Mining resource profiles from event logs

Pika, A.; Leyer, M.; Wynn, M.T.; Fidge, C.J.; ter Hofstede, A.H.M.; van der Aalst, W.M.P.

Published in:
ACM Transactions on Management Information Systems

DOI:
10.1145/3041218

Published: 01/03/2017

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Mining Resource Profiles from Event Logs

ANASTASIIA PIKA, Queensland University of Technology
MICHAEL LEYER, University of Rostock
MOE T. WYNN and COLIN J. FIDGE, Queensland University of Technology
ARTHUR H. M. TER HOFSTEDE, Queensland University of Technology and Eindhoven University of Technology
WIL M. P. VAN DER AALST, Eindhoven University of Technology and Queensland University of Technology

In most business processes, several activities need to be executed by human resources and cannot be fully automated. To evaluate resource performance and identify best practices as well as opportunities for improvement, managers need objective information about resource behaviors. Companies often use information systems to support their processes, and these systems record information about process execution in event logs. We present a framework for analyzing and evaluating resource behavior through mining such event logs. The framework provides (1) a method for extracting descriptive information about resource skills, utilization, preferences, productivity, and collaboration patterns; (2) a method for analyzing relationships between different resource behaviors and outcomes; and (3) a method for evaluating the overall resource productivity, tracking its changes over time, and comparing it to the productivity of other resources. To demonstrate the applicability of our framework, we apply it to analyze employee behavior in an Australian company and evaluate its usefulness by a survey among industry managers.

CCS Concepts: • **Information systems** → Information systems applications; • **Information systems applications** → Decision support systems;

Additional Key Words and Phrases: Resource profile, event log, mining resource behavior, evidence-based performance evaluation

ACM Reference Format:
Anastasiia Pika, Michael Leyer, Moe T. Wynn, Colin J. Fidge, Arthur H. M. ter Hofstede, and Wil M. P. van der Aalst. 2017. Mining resource profiles from event logs. ACM Trans. Manage. Inf. Syst. 8, 1, Article 1 (March 2017), 30 pages.
DOI: http://dx.doi.org/10.1145/3041218

1. INTRODUCTION

Employees are almost always involved in the execution of business processes, and they play a fundamental role in the creation of value in an organization. To better

This research was funded by the “Risk-Aware Business Process Management” project (ARC DP110100091), for which ethical clearance was obtained (approval no. 1100000288). Human participants were involved in the evaluation of the project results (by completing a survey), and ethical clearance for conducting the survey was obtained (approval no. 1400000420). We also used an anonymized dataset in our case study, so this component of the research was exempt from ethical review in conformity with the National Statement on Ethical Conduct in Human Research (2007; https://www.nhmrc.gov.au/guidelines-publications/e72).
Authors' addresses: A. Pika (corresponding author), M. T. Wynn, C. J. Fidge, and A. H. M. ter Hofstede, Queensland University of Technology, 2 George St, 4000, Brisbane, Australia; emails: {a.pika, m.wynn, c.fidge, a.terhofstede}@qut.edu.au; M. Leyer, University of Rostock, Ulmenstr. 69, 18057, Rostock, Germany; email: michael.leyer@uni-rostock.de; W. M. P. van der Aalst, Eindhoven University of Technology, PO Box 513, NL-5600 MB, Eindhoven, The Netherlands; email: w.m.p.v.d.aalst@tue.nl.

DOI: http://dx.doi.org/10.1145/3041218

understand areas of improvement of resource[1] and process performance, managers need objective information about the working behaviors of teams and individual employees. Often managers do not have such information, and they have to make decisions based on their subjective judgments [Manoharan et al. 2009]. This may result in inadequate performance appraisal systems, ineffective employee development strategies, and missed opportunities for performance improvement. For example, some teams may be overloaded, or important activities may be assigned to inefficient resources. Identification of such patterns of resource behavior can help improve individual employee and team performance and overall process performance.

Given this background, the research question that emerges is: How can resource behavior be analyzed in an objective way? Business processes are often supported by information systems that record information about process executions in event logs. Although resources play an important role in process executions, there are few methods currently available that allow organizations to extract knowledge about resource behavior from event logs, with some notable exceptions [Song and van der Aalst 2008; van der Aalst et al. 2005; Nakatumba and van der Aalst 2010; Ly et al. 2006], but these focus on organizational structures and do not consider changes in resource behavior over time (with the exception of the work of Nakatumba and van der Aalst [2010], in which changes in resource workload are considered). In a dynamic world, resources, as well as other process artifacts, change, and "it is important to be able to get [their] documented history" [Beheshti et al. 2016]. For example, the influence of contextual factors (e.g., holiday season, weather) on the efficiency of resources can be observed over time [Leyer 2011]. In addition, resource planning can be described as a dynamic decision system that is characterized by delays between actions and results [Sterman 1989]. If a new policy to guide resources in their work is set up, it is very likely that the effect in terms of efficiency can be observed with a delay, but the time of the delay is typically unknown. Thus, we present a framework for analyzing resource behavior through event log mining that allows tracking changes over time. Figure 1 depicts the main idea of our framework.

Analyzing resource behavior can be seen from different perspectives with measures being structured according to the most relevant aspects [de Leeuw and van den Berg 2011]. As we focus our analysis on the actors (resources) involved in undertaking activities within a process, three perspectives appear to be relevant [Sawhney and Chason 2005; White et al. 1999; de Leeuw and van den Berg 2011]:

—*The resource actions (i.e., what the resource has been doing).* This is a descriptive view of resource behavior. This leads to Research Question 1: How can the behaviors of resources be identified?
—*The effect of resource behaviors*. This covers the relationship between resource behaviors and outcomes and is thus a more analytic perspective. This leads to Research Question 2: How can the effects of resource behaviors be quantified?
—*Evaluation and comparison of resource productivity*. Here, a comparison is made with other resources over different time periods. This leads to Research Question 3: How can the overall productivity of resources be evaluated?

All three questions rely on the same input data but take different perspectives in analyzing resource behavior. As such, the analyses can be performed independently of each other. Although the results are related from a content point of view, the separation into the three perspectives allows a clear separation of the specific questions of interest and acknowledges that different methods are required to address the questions.

[1]In this article, we use the term *resource* as a generic term for *human resources*; it can refer to an individual employee or a group of employees (e.g., a role or a team).

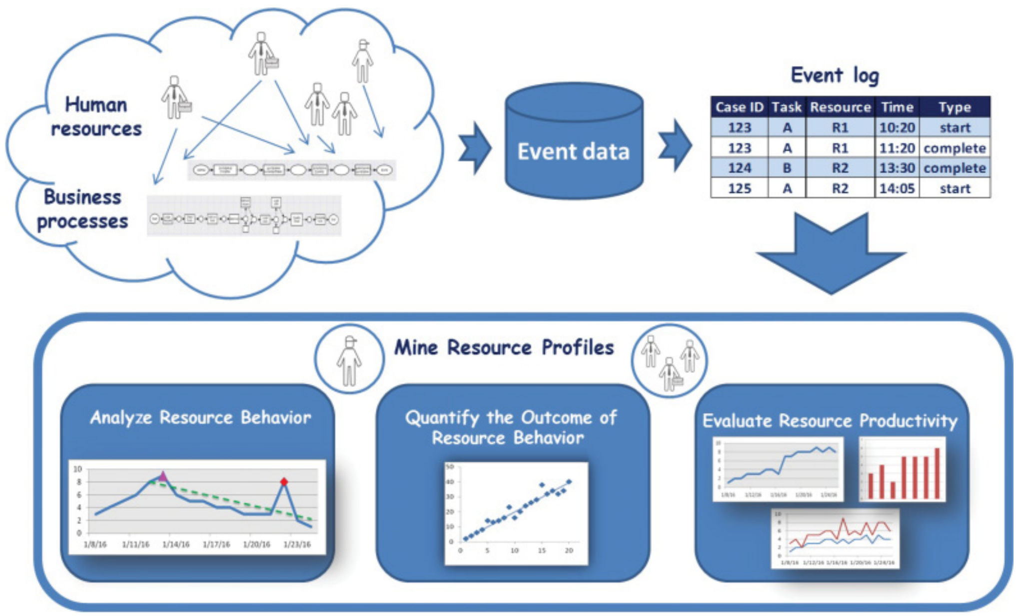

Fig. 1. Mining resource profiles (team or individual) from event logs.

In our earlier work [Pika et al. 2014], we presented a framework that allowed us to answer the first question by analyzing resource behavior indicators (RBIs) in the following categories of resource behavior: skills, utilization, preferences, productivity, and collaboration. However, it did not allow us to investigate the effects of this behavior, and it did not provide a means for evaluating overall resource productivity. In this article, we present an extended framework that includes a new method for investigating whether or not any relationships exist between given resource behaviors and outcomes by analyzing an event log (e.g., whether or not employees complete more work when they multitask) (Section 3.3). The proposed method is based on regression analysis and allows us to investigate relationships from the following three process perspectives: case, task, and time. Another technical contribution of this article is a new method for evaluating resource productivity by analyzing event logs (Section 3.4). The method uses the data envelopment analysis (DEA) technique [Bogetoft and Otto 2011] for this purpose. It allows us to compare the productivity of different resources (teams or individual employees) and to track productivity evolution over time. The framework is based on a set of predefined measures and also allows users to define their own indicators. We present the results of a new case study in which our framework was applied to analyze resource behaviors in an Australian company (Section 4.1) and the results of a new survey that evaluates the usefulness of the framework among managers (Section 4.2).

2. RELATED AND PREVIOUS WORK

In this section, we first introduce business processes and human resource behavior, then discuss process mining focusing on techniques for extracting human resource behavior. We conclude the section with a discussion of resource benchmarking.

Business processes. A business process is a coordinated sequence of activities that jointly realize a business goal and are performed "in an organizational and technical environment" [Weske 2007]. People are considered a core element of business process management (BPM) [Rosemann and vom Brocke 2015]. Business processes whose executions depend on knowledge workers performing knowledge-intensive tasks are referred to as knowledge-intensive processes [Di Ciccio et al. 2015]. Knowledge-intensive business services "have grown exponentially over the last three decades" [Freel 2016] and "have been the most dynamic component" in developed economies [Wirtz and Lovelock 2016]. Despite knowledge workers playing a prominent role in modern organizations, the BPM community traditionally provided "minor emphasis of collaboration aspects" [Di Ciccio et al. 2015].

Human resource behavior. Process outcomes depend on human resource behavior, which is considered to be the "most important element that can affect project success" [Thevendran and Mawdesley 2004]. Human behavior can be complex, as people "seek to promote their utility, through advancing their interests, preferences and ideas," and their decisions are influenced by their social contexts [Leftwich 2015]. Resource behavior can also evolve over time [Beheshti et al. 2016]. Hence, human performance planning and measurement can become a challenging task, which is typically handled by human resource management departments [Espinilla et al. 2013; Peretz and Fried 2012]. Traditional performance evaluation approaches often suffer from "lack of objectivity, prejudice or halo errors" [Espinilla et al. 2013], and they often have a "narrow, or uni-dimensional, focus" [Neely et al. 2000]. One way to overcome the lack of objectivity is to base performance evaluation on "the opinion of different groups of reviewers who socialize with evaluated employees" [Espinilla et al. 2013]. Various authors have proposed performance evaluation frameworks that are based on "a balanced set of measures"; they "suggest some areas in which measures of performance might be useful, but provide little guidance on how the appropriate measures can be identified" [Neely et al. 2000]. Such performance measures can be defined on an aggregate level, namely for teams, departments, or a company [Nudurupati et al. 2011], or for individual employees [Neely et al. 2005; Thompson and Goodale 2006; Dulebohn and Johnson 2013].

Process mining. Business processes are not always executed as expected "due to the high variability that may affect operational processes in real world scenarios" [Ceravolo et al. 2016]. Processes are often supported by information systems that record information about their executions in event logs [van der Aalst 2016; Ceravolo et al. 2016]. Process mining aims to extract knowledge about business processes from such event logs [van der Aalst 2016]. Early process mining techniques were aimed at discovering process models from event logs; later on, process mining algorithms were developed that also support the discovery of other process perspectives (e.g., organizational, case, or time perspectives) [van der Aalst 2016]. Some aspects of our framework were inspired by existing process mining approaches. In our own earlier work, we presented process risk indicators for the identification of case delays [Pika et al. 2013a, 2013b]. We showed that a high resource workload or the involvement of certain resources in a case can contribute to case delays. De Leoni et al. [2014] proposed a general framework for correlating process characteristics linked to events by applying decision tree analysis. Van der Aalst [2013] proposed the notion of process cubes, whereby events can be "organized using different dimensions" such as time windows or event or case types. Bose et al. [2013] proposed a framework for detecting changes in processes. Pika et al. [2016] presented a method for evaluating and predicting overall process risk that considers process evolution. However, the preceding approaches focus on business processes rather than resources.

Extracting human resource behavior. Several methods for extracting knowledge about some aspects of resource behavior from event logs have been proposed in the process mining field: Song and van der Aalst [2008] proposed techniques for extracting organizational models; Song and van der Aalst [2008] and van der Aalst et al. [2005] presented a method for extracting social networks from event logs; Huang et al. [2012] proposed some measures for resource preference, availability, competence, and cooperation; and Nakatumba and van der Aalst [2010] presented a method that investigates the effect of resource workload on service times. Several approaches have been devised that use event logs to derive resource allocation mechanisms: Liu et al. [2012] proposed an approach that mines resource allocation rules from event logs by applying a data mining algorithm, Cabanillas et al. [2013] proposed an approach for prioritizing potential task performers based on their preferences, Kumar et al. [2013] devised a model that considers compatibility between resources when assigning tasks and a technique for learning resource compatibility from logs, and Ly et al. [2006] presented a method for mining task assignment rules based on decision tree learning. These approaches focus on organizational structures, few specific measures, or resource allocation mechanisms (rather than different aspects of employee behavior), and they do not consider changes of resource behavior over time.

Resource benchmarking. We adopted DEA [Bogetoft and Otto 2011] to evaluate and compare resource productivity. DEA is a popular efficiency evaluation and benchmarking technique based on linear programming that is often used in operations research to compare the efficiency of companies [Bogetoft and Otto 2011], departments or company functions [Thanassoulis 1995; Tavakoli and Shirouyehzad 2013], and business processes [Dohmen and Leyer 2010]. A few works describe case studies in which DEA was applied to measure the efficiency of employees [Manoharan et al. 2009; Wagner et al. 2003; Koch-Rogge et al. 2014]. However, these case studies did not use event log data; they used data that was collected from different sources, manually preprocessed, and transformed into a form suitable for DEA analysis.

3. FRAMEWORK

Our framework consists of three modules, as depicted in Figure 1, with each module tackling the three research questions introduced in Section 1. The goal of the first module, called *Analyzing Resource Behavior*, is to discover what resources have been doing—that is, to gain objective information about their skills, utilization, working preferences, productivity, and collaboration patterns (Research Question 1 in Section 1). The goal of the second module, called *Quantifying the Outcome of Resource Behavior*, is to help managers to better understand the effects of resource behaviors on different process outcomes—for example, to check if the resource workload is affecting the duration of tasks executed by the resource or the quality of the work (Research Question 2 in Section 1). Our aim is not to learn whether or not certain resource behaviors are associated with particular outcomes. Instead, we aim to provide a method and a supporting tool that allows managers to observe what relationships exist between given resource behaviors and outcomes. Finally, the goal of the *Evaluating Resource Productivity* module is to provide a method for evaluating the overall productivity of a resource by comparing it to the productivity of other resources and tracking its evolution over time (Research Question 3 in Section 1). Resources are often involved in multiple activities with different levels of complexity. Looking separately at each task would be prohibitively time consuming. Managers would benefit from a method that can automatically evaluate the productivity of a resource considering different resource inputs and outputs. They can further explore some resources of interest (e.g., overperforming or underperforming) and look at individual indicators to investigate resource behaviors and their effects in detail (using other modules of the framework).

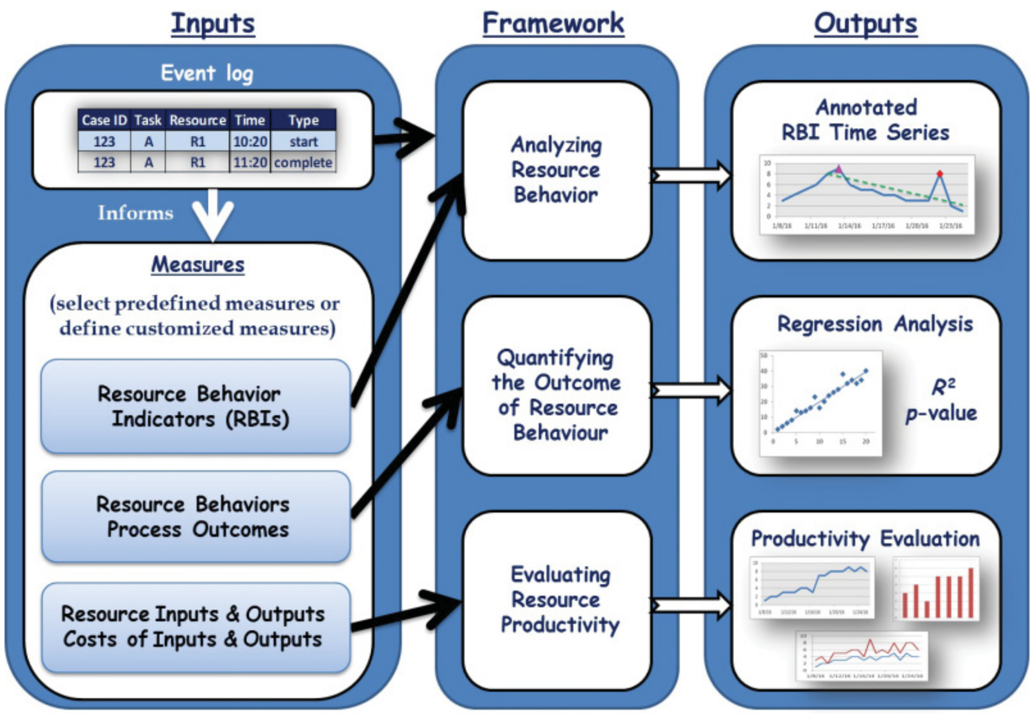

Fig. 2. Inputs and outputs of the framework for mining resource profiles from event logs.

Figure 2 depicts inputs and outputs of our framework. An input to the three modules is an event log that we discuss in detail in Section 3.1. Another input to the framework are resource behavior measures. These are different for the three modules, and we discuss them in Sections 3.2 through 3.4. We propose several measures of resource behavior and also provide an interface that allows users to define their own measures. Outputs of the framework are three different types of analysis visualized by corresponding charts. These will be discussed in Sections 3.2 through 3.4. In the following sections, we first describe definitions of the basic concepts that are used in the framework specifications. We then describe in detail the three modules of our framework.

3.1. Definitions

A key input to our framework is an event log that captures information about process executions.[2] Let $\mathcal{E}$ be the set of all events; an *event log* EL is a set of events (i.e., $EL \subseteq \mathcal{E}$). Events can have different attributes. We assume that each event has at least the following attributes (referred to here as *basic* attributes): $caseid$, $task$, $type$, $time$, and $resource$. We assume that *start* and *complete* event types are always recorded, and other event types (e.g., *create*) may also be recorded. We can treat an event log as a relation whose relation scheme is specified by the set of event attributes.

Let $\mathcal{C}$ be the set of all cases; a *case log* CL is a set of cases (i.e., $CL \subseteq \mathcal{C}$). Similarly, cases are characterized by different attributes. A case is uniquely identifiable by the case attribute $caseid$. A case log can also be treated as a relation whose relation scheme is specified by the set of case attributes.

[2]The framework implementation takes as input an XES event log, a standard format for event log data (http://www.xes-standard.org/).

caseid	task	time	type	resource	case_type
111	Investigate Incident	10/02/2006 10:00	start	Anne	urgent
111	Request Information	10/02/2006 10:15	start	Anne	urgent
112	Send Bill	11/02/2006 15:00	complete	Mike	normal
111	Investigate Incident	13/02/2006 13:45	complete	Anne	urgent
112	Update Record	13/02/2006 11:37	start	Mike	normal

Fig. 3. Example of an event log.

The value of attribute a of case c is denoted as c_a, whereas e_a denotes the value of attribute a of event e. For example, e_{caseid} is the case identifier of event $e \in \mathcal{E}$.

Events and cases can have other attributes (e.g., *cost* or *outcome*). As examples of such nonbasic attributes, we use the event attributes *sum* and *creator* and the case attributes *case_type* and *feedback*.

Figure 3 depicts an example of an event log that contains five events with five basic attributes and attribute *case_type*. We can see, for example, that in an urgent case with identifier "111," Anne started task "Request Information" at 10:15 on 10/02/2006. Note that in this example attribute *resource* refers to individual employees, but it can also refer to a role or a team.

We also derive from the basic event log the case attributes *case_duration* (the time difference between the timestamps of the last and first event of a case) and *case_resources* (the number of resources that were involved in a case). The following event attributes are derived from the basic event log: *task_duration* (the time difference between the corresponding activity's[3] *complete* and *start* events), *workload* (the number of task instances that were started but not completed by a resource involved in an event before the moment of event execution), and *workload_duration* (the time period during which the resource's workload did not change). For example, if a resource starts executing task a at time t_1 and then starts another task at time t_2 and task a is not completed yet at t_2, the resource's workload at time t_2 is 1 and the workload duration is $t_2 - t_1$. For each event, we also create the attribute *eventid*, which is a unique identifier of the event. If any of these derivable event or case attributes are recorded in a log, users may choose to use the log's values instead of deriving them from the basic attributes.

Let R be a set of resources, A a set of activities, t_1 and t_2 the beginning and the end of a given time slot, and r a given resource. Next, we define the functions that are later used in definitions of our methods:

—*Events completed during a given time slot [t1,t2)*:
$E_{CT}(t_1, t_2) \triangleq \{e \in EL \mid e_{time} \geq t_1 \wedge e_{time} < t_2 \wedge e_{type} = \text{`complete'}\}$
—*Events in which a given resource was involved during a given time slot*:
$E_{TR}(t_1, t_2, r) \triangleq \{e \in EL \mid e_{time} \geq t_1 \wedge e_{time} < t_2 \wedge e_{resource} = r\}$
—*Events completed by a given resource during a given time slot*:
$E_{CTR}(t_1, t_2, r) \triangleq E_{CT}(t_1, t_2) \cap E_{TR}(t_1, t_2, r)$
—*Cases completed during a given time slot*:
$$C_{CT}(t_1, t_2) \triangleq \{c \in \mathcal{C} \mid \exists e \in E_{CT}(t_1, t_2)[e_{caseid} = c_{caseid}] \wedge \nexists e' \in EL[e'_{caseid} = c_{caseid} \wedge e'_{time} > t_2]\}$$
—*Cases in which a given resource was involved*:
$C_R(r) \triangleq \{c \in \mathcal{C} \mid \exists e \in EL[e_{caseid} = c_{caseid} \wedge e_{resource} = r]\}$

[3]In this article, we use the terms *task* and *activity* interchangeably.

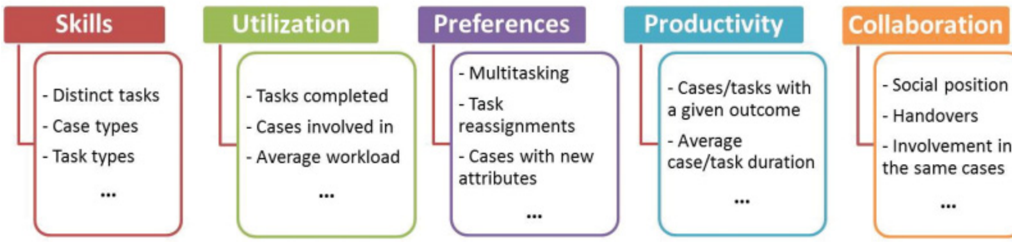

Fig. 4. Categories of resource behavior and examples of RBIs.

—*Cases in which a given resource was involved during a given time slot*:
$C_{TR}(t_1, t_2, r) \triangleq \{c \in \mathcal{C} \mid \exists e \in E_{TR}(t_1, t_2, r)[e_{caseid} = c_{caseid}]\}$.

3.2. Analyzing Resource Behavior

Our first goal is to be able to answer Research Question 1: What have resources been doing? To extract descriptive information about different aspects of resource behavior, we follow three main steps: (1) defining the RBIs, (2) extracting the RBI time series from the event logs, and (3) analyzing the RBI time series.

3.2.1. Defining RBIs. The kinds of resource behavior that managers may wish to analyze depend on the reasons for the analysis and the particular context. There may be many indicators of interest in a specific situation [Neely et al. 2005]. We refer to a collection of RBIs that are relevant in a particular context as a *resource profile*. Based on the literature, we propose the following categories of resource behavior [Pika et al. 2014]:

(1) Skills [Thompson and Goodale 2006] (What can a resource do?)
(2) Utilization [Neely et al. 2005] (What is a resource actually doing?)
(3) Preferences [Huang et al. 2012; van der Aalst 2010, 2015] (What working behavior does a resource often demonstrate?)
(4) Productivity [Murphy 1999] (How good is a resource at what it[4] does?)
(5) Collaboration [van der Aalst et al. 2005; Huang et al. 2012] (How does a resource work with other resources?)

We define a set of RBIs in each category (Figure 4). Many of our predefined RBIs can be extracted from basic event logs, whereas others require additional information (e.g., cost or feedback) to be recorded. Furthermore, some RBIs are generic, and others are only relevant in specific contexts; RBI values may be absolute or relative. We also provide an interface that allows users to define their own RBIs. For example, a manager may be interested in conformance and may wish to check whether a resource performs a task that should not be performed. In the remainder of this section, we discuss RBIs in each of the categories of resource behavior and provide formal definitions for a small selection of the RBIs. Let $RBI_n(t_1, t_2, r, [p_1 \dots p_n])$ denote the value of an RBI n during a given time slot, t_1 to t_2, for resource r and for an optional set of other parameters $p_1 \dots p_n$. For example, an optional parameter can refer to cost, customer feedback, or product category.

1. Skills: What can a resource do? Resources have different capabilities, and they acquire new skills at a different pace. Knowledge about resource capabilities can help in resource scheduling [Huang et al. 2012; Thompson and Goodale 2006] and resource development planning. We assume that a resource is capable of performing those types of activities that it has performed in the past; hence, RBIs in this category reflect

[4]In this article, we use the pronoun *it* when referring to resources, as the term *resource* can refer to an employee or a group of employees (e.g., a role or a team).

only "demonstrated" skills. Some example RBIs in this category include the number of distinct tasks completed by a resource, the fraction of distinct tasks completed by a resource with respect to the total number of distinct tasks completed, the number of completions of an activity with a given property by a resource, the fraction of completions of an activity with a given property by a resource with respect to the total number of completions of the activity with the given property, the number of case completions with a given property in which a resource was involved, the fraction of case completions with a given property in which a resource was involved with respect to the total number of case completions with the given property, and the fraction of completions of a given activity by a resource with respect to the total number of activity completions by the resource.

RBI 1.1. Distinct activities: The number of distinct activities completed by a given resource, r, during a given time slot, t_1 to t_2. This RBI is relevant in those working environments where resources learn new skills and get involved in more tasks over time.

$Distinct\_Activities(t_1, t_2, r) \triangleq |\{task \in A \mid \exists e \in E_{CTR}(t_1, t_2, r)[e_{task} = task]\}|$

RBI 1.2. Case types: The fraction of cases with a given value, p, of case attribute *case_type* completed during a given time slot, $[t_1, t_2)$, in which a given resource, r, was involved with respect to the number of cases completed during time slot $[t_1, t_2)$ in which the resource was involved (requires attribute *case_type* to be recorded in the log). Looking at the types of cases processed by a resource, managers can discover, for example, that the resource is only processing cases related to a specific product or that it is getting involved in more complex cases over time.

$Case\_Types(t_1, t_2, r, p) \triangleq$
$|\{c \in C_{CT}(t_1, t_2) \cap C_R(r) \mid c_{case\_type} = p\}|/|C_{CT}(t_1, t_2) \cap C_R(r)|$

RBI 1.3. Activity frequency: The fraction of completions of a given activity, a, by a given resource, r, during a given time slot, $[t_1, t_2)$, with respect to the total number of activity completions by resource r during time slot $[t_1, t_2)$. Managers can discover, for example, those activities that a resource frequently executes (hence, the activities in which the resource is more experienced) and those that it executes only occasionally.

$Activity\_Frequency(t_1, t_2, r, a) \triangleq$
$|\{e \in E_{CTR}(t_1, t_2, r) \mid e_{task} = a\}| \;/\; |E_{CTR}(t_1, t_2, r)|$

2. Utilization: What is a resource actually doing? Utilization RBIs measure how active a resource is without looking into the quality of its outputs. These indicators are inspired by measures in manufacturing (e.g., the number of units produced) [Neely et al. 2005]. Users can look at the number of activity instances completed by a resource, the fraction of activity instances completed by a resource with respect to the total number of activity instances completed, the number of completions of a given activity by a resource, the number of completed cases in which a resource was involved, the fraction of completed cases in which a resource was involved with respect to the total number of cases completed, or the average resource workload.

RBI 2.1. Activity completions: The number of activity instances completed by a given resource during a given time slot.

$Activity\_Completions(t_1, t_2, r) \triangleq |E_{CTR}(t_1, t_2, r)|$

RBI 2.2. Number of case completions: The number of cases completed during a given time slot in which a given resource was involved.

$Case\_Completions\_Number(t_1, t_2, r) \triangleq |C_{CT}(t_1, t_2) \cap C_R(r)|$

RBI 2.3. Fraction of case completions: The fraction of cases completed during a given time slot in which a given resource was involved with respect to the total number of cases completed during the time slot.

$$Case\_Completions\_Fraction(t_1, t_2, r) \triangleq |C_{CT}(t_1, t_2) \cap C_R(r)|/|C_{CT}(t_1, t_2)|$$

RBI 2.4. Average workload: The average number of activities started by a given resource but not completed at a moment in time. We do not consider the resource's working hours, and hence this RBI only yields a rough estimation of the resource's average workload.

$$Average\_Workload(t_1, t_2, r) \triangleq \sum_{e \in E_{TR}(t_1,t_2,r)} e_{workload} * e_{workload\_duration} / \sum_{e \in E_{TR}(t_1,t_2,r)} e_{workload\_duration}$$

3. Preferences: What working behavior does a resource often demonstrate? Resources have different working styles that may affect their performance [van der Aalst 2010, 2015; Huang et al. 2012]. We may learn about resources' preferences by checking if they often multitask, execute only similar tasks or take risks by performing a multitude of tasks, perform more work during certain weekdays, or reassign tasks to others.

RBI 3.1. Multitasking: The fraction of active time during which a given resource is involved in more than one activity with respect to the resource's active time. We do not consider the resource's working hours, and we assume that the resource is working on all tasks that they started. Hence, this RBI only yields a rough estimate of the resource's multitasking preference.

$$Multitasking(t_1, t_2, r) \triangleq \sum_{e \in E_{TR}(t_1,t_2,r), e_{workload}>1} e_{workload\_duration} / \sum_{e \in E_{TR}(t_1,t_2,r), e_{workload}>0} e_{workload\_duration}$$

RBI 3.2. New attribute values: The number of times a resource completed a task during a given time slot with an attribute value never seen before for any event (e.g., for the attribute *sum*). This RBI reflects the resource's propensity to execute new and hence risky tasks. For the attribute *sum*, the RBI can be formalized as follows:

$$New\_Attribute\_Values(t_1, t_2, r, sum) \triangleq |\{e \in E_{CTR}(t_1, t_2, r) \mid \nexists e' \in EL \; [e'_{sum} = e_{sum} \wedge e'_{time} < e_{time}]\}|.$$

RBI 3.3. Activity reassignments: The number of times when an activity started by a given resource during a given time slot was later completed by a different resource.[5]

$$Activity\_Reassignments(t_1, t_2, r) \triangleq |\{e \in E_{TR}(t_1, t_2, r) \mid e_{type} = start \wedge \exists e' \in EL[e'_{type} = complete \wedge e'_{resource} \neq r \wedge e'_{task} = e_{task} \wedge e'_{caseid} = e_{caseid} \wedge e'_{time} > e_{time}]\}|$$

4. Productivity: How good is a resource at what it does? Productivity RBIs aim to measure a resource's results—for example, its results in terms of the timeliness, cost, or quality of outputs (assuming that this information is recorded in the event log). Here we define RBIs for the number of activities/cases completed with a given outcome in which a resource was involved, the fraction of activities/cases completed with a given outcome in which a resource was involved with respect to the total number of activities/cases completed with a given outcome, the average value of a given outcome

[5]The definition is provided for cases without activity repetitions. If activities are repeated, one would need to use activity instance identifiers.

for cases/activities completed in which a given resource was involved, and the number of times when a given activity was repeated when completed by a resource.

RBI 4.1. Activity outcomes: The fraction of activities completed with a given outcome (e.g., *duration*) during a given time slot by a given resource with respect to the total number of activities completed by the resource during the time slot. For activity duration, the RBI can be formalized as follows:

$$In\_Time\_Activities(t_1, t_2, r, dur) \triangleq |\{e \in E_{CTR}(t_1, t_2, r) \mid e_{task\_duration} < dur\}| \;/\; |\{E_{CTR}(t_1, t_2, r)\}|.$$

RBI 4.2. Case outcomes: The fraction of cases completed during a given time slot with a given outcome (e.g., customer feedback *cf*) in which a given resource was involved with respect to the total number of cases completed during the time slot in which the resource was involved. This RBI requires a case outcome attribute to be recorded, such as *feedback* (a numeric value, e.g., a customer rating). For the attribute *feedback*, the RBI can be formalized as follows:

$$Satisfactory\_Cases(t_1, t_2, r, cf) \triangleq |\{c \in C_{CT}(t_1, t_2) \cap C_R(r) \mid c_{feedback} \geq cf\}| \;/\; |C_{CT}(t_1, t_2) \cap C_R(r)|.$$

RBI 4.3. Average duration of a given activity: The average duration of instances of a given activity completed during a given time slot by a given resource.

$$Average\_Activity\_Duration(t_1, t_2, r, a) \triangleq \sum_{e \in E_{CTR}(t_1,t_2,r), e_{task}=a} e_{task\_duration} / |\{e \in E_{CTR}(t_1, t_2, r) \mid e_{task} = a\}|$$

RBI 4.4. Average case duration: The average duration of cases completed during a given time slot in which a given resource was involved.

$$Average\_Case\_Duration(t_1, t_2, r) \triangleq \sum_{c \in C_{CT}(t_1,t_2) \cap C_R(r)} c_{case\_duration} / |C_{CT}(t_1, t_2) \cap C_R(r)|$$

RBI 4.5. Average customer feedback: The average customer feedback for cases completed during a given time slot in which a given resource was involved.

$$Average\_Customer\_Feedback(t_1, t_2, r) \triangleq \sum_{c \in C_{CT}(t_1,t_2) \cap C_R(r)} c_{feedback} / |C_{CT}(t_1, t_2) \cap C_R(r)|$$

5. Collaboration: How well does a resource work with other resources? It is important to measure the collaborative aspects of resource behavior. RBIs in this category can help us learn about a resource's collaboration patterns with some other resources (e.g., the number of handovers from/to a given resource or the number of times when two given resources were involved in the same cases) or obtain insight into a resource's overall social position within an organization (e.g., the number of other resources that executed a given activity, the average number of resources involved in the same cases with a given resource, or the fraction of resources involved in the same cases with a given resource during a given time slot with respect to the total number of resources active during the time slot).

RBI 5.1. Interactions between two given resources: The number of cases completed during a given time slot in which two given resources were involved.

$$Interactions\_Between\_Resources(t_1, t_2, r, r_2) \triangleq |C_{CT}(t_1, t_2) \cap C_R(r) \cap C_R(r_2)|$$

RBI 5.2. Social position: The fraction of resources involved in the same cases with a given resource during a given time slot with respect to the total number of resources active during the time slot.

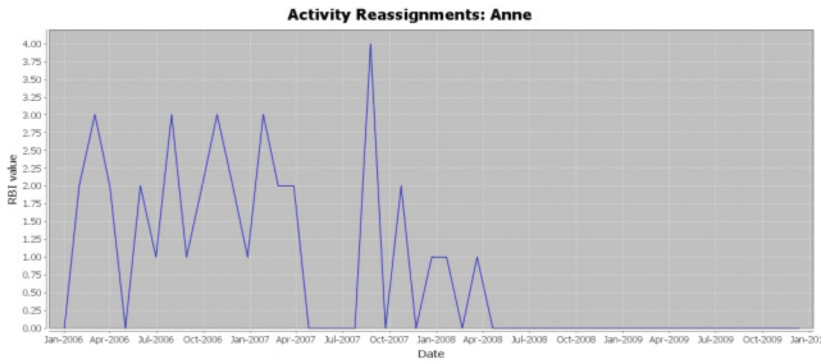

Fig. 5. Example of time series for RBI 3.3 "Activity reassignments".

$$Social\_Position(t_1, t_2, r) \triangleq |\{r_1 \in R \mid \exists c \in C_{TR}(t_1, t_2, r_1) \cap C_{TR}(t_1, t_2, r)\}|/|\{r_1 \in R \mid \exists c \in C_{TR}(t_1, t_2, r_1)\}|$$

RBI 5.3. Delegations: The number of times a resource, r, assigns a task to another resource, r_2. For this RBI, a corresponding attribute (e.g., *creator*) must be recorded in the log. For the attribute *creator*, the RBI can be formalized as follows:

$$Delegations(t_1, t_2, r, r_2) \triangleq |\{e \in E_{CTR}(t_1, t_2, r_2) \mid e_{creator} = r \wedge r \neq r_2\}|.$$

3.2.2. Extracting RBI Time Series from an Event Log. In the second step, we extract the RBI time series to track the evolution of a particular RBI over time. The RBI time series consists of RBI values extracted for a given period of time (e.g., per day or week) for a given resource. Let $RBI_n(t_1, t_2, r, [p_1 \ldots p_n])$ denote the value of an RBI n during a given time slot, t_1 to t_2, for a given resource, r, and optional parameters $p_1 \ldots p_n$; TS_{start} be the starting time point; $TS_{slotsize}$ be the sampling rate; TS_{size} be the number of time slots; and $Start(t)$ and $End(t)$ be functions that return the beginning and the end of a time slot for a given time t, respectively. An RBI time series is defined as

$$TS_{RBI_n} \triangleq \{(RBI_n(Start(t), End(t), r, [p_1 \ldots p_n]), t) \mid t \in \{TS_{start} + i * TS_{slotsize} \mid i \in \{0, 1, \ldots, TS_{size} - 1\}\}\}.$$

Here we use the following functions for the beginning and the end of a time slot: $Start(t) = t$ and $End(t) = t + TS_{slotsize}$; other functions can be defined.[6] (This gives flexibility to use different time series sampling methods. For example, one may define overlapping time slots.) The starting time point for the analysis, the time series sampling rate, and the number of time slots are the input parameters provided by a user. The time series sampling rate is an important parameter that can affect the analysis results. Selection of the time series sampling rate is a known problem often discussed in the literature [Lijffijt et al. 2012]. The time series sampling rate may depend on process characteristics (e.g., process granularity) and the type of analysis in which the user is interested. For example, if managers are interested in checking whether or not a team is less productive on Mondays, they will look at the daily RBI values. Figure 5 depicts an example of a time series (monthly values) for RBI 3.3, "Activity reassignments." We can observe that some tasks were reassigned from resource *Anne* until April 2008, and that there were no subsequent reassignments.

3.2.3. Analyzing the RBI Time Series. In the third step, we analyze the extracted RBI time series and visualize the analysis results. Time series charts accompanied by

[6] $Start(t) = t + a$ and $End(t) = t + TS_{slotsize} + b$, where a and b are input parameters.

trend lines can give many insights into the dynamics of resource behavior; however, they are not very convenient when the amount of available data is large. To deal with this issue, we use automatic techniques for time series analysis, such as the detection of outliers [van der Loo 2010] and change points [Ross and Adams 2012], and the time series comparison [Mann and Whitney 1947].

Many methods for change point detection have been developed; here we use nonparametric methods [Hawkins and Deng 2010; Ross and Adams 2012] that do not make assumptions about the distribution of the data. Users can choose nonparametric tests for detecting changes in location (Mann-Whitney), changes in scale (Mood), or arbitrary distributional changes (Kolmogorov-Smirnov, Cramer–von Mises) [Ross and Adams 2012].

Identifying outliers in the RBI time series (i.e., the points in time when RBI values are significantly different from typical values) can be helpful in the investigation of problems. We use an outlier detection method that first fits the distribution of the observations and then detects the observations that are unlikely to be generated by this distribution [van der Loo 2010].

Evaluation of the performance of a resource using RBIs can also be done by comparing the resource's behavior to the behavior of other resources. This allows users to quickly identify those resources whose behavior is different from others (e.g., underperforming or overperforming resources). To compare RBI time series, we use the nonparametric Man-Whitney U-test [Mann and Whitney 1947].

3.3. Quantifying the Outcome of Resource Behavior

Whereas the *Analyzing Resource Behavior* module described earlier allows us to extract descriptive information about different aspects of resource behavior, the goal of the *Quantifying the Outcome of Resource Behavior* module is to provide a method that allows managers to check whether or not a given resource's behaviors affect given outcomes (Research Question 2 in Section 1). For example, managers may be interested to learn how a resource's workload affects the quality of their work or if the case duration is influenced by the number of resources processing a case.

We use regression analysis to quantify the relationships between resource behaviors and outcomes. This is a statistical technique used for modeling the relationship between variables [Montgomery et al. 2012]. We use regression analysis because it is a popular technique often used for the investigation of relationships between different social and economic phenomena, and its results are easy to interpret for business users. The framework supports linear regression analysis [Montgomery et al. 2012], as well as nonparametric kernel-based regression, which does not make assumptions about data distribution [Racine and Li 2004].

To perform regression analysis, it is necessary to define the dependent variable and one or more independent variables. Our framework provides an interface for users to define the dependent and independent variables that they would like to analyze. It then extracts the values of the defined variables from an event log, fits a regression model, and provides the p-value, R^2 [Gujarati 2004] (coefficient of determination), and a plot of data and the fitted regression model (in the case of one independent variable).

In the context of business processes, we can look at resource behaviors and outcomes that relate to *cases* (e.g., the relationship between the percentage of tasks executed in a case by a given resource and the case duration or cost) and *tasks* (e.g., the relationship between a task outcome and the experience of a resource executing the task), or we can look at resource behaviors and outcomes during a given period of *time* (e.g., the relationship between the number of distinct tasks completed per week and the number of task instances completed during the week). Next we describe how to analyze whether or not relationships exist between given resource behaviors and outcomes for the three perspectives referred to here as the *case*, *task*, and *time* perspectives. The procedure consists of the following steps: (1) define a set of instances to be included in the analysis (i.e., cases, task instances, or time slots), (2) define the value of the dependent variable for a given instance, and (3) define the value of an independent variable for a given instance.

Case Perspective.

(1) Define a set of cases to be included in the analysis, $\mathcal{C}_{RA}$. The cases in which a user may be interested include all completed cases, cases in which a given resource was involved, or cases of a certain type. For example, $\mathcal{C}_{RA}$ may comprise all cases in which a given resource, r, was involved: $\mathcal{C}_{RA} \triangleq C_R(r)$.
(2) Define the value of the dependent variable DV for a given case $c \in \mathcal{C}_{RA}$ (e.g., one may want to check whether some resource behavior affects the duration or cost of cases). For example, $DV(c)$ may yield the duration of case c: $DV(c) \triangleq c_{case\_duration}$.
(3) Define the value of an independent variable IV_i for a given case $c \in \mathcal{C}_{RA}$ (e.g., one may want to check whether the percentage of tasks completed by a given resource in the case or the number of resources involved in the case affects the case outcome). For example, $IV_i(c)$ may yield the number of resources involved in case c: $IV_i(c) \triangleq |\{r \in R \mid \exists e \in EL[e_{resource} = r \wedge e_{caseid} = c]\}|$.

Task Perspective.

(1) Define a set of events that should be analyzed, $\mathcal{E}_{RA}$. A user may wish to include in the analysis the events related to completed tasks, instances of a given task, or tasks completed by a given resource. For example, $\mathcal{E}_{RA}$ may comprise all events related to completed instances of a given activity a: $\mathcal{E}_{RA} \triangleq \{e \in \mathcal{E} \mid e_{task} = a \wedge e_{type} = complete\}$.
(2) Define the value of the dependent variable DV for a given event $e \in \mathcal{E}_{RA}$. For example, $DV(e)$ may yield the task duration: $DV(e) \triangleq e_{task\_duration}$.
(3) Define the value of an independent variable IV_i for a given event $e \in \mathcal{E}_{RA}$ (e.g., the experience or workload of a resource that executed the related task). For example, $IV_i(e)$ may yield the resource's workload: $IV_i(e) \triangleq e_{workload}$.

Time Perspective.

(1) Define time series parameters, namely the starting time point for the analysis, sampling rate (e.g., a day or a week), and the number of time slots in which one is interested. For a given time slot, there can be only one value of the dependent variable and one value of each independent variable.
(2) Define the value of the dependent variable DV during a given time slot (t_1, t_2) (e.g., the number of task instances completed by a resource or the fraction of cases with a given outcome completed in which a resource was involved with respect to the total number of cases completed in which the resource was involved). For example, a user may look at the average duration of a task completed by a given resource, r: $DV(t_1, t_2, r) \triangleq \sum_{e \in E_{CTR}(t_1,t_2,r)} e_{task\_duration} / |E_{CTR}(t_1, t_2, r)|$.
(3) Define the value of an independent variable IV_i during a time slot (t_1, t_2) (e.g., the average resource workload). For example, a user may look at the number of distinct tasks completed by a resource r: $IV_i(t_1, t_2, r) \triangleq |\{a \in A \mid \exists e \in E_{CTR}(t_1, t_2, r)[e_{task} = a]\}|$.

The framework then extracts from an event log the values of the defined dependent variable DV and all independent variables IV_i for all cases in $\mathcal{C}_r$ (for the *case* perspective), events in $\mathcal{E}_r$ (for the *task* perspective), or time slots (for the *time* perspective), such as weekly values of the average duration of tasks completed by a resource and weekly values of the average workload of the resource. It runs regression analysis using the variable values and provides R^2, the p-value, the fitted regression model details (in the case of multiple independent variables), and a plot (when one independent variable is considered; an example is provided later in Figure 15) that allow us to identify whether or not a relationship exists between given resource behaviors and outcomes.

3.4. Evaluating Resource Productivity

The goal of the *Evaluating Resource Productivity* module is to provide a method for evaluating and comparing the productivity of human resources and tracking the evolution of productivity over time (Research Question 3 in Section 1). If resources only execute one task, use the same inputs needed to perform this task (e.g., information or materials), and their working hours are the same, then it is possible to compare the productivity of these resources by simply counting the number of task instances completed by each resource during a given time period. However, this scenario is not realistic because resources in modern organizations are often involved in multiple tasks with different levels of complexity.

To be able to evaluate resource productivity in such complex environments, we use the DEA [Bogetoft and Otto 2011] technique. DEA takes as input combinations of inputs and outputs that were previously observed and estimates the best practice (the "efficient frontier") [Bogetoft and Otto 2011]. This is a nonparametric method based on linear programming that can consider multiple inputs and outputs and is often used in operations research to compare the efficiency of companies [Bogetoft and Otto 2011], departments or company functions [Thanassoulis 1995; Tavakoli and Shirouyehzad 2013], and business processes [Dohmen and Leyer 2010], with typical inputs being money, materials, and people and typical outputs being units of production. DEA does not require knowledge of the relation between inputs and outputs and is not limited to a certain underlying stochastic function of the data [Dohmen and Leyer 2010; Cooper et al. 2011]. Thus, there are not many disadvantages compared to other efficiency evaluation methods like free disposal hull [Cooper et al. 2011]. It has been shown that it can be applied on a process level using process instances as decision-making units without severe limitations [Burger and Moormann 2008] and has been applied to assess resource behavior in general without using data from event logs [Koch-Rogge et al. 2014].

In many organizations, the amount of output (e.g., the number of completions of a given task) that should be produced by a resource in a given role using a given input (e.g., working hours) is unknown. The first challenge is to be able to identify productivity "best practice" for the role from data. The DEA technique extracts from data the "set of combinations of input and output such that the input can actually produce the output" (referred to as the *technology*) and estimates the best practice for the technology (referred to as the *efficient frontier*) [Bogetoft and Otto 2011]. When estimating the efficient frontier, users may select one of the four DEA models based on their return-to-scale assumptions: constant returns to scale (CRS), decreasing returns to scale (DRS), increasing returns to scale (IRS), or variable returns to scale (VRS) [Bogetoft and Otto 2011]. By default, we use the CRS model. In the context of business processes, a resource's working hours would typically be considered as an input. It seems reasonable to expect that the amount of a resource's output should be proportional to the amount of the resource's working time. For example, the number of tasks completed by a resource who works 40 hours per week should be twice as high as the number of tasks completed by a resource who only works 20 hours per week.

DEA also allows us to assign costs to inputs and outputs (if they are available) and to calculate optimal revenue output, optimal cost input, or optimal profit [Bogetoft and Otto 2011]. For example, a user may wish to assign costs to tasks with different levels of complexity and consider them when estimating productivity best practice.

After learning the efficient frontier—that is, the productivity best practice—we would like to measure the productivity of resources with respect to this frontier. DEA also allows us to measure efficiency as related to the efficient frontier [Bogetoft and Otto 2011]. The productivity score for a given resource during a given time period is a value from 0 to 1, with 1 being the maximum possible productivity. If a resource's productivity score during a given time period is 1, this means that the combination of the resource's inputs and outputs during this time period belongs to the efficient frontier.

Our framework allows users to define inputs and outputs (and their costs if they are known) for a given resource during a given time slot, estimate the efficient frontier using the defined inputs and outputs from an event log, and evaluate resource productivity as related to this frontier.

As the first step, users should define the values of the resource inputs and outputs for a given time slot. On many occasions, resources use the same inputs (e.g., tools or information) that they require to conduct their work. For such cases, there is no need to define the inputs, as they are constant. However, in some cases, the inputs are necessary. Resource inputs and outputs that can be handled by our framework depend on information recorded in an event log. Here we focus on resource inputs and outputs that can be extracted from a basic event log. We also provide examples of resource inputs and outputs that can be extracted from richer event logs; however, this is a direction for future work, as we explain in Section 5.2. Let $R_i(t_1, t_2, r, [p_1 \dots p_n])$ denote resource input and $R_o(t_1, t_2, r, [p_1 \dots p_n])$ denote resource output during a given time slot, t_1 to t_2, for resource r and for an optional set of other parameters $p_1 \dots p_n$, for example, for a given activity:

—*Resource inputs* that can be extracted from a basic event log include (1) the amount of time a resource works on a process during a given time slot (it should be considered as an input if it varies for different resources or if a casual worker's hours vary from week to week) and (2) the number of cases active during a given time slot (it should be considered as input if the number of active cases varies during different periods so that productivity scores are not affected during periods with fewer cases). For example, if we consider as input active cases, resource input can be defined as $R_i(t_1, t_2, r) \triangleq |\{c \in \mathcal{C} \mid \exists e \in EL[e_{caseid} = c_{caseid} \wedge e_{time} \geq t_1 \wedge e_{time} < t_2]\}|$. Resource skills and experience are examples of resource inputs that require a richer log.

—*Resource outputs* that can be extracted from a basic event log include (1) the number of instances of a given task completed by a resource during a given time slot, (2) the number of cases completed during a given time slot in which a resource was involved, and (3) the number of task instances completed by a resource during a given time slot. For example, if we consider as output task completions, resource output can be defined as $R_o(t_1, t_2, r) \triangleq |E_{CTR}(t_1, t_2, r)|$. An example of resource output that requires a richer event log is the number of cases of a given type (e.g., for a given product or of a certain complexity) completed during a given time slot in which a resource was involved.

As the second step, the user selects which data should be used to estimate the efficient frontier. Our framework provides three options:

(1) *Consider inputs and outputs for one resource during different time periods*. This way, the framework will estimate a "personal" efficient frontier. An assumption

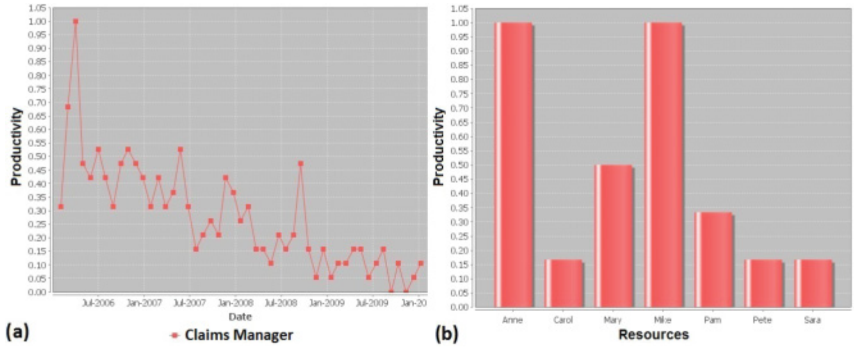

Fig. 6. Examples of outputs for resource productivity evaluation.

here is that the frontier does not change over time, although the productivity of a resource can change.

(2) *Use inputs and outputs for multiple resources during one time period*. An assumption here is that the efficient frontier is the same for the resources but their individual productivity can be different; hence, the user should only consider resources that do the same type of work.

(3) *Use inputs and outputs for multiple resources during different time periods*. An assumption here is that the efficient frontier is the same for the resources and does not change over time, but the resources can have different levels of productivity that can change over time.

Our framework uses defined inputs and outputs and extracts from an event log a set of observed input-output combinations for selected resources during selected time periods. It also allows users to define costs of inputs and outputs if they are available. It then estimates the efficient frontier using the set of input-output values, calculates resource productivity as related to the frontier, and produces a chart illustrating the productivity. Figure 6(a) shows an example of output for productivity evaluation for one resource during different time periods, Figure 6(b) shows an example of output for productivity evaluation for multiple resources during one time period, and an example of output for productivity evaluation for two resources during different time periods is provided later in Figure 16.

4. VALIDATION

This section presents two types of validation for our framework. We first apply our framework to analyze the behaviors of employees in an Australian company. The case study demonstrates real-life examples of the analyses that can be performed by the three modules of our framework (Sections 4.1.1 through 4.1.3) and is followed by a discussion of the insights gained from the case study (Section 4.1.4). We then present the results of an online survey conducted with managers to evaluate their opinions about the usefulness of the framework. The framework has been implemented as a plug-in[7] of the ProM[8] process mining framework as described previously [Pika 2015].

4.1. Analyzing Employee Behaviors in an Australian Company

We applied our framework to analyze employee behaviors in an Australian company. The company was interested in analyzing the behaviors of 34 selected employees to

[7]http://yawlfoundation.org/risk/files/MiningResourceProfiles.7z.
[8]http://www.promtools.org.

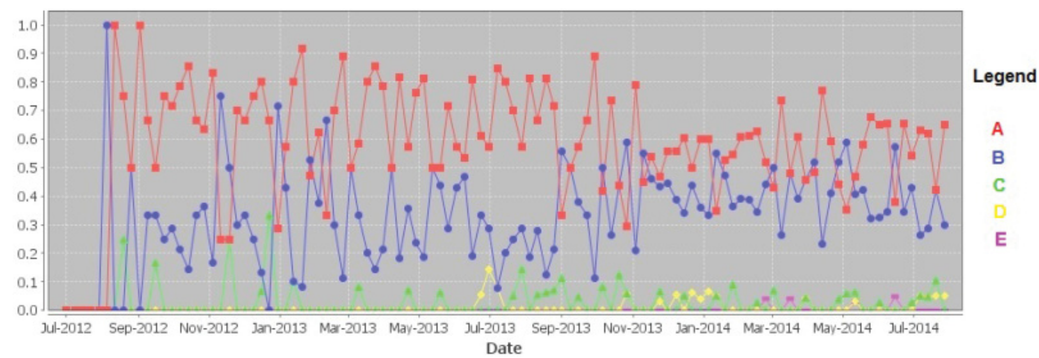

Fig. 7. RBI 1.2, "Case types," for employee R10: Fractions of cases related to five products.

(1) evaluate the performance of the employees (the company representative already had an idea about the performance levels of different employees and was interested to see if the analysis results would confirm the assumptions),
(2) identify opportunities for improvement of their performance, and
(3) identify opportunities for improving process performance.

The data was collected from different information systems, anonymized, and cleaned up based on input from process experts. It was an iterative process that involved several meetings with company representatives during which we discussed the data requirements and agreed on the data attributes used in the analysis. The resulting event log contained 17,750 cases and more than 700,000 events. Events in the log contained five basic attributes (i.e., *caseid*, *task*, *type*, *time*, and *resource*) and also attributes *case_type*, *feedback* (customer feedback), and *experience* (the number of days an employee had been employed by the company at the time of event occurrence). The original data contained information about roles and employees involved in events. We created two versions of the event log. In the first version of the log, we only selected role information and associated it with attribute *resource*. This version was used to demonstrate role-based analysis (RBI 2.1 in Section 4.1.1). In the second version of the event log, we only selected employee information and associated it with attribute *resource*. This version was used to analyze employee behaviors (all other analyses in the case study). We analyzed weekly behaviors of the 34 employees over a period of 2 years. The employees were only involved in one process represented in the event log used in the analysis. The company's feedback on the usefulness of the analysis results was collected via an unstructured interview. In the following sections, we first provide examples from the case study illustrating the three modules of our framework and then discuss the insights gained from the case study and show how they can help improve process and employee performance.

4.1.1. Module 1: Analyzing Resource Behavior. We demonstrate examples of RBIs from each category of resource behavior. An example of an indicator that reflects resource skills is depicted in Figure 7 (RBI 1.2, "Case types"). It shows the fractions of completed cases related to five products (referred to here as A, B, C, D, and E), in which employee R10 was involved with respect to the total number of completed cases in which R10 was involved. The chart reveals that employee R10 was typically involved in cases related to products A and B, but we can see that the employee was not very experienced in cases related to product E.

Figure 8 depicts the values of RBI 2.2, "Number of cases completed in which a resource was involved" (a utilization indicator), for employees R4 and R5 who play the

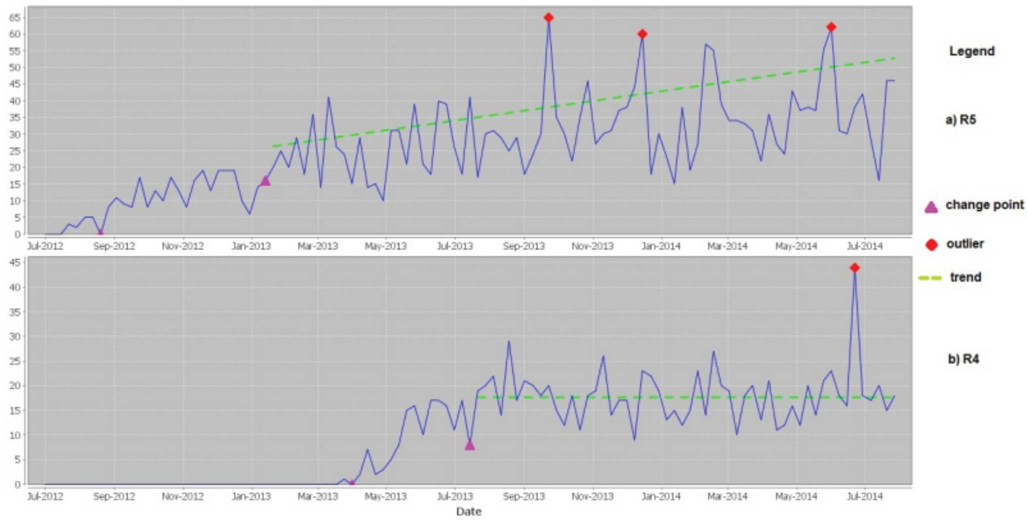

Fig. 8. RBI 2.2: Number of cases completed per week in which resources R5 (a) and R4 (b) were involved.

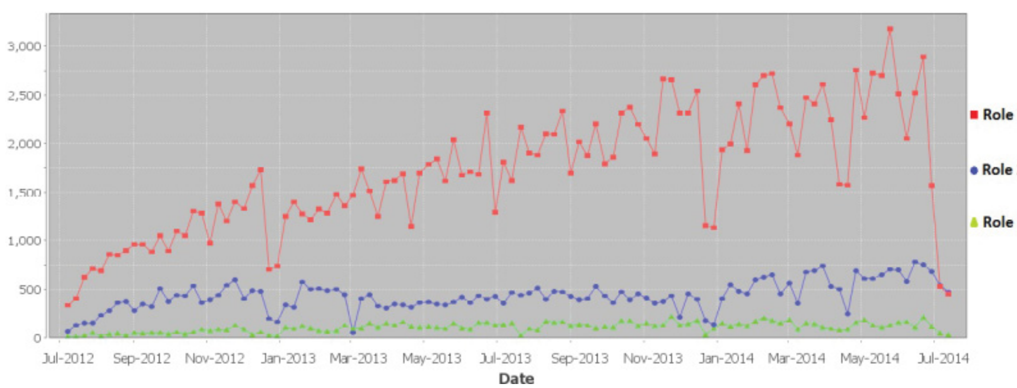

Fig. 9. RBI 2.1: Number of task instances completed per week by employees in Role 1, Role 2, and Role 3.

same role. We can observe an upward trend for employee R5: the number of cases in which R5 was typically involved gradually increased from 25 cases per week in January 2013 to 50 cases in July 2014, whereas the number of cases in which R4 was involved was stable starting from July 2013 at around 17 cases per week. Having analyzed this utilization indicator for the 34 employees, we could observe three types of behavior: some employees were becoming more active over time, other employees displayed stable behavior, and some employees showed irregular patterns being more or less active during different periods. Figure 9 depicts another example of a utilization indicator (RBI 2.1) for three roles, referred to here as Role 1, Role 2, and Role 3. Although the numbers of task instances completed during the week by Role 2 and Role 3 did not change significantly, employees in Role 1 tended to complete more task instances over time.

An example of an indicator that illustrates a resource's preferences is depicted in Figure 10. It shows the number of task reassignments from employees R3 and R15 (RBI 3.3, "Activity reassignments"). As we can see, starting from May 2013, the number of times a task assigned to employee R3 was completed by another employee increased,

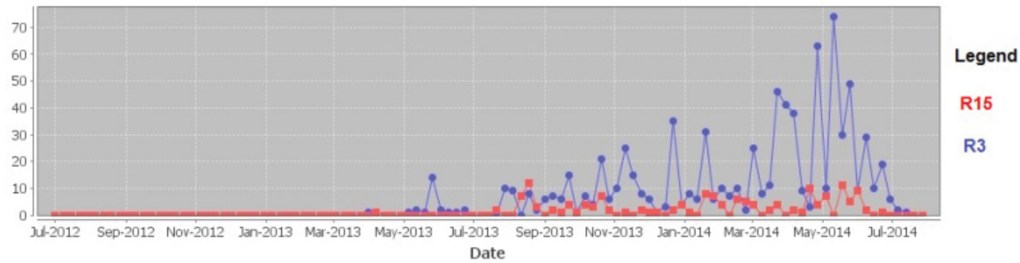

Fig. 10. RBI 3.3: Number of task reassignments from employees R15 and R3.

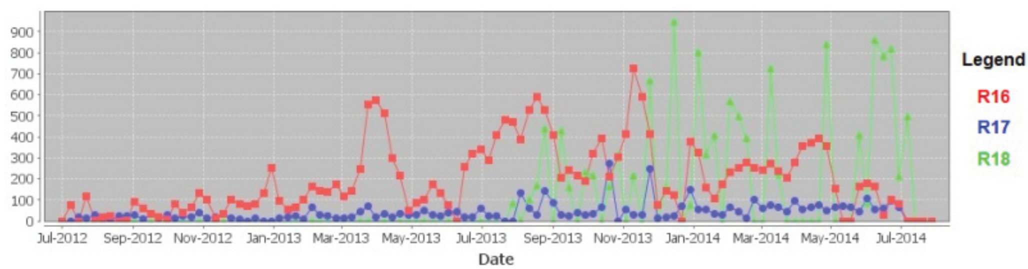

Fig. 11. RBI 4.3: Average duration of a given task completed by an employee.

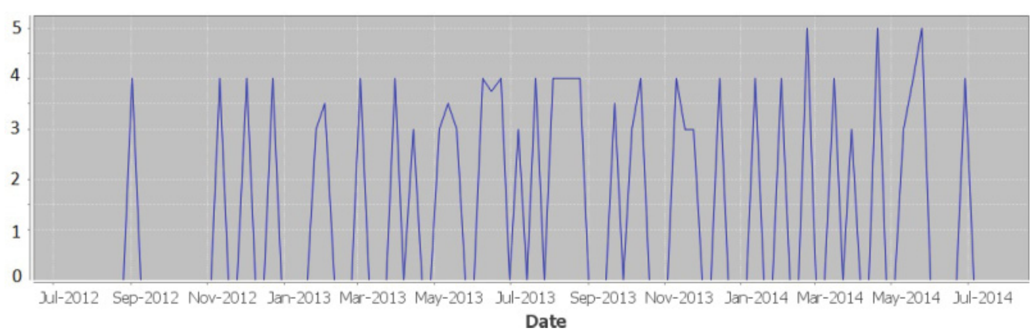

Fig. 12. RBI 4.5: Average customer feedback for cases managed by R11.

and during some weeks it was around 70. The number of task reassignments from employee R15 also increased from May 2013, but it typically did not exceed 10. We could observe significant differences in the number of task reassignments for different employees having the same role. Although for some of them the high number of task reassignments can be explained by a high workload, the reasons for frequent task reassignments from other employees should be investigated.

The values of RBI 4.3, "Average duration of a given activity" (a productivity indicator), for employees R16, R17, and R18 are depicted in Figure 11. We can clearly see that employee R17 was much faster in the execution of this activity (the average duration was typically below 100 hours) than the two other employees for whom the average duration of the task sometimes reached more than 500 hours. We could observe similar differences for other tasks and employees. The company will explore ways of improving process performance by creating groups of efficient employees for each task. Another example of a productivity indicator is depicted in Figure 12, showing RBI 4.5, "Average customer feedback," for cases managed by employee R11. The values of this indicator

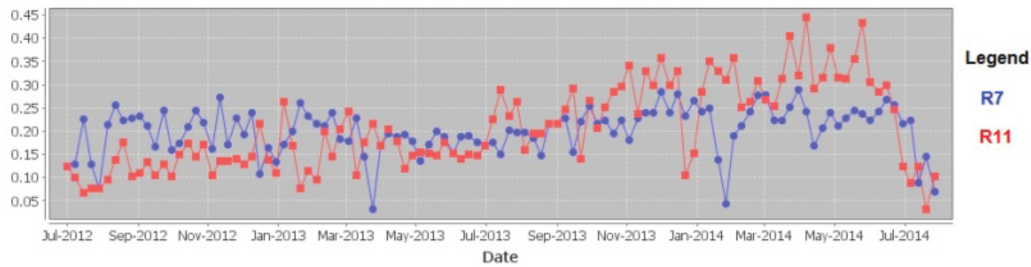

Fig. 13. RBI 5.2: Fraction of employees during a week involved in the same cases with an employee.

are based on a sample of cases rated by customers. The indicator values ranged between three and four during most weeks, which means that the quality of service was rated by customers "as expected" (value three) or "higher than expected" (value four) for cases managed by R11 (value zero means that no cases were completed during the week that were rated by customers and managed by R11). During some weeks, the average customer feedback was "much higher than expected" (value five).

Figure 13 depicts the values of RBI 5.2, "Social position," for employees R7 and R11. It is an example of a collaboration indicator and it is the fraction of resources active during a week who were involved in the same cases with a given resource with respect to the total number of resources active during the week. We can see in this figure that the values of the indicator were stable for employee R7 but gradually increased for employee R11 until June 2014, which means that employee R11 was becoming involved with more employees in the same cases over time.

To obtain a more complete picture of a resource's behavior, it is necessary to look at the resource's profile. As an example, Figure 14 depicts a profile of employee R30 that comprises five RBIs from different categories of resource behavior. We can observe that there was a change in the employee's behavior around March 2014. After this point in time, the employee performed fewer distinct tasks (Figure 14(a)) and was involved in more cases (Figure 14(b)) of shorter duration on average (Figure 14(d)). The fraction of employees involved in the same cases with R30 increased (Figure 14(e)). We can also observe that the number of task reassignments from employee R30 was typically not very high and that tasks were not reassigned from R30 before February 2013 (Figure 14(c)). The changes in resource behavior that we can observe in Figure 14(a), (b), (d), and (e) are related to changes in the employee's set of responsibilities in the organization.

4.1.2. Module 2: Quantifying the Outcome of Resource Behavior. An example of a process outcome that is available in a typical event log is duration. We consider duration from three perspectives, namely case, task, and time, and show examples from the case study for each perspective. Based on our knowledge of the process and the data, we did not identify any confounding factors that had to be eliminated in the procedure of performing the regression analysis.

We first looked at the *case* perspective and checked the relationship between the number of employees involved in a case (independent variable) and the case duration (dependent variable). All cases were included in the analysis. In our previous research, we found that in some processes, having higher numbers of resources involved in a case is often associated with long-running cases [Pika et al. 2013a]. We discovered for this process that there is only a weak relationship between the two variables (R^2 is 0.24, $p < .001$; the data and the fitted regression line are depicted in Figure 15(a).

We then looked at the *task* perspective and checked the relationship between the duration of a task (dependent variable) and the experience of the employee who completed the task (independent variable). Instances of all tasks were included in the analysis.

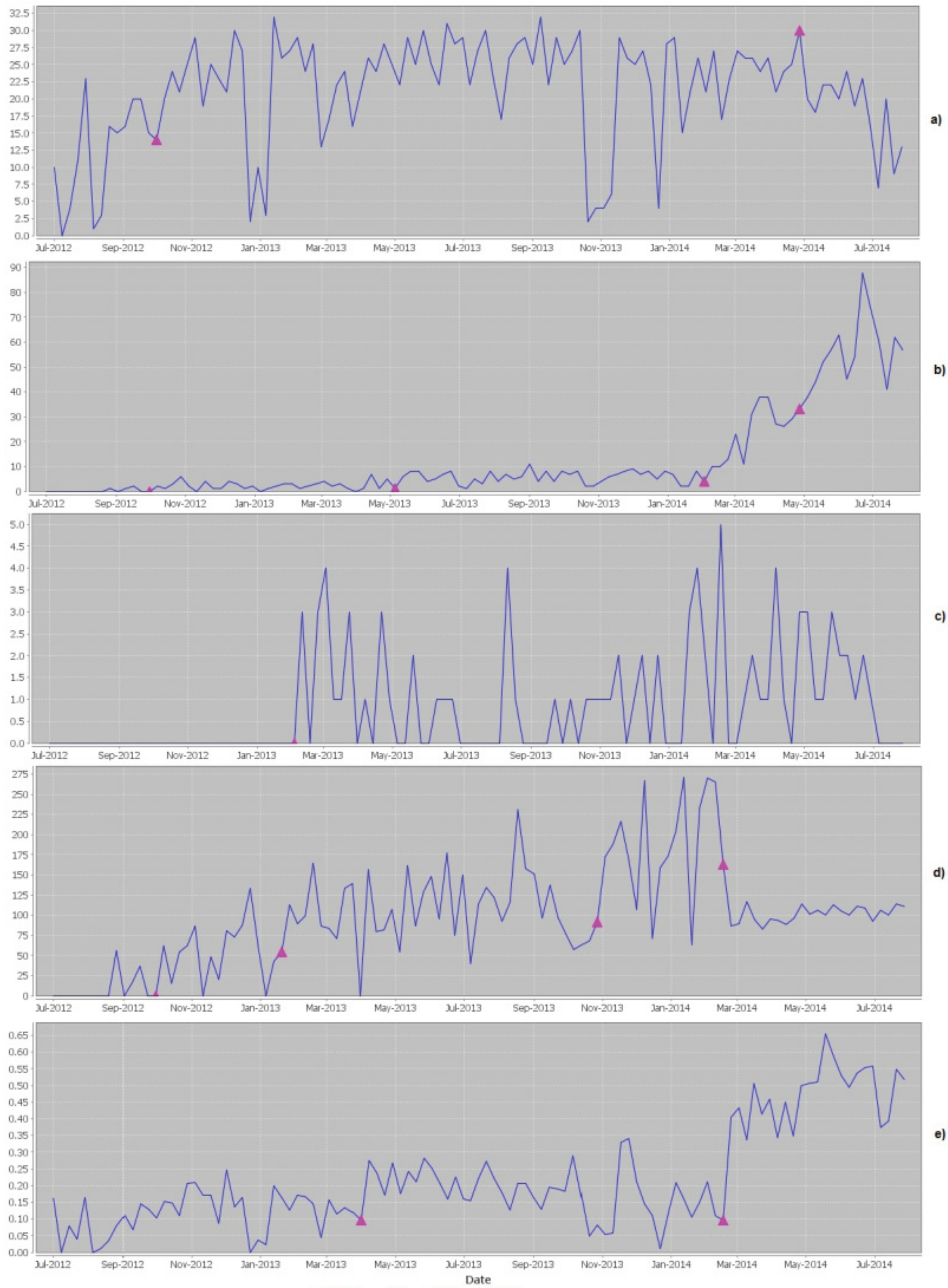

Fig. 14. Profile of employee R30. (a) RBI 1.1: Number of distinct tasks completed. (b) RBI 2.2: Number of cases in which employee R30 was involved. (c) RBI 3.3: Number of task reassignments from R30. (d) RBI 4.4: Average duration of cases in which R30 was involved. (e) RBI 5.2: Fraction of employees involved in the same cases with R30.

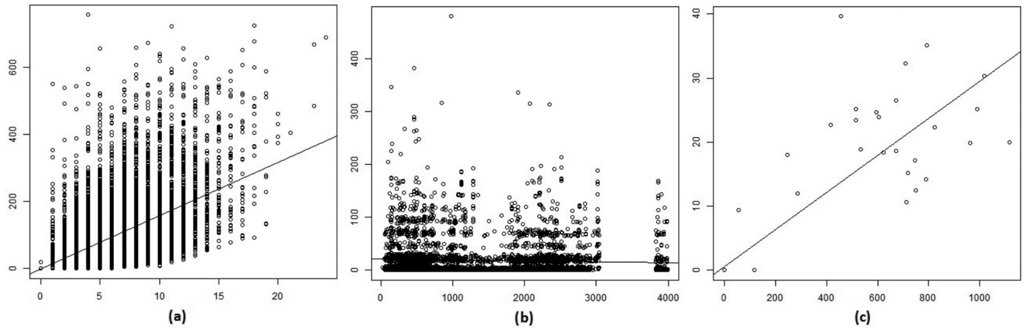

Fig. 15. (a) Relationship between the number of employees involved in a case(x-axis) and the case duration (y-axis). (b) Relationship between experience of an employee executing a task (x-axis) and the task duration (y-axis). (c) Relationship between the number of tasks completed by employee R5 per week (x-axis) and average duration of the tasks (y-axis).

Experience was estimated as the number of days the employee had been employed by the company at the time of task completion. We found that employee experience had no influence on the tasks' duration (R^2 is 0.002, $p < .001$) (Figure 15(b)).

We also looked at the *time* perspective and checked how the number of tasks completed by an employee during a week (independent variable) affected the average duration of the tasks (dependent variable) for employees R5 and R7. In both cases, workload affects task duration; however, the effect was strong for employee R5 (R^2 is 0.79, $p < .001$) (Figure 15(c)) and weak for employee R7 (R^2 is 0.23, $p < .001$). Based on these observations, the company will explore ways to balance the workload of employees whose efficiency suffers when they are overloaded.

4.1.3. Module 3: Evaluating Resource Productivity. We calculated weekly productivity scores for the 34 selected employees for a period of 109 weeks. We considered as an output the number of executions of a given activity. The company representative assigned the level of complexity to each activity (either high, medium, or low), which we used to assign output costs (either five, three, or one, respectively). We did not use inputs, as the employees had access to the same resources needed to perform their work (e.g., information). The 34 selected employees belonged to five different roles, and we learned from data the optimal revenue output (i.e., the estimated efficient frontier considering the costs of outputs (Section 3.4)) separately for each role. Then the weekly productivity scores for each employee were calculated as the division of the observed revenue output of an employee by the optimal revenue output calculated for the employee's role. The resulting productivity scores range from zero to one, with one being the maximum possible productivity.[9]

As an example, Figure 16 depicts the productivity scores for two employees in the same role. We can observe that the productivity of employee R5 was between 0.4 and 0.6 most of the time, and it gradually decreased from March 2014. The productivity of employee R9 was typically lower than the productivity of R5, and it varied during different periods. Having extracted the productivity scores for the 34 employees, we could observe that some employees in the same role had stable and similar productivity scores, whereas others displayed irregular productivity patterns. In some instances, the productivity scores extracted from the data supported the beliefs of the company representative about the employees' productivity, whereas the productivity

[9]This case study does not fully demonstrate the benefit of using DEA for productivity evaluation, as we did not use inputs and the levels of activity complexity were assigned by the company representative.

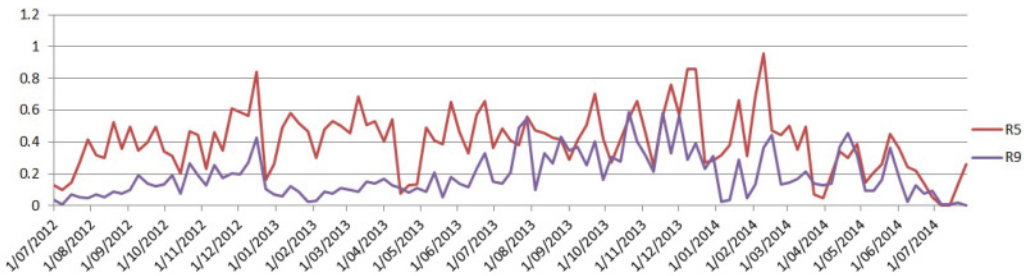

Fig. 16. Weekly productivity scores for employees R5 and R9.

scores of other employees were not as expected (i.e., the scores were higher or lower than expected).

4.1.4. Summary of the Case Study. This section summarizes our findings from the case study and discusses the lessons learned. The first goal was to perform an evidence-based evaluation of the performance of the selected 34 employees. The company was interested to see if our findings confirmed its assumptions. We extracted indicators in the five categories of resource behavior and evaluated the overall productivity of the employees. In most cases, the analysis results confirmed the assumptions identifying better and poorer performers. However, we found that some employees suspected of inconsistent performance actually displayed stable behavioral patterns (e.g., in terms of the number of cases and tasks completed and the average duration of the tasks). From the analysis of the behaviors of employees in this process, we could see that abrupt significant changes in employee behaviors were typically related to changes in role or changes in portfolio and outliers were often related to periods of vacations, whereas gradual changes in behaviors often required attention, as they indicated changes in employee performance. We also discovered that some employees were very active (e.g., involved in many cases and tasks); however, they became slower in the execution of some tasks over time, tasks were frequently reassigned from them, or they involved many employees in the cases they managed. This shows us that due care should be taken with the interpretation of results. Each indicator is an abstraction that highlights a certain aspect of resource behavior while discarding others. To obtain a more complete picture of employee performance, it would be necessary to look at different dimensions.

The second goal was to identify opportunities for improvement in the performance of the employees. We learned from the analysis that there were significant differences in the amount of time taken to complete a particular task for different employees; hence, the company will be exploring the possibility of assigning tasks to those employees who are more efficient in the execution of a particular task. We also discovered that for some employees, the speed of task processing was significantly affected by their workload. The company plans to explore ways to balance the workload of such employees.

Finally, the third goal was to identify opportunities for improving process performance. On the basis of the results, it is expected that case throughput times may be reduced by defining pools of efficient employees for each task and assigning tasks to efficient resources. Balancing employee workload should also contribute to shorter case processing times. We also identified differences in the number of employees processing cases handled by different case managers. The company will be looking into these cases in more detail to identify the reasons and suggest best practices for case resourcing.

Applying the framework to analyze the behaviors of employees in an Australian company, we demonstrated the different types of analysis that the framework allows us to perform. We showed how the framework can help to evaluate resource performance

Table I. As-Is Versus Would Be Helpful with Regard to Resource Performance Evaluation

	As-Is (M, SD)	Would Be Helpful (M, SD)
Resource Skills	2.88, 0.633	5.78, 0.909
Resource Utilization	3.12, 0.739	5.80, 1.056
Resource Preferences	2.59, 0.670	5.57, 1.151
Resource Productivity	2.91, 0.726	6.02, 0.869
Resource Collaboration	2.36, 0.533	5.45, 1.214

M, mean; SD, standard deviation; "As-Is" scale: 1–4 ("Please indicate if [*resource behavior category name*] is currently analyzed in your organization when evaluating resource performance": (1) no opinion, (2) not analyzed, (3) analyzed based on personal judgments, (4) analyzed using information provided by information systems); "Would Be Helpful": 7-point Likert scale.

in a more objective way and to identify opportunities for improving the performance of employees and the business process.

4.2. Survey on the Usefulness of the Framework

We conducted an online anonymous survey to evaluate the perceived usefulness of our proposed framework among managers (as they are the intended users of the framework). The survey described the functionality of each module and showed screenshots of examples of its application using a real dataset[10] (the survey questions are available at yawlfoundation.org [11]). Questions were then asked regarding usefulness of the framework features, current status of analyzing resource behavior in respondents' organizations, and process orientation (managers with higher process orientation are expected to see a higher usability of resource performance analyses [Forsberg et al. 1999]).

The questionnaires were distributed using professional social media (e.g., LinkedIn[12]). They were also sent via email to employees in the financial services industry in Germany using a non–public university database of industry contacts available to a member of the research team (the database includes contacts of 2,286 employees, of whom 34.4% (786) are expected to have managerial positions [Leyer and Moormann 2014]). We asked employees with managerial functions to fill out the questionnaire anonymously. The questionnaire was available in English and in German. In total, 42 managers answered the questionnaire fully (11 in English, 31 in German). The participants had a professional experience of 10.9 years on average (SD: 7.56) and supervised 31.9 (SD: 55.27) employees on average. There was no statistically significant difference between the German and English survey responses; thus, the answers were analyzed jointly. Among the respondents, 59.52% indicated that they worked in the finance and banking industry; others specified the following industries: information technology, consulting services, manufacturing, education, health, and the restaurant business.

The participants reported a limited application of resource behavior analysis in the five categories (the "As-Is" column in Table I) but would find such analyses to be very helpful (the "Would Be Helpful" column in Table I). Overall, the modules (including features within the modules) were considered as very relevant by the participants (Table II). The most relevant features (with the percentage of respondents who selected the feature shown in brackets) are analyzing indicators of resource behavior in the categories of productivity (71.4%), skills (66.7%), and utilization (61.9%); estimating trends (52.4%); comparing the overall resource productivity scores for multiple resources

[10]The survey did not use the dataset from the case study company.
[11]http://yawlfoundation.org/risk/files/Survey.pdf.
[12]http://www.linkedin.com.

Table II. Perceived Usefulness of the Framework and Its Modules

Module	(Mean, Standard Deviation)*
Framework	(5.71, 1.031)
Analyzing Resource Behavior	(5.91, 0.878)
Quantifying the Outcome of Resource Behavior	(5.83, 0.824)
Evaluating Resource Productivity	(5.86, 1.026)

*7-point Likert scale.

during different time periods (52.4%), and finding correlations between resource behaviors and outcomes for the time perspective (50.0%).

There is a weak statistically significant influence of process orientation on the evaluation of the modules (ns, nonsignificant): *Analyzing Resource Behavior* (ns (0.427), $R^2 = -0.020$), *Quantifying the Outcome of Resource Behavior* ($p < .05$ (2.039), $R^2 = 0.072$), *Evaluating Resource Productivity* ($p < .05$ (2.227), $R^2 = 0.088$), overall evaluation of the framework (ns (1.942), $R^2 = 0.065$). As a consequence, it can be stated that the degree to which the framework was considered relevant by a manager is almost independent of the degree of process orientation of their environment. In addition, the number of subordinates and length of working experience did not have any statistically significant influence on the degree to which the framework was considered relevant.

The survey revealed that only a limited amount of resource analysis is conducted in companies. It confirmed that managers find our framework useful for resource performance evaluation and that their opinion of potential usefulness of the framework is independent of the level of process orientation of their environments, their length of experience, or the number of subordinates.

5. DISCUSSION

In this section, we first discuss the contributions of our work and compare it to existing approaches. We then summarize assumptions and limitations of our framework and discuss the ways in which our work can be extended.

5.1. Contributions

Human performance measurement and planning are challenging tasks [Espinilla et al. 2013; Peretz and Fried 2012], as human behavior is complex [Leftwich 2015] and changes over time [Beheshti et al. 2016]. Despite the important role human resources play in business processes [Rosemann and vom Brocke 2015] and the increase in knowledge-intensive tasks in modern organizations [Di Ciccio et al. 2015; Freel 2016; Wirtz and Lovelock 2016], the BPM community has paid little attention to collaborative aspects of BPM [Di Ciccio et al. 2015]. On the other hand, traditional resource performance evaluation approaches proposed in the management literature are criticized for lack of objectivity [Espinilla et al. 2013] and narrow focus [Neely et al. 2000]. In this article, we presented a framework for mining resource profiles from process execution data. Knowledge about resource behaviors that can be extracted by our framework enables evidence-based performance evaluation and may provide insights for better workload planning, effective human resource development strategies, and resource and process performance improvement, as we demonstrated in Section 4.

The framework presented in this article is an extension of our earlier work [Pika et al. 2014], in which we proposed a method for analyzing resource behavior. Here we presented an extended framework that also includes a new method for quantifying the outcome of resource behavior (Section 3.3) and a new method for evaluating and comparing resource productivity (Section 3.4). We presented the results of a new case study and a survey evaluating the usefulness of our framework (Section 4).

Some technical aspects of our framework were inspired by existing process mining approaches [Pika et al. 2013a, 2013b; de Leoni et al. 2014; van der Aalst 2013; Bose et al. 2013]: we use indicators extracted from event log data, our framework allows the correlation of resource behaviors and outcomes, it allows analyzing resource behavior from different perspectives, and it considers changes in resource behavior over time; however, unlike the preceding approaches, the focus of our framework is resource behavior analysis rather than business processes.

We adopted DEA [Bogetoft and Otto 2011] for resource productivity evaluation. DEA was previously used to evaluate productivity of employees [Manoharan et al. 2009; Wagner et al. 2003; Koch-Rogge et al. 2014]. However, in these case studies, data was collected from different sources and manually transformed into a form suitable for DEA analysis. Our goal, on the other hand, was not to demonstrate the use of DEA in a particular company but to present a method that allows us to define resource inputs and outputs, extract their values from an event log, and evaluate resource productivity.

Existing organizational mining approaches extract organizational structures [Song and van der Aalst 2008; van der Aalst et al. 2005], devise resource allocation mechanisms [Liu et al. 2012; Cabanillas et al. 2013; Kumar et al. 2013; Ly et al. 2006], or focus on few specific resource behaviors [Huang et al. 2012; Nakatumba and van der Aalst 2010]. In contrast to these approaches, in this article we presented a framework that focuses on employees, allows analyzing their behaviors from different perspectives, is extensible (allows definition of new measures), and considers the evolution of resource behaviors over time.

5.2. Assumptions, Limitations, and Future Work

We assume that the data recorded in an event log accurately represents a company's business process. If the event log contains incomplete or inaccurate information (e.g., some events or attributes are not recorded or the timestamps of events do not represent the time of event occurrence), then the analysis results may be incorrect. We assume that resources are involved in one process only; otherwise, logs from different processes should be merged, ensuring that the case identifiers in the combined log are unique. Alternatively, the proportion of time a resource spent on the process under analysis should be known to ensure fair comparisons across resources.

We provided examples of measures of resource behavior in each of the three modules that can be mined from typical event logs. The option of defining more comprehensive sets of resource behavior measures relevant in different contexts can be explored.

We evaluated the usefulness of our framework by showing examples of its application and conducting an online survey among managers. We also conducted a real case study using a dataset containing the events from a core business process in an Australian company recorded over a 2-year period. We collected feedback from the case study company via an unstructured interview. A manager from the case study company was not asked to participate in the online survey. To evaluate ease of use of the corresponding software artifact, a study can be conducted in which users gain hands-on experience with the tool and provide their feedback.

The framework is based on the analysis of event log data; hence, it can only provide knowledge about the process-related behaviors of resources. Such knowledge can potentially be used in an inappropriate way. For example, a narrow set of resource behaviors may be used to instigate disciplinary action. The possibility of eliminating potential framework misuse should be further investigated. Countries may have different regulations regarding privacy and personal data protection. For example, an employer may need to inform employees about data collection and analysis [van der Aalst et al. 2005], or permission of the relevant employee union may be required (as is the case in Germany). A direction for future work is providing users with an option to select types of resource analysis depending on the regulations that exist.

6. CONCLUSIONS

Managers need objective information about the different working behaviors of resources to better understand areas of performance improvement and to make more informed resource-related decisions. Such knowledge may enable better workload planning, improved performance appraisal systems, and more efficient resource development strategies and provide insights for resource and process performance improvement.

The framework we presented in this article contributes to the process mining field, specifically to the organizational mining area. In contrast to prior approaches, our framework allows users to extract knowledge about different aspects of the behavior of individual employees and teams, either using predefined indicators or defining their own measures, to investigate the impact of given resource behaviors on process outcomes, evaluate resource productivity, and understand changes in resource behavior over time.

We performed two types of evaluation. We conducted a case study applying our framework to an event log from an Australian company to analyze employee behaviors in the company. We demonstrated how the framework can be used to gain insights into performance improvement. We also conducted an online survey of managers to evaluate the perceived usefulness of the framework and its modules. We found that most respondents did not evaluate many aspects of resource behavior, or they did it solely based on their subjective judgments. They stated that the methods provided by our framework would be useful and relevant to them.

ACKNOWLEDGMENTS

We would like to thank the process experts from the case study company for their help and insights shared with us during the case study.

REFERENCES

S. M. R. Beheshti, B. Benatallah, S. Sakr, D. Grigori, H. M. Motahari-Nezhad, M. C. Barukh, A. Gater, and S. H. Ryu. 2016. *Process Analytics: Concepts and Techniques for Querying and Analyzing Process Data*. Springer.

P. Bogetoft and L. Otto. 2011. *Benchmarking with DEA, SFA, and R*. Springer, New York, NY.

J. C. R. P. Bose, W. M. P. van der Aalst, I. Zliobaite, and M. Pechenizkiy. 2013. Dealing with concept drifts in process mining. *IEEE Transactions on Neural Networks and Learning Systems* 25, 1, 154–171.

A. Burger and J. Moormann. 2008. Detecting intrinsic inefficiency on process level: Benchmarking of transactions in banking. In *Proceedings of the International Conference on Business Process Management*. 146–157.

C. Cabanillas, J. M. García, M. Resinas, D. Ruiz, J. Mendling, and A. Ruiz-Cortés. 2013. Priority-based human resource allocation in business processes. In *Service-Oriented Computing*. Lecture Notes in Computer Science, Vol. 8274. Springer, 374–388.

P. Ceravolo, A. Azzini, E. Damiani, M. Lazoi, M. Marra, and A. Corallo. 2016. Translating process mining results into intelligible business information. In *Proceedings of the 11th International Knowledge Management in Organizations Conference on the Changing Face of Knowledge Management Impacting Society*. ACM, New York, NY, 14.

C. Di Ciccio, A. Marrella, and A. Russo. 2015. Knowledge-intensive processes: Characteristics, requirements and analysis of contemporary approaches. *Journal on Data Semantics* 4, 1, 29–57.

W. W. Cooper, L. M. Seiford, and J. Zhu. 2011. Data envelopment analysis: History, models, and interpretations. In *Handbook on Data Envelopment Analysis*. Springer, 1–39.

S. de Leeuw and J. P. van den Berg. 2011. Improving operational performance by influencing shopfloor behavior via performance management practices. *Journal of Operations Management* 29, 3, 224–235.

M. de Leoni, W. M. P. van der Aalst, and M. Dees. 2014. A general framework for correlating business process characteristics. In *Business Process Management*. Lecture Notes in Computer Science, Vol. 8659. Springer, 250–266.

A. Dohmen and M. Leyer. 2010. Combining data envelopment analysis and business process simulation for analyzing the efficiency of business processes. In *Proceedings of the International MultiConference on Engineers and Computer Scientists*. 2134–2138.

J. H. Dulebohn and R. D. Johnson. 2013. Human resource metrics and decision support: A classification framework. *Human Resource Management Review* 23, 1, 71–83.

M. Espinilla, R. de Andrés, F. J. Martínez, and L. Martínez. 2013. A 360-degree performance appraisal model dealing with heterogeneous information and dependent criteria. *Information Sciences* 222, 459–471.

T. Forsberg, L. Nilsson, and M. Antoni. 1999. Process orientation: The Swedish experience. *Total Quality Management* 10, 4–5, 540–547.

M. Freel. 2016. *Knowledge-Intensive Business Services: Geography and Innovation*. Routledge.

D. Gujarati. 2004. *Basic Econometrics*. McGraw-Hill.

D. M. Hawkins and Q. Deng. 2010. A nonparametric change-point control chart. *Journal of Quality Technology* 42, 2, 165–173.

Z. Huang, X. Lu, and H. Duan. 2012. Resource behavior measure and application in business process management. *Expert Systems with Applications* 39, 7, 6458–6468.

M. Koch-Rogge, G. Westermann, C. Wilbert, and R. Willis. 2014. How are they doing? The application of data envelopment analysis to develop a strategic performance appraisal on individual level. In R. Sanchez and A. Heene (Eds.), *A Focused Issue on Building New Competences in Dynamic Environments*. Vol. 7. Emerald Group Publishing, Bingley, UK, 193–218.

A. Kumar, R. Dijkman, and M. Song. 2013. Optimal resource assignment in workflows for maximizing cooperation. In *Business Process Management*. Lecture Notes in Computer Science, Vol. 8094. Springer, 235–250.

A. Leftwich (Ed.). 2015. The political approach to human behaviour: People, resources and power. *What Is Politics: The Activity and Its Study*. Polity Press, Cambridge, UK, 100–118.

M. Leyer. 2011. Towards a context-aware analysis of business process performance. In *Proceedings of the 15th Pacific Asia Conference of Information Systems*. http://www.pacis-net.org/file/2011/PACIS2011-103.pdf.

M. Leyer and J. Moormann. 2014. How lean are financial service companies really? Empirical evidence from a large scale study in Germany. *International Journal of Operations and Production Management* 34, 11, 1366–1388.

J. Lijffijt, P. Papapetrou, and K. Puolamaki. 2012. Size matters: Finding the most informative set of window lengths. In *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD'12)*. 451–466.

T. Liu, Y. Cheng, and Z. Ni. 2012. Mining event logs to support workflow resource allocation. *Knowledge-Based Systems* 35, 320–331.

L. T. Ly, S. Rinderle, P. Dadam, and M. Reichert. 2006. Mining staff assignment rules from event-based data. In *Proceedings of the Business Process Management Workshops*. 177–190.

H. B. Mann and D. R. Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics* 18, 1, 50–60.

T. R. Manoharan, C. Muralidharan, and S. G. Deshmukh. 2009. Employee performance appraisal using data envelopment analysis: A case study. *Research and Practice in Human Resource Management* 17, 1, 92–111.

D. C. Montgomery, E. A. Peck, and G. G. Vining. 2012. *Introduction to Linear Regression Analysis*. John Wiley & Sons.

P. Murphy. 1999. Service performance measurement using simple techniques actually works. *Journal of Marketing Practice: Applied Marketing Science* 5, 2, 56–73.

J. Nakatumba and W. M. P. van der Aalst. 2010. Analyzing resource behavior using process mining. In *Business Process Management Workshops*. Lecture Notes in Business Information Processing, Vol. 43. Springer, 69–80.

A. Neely, M. Gregory, and K. Platts. 2005. Performance measurement system design: A literature review and research agenda. *International Journal of Operations and Production Management* 25, 12, 1228–1263.

A. Neely, J. Mills, K. Platts, H. Richards, M. Gregory, M. Bourne, and M. Kennerley. 2000. Performance measurement system design: Developing and testing a process-based approach. *International Journal of Operations and Production Management* 20, 10, 1119–1145.

S. S. Nudurupati, U. S. Bititci, V. Kumar, and F. T. S. Chan. 2011. State of the art literature review on performance measurement. *Computers and Industrial Engineering* 60, 2, 279–290.

H. Peretz and Y. Fried. 2012. National cultures, performance appraisal practices, and organizational absenteeism and turnover: A study across 21 countries. *Journal of Applied Psychology* 97, 2, 448.

A. Pika. 2015. Mining process risks and resource profiles. Ph.D. Dissertation. Queensland University of Technology. http://eprints.qut.edu.au/86079/.

A. Pika, W. M. P. van der Aalst, C. J. Fidge, A. H. M. ter Hofstede, and M. T. Wynn. 2013a. Predicting deadline transgressions using event logs. In *Business Process Management Workshops*. Lecture Notes in Business Information Processing, Vol. 132. Springer, 211–216.

A. Pika, W. M. P van der Aalst, C. J. Fidge, A. H. M. ter Hofstede, and M. T. Wynn. 2013b. Profiling event logs to configure risk indicators for process delays. In *Advanced Information Systems Engineering*. Lecture Notes in Computer Science, Vol. 7908.Springer, 465–481.

A. Pika, W. M. P. van der Aalst, M. T. Wynn, C. J. Fidge, and A. H. M. ter Hofstede. 2016. Evaluating and predicting overall process risk using event logs. *Information Sciences* 352, 98–120.

A. Pika, M. T. Wynn, C. J. Fidge, A. H. M. ter Hofstede, M. Leyer, and W. M. P. van der Aalst. 2014. An extensible framework for analysing resource behaviour using event logs. In *Advanced Information Systems Engineering*. Lecture Notes in Computer Science, Vol. 8484. Springer, 564–579.

J. Racine and Q. Li. 2004. Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics* 119, 1, 99–130.

M. Rosemann and J. vom Brocke. 2015. The six core elements of business process management. In *Handbook on Business Process Management 1*. Springer, 105–122.

G. J. Ross and N. M. Adams. 2012. Two nonparametric control charts for detecting arbitrary distribution changes. *Journal of Quality Technology* 44, 2, 102–116.

R. Sawhney and S. Chason. 2005. Human behavior based exploratory model for successful implementation of lean enterprise in industry. *Performance Improvement Quarterly* 18, 2, 76–96.

M. Song and W. M. P. van der Aalst. 2008. Towards comprehensive support for organizational mining. *Decision Support Systems* 46, 1, 300–317.

J. D. Sterman. 1989. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science* 35, 3, 321–339.

M. M. Tavakoli and H. Shirouyehzad. 2013. Application of PCA/DEA method to evaluate the performance of human capital management: A case study. *Journal of Data Envelopment Analysis and Decision Science* 2013, 1–20.

E. Thanassoulis. 1995. Assessing police forces in England and Wales using data envelopment analysis. *European Journal of Operational Research* 87, 3, 641–657.

V. Thevendran and M. J. Mawdesley. 2004. Perception of human risk factors in construction projects: An exploratory study. *International Journal of Project Management* 22, 2, 131–137.

G. M. Thompson and J. C. Goodale. 2006. Variable employee productivity in workforce scheduling. *European Journal of Operational Research* 170, 2, 376–390.

W. M. P. van der Aalst. 2010. Business process simulation revisited. In *Enterprise and Organizational Modeling and Simulation*. Lecture Notes in Business Information Processing, Vol. 63. Springer, 1–14.

W. M. P. van der Aalst. 2013. Process cubes: Slicing, dicing, rolling up and drilling down event data for process mining. In *Asia Pacific Business Process Management*. Lecture Notes in Business Information Processing, Vol. 159. Springer, 1–22.

W. M. P. van der Aalst. 2015. Business process simulation survival guide. In *Handbook on Business Process Management 1, International Handbooks on Information Systems*, J. vom Brocke and M. Rosemann (Eds.). Springer-Verlag, Berlin, Germany, 337–370.

W. M. P. van der Aalst. 2016. *Process Mining: Data Science in Action*. Springer-Verlag, Berlin, Germany. http://www.springer.com/978-3-662-49850-7.

W. M. P. van der Aalst, H. A. Reijers, and M. Song. 2005. Discovering social networks from event logs. *Computer Supported Cooperative Work* 14, 6, 549–593.

M. P. J. van der Loo. 2010. Distribution based outlier detection for univariate data. Discussion Paper 10003. Statistics Netherlands.

J. M. Wagner, D. G. Shimshak, and M. A. Novak. 2003. Advances in physician profiling: The use of DEA. *Socio-Economic Planning Sciences* 37, 2, 141–163.

M. Weske. 2007. Architectures and methodologies. In *Business Process Management: Concepts, Languages, Architectures*. Springer-Verlag, Berlin, Germany, 304–355.

R. E. White, J. N. Pearson, and J. R. Wilson. 1999. JIT manufacturing: A survey of implementations in small and large US manufacturers. *Management Science* 45, 1, 1–15.

J. Wirtz and C. Lovelock. 2016. *Services Marketing: People, Technology, Strategy* (8th ed.). World Scientific.

Received April 2016; revised September 2016; accepted December 2016