A consolidated process for software process simulation: State of the Art and Industry Experience

Nauman Bin Ali, Kai Petersen

38th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)

2012 Cesme

A consolidated process for software process simulation: State of the Art and Industry Experience

Nauman Bin Ali, Kai Petersen
Blekinge Institute of Technology, Karlskrona, Sweden
nauman.ali, kai.petersen @bth.se

***Abstract*—Software process simulation is a complex task and in order to conduct a simulation project practitioners require support through a process for software process simulation modelling (SPSM), including what steps to take and what guidelines to follow in each step. This paper provides a literature based consolidated process for SPSM where the steps and guidelines for each step are identified through a review of literature and are complemented by experience from using these recommendations in an action research at a large Telecommunication vendor. We found five simulation processes in SPSM literature, resulting in a seven-step process. The consolidated process was successfully applied at the studied company, with the experiences of doing so being reported.**

***Keywords*-Software process simulation; empirical;**

I. Introduction

Software processes are complex, dynamic and non-deterministic in nature and it is therefore appropriate to study these using simulation. SPSM attempts to imitate the real-world software processes and enables investigations using the model for various purposes ranging from training, education, planning, strategic management etc. [1] [2]. The potential of SPSM is well established and it is time to make an engineering discipline from this art form [3]. To achieve this various systematic processes to develop simulation models have been proposed in literature.

Today, when a practitioner takes on the task of SPSM the immediate challenge is the choice of process for SPSM study because of the number of available process prescriptions in literature that claim successful results. These processes target certain simulation approaches e.g. Rus et al. [3] for discrete event simulation (DES) and Pfahl and Lebsanft [4] for System Dynamics (SD). Yet other proposed processes are different at the level of formalism in the process description, use of terminology or targeted experience level of modeller and size of the organization undertaking SPSM.

Some of these proposals suggest that their process is extensible to simulation approaches other than the one they were originally proposed for. It is not obvious to an inexperienced modeller as to what part of the process will change if they use the same process for a different approach (e.g. using the process for SD to develop a DES model) or context (e.g. using the same process in a small enterprise). Similarly, the existence of these variations is confusing for a practitioner with no simulation background because if the same process can be used why do we have a number of proposals. Therefore, it is important to analyse and use these processes to identify deficiencies and improvement opportunities. An aggregation of best practices and a consolidated simulation process will support wide spread adoption of process simulation in software industry.

Ahmed et al. [5] found in a survey that software process simulation modellers consider simulation-modelling process a major challenge in SPSM. They also present a possible explanation for the lack of SPSM process discussion in literature saying that simulation modellers do not report on the modelling process as either they already have a satisfactory process or are not interested in it. Today, however, considering the amount of SPSM literature, the need is to combine the lessons learnt and make use of the collective experience gained over more than two decades of SPSM research and practice.

The contribution of this study is two fold. First a literature based consolidated simulation process is created along with supplementary recommendations for each step. Secondly, this process, which utilizes experience of various simulators (thus less influenced by individual preferences), is applied in an industrial simulation project. We also compare the results to the research from outside the SPSM venues.

Section II summarizes the existing processes for SPSM and their industrial evaluation. Section III describes the research methods used in this study. Section IV presents the aggregated guidelines and consolidated simulation process. Section V reflects on the findings of the literature review and action research. Section VI concludes the paper.

II. Related work

Pfahl and Günther [6] proposed a methodology (called IMMoS) that complements SD with existing modelling methods and quantitative measurement. One of the four components of this model is a process description for SPSM, which in turn has four phases and seventeen sub-activities [4] [7]. Angkasaputra and Pfahl [8], analysed this process against agile practices and made improvements to make it more agile. Zhai et al. [9] however claim that it is difficult to combine information learned from empirical studies into the simulation results generated by IMMoS. Furthermore, they claim that it is difficult to build a SD model in practice using this approach, however, the paper does not state a rationale for that conclusion.

Rus et al. [3] create two clusters of the simulation process activities: engineering and management activities. They presented a process for DES but also claimed that it can be used for other types of simulations. The process indicates its documentation-focus with emphasis on specification and approval of changes. Müller and Pfahl [10] summarize five major steps for SPSM study and refer to [6] and [3] for detailed guidelines for SD and DES respectively.

Ahmed et al. [11] proposed a process based on interviews of simulation modellers. This process was aimed at novice modellers and it was evaluated in a controlled experiment with students as subjects.

Silva Filho et al. [12] present a simulation approach that is targeted for Small and Medium Enterprises. The process consists of these steps: The user selects the variables related to the subject to be addressed; Data collection; Preprocessing data; Causal model building; Simulation model assembly and evaluation.

Kellner et al. [2] did not present a process for SPSM but identified major steps in a simulation study by answering the questions: Why simulate? (identifying the purpose and uses of SPSM); What to simulate? (the parameters and scope of simulation); How to simulate? (techniques and approaches for SPSM).

Madachy [1] proposes the use of Win Win Spiral model for model development and maps five major phases of a simulation study to this process. The book also provides comprehensive guidance especially focusing on use of SD in SPSM.

There is a large number of industrial SPSM studies however to the best of our knowledge, among the explicitly documented simulation processes only IMMoS [6] and the process suggested in [3] have been used in industry (all other industrial studies do not explicitly report the process they followed). In both cases the original authors of the process description used it. The reason to mention this is that when one describes the process one followed a lot of the information may be implied for the authors and thus unconsciously missed from the description. Similarly, one may have consciously left certain details or decisions involved out of the process description considering them too obvious.

The high degree of similarity in these guidelines suggests that these approaches may be combined to develop a common process. Such a process will facilitate understanding, provide guidance and promote a systematic manner of approaching SPSM. It will also open new avenues for reuse. As having a shared method with identified activities and artefacts for SPSM will be the basis to identify how simulation models can be developed for and with reuse [13] and will reduce effort and cost of SPSM.

III. Research Methodology

The aim of this study is to identify process descriptions for SPSM and the guidelines and best practices for their operational use. For this purpose following research questions were asked:

RQ1: *Which process descriptions and guidelines are reported for SPSM in industry?*

RQ2: *Is the literature based consolidated process for SPSM useful in an industrial study?*

A. Research methods

1) Literature review: We conducted a thorough literature search where results from Zhang et al. [14] systematic review were used as the foundation. Unfortunately, the classification of articles relevant to methodology was not available for the second phase of their study. Therefore we resorted to reading the titles and abstracts to select the articles for data extraction from the primary studies of the second phase of their systematic review. We supplemented this with snowball sampling where all the references in these articles that were found relevant to the topic were also considered potential primary studies.

2) Action Research: The analysis of findings of the literature review resulted in a consolidated process for conducting a simulation study (see Section IV). We evaluated this process in a large telecommunication vendor that has a system of systems with high complexity. The simulation study was conducted as action research, where the authors of this paper took on the role of analysts doing process simulation in the company. The goal of the simulation was to assist product management develop a better understanding of the testing process of the large system to which all the subsystems in the product are integrated. This choice of the system was motivated by the complex interactions and dependencies involved in its testing process. It was therefore appropriate to use simulation to better understand this process. The simulation model along with other relavant information is available at [15].

The consolidated simulation process was used as it comprehensively covers the proposals and best practices found in literature. Furthermore, we have reported our experience of using this process from a practitioner's point of view who may have little or no prior simulation experience.

3) Validity threats: There is a risk of overlooking some relevant literature, we tried to minimize this risk by basing the study on a published systematic literature review. We further supplemented it with snowball sampling. Inclusion and exclusion of primary studies was done by first author alone however we tried to reduce the bias in selection by having an objective criteria. Both authors were involved in interviews, presentation and review meetings and compared notes and observations in after each meeting. Second author has been working with the company in a certain role for considerable time now and may have a certain view of the process. However, through triangulation using multiple data sources and multiple practitioners from different teams we tried to overcome any such biased perspective.

IV. Literature Based Consolidated process for SPSM

We started with identification of process steps from individual articles and coded them. Using this bottom up approach we reduced our influence on the formulation of the overall consolidated process. A common trait in all prescriptions of SPSM processes found in this study is the iterative and incremental nature of the process [11] [10] [1]. The consolidation of simulation processes in literature resulted in a process with seven steps and it is presented in first column of Table I. We present a brief description of each step, the sources which recommend them, the model heuristics [1] and guidelines applicable for each of the steps (in Sections IV-A to IV-G). We also report our experience from using this in an industrial study.

Here are some of the guiding principles that are applicable in general to a SPSM:

- Start small and later on add more content to the simulation model, look for analogies rather than starting the model-building from scratch and work over an extended period of time instead of developing the model in one go [5] and maintain frequent contact with the clients [5].
- No model is perfect (i.e. it is an imitation and has a trade off between complexity and realism. [1].
- All models are incomplete (as the goal is to model the behaviour under study that is necessary to answer the questions of interest) [1].
- A model is not reality (therefore all results should be critically analysed) [1].
- It is possible to build many different models of a single process (the perspective of stakeholders, level of detail and assumptions will make the model of same process look very different) [1].
- Continually challenge the model (the credibility of the model can only be increased through further V&V) [1].
- The models are there for decision support, not to make your management decisions for you [1].
- To facilitate effective communication use simple diagrams to communicate with others until they seek more detail (all stakeholders may not understand the equations and it may not be necessary to present those details) [1].

A. *Problem definition, proposed in [2] [10] [3] [11] [1] [6]*

The starting point for SPSM is the identification of a problem to be investigated. The goals of SPSM will influence important simulation decisions like the scope of simulation, level of detail, requirements on accuracy, which will in turn determine the required level of validation [17]. The common purposes for SPSM are [2]: Strategic management; Planning, control and operational management; Process improvement and technology adoption; Understanding; Training and learning. The aim in this step is to identify the key questions that need to be addressed.

Table I
Major steps in a simulation study

Steps in consolidated process for SPSM	Shannon's simulation process [16]
A. Defining the problem B. Designing the model 1. Deciding the model scope 2. Specifying the reference behaviour 3. Identifying the result variables 4. Model conceptualization 5. Input parameters identification 6. Technical feasibility C. Choosing the simulation approach D. Implementing an executable model 1. Choosing the simulation tools and techniques 2. Model calibration E. Verification and Validation 1. Sensitivity analysis F. Simulation based investigations G. Analysis and documentation	1) Problem definition 2) Project planning 3) System definition and determining the boundaries 4) Conceptual model formulation 5) Preliminary experimental design 6) Input data preparation 7) Model translation in a simulation language 8) Verification and Validation 9) Final experimental design 10) Experimentation 11) Analysis and interpretation 12) Implementation and documentation

- Identify the users of the simulation model [6] [11].
- Define the usage scenarios (the use-cases for the simulation model e.g. for question "How does missing the testing hand-off affect the cost and schedule of the release?", a corresponding scenario would be: "Keeping all inputs constants, miss the internal hand-off between testing phases for different types of requirements and observe the values for cost and duration for each of the cases.") of the simulation model [3].
- Develop test cases that will clarify the model's purpose and will be later used for validation of the simulation model itself and its results [3].

Guidelines:

- Consider the audience, desired policy set (e.g. company rules such as having a certain level of quality before moving to the next development stage), and level of detail of the implementation when defining the model purpose [1].
- A model is created to answer specific questions [1], i.e. the SPSM undertaking should be goal-driven [4].
- Identify the users of the simulation model to allow their involvement in the simulation process. It will ensure that their needs are met and increase chances of adoption of simulation model in practice [4].
- Use GQM for defining goals, questions and the appropriate metrics for simulation [3] [1].
- A well-specified purpose will avoid misuse of the model (e.g. a simple model was created for training, but later used for prediction) [4].

Our experience: We started the project with a presentation of the capabilities of process simulation and the main steps in developing a simulation model. We also emphasized the importance of having a focused specific goal for a simulation study. This helped convey the message upfront about the potential benefits of SPSM for the organization and the limitations.

The success of the initial presentation was also visible in the problem identified by the panel of managers from the company, as it was very focused and concentrated on limited phases of the overall development process.

In our case the initial set of users of the model was among the sponsors of this study. But after the initial prototype was presented, their confidence grew in SPSM and its potential. This gave us access to more end-users of the current simulation model (although with similar role and responsibilities in the organization).

As this SPSM study aimed at understanding and training of the managers it was not reasonable to start a new top-down measurement program to support and achieve the goals of the study. We elicited a very focused goal and used GQM for documentation purpose. Instead of starting a new measurement program, we revised the goal and scope of the study to determine what questions can be answered given the metrics that are available, can be derived or accurately estimated.

Therefore, we renegotiated the goal and scope of the study based on the priorities from the client, existing metrics and trade-off on accuracy of results. Perhaps it is the lenient requirements on the accuracy of the model (as the goal of this study was to use the model for training and developing an improved understanding of the test process) that allowed this way of working.

Moreover, creating at least high-level usage scenarios helped to ensure that the modellers correctly interpreted the goals and specific questions that the model should be able to answer.

B. Model Design

1) Model Scope, proposed in [2] [11] [1]: Depending on the aim of the study, the boundaries of system being modelled are defined. The scope is defined in terms of the organizational breadth, product or project team, portion of lifecycle, single or concurrent projects and the time span that needs to be modelled. The system for one study may well be a subset of a larger system.
Guidelines:

- Independent of the scope of the study, to keep the customers interested, modelling should proceed iteratively ensuring the delivery of first executable model as soon as possible [4].
- The scope is largely dictated by the purpose hence the model purpose should be very focused to ensure that it does not become over complex that in turn threatens its validity and credibility [4].

Our experience: This problem statement identified earlier was discussed further in a subsequent meeting to elaborate and characterize the scope of the simulation model. It was decided to focus on the testing process at the company and understand the cost of delay when a requirement misses the system integration test cycle.

While deciding the scope of the system that will be modelled it was important to see what level of dependence exists between the subsystem being modelled and the overall system. In this case, it was the testing process that was the system of interest but we could not capture the implications of decisions at this level without considering the overall development process. Therefore we decided to have an abstract view of other processes and relatively more detailed view of the testing process.

2) Reference behaviour, proposed in [10] [1]: Müller and Pfahl emphasize the importance of specifying observed problematic behaviour or the desired behaviour (called the reference behaviour) [10]. It can also help to identify the model output parameters and can be used later as input for model validation.
Guidelines:

- The reference behaviour should be defined dynamically (plotting behavioural change over time) considering behavioural patterns (e.g. when testing is late the release is delayed), even when not having hard data available in the company [1].
- Instead of striving for absolute measures focus on relative measures (e.g. increase defect count in %) [1].

Our experience: In the case company there was already a measurement based information visualization tool that can graphically represent the software process output. This was especially useful for us to document the existing behaviour of the system. Apart from the benefits of reference behaviour for validation and output parameter identification, we found that being able to replicate the reference behaviour had very positive impact on the interest levels of simulation-model users. As in our case the users were excited to see that the system behaviour was replicated and gave credibility to the model. This increased interest was also confirmed by the fact that they enthusiastically took it up in the project management organization.

3) Result variables, proposed in [2] [3] [11]: The information elements required to address the problem statement by answering the key questions identified earlier. Results variables can influence the process abstraction or level of detail because e.g. if instead of end-to-end values one is interested in intermediate values of a variable then a certain granularity level of the model will be necessary.
Guidelines: Use of GQM to identify the output variables necessary to address the goals of the simulation [3] [1].
Our experience: Creating usage scenarios in ”Problem identification” IV-A helped in identifying variables of

interest. It also helped in categorizing them as input, output and control variables for the simulation model. For example, when we specified a usage scenario like, What is the state of workload, if the workforce allocation is kept constant and failure rate increases by 10 percent?

4) Conceptual modelling, proposed in [2] [10] [3]: Process elements, information flows, decision rules and behaviour that have an influence on the result variables and are relevant to the simulation goal are identified for modelling. This step will involve making use of both the explicit and tacit knowledge about the software process. Some other activities in this step are as follows:

- Create static process models to understand the flow of information and transformation of artefacts in various activities [3].
- Create influence diagrams where (positive or negative) influence of various parameters on each other is depicted [3].
- Create causal loop diagrams (more pertinent for SD) [1] [6].
- Collect and analyse available empirical data for deriving the quantitative relations identified in the influence diagrams [3].
- Distinguish the parameter type as calibration, variable or project specific input parameter [3].

Guidelines:

- Consult domain experts as they have knowledge beyond the documentation and can judge the relevance of information to the problem under study [10].
- Interview domain experts during modelling to avoid missing important aspects of the process and reduce the threat of misunderstanding [4].
- Define a clear, operational purpose of the model (it will help to guide the remaining steps of model development and use) [1].
- Do not try to model the system, i.e. include only the entities and relations necessary to generate the behaviour of interest. Using top-down iterative approach, add details only when necessitated [1]. Start small and later on add more content to the simulation model [5].
- Using the goal and scope of model, aggregate and abstract to an appropriate degree hiding unnecessary details [1].
- KISS (keep it simple, stupid), both the model and communication with users [1].

Our experience: What cause/effect relations should be modelled? This is a difficult choice as well and again one of the most influencing factors is the aim of the simulation. Other than the obvious factors e.g. rework positively reinforces workload, it is rather difficult for a beginner to anticipate more complex feedback loops. Looking at the existing simulation models was helpful to evaluate the relevance of various influencing factors and how others have mathematically modelled them. Had there been an aggregated collection we could easily choose which of these relations are important to map for the simulation goal at hand.

The next difficulty in simulation is the level of abstraction at which to model the process so that there is a balance between the effort expended and realism depicted in the model. For a beginner it is not an easy decision. It was very helpful to use a top down approach i.e. to start with an abstract representation and then add detail later. We did multiple iterations and if the customer thought that a certain aspect of the development process was over simplified then we looked at how adding a certain detail will effect the goal of simulation. If we found that adding this detail is necessary to achieve the aims of simulation, we incorporated it into the model.

We analysed the empirical data early on in the life cycle rather than waiting till the process modelling is done [3]. It helped to face the issues regarding data availability etc. early on and we did revise the scope of the model through discussions with the stakeholders without spending any unnecessary effort.

5) Input parameters, proposed in [2] [3]: These are the independent variables and their choices is largely dictated by the choice of desired result variables. In some cases these variables can take on constant values.

Guidelines:

- Start designing the model early on, independently of the availability of data for calibrating the independent variables [1]. Also note that it is often not possible to measure all variables of relevance accurately.

Our experience: Even if an industrial practitioner takes on the task of SPSM, no one stakeholder has access to all the relevant information required. Also the accessibility of different stakeholders is another constraint in industrial settings. Therefore we devised a template with these two considerations in mind. The template shown in Table II assisted us to effectively acquire the data required for calibration of the simulation model. While filling it out we noticed that certain metrics were not available or that we have to consult other departments.

In cases where the information was not available this initiated the discussion with domain experts about appropriate level of detail for simulation or acquiring estimates for the values. Doing the data source elicitation early on during the study helped to revise the goals of the simulation as well as the appropriate level of detail. For example, in our case the company was interested in seeing the effect of putting in more resources at a certain test level and how that will reduce the required effort at other levels. However, when we looked at the fault slippage data we found that it was abandoned/unreliable and thus we will not be able to calibrate the model. Since this was not the highest priority issue for the stakeholders they decided to skip this as a goal for this simulation study.

Table II
DATA COLLECTION

Item	Description
Variable name	Name the variable e.g. Size
Description	A short description of the variable.
Variable type	Depending on how we want to analyse this variable as independent, dependent or control variable.
Stakeholders (roles)	The stakeholders who are responsible for this variable or who could give more information about this variable.
Activities	The activities in the software development life-cycle that will influence, produce or use this variable's values.
Software artefacts	Software artefacts where this variable can be computed from or is stored in.
Unit of measurement	The unit of measurement.
Simulation values	This may be different from the value above e.g. in the absence of exact numerical values we may have Small, medium, large or Intervals to estimate size for simulation.
Typical and extreme values	The typical, minimum and maximum value of this variable.
Distribution	The distribution of this variable.
Tolerance	The acceptable tolerance in the estimation, in case of a dependent variable.

6) Technical feasibility, proposed in [6] [11]: Before the initial simulation model development technical feasibility should be checked [6] [11].
Our experience: As novice modellers we did not feel confident to comment on the "technical feasibility" before the initial model development as recommended by [6]. We tried to increase the chances of correctly assessing the technical feasibility by delaying the decision till initial prototyping of the model.

C. Simulation approach, proposed in [2]

As software processes have both discrete and continuous aspects, e.g. there are discrete quantities like lines of code, defects, number of requirements etc. and continuous aspects like staff experience, motivation level, commitment, cohesion in the teams etc. are continuous. So different types of simulation approaches are applicable in SPSM [1]. The choice of an appropriate simulation approach will therefore depend on the aim of the study and the modelled system. A multitude of simulation approaches are available e.g. Zhang et al. [14] found 10 approaches that have been used for SPSM. The most commonly used approaches are SD and DES. The choice is again influenced by the purpose, scope and result variables among others.
Guidelines: Continuous simulation is the approach of choice when the analysis does not require the low-level process details e.g. for strategic analyses and long-term behaviour. SD models have levels and flows of entities. These entities are not individually identified and are not traced through the process [2]. DES is the approach of choice for detailed process analysis, relatively short-term analysis. DES models contain distinct entities with attributes that move through the process [2].
Our experience: The decision of the simulation modelling approach is difficult for a practitioner with no simulation background. There is a long list of possible approaches that have been used for SPSM. We looked at literature and short listed SD, DES or Hybrid simulation with both continuous and discrete simulation because of the frequency of use and availability of tool support. By keeping the aims of the study in mind where we want to develop a better understanding of the process and use it as a training tool, we decided to use SD as we were interested in the macro view of the process and wanted to look at the overall flow of artefacts through the process. Furthermore the general implication of ”missing the iteration" (i.e. fails at a certain intermediate test level and has to go back to development and wait for the next iteration) was more important for the users than what happens when a particular requirement misses the iteration. We felt that the case of taking hybrid simulation is very compelling as we could relate with the reasons mentioned in literature. In addition to the aforementioned reasons for starting with SD modelling, we are SPSM novices, unfamiliar with both discrete event and continuous simulation. However, we also decided to stay open to switch to a different choice later if we will be unable to achieve the goal with the current approach.

D. Implementation of executable model, proposed in [10] [3] [11]

The choice of simulation approach chosen for SPSM will determine the details of this step and exactly how the model will be implemented. Rus et al. [3] are the sole proposers of creation of "High level" and "Detailed design" of the simulation model. Perhaps this is unique to DES as none of the other process descriptions have suggested such design activity.
Guidelines:

- Do not add all factors at first, but create the model in an iterative manner adding factors incrementally [1].
- Relationships between elements of the model should also be added iteratively. [1].
- The use of relative measures and normalizing them helps in scaling the model. [1].
- Always keep the model in a state that it could be simulated (or tested) [1].
- Actively involve the actors being modelled in the model building process [1].

Our experience: We found follow-up meeting with the users very helpful where the implemented model was presented. The ability of the simulation tool to dynamically change values of parameters and show the effect on output variables was also very helpful for the users. It helped them to understand what was going on and reflect whether it made sense in their context.

1) Simulation tools and techniques, proposed in [2] [11]: A simulation can be deterministic (which require only one simulation run), stochastic or hybrid (which require multiple simulation runs and rely on statistics to analyse result variables). A number of tools are available today to facilitate these simulation techniques. Many come with graphical interface (facilitating walk-through), output visualization, support for interactive simulation, sensitivity analysis and connectivity with third party applications.

Guidelines: Madachy [1] lists the price of the tool, its documentation, training opportunities, support and maintenance, computer platform and user familiarity as important criteria for choosing a simulation tool.

Our experience: The choice of simulation tool, there is again a long list of simulation tools and the features they support. To the best of our knowledge there are no comparative studies where these tools may have been evaluated objectively for the learning curve, ease of use, effort required to model and interpret results. It will also be useful if we knew the bare minimum features that should be available and are useful for each simulation approach so that a practitioner may make an informed decision. We choose the Vensim tool because of the graphical interface, its ability to allow runtime changes to parameters, visual representation of result variables and state of the system. Some other tools also provided similar features but as we choose SD and Vensim happens to be the tool of choice for this type of simulation [14]. This also motivated our choice.

2) Model calibration, proposed in [1] [2]: In this step the model is calibrated against real world data. However, the lack of data for model calibration and validation is a common issue in real-world settings [2]. The desired data is often poorly defined, inaccurate or missing altogether.

Guidelines: Here are some suggestions from Raffo and Kellner [17] to deal with the typical situations that modellers have to face:

- Consult domain experts to deduce the accuracy and relevance of data.
- If the desired metric is not available but a similar one is then make calculated adjustment, make adjustments by consulting experts, use source documents if you need finer granularity instead of aggregates, adjust decision variables to capitalize on available data or adjust the model scope.
- If the desired metric is not available and there are no similar ones: reconstruct the metric using other data sources, estimate using expert opinion, look for data in literature, drop the variable altogether.

Our experience: The issues identified above and the mitigation techniques proved fairly comprehensive in our case. We faced most of the issues highlighted above and chose relevant mitigation techniques. Here is the list of some other challenges that we faced during calibration:

- Always consult domain expert for the real meaning of the data fields even if they have a perfectly straight forward name. In the case organization while eliciting the units of the variable "Size" that it was the estimated effort in person hours instead of a size metric. The template presented in Table II proved useful here to surface such misunderstanding before it could cause any major problems
- We found that there were certain "silos" of information e.g. the information about the requirements was in one database and defects reports in another and there is not explicit connection between them. So although both data points exist we cannot use it as there is no traceability between them.
- We found during discussion with domain experts that certain information in the databases was unreliable (although it had legitimate values). It was unreliable as it was difficult to compute in reality and was filled with typical guesses only because it is a mandatory field.
- Since data required was distributed in various data sources across various departments and it was difficult to find accurate descriptions for data-fields in documentation. We had to frequently find the right experts and consult them, which was fairly time consuming.
- Another issue which was not as big as unreliable or missing data but still made the task of calibration difficult was the inability of various interfaces (that were custom built for a certain purpose) to export data to a file that may be consumed in other tools.
- The choice of that time span of historical data to use for calibration of model needs discussion with domain experts. As we have to use a time period long enough to model the true behaviour rather than any temporal trend and yet capture the current reality. So the knowledge of underlying changes e.g. if anything from the development process, programming language or development platform has changed then we need to be aware of that.

E. Verification and Validation, proposed in [2] [3] [11] [10] [6] [1]

A simulation model should undergo validation to the extent possible [2] and V&V should go on throughout the simulation study [3].

Guidelines: Seven V&V activities are presented in [17]. Please refer to Ahmed et al. [18] for a comprehensive framework (where they have combined existing work on V&V) for evaluation of software process simulation models that may be used as a guide as well.

- Face validity: by using inspections and structured walk-through of the model [2] [3] [11].
- False expectations (e.g. perfect predictions on the first model run) should be avoided, rather patterns should be reviewed for qualitative similarity instead [1].
- Consider constraints from the real world when conducting sensitivity analysis for a policy or a combination thereof to make sure that the model reflects real world behaviour [1].
- Model validity is a relative matter [1] as it is difficult to thoroughly validate a software process simulation model because of data required is not available and

validation throughout the process is costly [17].

Before the full-scale development of the simulation model following steps should be done:

- Validation of the SPSM Requirements [3].
- Review of causal diagrams [6].
- In the review meeting describe the problem that will be addressed with SPSM, delineated system boundary and what will be the expected output of the study [1].
- Using the prototype to show how a user of the model will manipulate the input and control variables in the simulation [1].

Our experience: We used structured walk-through approach for face validity of simulation model. We found that presenting more details in increments was helpful. Here is the order we followed in validation meetings:

- Repeat the goals of simulation model.
- Discuss specific questions that will be answered by the simulation.
- Discuss usage scenarios [3] of simulation model.
- Discuss influence diagram.
- Discuss a high-level view of the model (hiding the implementation details) jump to the detailed version only if required.
- Discuss all assumptions and simplifications done in modelling the process and reasons for making them.

We found developing concrete questions and usage scenarios was very helpful to validate that we (as modellers) properly understood the goal. A walk-through of influence diagrams and causal loops instigated a lot of discussion about which cause effects relations are visible in the organization's context. Stakeholders also reflected on how some cause-effect relations are strongly visible in their organization. It was interesting to see that unlike our assumption that a higher schedule pressure would create a higher failure rate when the system is delivered to testing didn't hold true in the organization. Understanding the reasons was beyond the scope of this study but one reason may be the use of reviews, functional testing and automated test suites by development before the code is delivered for system and systems integration testing. We found that explicitly documenting and presenting the assumptions and simplifications we made was very useful. As this helped get the stakeholders perspective if these simplifications were unrealistic in their context. Also whether a certain abstraction will hinder us to capture certain important detail of reality that is important for the aim of the simulation.

We followed the presentation with a structured interview to assess the validity of simulation model. The question used in this interview are available at [15]. We used face validity of graphical models, replication of trends in reference behaviour, verification of model inputs and outputs (whether the calculations are correct) and qualitative assessment of reasonableness of model output. These were deemed sufficient, as the purpose of this model was understanding and training.

1) Sensitivity Analysis, proposed in [2] [11] [1]: It is applicable to all types of simulation approaches [2]. It explores the effect of changing certain parameter values on result variables. It helps identify how significant is the effect of a certain parameter on the results of the simulation. The extent of statistical analysis however can be adjusted according to the use of data and the model [17].

Guidelines: Madachy [1] discusses application of numerical, behavioural and policy sensitivity analysis on simulation models. Begin with changing values of one parameter at a time keeping others constant [1].

Statistical analysis is facilitated by good interconnectivity (automated data transfer) between the simulation and statistics tool [19].

For stochastic models use Monte-Carlo analysis [1].

Our experience: Since the purpose of the model is for understanding the interaction of various variables in the testing process and how different decisions are intertwined accuracy of results is not the highest priority. It was sufficient if the model manages to replicate the real life behaviour and trends in output variables given the calibrations from the industry. Therefore we developed a deterministic model of the process. So for sensitivity analysis we only used three cases for each of the parameter values (minimum, typical and maximum). It was helpful to start with altering one variable at time as it was possible to see the effects of that and interpret whether it still made sense in all three cases or not. Since the extreme and typical values were computed from the industrial data we were confident that model is valid at least for the "relevant" range of values.

F. Simulation based Investigations, proposed in [10] [11] [1]

We purposefully replaced "experiments" with "investigations" as we agree with Wernick and Hall [20] that the term "experiment" is misleading in the context of SPSM. In this phase we are interested to study the modelled system. We design and conduct investigations to answer the key questions raised in the start of the simulation study.

Guidelines: Raffo and Kellner [17] explored four approaches to evaluate alternatives using simulation results: simple comparison of performance measure deltas, comparison of performance measure deltas using utility functions, comparison of performance measure deltas using financial measures or comparison of overall performance measure values using Data Envelopment Analysis.

Our experience: The aim of this study was to train managers to understand the implications of various decisions during the testing process therefore the object evaluation of alternative policies or decisions were not done. Furthermore the accuracy and credibility of the model (since it was developed for understanding and training) does not allow such objective assessments.

G. Application and documentation, proposed in [11] [6] [1]

Report the findings of simulation based investigations and put them in context, apply and use the results and document the simulation model and its usage. A proper documentation will ensure that the model is not misused and will support its maintenance in the future.

Guidelines: Madachy [1] presented a report template that is a good starting point and can be customized with level of detail depending on the purpose of simulation study.

- Always show equations (improves maintainability and supports reuse at least at this level) [1].
- Like any software, well-commented equations will be easier to understand [1].
- Document assumptions, calibration values and rationales [1].
- Show output of multiple cases and discuss them [1].

Our experience: We used Madachy's [1] report template as a checklist when documenting the simulation model. We found that just presenting the values of result variables is not enough for the users. The results need to be put into context where the limitations of the results and their implications are interpreted and discussed. Since simulations are an abstraction of reality the results were analysed critically and other competing explanations for the results were sought and presented to the users. This increased the credibility of the simulation results and initiated healthy discussions about the results of the simulation.

V. Discussion

We can see that the formalism that was earlier visible in the SE field in general also influenced the SPSM methodologies where a lot of emphasis was on specifications e.g. [3]. Most of the methodologies found in this study are incremental and iterative in nature e.g. [11] [10] [1] [6].

The overall process for SPSM is very similar for different simulation approaches. The only major differences occur in the implementation phase, where the tactics and guidelines will be particular to a certain simulation technique or tool.

Winter Simulation Conference (WSC), since 1967 has become a premier forum for information on system simulation with principal focus on discrete-event, and combined discrete-continuous simulation in all disciplines. We compare the simulation process presented by Shannon [16] in this forum (as shown in second column of Table I) to the SPSM-literature based consolidated process description (as shown in the first column of Table I). We choose this as a comparison point as it is frequently cited (121 citations in Google Scholar) in diverse areas e.g. mining and healthcare. Compared to other engineering disciplines, absence of established physical laws raises unique challenges for model calibration and validation in SPSM. However, besides these differences, looking at the similarity between two processes in Table I) we may conclude that the overall simulation process is independent of the organizational context, simulation approach and experience of the modellers. The similarity between these two independently created processes also adds confidence to our literature based consolidated simulation process as it indicates the stability of the process.

Acknowledging the similarity of process and applicability of guidelines opens new possibilities to learn from other disciplines that have been using simulations far longer than relatively new SPSM discipline. For example, the guidelines and tips given by Robinson [21] for conceptual modelling activity, a vital part of a simulation study, are also valid and applicable for SPSM and have some overlap with Madachy's "Modelling heuristics" [1].

Furthermore, the availability of process guidelines will only take us so far. In addition we need to document and make the collections of our combined knowledge available in more reusable way. Till the dream of universal software process that is customizable to every context or use of patterns in SPSM becomes a reality we must facilitate the development of simulation models from at least reuse at design level. Software engineers do not have established laws like other engineering fields where the models are based on established physical laws. This poses challenges for both calibration and validation of the models [22]. In these circumstances, attempts like: The compendium of 86 "fundamental rules of software engineering" [23] can be a useful starting point. Such lists can be expanded or reduced with empirical evidence as it evolves. Also a collection of parameter values used for calibration along with the good description of the context of their use will considerably help in model calibration in the dearth of real data.

VI. Conclusion

This paper presents a consolidated process for SPSM and presented the steps of the process and provided guidelines for each step. The process identified in literature is completed by our own experience of using it in an industrial case. This led to the following answers for our research questions:

RQ1: *Which process descriptions and guidelines are reported for SPSM in industry?* We identified five different processes for SPSM. Although literature in SPSM differentiates between process descriptions for different simulation approaches but through analysis of these descriptions we found that the overall process of SPSM is independent of the simulation approach. The choice of a particular simulation approach will largely affect the specific operational tactics in implementation phase of SPSM life-cycle. Various authors have highlighted the lack of process descriptions for SPSM [3] [11] and then went on to propose simulation models to overcome this gap. Some differentiated the process models on the basis of formalism, level of experience of modellers and the size of the organizations undertaking a SPSM. However, we found that there is remarkable similarity between these proposals. It was also interesting to see that the combined process developed from consolidating these guidelines looks very similar to the simulation process proposed in a venue that is not focused on software domain.

RQ2: *Is the consolidated simulation process useful in an industrial study?* In our experience the existing process descriptions are ample for overall design and execution of SPSM in industry. However, the need is to supplement these guidelines with practical tactics relevant to the software domain, which may then be validated and improved with empirical evidence. We also found that the guidelines might vary with respect to the simulation goal. For example, our goal was training and hence e.g. influenced the level of detail and how sensitivity analysis was done, which would receive more focus when creating a model for project planning.

In future work the experiences obtained in this action research need to be extended by experiences in different model building contexts (e.g. models for prediction purposes).

Acknowledgment

This work has been supported by ELLIIT, the Strategic Area for ICT research, funded by the Swedish Government.

References

[1] R. J. Madachy, *Software Process Dynamics*. Wiley-IEEE Press, 2008.

[2] M. I. Kellner, R. J. Madachy, and D. M. Ra, "Software process simulation modeling : Why ? What ? How ?" *Journal of Systems and Software*, vol. 46, 1999.

[3] I. Rus, H. Neu, and J. Münch, "A systematic methodology for developing discrete event simulation models of software development processes," in *Proceedings of Software Process Simulation Modeling Workshop (ProSim 2003)*. Citeseer, 2003.

[4] D. Pfahl and K. Lebsanft, "Integration of system dynamics modelling with descriptive process modeling and goal-oriented measurement." *Journal of Systems and Software*, vol. 46, pp. 2–3, 1999.

[5] R. Ahmed, "Software process simulation modelling: A survey of practice," *Journal of simulation*, vol. 2, no. 2, pp. 91 – 102, 2008.

[6] D. Pfahl and G. Ruhe, "IMMoS: a methodology for integrated measurement, modelling and simulation," *Software Process: Improvement and Practice*, vol. 7, no. 3-4, pp. 189–210, Sep. 2002.

[7] ——, "IMMoS: a methodology for integrated measurement, modelling and simulation," *Proceedings of Software Process Simulation Modeling Workshop (ProSim 2003)*, 2003.

[8] N. Angkasaputra and D. Pfahl, "Towards an Agile Development Process of Software Process Simulation," in *6th International Workshop on Software Process Simulation and Modeling (ProSim)*, 2005.

[9] J. Zhai, Q. Yang, F. Su, J. Xiao, Q. Wang, and M. Li, "Stochastic process algebra based software process simulation modeling," in *International Conference on Software Process*. Springer, 2009, pp. 136–147.

[10] M. Müller and D. Pfahl, *Simulation Methods*, 1st ed., F. Shull, J. Singer, and D. I. K. Sjø berg, Eds. London: Springer, 2008.

[11] R. Ahmed, T. Hall, P. Wernick, and S. Robinson, "Evaluating a rapid simulation modelling process (RSMP) through controlled experiments," in *2005 International Symposium on Empirical Software Engineering*, Piscataway, NJ, USA, 2005.

[12] C. Silva Filho, A. Cavalcanti da Rocha, and Others, "Towards an Approach to Support Software Process Simulation in Small and Medium Enterprises," in *Software Engineering and Advanced Applications (SEAA), 2010 36th EUROMICRO Conference on*. IEEE, 2010, pp. 297–305.

[13] H. Neu and I. Rus, "Reuse in software process simulation modeling," in *Proceedings of Software Process Simulation Modeling Workshop ProSim03*. Citeseer, 2003, pp. 1–4.

[14] H. Zhang, B. Kitchenham, and D. Pfahl, "Software process simulation modeling: Facts, trends and directions," in *Software Engineering Conference, 2008. APSEC'08. 15th Asia-Pacific*. IEEE, 2008, pp. 59–66.

[15] "Additional information for this SPSM study," http://www.bth.se/com/nal.nsf/pages/spsm-info, [Acc. Mar. 2012].

[16] R. E. Shannon, "Introduction to the art and science of simulation," *Proceedings of the 1998 Winter Simulation Conference*, pp. 7–14, 1998.

[17] D. M. Raffo and M. I. Kellner, "Empirical analysis in software process simulation modeling," *Journal of Systems and Software*, vol. 53, no. 1, pp. 31–41, 2000.

[18] R. Ahmed, T. Hall, and P. Wernick, "A Proposed Framework for Evaluating Software Process Simulation Models," in *4th International Workshop on Software Process Simulation and Modeling (ProSim)*, 2003.

[19] W. Wakeland, R. Martin, and D. Raffo, "Using design of experiments, sensitivity analysis, and hybrid simulation to evaluate changes to a software development process: a case study," *Software Process: Improvement and Practice*, vol. 9, no. 2, pp. 107–119, 2004.

[20] P. Wernick and T. Hall, "Getting the best out of software process simulation and empirical research in software engineering," in *Proceedings of the Second International Workshop on Realising Evidence-Based Software Engineering*, ser. REBSE '07. Washington, DC, USA: IEEE Computer Society, 2007, pp. 3–.

[21] S. Robinson, "Conceptual modeling for simulation: issues and research requirements," in *Proceedings of the 38th conference on Winter simulation*, ser. WSC '06. Winter Simulation Conference, 2006, pp. 792–800.

[22] A. Christie, "Simulation: An enabling technology in software engineering," *CROSSTALK–The Journal of Defense Software Engineering*, vol. 12, no. 4, pp. 25–30, 1999.

[23] E. O. Navarro, "The Fundamental Rules of Software Engineering," 2002. [Online]. Available: http://www.ics.uci.edu/~emilyo/SimSE/se_rules.html