Janez Povh (ed.)

Applied modelling and computing in social science

PL ACADEMIC RESEARCH

Biljana Mileva Boshkoska, Nadja Damij
Faculty of Information Studies
Sevno 13, 8000 Novo mesto, Slovenia
[biljana.mileva, nadja.damij]@fis.unm.si

An outline for business process modelling using spreadsheets

Abstract: Process simulation is the first step towards the redesign of business process models. However, the leading process simulation tools are usually expensive and often unavailable for common users. On the other hand, almost every computer user has experienced the usage of spreadsheets at least once. Their vast usage is due to its simplicity and to its availability in companies. This paper presents an approach for the usage of spreadsheets for process simulation. We argue that for modelling of simple business processes, the usage of spreadsheets provides just as reliable results as any other expensive simulation tool.

Keywords*:* business process modelling, spreadsheets

1 Introduction

Business process simulation (BSP) has an essential role in the process of improvement of business process management (BPM) in the public and private sector. BPS is a tool which companies use to get feedback regarding their performances under different conditions and thus enables the redesign of the business process. Regardless of the fact that BPS is acknowledged as relevant and highly applicable, the use of simulation is limited in reality (Nakatumba, Rozinat and Russell 2008). We have managed to observe two main reasons.

The first reason is that there are more than 100 available tools for BPS, each one with different specifics; however, there is a lack of guidance on how to choose a particular tool for simulating a specific task. The second reason is that most of these tools are either too expensive to buy for small companies and start-ups or require a great knowledge for their set up. This paper provides an alternative to the expensive simulation tools by providing an outline for using spreadsheets for simulation of business processes. The reason that we have chosen spreadsheets is because of their vast usage today. They are easy to install and do not require high programming experience to use them.

2 Business Process Modelling and Simulations

In general, a process consists of interconnected activities that require certain resources for their implementation. Thus, to understand the real processes, one has to observe their behaviour and consequently the behaviour of activities and process resources. The observation is usually performed in three steps. Firstly, the process is modelled using some well-known technique. Next, the model is validated in order to find out how close the model reflects the real process. Then the model is analysed by using "what-if" questions to test the options of interest and the functional possibilities functioning of the process. Finally, the model is used for simulation of different scenarios in order to observe the behaviour of the process. The simulations are based on different scenarios using a set of data and assumptions about the process activities and resources (Banks et al. 2001).

Business processes (BP) are modelled with the aim of analysing their current states within the organisation, as well as improving them through the execution of potential "what-if" simulation scenarios (Saven, Olhager 2002). The use of scenario-based what-if analyses enables the design team to test various alternatives and choose the best one (Laguna, Marklund 2005). BP models are simulated with discrete-event simulation (DES) tools. When using DES, the state variable changes only at a discrete set of points in time. When using DES for BPM, the following basic elements have to be considered:

- State: A collection of variables that contain all the information necessary to describe the system at any time;
- Entity: Any object or component of the system which requires explicit representation in the model (i.e. a server, a customer, a machine);
- Attribute: The properties of a given entity (i.e. the priority of a waiting customer, the routing of a job through a job shop);
- Activity: A duration of time of specified length (i.e. a service time or arrival time), which is known when it begins (though it may be defined in terms of a statistical distribution);
- Delay: A duration of time of unspecified indefinite length, which is not known until it ends (i.e. a customer's delay in a last-in, first-out waiting line which, when it begins, depends on future arrivals);
- Clock: A variable representing simulated time.

All of the above described elements can be modelled using spreadsheets.

3 Process Simulation using spreadsheets

3.1 An overview of process modelling using spreadsheets

There are many research papers that describe modelling and simulations using spreadsheets. For example, spreadsheets are used for simulations in healthcare (Klein, Reinhardt 2012; Hartvigsen 2001), in pharmacokinetic processes (Meineke, Brockmöller 2007), and for the simulation of optimisation processes of control parameters (Papa, Mrak 2010). Many companies use spreadsheets for Monte Carlo simulations (Croll 2012); Emmett, Goldman 2004) or optimisation purposes (Croll 2012). However, spreadsheets are used for the simulation of complex physical phenomena, such as the calculation and simulation of the magnetic field of solenoids (Derby, Olbert 2010), or for finite difference time domain simulations for propagating electromagnetic waves (Ward, Nelson 2005), and for the simulation of microprocessor systems (El-Hajj et al. 2000). They are used for the simulation of functions by simulating basic blocks that are interconnected into a graphical interface (El-Hajj 1999) and for the simulation of the work of components of an analog computer so that a simulation of the whole system is obtained (Karim Y. Kabalan 1997). A recent publication in bioinformatics that uses spreadsheets simulates the consequences of varying parameters on measurement results of trace amounts of non-authorised genetically modified organisms (GMO) in feed (Gerdes, Busch, Pecoraro 2014). The processes are usually modelled with equations that are interconnected or by diagrams obtained by interconnecting different shapes offered by the spreadsheets.

3.2 Process Simulation using spreadsheets

Each of the basic DES elements described above, a state, entity, attribute, activity, delay, and clock can be modelled using spreadsheets. We first start with defining the following three elements in spreadsheets: an activity, a transaction, and a gate.

An *activity* has a duration and requires resources for its execution. *Transactions* are flow units that move through the process model from one activity to the next one (iGrafx Help system, 2013). For example, in a model that describes the healthcare process, patients are an example of transactions. They represent entities from the aspect of the defined elements in DES. A *gate* is a modelling element that is activated at a particular time of the day. It means that the transactions that arrive at the gate have to wait until the gate's door opens at the defined time. It represents a *certain delay* in the sense of DES. Uncertain delays are described later. A gate does not have time duration nor resources. Therefore, it is not considered as an activity.

Each modelling element is defined in one spreadsheet. Each row in these spreadsheets represents the *state* of the modelling element at a particular moment of time. A process in spreadsheets is defined by interconnecting the spreadsheets with formulas that lead to the simulation of transactions that move from one modelling element to the next one. The configuration data for each modelling element, such as duration time, and all initial set ups, such as number of starting transactions, and their interconnected start up time are given in a separate worksheet called the *Input Data* worksheet. One example of an *Input Data* worksheet is given in Table 1. Table 1 refers to a process model of the abdominal surgery clinic in Ljubljana, Slovenia. All consecutive tables in this paper refer to the same healthcare process.

Table 1: Excel simulation input data worksheet

Table 2a

Number of beds	75

Table 2b

Parameters of triangular distribution	Activity 18	Activity 24
a	1	3
b	5	5
c	4	4

Table 2c

Activity ID	Activity Name	Resource	Duration (Minutes)
01	Confirm surgery date	1	10
02	Patient reception	1	15
03	View and organize documents	1	10
04	Fulfill care documents	1	15
05	Explain surgery	1	10
06	Sign surgery documents	1	5
07	Sign anaesthesia documents	1	5
08	Prescribe medications	1	10
Gate1			Every morning
09	Prepare patient for surgery	1	30
10	Carry out anaesthesia	1	Between(30;45)
11	Carry out surgery	1	Between(60;240)
12	Wake up patient	1	30
13	Place in intensive care	1	30
14	Prescribe therapy	1	10
15	Addition-al tests	1	10
16	Order additional tests	3	10
17	Carry out care	1	Between(90;180)
18	Check recovery	2	TringDist(1;5;4)
IntG			Every morning
19	Place in Clinic	1	20
20	Prescribe therapy	2	10
21	Addition-al tests	1	10
22	Order additional tests	2	10
23	Carry out care	1	Between(90;240)
CliG			Every morning
24	Check recovery	2	TringDist(3;5;4)
Gate2			Every morning
25	Inform about release	1	10
26	Prepare release report	1	30
27	Organize transport	1	10
28	Need transport	1	10

The *Input Data* organises the configuration parameters into three tables. The first table, Table 1a, defines the constraints of the model, which for the simulation model of abdominal surgery clinic is presented with the maximum number of available beds. The second table, Table 1b, provides various data about the parameters of the modelled activities, described with different distributions.

Table 1b gives the parameters of the triangular distributions that are required in order to determine the duration of some of the activities, such as activities 18, 24, and others. For example, a triangular distribution determines the duration of the patient's stay in intensive care (activity 18 in Table 1c). It has the following parameters: minimum stay of the patient in the intensive care is one day, maximum stay is five days, and the most probable stay is four days. The actual duration is estimated with the function "rand()" that generates a pseudo-random number uniformly distributed in the interval (0,1), which is scaled to the desired distribution according to the given parameters in Table 1b. This transformation is known as a Probability Integral Transformation (Leemis, Park 2005). This way, we model a *delay* from the aspect of the mentioned DES elements.

The last table, Table 1c, shows the *attributes* for each activity, such as ID, number of resources, and duration.

Each of the spreadsheets modelling elements is described in a separate worksheet. The activities worksheets start with a column that states the actual time in the simulation. This column is a model of the *clock* in the DES modelling elements.

3.3 Types of modelling elements

Each of the three modelling elements – gate, activity, and transaction is modelled with a separate spreadsheet. Each row in the spreadsheet represents the state of the modelling element at a particular moment of time.

3.3.1 Gate

The model of a gate in a worksheet comprises the following four columns:

(A) states the time when the gate is activated;
(B) gives the number of transactions that are currently waiting at the gate;
(C) lists all transactions that wait at the gate; and
(D) lists all transactions which are ready to be propagated to the next activity.

An example of a gate is presented in Figure 1. The first column is the *certain delay* that shows the activation time moment of the gate element.

Figure 1: Example of a gate modelling element: Gate1

	A	B	C	D
1	Time in minutes	Patients waiting	All transactons that arrived	Transactions waiting at 8:00 AM
2	481	25	1:2:3:4:5:6:7:8:9:10:11:12:13:14:15:16:17:18:19:20:21:22:23:24:25	1:2:3:4:5:6:7:8:9:10:11:12:13:14:15:16:17:18:19:20:21:22:23:24:25
3	961	0	1:2:3:4:5:6:7:8:9:10:11:12:13:14:15:16:17:18:19:20:21:22:23:24:25:26	26:27:28:29:30:31:32:33:34:35:36:37:38:39:40:41:42:43:44:45:46:47:48:49:50:51:52:53:54:55
4	1441	0	1:2:3:4:5:6:7:8:9:10:11:12:13:14:15:16:17:18:19:20:21:22:23:24:25:26	56:57:58:59:60:61:62:63:64:65:66:67:68:69:70:71:72:73:74:75

3.3.2 Activity

Figure 2 presents an example of an activity. The worksheet that models an activity is comprised of the following eight columns (A-H):

(A) States the simulation minute of the time which applies to the whole row in the worksheet;

(B) Stores whether a transaction has arrived in the indicated minute; "one" denotes arrival of a transaction, and "zero" denotes that no transaction has arrived;

(C) Shows until when a resource is busy; meaning until which minute;

(D) States the availability of the resource; "one" denotes that the resource is busy, and "zero" denotes that the resource is available;

(E) Gives a cumulative number of transactions that are waiting in a queue for processing;

(F) Indicates whether the transaction defined in row I is finished or not. If it is finished then one is inserted in cell (I, F), which means that the current transaction continues in the next activity. Otherwise, "zero" is written in cell (I,F);

(G) Lists all transactions that are waiting in a queue for processing at the current activity. The transaction that will be the first processed is first in the queue.

(H) The last column lists the transactions that were processed by the current activity. Consequently, these activities are propagated to the next worksheet (activity). The values are given as integer numbers, representing the transaction IDs.

(I) When a gate follows an activity, we need to list all transactions that have finished with the activity.

Generally, an activity input is a transaction that has finished its execution from the preceding activity. In continuation we describe, which may have multiple sources of incoming transactions and a decision activates.

Figure 2: Spreadsheet of "Confirm surgery date" activity

	A	B	C	D	E	F	G	H
	Time in minutes	Arrival of transaction	Resource busy until	Resource Busy (1 = busy; 0 = free)	Transactions waiting for execution	Finished transaction	List of transactions that started	Transaction ID ended
1	1	1	11	1	0	0	1	0
2	2	1	11	1	1	0	1:2	0
3	3	1	11	1	2	0	1:2:3	0
4	4	1	11	1	3	0	1:2:3:4	0
5	5	1	11	1	4	0	1:2:3:4:5	0
6	6	1	11	1	5	0	1:2:3:4:5:6	0
7	7	1	11	1	6	0	1:2:3:4:5:6:7	0
8	8	1	11	1	7	0	1:2:3:4:5:6:7:8	0
9	9	1	11	1	8	0	1:2:3:4:5:6:7:8:9	0
10	10	1	11	1	9	0	1:2:3:4:5:6:7:8:9:10	0
11	11	1	11	0	10	1	2:3:4:5:6:7:8:9:10:11	1

3.3.3 Different types of activities

The model distinguishes two main modifications of an activity: a decision activity and activity with multiple sources of incoming transactions. In the latter case, the model defines two modifications of the activity depending on the type of the modelling element that generates the incoming transaction to the activity:

1. one transaction is generated from a preceding activity and another is generated from a preceding decision activity, or
2. one transaction is generated from a preceding activity and another is generated from a preceding gate.

3.3.4 Decision activity

A decision activity is used to check whether a certain condition is fulfilled and on the basis of the outcome, decides for further actions. The decision activity usually has two paths. One path is indicated by "YES" and means that the condition is fulfilled. "NO" denotes the other path.

Figure 3 is an example of a model of a decision activity. In addition to the columns defined for an activity, it comprises the following columns:

(I) generates a random number using the "RAND()" function;
(J) gives a decision YES or NO based on the obtained number in (A);
(L) auxiliary column that keeps record of the transactions that ended with a decision "YES", and
(M) auxiliary column that keeps record of the transactions that ended with a decision "NO".

Figure 3 Example of a decision activity 18

	A	B	C	D	E	F	G	H	I	J	L	M
Row No.	Time in minutes	Arrival of transaction	Resource busy until	Resource Busy (1 = busy; 0 = free)	Transactions waiting for execution	Finished transaction	List of transactions that started	TransactionID ended	Decision	20% YES; 80% NO	Auxilary column YES; follow the gate INT	Auxilary column NO; follow the activity 19
163	1316	1	1326	0	0	0	49	0	0,56087	NO		
314	1448	1	1458	0	0	0	74	0	0,88199	YES	49	
446	1552	1	1562	0	0	0	99	0	0,010479	NO	49	
550	1722	1	1732	0	0	0	124	0	0,605636	NO	49	
720	1878	1	1888	0	0	0	49	0	0,661489	NO	49	
876	2020	1	2030	0	0	0	149	0	0,233557	NO	49	
1176	2178	1	2188	0	0	0	174	0	0,005266	NO	49	
1305	2307	1	2317	0	0	0	199	0	0,796961	NO	49	
1422	2424	1	2434	0	0	0	224	0	0,51937	NO	49	
1577	2579	1	2589	0	0	0	249	0	0,974492	YES	49	
1748	2750	1	2760	0	0	0	274	0	0,023592	NO	49	
1926	2928	1	2938	0	0	0	299	0	0,518122	NO	49:274	

3.3.5 Activity with incoming transactions from two preceding activities

Figure 4 represents the activity 17, which is a successor of two activities: the decision activity 15 and the activity 16. The arrival of transactions is simulated with three columns: columns B and C show arrival of the transaction from activities 15 and 16, respectively, and column D shows if the transaction has arrived from either activity 15 or 16.

The time for execution of this activity is uncertain, in the sense that it may take any value between 90–180 minutes (see activity 17 in Table 1c). Consequently, to simulate the time execution, a random number is generated in the interval (0–1) in each minute, which is scaled to the required interval.

Figure 4 Example of an activity 17 with two inputs: from activity 15 and 16

	A	B	C	D	E	F	H	I	J	M	N
1	Time in minutes	Arrival of transaction from 15	Arrival of transaction from 16	Arrival of transaction from 15 or 16	Resource busy until	Resource Busy (1 = busy; 0 = free)	Transactions waiting for execution	Finished transaction	Uniform distribution	List of transactions that started	Transaction ID ended
2	4804	0	0	0	4888	1	1	0	180	24:21	0
3	4805	0	0	0	4888	1	1	0	109	24:21	0
4	4806	0	0	0	4888	1	1	0	96	24:21	0
5	4807	0	0	0	4888	1	1	0	169	24:21	0
6	4808	0	0	0	4888	1	1	0	95	24:21	0
7	4809	1	0	1	4888	1	2	0	170	24:21:18	0
8	4810	0	0	0	4888	1	2	0	155	24:21:18	0
9	4811	0	0	0	4888	1	2	0	138	24:21:18	0
10	4812	0	0	0	4888	1	2	0	159	24:21:18	0

3.3.6 Activity with incoming transactions from preceding activity and a gate

The second modification of an activity with multiple incoming transactions discusses the case when the incoming transactions come from: a preceding gate and a preceding activity. Figure 5 shows Activity 14 as an example of this case. Columns H and I are the newly introduced ones in the model of such an activity. Column H checks whether it is time to remove transactions from the Gate *IntG*.

Figure 5 Example of an activity 14 with two inputs: from activity ID 13 and the gate activity ID IntG

A	B	C	D	E	F	H	I	J	M	N
Time in minutes	Arrival of transaction	Resource busy until	Resource Busy (1 = busy; 0 = free)	Transactions waiting for execution	Finished transaction	Check if time to withdrow transactions from IntG	Transaction has arrived from IntG	Transaction from 13 or from IntG	List of transactions that started	TransactionID ended
726	1	736	1	0	0	0	0	1	1	0
960	0	938	0	0	0	1	0	0		0
1440	0	1352	0	1	0	1	3	1	3	0
1920	0	1801	0	0	0	1	0	0		0
2400	0	2354	0	0	0	1	0	0		0
2880	0	2799	0	3	0	1	10:11:12	1	10:11:12	0
3360	0	3345	0	2	0	1	13:14	1	13:14	0
3840	0	3810	0	2	0	1	12:16	1	12:16	0
4320	0	4296	0	1	0	1	17	1	17	0

3.3 Outputs of the model

The output of the simulation model provides calculations and graphical representation of distributions of the following parameters:

1. Minimum cycle time
2. Average cycle time
3. Maximum cycle time
4. Average work time
5. Average wait time

Work time is the simulated amount of time that activities are processed. It is directly calculated as a sum of the duration times of the activities.

Wait time is the simulated amount of time that transactions are waiting. The waiting may occur due to different reasons such as waiting for a resource or because of a gate.

Cycle time is the simulated amount of time that a transaction spends within the process starting with the first activity and ending with the last one. A cycle time may differ from one activity to another due to the different paths in the process and to the different activity durations calculated from their distributions. The cycle time is a sum of work time and wait time.

The following equation connects the last three parameters:

$$\text{Cycle time} = \text{Work time} + \text{Wait time}.$$

Maximum (minimum) Cycle time refers to the highest (lowest) accumulated time for the transactions at hand, at any point in time.

The term *average* refers to the total time that all transactions spent in each activity divided by the number of completed transactions for each activity.

4 Conclusion

Simulation is a powerful tool for process analysis because there are currently more than 100 different software tools developed for simulation. The main difficulty that practitioners of BP face is choosing the right simulation tool for a given business process because of two reasons: the prize of the offered tools and the limited literature on comparison of these tools. In this paper, we propose an outline for using spreadsheets for discrete event simulation of business process models. Spreadsheets are widely accessible to practically all computer users and also require little programming knowledge; however, they require a suitable initial time to build the model. We propose for the usage of spreadsheets in cases with a medium complexity of processes and financial limitations.

5 Acknowledgements

This work is supported by Creative Core FISNM-3330–13-500033 'Simulations' project funded by The European Regional Development Fund of the European Union.

References

Banks, Jerry / John S. Careson / Nelson Barry L. / Nicol David M.: *Discrete-Event System Simulation*. New Jersey: Prentice-Hall, Inc. 2001.

Croll, Grenville J.: "Spreadsheets and Long Term Corporate Survival." *Proceedings of European Spreadsheet Risks Interest Group* 2012, pp. 77–96.

David Hartvigsen Mendoza: *SimQuick Process Simulation with Excel*. College of Business Administration, University of Notre Dame. Prentice Hall, Upper Saddle River, NJ 07458. 2001.

Derby, Norman / Stanislaw Olbert: "Cylindrical magnets and ideal solenoids." *American Journal of Physics* 78 (3), 2010, pp. 229–235.

El-Hajj, Ali: " Functional Simulation Using Spreadsheets." *Simulation* 73 (2), 1999, pp. 80–90.

El-Hajj, Ali / Karim Y. Kabalan / Maher N. Mneimneh / Feras Karablieh: "Microprocessor Simulation and Program Assembling Using Spreadsheets." *Simulation* 75 (2), 2000, pp. 82–90.

Emmett, Hilary L. / Lawrence I. Goldman: "Identification of Logical Errors through Monte-Carlo Simulation." *Proc. European Spreadsheet Risks Int. Grp. (EuSpRIG)*. 2004.

Gerdes, Lars/ Ulrich, Busch / Sven, Pecoraro: "A statistical approach to quantification of genetically modified organisms (GMO) using frequency distributions." *BMC Bioinformatics* 15 (1), 2014, pp. 1–12.

Hartvigsen, David: *SimQuick Process Simulation with Excel*. Mendoza College of Business Administration, University of Notre Dame, Prentice Hall, Upper Saddle River, NJ 07458. 2001.

2011. "http://www.igrafx.com/company/news/article/igrafx-spin-off-establishes-leading-independent-provider-of-business-process-analysis-solutions." Accessed January 20, 2015.

2013. *iGrafx Help system*. iGrafx LLC.

Kadjo, Akinde / Purnendu K. Dasgupta: "Tutorial: simulating chromatography with Microsoft Excel Macros." *Anal Chim Acta*. 773, 2013, pp. 1–8.

Karim Y. Kabalan / Ali El-Hajj / Hassan Diab / Souhier Fakhreddine: "Analog Computer Simulation Using Spreadsheets." *Simulation* 68 (2), 1997, pp. 101–106.

Klein, Michael G. / Gilles, Reinhardt: "Emergency Department Patient Flow Simulations Using Spreadsheets." *Simulation in Healthcare* 7 (1), 2012, pp. 40–47.

Laguna, Manuel / Marklund, Johan: *Business Process Modeling, Simulation, and Design*. New Jersey: Pearson Education, Inc. 2005

Leemis, Lawrence / Steve Park: *Discrete-event simulation: a first course*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc. 2005.

Macho, Siegfried: "Cognitive modeling with spreadsheets." *Behav Res Methods Instrum Comput* 34 (1), 2002, pp. 19–36.

Meineke, Ingolf / Jürgen Brockmöller: "Simulation of complex pharmacokinetic models in Microsoft EXCEL." *Computer Methods and Programs in Biomedicine* 88 (3), 2007, pp. 239–245.

Mileva Boshkoska, Biljana / Damij, Talib / Jelenc, Franc / Damij, Nadja: "Abdominal Surgery Process Modeling Framework for Simulation Using Spreadsheets." *Computer Methods and Programs in Biomedicine*. 2015 (submitted).

Nakatumba, Joyce / Rozinat, Anne / Russell, Nick: "Business Process Simulation: How to get it right?" In *International Handbook on Business Process Management*. Springer-Verlag. 2008.

Papa, Gregor / Peter Mrak: "Optimization of cooling appliance control parameters." *2nd International Conference on Engineering Optimization*. Lisbon, Portugal. 2010.

Saven, Ruth Sara / Olhager, Jan: "Integration of Product, Process and Functional Orientations: Principles and a Case Study." *Preprints of the International Conference on Advanced Production Management Systems, APMS, IFIP*. The Netherlands. 2002.

SN, Wu: "Simulations of the cardiac action potential based on the Hodgkin-Huxley kinetics." *Chinese Journal of Physiology* 47 (1), 2004, pp. 15–22.

Ward, David W. / Keith A. Nelson: "Finite Difference Time Domain (FDTD) Simulations of Electromagnetic Wave Propagation Using a Spreadsheet." *Computer Applications in Engineering Education* 13 (3), 2005, pp. 213–221.

Yusuf, Jafry / Fredrika, Sidoroff / Roger, Chi: "A Computational Framework for the Near Elimination of Spreadsheet Risk." *Proc. European Spreadsheet Risks Int. Grp. (EuSpRIG)*, 2006, pp. 85–89.