and conditional efficiency estimates on environmental factors, along with their MATLAB codes.

3 Method and data

Bibliometrics is defined as "the study of the quantitative aspects of the production, dissemination, and use of recorded information" (Van Leeuwen, 2004). It is a research area that uses quantitative methods to analyse journals, books, articles, and other publications and provides an informative overview of such bibliographic materials (Broadus, 1987; Laengle

et al., 2017; Merigó et al., 2018; Pritchard, 1969). There are a wide range of bibliometric indicators including the number of papers, citations, co-authorships, and keywords frequency. Of the varied approaches to bibliometric analysis, the study of citations is pronounced receiving increased attention (Kaffash & Marra, 2017; Lampe & Hilgers, 2015). Citation analysis provides critical information about emerging knowledge trends in a discipline (Kaffash & Marra, 2017).

This study focuses on the bibliometric analysis of the use of bootstrap approach in non-parametric efficiency analysis and statistical advancements developed since then. To conduct this bibliometric analysis, we first identify the methodological papers in the field which are found in a recent comprehensive survey conducted by Simar and Wilson (2015). Those papers with an average of at least 5 citations per year are then selected as possessing influential methodological content. We also identified the most influential methodological papers by choosing the papers with at least 50 citations per year. Using a range of bibliometric indicators including the cites per paper, cites per year, productive journals, frequent keywords and network of co-authors and countries contributed to both development and application of statistical approaches in non-parametric efficiency estimation, we identify important trends. This analysis reveals useful information about successful applications of new methods, past trend, and future outlook.

Specifically, we provide graphical visualisation of the relevant bibliometric materials by using visualisation of similarities (VOS) viewer software (Van Eck & Waltman, 2010). Recently, the VOS viewer has been used as a powerful visualization tool in bibliometric analysis of various research fields (e.g., Baier-Fuentes et al., 2019; Laengle et al., 2017; Merigó et al., 2018; Muhuri et al., 2019; Türkeli et al., 2018; Yeung et al., 2017). This software provides a network representation of bibliometric indicators such as co-authorship (Peters & Van Raan, 1991) and co-occurrence of keywords (Callon et al., 1983). Co-authorship counts the number of co-authored publications among two authors and co-occurrence counts the number of publications in which two keywords used together (Eck & Waltman, 2020).

The Scopus database owned by Elsevier Ltd is used to extract the relevant data and information including citation counts and journal ranking. It is argued that Scopus covers a greater number of journals compared to the main alternative, the Web of Science (WoS) (Baier-Fuentes et al., 2019; Mongeon & Paul-Hus, 2016). Consequently, we choose Scopus over the WoS, as it covers a greater number of publications cited our selected methodological papers and provides more flexible output formats for our bibliometric analysis. The data was collected during 2021. The covered search period was from 1998, when the first article published by Simar and Wilson (1998), to the end of 2020.

4 Influential methodological papers

Figure 2 illustrates the influential methodological papers that introduce or extend statistical approaches in the context of non-parametric models. The left vertical axis indicates the number of total citations while the right axis represents the average citation per year. As can be seen, the two most cited papers are Simar and Wilson (2007) and Simar and Wilson (1998) with 1586 and 1098 total citations, respectively. Figure 2 also reveals the average number of citations per year for these two papers are significantly higher than the other methodological papers. Therefore, for this bibliometric analysis we mainly focus on these two most influential articles, analysing the trend of their citations, and other relevant bibliometric indicators.

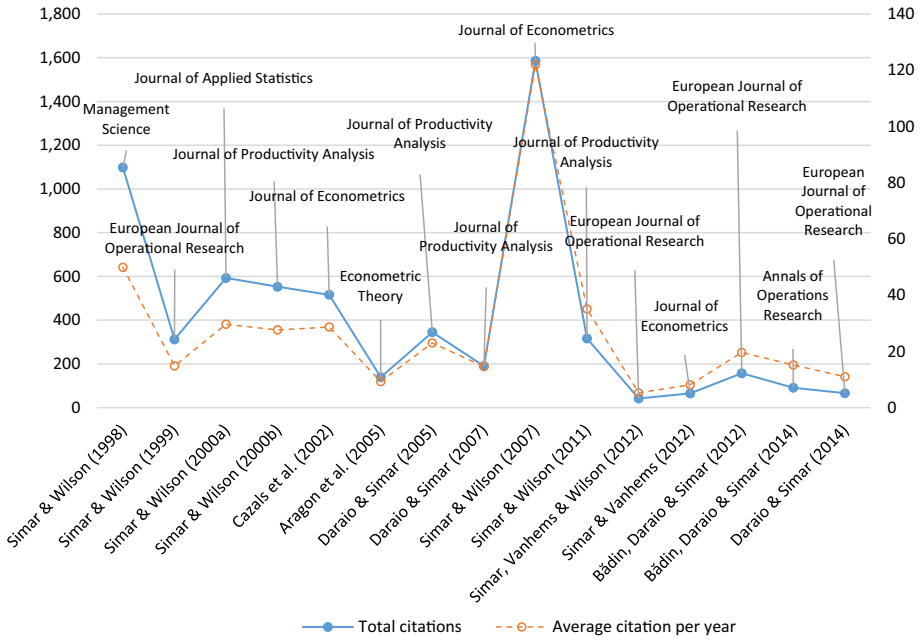

Fig. 2 Influential methodological papers

As shown in Fig. 2, the European Journal of Operational Research, Journal of Productivity Analysis and Journal of Econometrics have published the majority of methodological articles.

5 Bibliometric analysis

Figure 3 shows the trend of citations of those most influential methodological papers. As can be seen, the number of citations of Simar and Wilson (1998) is less than 20 per annum until 2007, however, a surge in citations is evident after 2008. It is likely a lack of knowledge about the method and user-friendly software applications are responsible for this10-year gap between the introduction of the method and its widespread application. On the contrary, Simar and Wilson (2007) was highly cited in a short span of time. For example, it was cited 93 times in 2010 only three years after its publication. This high citation level may be a result of the release of the FEAR package in R by Wilson (2008), combined with a decade effort by Simar and Wilson to promote their bootstrap approach. Figure 3 also details that while the number of citations of Simar and Wilson (1998) is relatively flat after its peak in 2013, the number of citations for Simar and Wilson (2007) continues its ascent reaching 238 citations in 2020.

Table 1 provides the list of top fifteen authors citing Simar and Wilson's papers published in 1998 or 2007. As is described, most authors are domiciled in European countries. Barros[6] as the most productive author has 50 published papers and is the listed first author in 27 papers. His prevalence suggests that he has played a significant role in both the promotion and application of the bootstrapping method. His research in the context of non-parametric

[6] Professor Carlos Barros, sadly, passed away in 2017.

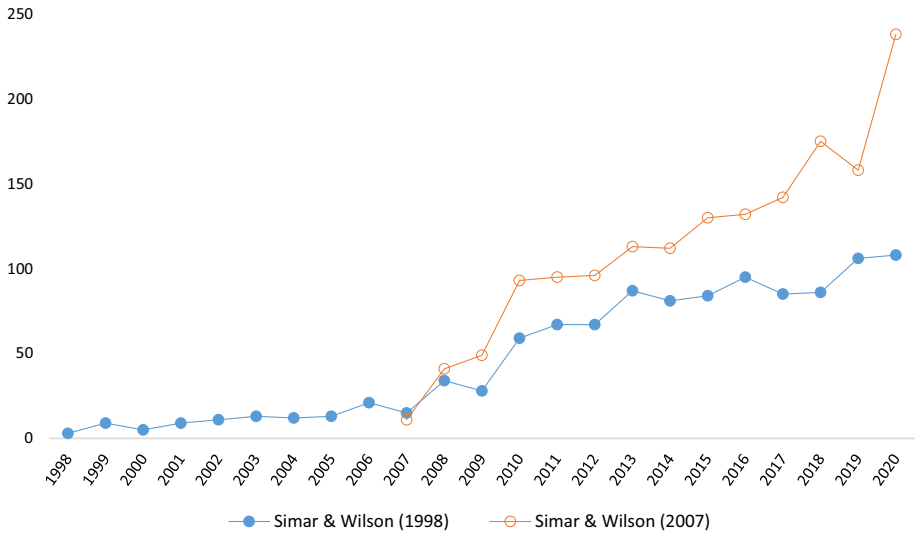

Fig. 3 Trend of citations of Simar and Wilson (1998) and Simar and Wilson (2007)

efficiency analysis sees him collaborate with 31 individual co-authors. Collaborations with Wanke and Assaf (3rd and 7th authors in Table 1) are the highest, with 18 and 13 co-authored papers, respectively. Not surprisingly, Simar is the second productive author with 47 articles. Simar has 21 individual co-authors, and he authors with Wilson (5th author in Table 1) with 23 co-authored papers. The third most productive author is Wanke with 42 papers and 39 individual co-authors. To provide a comprehensive picture of co-authorship connections, we analyse co-authorship network of key researchers.

Figure 4 reveals the co-authorship network of researchers citing Simar and Wilson (1998) or Simar and Wilson (2007). Further, it highlights how the productive authors listed in Table 1 are connected and disseminated their knowledge to other researchers worldwide.[7] To draw Fig. 4, the co-authorship network was restricted to researchers with at least 3 articles, resulting in 466 authors. Removing authors with limited connections, resulted in 236 authors in the final network map as demonstrated in Fig. 4. Note that the size of circles reflects the number of documents published by each author. As is seen, key authors form the clusters (shown with different colours) and/or act as a bridge between two or more clusters. For example, while Barros is the key person in the orange cluster, he connects to Wanke who performs as a bridge between the orange and green clusters.

Figure 4. can also provide us with useful information on how a new method or approach in efficiency analysis is disseminated among researchers in different countries or institutions. Interestingly, Fig. 4. shows Simar and Wilson's co-authorship cluster (shown in dark blue) is relatively smaller than some other clusters and they have directly worked with a limited number of co-authors. However, further investigation reveals Simar has run many workshops, seminars and trained European researchers.[8] Thus, it suggests, lectures and seminar involvement may be a useful approach to introduce new methods in efficiency analysis.

[7] There are two authors (Lu W.-M. and Guccio C.) in Table 1 that cannot be identified in Fig. 4 given that they have no direct connection with other authors.

[8] See the list of 142 talks or lectures provided by Simar between 1998 and 2014.http://www.dis.uniroma1.it/diagest/sites/default/files/SIMAR%20CV_roma.PDF

Table 1 Top fifteen authors cited Simar and Wilson (1998) or Simar and Wilson (2007)

Author	Affiliation	Total	Number of Publications				
			1st author	2nd author	3rd author	4th author	5th + author
Barros C.P	Technical University of Lisbon, Portugal	50	27	21	2	–	–
Simar L	Catholic University of Leuven, Belgium	47	23	15	8	1	–
Wanke P.F	COPPEAD Graduate Business School, UK	42	26	7	6	2	1
Oude Lansink A	Wageningen University, Netherlands	35	2	17	8	5	3
Wilson P.W	Clemson University, USA	33	2	21	8	2	–
Tzeremes N.G	University of Thessaly, Greece	32	1	19	8	4	–
Assaf A.G	University of Massachusetts-Amherst, USA	31	24	5	1	–	1
Halkos G	University of Thessaly, Greece	28	25	3	–	–	–
Marques R.C	Technical University of Lisbon, Portugal	26	3	18	3	2	–
Guccio C	University of Catania, Italy	24	11	13	–	–	–
Lu W.-M	National Defense University, Taiwan	23	7	9	6	–	1
Zelenyuk V	University of Queensland, Australia	20	2	8	7	3	–
Tortosa-Ausina E	Jaume I University, Spain	19	3	1	8	7	–
Daraio C	Sapienza University of Rome, Italy	19	7	10	2	–	–
de Witte K	Maastricht University, Netherlands	18	8	9	–	1	–

In addition, direct trainings via PhD candidates or post-doctoral fellowships also seems effective. For example, Fig. 4. shows that Zelenyuk (twelfth in Table 1) played a key role as a bridge between the purple cluster and several other clusters. Zelenyuk has been a post-doctoral research fellow at the Catholic University of Leuven, working with Simar between Sep 2004 and Aug 2005.

Figure 5 presents co-authorship network map of countries. This network is based on 70 countries with at least 3 documents (total number of countries was 114). Croatia has also been removed from the map, having no link with other countries. As can be seen, the USA and UK have the highest number of documents and interestingly, many developing countries in Fig. 5, indicative of the global application of the bootstrapping method. Figure 5 also details

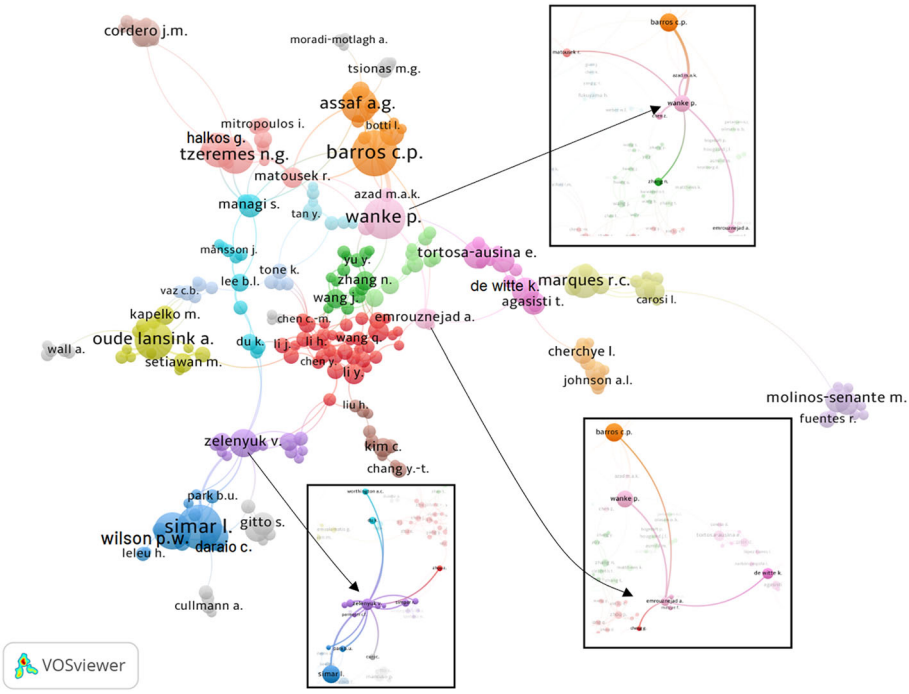

Fig. 4 Co-authorship network

how researchers from different countries are connected and formed clusters. For example, the link between Belgium and Italy demonstrated the close relationship between researchers in these countries, as two out of fifteen highly productive authors are based in Italy (as shown in Table 1). Further, our investigation reveals Daraio (14th in Table 1) was a former PhD candidate of Simar, again highlighting the importance of direct connection and training of PhD candidates, which in turn disseminates techniques and develops new knowledge.

Table 2 provides the list of academic journals with at least ten articles citing Simar and Wilson (1998) or Simar and Wilson (2007). It reveals the European Journal of Operational Research, Journal of Productivity Analysis and Applied Economics with 110, 89 and 53 are, respectively, the three most utilised journals. This result consistent with Fig. 2 which shows the European Journal of Operational Research and Journal of Productivity Analysis had the highest number of methodological papers on statistical developments in the context of non-parametric methods. Note that the percentage and cumulative percentage columns are based on the total number of 2,306 documents including journal and conference publications cited Simar and Wilson (1998) or Simar and Wilson (2007). As shown in Table 2, 35 listed journals cover more than a third of total documents. Therefore, it highlights industries in which the bootstrap approach has been employed, including health, air transport, energy, environment, and banking.

Figure 6 shows the most frequently used all keywords in papers cited by Simar and Wilson (1998) or Simar and Wilson (2007). Choosing keywords with at least 5 recurrences, resulted in 846 keywords. Removing common keywords including DEA, efficiency and bootstrapping provided increased clarity of the co-occurrence of keywords network.

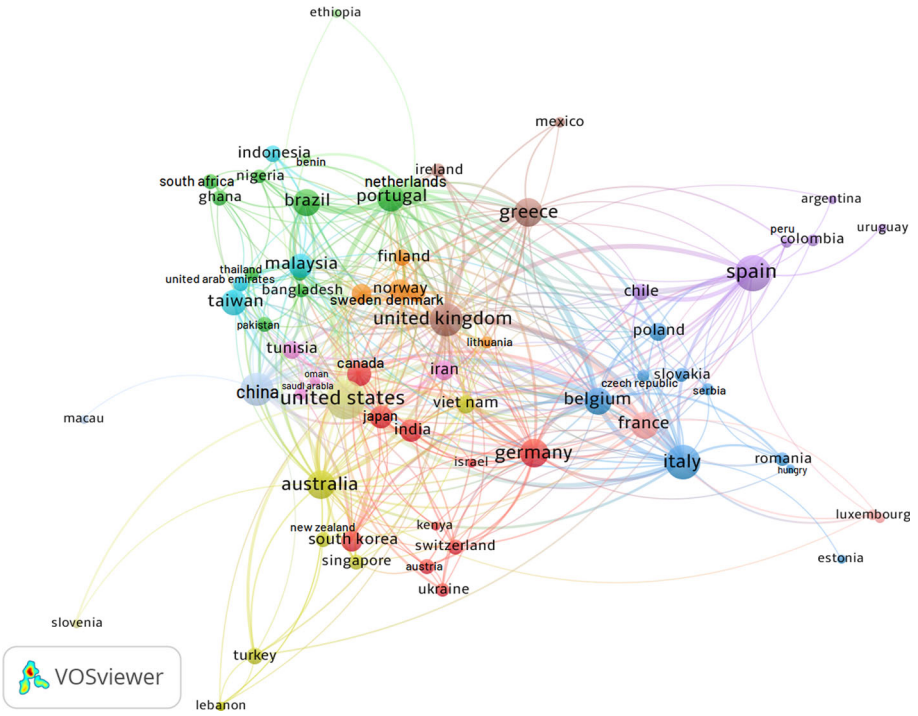

Fig. 5 Co-authorship country network

As shown in Fig. 6, technical efficiency and productivity are the commonly used keywords located in the centre of network. A continuum of colours is used to distinguish between the average year of publication for the keywords. Accordingly, a shift from traditional topics such as banking, air transportation and health to new areas such as environmental efficiency, water treatment and sustainability can be seen. A co-occurrence keywords network map can highlight the most common research areas in different countries or regions. We highlight the relevant parts of Fig. 6 for the top five countries in the appendix to provide a clear picture of the link between each country and keywords. Specifically, Figs. 7, 8, 9, 10 and 11 illustrate the keywords networks of China, Italy, Spain, Brazil, and United States, respectively. As can been, these figures show both the average year of publications and research areas. For example, comparing Figs. 7 and 11 shows China, as a keyword has been used in more recent studies than the United States. In other words, there is more recent attention on efficiency studies in China. The comparison also reveals more topics are relevant to China than other countries. Specifically, efficiency studies related to China have addressed a wide range of research topics, especially (yellow coloured) recent research areas such as eco-efficiency, sustainability, and wastewater treatment.

Figs. 7, 8, 9, 10 and 11 also assist us to identify the areas where the frontier methods have been applied in each country. For example, Fig. 7 shows that in China the focus of efficiency analysis has been more on areas such as airline, airport, agriculture, energy efficiency, health care, retailing, sustainability and environmental efficiency. Similarly, Figs. 8, 9, 10 and 11 reveal the research areas in other countries respectively as follows: In Italy, they include: airline, airport, health care, hospitals, local government, tourism, and water industry. Efficiency

Table 2 Journals with at least ten articles citing Simar and Wilson (1998) or Simar and Wilson (2007)

Journal	Cite Score 2020*	Number of papers	Percentage (%)	Cumulative Percentage (%)
European Journal of Operational Research	9.5	110	4.77%	4.77%
Journal of Productivity Analysis	3.1	89	3.86%	8.63%
Applied Economics	2.3	53	2.30%	10.93%
Sustainability (Switzerland)	3.9	36	1.56%	12.49%
Journal of Cleaner Production	13.1	35	1.52%	14.01%
Socio-Economic Planning Sciences	4.9	35	1.52%	15.52%
Health Care Management Science	4.6	30	1.30%	16.83%
Journal of Air Transport Management	6.5	30	1.30%	18.13%
Annals of Operations Research	5.2	29	1.26%	19.38%
Journal of the Operational Research Society	4.1	29	1.26%	20.64%
Energy Policy	10.2	28	1.21%	21.86%
Energy Economics	10	26	1.13%	22.98%
Omega (United Kingdom)	11.7	23	1.00%	23.98%
Economic Modelling	4.9	21	0.91%	24.89%
International Series in Operations Research and Management Science	1.1	19	0.82%	25.72%
Transportation Research Part A: Policy and Practice	8.5	19	0.82%	26.54%
Journal of Environmental Management	9.8	17	0.74%	27.28%
Utilities Policy	4.1	17	0.74%	28.01%
International Journal of Production Economics	12.2	15	0.65%	28.66%
Journal of Banking and Finance	4	15	0.65%	29.31%
Empirical Economics	2	14	0.61%	29.92%
Expert Systems with Applications	12.7	14	0.61%	30.53%
Health Policy	4.3	14	0.61%	31.14%

Table 2 (continued)

Journal	Cite Score 2020*	Number of papers	Percentage (%)	Cumulative Percentage (%)
International Transactions in Operational Research	6.2	14	0.61%	31.74%
Transport Policy	6.9	12	0.52%	32.26%
Managerial and Decision Economics	1.2	11	0.48%	32.74%
Research in International Business and Finance	4.9	11	0.48%	33.22%
Agricultural Economics	3.8	10	0.43%	33.65%
Benchmarking	5	10	0.43%	34.08%
Economics Bulletin	0.7	10	0.43%	34.52%
Energy	11.5	10	0.43%	34.95%
Journal of Economic Studies	2.3	10	0.43%	35.39%
Operational Research	3.8	10	0.43%	35.82%
Operations Research	4.3	10	0.43%	36.25%
Transportation Research Part E: Logistics and Transportation Review	9.3	10	0.43%	36.69%

*Cite Scores were extracted from Scopus on 12 January 2022

studies conducted in Spain are mainly on agriculture, education, energy, food industry, hospital, local government, sustainability, port operation and water industry. Studies in Brazil also focus on areas such as airport, agriculture, electricity, energy, local government, port, public health, and railway transport. Finally, efficiency analysis in the United States covers areas like airline, airport, electricity, energy, government, health care, higher education, hospital, power plans, sustainability, utility sector and water. As can been, among all five countries health related studies is the most common research area followed by airline, airport, and environment.

6 Software applications

Darairo et al. (2019) provide a comprehensive review of software options to conduct efficiency and productivity analysis. Given that our emphasis is on the bootstrap approach, we only focus on those software applications which are capable of handling bootstrap method in non-parametric efficiency context. Due to the increasing use of the bootstrapping approach, recently more software applications and packages have been developed mainly by academics in the field. While some packages like FEAR designed by Wilson (2008) have been specifically developed to run the bootstrap DEA models, some other existing DEA software applications have added bootstrapping as a new option. Table 3 provides a list of common software applications and packages capable of running bootstrap methods suggested by Simar and Wilson (1998) and/or Simar and Wilson (2007).

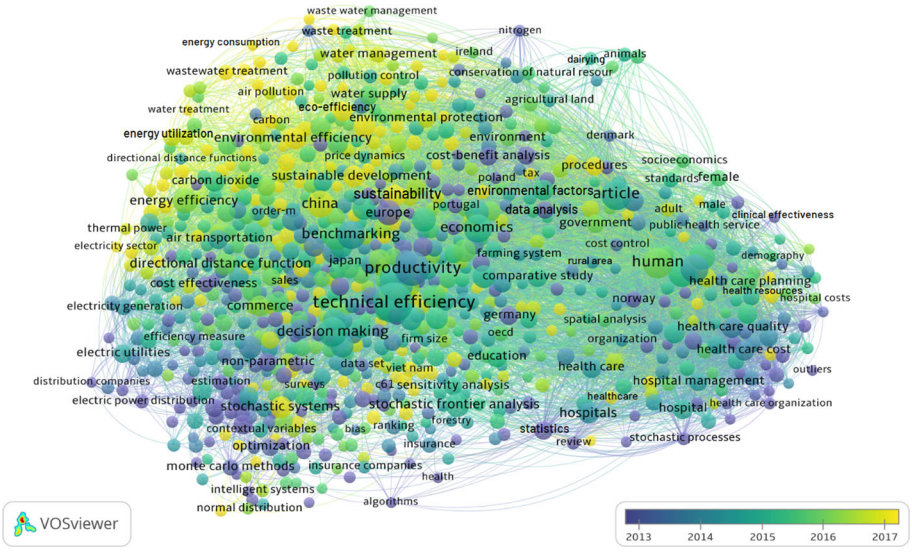

Fig. 6 Co-occurrence of keywords

Table 3 Software applications

Software	Package	Developer(s)	Reference	Citation*
R	FEAR	Paul Wilson	Wilson (2008)	916
R	Benchmarking	Peter Bogetoft & Lars Otto	Bogetoft et al. (2019)	224
R	rDEA	Jaak Simm & Galina Besstremyannaya	Simm et al. (2016)	121
Ms Access	MaxDEA	Cheng Gang & Qian Zhenhua	Cheng and Qian (2014)	79
Stata	Simarwilson	Oleg Badunenko & Harald Tauchmann	Badunenko and Tauchmann (2018)	61
PIM-DEA		Emrouznejad & Thanassoulis	Emrouznejad and Thanassoulis (2005)	31
Ms Excel	DEAFrontier	Joe Zhu	Zhu (2014)	25
Matlab	Data Envelopment Analysis Toolbox	Inmaculada C. Álvarez, Javier Barbero and José L. Zofío	Álvarez et al. (2020)	23
R	DEAboot	Joe Atwood & Saleem Shaik	Atwood and Shaik (2015)	8

*The name of software along with "bootstrap","Simar" and "Wilson" were used to search the number of documents has mentioned each package. This search was conducted using the Google Scholar on 14th February 2022

As shown in Table 3, there exist a wide range of platforms from Ms Excel to R which are used to develop packages capable of running the bootstrap method. R is a popular free open source statistical software and widely used by academics to create novel statistical packages. There are more bootstrap packages in R than other platforms and the FEAR package is by far the most used program.

7 Conclusions

Simar and Wilson (1998) propose a bootstrap DEA approach which addresses one of the main shortcomings of non-parametric efficiency analysis methods such as DEA and FDH. The method provides statistical properties of efficiency estimates including bias and confidence intervals. In the last two decades, the statistical approaches including the bootstrap method have been widely used in DEA efficiency estimations. Simar and Wilson (1998) and Simar and Wilson (2007) have been highly cited by authors following these publications, yet, despite the widespread popularity of their method, to the best of our knowledge, no comprehensive review and bibliometric analysis have been undertaken. Addressing this gap, this study not only reviews and summarises the influential statistical approaches, but it also provides a bibliometric analysis of the two most influential papers in the field, totalling 1586 and 1098 citations, respectively by 2020.

We utilised a range of bibliometric indicators including the cites per paper, cites per year, productive journals, frequent keywords and network of co-authors and countries revealing useful information about the most influential methods and potential factors driving their success. For example, we identified the most productive authors, highlighting their role in disseminating the bootstrap approach. Further, we revealed the most common research areas in different countries and trends using keyword co-occurrence network analysis. Our results highlight how future methodological advancements can be highlighted, promoted effectively disseminated. Finally, we revealed the important role workshops, seminars and lectures, co-authorship and user-friendly software applications play in enhancing the use of methodological advancements in the field.

Acknowledgment Authors would like to thank Professor Paul W. Wilson (at Department of Economics, Clemson University, USA) for his review of the earlier version of this article. His valuable comments and suggestions helped us to improve the quality of the article.

Appendix

See Figs. 7, 8, 9, 10, and 11.

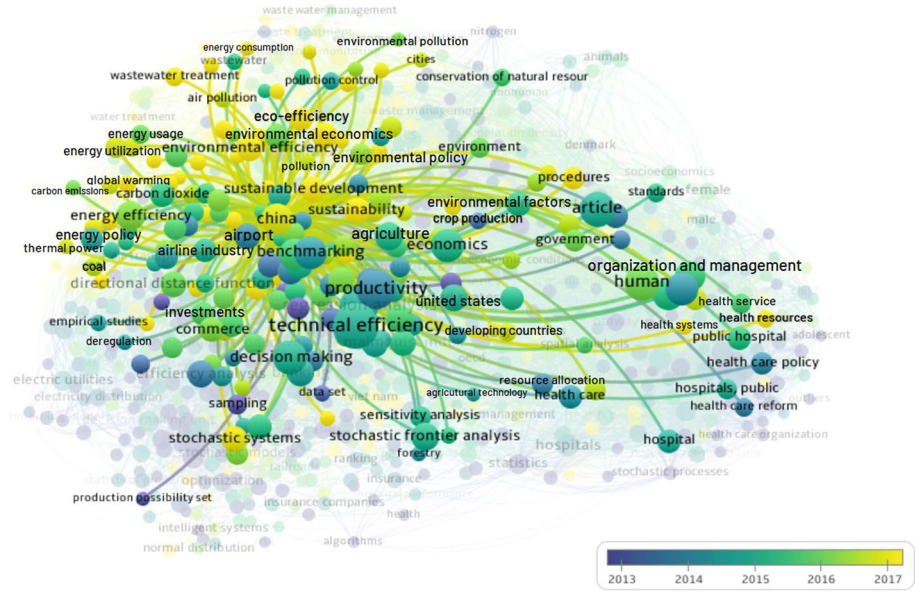

Fig. 7 Keywords network related to China

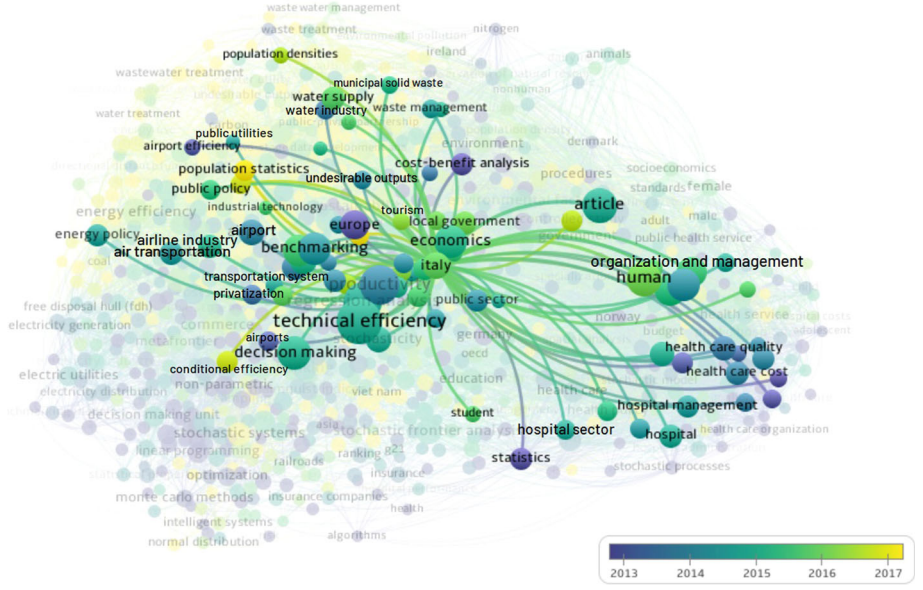

Fig. 8 Keywords network related to Italy

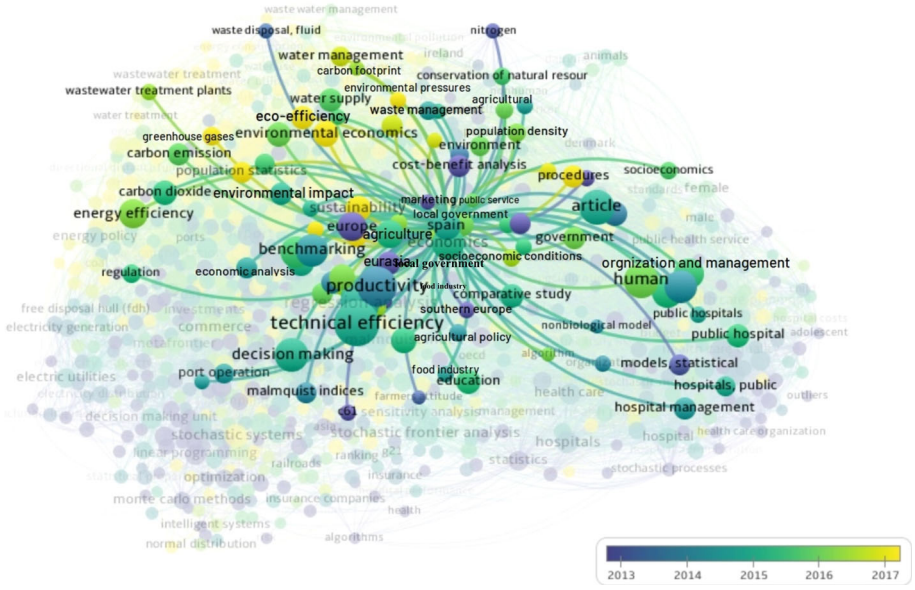

Fig. 9 Keywords network related to Spain

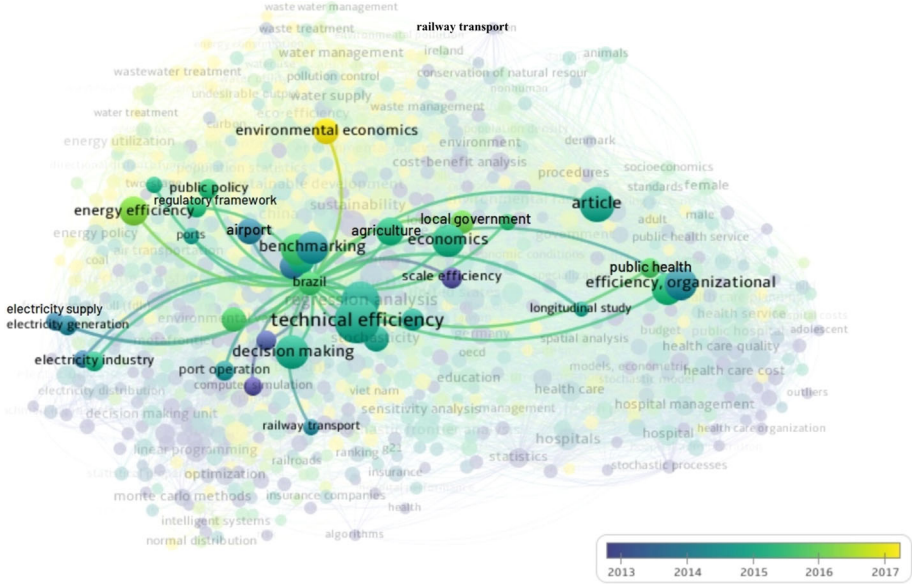

Fig. 10 Keywords network related to Brazil

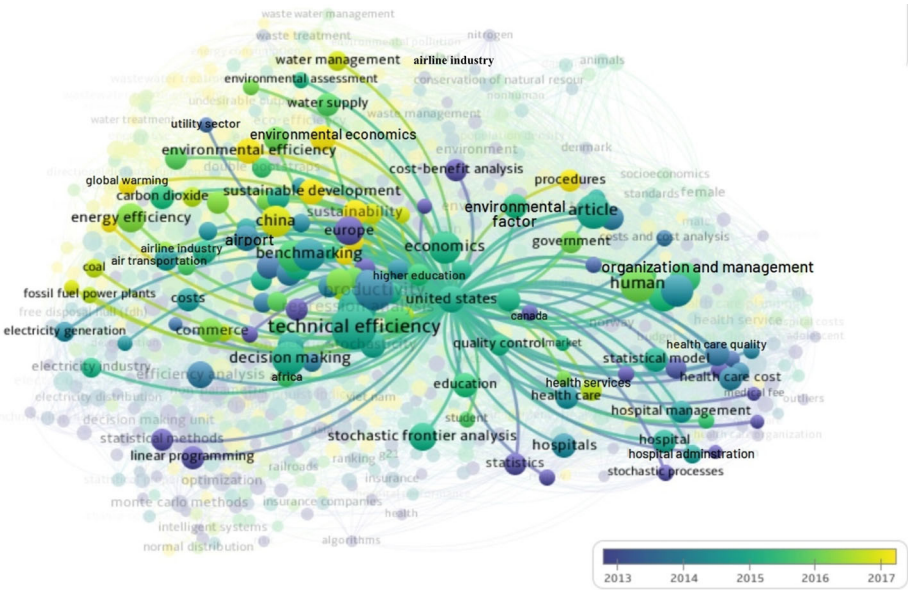

Fig. 11 Keywords network related to the United States

References

Agrell, P. J., Mattsson, P., & Månsson, J. (2020). Impacts on efficiency of merging the Swedish district courts. *Annals of Operations Research, 288*(2), 653–679.

Alberta Oliveira, M., & Santos, C. (2005). Assessing school efficiency in Portugal using FDH and bootstrapping. *Applied Economics, 37*(8), 957–968.

Álvarez, I. C., Barbero, J., & Zofío, J. L. (2020). A data envelopment analysis toolbox for MATLAB. *Journal of Statistical Software*, *95*(3), 1–49.

Andersson, C., Antelius, J., Månsson, J., & Sund, K. (2017). Technical efficiency and productivity for higher education institutions in Sweden. *Scandinavian Journal of Educational Research, 61*(2), 205–223.

Aragon, Y., Daouia, A., & Thomas-Agnan, C. (2005). Nonparametric frontier estimation: A conditional quantile-based approach. *Econometric Theory, 21*(2), 358–389.

Atwood, J., & Shaik, S. (2015). Package 'DEAboot'

Bădin, L., Daraio, C., & Simar, L. (2010). Optimal bandwidth selection for conditional efficiency measures: A data-driven approach. *European Journal of Operational Research, 201*(2), 633–640.

Bădin, L., Daraio, C., & Simar, L. (2012). How to measure the impact of environmental factors in a nonparametric production model. *European Journal of Operational Research, 223*(3), 818–833.

Bădin, L., Daraio, C., & Simar, L. (2014). Explaining inefficiency in nonparametric production models: The state of the art. *Annals of Operations Research, 214*(1), 5–30.

Badunenko, O., & Tauchmann, H. (2018). SIMARWILSON: Stata module to perform Simar & Wilson (2007) efficiency analysis.

Baier-Fuentes, H., Merigó, J. M., Amorós, J. E., & Gaviria-Marín, M. (2019). International entrepreneurship: A bibliometric overview. *International Entrepreneurship and Management Journal, 15*(2), 385–429.

Banker, R., & Maindiratta, A. (1992). Maximum likelihood estimation of monotone and concave production frontiers. *Journal of Productivity Analysis, 3*(4), 401–415.

Banker, R., & Natarajan, R. (2008). Evaluating contextual variables affecting productivity using data envelopment analysis. *Operations Research, 56*(1), 48–58.

Banker, R., Natarajan, R., & Zhang, D. (2019). Two-stage estimation of the impact of contextual variables in stochastic frontier production function models using data envelopment analysis: Second stage OLS versus bootstrap approaches. *European Journal of Operational Research, 278*(2), 368–384.

Boame, A. K. (2004). The technical efficiency of Canadian urban transit systems. *Transportation Research Part E: Logistics and Transportation Review, 40*(5), 401–416.
Bogetoft, P., Otto, L., & Otto, M. L. (2019). Package 'benchmarking'.
Broadus, R. (1987). Toward a definition of "bibliometrics." *Scientometrics, 12*(5–6), 373–379.
Brümmer, B. (2001). Estimating confidence intervals for technical efficiency: The case of private farms in Slovenia. *European Review of Agricultural Economics, 28*(3), 285–306.
Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Information (international Social Science Council), 22*(2), 191–235.
Cazals, C., Florens, J.-P., & Simar, L. (2002). Nonparametric frontier estimation: A robust approach. *Journal of Econometrics, 106*(1), 1–25.
Chambers, R. G., Chung, Y., & Färe, R. (1998). Profit, directional distance functions, and Nerlovian efficiency. *Journal of Optimization Theory and Applications, 98*(2), 351–364.
Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429–444.
Cheng, G., & Qian, Z. (2014). *MaxDea pro 6.3 manual*. Beijing Realworld Software Company Ltd: Beijing.
Cooper, W. W., Huang, Z., Lelas, V., Li, S. X., & Olesen, O. B. (1998). Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA. *Journal of Productivity Analysis, 9*(1), 53–79.
Daouia, A., & Simar, L. (2007). Nonparametric efficiency analysis: A multivariate conditional quantile approach. *Journal of Econometrics, 140*(2), 375–400.
Daouia, A., Simar, L., & Wilson, P. W. (2017). Measuring firm performance using nonparametric quantile-type distances. *Econometric Reviews, 36*(1–3), 156–181.
Daraio, C., & Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: A probabilistic approach. *Journal of Productivity Analysis, 24*(1), 93–121.
Daraio, C., & Simar, L. (2007a). *Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications*. Berlin: Springer Science & Business Media.
Daraio, C., & Simar, L. (2007b). Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach. *Journal of Productivity Analysis, 28*(1–2), 13–32.
Daraio, C., & Simar, L. (2014). Directional distances and their robust versions: Computational and testing issues. *European Journal of Operational Research, 237*(1), 358–369.
Daraio, C., Simar, L., & Wilson, P. W. (2010). Testing whether two-stage estimation is meaningful in nonparametric models of production. *Universite Catholique De Louvain (Discussion Paper)*.
Daraio, C., Simar, L., & Wilson, P. W. (2018). Central limit theorems for conditional efficiency measures and tests of the 'separability'condition in non-parametric, two-stage models of production. *The Econometrics Journal, 21*(2), 170–191.
Daraio, C., Simar, L., & Wilson, P. W. (2020). Fast and efficient computation of directional distance estimators. *Annals of Operations Research*, *288*(2), 805–835.
Darairo, C., Kerstens, K., Nepomuceno, T. C. C., & Sickles, R. C. (2019). Productivity and efficiency analysis software: An exploratory bibliographical survey of the options. *Journal of Economic Surveys, 33*(1), 85–100.
Davidova, S., & Latruffe, L. (2007). Relationships between technical efficiency and financial management for Czech Republic farms. *Journal of Agricultural Economics, 58*(2), 269–288.
De Witte, K., & Marques, R. C. (2010). Designing performance incentives, an international benchmark study in the water sector. *Central European Journal of Operations Research, 18*(2), 189–220.
Deprins, D., Simar, L., & Tulkens, H. (2006). Measuring labor-efficiency in post offices. In P. Chander, J. Drèze, C. Knox Lovell, & J. Mintz (Eds.), *Public goods, environmental externalities and fiscal competition* (pp. 285–309). Springer.
Eck, N. v., & Waltman, L. (2020). VOSviewer Manual: Manual for VOSviewer Version 1.6. 14: Leiden: CWTS.
Efron, B. (1992). Bootstrap methods: another look at the jackknife. In S. Kotz, & N. L. Johnson (Eds.), *Breakthroughs in statistics* (pp. 569–593). Springer.
Eling, M., & Luhnen, M. (2010). Efficiency in the international insurance industry: A cross-country comparison. *Journal of Banking & Finance, 34*(7), 1497–1509.
Emrouznejad, A., & Thanassoulis, E. (2005). Performance improvement management. *DEASoft, PIM Ltd*.
Emrouznejad, A., & Yang, G.-L. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences, 61*, 4–8.
Essid, H., Ouellette, P., & Vigeant, S. (2014). Productivity, efficiency, and technical change of Tunisian schools: A bootstrapped Malmquist approach with quasi-fixed inputs. *Omega, 42*(1), 88–97.
Färe, R., & Grosskopf, S. (2006). *New directions: Efficiency and productivity* (Vol. 3). Berlin: Springer Science & Business Media.

Fukuyama, H., & Tan, Y. (2021). Corporate social behaviour: Is it good for efficiency in the Chinese banking industry? *Annals of Operations Research, 306*(1), 383–413.
Galariotis, E., Kosmidou, K., Kousenidis, D., Lazaridou, E., & Papapanagiotou, T. (2021). Measuring the effects of M&As on Eurozone bank efficiency: An innovative approach on concentration and credibility impacts. *Annals of Operations Research, 306*(1), 343–368.
Halkos, G. E., & Tzeremes, N. G. (2013). Estimating the degree of operating efficiency gains from a potential bank merger and acquisition: A DEA bootstrapped approach. *Journal of Banking & Finance, 37*(5), 1658–1668.
Hatami-Marbini, A., Emrouznejad, A., & Tavana, M. (2011). A taxonomy and review of the fuzzy data envelopment analysis literature: Two decades in the making. *European Journal of Operational Research, 214*(3), 457–472.
Hawdon, D. (2003). Efficiency, performance and regulation of the international gas industry—a bootstrap DEA approach. *Energy Policy, 31*(11), 1167–1178.
Johnes, J. (2006). Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economics of Education Review, 25*(3), 273–288.
Kaffash, S., & Marra, M. (2017). Data envelopment analysis in financial services: A citations network analysis of banks, insurance companies and money market funds. *Annals of Operations Research, 253*(1), 307–344.
Kao, C. (2014). Network data envelopment analysis: A review. *European Journal of Operational Research, 239*(1), 1–16.
Kneip, A., Park, B. U., & Simar, L. (1998). A note on the convergence of nonparametric DEA estimators for production efficiency scores. *Econometric Theory, 14*, 783–793.
Kneip, A., Simar, L., & Wilson, P. W. (2008). Asymptotics and consistent bootstraps for DEA estimators in nonparametric frontier models. *Econometric Theory, 24*, 1663–1697.
Kneip, A., Simar, L., & Wilson, P. W. (2015). When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores. *Econometric Theory, 31*, 394–422.
Kneip, A., Simar, L., & Wilson, P. W. (2016). Testing hypotheses in nonparametric models of production. *Journal of Business & Economic Statistics, 34*(3), 435–456.
Kohl, S., Schoenfelder, J., Fügener, A., & Brunner, J. O. (2019). The use of Data Envelopment Analysis (DEA) in healthcare with a focus on hospitals. *Health Care Management Science, 22*(2), 245–286.
Laengle, S., Merigó, J. M., Miranda, J., Słowiński, R., Bomze, I., Borgonovo, E., et al. (2017). Forty years of the European Journal of Operational Research: A bibliometric overview. *European Journal of Operational Research, 262*(3), 803–816.
Lampe, H. W., & Hilgers, D. (2015). Trajectories of efficiency measurement: A bibliometric analysis of DEA and SFA. *European Journal of Operational Research, 240*(1), 1–21.
Liu, J. S., Lu, L. Y., Lu, W.-M., & Lin, B. J. (2013). A survey of DEA applications. *Omega, 41*(5), 893–902.
Lothgren, M., & Tambour, M. (1999). Bootstrapping the data envelopment analysis Malmquist productivity index. *Applied Economics, 31*(4), 417–425.
McDonald, J. (2009). Using least squares and tobit in second stage DEA efficiency analyses. *European Journal of Operational Research, 197*(2), 792–798.
Merigó, J. M., Pedrycz, W., Weber, R., & de la Sotta, C. (2018). Fifty years of Information Sciences: A bibliometric overview. *Information Sciences, 432*, 245–268.
Merkert, R., & Hensher, D. A. (2011). The impact of strategic management and fleet planning on airline efficiency–A random effects Tobit model based on DEA efficiency scores. *Transportation Research Part a: Policy and Practice, 45*(7), 686–695.
Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics, 106*(1), 213–228.
Moradi-Motlagh, A., & Babacan, A. (2015). The impact of the global financial crisis on the efficiency of Australian banks. *Economic Modelling, 46*, 397–406.
Muhuri, P. K., Shukla, A. K., & Abraham, A. (2019). Industry 4.0: A bibliometric analysis and detailed overview. *Engineering Applications of Artificial Intelligence, 78*, 218–235.
Olesen, O. B., & Petersen, N. (1995). Chance constrained efficiency evaluation. *Management Science, 41*(3), 442–457.
Olesen, O. B., & Petersen, N. C. (2016). Stochastic data envelopment analysis—A review [Review]. *European Journal of Operational Research, 251*(1), 2–21.
Park, B. U., Simar, L., & Weiner, C. (2000). FDH efficiency scores from a stochastic point of view. *Econometric Theory, 16*, 855–877.
Peters, H., & Van Raan, A. (1991). Structuring scientific activities by co-author analysis: An expercise on a university faculty level. *Scientometrics, 20*(1), 235–255.

Porembski, M., Breitenstein, K., & Alpar, P. (2005). Visualizing efficiency and reference relations in data envelopment analysis with an application to the branches of a German bank. *Journal of Productivity Analysis, 23*(2), 203–221.
Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation, 25*(4), 348–349.
Rossi, M. A., & Ruzzier, C. A. (2000). On the regulatory application of efficiency measures. *Utilities Policy, 9*(2), 81–92.
Salim, R., Arjomandi, A., & Seufert, J. H. (2016). Does corporate governance affect Australian banks' performance? *Journal of International Financial Markets, Institutions and Money, 43*, 113–125.
Simar, L. (1996). Aspects of statistical analysis in DEA-type frontier models. *Journal of Productivity Analysis, 7*(2), 177–185.
Simar, L. (2003). Detecting outliers in frontier models: A simple approach. *Journal of Productivity Analysis, 20*(3), 391–424.
Simar, L., & Vanhems, A. (2012). Probabilistic characterization of directional distances and their robust versions. *Journal of Econometrics, 166*(2), 342–354.
Simar, L., Vanhems, A., & Wilson, P. W. (2012). Statistical inference for DEA estimators of directional distances. *European Journal of Operational Research, 220*(3), 853–864.
Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science, 44*(1), 49–61.
Simar, L., & Wilson, P. W. (1999). Estimating and bootstrapping Malmquist indices. *European Journal of Operational Research, 115*(3), 459–471.
Simar, L., & Wilson, P. W. (2000a). A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics, 27*(6), 779–802.
Simar, L., & Wilson, P. W. (2000b). Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis, 13*(1), 49–78.
Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics, 136*(1), 31–64.
Simar, L., & Wilson, P. W. (2011a). Inference by the m out of n bootstrap in nonparametric frontier models. *Journal of Productivity Analysis, 36*(1), 33–53.
Simar, L., & Wilson, P. W. (2011b). Two-stage DEA: Caveat emptor. *Journal of Productivity Analysis, 36*(2), 205–218.
Simar, L., & Wilson, P. W. (2015). Statistical Approaches for Non-parametric Frontier Models: A Guided Tour. *International Statistical Review, 83*(1), 77–110.
Simar, L., & Wilson, P. W. (2020). Technical, allocative and overall efficiency: Estimation and inference. *European Journal of Operational Research, 282*(3), 1164–1176.
Simm, J., Besstremyannaya, G., & Simm, M. (2016). Package 'rDEA'.
Tiemann, O., & Schreyögg, J. (2009). Effects of ownership on hospital efficiency in Germany. *Business Research, 2*(2), 115–145.
Tiemann, O., & Schreyögg, J. (2012). Changes in hospital efficiency after privatization. *Health Care Management Science, 15*(4), 310–326.
Tortosa-Ausina, E. (2002). Bank cost efficiency and output specification. *Journal of Productivity Analysis, 18*(3), 199–222.
Türkeli, S., Kemp, R., Huang, B., Bleischwitz, R., & McDowall, W. (2018). Circular economy scientific knowledge in the European Union and China: A bibliometric, network and survey analysis (2006–2016). *Journal of Cleaner Production, 197*, 1244–1261.
Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics, 84*(2), 523–538.
Van Leeuwen, T. (2004). Descriptive versus evaluative bibliometrics. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 373–388). Springer.
Wilson, P. W. (1993). Detecting outliers in deterministic nonparametric frontier models with multiple outputs. *Journal of Business & Economic Statistics, 11*(3), 319–323.
Wilson, P. W. (1995). Detecting influential observations in data envelopment analysis. *Journal of Productivity Analysis, 6*(1), 27–45.
Wilson, P. W. (2008). FEAR: A software package for frontier efficiency analysis with R. *Socio-Economic Planning Sciences, 42*(4), 247–254.
Witte, K. D., & López-Torres, L. (2017). Efficiency in education: A review of literature and a way forward. *Journal of the Operational Research Society, 68*(4), 339–363.
Worthington, A. C., & Lee, B. L. (2008). Efficiency, technology and productivity change in Australian universities, 1998–2003. *Economics of Education Review, 27*(3), 285–298.

Yang, L., & Zhang, X. (2018). Assessing regional eco-efficiency from the perspective of resource, environmental and economic performance in China: A bootstrapping approach in global data envelopment analysis. *Journal of Cleaner Production, 173*, 100–111.

Yeung, W., Goto, T. K., & Leung, W. K. (2017). A bibliometric review of research trends in neuroimaging. *Current Science, 112*(4), 725–734.

Zhu, J. (2014). *DEAFrontier software*. In *Quantitative models for performance evaluation and benchmarking* (pp. 399–407). Cham: Springer.