Software Process Simulation over Decade:
Trends Discovery from A Systematic Review

He Zhang
National ICT Australia
School of Computer Science
and Engineering, UNSW
he.zhang@nicta.com.au

Barbara Kitchenham
School of Computer Science
and Mathematics,
Keele University, UK
barbara@cs.keele.ac.uk

Dietmar Pfahl
Simula Research Laboratory
Department of Informatics,
University of Oslo, Norway
dietmarp@simula.no

ABSTRACT
Software Process Simulation (SPS) research has increased since 1998 when the first ProSim Workshop was held. This paper aims to discover how SPS has evolved during the past 10 years based on the preliminary results from the systematic literature review of SPS publications from 1998 to 2007. Trends over the period showed that interest in continuous modelling was decreasing and interest in micro-processes was increasing. Hybrid models were based primarily on system dynamics and discrete event simulation and were all implemented by vertical integration.

Categories and Subject Descriptors
D.2.8 [Software Engineering]: Management—Software process models

General Terms
Management

Keywords
Software process simulation, systematic literature review

1. INTRODUCTION
Software Process Simulation (SPS) was introduced into the software engineering domain by Abdel-Hamid’s pioneering work [1]. In the last two decades, it has been emerging as an effective tool to help evaluate and manage changes made to software projects and organisations. As a major research event, the ProSim series conference has taken place since 1998, and focuses on up-to-date theories and practice of SPS research in addressing real-world problems.

Recently, we undertook a systematic review to revisit and update the state-of-the-art of SPSM research, to summarise the 10-years (1998-2007) progress, and to discover trends for our future research activities in this domain. The process and preliminary results of this review has been reported in [5]. This paper presents the underlying trends over the decade derived from the review results.

2. THE SYSTEMATIC REVIEW
This section briefly describes the method applied in this review and the studies included.

2.1 Method
This research follows Kitchenham’s methodological guidelines for systematic reviews [3]. Three researchers were involved, including one principal and one secondary reviewer, plus the third acting as the expert panel. The follows are the research questions predefined for this review:

Q1. What are the purposes or motivations for SPS in the last decade’s practice?
Q2. Which simulation paradigms have been applied in the last decade, and how popular are they in SPS?
Q3. Which simulation tools are available for SPS and have been in use in the decade? And how popular are they?
Q4. On model level, what are problem domains and model scopes focused on by simulation models?
Q5. On parameter level, what are the output variables considered when developing simulation models?
Q6. Which simulation paradigm is the most appropriate for a specific SPS purpose and scope?

The time frame of sources for this study is constrained to the period from 1998 to 2007. As the ProSim series are regarded as the most important forum of SPS, the related sources, i.e. ProSim, ICSP, Journal of Systems & Software (JSS) and Software Process Improvement & Practice (SPIP), are the primary data sources for this research.

2.2 Studies
The pilot review identified four categories of primary study:

A. Software process simulation models or simulators;
B. Process simulation modelling paradigms, methodologies, and environments;
C. Applications, guidelines, and solutions for adopting process simulation in software engineering practice;
Table 1: Sources identified as primary studies

<table>
<thead>
<tr>
<th></th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proc.</td>
<td>15</td>
<td>13</td>
<td>21</td>
<td>0</td>
<td>32</td>
<td>27</td>
<td>24</td>
<td>8</td>
<td>8</td>
<td>32</td>
<td>148</td>
</tr>
<tr>
<td>JSS</td>
<td>11</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>SPIP</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td>selected</td>
<td>13</td>
<td>9</td>
<td>14</td>
<td>7</td>
<td>16</td>
<td>10</td>
<td>13</td>
<td>6</td>
<td>8</td>
<td>96</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Paradigms applied in simulation models

<table>
<thead>
<tr>
<th></th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>DES</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>SBS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>KBS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>QSIM</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>RPG</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>ABS</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>DTS</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Stig.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Hybrid</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SD: system dynamics  DES: discrete-event simulation
SBS: state-based simulation  KBS: knowledge-based simulation
RPG: role-playing game  ABS: agent-based simulation
Stig.: stigmergy  QSIM: qualitative/semi-quantitative simulation
Emrg.: emergent  DTS: discrete-time simulation

D. Experience reports of SPS research and practice.

These four categories are not a mutually exclusive. They focus on different aspects of SPS research, and may give answers to the research questions from different perspectives.

In total, 209 papers have been published in the ProSim sources. They form a comprehensive body of knowledge of software process simulation. By carefully reviewing their titles, abstracts, keywords, conclusions, 96 articles were selected and identified as the primary studies. The number of papers per year and source are summarised in Table 1.

3. TRENDS DISCOVERY

Trend means “the general movement over time of a statistically detectable change” (Merriam-Webster Dictionary). This section attempts to detect the interesting ‘movement’ or ‘change’ derived from the review results over the decade.

3.1 Paradigms

In 1998, there were only three simulation paradigms employed by the Category A studies (models) published in the first ProSim workshop (Table 2). They were SD, DES and KBS. Kellner, Madachy and Raffo (KMR) discussed four types of simulation in their seminal paper in ProSim community [2]. However, our review identified ten simulation paradigms from the primary studies.

Trend 1 System dynamics and discrete-event simulation formed the main stream of SPS paradigms.

Trend 2 New simulation paradigms continued to be introduced into SPS research.

Trend 3 Continuous modelling gradually lost its dominant position in SPS research in comparison with discrete approaches during the decade.

When revisiting the top two rows of Table 2, both SD and DES supported simulation studies of software process...
each year. Although the total number of SD models is larger than any others', it is not difficult to observe a decreasing trend of SD application in the decade, by absolute number or percentage. The studies using SD dominated in the early years (prior to 2000). Recently, the number of published studies using SD has decreased and approached the number of studies using DES. It implies that simulation research has become more interested in micro-process modelling.

Process level simulation models are based on the conventional life-cycle or sequential process modelling. However, it is difficult to simulate team effects by process level modelling through aggregating individuals. This becomes an issue when modelling a process with many participants, such as the open source developer community. Accordingly, modelling and simulation technology at entity level requires quite different characteristics from the technology used on process level. Simulation at entity level consumes more resources to track all entities and their relationships individually than the approaches using aggregation.

Table 4 indicates for each simulation paradigm the inherently supported research granularity level(s) as identified from the review. Here, RPG is a special case, its supported granularity level often depending on its combined approach.

Trend 4 Most of newly introduced paradigms enhanced the research capability at the micro-process level.

All newly introduced paradigms are listed at the right side of the vertical line (between KBS and QSIM) in Table 4 (compared with the four discussed by KMR at leftmost). If we define the scope of micro-process to cover the entity granularity level, then most of them are capable of the process or entity level research corresponding to the micro-process.

Trend 5 In recent years, micro-processes have been attracting more simulation research.

Table 5 records the number of published studies on each granularity level over years. When applying the extended concept of micro-process, since 2004 the number of micro-process simulation models has been no less than macro-process models. It is also an enhancement of Trend 3.

3.3 Integration

Like the interaction between macro-process and micro-process [4], process modelling on the three granularity levels is not mutually exclusive, and sometimes combined.

Hybrid modelling employs more than one paradigm in developing a process simulation model. The systematic review concludes that hybrid process simulation models have attracted interest as a possibility to avoid the limitations of applying single modelling method, and more realistically capture complex real world software processes.

Trend 6 System dynamics and discrete-event simulation were the most common combination for hybrid modelling.

The bottom line of Table 2 shows the number of hybrid models published per year. All these 10 models were at least based on the combination of SD and DES, or even more.

Trend 7 Hybrid process simulation modelling concentrated on vertical integration.

There exist two broad approaches to develop a hybrid process model: vertical integration, which primarily implements discrete modelling at the lower process level, then continuously calculates the process factors and incorporates the feedback loops at system level; and horizontal integration, in which the sub-processes or phases within a large scale or complex software process may be modelled using different approaches respectively and sequentially, and the data flow has to be converted at the interface in transition. We found all the hybrid models reported in primary studies were implemented by vertical integration.

4. CONCLUSION

The preliminary results from the first systematic review of software process simulation we conducted has been reported in [5]. This paper further presents the in-depth findings of SPS research by discovering the trends over the past decade. In addition, we also gain the new insights about SPS research on modelling granularity and integration.

The results from this review can help both insiders’ and outsiders’ observation and understanding of SPS research. These trends combining the preliminary facts can suggest possible directions for future SPS research.

5. REFERENCES