Applications of DEA and SFA in benchmarking studies in forestry
state-of-the-art and future directions
Strange, Niels; Bogetoft, Peter; Aalmo, Giovanna Ottaviani; Talbot, Bruce; Holt, Anders Holm; Astrup, Rasmus

Published in: International Journal of Forest Engineering

DOI: 10.1080/14942119.2021.1914809

Publication date: 2021

Document version
Publisher's PDF, also known as Version of record

Document license: CC BY-NC-ND

Citation for published version (APA):
Applications of DEA and SFA in benchmarking studies in forestry: state-of-the-art and future directions

Niels Strange, Peter Bogetoft, Giovanna Ottaviani Aalmo, Bruce Talbot, Anders Holm Holt, and Rasmus Astrup

*Department of Food and Resource Economics, University of Copenhagen, Copenhagen, Denmark; †Department of Economics, Copenhagen Business School, Frederiksberg, Denmark; ‡Division of Food Production and Society, Department of Society and Economy, Norwegian Institute for Bioeconomy Research (NIBIO), Norway; §Stellenbosch University, Matieland, Stellenbosch, South Africa; ‡Skovdyrforeningen Syd, Barkop, Denmark; ‡Division of Forest and Forest Resources, National Forest Inventory, Norwegian Institute for Bioeconomy Research (NIBIO), Norway

ABSTRACT

The forestry sector is constantly looking for ways to make data-driven decisions and improving efficiency. The application of Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) allow the users to go beyond looking at simple key performance indicators. Benchmarking is one of the most common tools in business for improving efficiency and competitiveness. This study searched for benchmarking studies in Web of Science until December 2020. It reviewed 56 benchmarking studies in forestry and discusses the potential advantages of using benchmarking in forestry. More than 80% of the studies apply DEA. This review found that almost half of the benchmarking studies in forestry have attempted to estimate the efficiency of forest management organizations at regional scale, mostly being public or state-owned forest districts. A bit more than one-third of the studies have focused on benchmarking forest industries and one-fifth, benchmarking of forest operations. Forest management organizations mainly applied benchmarking for internal comparison and forest industries entirely focused on competitive benchmarking. Surprisingly, in most cases the studies do not necessarily overlap geographically with forest rich countries (e.g., Russian Federation or Brazil). A number of studies address multiple criteria. The future potential for applying automatic data transfer from harvest machines to interactive benchmarking systems are discussed. Finally, the paper discusses the advantages and weaknesses of benchmarking and future research on improving usefulness and usability of benchmarking in forest businesses.

Introduction

Improving performance and efficiency is a fundamental goal across all sectors including forestry. Traditionally, key performance indicators (KPIs) have been applied in the forestry sector to track business performance and efficiency (Rafiei et al. 2014; Proto and Zimbattali 2016). Depending on the goal, specific KPIs are often expressed as the ratio of an output to an input where naturally the higher KPI the better. Each KPI is then a partial measure of performance. However, KPIs have limitations because they are based on several explicit assumptions, e.g., input and output can be scaled proportionally (i.e., constant returns-to-scale). This means that production does not depend on the size of the production unit, and that a small production unit can generate the same income per unit produced as do larger units. Furthermore, KPIs usually do not capture relationships associated with the substitution of inputs and outputs. In order to address this limitation of KPIs, other benchmarking approaches have been developed. In modern benchmarking the two main approaches are Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA). DEA has the advantage that it is non-parametric, it does not require a functional form specification and it is easy to compute using linear programming. Furthermore, it directly allows for learning from dominating peers. SFA usually requires a parametric functional form specification and distributional assumptions on the noise and inefficiency term, but has the advantage that it naturally handles noise in the data (Kumbhakar and Lovell 2003; Bogetoft and Otto 2011). The potential of such applications of benchmarking has so far merely been discussed in the research community (Aalmo and Baardsen 2015; Wang et al. 2019; Obi and Visser 2020; Aalmo et al. 2020).

The main aim of this paper is to review the applications of DEA and SFA for benchmarking in the forestry sector. Earlier literature reviews on their applications can be found in Salehirad and Sowlati (2005) and Sowlati (2005). Since then a number of studies applying SFA and DEA in the forestry sector have been published. The review is followed by a discussion of the potential for operators, machine owners and forest owners to make use of the vast amount of data modern harvesters store from harvest operations for benchmarking. However, most often benchmarking analysis is made on a static data set not allowing for more frequent updating of the data which forms the production set (Agrell and Bogetoft 2017). Therefore, we discuss the potential of automated data
capture from machines, online monitoring of productivity and interactive benchmarking as a tool for improving productivity in forest operations.

**Benchmarking with DEA and SFA**

Benchmarking has been applied as a management tool to improve performance by identifying and applying best documented practices. Managers compare the performance of their respective organizations, products and processes externally with competitors and best-in-class companies and internally with other operations within their own firms that perform similar activities.

In contrast to KPI’s, the starting point of benchmarking is the “production process” describing the physical relationship between resources (called “inputs”) and products/outcomes (called “outputs”). The granularity of this specification can be very high (i.e. all inputs contribute in production of all outputs), but can also be more refined e.g., (i) one can specify different sub-processes with their own process-specific inputs, shared inputs and (intermediate) outputs; (ii) one can specify an output-specific production process for every output with its own output-specific inputs and joint inputs. The granularity of this specification mostly depends on the information available to the analyst, who also needs to consider a sufficient level of discrimination (Dyson et al. 2001).

Together, the inputs and outputs form a production set on which a number of plausible assumptions in the form of mathematical axioms are postulated. Together with the observations, these assumptions serve to approximate the production set. The efficient frontier of the defined production set represents all the “technically efficient” input-output combinations. This frontier gives the maximal outputs that can be produced for some given underlying inputs, or the minimum input required to produce a given output (Bogetoft 2012, p. 8). One of the first DEA models (named DEA-CCR), suggested by Charnes et al. (1978) assumes a constant returns to scale between inputs and outputs. A second and frequently applied model is the DEA-BCC model, which assumes variable returns to scale (Banker et al. 1984). Often technical efficiency from DEA-CCR and DEA-BCC models are applied to estimate the scale inefficiency due to increasing or decreasing returns to scale (Cooper et al. 2000). The actual benchmarking is done by comparing the observation under evaluation against the efficient frontier by measuring the distance between the evaluated observation and the frontier using a distance function. Depending on the specification, this distance function effectively gauges for potential improvements in (i) inputs keeping outputs constant, (ii) outputs keeping inputs constant or (iii) both. Approximation of the specific production set can be done using either DEA or SFA. Productivity may change over time and to capture such changes in efficiency analysis the most popular method has been the Malmquist index, which measures how much a firm improves from one period to the next (Färe et al. 1994).

**Materials and methods**

Drawing upon the literature and practice of systematic reviews in social sciences (Gough et al. 2012; Petticrew and Roberts 2012), a systematic review methodology was applied to identify and collect relevant studies. The systematic review process consisted of identification of scope of the review and research questions, development of the search strategy, literature search and selection of relevant publications and data extraction and synthesis. The scoping focused on identifying all applications of DEA and SFA studies in forestry. Based on this we developed a search for relevant DEA and SFA studies in Web of Science (December 2020) using Term Searched = TS = (“forest* OR wood OR sawmill OR “wood industr*” OR ”forest sector”*) AND (“stochastic frontier analysis” OR SFA OR DEA OR ”data envelopment analysis”) as keywords. This stage of the search revealed 360 references, and we extracted a final set of references by excluding conference proceedings, theoretical or methods papers without case applications, and papers mentioning DEA or SFA studies but not implementing an efficiency analysis on a forestry relevant case. We synthesized these studies’ by their general characteristics, country case, methodological approaches and data. Furthermore, the studies were assigned a benchmarking type according to whether they were performance, process or strategic benchmarking, and if they were applied for internal or competitive benchmarking (Andersen and Pettersen 1996, p. 7). Performance benchmarking determines how well a company compares to others. Process benchmarking is a comparison of methods and practices to learn and improve a company’s processes. Strategic benchmarking is a comparison of the strategic choices and disposition made by other companies to improve a company’s own strategic planning. Internal benchmarking is a comparison within the same company or organization, and competitive benchmarking compares a company with competitors (ibid.).

**Results**

The final selection included 56 studies that were classified as to whether they applied DEA, SFA or both, together with a brief description of the subject they analyzed. Furthermore, we noted the returns to scale assumptions of the model (technology), efficiency measures, benchmarking types, inputs, outputs, country case, and data sources (see Figure 1 and Supplementary Information Table T1).

**Case countries and focus area of the benchmarking**

We identified the case countries and categorized the studies according to the focus area of the benchmarking.

DEA and SFA studies in forestry have been conducted in many countries. However, the most numerous investigations have been applied in United States, Canada, Taiwan, China, and Norway. The results indicate that the applications of benchmarking approaches such as DEA and SFA in forestry are concentrated on a few countries (Table 1). Most of the studies, like Kao and Yang (1991) and LeBel (1996) verify the viability of DEA within forestry. The study subjects range from the operational assessment of forest workers’ performance to the efficiency of sawmills and forest management organizations. These studies use DEA or SFA as methods to measure efficiency and establish a foundation for comparison. The
efficiency assessment becomes the first step for subsequent analysis. Examples include the identification of potential mergers (Bogetoft et al. 2003), the identification of scale efficiency (Nyrud and Bergseng 2002) and statistical analysis and verification (Viitala and Hänninen 1998). To be able to better distinguish between these studies, they have been categorized according to the subjects they address, which is the efficiency estimation of three main focus areas: benchmarking of forest management organizations, benchmarking of forest industries, and benchmarking of forest operations (Figure 1, panel A). More than 46% of the 56 studies attempted to estimate the efficiency of forest management organizations, with the majority addressing public or state-owned forest districts (Korkmaz 2011; Lin and Ge 2019; Zadmirzaei et al. 2019). A relative large number of studies (36%) have investigated the application of DEA within the forest industry. These studies have been conducted in a highly competitive environment, with marginal profit margins (Salehirad and Sowlati 2005). The primary focus of these studies was the productivity of individual pulp and paper plants and sawmills that produced a range of different products. For such reasons a narrow focus on input and outputs, which best describe a cost efficient production, was applied. For example Yin (1998) investigated the application of DEA to measure the productivity efficiency of 44 mills. This study found that most of the mills were technically efficient, but not allocative efficient. Similar performance studies were conducted by Nyrud and Bergseng (2002) and Nyrud and Baardsen (2003), each using DEA as a part of an analysis to estimate productivity growth in technical and scale efficiency over multiple periods for a large number of sawmills. Finally,
a study conducted by Penfield et al. (2014) combined DEA analysis with a survey of 15 sawmills, where the DEA was used to verify the results of the questionnaire. Approximately 18% of the 56 studies estimated the efficiency of operational forest management and logging. A few studies have applied DEA to investigate forest contractors’ efficiency (LeBel 1996; Obi and Visser 2018). LeBel (1996) made a very comprehensive study of 23 competing logging contractors, applying DEA to evaluate annual performance over 6 years. They also studied the factors determining efficiency, like how local conditions affected productivity. Šporčić et al. (2009b) applied DEA to research the maintenance of 13 mechanized forestry units that executed forest operations in the Croatian State Forest. They assessed the mechanized units’ efficiency to operate when waste (waste tyres, solid waste, and waste oil) from the mechanized units was taken into account, along with financial measures. Their results show potential for the different mechanized units to minimize waste compared to the best performers. They concluded that DEA was applicable in a context where the ecological and financial efficiency of organizational units were aggregated. This approach could be used as the first part of a process to identify the best performing units, while another conclusion was that the method was highly sensitive to outliers and that DEA can identify best performance, but not provide the answers. More than 81% of the 56 studies applied DEA, 14% SFA, and the remaining 5% both approaches for studying efficiency (Figure 1, panel B).

Efficiency measures

The focus of the majority of the studies is on evaluating technical efficiency (Figure 1, panel C). Almost 20 studies investigated scale efficiency (Viitala and Hänninen 1998; Yin 1998, 2000; Nyrud and Bergseng 2002; Bogetoft et al. 2003; Salehirad and Sowlati 2005). For example, Kao and Yang (1991) estimated the technical and scale efficiency of 13 national forest districts based on their ability to produce wood, their recreational value and their ability to prevent soil erosion. They showed that, by disaggregating efficiency into technical and scale efficiency, forest managers can identify inefficiency and propose policies that improve efficiency of the forest districts. Martinc et al. (2009) conducted a similar study using a much larger dataset. The authors grouped the forest offices by basic structural characteristics, like region, number of hectares and growing stock. Results showed clear differences between technical and scale efficiency amongst groupings. Shiba (1997) also conducted a technical efficiency study of a forest owners association. The author made three partial models, each using several combinations of inputs and outputs, to test their effect on efficiency score. This study concluded that a key component in DEA is to choose the right inputs and outputs that best describe efficiency, and that DEA allows for the identification of Decision Making Units (DMUs) with the highest efficiency scores to gain further information about how to achieve the same performance. The same conclusions were reached by Korkmaz (2011) by applying three partial models, and a fourth model that included all inputs and outputs. Viitala and Hänninen (1998) conducted similar studies of 19 public forestry boards within the same organization. Because the managers in each forestry board could allocate resources freely to each input, the estimation of efficiency and scale efficiency where done so that each input that was specified in costs and regarded as an activity had its own efficiency and scale efficiency score. Their conclusion was that because of increasing returns to scale, it would be feasible to merge some of the units to improve efficiency. Furthermore, they concluded that DEA proved useful but that it is prone to extreme values, so that verification of results is necessary. Similar studies assessing organizational efficiency and scale efficiency for eventual merger propositions based on increasing returns to scale were done by Kao and Yang (1991). Bogetoft et al. (2003) estimated the technical efficiency of 14 private forests cooperatives. Their efficiency assessment led them to propose a potential merger of specific districts based upon further decomposition of the efficiency measure, so that increases in efficiency can be based upon changes in returns to scale, technology and increased knowledge about the organization.

LeBel (1996) applied cost information for the input and could compute the cost efficiency. Other studies that considered contractors applied SFA to estimate technical efficiency, such as Aalmo and Baardsen (2015) and Bonds and Hughes (2007). Aalmo and Baardsen (2015) conducted an operational study that used the SFA approach to estimate the technical efficiency of 11 harvesting crews, to identify sources of inefficiency among environmental parameters such as terrain roughness, steepness, weather conditions and snow-depth that influenced the productivity of steep terrain harvesting. Their conclusion was that the SFA method could be successfully applied to measure inefficiencies in these logging operations. Carter and Cubbage (1995) conducted a related study where SFA was used to estimate technical efficiency and sources of inefficiency within the Southern U.S. pulpwod timber harvesting industry by assessing efficiency of forest
contractors on a crew level. In the forest management organization studies where the efficiency of forest districts was assessed, technical efficiency was the main efficiency measure applied, one example is the study by Šporčić (2009a). In studies assessing the efficiency of individual sawmills, technical efficiency was the main efficiency measure applied (Nyrud and Bergseng 2002; Penfield et al. 2014). Yin (1998, 2000) were able to obtain technical and cost information to decompose efficiency into technical and allocative efficiency. As such, these studies were able to estimate the cost and allocative efficiency of different sawmills. We found it is rather common to apply panel data to study productivity change over time. More than 16 studies estimate and compare the efficiency scores of several years or periods (Nyrud and Bergseng 2002; Bogetoft et al. 2003; Nyrud and Baardsen 2003; Vahid and Sowlati 2007; Helvoigt and Adams 2008, 2009; Kao 2013; Zadmirzai et al. 2015; Obi and Visser 2017a; Xiong et al. 2018; Zhang et al. 2018; Kropivšek and Grošelj 2019; Shahi and Dia 2019; Wang et al. 2019; Chen et al. 2020; Toma et al. 2020). Six studies applied Malmquist indices to evaluate the changes in productivity over time (Kao 2000; Sowlati and Vahid 2006; Salehirad and Sowlati 2007; Lin and Ge 2019; Kropivšek et al. 2019; Obi and Visser 2020). To reduce the risk of using extreme values from a particular year some studies use data from several years to estimate average values (Kovalčík 2018, 2020).

Returns to scale

Depending on the scope of the intended research, studies in the forestry sector have proposed different approaches for applying DEA and SFA. All of the DEA studies applied the Constant Returns to Scale (CRS) and Variable Returns to Scale (VRS) technologies, with the exception of Shiba (1997) and Penfield et al. (2014), which only applied one of them. The efficiency measures produced by the CRS and VRS technologies are often compared in the studies and different types of sensitivity analyses are conducted to verify and compare the efficiency from these two technologies, as in Bogetoft et al. (2003). The choice of scale assumption is crucial in setting up a model.

Choice of DMUs and inputs and outputs

As mentioned by Bogetoft (2012), a major issue to consider when applying DEA is the choice of DMUs and appropriate inputs and outputs. These must have relevance and the DMU should have a certain managerial control. An example of this is Viitala and Hänninen (1998), where each manager on the forest board could distribute resources freely amongst activities (outputs). In this example, each forest board is regarded as a DMU. The same can be seen in the sawmill studies, where each sawmill is considered a DMU (Yin 1998). Another common approach is to pool data for the same districts or units over a number of years, so the number of DMUs increases (Diaz-Balteiro et al. 2006; Kovalčík 2020; Gutiérrez and Lozano 2020). Bogetoft et al. (2003) and LeBel (1996) pooled data over 3 and 6 years, respectively, so that a contractor’s annual performance represented a DMU. While there is no fixed rule, an often encountered rule-of-thumb is that the number of DMUs relative to 3 times the sum of input and output variables should be at least 1 (Dyson et al. 2001). It was found that more than 72% of the studies complied with this criterion (Figure 1, panel D).

Benchmarking categories

According to the classification developed by Andersen and Pettersen (1996), every study can be categorized as performance benchmarking because they all assess efficiency, to identify the best performers. Studies of the sawmill industry can be categorized as competitive benchmarking, as data from several companies was included in each study. It was not possible to categorize any of the studies as process benchmarking or strategic benchmarking. The majority (81%) of forest management organization studies can be characterized as internal benchmarking, because the different units are compared within the same organization (Figure 1, panel E). All studies concerned government or public organizations, with the exception of Bogetoft et al. (2003), which investigated a private cooperative. Data used in these forest management organizations studies was mostly obtained directly from forest bureaus and district offices (Shiba 1997; Kao 2000; Korkmaz 2011). LeBel’s (1996) study of forest contractors can be characterized as competitor benchmarking. Besides the assessment of performance, field inspections were conducted to assess equipment and the working process, like the sorting of timber. According to Andersen and Pettersen (1996) framework, this study could potentially also be categorized as process benchmarking. All benchmarking studies of forest industries were assessed as competitive. Benchmarking studies on forest operations were assessed as 40% internal and 60% competitive (Figure 1, panel E).

Information sources

Data applied in the forest management organization studies was primarily based on data from forest districts, while the forest industrial studies mainly applied data from national statistics (Yin 1998; Nyrud and Baardsen 2003), consultancies or directly from sawmills (Penfield et al. 2014). The data obtained for the forest contractors came from harvest operations. The number of inputs and outputs ranged from two to more than nine. Data entered as cost or revenues or biophysical quantities (e.g., volume of forest various forest products, volume stock, numbers of workers). A limited number of studies include environmental variables which may affect productivity (Aalmo and Baardsen 2015; Aalmo et al. 2020) (see Supplementary Information Table S1).

Discussion

General discussion of review results

Focus of the benchmarking studies

The forestry sector strives for increased efficiency in production and environmental performance. It is crucial for the sector that it is continuously able to increase competitiveness with reduced costs and lower environmental impacts. One approach to understand what drives production efficiency is applying
advanced benchmarking. The current study reviews the applications of Data Envelopment Analysis and Stochastic Frontier Analysis in forestry. More than 81% of the studies apply DEA only, 5% apply SFA only, and 14% both DEA and SFA. Therefore, the scope of this discussion is mainly focused on DEA. We found that the majority of studies (about half) have investigated the efficiency of forest management organizations at a regional or national scale with the main part addressing public or state-owned forest districts. More than one-third of the studies focused on benchmarking forest industries and one fifth benchmarked forest operations. This may also reflect that benchmarking is one of the most common tools applied by business for comparing their competitiveness and making strategic decisions about organizational development (Bain & Company 2018). We found that the number of studies focusing on internal benchmarking was comparable to the number of studies focusing on competitive benchmarking, but it varied between three categories of focus areas of benchmarking. Forest management organization studies addressed internal benchmarking more frequently in evaluating the performance within their organization or company. Benchmarking of forest industries was often framed as competitive benchmarking, as data from several companies was included in each study. Internal benchmarking is useful for initiating a process of identifying, sharing, and using the knowledge and practices inside an organization and to learn how the organization may improve its performance (O’Dell and Grayson 1998). Competitive (external) benchmarking is important for managers and boards to compare the performance of their respective organizations with competitors and best-in-class companies. Selection criteria for when to implement either have been discussed by Southard and Parente (2007).

DEA and SFA applications in forestry compared to other domains
We found 56 applications of DEA or SFA in the forestry domain. However, a similar Web of Science search, which only included the search term TS = (“stochastic frontier analysis” OR SFA OR DEA OR ”data envelopment analysis”) resulted in 33,113 records. The major categories included operations research and management science (4657 records), economics (4268 records), management (4107 records) and environmental science (2565 records). This may indicate that benchmarking studies within the forestry sector are less frequent than in other sectors and research domains.

We found 56 studies, which complied with our inclusion criteria. However, a total of 96 studies (of 33,113 records) in Web of Science, which use the terms of stochastic frontier analysis, data envelopment analysis, SFA OR DEA have been classified as belonging to the forestry research field. The difference is mainly caused by the inclusion criteria that the study should include an application of DEA or SFA on forest-related data. Not surprisingly, research fields such as business economics (25.8%), engineering (17.3%), operations research management science (14.2%), and computer science (9.5%) dominate the literature. Interestingly, agriculture represents 6% of all studies compared to only 0.3% presenting evaluations in forestry. A recent study on dynamic DEA found that most studies address efficiency analysis in the agriculture and farming sector, followed by banking and energy sectors (Mariz et al. 2018). This may indicate that the uptake of modern benchmarking may also be higher in the agricultural sector than in the forestry sector. Examples document that benchmarking has been applied extensively in agricultural practice. One such example is the UK Farm Business Survey (farmbusinesssurvey.co.uk) which provides online information on the physical and economic performance of farm businesses in England and Wales. Farmbench (ahdb.org.uk/farmbench) is another online tool which farmers can anonymously sign up to and compare key performance indicators (KPI) with neighboring, local or national farmers.

Geographical distribution of studies in forestry and other domains
The geographical distribution of the 33,113 records is presented in Table 2. Comparing this with Table 1 we see overlaps with some countries, e.g. United States, China, Canada, and Iran. However, it is also clear that some continents and a lot of countries are not covered in the forestry literature. Missing countries include, e.g., the Russian Federation and Brazil, which account for more than 32% of the world’s forest (FAO 2020) and a forestry sector, which is important for their national economies. This could also indicate a mismatch between supply of benchmarking studies and demand for improving efficiency analysis in the forestry sector. Such geographical mismatches have been studied in other research domains, e.g. climate change (Pasgaard and Strange 2013; Pasgaard et al. 2015). It should be noted that the Web of Science country classification is based on a search in the affiliation field. This means for some of the records the author addresses may not reflect the case country of the research. However, all 56 applications of DEA or SFA in forestry did include author addresses from the case country. We applied English search terms, which may cause a language bias and lead to an underestimation of publication written in non-English languages.

Identifying relevant input, output, and environmental variables
The studies presented a large range of input and output variables as well as environmental variables. Selecting the best possible variables and aggregation of data may be facilitated through discussions with various stakeholders who understand the relevance of various decision-making units (DMUs) and the production process one wants to model. Expert knowledge may also be necessary to select potential useful variables for the benchmarking analysis. DMUs can be identified at multiple levels: from a highly aggregated level (e.g., harvester operator organization) to a very low level (e.g., individual harvester operators). Furthermore, the DMUs must be comparable in the sense that they perform similar tasks and transform the same types of inputs to the same types of outputs, cf. also Golany and Roll (1989).

Ideally, the selection of input and output variables should characterize the technology, which describes the production. However, sometimes the criteria for the selection of inputs and outputs are primarily based on availability of data in databases
and relevant literature (Obi and Visser 2017b). Bogetoft (2012) recommends that the selection should be comprehensive and suggest organizing the variables into a hierarchy, where higher at the top we have the overall inputs and outputs level variables reflecting the properties of lower level ones. If only lower level variables are available, one should assure that all facets of the overall variables are taken into account when selecting candidate variables. Generally, the studies apply inputs (e.g., labor, machine time, fuel) which represent the resources used in the production process and outputs (e.g., cubic meters of harvested wood of different assortments) which represent the resulting outcomes. The underlying assumption is that more inputs lead to more outputs and thus a positive correlation between inputs and outputs. Environmental variables (e.g., terrain ruggedness (Aalmo et al. 2020)) represent operating conditions that directly affect the production process, but which are beyond the control of the DMUs. In the modeling, environmental variables can sometimes be dealt with as non-controllable inputs (if this facilitates the production) or outputs (if they complicate the production), while in other cases one must specifically account for these environmental conditions via the potential peer units such that DMUs are only benchmarked against other DMUs operating in similar or worse environments. Related studies have applied DEA to investigate the environmental efficiency of the forestry sector, recognizing that forest production also involves the production of environmental externalities (Lin and Ge 2019).

There are a number of issues to consider in the selection of variables: (i) the number of variables relative to the number of DMUs; (ii) factor measurement and (iii) correlation of variables. The first issue is particularly relevant in DEA. If the number of inputs and outputs is large relative to the number of DMUs then DEA may lose discriminatory power. We found when applying a rule of thumb that the ratio between number of DMUs and 3 times the sum of inputs and outputs should be higher than 1 (Dyson et al. 2001), that more than 72% of the studies did meet this criteria. If the ratio is smaller than 1 the risk is that the discriminatory power is too small. Some of the studies suggest alternative procedures to overcome too few DMUs, e.g., by using overlapping windows of data and window analysis (Helvoigt and Adams 2008; Kropivšek and Grošelj 2019). A number of studies emphasized the need to ensure a sufficient number of DMUs compared to the number of input and output variables (Diaz-Balteiro et al. 2006; Šporčić et al. 2009a). The second issue is that one must be careful not to mix indices, percentages, or ratios with absolute numbers. Improper mixing can result in incorrect efficiency results. All studies complied with this criteria except Kropivšek and Grošelj (2019) which uses leverage and liquidity ratios mixed with e.g. size of assets, capital, number of employees, net sales, gross profit, and gross value added. Third, in an effort to reduce the number of variables in order to increase discriminatory power one might be tempted to remove highly correlated inputs (outputs). Dyson et al. (2001) warn against this practice as removing highly correlated inputs (outputs) can substantially alter the efficiency analysis. One can only safely remove one of two highly correlated inputs (outputs) if one is a scalar multiple of the other.

### Selection of DEA models

An important property of a benchmarking approach is its ability to reflect and respect the characteristics of the technology that we aim at modeling (Bogetoft 2012). One weakness of DEA is that it does not include any statistical tests to help the user to select the most appropriate model. Use of stakeholders and expert knowledge may be insufficient to select the most robust models. It is generally recommended to use a flexible
model in the sense that its structure can adapt to data, and cope with changes in the number of variables and range sizes (including procedures for excluding outliers), covariance between input variables or output variables (Smith 1997). Several accompanying methods have been developed to assist in the identification of more robust models. They include the F-tests, Principal Component Analysis, regression-based tests and bootstrapping approaches which are applied to evaluate the inclusion and exclusion of input/output variables (Pastor et al. 2002). Others point at applying probabilistic/combinatorial approaches that provide DEA efficiency scores which are robust against the selection of inputs/outputs to be included in the model (Landete et al. 2017). None of the studies included in this review report selection procedures and test for robustness.

**Potential for making benchmarking useful in forestry**

This review revealed that the majority of studies used data from either public statistics, reports from forest or company offices, surveys or field plots and harvest operations. While this type of analysis allows the analyst to investigate a range of operational, tactical or strategic decisions, the analysis is based on historical and potentially outdated information. Furthermore, the studies rarely involve industrial partners or report any stakeholder interaction. Only a few examples of the reviewed studies involve authors with affiliations from the private sector parties, e.g., consultancies (Helvoigt and Adams 2008, 2009) or the bank sector (Zhang et al. 2018). Reviewing the 56 studies only revealed that Obi and Visser (2017a, 2017b, 2018, 2020), Shahi and Dia (2019), Aalmo and Baarsden (2015), and Gutiérrez and Lozano (2020) discuss how their studies could provide some useful decision support to forest companies, forest industries policymakers, and general industry stakeholders involved in the measurement and overall improvement of performance. This may indicate that DEA and SFA studies are mainly prepared by specialists with high technological skills. Even though most managers have experience with key performance indicators, DEA and SFA may be difficult to comprehend and implement for managers. The potential gain from using benchmarking is then constrained by both applying outdated information and users/decision makers having insufficient technical and theoretical knowledge to run the analysis on their own. However, taking benchmarking to the next level, the forestry sector may benefit significantly from developing software solutions which combine state-of-the-art benchmarking methods with an easy and intuitive user interface (Bogetoft 2012). Digital and technological progress has created new opportunities for the forestry sector (Müller et al. 2019). Today data on harvested volumes, assortment and tree species distributions, log sizes, fuel consumption and other types of information are stored by on board computers in harvesters. Data that are highly relevant for characterizing the productivity (Olivera et al. 2016) of forest operations. Transmission of such data to databases may allow for day-to-day benchmarking analysis of forest operations, a daily comparison of performance against past operations. Such data capture and integration with modern benchmarking software would make it possible to create learning labs for the machine operator or owner who can explore operational, tactical and strategic decisions (Bogetoft 2012). They could serve as online monitoring and interactive learning tools and allow the user to explore various features of a firm (Post and Spronk 1999; Bogetoft et al. 2006). Interactive benchmarking would allow the user to ask a lot of what-if questions. This may include selecting and testing various data sets, the KPIs of the firm, choosing a model and deciding who to be compared to, identifying peers in the sector or inside the firm, evaluating dynamic development over time, and extracting reports which may form part of the information needed to make long-term strategic decisions. Comparisons to other sectors, e.g., agriculture, and observations of tremendous gains across sectors in efficiency due to digitization document large potentials for advancing modern benchmarking in the forestry sector.

Another potential future application of benchmarking may be to facilitate incentive provision. The operator may also be the machine owner, however often the machine owner may have several operators in the team or sub-contract with other machine owners. Even if a benchmarking proves that efficiency gains could be obtained by the machine operators and this information is part of the operational plan, there are no guarantees that the operators will follow the plan. The operators involved in the plan must also have incentives to implement it, and the incentives must be attractive enough to make the operators participate. The benchmarking literature describes this relationship between the benchmarking model and the motivational aspect as either implicit or explicit. An implicit or informal relationship is found when the mere existence of a model improves behavior because performance now receives more attention in the minds of the operators. A more explicit and formalized relationship is found when the payment to an operator is made a direct function of the benchmarking results (Bogetoft 2012). Benchmarking and so-called yardstick competition has been suggested for regulating such asymmetric incentive structures (Agrell et al. 2005, 2007), mostly for regulating large monopolistic companies. However, it may be relevant for future work to address how this and other benchmarking approaches could be applied for increasing competition and efficiency within and between organizations or firms in the forestry sector.

**Conclusion**

The efficiency of the forestry sector has increased for the last 50 years, mainly as a result of technological development and mechanization of processes. Benchmarking may be an important part of exploring new opportunities for increasing efficiency in the forestry sector. This review found that almost half of the benchmarking studies in forestry have attempted to estimate the efficiency of forest management organizations at a regional scale, mostly being public or state-owned forest districts. A bit more than one-third of the studies have focused on benchmarking forest industries and one-fifth benchmarking of forest operations. Surprisingly, the studies are in most cases not geographically overlapping with forest rich countries. We also found that almost half of the studies evaluate the productivity change over time. Most often, the productivity data is based on annual reports or longer periods (months to several years). We suggest that the forestry sector may benefit further from the potential advantages of digitization and use of interactive benchmarking based on frequently updated
performance data automatically captured from harvesters and forwarders. The challenge is to establish incentives and make the benchmarking tools user-friendly and interesting learning labs for the potential users in the forestry sector.

Acknowledgements
This work has received funding from the Bio-Based Industries Joint Undertaking under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 720757.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Funding
This work was supported by the Bio Based Industries Joint Undertaking under the European Union’s Horizon 2020 Research and Innovation Programme [720757]; Bio Based Industries Joint Undertaking under the European Union’s Horizon 2020 Research and Innovation Programme [720757]; Bio Based Industries Joint Undertaking under the European Union’s Horizon 2020 Research and Innovation Programme [720757]; Bio Based Industries Joint Undertaking under the European Union’s Horizon 2020 Research and Innovation Programme [720757]; Bio Based Industries Joint Undertaking under the European Union’s Horizon 2020 Research and Innovation Programme [720757].

ORCID
Niels Strange http://orcid.org/0000-0002-8842-8022
Peter Bogetoft http://orcid.org/0000-0002-2173-2791
Bruce Talbot http://orcid.org/0000-0003-1935-5429

References