Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 200 (2022) 1451–1460

Procedia Computer Science

www.elsevier.com/locate/procedia

3rd International Conference on Industry 4.0 and Smart Manufacturing

Workers benchmarking using multi-directional efficiency analysis in a manufacturing production system

Eugénio M. Rocha[a,b,*], Ângela F. Brochado[c], Ana Moura[c,d]

[a]Department of Mathematics, University of Aveiro, 3810-193, Portugal
[b]Center for Research and Development in Mathematics and Applications (CIDMA), Aveiro, 3810-193, Portugal
[c]Department of Economics, Management, Industrial Engineering and Tourism, University of Aveiro, 3810-193, Portugal
[d]Systems for Decision Support Research Group (GOVCOPP), Aveiro, 3810-193, Portugal

Abstract

The human factor plays a relevant role in all manual or partially automatic production systems, specially, the ones showing reliable and balanced dynamics. In the literature, parametric or survey-based models are quite common for performance evaluation of production workers. In this work, multi-directional efficiency analysis is used instead, for root cause analysis of product reworks and bottlenecks occurrence, according to four worker-related parameters: experience time, wage, delay time and response time. The approach allows to identify individual inefficiencies per tuple worker/working shift and to cluster them according to similar inefficiency parameters. In addition, this work opens a path to new applications of multi-directional efficiency analysis to problems in the manufacturing industry.

Peer-review under responsibility of the scientific committee of the 3rd International Conference on Industry 4.0 and Smart Manufacturing

Keywords: multi-directional efficiency analysis; bottleneck; workers performance clustering; manufacturing production system

1. Introduction

Production throughput is highly influenced by different types of direct/indirect inputs: materials/parts in processing and/or in queue, workstations, workers, among others. Therefore, it is important to regularly measure and quantify poor performances in a production system, taking into account that its sole measurement is not enough. According to [1], technical inefficiencies can be interpreted as a result of a lack of motivation or effort of workers. On the other

* Corresponding author. Tel.: +351-234-370-359 ; fax: +351-234-382-014.
E-mail address: eugenio@ua.pt

1877-0509 © 2022 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 3rd International Conference on Industry 4.0 and Smart Manufacturing
10.1016/j.procs.2022.01.346

hand, [2] perceived the term differently, considering that it could result from a lack of knowledge or managerial ability, where improvements could be obtained through learning processes.

The evaluation of a worker's performance is a regular assessment procedure that some companies conduct, not only for effective job evaluation but also to allow workers to set expectations and measure their success. The information gathered can help decision making in regards to pay raises, promotions, bonuses, but also work assignments or work tasks improvement, as well as general quality enhancement. Performance evaluations should be judged in alignment with specific goals using defined metrics. However, many times they may show subjectivity and even some controversy. Several studies exhibit different strategies and metrics to better understand job performance [3]. Some strategies combine what is done in reality with what could be done in ideal circumstances. Other studies try to distinguish between maximal and typical performance, by examining the effects of workers' differences on the maximal–typical performance relationship [4, 5, 6].

Productivity in general or throughput can be directly associated with the occurrence of operational bottlenecks or/and reworks. An operational bottleneck is anything that diminishes productivity on the factory floor, and a rework or a re-processed part is a product that had to be re-worked in a workstation because of incorrect or incomplete processing. In parts manufacturing and assembly, workers tend to be the rate-limiting factor across various steps, meaning that throughput is directly affected by their behavior. So, managing bottlenecks is quite often a matter of managing people – through appropriate staffing and task leveling. Most authors associate task execution delays, in production lines, with variability. In a production process, variability is one of the primary causes for bottleneck occurrence, line efficiency degradation and inventory excess [7]. In line with the Industry 4.0 topic, the most effective way to determine the causes for throughput decrease, and so bottleneck occurrence, is via real-time monitoring and post-evaluation of production data by the integration with a manufacturing execution system (MES) [8].

Considering all aspects above, in this work, it is presented a novel contribution of a benchmark analysis of worker's performance, using multi-directional efficiency analysis (MEA), by linking bottlenecks and reworks occurrence (outputs) with the effect of real worker's resources (inputs). Therefore, using a data-driven approach, it was possible to benchmark the worker's associated inputs such as wage (euros), experience time (hours), response time (seconds) and delay time index, with production outputs as bottleneck information and number of re-processed parts. The delay time index is a penalization value that measures the lapse between the planned shift start time and the time a worker effectively started working, by using a Gaussian function as weight. The response time measures the average time the worker takes to start a job in a workstation compared to the time the part/product is available to be processed.

Plenty of benchmarking research has been supported by the use of data envelopment analysis techniques (DEA). However, for this study, it was intended to use a more tailored and detailed method, to not only study efficiency outcomes, but also inefficiencies, which is one of the strengths of MEA. Additionally, up to our knowledge, bottleneck information has not yet been used to benchmark worker's performance, contributing with another innovative aspect.

In terms of structure, the paper is organized as follows. In Section 2, some clarification is provided about the main topics addressed in the study: the importance of workers' performance evaluation, as well as conducting benchmark studies. In sub-section 2.2., a bottleneck identification model for a manufacturing line is explained, where the worker's influence is now added as complementary variable of the human-machine interaction in the context of bottleneck identification (an extension from previous work [9]). Section 3 presents the Methodology used (multi-directional efficiency analysis) for the performance evaluation study, providing, additionally, an explanation about the input and output variables considered for the model. Section 4 shows the discussion of the results, and Section 5 closes the present study addressing the conclusions and next steps for future investigation.

2. Definitions and concepts

2.1. Benchmarking and performance evaluation

In 1979, Xerox introduced benchmarking as a total quality management tool, that has been broadly adopted by many industries, as the manufacturing industry [10]. According to [11], it has been proven to be effective for continuous improvement achievement in diverse management areas, so as in production refinement. In manufacturing production, benchmarking allows a better understanding of strengths and weaknesses of processes, through a comparative analysis between system variables. It aims to pinpoint, understand and apply the best practices to support and help firms to maximize its performance [12] and minimize inefficiencies .

Often times, the terms *benchmarks* and *benchmarking* get confused [13], so it is important to clarify the difference between them. Benchmarking is defined as an improvement process while benchmark is a performance level [14]. To perform benchmarking, researchers found that some companies focus on setting key performance indicators (KPIs) based on their organizational strategy [15]. The objective is to increase production throughput, and ultimately customer satisfaction through a lower delivery time. To achieve both goals, what usually managers think about first, while looking up to costs and the economic spectrum, is to produce as large as possible a certain output from a given set of inputs. Therefore, companies want to minimize inputs and maximize outputs, in order to obtain not only better throughput but also fewer costs. For an effective benchmarking implementation, adequate tools, system, and IT support are necessary [16]. Therefore, real data coming from a manufacturing execution system (MES) must be used and analysed for a true and real understanding of the system flow.

2.2. Bottleneck identification

There are numerous definitions for bottleneck in the literature [17], in which all of them agree to the fact that it introduces a negative impact on the output of a production system. Therefore, system monitoring and root cause analysis is crucial for bottleneck control and detection, and ultimately throughput improvement. The process of root cause identification, to subsequently eliminate or reduce problems, is crucial to understand if, e.g., in a workplace with a machine-worker interaction, the worker is one of the causes for bottlenecks occurrence, and which aspects of his performance may be enhanced. In Figure 1 (middle layer), a pictorial representation of a typical manufacturing line is shown, in parallel with the corresponding actions happening on the sensoring level (upper layer) and the formal representation level (lower layer).

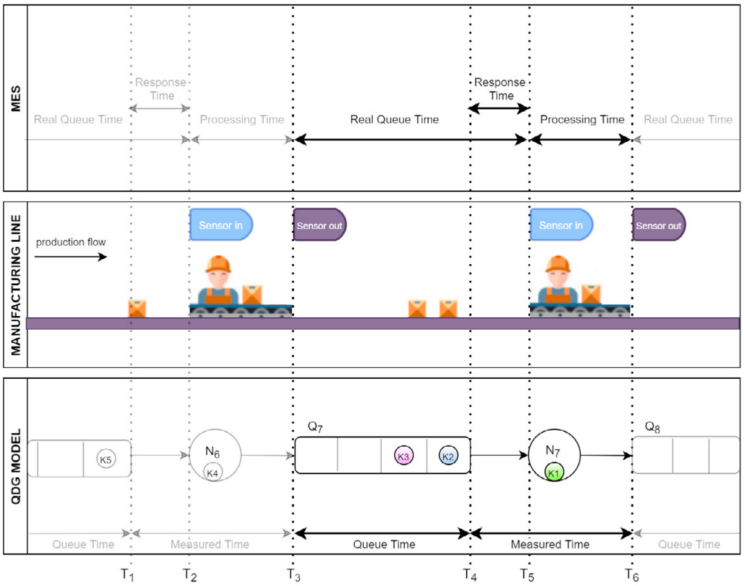


Fig. 1: Manufacturing line and correspondent abstract representation through the Queue Directed Graph (QDG) model, with additional information about relevant metrics.

At the upper layer, a representation of the manufacturing execution system (MES) information is designed in connection with the events occurred in the manufacturing line. The MES is the information system that helps to monitor, supervise, and control processes in the production system. It receives real-time data from the production line which includes machines, sensors, and edge devices. Each workstation has a sensor placed at the entrance (*Sensor in*) and at the exit (*Sensor out*), aimed to acquire data and transmit it to the MES. Based on the collected data, for every part that enters the production line, the MES determines the *Processing Time* in a certain workstation and the *Real*

Queue Time in the correspondent queue. However, the *Real Queue Time* does not represent exclusively the amount of time a part is in a queue. Because workers are responsible for transport and manipulation of parts between queues and workstations, there will always be a small percentage of time associated with human delay in their response. Therefore, according to MES calculations, the *Real Queue Time* is obtained by the sum of the effective queue time and the so-called *Response Time*, represented in Figure 1.

In the past literature, for studies addressing real case studies with companies with a MES, the metrics above were used for bottleneck identification, depending on either queue behavior or workstation behavior. Most recently, [9] developed a data-driven bottleneck identification approach, where bottlenecks were identified based on minimal data information. "Minimal information" is a set of records with an identification (ID) for the location (workstation), an ID for the part, and a timestamp corresponding to the departure time of a part from a workstation (given exclusively by the Sensor out, represented in Figure 1). Based on these records (see Figure 1 (lower layer)), for each part that enters the production line, it is possible to mathematically calculate the timestamps of: (1) queue entry, (2) queue exit; and (3) workstation exit. Physical workstations are abstractly represented by nodes (elements N6 and N7), related queues (elements Q7 and Q8), and products/ parts (elements K1, K2, and K3). From these information, bottleneck measures can be derived as the average active period method (AAPM).

Introduced by [18], AAPM uses the average of the *active* periods of a process and detects fluctuations to it. Formally, this may be defined as $\Psi_{AAPM} = argmax_{i\in\{1,\ldots,s_N-1\}} \Phi_i$ with $\Phi_i = mean(\{\zeta^{node}_{w,i} \in \mathbb{R}^+_0\})$, where $\zeta^{node}_{w,i}$ is the active time of a node i at the instant of time w. Commonly, the notion of *active* periods is associated to the processing times obtained by hardware devices. In fact, AAPM can be seen as a family of methods, where variants are obtained by changing the underline notion of *active* period. We consider the following useful variants: the **average measured period method (AMPM)** where the average is made over the periods of *measuredTime*; and the **average queue period method (AQPM)** where the average is made over the periods of time where the queue is occupied by at least one token (so-called *queueTime*). Such variant will be calculated and used as output variables in the benchmark method.

Taking a closer look at Figure 1, a clear difference can be pinpointed between the Queue Directed Graph (QDG) metrics (*Queue Time* and *Measured Time*) and the MES metrics (*Real Queue Time* and *Processing Time*. In the QDG model, the human component, measured by the *Response Time (RT)* is aggregated with the *Measured Time*. Consequently, this paper extends previous work done in [9], by considering the existence of a worker's response time, which as not yet been considered in the literature by any method of bottleneck identification. So, by setting the difference between the *Measured Time* and the *Processing Time*, it is possible to compute the amount of time a worker spent to "respond" to a part in a queue. Thereafter, the *RT* variable will be further utilized as an input for the present benchmark analysis.

We point out that the approach proposed in this work is general enough to be implemented over any definition of bottleneck, available in the literature as [18, 19], and references wherein.

3. Methodology

3.1. Multi-directional efficiency analysis

The multi-directional efficiency analysis (MEA) method is based on the potential improvements approach proposed by [20], and initially implemented by [21] for efficiency determination of Danish dairy farms. It is a non-parametric approach, belonging to a group of advanced data envelopment analysis (DEA) techniques, which selects benchmarks with the aim of allowing the estimated output to determine the direction of improvement [22]. This means that the MEA model selects both input reduction and output expansion, according to improvements potential, related to each separated input and output of the model. Therefore, a clear advantage of MEA, is the fact that it can simultaneously provide multiple efficiency estimations of input-output [23], [24], providing an analysis about the relative importance or contribution of the input factors for efficiency maximization or inefficiencies minimization.

MEA contrasts with the proportional improvements approach introduced by [2] and better known as DEA, modeled by [25] and later extended by [26]. The DEA approach is a highly used non-parametric methodology which formulates and solves a linear programming problem, to compute the relative efficiency of a set of homogeneous units, named decision making units (DMU). It produces an efficiency score and a target operating point that lies on the efficient frontier for each DMU. This approach is computed in a way that usually it uses the same, or less inputs, to produce identical or more outputs (it requires a fixed consumption mix). These aspects trigger several reasons for selecting MEA over the traditional DEA model. As briefly explained above, one of the reasons lies on the fact that DEA is restricted to the radial or proportional contractions of inputs (or, on the other hand, expansions of outputs [27]). Practically, this means that in a situation where 3 inputs have been used in a certain proportion (e.g. 1:3), the savings will show this same proportion. This aspect comes with a range of procedural concerns to be explored and fixed, including those relating to the homogeneity of the units under evaluation, the input/output set chosen, the measurement of those selected variables and the weights attributed to them [28]. Therefore, the improved MEA methodology determines efficiency/slack for each factor separately, presuming that all other input factors remain unchanged. For this reason, MEA better reflects the fact that there is no 'one-size-fits-all' and so diverse combinations of earns and losses can contribute to superior performance [29].

In what follows, we give a general description of the MEA model used and fixed notation. Let $n = (w, s) \in \mathcal{N}$ be a tuple identifying the worker $w \in W$, and the working shift $s \in S$, which we call a worker/working shift tuple, and $[m]$ denotes the set $\{1, ..., m\}$, for some $m \in \mathbb{N}$. We consider that any given tuple $n \in \mathcal{N}$ produces $J \in \mathbb{N}$ outputs $y_j(n)$, $j \in [J]$, using $I \in \mathbb{N}$ inputs $x_i(n)$, $i \in [I]$, where the first $1 < D \leq I$ inputs are the so-called discretionary inputs, i.e. variables that enter into the optimization process, because the non-discretionary inputs are variables that cannot be changed. Hence, $x(n) \in \mathbb{R}^I$ is the vector of all the inputs and $y(n) \in \mathbb{R}^J$ is the vector of all the outputs, for a given worker/working shift tuple $n \in \mathcal{N}$. Our dataset $Z = \{z(n)\}_{n\in\mathcal{N}}$ is the set of values $z(n) = (x(n), y(n))$ for all $n \in \mathcal{N}$. Considering the variable returns to scale (VRS) model for the efficiency measurement of decision making units, see [30], we define the set $\Lambda^N = \left\{\lambda \in \mathbb{R}^{\mathbb{N}} : \sum_{n=1}^{N} \lambda_n = 1\right\}$.

The MEA score for a specific observation $z(\bar{n}) = (x(\bar{n}), y(\bar{n}))$ is found by solving the following linear optimization programs:

Problem $P_m^{\alpha}(z,\bar{n})$:	Problem $P_j^{\beta}(z,\bar{n})$:	Problem $P^{\gamma}(\alpha,\beta,z,\bar{n})$:
$\min \alpha_m(\bar{n})$ *such that*	$\max \beta_j(\bar{n})$ *such that*	$\max \gamma(\bar{n})$ *such that*
$\sum_n \lambda_n x_m(n) \leq \alpha_m(\bar{n})$,	$\sum_n \lambda_n x_i(n) \leq x_i(\bar{n}),\ i \in [I]$,	$\sum_n \lambda_n x_i(n) \leq x_i(\bar{n}) - \gamma(\bar{n})(x_i(\bar{n}) - \alpha_i^*(\bar{n})), i \in [M]$,
$\sum_n \lambda_n x_i(n) \leq x_i(\bar{n}),\ i \in [I], i \neq m$,	$\sum_n \lambda_n y_s(n) \leq \beta_j(\bar{n}),\ s \in [J]$,	$\sum_n \lambda_n x_i(n) \leq x_i(\bar{n}), i \in [I] \setminus \{m\}$,
$\sum_n \lambda_n y_l(n) \leq y_l(\bar{n}),\ l \in [J]$,	$\sum_n \lambda_n y_l(n) \leq y_l(\bar{n}), l \in [J], l \neq j$,	$\sum_n \lambda_n y_l(n) \geq y_l(\bar{n}) + \gamma(\bar{n})(\beta_l^*(\bar{n}) - y_l(\bar{n})), l \in [L]$,

where $\lambda \in \Lambda^n$, $\alpha_m^*(\bar{n})$ and $\beta_j^*(\bar{n})$ are the optimal solutions to the problems $P_m^{\alpha}(z,\bar{n})$ and $P_j^{\beta}(z,\bar{n})$ respectively. The ideal point of $(x(\bar{n}), y(\bar{n})$ is given by the MEA output vector $\zeta(n) \doteq (\alpha_1^*(n), ..., \alpha_d^*(n), ..., x_I(n), \beta_1^*(n), ..., ,\beta_J^*(n))$. Often, some of the input variables may be discretionary (their values can be changed) and others may be not discretionary (are fixed). From now on, the discretionary variables are represented by the first indices d, $1 < d < I$. Thus, $i \in [D]$ indicates the discretionary inputs and $i \in [I] \setminus d$ indicates the non-discretionary inputs. In this setting, the MEA for a specific observation $z(\bar{n}) = (x(\bar{n}), y(\bar{n}))$ consists of $(|D| + |J| + 1) \times N$ linear programs, which includes one problem $P_d^{\alpha}(z,\bar{n})$ for each discretionary input $d \in [D]$, one problem $P_j^{\beta}(z,\bar{n})$ for each of the output dimensions $j \in [J]$ and one problem $P^{\gamma}(\alpha,\beta,z,\bar{n})$.

For a given dataset $z = \{z(n)\}_{n\in\mathcal{N}}$ the MEA score of each $n \in \mathcal{N}$ is given by

$$MEA_z(n) = \frac{\frac{1}{\gamma^*(n)} - \frac{1}{D}\sum_{i=1}^{D} \frac{x_i(n)-\alpha_i^*(n)}{x_i(n)}}{\frac{1}{\gamma^*(n)} + \frac{1}{J}\sum_{j=1}^{J} \frac{\beta_j^*(n)-y_j(n)}{y_j(n)}}, \tag{1}$$

where $\alpha_i^*(n)$, $\beta_j^*(n)$ and $\gamma^*(n)$ represent the corresponding optimal solutions to the linear optimization problems $P_i^{\alpha}(z,n)$, $P_j^{\beta}(z,n)$ and $P^{\gamma}(z,n,\alpha^*,\beta^*)$.

The MEA score is then obtained by the directional contribution of each input and each output variable. In fact, for the input $i \in [I]$ the contribution in the unit $z(\bar{n})$ is given by

$$\mathrm{mEff}_i(n) = \frac{x_i(n) - \gamma(n)(x_i(n) - \alpha_i^*(n))}{x_i(n)}\chi_{[D]}(i), \tag{2}$$

where $\chi_{[D]}$ is the characteristics function of the set $[D]$. That means $\chi_{[D]}(i) = 1$, if $i \in [D]$ and $\chi_{[D]}(i) = o$ if $i \notin [D]$.

For the outputs $j \in [J]$ the contribution is given by

$$\mathrm{mEff}_j(n) = \frac{y_j(n)}{y_j(n) + \gamma(n)(\beta_j^*(n) - y_j(n))}. \tag{3}$$

One interesting feature about MEA is that the inefficiency of each input can be analyzed individually. In fact, using the ideas in [30], we calculate the following inefficiency index. For a given dataset $z = \{z(n)\}_{n\in\mathcal{N}}$ the inefficiency index for each given input index $i \in [I]$ and tuple $n \in N$ is given by

$$\mathrm{mIneff}_i(n) = \frac{\sum_{n=1}^{N} \gamma(n)(x_i(n) - \alpha_i^*(n))}{\sum_{n=1}^{N} x_i(n)}. \tag{4}$$

We refer to the inefficiency index to determine the number of times each input was used inefficiently, since our particular interest is to assess if global efficiency can be improved with less inputs.

3.2. Input and output variables

At this section, we present the inputs and outputs selected for the MEA model. As inputs, the following variables were considered: (i) *Wage (W)*, in euros; (ii) *Experience time (ET)*, in hours, as the quantity of training hours invested by the company and benefited by the worker; (iii) *Response time (RT)*, in seconds, as the time a worker spends to react or respond to the presence of parts in a queue; (iv) *Delay time index (DT)*, per unit, as a penalization value that measures the lapse between the planned shift start time and the time a worker effectively started working, by using a Gaussian function as weight.

The above mentioned inputs where primarily chosen based on an economic factor. The purpose is to understand which variables have greater influence on production outcomes, such as bottleneck occurrence and product reworks. By this, we mean that either W, ET, RT and DT are, per se, direct/ indirect costs for the company, and can either be interpreted as resources. That is, worker's resources (W and ET) and company resources (RT and DT). To the best of our knowledge, regarding workers' performance evaluation, no paper has been found in the literature addressing the above-mentioned variables used in the MEA model.

Information relative to the mean and standard deviation values of each input and dataset is present in Table 1. Every dataset is unique, and it detains information correspondent to a single working shift (WS). For the study, the same 14 workers are analysed, so the average and standard deviation values for wages and experience time remain equal to all datasets.

WS	*W (€)*	*ET (hours)*	*RT (secs)*	*DT (pu)*
d01			1.16 ± 0.99	0.26 ± 0.14
d02			0.59 ± 0.45	0.17 ± 0.13
d03	666.21 ± 46.32	1060.71 ± 586.66	0.59 ± 0.53	0.21 ± 0.14
d04			0.96 ± 0.88	0.19 ± 0.15
d05			1.06 ± 0.96	0.18 ± 0.13

Table 1: Mean and standard deviation of the information of the dataset from each working shift (WS) and each input variable.

For production performance analysis, the occurrence of bottlenecks and/or machine reworks are usually aspects that concern managers. According to [19], for bottleneck identification, the majority of methods found in the literature can be split into two main categories: the workstation state, evaluating the production time of a part in a workstation, or the queue state, evaluating the time of permanence of a part in a queue, (the *Queue Time* and the *Measured Time*, respectively, explained in section 2.2., Figure 1). However, when calculating these metrics for human-machine types of workstations, two influence factors can be considered: the machine's workload and worker's workload. To determine the impact of bottlenecks occurrence related to the worker's performance, then only the percentage of time of the worker's workload should be accounted for and applied to the bottleneck results. To this extent, and with previous knowledge about worker's workload percentages, such information was also added to the results, bringing more credibility into the analysis.

Considering that all machines perform with the same accuracy level, the number of workstation reworks can provide relevant information about workers quality standards. If a workstation has the highest number of reworks, the assigned worker should be "penalized" in his efficiency score, when compared to other workers in others workstations.

So, it is intended to determine which workers have the lowest associated values of W, ET, RT, DT, and, in combination, are able to "create" fewer bottlenecks and reworks in the production system, maximizing performance. Following the logic of benchmarking theory, it is intended to minimize inputs and maximize outputs. Acknowledging this, the outputs selected for the study are variables that managers want to minimize, so their complementary values are determined. Hence, let $i \in \{0 \ldots, j+1\}$ be the unique ID that characterizes a worker for a given working shift, where j is the total number of records analyzed (14 workers x 5 work shifts, so $j = 70$). Thus, taking into account the definition of AAPM (see [18]), they may be written as

$$cAMPM_i = max\{AMPM_k : k \in \{0, \ldots, j+1\}\} - AMPM_i, \quad (5)$$

$$cAQPM_i = max\{AQPM_k : k \in \{0, \ldots, j+1\}\} - AQPM_i, \quad (6)$$

$$cReworks_i = max\{Reworks_k : k \in \{0, \ldots, j+1\}\} - Reworks_i. \quad (7)$$

The wages were randomly allocated according to the information provided by [31], as well as the values for the experience time assigned to each worker. The response times were calculated according to the information present in sub-section 2.2. The delay time index DT(t) was calculated as:

$$\mathrm{DT}(t) = G_\sigma(t_b) - G_\sigma(t) \quad \text{where} \quad G_\sigma(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(t-t_b)^2}{\sigma^2}\right), \quad (8)$$

for a given standard deviation σ that accounts for the penalization curve spread.

4. Results discussion

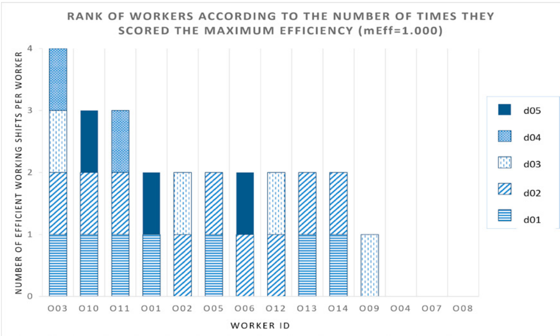
Table 2 presents a general characterization of the datasets used for the MEA algorithm. Four of these datasets were selected by a company manager and labelled as “Good” and “Bad”, according to the best/worst quality percentages and best/worst production numbers, as pictured in Table 2. Dataset d01 was randomly selected for control. The study of three types of datasets represents an advantage for the problem analysis, by comparing which parameters have greater impact in better or worst production quality percentages and/or production numbers.

WS	*Dataset Type*	*Date*	*Duration*	*Quality %*	*Reworks %*	*Production*	*Workers*
d01	Random	23/Nov/2020	08:08:21	94.07%	0.726%	112	
d02	Good	08/Feb/2021	08:08:33	97.98%	1.260%	132	
d03	Good	09/Feb/2021	08:05:13	97.78%	0.782%	132	14
d04	Bad	17/Feb/2021	07:53:47	72.59%	0.813%	98	
d05	Bad	18/Feb/2021	07:59:28	81.48%	0.416%	110	

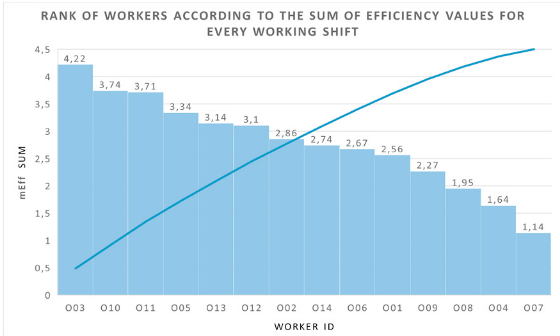
Table 2: Characterization of the datasets used in the benchmarking.

Figure 2 (a) and Figure 3 (a) show, respectively, the ranking results of workers and working shifts, according to the number of times maximum efficiency was determined by the algorithm. Looking to the example of worker ID=O03, the maximum efficiency $\mathrm{mEff}_i(\mathrm{O03}, s) = 1.000$, was scored in $s = \{1, 2, 3, 4\}$, datasets d01, d02, d03 and d04. On the other hand, Figure 2 (b) and Figure 3 (b) show, respectively, the ranking results of workers and working shifts, in association with the sum of the efficiency values predetermined by the MEA. So, particularly looking to Figure 3 (b) and comparing it to Figure 3 (a), d05 has a higher global sum of mEff values than d03. However, the number of workers in d05 with a maximum efficiency rate is smaller when compared to d03. With this information, it can be said that in d05, the group of 14 workers had worked more evenly balanced, once there were fewer workers working with maximum efficiency, and a higher sum of total efficiency was registered. Taking another look at the dataset characterization and building a connection between the results of Figure 3, it is interesting to check that the ”bad” datasets are also ranked in the bottom positions. So, this benchmarking analysis also validates, as well as it is equally validated, by the characterization type made by the line managers, for each dataset. In addition to the historical data information in which managers based their categorization, there is now also a relationship between the workers’ evaluation parameters, which had not yet been considered.

To extract further information from the benchmark study, a clustering approach was carried out for each pair worker/working shift. According to [32, 33], clustering allows the allocation of observations into groups such that similarity is maximized within-groups and minimized between-groups. In this work, clustering is used to group

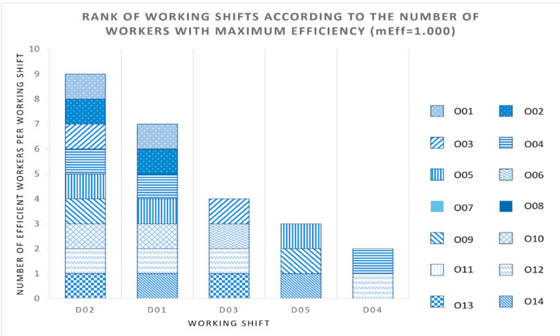


(a)

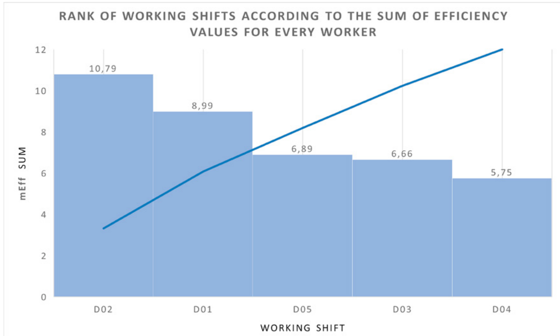


(b)

Fig. 2: Worker's rank according to MEA algorithm calculations.

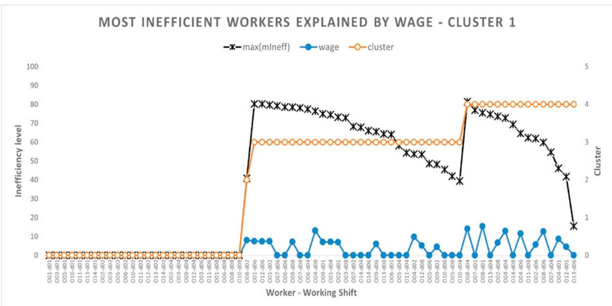


(a)

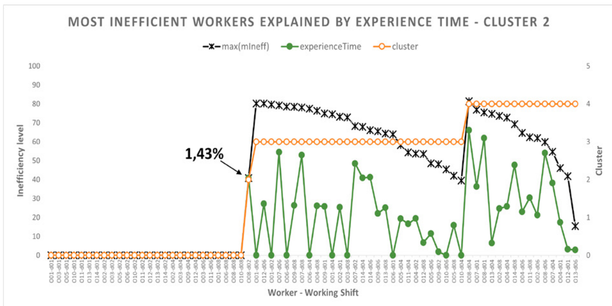


(b)

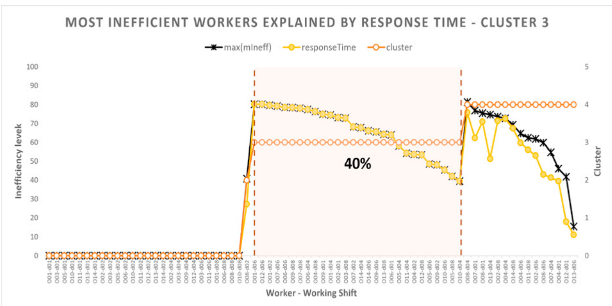
Fig. 3: Working Shift's rank according to MEA algorithm calculations.



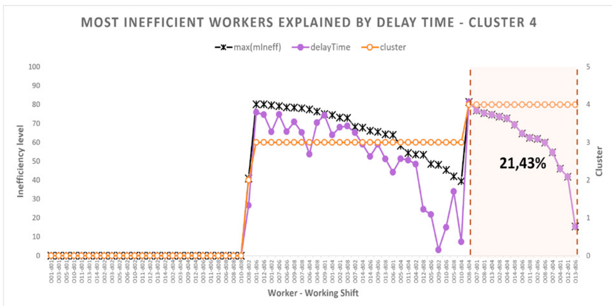
(a)



(b)



(c)



(d)

Fig. 4: Clustering by inefficiency.

tuples whose maximum individual inefficiencies, calculated by $\mathrm{mIneff}_i(n)$, are explained by the same input variables. Hence, Figure 4 shows four clusters, where the maximum inefficiency value is equal to: (a) $\mathrm{mIneff}_{wage}(n)$; (b) $\mathrm{mIneff}_{experiencetime}(n)$; (c) $\mathrm{mIneff}_{responsetime}(n)$; (d) $\mathrm{mIneff}_{delaytime}(n)$.

According to the results, both response time and delay time are responsible for more than 60% of worker/working shift maximum inefficiency levels. Solely for tuple (O08, d02), maximum inefficiency is explained by the experience time, and no case of inefficiency is explained by the worker's wage. We may conclude that the workers' response time to parts in queue, and the delay time to the start of the working shift are the two main reasons for throughput decrease.

5. Conclusions and future work

In the context of the aims of the industry 4.0, this work shows the importance of considering, not only IoT data, but also workers' behaviour (the so-called human factor) on real manufacturing production environments. Commonly, bottleneck analysis is used for tasks balancing between workstations, root cause analysis, maintenance reasons, study of machine limitations, were here, instead, it was used to assess workers' performance. To reduce overall costs and to identify potential workers qualification requirements, an approach was developed, capable of analysing data, extracted directly from a company MES, by using a multi-directional efficiency analysis methodology to quantify the impact of machine/worker operations on the production outcomes (by choosing specific inputs and outputs). The methodology also provides worker/shift clusters, obtained from the relative ranking and the measures of inefficiency, that help to better understand the great variability of real production lines.

As proven in this work, the response time (RT) is one of the primary causes for production line inefficiency, where delay time index (DT) is the second main cause. Having this kind of information, managers can make decisions with very few subjectivity involved about workers evaluation, which is one of the main contributions of this paper.

Since workers present distinct behaviors on different days, the proposed approach can be used to alert managers for a potential problem happening with a worker (e.g., motivational, skill related), when its ranking position is consistently decreasing. Additionally, to stimulate the rise of efficiency scores calculated with the MEA method, the approach suggested by [34] is recommended: the development of an appropriate incentive scheme. The idea of the author, adapted to the present problem, is to award workers who are willing to reduce their response time and delay time index, through different types of incentives. The goal: to increase efficiency, improve overall production but also to create a more competitive environment, where workers feel like their individual efforts may be compensated.

In a future work, it would be interesting to merge this quantitative approach with qualitative data, collected directly from workers. The aim would be to formulate a more broad, human-related, fair and democratic view of the problem. Furthermore, a system that is able to establish a correlation between the obtained results and absenteeism, or even predict workers health degradation would be a major advance into a better production system.

Acknowledgements. The present study was developed in the scope of the Project Augmented Humanity (PAH) [POCI-01-0247-FEDER-046103], financed by Portugal 2020, under the Competitiveness and Internationalization Operational Program, the Lisbon Regional Operational Program, and by the European Regional Development Fund. The first author was partially supported by the Center for Research and Development in Mathematics and Applications (CIDMA), through the Portuguese Foundation for Science and Technology, reference UIDB/04106/2020. The second author has a PhD grant in the scope of PAH.

Data availability statement. The bottleneck results used in the present study are confidential information of Bosch company manufacturing systems, so they are not publicly available.

References

[1] Leibenstein, H. (1966). Allocative Efficiency vs. "X-Efficiency." The American Economic Review, 56(3), 392–415. http://www.jstor.org/stable/1823775

[2] Farrell, M. J. (1957). The Measurement of Productive Efficiency. Journal of the Royal Statistical Society. Series A (General), 120(3), 253. https://doi.org/10.2307/2343100

[3] Klehe, U.-C., & Latham, G. (2006). "What Would You Do—Really or Ideally? Constructs Underlying the Behavior Description Interview and the Situational Interview in Predicting Typical Versus Maximum Performance". *Human Performance*, **19**, 357–382. https://doi.org/10.1207/s15327043hup1904_3

[4] Deadrick, D. L., Gardner, D. G. (2008). ”Maximal and typical measures of job performance: An analysis of performance variability over time.” *Human Resource Management Review* **18** (3), 133–145. https://doi.org/10.1016/j.hrmr.2008.07.008

[5] Klehe, U.-C., & Anderson, N. (2007). ”The role of typical and maximum performance in personnel selection”. *Tetrahedron Letters - TETRAHEDRON LETT*, 11–38.

[6] Katiraee, N., Calzavara, M., Finco, S., Battini, D., Battaïa, O. (2021). ”Consideration of workers' differences in production systems modelling and design: State of the art and directions for future research.” *International Journal of Production Research*, **59** (11), 3237–3268. https://doi.org/10.1080/00207543.2021.1884766

[7] Sawhney, R., Thakur, K. A., Maleki, S., Marella, L., Upreti, G. (2011). Improving Efficiency and Reliability of Manufacturing Systems by Variability Placement. February 2015. https://doi.org/10.13140/RG.2.1.4720.1364

[8] Liang, C., Li, Q. (2008). Enterprise information system project selection with regard to BOCR. International Journal of Project Management, 26(8), 810–820. https://doi.org/10.1016/j.ijproman.2007.11.001

[9] Brochado, A.F., Rocha, E.M., Almeida, D., de Sousa, A., Moura, A. (2021) “Data-Driven Bottleneck Identification with Minimal Information.” *Manuscript submitted for publication.*

[10] Camp, R. C. (1989). Benchmarking: the search for industry best practices that lead to superior performance. Asq Press.

[11] Altham, W. (2007). Benchmarking to trigger cleaner production in small businesses: drycleaning case study. Journal of Cleaner Production, 15(8–9), 798–813. https://doi.org/10.1016/j.jclepro.2006.07.005

[12] Boxwell Jr, R. J. (1994). Benchmarking for competitive advantage.

[13] Bolli, A., Emtairah, T. (2001, January). Environmental benchmarking for local authorities: From concept to practice - European Environement Agency. https://www.eea.europa.eu/publications/Environmental_issues_No_20

[14] Camp, R. C. (1995). Business Process Benchmarking: Finding and Implementing Best Practices. Milwaukee: ASQC Quality Press. https://www.amazon.com/BENCHMARKING-Implementing-organizations-Corporations-benchmarking/dp/B003UO56BG

[15] Meybodi, M. Z. (2009). Benchmarking performance measures in traditional and just-in-time companies. Benchmarking, 16(1), 88–102. https://doi.org/10.1108/14635770910936531

[16] Lai, M. C., Huang, H. C., Wang, W. K. (2011). Designing a knowledge-based system for benchmarking: A DEA approach. Knowledge-Based Systems, 24(5), 662–671. https://doi.org/10.1016/j.knosys.2011.02.006

[17] Roser, C., Nakano, M. (2015). A quantitative comparison of bottleneck detection methods in manufacturing systems with particular consideration for shifting bottlenecks. IFIP Advances in Information and Communication Technology, 460, 273–281. https://doi.org/10.1007/978-3-319-22759-7_32

[18] Roser, C., Nakano, M., Tanaka, M. (2002). Detecting Shifting Bottlenecks. International Symposium on Scheduling, 59–62.

[19] Thürer, M., Ma, L., Stevenson, M., Roser, C. (2021). Bottleneck detection in high-variety make-to-Order shops with complex routings: an assessment by simulation. Production Planning Control, 1–12. https://doi.org/10.1080/09537287.2021.1885795

[20] Bogetoft, P., Hougaard, J. L. (1999). Efficiency Evaluations Based on Potential (Non-Proportional) Improvements. Journal of Productivity Analysis, 12(3), 233–247. https://doi.org/10.1023/A:1007848222681

[21] Asmild, M., Hougaard, J. L., Kronborg, D., Kvist, H. K. (2003). Measuring inefficiency via potential improvements. Journal of Productivity Analysis, 19(1), 59. https://www.proquest.com/scholarly-journals/measuring-inefficiency-via-potential-improvements/docview/201701098/se-2?accountid=26357

[22] Bogetoft, P., Hougaard, J. L. (2004). Super efficiency evaluations based on potential slack. European Journal of Operational Research, 152(1), 14–21. https://doi.org/10.1016/S0377-2217(02)00642-2

[23] Asmild, M., Baležentis, T., Leth Hougaard, J. (2016). Multi-directional program efficiency: the case of Lithuanian family farms. Journal of Productivity Analysis. https://doi.org/10.1007/s11123-014-0419-6

[24] Manevska-Tasevska, G., Hansson, H., Asmild, M., Surry, Y. (2021). Exploring the regional efficiency of the Swedish agricultural sector during the CAP reforms – multi-directional efficiency analysis approach. Land Use Policy, 100, 104897. https://doi.org/10.1016/j.landusepol.2020.104897

[25] Charnes, A., Cooper, W. W., Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 2(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

[26] Banker, R. D., Charnes, A., Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Management Science, 30(9), 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

[27] Asmild, M., Matthews, K. (2012). Multi-directional efficiency analysis of efficiency patterns in Chinese banks 1997-2008. European Journal of Operational Research, 219(2), 434–441. https://doi.org/10.1016/j.ejor.2012.01.001

[28] Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., Shale, E. A. (2001). Pitfalls and protocols in DEA. European Journal of Operational Research, 132(2), 245–259. https://doi.org/10.1016/S0377-2217(00)00149-1

[29] Hahn, G. J., Brandenburg, M., Becker, J. (2021). Valuing Supply Chain Performance Within and Across Manufacturing Industries: A DEA-based Approach. International Journal of Production Economics, 108203. https://doi.org/10.1016/j.ijpe.2021.108203

[30] Bogetoft, P., Otto, L. (2010). Benchmarking with dea, sfa, and r (Vol. 157). Springer Science Business Media.

[31] Direção de Serviços de Apoio Técnico, do Ministério do Trabalho e Segurança Social. (2021). Boletim do Trabalho e Emprego - “Enquadramentos e tabela de remunerações mínimas mensais.” http://bte.gep.msess.gov.pt/completos/2021/bte1_2021.pdf

[32] Hair, Joseph F., et al. Multivariate Data Analysis: A Global Perspective. 7th ed. Upper Saddle River: Prentice Hall, 2009. Print.

[33] Agard, B., Penz, B. (2009). A simulated annealing method based on a clustering approach to determine bills of materials for a large product family. International Journal of Production Economics, 117(2), 389–401. https://doi.org/10.1016/j.ijpe.2008.12.004

[34] Bogetoft, P. (1994). Incentive efficient production frontiers: an agency perspective on DEA. Management Science, 40(8), 959–968. https://doi.org/10.1287/mnsc.40.8.959