Association for Information Systems
AIS Electronic Library (AISeL)

Wirtschaftsinformatik Proceedings 2013 Wirtschaftsinformatik

2013

Experimental Evaluation of a Process Benchmarking Tool in a Green Business Process Management Context

Matthias Gräuler
University of Osnabrück, Accounting and Information Systems, Osnabrück, Germany, matthias.graeuler@uos.de

Frank Teuteberg
University of Osnabrück, Accounting and Information Systems, Osnabrück, Germany, frank.teuteberg@uos.de

Experimental Evaluation of a Process Benchmarking Tool in a Green Business Process Management Context

Matthias Gräuler and Frank Teuteberg

University of Osnabrück, Accounting and Information Systems, Osnabrück, Germany
{matthias.graeuler,frank.teuteberg}@uos.de

Abstract. Using a combination of metamodels, ontologies, green performance indicators and metrics, we apply a novel approach in Semantic Business Process Benchmarking to the area of Green Business Process Management (Green BPM). Up to now, process benchmarking has mainly been a manual process; the approach described and empirically evaluated in this paper partially automates the time-consuming and costly process analyses while introducing more flexibility regarding varying terminology, level of abstraction and modeling notation. Also, overviews of literature relevant to the field of Green Semantic BPM and commonly applied metrics in a Green BPM context are given.

Keywords: Semantic Process Management, Ontologies, Sustainability, Green Business Process Management, Benchmarking

1 Introduction

Over the last two decades, many research efforts aimed at developing more energy efficient technologies, alternative energy sources and ecological end-of-life for products. We consider these efforts to be highly relevant and desirable in mitigating our collective impact on the environment, however, the possibilities offered by improving existing operational practices are often ignored. Business Process Management (BPM) offers an integrated, holistic approach to the management of sustainability change [1], which is needed to change these practices efficiently. By looking at the current state of an organization's business processes (BP) as well as possible improvements through comparison with other organizations' BP or reference processes proposed by researchers, considerable improvements can be implemented.

Business Process Benchmarking is commonly used to identify areas in which organizations can improve their efficiency [2], but it is primarily a time-consuming, manual process performed by domain experts. Therefore, as Drew pointed out in 1997, "ways must be found for doing [benchmarking] faster, more effectively and economically, without sacrificing rigour [sic] or integrity of approach [sic]" [3]. Benchmarking primarily suffers from two difficulties that need to be overcome:

- There are many different modeling notations for BPs (e.g. Event-Driven Process Chains (EPC) [4], Business Process Model and Notation (BPMN) [5]), which can-

11th International Conference on Wirtschaftsinformatik,
27th February – 01st March 2013, Leipzig, Germany

not simply be compared because of syntactic incompatibility and varying syntactic richness.
- Even if a process or a process step describes the same chain of activities in two different models, in the same language and perspective, the terminology and semantics of the models may differ, which prohibits a direct comparison [6]. This may stem from the fact that process models can be created from various perspectives at different levels of abstraction [7], because models are created for varying purposes, not all of them requiring indication of each atomic activity, and different organizational entities use different terminology for describing the same domain.

With the help of semantic BPM, researchers try to overcome several issues with BPM (e.g. work done by domain experts, mostly interpreting unstructured information) by providing a common terminological reference point [8-9]. Up to now, there is no generally acknowledged meta-model covering all aspects of process models [7].

In this paper, we follow the suggestion to annotate BPs with corresponding effects of individual activities on the environment (e.g. emissions, waste), which are accumulated along the process flow [6], [10-13]. Using this approach, two or more (sub-)processes can be benchmarked to find ways to improve efficiency. The proposed approach must not be confused with life cycle assessment (LCA) [14], as it is by no means intended to replace but supplement LCA. During the course of our research, we acquired practical examples of BPs in different modeling notations, annotated them with semantic information and demonstrated how process benchmarking can be performed semi-automatically with only a small degree of manual modifications using the software package SEMAT. To achieve the necessary degree of rigor, we started with a systematic literature review [15] to identify the current state of Semantic and Green BPM and also appropriate metrics applicable to Green BPM. To achieve our goals, we laid out a research agenda with the following research questions:

RQ 1: What is the current status quo of Semantic and Green BPM?
RQ 2: What metrics are applied to Green BPM?
RQ 3: Does SEMAT support Green BPM by assisting queries?

This paper is structured as follows: Section 2 details the literature review and lists metrics for Green BPM found during our literature review. In section 3, the proposed benchmarking process is explained in detail and an example is given. Section 4 gives an overview over the design for the experimental validation of the tool, whose results will be shown in section 5. The paper closes with the 6th section, concluding the paper and providing starting points for future research.

2 Literature Review and Related Work

As we wanted to focus on high-quality literature to maximize the reliability of our findings [15-16], we examined the top 20 journals ranked by the AIS as well as the proceedings of the A-ranked conferences ICIS, ECIS, AMCIS, ACIS, PACIS according to the ERA conference rating and additionally WI. We used the following search terms in varying combinations: *Benchmarking, Business, Green, Management, Modeling, Ontology, Process, Semantic, Sustainable* and *Sustainability* to investigate the current works on Green BPM as well as the progress made in the field of Semantic BPM underlying our research. We read title and abstract of each paper and determined whether it was relevant for our research or not (i.e. at least two of the following seven criteria have to be met). We also conducted a forward and backward search to identify more relevant literature. Table 1 summarizes the subset of results we consider to be relevant. In total, we identified 31 papers relevant to our research. The prominence of a certain topic is indicated by the number of asterisks, with a maximum of three asterisks. If a topic is not focused in a publication, it is denoted with a minus. The topics used for comparison to our approach are:

- *Business Processes (BP):* Does the paper focus on business processes?
- *Benchmarking (Ben):* Does the paper propose a process benchmarking approach to analyze and compare individual process models efficiently and effectively?
- *Ecological and/or Social Sustainability (Sust):* Does the paper focus on ecological and/or social sustainability?
- *Semantic Approach (SA):* Does the paper employ a semantic approach to analyze business processes?
- *Research Agenda (RA):* Does the paper propose a research agenda or directions for future research?
- *Evaluation Approach (EA):* Do the authors present an evaluation approach?
- *Focuses on…:* On which problem domain does the study focus?

The results of the literature review are displayed in table 1. The results were loosely clustered by the focused topics; due to the different dimensions examined in the review, a distinct clustering was neither possible nor desirable. As can be seen from our literature review, the combination of research topics (i.e. BPM, Benchmarking, Green IS and the Semantic Web) is largely unexplored, which necessitates this combination.

Three of the papers found in the literature review were written by the same research group [10], [17-18]. As can be seen in table 1, these works are closely related to our research. Similar to our approach, they propose annotating process models with semantic effects regarding e.g. energy efficiency or resource consumption and using the software tool ProcessSEER [19] to obtain the cumulative effects at the end-event. However, in contrast to our method, their approach only supports BPMN and, to our knowledge, there are no efforts to change this. Furthermore, their approach remains largely invalidated. At the time of writing, we were unable to find related papers published after 2010, which indicates that their research in this area has ceased.

While analyzing the literature, we found several Green BPM metrics (cf. table 2). This list of metrics is not meant to be exhaustive, it is rather intended to illustrate that Green BPM should not simply focus on single aspects but needs to be a holistic approach including interdependencies among different metrics. We focused on the social and ecological dimensions of the “triple bottom-line”, omitting the economic perspective, because contrary to the other two dimensions, this perspective has already been explored to a much more mature extent [20].

Table 1. Results of the Literature Review

Ref	BP	Ben	Sust	SA	RA	EA	Focuses on…
[10]	***	***	***	***	-	-	…extending existing BPM technology to enable organizations to be informed about their processes' carbon footprint
[18]	***	***	**	***	-	-	…an algebraic framework for green BPM
[21]	***	*	***	-	***	-	…using BPM techniques to leverage Green IT initiatives
[22]	***	***	-	***	-	-	…an approach for pattern-based process model analysis
[6]	***	***	-	***	*	*	…semantic interoperability of BP models
[9]	***	**	-	***	*	*	…the semantic interoperability of BPs
[23]	***	**	-	**	-	-	…semi-automatic checking of semantically analyzable BP models
[24]	***	*	-	***	-	-	…compliance checking in financial institutions
[25]	***	*	-	-	**	-	…a knowledge-based system aiding process redesign
[17]	***	-	***	*	***	-	…issues and a research agenda for green BPM
[1]	***	-	***	-	***	-	…call for action for investigating BPM to create more sustainable organizations
[26]	***	-	***	-	**	-	…research agenda for green BPM
[27]	**	-	***	-	***	-	…various areas of Green IT, e.g. Green BPM
[20]	**	-	**	-	**	-	…economic sustainability in BPM
[28]	***	-	***	-	-	**	…proposal of an approach for measuring CO_2 emissions during BP execution
[29]	***	-	***	-	-	-	… the evaluation and comparison of process designs
[12]	***	-	-	***	*	-	…combination of semantic web services and BPM to create the concept of semantic BPM
[30]	***	-	-	**	**	-	…the necessity of methodological elaborations of BPM
[31]	***	-	-	*	**	-	…inter-organizational BP design
[32]	***	-	-	***	-	*	…partially automatic planning and modeling BPs
[7]	***	-	-	***	-	-	…usage of an ontology framework to reduce complexity of e.g. process modeling
[11]	***	-	-	***	-	-	…ontologically representing the business and IS perspective on BPs, and on translating between these two perspectives
[13]	***	-	-	***	-	-	…semantic annotation of EPCs to specify the semantics of individual process model elements
[33]	***	-	-	**	-	-	…domain-specific semantic BP modeling language for banks
[34]	***	-	-	***	-	-	…process verification using process logic
[35]	***	-	-	***	-	-	…analyzing process mining
[36]	***	-	-	***	-	-	…a domain ontology based approach to support BP design
[37]	***	-	-	***	-	-	…an integrated model for inter-organizational BP integration
[38]	***	-	-	**	-	-	…finding similarities and contradictions in BP models
[39]	***	-	-	-	*	*	…BPM success
[40]	-	-	***	-	-	*	…sustainable performance measurement at airlines
[41]	*	*	**	-	-	**	…description of a tool for managing material flow networks

Hoesch-Klohe and Ghose [18] and previously Teuteberg et al. [6], [42] propose to annotate each activity in a business process model and to accumulate the values along the process chain, taking the different possible paths of process execution into account. The metrics listed in table 2 can be attached to any construct of any modeling language describing an action (e.g. functions in EPCs or activities in UML-AD and BPMN). Some of them can easily be quantified (e.g. Energy Consumption, Waste Generation), whereas others can only be measured in a qualitative scale (e.g. Environmental Performance) [18].

Table 2. Metrics for Green BPM

Dim	Metric / Explanation / Example	Reference(s)
Env	Air Quality indicated by e.g. Air Quality Index	[18]
Env	Congestion leads to unnecessary consumption of resources	[40]
Env	Emissions of greenhouse gases, ozone depleting substances or other emissions in e.g. CO_2-equivalents	[1, 10, 17, 18, 26–28, 40, 43]
Env	Energy Efficiency/Consumption in e.g. kWh/unit	[1, 17, 21, 26–29, 43]
Env	Environmental Performance: Qualitative measure, representing a variety of measures; performance could range from A to D	[18]
Env	Fuel Efficiency/Consumption in e.g. km/100 liter fuel	[21, 28, 40]
Env	Odour emission in e.g. olf or decipol	[44]
Env	Paper Consumption in e.g. sheets/employee	[1, 27, 28]
Env	Radiation in e.g. sievert	[44, 45]
Env	Waste Generation in e.g. kg/unit	[18, 28, 40, 43]
Env	Water Consumption in e.g. liter/unit	[18, 26, 43]
Env	Water Discharge in liters	[43]
Env,	Noise Generation measureable in e.g. decibel	[40, 44]
Soc	Probability of accidents/casualties	[44]
Soc	Training and Development required for new employees	[40]
Soc	Workforce size indicates the number of employees needed	[43]

3 Green Process Benchmarking Approach – An Example

Figure 1 illustrates the actions performed during the benchmarking process using an example from the printing industry and applying some of the metrics introduced in table 2. This example was selected because it describes the production of photobooks, a rapidly growing market, and provides information about CO_2 emissions. Furthermore, the used production process could be inferred from the descriptions.

We used a process described by Hausmann [46] to create two very simple process models: one using EPC and one using BPMN. We annotated each step with the effects on the environment (i.e. energy consumption, emissions). The process models are transferred into a machine-readable form, such as a Comma Separated Values (CSV) file, so they can easily be imported by the software created for this benchmarking approach: SEMAT. Although essentially accomplishing the same outcome, both processes exhibit varying degrees of abstraction and different terminology. They are also modeled in an entirely different notation. To overcome the problems associated with this kind of benchmarking endeavor, we use a combination of methods.

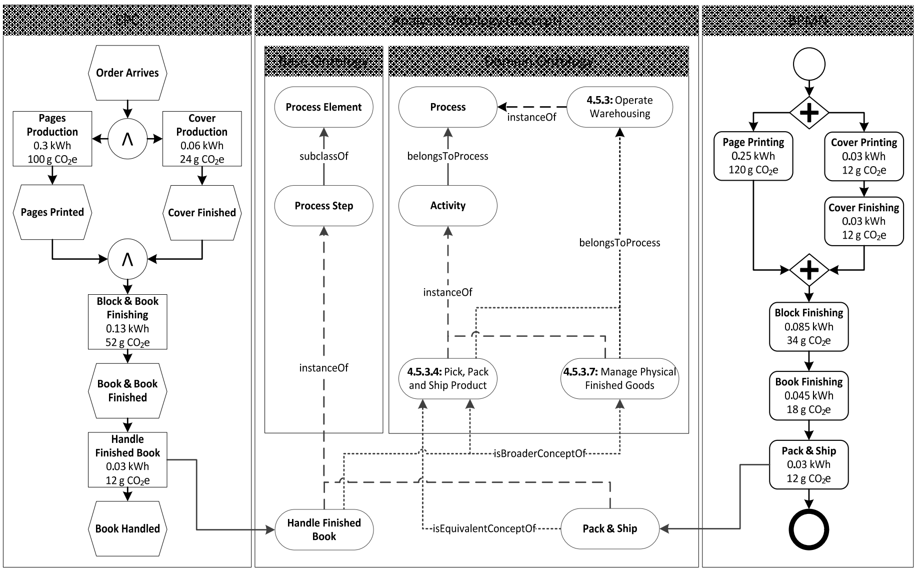

Fig. 1. Example for the Benchmarking Process

First, a domain ontology describing the benchmarking domain has either to be created or, in case this benchmarking approach has been used in the specific domain before, selected from a repository (Step 1). For the example in figure 1, we used the Process Classification Framework [47] that was transformed into the desired artifact. Then, a base ontology has to be created or reused (Step 2). The base ontology must be able to represent all relevant constructs of the used modeling languages. Furthermore, it can be reused if it provides support for the modeling languages used in the process models to be benchmarked. Figure 1 shows that the element types "function" (EPC) and "task" (BPMN) are syntactically equivalent and, therefore, mapped to the base ontology's concept "Process Step". Even though domain and base ontologies are merged into one analysis ontology later, at this point, they should be kept separate because the domain ontology can be reused regardless of the used modeling languages and the base ontology can be reused in any domain. The third step introduces the process models into the benchmarking process. A domain expert maps the elements of the process models to instances of the domain ontology (Step 3). In contrast to Höfferer (2007) [9], who suggested only one mapping type, we use three mapping types: one each if the element is equivalent to the domain concept, if it is represented in a broader or in a narrower sense. By resorting to a domain experts' knowledge for mapping, we are able to overcome the problems raised by inconsistent terminology and different degrees of abstraction. Subsequently, the sustainability performance information needs to be captured and inserted into the ontology (Step 4). Using the specifically designed tool SEMAT, the process models are imported from text-files and transformed into an instance ontology of the base ontology. This instance ontology is then mapped to the base ontology using the mappings made in step three and, subsequently, merged into a single analysis ontology (Step 5). The analysis ontology can easily be queried using query languages such as SPARQL [48] and by using predefined metrics supplied by SEMAT (Step 6).

Figure 2 demonstrates how the results of the analysis are displayed. The selected model (the process modeled in EPC notation from figure 1) and metrics are highlighted in the panel at the top. The table at the bottom shows the actual process performance and accumulates them along the process flow, enabling analysts to quickly assess process performance. As can be seen, any manufacturing process is generalized into the concept of "Produce Product", which was taken from the PCF taxonomy.

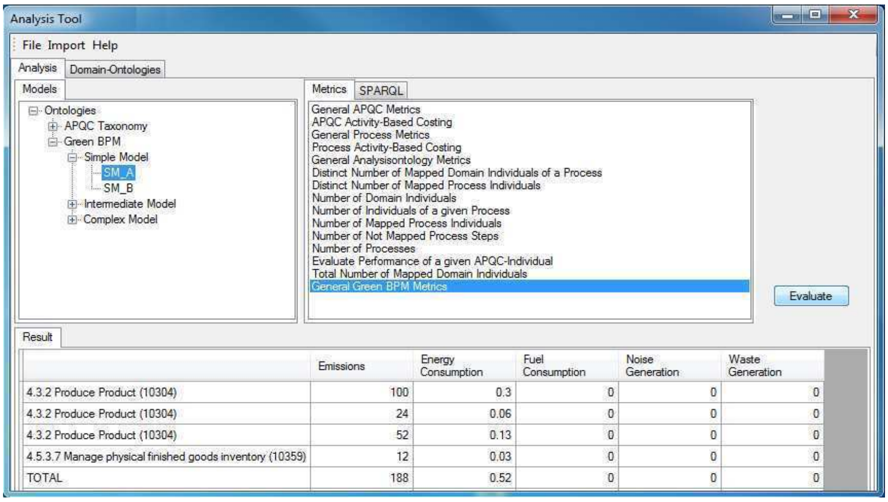

Fig. 2. Screenshot of the analysis results presented by SEMAT

4 Evaluation Concept

The previously described benchmarking process can be split into two parts: One that has to be performed once per process (steps 1-3) and one that has to be carried out every time the process's performance information changes (steps 4-6). It should be noted that steps 1-3 are more time-consuming than steps 4-6, especially when an appropriate domain or base ontology has to be created. The scope of this research is limited to the evaluation of the advantageousness of our proposed methodology in respect of querying the data. Therefore, we decided to only evaluate step six, as this is the most frequent step and does not require extensive domain or process knowledge. It must be taken into account that in a manual comparison steps 1 to 5 are unnecessary and, have thus been ignored. Therefore, both groups operate within the same context and boundaries. This step was examined using an experimental setup.

Prior to the experiment (can be requested from the corresponding author), all test persons were screened whether they possessed sufficient knowledge about the modeling notations used later in the experiment; also, specific information, such as *Computer Self-Efficacy (Comp. SE)* and a self-assessment of *BPM Knowledge,* were collected using 7-point Likert-scales. Then, the test persons were randomly divided into two groups that received different treatments. Group A (Manual) was given three more complex process models, each in a different notation (i.e. EPC and BPMN) and with annotated emissions, resource and energy consumption. Group B (SEMAT) was given a short introduction to our software used in the proposed benchmarking process and then received the same treatment as group A, except that they used analysis results from the software to solve the same tasks on identical process models.

The constructs *Comp. SE* and *BPM Knowledge* are based on a self-assessment by the test persons. According to the self-efficacy theory, the expectations of personal efficacy determine whether a certain behavior will be initiated, how much effort will be expended and how long it will be sustained in the face of obstacles [49]. Higgins and Compeau found that *Comp. SE* has a significant impact on individuals' expectations of the outcomes of using computers, their emotional reactions to computers as well as their actual computer use [50]. Another theory that was applied in our research model is the Theory of Task-Technology-Fit (TTF), which asserts that IS are means for users to complete tasks (i.e. summarizing process performance). The higher the TTF, the better the performance of the users (i.e. fast completion, high accuracy) [51]. The variables *Perceived Usefulness* (Perc. Usef.) and *Perceived Ease Of Use (Perc. EoU)* were taken from the Technology Acceptance Model and it's successors [52]. These variables are assumed to be fitting, as they are commonly used to empirically analyze an IT artifact. It was considered whether the Expectation Confirmation Theory [53-54] applies in the context of the experiment. However, we decided not to include expectations and their (dis-)confirmation because the majority of the test persons have experience with BPM tools, but not to an extent in which an evaluation of these variables can be considered useful and worthwhile. Using this approach, we were able to test the following hypotheses:

H1 Test persons who think of themselves as more computer proficient will…
- *H1a* …be more accurate when using SEMAT
- *H1b* …finish tasks faster when using SEMAT

H2 Test persons who indicated that they have extensive knowledge of BPM will…
- *H2a* …be more accurate when not using SEMAT
- *H2b* …be more accurate when using SEMAT
- *H2c* …finish tasks faster when not using SEMAT
- *H2d* …finish tasks faster when using SEMAT

H3 Test persons using SEMAT for the analysis will…
- *H3a* …be more accurate than those who did not use SEMAT
- *H3b* …finish tasks faster than those who did not use SEMAT

H4 The test persons' rating for Perc. EoU will be higher if they…
- *H4a* …are more accurate when using SEMAT
- *H4b* …finish tasks faster when using SEMAT

H5 The test person's rating for Perc. Usef. will be higher if they…
- *H5a* …are more accurate when using SEMAT
- *H5b* …finish tasks faster when using SEMAT

H6 The Perc. EoU will have a positive effect on Perc. Usef. when using SEMAT

H7 The test person's satisfaction will be positively related with…

H7a …Perc. Usef. when using our SEMAT
H7b …Perc. EoU when using our SEMAT

These hypotheses and the respective constructs are displayed in figure 3. All constructs in squares are measured by 7-point Likert-scales using at least four items to measure each construct. The constructs in oval shapes are measured during the experiment. Whereas most constructs' values were calculated using the arithmetic mean of the items measuring them (equally weighed), the variable *Time* was measured by the time needed to fulfill the assigned tasks. Respectively, *Accuracy* was measured by the number of correctly solved tasks.

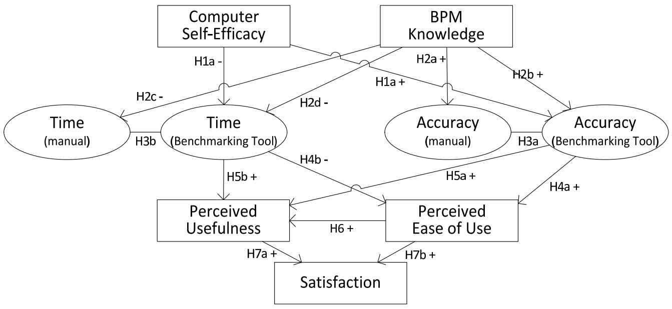

Fig. 3. Research Model

5 Results

In total, we collected 24 datasets from the participants, 18 of them (under-)graduate students and 6 Ph.D. candidates. 11 participants were assigned to the group performing a manual analysis, 13 to the group using SEMAT. Table 3 gives an overview of the descriptive statistics as well as reliability coefficients for each construct measured by Likert-scales. To display sufficient convergent validity, three criteria should be met: A minimum of (1) 0.7 for Cronbach's Alpha [55]; (2) 0.7 for composite reliability [56]; and (3) 0.5 for Average Variance Extracted (AVE) [57]. Our data demonstrate compliance with these criteria. Additionally, we conducted a factor analysis for each construct. All item loadings were above the .5 threshold with eigenvalues above 1.0 indicating good convergent validity. There were no correlations above .9, indicating that multicollinearity is not an issue [58].

To examine hypotheses H1 and H2, we used the correlations of several factors with the *Time* needed to complete a task and the *Accuracy* of the results; these are shown in table 4. In H1 we hypothesized that persons with a higher rating in *Comp. SE* will (H1a) be more accurate and (H1b) finish tasks faster when using SEMAT. We were not in a position to demonstrate these relationships, therefore, both hypotheses were rejected. However, this may indicate that *Comp. SE* only has a minor influence on benchmarking performance using SEMAT, which is desirable and can be assisted with further improvements regarding usability.

Table 3. Descriptive Statistics and Reliability Coefficients

	Group	Items	Mean	SD	Alpha	Comp. Rel.	AVE
Comp. SE	Manual	4	6.205	.706	.827	.895	.682
	SEMAT		5.846	.681			
BPM Knowledge	Manual	4	4.522	1.186	.872	.912	.723
	SEMAT		4.885	1.368			
Perc. Usef.	Manual	5	4.900	.939	.832	.885	.609
	SEMAT		5.508	.755			
Perc. EoU	Manual	7	5.571	.774	.872	.909	.592
	SEMAT		5.956	.770			
Satisfaction	Manual	4	4.975	1.204	.895	.928	.763
	SEMAT		5.558	.693			

Table 4. Correlations of various constructs with *time* and *accuracy*

Construct		**Comp.SE**	**BPM Knowledge**		**Perc. EoU**	**Perc. Usef.**
Group		SEMAT	SEMAT	Manual	SEMAT	SEMAT
Time Total	Corr.	.068	-.023	-.159	-.159	-.230
	Sign.	.787	.928	.528	.528	.365
Accuracy Total	Corr.	-.292	-.309	.654	.358	.092
	Sign.	.224	.197	.008	.130	.703

Equally, we were unable to support H2, which indicated that people with more *BPM Knowledge* will finish the tasks faster and more accurately. Only H2a could be supported by a strong correlation ($p < .01$). Accordingly, these results suggest that *BPM Knowledge* only plays a minor role when using SEMAT, and therefore a lower degree of knowledge is necessary compared to performing manual benchmarking, where a strong correlation is present. While this is desirable as well, it implies that our tool – or BPM software in general – should assist the user wherever possible and mitigate the probability of human error.

Further, in H3 we hypothesized that persons using SEMAT will be more accurate and faster than those who did not. The descriptive statistics for *Time* and *Accuracy* shown in table 5 indicate that in fact, SEMAT users are on average 14.3% faster while scoring 6.7% higher. These results imply that SEMAT – and perhaps other BPM tools – increase efficiency and the quality of benchmarking results. We expect a greater difference between mean values with increasing task complexity. However, a T-test revealed that these differences in mean-values are not statistically significant and therefore H3 was rejected.

In H4 and H5 we hypothesized that, when using SEMAT, Time and Accuracy are positively related with the Perc. EoU and Perc. Usef. To test these hypotheses, we examined the correlations of said constructs; the results can be reviewed in table 4. Although the correlations imply that the hypotheses are in fact true, these correlations are not statistically significant. Subsequently, these hypotheses were rejected as well.

Table 5. Descriptive statistics for *Time* and *Accuracy*

	Group	Time Task 1	Time Task 2	Time Task 3	Time Total	Accuracy Task 1	Accuracy Task 2	Accuracy Task 3	Accuracy Total
Mean	Manual	396.400	209.930	675.408	1255.232	5.182	8.455	18.455	32.091
	SEMAT	309.454	177.132	606.720	1076.300	5.846	8.923	19.615	34.385
SD	Manual	119.604	42.641	121.712	221.511	1.834	1.809	4.083	6.156
	SEMAT	130.623	61.341	137.136	264.322	0.555	0.277	3.841	4.407

Table 6. Correlations among the constructs *Perc. Usef.*, *Perc. EoU* and *Satisfaction* for persons that used SEMAT for the analysis

		Perc. EoU	**Satisfaction**
Perc. Usef. SEMAT	Corr.	.622	.823
	Sign.	.004	.000
Perc. EoU SEMAT	Corr.	1.000	.677
	Sign.	-	.002

Further, we hypothesized that people's *Perc. EoU* will have a positive effect on *Perc. Usef.* (H6). This hypothesis was tested by means of an investigation into the constructs correlations. As can be seen in table 6, there is a high positive correlation, with a statistical significance at the .005-level. Therefore, this hypothesis can be accepted. The last hypothesis suggests that a person's *Satisfaction* will be positively related with the *Perc. Usef.* (H7a) and *Perc. EoU* (H7b), provided they used SEMAT for the analysis. Again, table 6 shows the correlations and their significance used for testing these hypotheses. Hypothesis H7a is supported by the strong correlation between the constructs *Perc. Usef.* and *Satisfaction* (.823; $p < .001$) and hypotheses H7b by the correlation between *Perc. EoU* and *Satisfaction* (.677; $p < .005$).

We used regression analysis to calculate the R^2 values. As to be expected from the previously reported results, most R^2 values were unsatisfactory and/or statistically insignificant. However, we found an R^2 value of .598 ($p < .01$) with *Perc. EoU* as the independent and *Perc. Usef.* as the dependent variable and an R^2 value of .839 ($p < .001$) with *Perc. EoU* and *Perc. Usef.* as the independent variables and *Satisfaction* as the dependent variable.

Additionally, we gathered the participants' opinions on SEMAT. The participants found that the software was easy to understand (38.5%), very helpful to determining process information (30.8%) and helped them to concentrate on what is important to accomplish the tasks (23.1%). On the other hand, they indicated that they would appreciate more extensive support with aggregating process information (30.8%) and automated checks whether the entered information is valid (23.1%).

6 Conclusion and Future Research

In this paper, we describe an approach for benchmarking sustainability-related metrics using specially engineered software. We show our approach to evaluate SEMAT in an experiment with 24 participants and share the gathered data. Regarding the population and size of the sample, even though all participants had training in BPM, it would have been beneficial to target BPM professionals in a natural use-setting with our experiment. While a sample size of 24 is sufficient to make statistically relevant statements, a bigger sample always improves statistical power [59]. Another way to improve the validity of the results would be to add more levels of task complexity. Also, in future research it would be beneficial to draw comparisons with other established benchmarking approaches.

SEMAT is helpful to improve sustainability efforts, but still possesses some issues to address, one of them being the accurate handling of measurements. Even if metrics are applied consistently, some organizations may capture, e.g. fuel efficiency in miles per gallon whereas others may capture it in liters per 100 kilometers. To ensure correct benchmarking results, these measurements should automatically be translated without introducing the risk of human error. This can be achieved by another ontology that is not only able to perform the aforementioned calculations but also to, e.g., translate any greenhouse gas into CO_2 equivalents.

The creation of the domain and base ontology itself is rather time-consuming. Therefore, ways need to be found to accelerate these processes and make them less prone to human error. This can be achieved by maintaining a centralized base ontology that is appropriate for a multitude of notations and languages. This paper only examines the benefits of querying the analysis ontology, which is only one out of six steps in the methodology. Once the mentioned improvements are implemented, the entire benchmarking process should be evaluated. We show that, on average, benchmarking efforts using our tool are less time-consuming than doing the work manually, although the differences are not significant enough and should therefore be repeated, addressing the issues mentioned above.

Also, our research has several implications for the creation of future (Green) BPM tools. Currently, there is no feature implemented that can tell whether, e.g., a process that produces fewer emissions (metric A) is preferable to a process that produces less noise (metric B) (i.e. allowing preference rankings between different metrics A and B). Such a preference ranking varies in each organization and location and must currently be decided manually. In future, Green BPM software enabling this decision support and containing appropriate formulae and thresholds for preference rankings could be implemented. We also show that *Perc. Usef.* and *Perc. EoU* have a significant impact on *Satisfaction*. Therefore, developers of (green) BPM software should focus on creating easy to use, well documented software that does not distract the users with unneeded information. Furthermore, this software should be easily accessible for less experienced analysts. The variety of metrics that can be applied to Green BPM shows that software tools must exhibit a certain degree of flexibility in regard to what information can be annotated to process models.

One of the shortcomings of any BPM measure, the rather time-consuming collection of process information, could be addressed by integrating various data sources. Some of which may be used to generate and/or update process performance information, further increasing efficiency of the benchmarking approach.

Acknowledgements

The authors would like to thank the participants in the experiment as well as the other project members, specifically Ms. Marita Imhorst and Mr. Martin Kluth, who provided valuable insights, help and substantive feedback during the research process.

This work is part of the project IT-for-Green (Next Generation CEMIS for Environmental, Energy and Resource Management). The IT-for-Green project is funded by the European regional development fund (grant number W/A III 80119242).

References

1. Seidel, S., Vom Brocke, J., Recker, J.: Call for Action: Investigating the Role of Business Process Management in Green IS. Sprouts 11 (4), 1–7 (2011)
2. Watson, G.H.: Strategic Benchmarking - How To Rate Your Company's Performance Against The World's Best. Wiley and Sons, New York (1993)
3. Drew, S.A.W.: From Knowledge to Action: The Impact of Benchmarking on Organizational Performance. In: Long Range Planning 30 (3), 427–441 (1997)
4. Scheer, A.-W.: ARIS - Business Process Modeling. Springer, Berlin (2000)
5. Object Management Group: Business Process Model and Notation (BPMN) - Version 2.0
6. Teuteberg, F., Kluth, M., Smolnik, S., Ahlemann, F.: Semantic Benchmarking of Process Models - An Ontology-Based Approach. In: ICIS 2009 Proceedings, paper 89. AIS (2009)
7. Markovic, I., Hasibether, F., Jain, S., Stojanovic, N.: Process-Oriented Semantic Business Modeling. In: 9th Wirtschaftsinformatik 2009, pp. 683-693 (2009)
8. Thomas, O., Fellmann, M.: Semantic Business Process Management: Ontology-Based Process Modeling Using Event-Driven Process Chains. International Journal of Interoperability in Business Information Systems 2, 29–44 (2007)
9. Höfferer, P.: Achieving Business Process Model Interoperability Using Metamodels and Ontologies. In: European Conference on Information Systems St. Gallen (2007)
10. Hoesch-Klohe, K., Ghose, A., Lê, L.-S.: Towards Green Business Process Management. In: IEEE International Conference on Services Computing, pp. 386–393 (2010)
11. Hepp, M., Roman, D.: An Ontology Framework for Semantic Business Process Management. In: Proceedings of Wirtschaftsinformatik 2007 Karlsruhe (2007)
12. Hepp, M., Leymann, F., Bussler, C., Domingue, J., Wahler, A., Fensel, D.: Semantic Business Process Management: A Vision Towards Using Semantic Web Services for Business Process Management. In: IEEE International Conference on e-Business Engineering, pp. 535–540 (2005)
13. Thomas, O., Fellmann, M.: Semantic Process Modeling - Design and Implementation of an Ontology-Based Representation of Business Processes. Business Information Systems Engineering 2, 15–28 (2009)
14. International Organization for Standardization: ISO 14040: Environmental Management - Life Cycle Assessment - Principles and Framework, Geneva (2006)
15. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26, 13–23 (2002)
16. Levy, Y., Ellis, T.J.: A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. Informing Science: International Journal of an Emerging Transdiscipline 9, 181–212 (2006)

17. Ghose, A., Hoesch-Klohe, K., Hinsche, L., Le, L.-S.: Green Business Process Management: A Research Agenda. Australasian Journal of Information Systems 16, 103–117 (2009)
18. Hoesch-Klohe, K., Ghose, A.: Carbon-Aware Business Process Design in Abnoba. In: LNCS, Vol. 6470, pp. 551–556. Springer (2010)
19. Hinge, K., Ghose, A., Koliadis, G.: Process Seer: A Tool for Semantic Effect Annotation of Business Process Models. In: IEEE International EDOC Conference Auckland (2009)
20. Cleven, A., Winter, R., Wortmann, F.: Process Performance Management as a Basic Concept for Sustainable Business Process Management – Empirical Investigation and Research Agenda. In: LNBIP, Vol. 66, pp. 479–488. Springer (2011)
21. Houy, C., Reiter, M., Fettke, P., Loos, P., Geb, D.: Towards Green BPM – Sustainability and Resource Efficiency through Business Process Management. In: LNBIP, Vol. 66, pp. 501–510. Springer (2011)
22. Becker, J., Weiß, B., Winkelmann, A.: Automatic Identification of Structural Process Weaknesses – Experiences with Semantic Business Process Modeling in the Financial Sector. In: Wirtschaftsinformatik, Zürich (2011)
23. Becker, J., Räckers, M., Winkelmann, A.: Pattern-Based Semi-Automatic Analysis of Weaknesses in Semantic Business Process Models in the Banking Sector. In: European Conference on Information Systems, Pretoria (2010)
24. Becker, J., Bergener, P., Delfmann, P., Eggert, M., Weiß, B.: Supporting Business Process Compliance in Financial Institutions – A Model-Driven Approach. In: Wirtschaftsinformatik (2011)
25. Nissen, M.E.: Redesigning Reengineering Through Measurement-Driven Inference. MIS Quarterly 22, 509–535 (1998)
26. Nowak, A., Leymann, F., Mietzner, R.: Towards Green Business Process Reengineering. In: LNCS, Vol. 6568, pp.187–192. Springer (2011)
27. Loos, P., Nebel, W., Marx Gómez, J., Hasan, H., Watson, R.T., vom Brocke, J., Seidel, S., Recker, J.: Green IT: A Matter of Business and Information Systems Engineering?. Business & Information Systems Engineering 3 (4), 245–252 (2011)
28. Recker, J., Rosemann, M., Gohar, E.R.: Measuring the Carbon Footprint of Business Processes. In: LNBIP, Vol. 66, pp. 511–520. Springer (2011)
29. Kundisch, D., Herrmann, P., Meier, C.: Sustainable Process Management - Status Quo and Perspectives. In: LNBIP, Vol. 46, pp. 94–106. Springer (2010)
30. Recker, J.: Towards an Understanding of Process Model Quality. Methodological Considerations. In: European Conference on Information Systems, Göteborg (2006)
31. Legner, C., Wende, K.: The Challenges of Inter-Organizational Business Process Design - A Research Agenda. In: European Conference on Information Systems, St. Gallen (2007)
32. Heinrich, B., Bolsinger, M., Bewernik, M.: Automated Planning of Process Models : The Constructions of Exclusive Choices. In: International Conference on Information Systems, Phoenix (2009)
33. Becker, J., Weiß, B., Winkelmann, A.: Transferring a Domain-Specific Semantic Process Modeling Language: Findings from Action Research in the Banking Sector. In: European Conference on Information Systems, Pretoria (2010)
34. Bi, H., Zhao, L.: Process Logic for Verifying the Correctness of Business Process Models. In: International Conference on Information Systems, Turku (2004)
35. De Medeiros, A., van der Aalst, W.M.P., Karla, A., Pedrinaci, C.: Semantic Process Mining Tools: Core Building Blocks. In: European Conference on IS (2008)
36. Hua, Z., Zhao, J.L., Storey, V.C.: Exploring a Domain Ontology Based Approach to Business Process Design. In: International Conference on IS, St. Louis (2010)

37. Otto, B., Wäsch, J.: A Model for Inter-Organizational Business Process Integration. In: Wirtschaftsinformatik, Dresden (2003)
38. Simon, C., Mendling, J.: Integration of Conceptual Process Models by the Example of Event-driven Process Chains. In: Wirtschaftsinformatik, Karlsruhe (2007)
39. Hailemariam, G., vom Brocke, J.: What Is Sustainability in Business Process Management? A Theoretical Framework and Its Application in the Public Sector of Ethiopia. In: LNBIP, Vol. 66, pp. 489–500. Springer (2011)
40. Alemayehu, W., Vom Brocke, J.: Sustainability Performance Measurement – The Case of Ethiopian Airlines. In: LNBIP. 66, pp. 467–478. Springer (2011)
41. Joschko, P., Page, B., Wohlgemuth, V.: Combination of job oriented simulation with ecological material flow analysis as integrated analysis tool for business production processes. In: Proceedings of the 2009 Winter Simulation Conference, pp 1456–1465. IEEE (2009)
42. Teuteberg, F., Kluth, M., Ahlemann, F., Smolnik, S.: Semantic Process Benchmarking to Improve Process Performance. Benchmarking: An International Journal (2013)
43. Global Reporting Initiative: Sustainability Reporting Guidelines 3.1. Global Reporting Initiative (2011)
44. De Haes, U.: Towards a Methodology for Life Cycle Impact Assessment. SETAC-Europe, Brussels (1996)
45. Jolliet, O., Margni, M., Charles, R., Humbert, S., Payet, J., Rebitzer, G., Rosenbaum, R.: IMPACT 2002+: A New Life Cycle Impact Assessment Methodology. The International Journal of Life Cycle Assessment 8, 324–330 (2003)
46. Hausmann, M.: Sustainability of thc CEWE PHOTOBOOK. In: International Conference on Digital Printing Technologies and Digital Fabrication, pp. 736–740 (2011)
47. APQC: APQC Process Classification Framework (PCF)
48. Prud'hommeaux, E., Seaborne, A.: SPARQL Query Language for RDF - W3C Recommendation
49. Bandura, A.: Self-Efficacy: Toward a Unifying Theory of Behavioral Change. Advances in Behaviour Research and Therapy 1, 139–161 (1978)
50. Higgins, C.A., Compeau, D.R.: Computer Self-Efficacy: Development of a Measure and Initial Test. MIS Quarterly 19, 189–211 (1995)
51. Goodhue, D.L.: Understanding User Evaluations of Information Systems. Management Science 41, 1827–1844 (1995)
52. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 13, 319–340 (1989)
53. Oliver, R.: Effect of expectation and disconfirmation on postexposure product evaluations: An alternative interpretation. Journal of Applied Psychology 62, 480–486 (1977)
54. Oliver, R.: A cognitive model of the antecedents and consequences of satisfaction decisions. Journal of Marketing Research 1, 460–469 (1980)
55. Kline, P.: The Handbook of Psychological Testing. Routledge, London (1998)
56. Hair, J.F.J., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L.: Multivariate Data Analysis. Pearson Prentice Hall, Upper Saddle River (2006)
57. Fornell, C., Larcker, D.F.: Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 18, 39–50 (1981)
58. Burns, R.B., Burns, R.A.: Business Research Methods and Statistics using SPSS. SAGE Publications, London (2008)
59. Cohen, J.: Statistical Power Analysis for the Behavioural Sciences. Lawrence Erlbaum, New York (1988)