Business Process Simulation under Deep Uncertainty

Case study at ING Arrears Management

Master thesis – Public version
December 2013

Tim Leendert Markensteijn
Student number: 1310798
Systems Engineering, Policy Analysis and Management
Faculty of Technology, Policy and Management
Delft University of Technology

Graduation Committee

Chair: Prof.dr.ir. A. Verbraeck	- TU Delft, faculty of TPM, section systems engineering
First supervisor: Dr. M.A. Oey	- TU Delft, faculty of TPM, section systems engineering
Second supervisor: Dr. E. Pruyt	- TU Delft, faculty of TPM, section policy analysis
External supervisor: W.P.J. Kuling	- ING, COO, Business Change/Blackbelts
External supervisor: J. Meester	- ING, COO, Business Change/Blackbelts

Table of contents

Abbreviations

BPM	- Business Process Management
BPS	- Business Process Simulation
CFA	- Call For Action
DSISR	- Design Science Information Systems Research
DES	- Discrete Event Simulation
EMA	- Exploratory Modeling and Analysis
FTE	- Full Time Equivalent
ING	- Internationale Nederlanden Groep
KDE	- Kernel Density Estimate
KPI	- Key Performance Indicator
LHS	- Latin Hypercube Sampling
MCS	- Monte Carlo Sampling
PRIM	- Patient Rule Induction Method
SEPAM	- Systems Engineering, Policy Analysis and Management

Acknowledgements

This thesis is written in light of the final examination for acquiring a MSc. Degree in Systems Engineering, Policy Analysis and Management. The presented research has been conducted during a graduate internship at ING Business Change / Blackbelts under academic guidance from the respective Systems Engineering and Policy Analysis sections of the Faculty of Technology, Policy and Management at the Delft University of Technology.

First, I would like to thank Prof.dr.ir. Alexander Verbraeck for supervising this graduation project. Next, I would like to thank Dr. Michel Oey for his academic guidance throughout the project and helping me structure the research. Dr. Erik Pruyt for comments and discussion on the application of Exploratory Modeling and Analysis. Dr.ir. Jan Kwakkel on is help in the use of data mining techniques in the EMA workbench as well as his patience in explaining the technical constructs of EMA. Wendell Kuling for supervising on behalf of ING and his dedication to make this research an academic success as well as a success for ING. Jaap Meester for his daily guidance throughout the project, countless discussions and reflections on the use of simulation models at ING Arrears Management and ING Business Change / Blackbelts. Jessica Sun, Rene v/d Guchte and Marie Silvie Verrier for their critical reflection and comments on work presented at ING Arrears Management.

Furthermore, I would like to thank Rienk Bijlsma and Gerbert Heijkoop for their technical support on the use of Arena and Scenario Navigator far beyond the regular support.

Pim Bellinga for his help in installing the EMA workbench and comments on PRIM analyses, Jolanda Koot and Cecile Crobach for their fast and valuable reviews.

Finally, I would like to express my gratitude towards my parents, Jacomien Markensteijn-Sandifort and Leo Markensteijn for their unconditional support during my entire student career in Delft.

Tim Markensteijn

Summary

Organizations are in ever changing environments which results in the need for constant adaptation of business processes and structures. Continuous business process improvements can result in cost savings as well as higher efficiency and effectiveness. In some cases business process improvements can be realized through experience and competent management. However, in more complex processes, decision makers may require some form of decision support. A popular decision support method is business process simulation (BPS) One of the most commonly used applications of BPS is Discrete Event Simulation (DES). This application is also used at ING for business process management purposes. DES can be a very powerful method in case much data is available in the target system and the process is transparent. However, if a system contains deep uncertainties, a different approach is required. Deep uncertainty exists in business processes in case process analysts and other stakeholders do not know or cannot agree upon the structure of a process, the value of key variables in a process, and the valuation of desired outcomes. A possible approach in dealing with deep uncertainty is Exploratory Modeling and Analysis (EMA). EMA can be used to explore possible futures based on simulation models.

Existing methods for dealing with uncertainty in discrete event simulations are largely limited to variations in input variables. Hence, it seems undesirable to use DES in highly uncertain environments. However, based on promising applications of EMA in other modeling fields (system dynamics and agent based modeling), the questions arises whether or not it is possible and if so, desirable to apply an EMA approach on DES studies. So far, no attempts have been made to apply an EMA approach on DES studies, resulting in the central research question for this thesis:

How can an Exploratory Modeling and Analysis (EMA) approach be applied on Discrete Event Simulation (DES) in order to help decision makers design business processes and develop adaptive polices under uncertainty?

To answer this research question an approach is proposed based on traditional DES modeling from an EMA point of view. This approach is tested in a case study at ING Arrears Management where there is a need for decision support during the development of new processes in an uncertain environment. Hence, the main objective of this thesis is to experiment with applying EMA on DES in an uncertain business process environment and to elicit the basic methodological principles for doing so. Uncertainties are identified at ING Arrears Management, aggregate simulation models are used to produce large databases with thousands of scenarios depicting a solution space full of plausible future scenarios in terms of business process performance at ING Arrears Management. This solution space is explored through an EMA methodology called scenario discovery. In scenario discovery, the Patient Rule Induction Method (PRIM) is applied to find danger zones in the solution space. PRIM is essentially a bump hunting algorithm that identifies areas in the solution space that contain a high density of cases of interest. These high density areas are interpreted as danger zones that could jeopardize the achievement of business objectives at ING Arrears Management.

The application of the proposed approach towards applying EMA on DES resulted in the identification of several danger zones that form a starting point for the development of adaptive policies at ING Arrears Management for the purpose of avoiding the identified danger zones. Furthermore, bottlenecks were identified as well potential capacity issues in various sub-processes. However, numerous potentially dangerous scenarios remain unexplained through the application of PRIM analysis. Therefore, based on the case study, it can be concluded ING Arrears Management was partly helped in designing efficient new business processes in an uncertain business process environment.

Even though the case study at ING Arrears Management was not completely solved through the application of the proposed approach, it can be concluded that the approach shows great potential compared to a traditional DES approach. Not only in the appropriate use of tools and techniques for EMA, but also in the application of an iterative approach in practice that resulted in helping decision makers at ING Arrears Management in identifying gaps, risks and weak spots in their proposed business processes. Considering the added value of an application of EMA at ING Arrears Management, it can be concluded that a partial proof of concept for the proposed approach has been acquired. However, the (partial) proof of concept is based on a single case study. For this reason, extrapolation of conclusions towards business processes under uncertainty in general must be done with great care.

Considering the proof of (partial) proof of concept provided in this research is only valid for the case study presented in this report, the most important recommendation is to apply an EMA approach on DES on other cases where business processes under (deep) uncertainty can be identified. When choosing case studies for future research, it is recommended to select case studies in which an attempt can be made to study identified methodological obstacles including probabilistic information in DES models, application of other data mining and machine learning techniques, and further development of integrated technical tools for applying EMA on DES.

1. Introduction

The aim of this introductory chapter is to introduce the background, events and (gaps in) literature leading to the study presented in this thesis report. First, a case at the ING Arrears Management department will be introduced. Second, Discrete Event Simulation will be introduced as a widely recognized form of Business Process Simulation both in literate and within ING. Third, Exploratory Modeling and Analysis will be introduced as a potential approach towards dealing with deep uncertainty in Business Processes at ING Arrears Management. Fourth, the research methodology that serves as a common thread in this report will be introduced. Last, a brief overview of the thesis outline will be given.

1.1. Introduction ING Arrears Management department

The financial crisis which peaked in 2008 resulted in a great recession. One of the key organizations that is striving to prevent such a crisis from occurring in the future is the Basel Committee on Banking Supervision (BCBS) which aims "to strengthen the regulation, supervision and practices of banks worldwide with the purpose of enhancing financial stability." (BCBS, 2013). As part of its ongoing efforts to battle the effects of the financial crisis, the BSBS achieved the Basel III accord (Blundell-Wignall and Atkinson, 2010). In this Basel II accord, the BCBS proposed several measures "to strengthen global capital and liquidity regulations with the goal of promoting a more resilient banking sector" (BCBS, 2009). As a result of the Basel III accord, banks have to make provisions in case mortgage customers are no longer able to afford their mortgage. Hence, ING will have to adapt its policies (i.e. make provisions) to meet the Basel III accords requirements.

[redacted]

The next paragraph will shortly describe the scope of the new processes at ING Arrears Management regarding the research presented in this thesis. Next, Discrete Event Simulation (DES) will be introduced as a method for decision support through Business Process Simulation (BPS) and Exploratory Modeling and Analysis (EMA) will be introduced as a methodology towards accounting for deep uncertainty that may be present at ING Arrears Management.

1.2. Research scope at ING Arrears Management

Currently, the ING Arrears Management department is grossly divided in two sections: Maintenance and parting. The aim of maintenance is to help customers through payment arrangement and other means for the purpose of reducing and eventually completely ending their mortgage arrear. In case maintaining the customer is not possible, the parting section of the ING Arrears Management department will be responsible for parting with the customer. The focal point of this research will be on maintenance. More specifically, the scope can be summarized to include all ING customers with a mortgage arrear (identified through a weekly query on customer payment data) until the customer is either financially healthy or the decision is made by ING to part with the customer. As soon as this decision is made, the parting process starts which is considered out of scope. Furthermore, support (legal, business support) and management activities are considered out of scope.

1.3. Discrete Event Simulation and its shortcomings under uncertainty

Organizations are in ever changing environments which results in the need for constant adaptation of business processes and structures (de Vreede et al., 2003), as is the case at ING Arrears Management. Continuous business process improvements can result in cost savings as well as higher efficiency and effectiveness. The field of improving business processes is commonly known as business process management (BPM) (Jansen-Vullers and Netjes, 2006). In some cases business process improvements can be realized through experience and competent management. However, in more complex processes, decision makers may require some form of decision support. A popular decision support method is business process simulation (BPS) (Jansen-Vullers and Netjes, 2006).

There are many different types of simulation applications in business processes. One of the most popular and commonly used applications is Discrete Event Simulation (DES) (Jahangirian et al., 2010). DES can be very useful in case mathematical modeling is not possible or impractical from an analytical point of view, and in case experimenting in the real system is very expensive, dangerous, time consuming or even impossible (Boersma and Hoenderkamp, 1981; Neelamkavil, 1987; Shannon, 1975). DES can be a powerful method in optimizing business processes. ING uses DES models regularly as part of process improvement projects.

However, some important requirements must be met in order for DES to be applied successfully. Firstly, there must be consensus on what the actual system, and consequently the simulation model, looks like. Secondly, reliable data is required from the real world system. Last, consensus is required on the desired outcomes of business process improvements. Ideally, the above mentioned requirements are met. In case the process improvement concerns alterations of existing business processes, at least the first two requirements should be met. If the target business process does not exist yet, it is very difficult to acquire accurate data and to be sure about the optimal process layout. Hence, parametric and structural uncertainties are expected to be present in a non-existent system. Parametric uncertainty can be defined as uncertainty within the numerical value of a specific variable. Structural uncertainty refers to uncertainty in how variables relate to one another, i.e. there is uncertainty in the systems structure.

A way of dealing with these uncertainties is to estimate as best as possible what the required data could be. Other methods have been described in literature related to various modeling and simulation applications. These methods include stochastic input variation, sensitivity analysis, parametric programming, and robust optimization (Li and Ierapetritou, 2008). Stochastic input variation is a simple method where deterministic input variables are replaced with stochastic distributions, in case of DES this means using discrete probability functions in input variables. Sensitivity analysis can be used to determine what the effect is of variations in input variables on output variables of interest, i.e. sensitivity analysis is used to determine how sensitive key performance indicators are to changes in exogenous variables (Boersma and Hoenderkamp, 1981; Li and Ierapetritou, 2008). Parametric programming can be used to account for stochastic, parameter, and model uncertainties as a result of discrete probability functions for input variables (Zouaoui and Wilson, 2003). Robust optimization strives to find solutions to modeling problems where the output is robust to the uncertainty in input data. A solution is considered to be solution robust if it remains "close" to optimal for all scenarios of the input data, and a solution is considered to be model robust if it remains "almost" feasible for all scenarios based on input data (Mulvey et al., 1995). To a certain level of uncertainty the methods mentioned above are very useful in accounting for uncertainties in a target system. However, in case deep uncertainty is present in a target system, the methods described above are not sufficient. According to Lempert et al. (2003, p. 3), systems are considered to contain deep uncertainties in case "analysts do not know, or the parties to a decision cannot agree on:

(1) the appropriate conceptual models that describe the relationships among the key driving forces that will shape the long-term future,
(2) The probability distributions used to represent uncertainty about key variables and parameters in the mathematical representations of these conceptual models, and/or
(3) How to value the desirability of alternative outcomes."

In case the level of uncertainty is limited to the second point mentioned above, stochastic distributions, parametric programming, or sensitivity analysis will be sufficient to deal with uncertainties. However, in case a system contains deep uncertainties, a different approach is required. This approach will be elaborated upon in the next paragraph.

1.4. Exploratory Modeling and Analysis to deal with deep uncertainty

According to Bankes (1993), a distinction can be made in consolidative modeling and exploratory modeling. Consolidative modeling is the construction of a model based on known facts and then using this model as a surrogate for the real system. Traditional discrete event simulation modeling can be perceived as a consolidative modeling approach. "Exploratory Modeling is using computational experiments to assist in reasoning about systems where there is significant uncertainty … and to explore the implications of varying assumptions and hypotheses" (Bankes, 1993, p. 435). Hence, if significant uncertainty is present in business processes, exploratory modeling and analysis is a promising approach in recognizing uncertainty and dealing with it appropriately to ensure more robust decision support and a solid foundation for adaptive policy making. In essence, the most important difference between consolidative modeling and exploratory modeling can be found in the objective of the respective modeling studies. In most cases of consolidative modeling studies, the objective is to predict system behavior for the purpose of optimization and/or to determine the best possible alternative from a set of proposed alternative policies or strategies. In short, the objective is to predict future behavior (Shannon, 1975). However, predictions are only possible if sufficient information is available to make predictions (Bankes, 1993). Exploratory modeling is not aimed at prediction. Rather it is meant to explore possible future outcomes. EMA can help organizations in changing environments to prepare for future challenges by guiding their adjustment capabilities and adaptability (Kwakkel and Pruyt, 2013). Exploring possible futures can be achieved by clearly mapping (using visualizations) the plausible solution space and to identify areas (i.e. KPI values as a result of combinations of input variable values) in this solution space that lead to a certain undesired behavior. In short, the objective is to find undesired outcomes. Once these undesired outcomes, (i.e. pitfalls in the solution space) have been identified, adaptive policies can be developed for the purpose of preventing undesired outcomes. The adaptive policies can vary from slightly adjusting measures to make sure KPI values stay on the desired track to drastic measures in case of potential system breakdowns.

Exploratory Modeling and Analysis has recently been applied for adaptive robust design under deep uncertainty (Hamarat et al., 2013). Several system dynamics studies based on Exploratory Modeling Analysis under deep uncertainty show successful results in developing robust policies (Kwakkel and Pruyt, 2013; Pruyt and Hamarat, 2010; Pruyt and Kwakkel, 2012; Pruyt et al., 2011). A similar approach towards discrete event simulation could be very useful in case deep uncertainty exists in target systems. In summary, applying an EMA approach on DES can be very useful in complex cases where uncertainties are present, data availability is limited and desired outcomes are debated upon. The purpose of applying EMA on DES is to acquire a complete "what if?" overview based on a potentially large number of plausible scenarios. A scenario for the purpose of this research is defined as "a product that describes some possible future state and/or that tells the story about how such a state might come about" (Bishop et al., 2007, p. 8). The complete "what if?" overview of plausible scenarios will form a solid foundation for adaptive policy making. The next paragraph will further introduce the case at ING Arrears Management in light of its deeply uncertain environment.

1.5. EMA to deal with deep uncertainty at ING Arrears Management

Conceptual models of proposed new processes have been defined by ING Arrears Management. However, the final process structure is debated upon and the process does not exist in its final form yet. As the final process does not exist yet, it is impossible to accurately determine its performance simply because there is no process to measure. Despite the lack of available data and impossibility of acquiring this data in an early stage, it is desirable for ING Arrears

Management to acquire insight in the impact the proposed new processes will have on the risk costs and more specifically (for the purpose of this research), the efficiency of the proposed new processes.

[redacted] Looking back at the three criteria for deep uncertainty as formulated by Lempert et al. (2003), introduced in paragraph 1.3, it can be concluded that the target system at ING Arrears Management is deeply uncertain. These conditions call for a different approach that allows for decision support while taking into account the deeply uncertain nature of the environment at ING Arrears Management. Possible future scenarios will be explored for the purpose of identifying focal areas as a starting point for adaptive policies in service of preventing undesired system behavior. The next paragraph will introduce a theoretical research framework that will function as the backbone of this research report.

1.6. Research methodology: Design Science in Information Systems Research

As the objective is to experiment with applying EMA on DES in a business environment for the purpose of solving practical problems at ING Arrears Management, there is a need for a design oriented research methodology. However, considering the scientific aim of this research, a solid theoretical foundation is also required to make sure principles for applying an EMA approach on DES can be elicited. Besides design oriented and well founded in theory, the methodology needs to balance the need for theory development as well as feasible and usable application in practice. Hence, iterations based on interaction between theory and practice is required.

A research methodology that seems to meet the requirements stated above is Design Science in Information Systems Research (DSISR) (Hevner et al., 2004). Hevner et al. (2004) base their research methodology on design science and behavioral science which respectively seek to "extend the boundaries of human and organizational capabilities by creating new and innovative artifacts" and "develop and verify theories that explain or predict human or organizational behavior". The basis for design science in information systems research is not necessarily limited to behavioral science. Kuechler and Vaishnavi (2008), take a slightly broader approach by naming kernel theories besides design theories as a basis for design science in information systems research. Kernel theories can also include behavioral theories, but also natural or sociological theories (Kuechler and Vaishnavi, 2008).

As the main objective of this research is explore a combination of EMA and DES which should help explain organizational behavior for the purpose of improving organizational capabilities, design and behavioral science are good starting points for this research. Furthermore, Hevner et al. (2004) emphasize that information systems research is not a step by step approach, rather it is an iterative process where the development and building of theories and artifacts is assessed through methods for justification and evaluation after which the theories and artifacts in turn can be refined based on the evaluation results. The information systems research can be influenced through people, organizations and technology in the environment (relevance) and foundations and methodology in the knowledge base (rigor) (Hevner et al., 2004). Finally, the information systems research can be applied in the environment and additions can be made to the knowledge base. In this research the environment mostly concerns ING and the knowledge base is related to business process management, DES and EMA. After the publication of the information systems research framework, Hevner (2007) extended the framework by identifying the rigor, relevance and design cycles. The relevance cycle can initiate the design cycle by providing design requirements and evaluation criteria, next the output of the design science research can be tested in the environment (Hevner, 2007). The rigor cycle provides foundations and methodologies as input for the design science research which in turn can result in novel input for the knowledge base (Hevner, 2007). The design cycle emphasizes the iterative nature of the development of theories and artifacts and their evaluations. The resulting design science research framework, adapted from Hevner et al. (2004) and Hevner (2007), is depicted in figure 1.1.

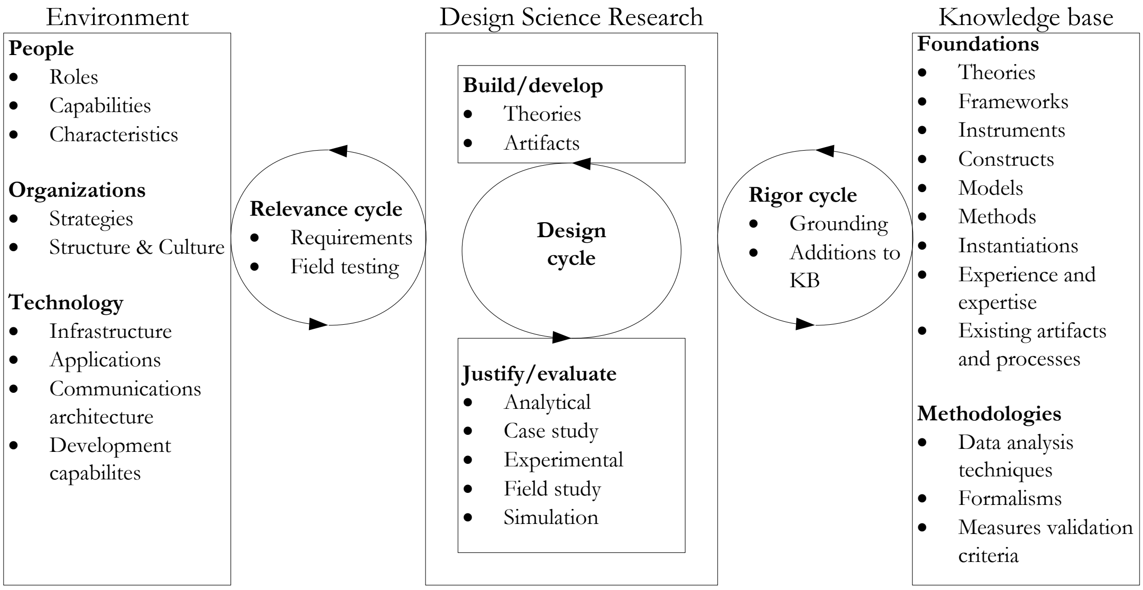

Figure 1.1: design science research framework adapted from Hevner et al. (2004) and Hevner (2007).

The knowledge base for this research project will largely be based on business process simulation and exploratory modeling analysis, including different methodologies required in both subjects. The relevance of this research project will be tested in the ING Business Change/Blackbelts and ING Arrears Management departments (the environment). The actual core of this design science research will be on a modeling approach based on traditional DES from an EMA point of view.

Based on the research methodology presented in this paragraph, the thesis outline is constructed. This outline is discussed below.

1.7. Thesis outline

The thesis will be structured following the design science in information systems research framework. Chapter 1 functions as an introductory chapter including an introduction based on literature studies in the field of EMA and DES. In chapter 2, the research problem and knowledge gap stemming from the literature review will be explained and research questions will be elicited based on this knowledge gap. In chapter 3, an approach to apply EMA on DES is introduced including a discussion on the expected challenges in applying this approach in practice. In chapter 4, the environment including requirements will be described. The environment will involve the ING Arrears Management department as an organization including its available people, capabilities, and technologies. Chapter 5 will focus on the design cycle including the application of the detailed design approach described in chapter 3. Chapter 6 concerns the presentation of model results, analyses, and implications of exploration. In chapter 7, the case study will be evaluated as a means for acquiring a proof of concept. Furthermore, the proposed approach towards applying EMA on DES presented in chapter 3 will be evaluated in chapter 7. Chapter 8 holds the conclusion including summarized answers to the research questions and recommendations for future research. Finally, in chapter 9, there will be a reflection on the process, the case study at ING Arrears Management and on the approach of applying EMA on DES.

2. Research problem

The introduction provides an overview of discrete event simulation and its merits and limitations related to accounting for uncertainties. Furthermore, an introduction is given on EMA as an approach towards dealing with uncertainties is given as well as state of the art examples of EMA applied in practice. In this chapter, the problem will be explored in more detail and the knowledge gap will be identified. The scientific and societal relevance of this research will be explained. Finally, the research objectives and deliverables will be stated.

2.1. Problem exploration and knowledge gap

Considering the current methods for dealing with uncertainty in DES are largely limited to variations in input variables through several methods, it seems the possibility of applying DES in highly uncertain systems is undesirable. However, based on promising applications of EMA in System Dynamics (ESDMA), the questions arises whether or not it is possible and if so, desirable to apply an EMA approach on DES studies. So far, no attempts have been made to apply an EMA approach on DES studies, hence a knowledge gap based on a literature review in the fields of DES and EMA can be formulated:

The extent to which it is possible, practical and desirable to apply an EMA approach on DES for the purpose of designing non-existent real world complex business processes and adaptive policies to ensure robust implementation and operation of these processes.

A few problematic elements can be filtered from the knowledge gap. Firstly, the technical possibility to generate thousands scenarios based on DES models needs to be explored as no integrated tools exist at this point for applying EMA on DES. Secondly, it is questionable whether or not it is practical or even necessary to produce thousands of plausible scenarios as complex DES models can easily take a few minutes or more for one replication to run. Producing thousands of scenarios could mean days of simulation runtime. Thirdly, even if it is necessary and possible to generate a large solution space through a large number of scenarios, it will be challenging to keep track on the causality between input variables and output variables, i.e. KPI's in the solution space. Fourth, as EMA has never been applied on DES, it is unclear at this point to what extent it would lead to a better form of decision support compared to a traditional DES approach. Lastly, there may be methodological obstacles that make it difficult or undesirable to apply an EMA approach on DES studies. These methodological issues will be identified during a case study at ING Arrears Management.

Taking into account the novel nature of applying and EMA approach on DES, the ambition of this research will not be to prove such an approach is successful in all cases where business processes involve deep uncertainties, nor will the aim be to develop an integrated toolset to perform EMA on DES. Rather, this research aims at experimenting with EMA as a methodology applied on a specific case study at ING Arrears Management involving DES as a business process simulation method. While experimenting with this case study, lessons learned and methodological principles can be extracted to reflect upon EMA as a methodology applied on DES. As the purpose of this research is to explore the potential of applying an EMA approach on DES, the practical problem extracted from a case study at ING will be the starting point. The main reason for choosing a practical problem as starting point is that any principles and lessons learned from a scientific point of view will be meaningless unless the practical problem is solved through the application of these principles and lessons in practice.

2.2. Scientific and societal relevance

From a scientific point of view it is valuable to bridge the knowledge gap and solve the problems stated in the previous paragraph as it would open a new range of possibilities for the application of an EMA approach in a different modeling discipline. From a societal point of view, solving these problems could lead to a starting point for the development of practical tools and techniques for designing new business processes under deep uncertainty. The basics for this societal relevance will be tested in a real life case at the ING Arrears Management department. Hence, the results in this study may contribute to identifying a foundation for adaptive policies.

2.3. Research objective and deliverables

The main of objective of this thesis is to experiment with applying EMA on DES in a business process environment and to elicit the basic methodological principles for doing so. In order to achieve this objective, different approaches must be tested while keeping in mind that the practical problem should be solved. Furthermore, an iterative approach is required to synchronize the theoretical background of EMA and DES with the practical business environment at ING. Deliverables related to the research objective include a simulation model of the new business processes at ING Arrears Management for exploratory purposes, and an advice regarding potential danger zones as a result of uncertainties in the new processes. Furthermore, methodological principles and lessons learned for applying DES on EMA in business processes will be documented for future application. Lastly, a part of this thesis will be presented in a research article for academic purposes.

The deliverables are summarized in table 2.1 for three different stakeholders: TU Delft (academic deliverables), ING Blackbelts, and ING Arrears Management.

TU Delft	**ING Blackbelts**	**ING Arrears Management**
Thesis report	Recommendations on how and when to apply an EMA approach on DES in future projects	DES model of new processes
Scientific article		Advice regarding danger zones as a basis for adaptive policies
Case study on EMA approach applied on DES		

Table 2.1: Scientific deliverables and deliverables for ING.

Based on the problem exploration in this chapter, the research questions are formulated in the next paragraph.

2.4. Research question and sub questions

Based on the research problem and objective, research questions can be formulated. A distinction will be made between a scientific research question and a societal/practical research question related to ING.

Research question ING: How can Exploratory Modeling and Analysis, applied on DES models, help ING in designing efficient new business processes which are robust under uncertainty?

ING sub questions:

- What are the most important uncertainties in new processes at the ING Arrears Management department?
- What do simulation models for new processes in the ING Arrears Management department look like?
- Which combination of uncertainties lead to potential danger zones that could jeopardize chosen business objectives at ING Arrears Management?

Scientific research question: How can an Exploratory Modeling and Analysis (EMA) approach be applied on Discrete Event Simulation (DES) in order to help decision makers design business processes and develop adaptive polices under uncertainty?

Scientific sub questions:

- What does a methodology to apply an EMA approach on DES look like?
- When is it desirable to apply an EMA approach on DES?
- What are the main strengths and weaknesses of applying an EMA approach on DES?

The research questions stated above will be answered throughout this research report by following the research method as presented in chapter 1. Sub-questions will be repeated at the beginning of a chapter in case that chapter aims to answer the respective sub-question.

3. DES from an EMA perspective

This chapter aims to answer the first scientific research sub-question: What does a methodology to apply an EMA approach on DES look like?

In order to answer this question, firstly, the use of DES as a separate method and EMA for the purpose of this research will be justified. Afterwards, based on a traditional DES modeling framework, an approach of DES under deep uncertainty from an EMA perspective will be introduced including the potential challenges.

3.1. Justification of DES and EMA as an appropriate research method and methodology

As mentioned above, before proceeding with a detailed research approach based on DES and EMA, the choice for DES as an appropriate BPS method in the case at ING Arrears Management will be justified as well as the choice for EMA in terms of dealing with deep uncertainty.

Justification of DES modeling:

The justification of DES as an appropriate method in this case consists of two steps. First, it needs to be determined whether or not a simulation model is required in this case. Second, the choice for DES as a modeling method needs to be justified. The view of Epstein (2008) is that everyone makes models to a certain extent. However, most models are implicit (mental) models where assumptions are hidden and the system logic is unknown. A model becomes explicit as soon as the assumptions are laid out in detail such that they can be analyzed by others (Epstein, 2008). As multiple stakeholders are involved in this study and the model user is not the same stakeholder as the modeler, it is desirable to use an explicit model for communication purposes.

Explicit models are not necessarily computer simulation models. Computer simulation models can be used in case mathematical modeling is not possible or impractical from an analytical point of view, and in case experimenting in the real system is very expensive, dangerous, time consuming or even impossible (Boersma and Hoenderkamp, 1981; Neelamkavil, 1987; Shannon, 1975). In the case of new processes at ING Arrears Management, it will be very expensive and time consuming to test the new processes on full scale. Even though smaller scale tests with a limited amount of customers are done on the real system, full scale tests with the real system will be too expensive and time consuming. For this reason the choice is made to use a computer simulation model.

As the subject system entails business processes, it seems obvious to select a BPS method. As mentioned in the introduction, DES is one of the most popular and commonly used methods in BPS (Jahangirian et al., 2010). An extensive recognition in literature is not the only reason for choosing DES as a BPS method, there is also a practical reason. The ING Blackbelts department uses DES modeling studies, specifically through the use of the Arena software package, to model existing processes within ING bank for multiple purposes. These purposes, among others, include bottleneck detection, throughput time reduction or design alternative studies. Considering the practical experience of ING with DES and the background of DES as an appropriate tool for BPS in literature, DES is chosen in this study as a simulation modeling method.

Justification of EMA approach:

The choice for EMA is the result of a need to mitigate shortcomings of a traditional DES approach towards business process decision support. Hence, the most important requirement in choosing an appropriate methodology is that this methodology does in fact provide means to deal with deep uncertainty. Therefore, the chosen methodology must be able to account for structural uncertainty, parametric uncertainty, and uncertainty in the desirability of outcomes. While many methods are described in literature that account for either one of these types of uncertainties (Ben-Tal and Nemirovski, 2002; Li and Ierapetritou, 2008; Mulvey et al., 1995; Zouaoui and Wilson, 2003), not many exist that account for all of them. EMA is a methodology that shows great promise in System Dynamics studies (Bryant and Lempert, 2010; Hamarat et al., 2013; Kwakkel and Pruyt, 2013; Lempert et al., 2003; Pruyt and Hamarat, 2010; Pruyt and Kwakkel, 2012; Pruyt et al., 2011). Recently, also applications of EMA on Agent Based Modeling have been studied (Belinga, 2013; Kwakkel and Pruyt, 2013). Considering the promising results of EMA in other modeling fields on the one hand and the lack of other methodologies described in literature that are capable of dealing with all types of uncertainty involved in deep uncertainty on the other hand, EMA is chosen as a methodology to apply on DES in this study.

3.2. Traditional (Discrete Event) Simulation approach

Simulation and modeling have been growing as numerical problem solving techniques since the Second World War, this growths has been accelerated by advances in computer technology from the 60s until now (Neelamkavil, 1987). Shannon (1975, p. 2) defines simulation as "the process of designing a model of a real system and conducting experiments with this model for the purpose either of understanding the behavior of the system or of evaluating various strategies". As mentioned above, simulation can be applied in case more simple analytical or mathematical approaches are not sufficient and experimenting in the real system is impossible or undesirable.

Simulation approaches have been described in many different simulation handbooks (Banks, 1998; Boersma and Hoenderkamp, 1981; Hoover and Perry, 1989; Law and Kelton, 1982; Neelamkavil, 1987; Shannon, 1975). Most handbooks describe simulation modeling studies in steps including a depiction of these stepwise approaches in a flowchart. None of the depicted simulation modeling steps schemes is exactly similar. However, there are some steps that can be found in most simulation handbooks (i.e. found in more than half of the simulation handbooks referred to above). These steps include:

- Problem definition
- Data collection
- Model definition / conceptualization
- Model specification / translation

- Verification and/or validation
- Experimental design
- Execution of experiments
- Analysis of outcomes
- Documentation and implementation of results

Verbraeck and Valentin (2006) summarized the simulation modelling approaches mentioned in various simulation handbooks in a visualization of simulation modeling steps. This stepwise approach towards discrete event simulation modeling studies will be taken as a starting point for the detailed research approach applied in this study. The simulation modeling steps as summarized by Verbraeck and Valentin (2006) are depicted in figure 3.1.

The modeling approach, depicted in figure 3.1, is traditionally executed in a way that resembles the waterfall approach, well known in systems engineering literature (Sage and Armstrong, 2000). The resemblance lies in the notion that once a step is completed, the next step in the model approach is started without returning to a previous step. In some model approaches, there are stage gates (for example in whether or not a model is considered valid) that have to be passed before the next step can be taken (Law and Kelton, 1982; Neelamkavil, 1987; Shannon, 1981). Nevertheless, the general approach is stepwise and iterations are fairly rare in case stage gates are passed.

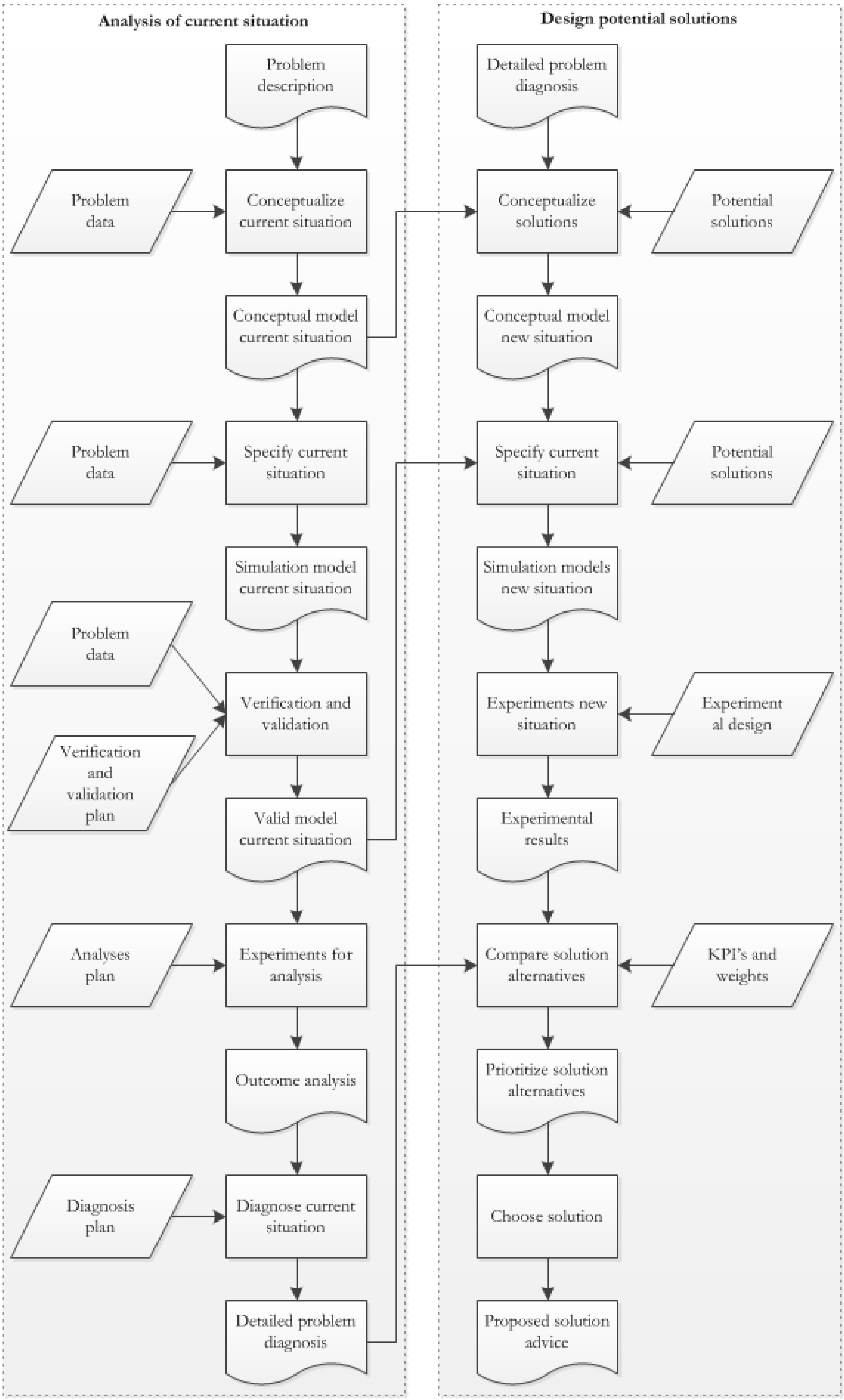

Figure 3.1: DES stepwise approach visualization adopted from Verbraeck and Valentin (2006).

Figure 3.1 shows all modeling steps enumerated above as well as additional solution design steps. The latter will largely be omitted as the purpose of this study is exploration rather than optimal solution design. The main reason for this is that exploration is more likely to result in robust strategies than optimization. Robustness is preferred to optimization as criterion for good strategies in deeply uncertain environments (Lempert et al., 2006). However, despite omitting optimization and solution design steps, most steps in simulation handbooks as well as the framework lay-out will be used as a starting point for the more detailed research approach in the next paragraph.

3.3. Detailed research approach: DES from an EMA point of view

Based on the simulation steps mentioned in the previous paragraph on the one hand and requirements for exploratory modeling and analysis as described in various EMA related sources (Bankes, 1992, 1993; Bishop et al., 2007; Bryant and Lempert, 2010; Hamarat et al., 2013; Kwakkel and Pruyt, 2013; Lempert et al., 2006) on the other hand, a more detailed research approach can be formulated from an EMA point of view. Specific additions related to the use of practical tools are based on ESDMA studies (Hamarat et al., 2013; Kwakkel and Pruyt, 2013; Pruyt and Hamarat, 2010; Pruyt and Kwakkel, 2012; Pruyt et al., 2011), EMA lectures, the EMA workbench developed by Kwakkel (2011), and information on Arena software and Scenario Navigator software provided by Systems Navigator.

The proposed research approach is depicted in figure 3.2 and more detailed descriptions for specific elements in the research approach are provided in table 3.1.

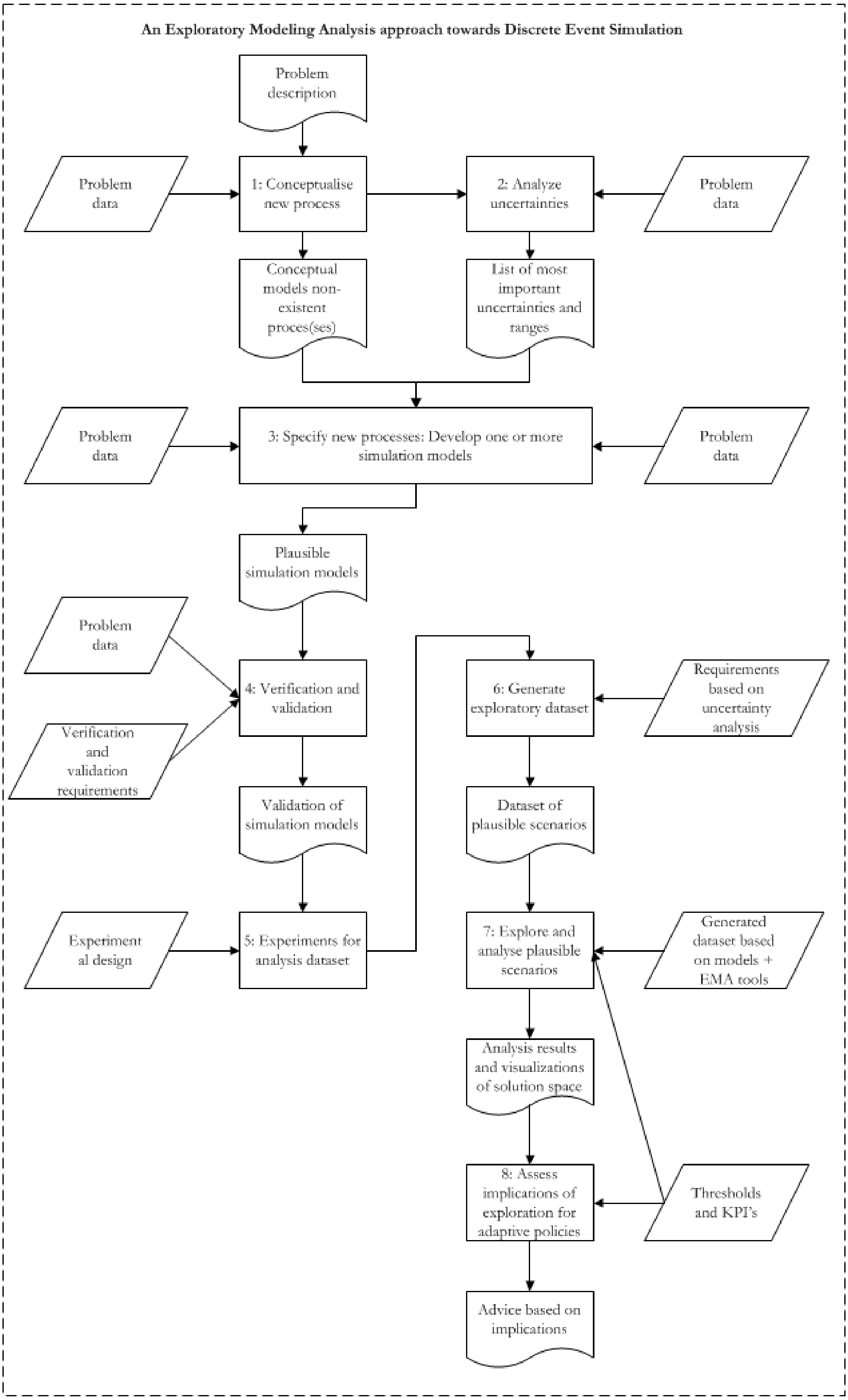

Figure 3.2: Detailed research design visualization adapted from Verbraeck and Valentin (2006).

The steps corresponding to the steps in figure 3.2 are summarized in table 3.1 including the required approach, data, collection method and tools used in a specific step.

Step	**Approach**	**Data**	**Collection method**	**Tools**
1: Conceptualize new processes	Draw flowcharts of new processes	Proposed process lay-out	Interview/works hop	MS Visio
2: Analyze uncertainties	Identify known factors, unknown factors and debates	Focus on missing data	Interview/works hop, database(s).	-
3: Specify new processes: Develop simulation models	Feed conceptual model(s) with known data	Numerical factor values and structural process lay-outs	Interview/works hop, database(s), pilot measurements	Arena
4: Verification & validation	Several verification tests and iterative face validation	Expert opinion	Interview/works hop	Arena
5: Experiments for analysis dataset	Determine number of uncertainties to be tested	Parametric and structural uncertainties	Acquired in 2	-
6: Generate exploratory dataset	Sampling of dataset based on uncertainties	Parametric and structural uncertainties	Acquired in 2	EMA Workbench, Scenario Navigator, Arena, SQL Server, Excel (@Risk add-on)
7: Explore and analyze plausible scenarios	PRIM analysis	KPI's and threshold values	Interview/works hop	EMA workbench
8: Asses implications of exploration	dim plots visualizations, and danger zone advise	Generated dataset (6), KPI's and threshold values (7).	Acquired in step 6 and 7	EMA workbench

Table 3.1: Summary of steps, approaches, required data, collections methods and tools.

Most modeling studies consists of the following phases: Conceptualization, specification, verification/validation, experimental design, results analysis (Banks, 1998; Boersma and Hoenderkamp, 1981; Hoover and Perry, 1989; Law and Kelton, 1982; Neelamkavil, 1987; Shannon, 1975). There are several differences in steps to be taken during phases in an exploratory modeling process and a consolidative modeling process. All steps that are mentioned in figure 3.2 and table 3.1 will be briefly discussed in terms of differences compared to a traditional modeling study and potential challenges from an EMA point of view.

1) Conceptualize new processes

Conceptualization in consolidative modeling is likely to be more detailed as the aim is to build a model that accurately matches reality. In case of exploratory modeling, one or an ensemble of more simple models instead of one large complex model can be built. These models do not accurately reflect reality, however they are all plausible as representations of the real system (Bankes, 1993). Different models may be required as a result of present uncertainties. These uncertainties can find their origin in the lack of available information to construct a conceptual model or disagreement among stakeholders in regard to what the model structure looks like. In both cases it may be desirable to use multiple models for exploration purposes that defer either in aggregation level, model structure or both.

2) Analyze uncertainties

In consolidative modeling, the data collection during the conceptualization phase consists of analyzing objects, factors and other components from the real system as a basis for a conceptual model. In case of EMA, it is likely that the real system does not exist (yet). Therefore, it is important to determine what exactly is known in regard to the proposed system and what is unknown and to what extent is it unknown. I.e. it is important to determine the amount, type, level and location of uncertainty in a target system. This should be done in parallel with the conceptual model because uncertainties may influence the conceptual model lay-out.

3) Specify new processes: Develop simulation models

Specification in consolidative modeling requires a large amount of data which is preferably measured in the real world system. Exploratory modeling may involve guessing ranges of data in case real world data is unavailable and the required variable data is considered uncertain (Bankes, 1993). Hence, even when specifying a single model that accurately reflects reality is not possible, but where relevant information exists, EMA can be a useful approach (Kwakkel and Pruyt, 2013). While trying to identify uncertainties, it is important to involve different experts. By involving multiple experts, a more complete and reliable insight in uncertainties can be acquired. The product of the specification phase in consolidative modeling is a simulation model fed with measured data. The product of a similar phase from an EMA perspective will be a model fed with measured data to the extent it is possible and additionally, it will include value ranges for uncertainties that are either chosen by experts or determined in another way.

4) Verification and validation

Verification is largely similar in consolidative and exploratory modeling. However, validation in exploratory modeling is slightly different as historical data often does not exist. Considering the lack of data in uncertain systems, strong validation (i.e. comparing the simulation model to reality for the purpose of checking whether the simulation model reflects the real system correctly) is impossible (Bankes, 1993). In the case at ING Arrears Management, "the real system" does not exist yet. Hence, validation seems impossible in this case according to the definition of validation stated above. However, a different well known definition of validation in literature allows for other possibilities compared to rendering validation impossible in case of models based on non-existent systems. Schlesinger et al. (1979) define validation as "substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model". Considering that the intended use of a simulation model in this study is exploration rather than prediction, a satisfactory range of accuracy does not necessarily mean correct reflection of the real system is required for validation. Schlesinger et al. (1979) further note that this satisfactory range of accuracy should not be achieved through absolute validity as this is most likely very costly and time consuming. Rather, modelers should focus on acquiring sufficient confidence in the simulation model for its intended purpose (Shannon, 1981). Hence, the question remains how to acquire sufficient confidence in the simulation model for its intended purpose.

Looking at a simplified version of the modeling process (see figure 3.3. adopted from Sargent (1996)) in light of validating simulation models for ING Arrears Management for exploratory purposes, data validation and conceptual model validation are expected to be the most challenging because data is largely not available and the process structure has not been defined accurately yet.

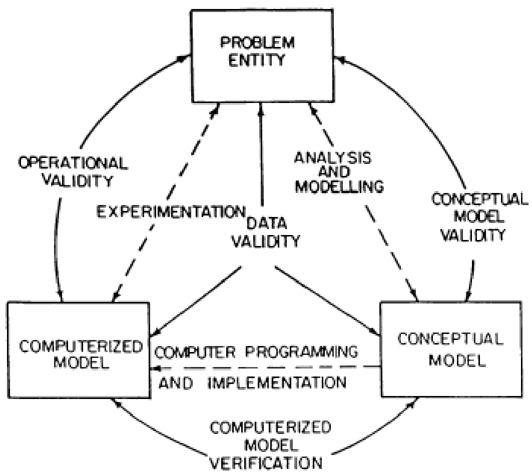

Figure 3.3. Validation and verification in the modeling process, adopted from Sargent (1996)

Considering parametric uncertainty in input variables, data validity is expected to be challenging. Furthermore, conceptual model validity is difficult as there may be no consensus on what the final system at ING Arrears Management will look like. To deal with these challenges in the case at ING Arrears Management, all model iterations will be subject to face validation in both conceptual (and simulation) model changes, input data, and model outcomes. This way, the assumptions of both modeler and users at ING can be made explicit. By making these assumptions explicit, even in cases where strict model validation is impossible, EMA can provide new knowledge (Kwakkel and Pruyt, 2013).

Lastly, it is important to note that in the paradigm of exploratory modeling, the main question should not be on the relative validity of simulation models, but on the most suitable strategy for using these models given their limitations (Bankes, 1993).

5) Experiments for analysis dataset

Experimental design in consolidative modeling is meant to design experiments which produce a limited amount of results for the purpose of optimization or alternatives evaluation. Experimental design in exploratory modeling is aimed at designing experiments which can produce a potentially large dataset to resemble the complete solution space. The size of the solution space depends on the amount of uncertainties present in the system to be explored. In consolidative modeling the experimental outcomes are directly interpretable in contrast to exploratory modeling where further analysis is required. Solution design in consolidative modeling is aimed at optimizing a certain desirable KPI in contrast to exploratory modeling which is used to find potential causes of (un)desirable behavior for the purpose identifying grounds for adaptive policy measures. I.e. the purpose is to find danger zones in the solution space that result in business goals that will likely not be achieved. In order to be able to find these danger zones it is important to make sure causality is taken into account in an early stage while applying an exploratory modeling approach. Establishing causality is crucial because it is impossible to formulate adaptive policies based on a solution space without knowing which combination of uncertain values lead to a specific outcome in the solution space. This makes sense considering the fact that the solution space consists of plausible scenarios and scenarios are not only defined as a possible future state, but also the story of how such a state might come about (Bishop et al., 2007). Hence, without knowing causality, it is impossible to determine which factors could be influenced through adaptive policies. This potential obstacle will be avoided by labeling scenario's in the input database as well as in the output database such that they can be merged into a database that is suitable for analysis.

Even when all scenarios in terms of input and output are labelled and merged, it is important to realize that DES typically involves probabilistic information. Whereas probability theory is an integral part of consolidative modeling and optimization through traditional decision analytic methods (Lempert et al., 2006), inclusion of probabilistic information might contaminate the scenario results regarding causality in an EMA approach (Bryant and Lempert, 2010). I.e. observed scenarios that are assumed to be caused by combinations of uncertain input variables may in fact be caused by the probabilistic nature of a DES model.

6) Generate exploratory database

This is essentially a step that is unique to exploratory modeling compared to consolidative modeling. Results in consolidative models are usually limited to one or several scenarios. However, for the purpose of exploration, a database of scenario results must be generated. The size of the database depends on the amount of uncertainties in the target system, the value ranges of these uncertainties and the chosen method of representing the full solution space. There are several options for design choices in regard to database generation. Three options that will briefly be discussed are:

- Full factorial design
- Monte Carlo Sampling (MCS)
- Latin Hypercube Sampling (LHS)

Full factorial design means that a database will be generated containing all possible scenarios. On first sight this seems like a good option as it would be the best representation of the solution space because a full factorial design actually represents the complete solution space. However, full factorial designs rapidly become impractical as the number of uncertainties increases. For example, 10 uncertainties with 2 possible values would result in a 2^{10} = 1.024 scenario large solution space. Hence, 4 values per uncertainty would make the solution space already larger than 1 million scenarios. Considering the case at ING Arrears Management contains more than 10 uncertainties and value ranges per uncertainty, the full solution space moves towards infinity.

Sampling is a process by which values are randomly drawn from a certain distribution (Palisade Corporation, 2010). Sampling can be done through different methods, MCS and LHS will be discussed briefly (See appendix B for illustrations of MCS and LHS). The main difference between MCS and LHS is that in MHS values are drawn from a distribution independently and in LHS the distribution is divided in equal ranges (the number of ranges is equal to the chosen size of the sample). From these ranges a value is drawn randomly. LHS has the advantage of covering the full range of possible values even in case of outliers. However, this is only an advantage in relatively small samples. As soon as the sample size increases, the processing time for LHS increases and the coverage of MCS also increases as more values are drawn. Therefore, in case of very large samples it may be preferable to use MCS. For the purpose of this research LHS is used to generate uncertainty input value databases, this way exploration of the full uncertainty range (i.e. good representation of the full solution space) is achieved.

7) *Explore and analyze plausible scenarios*

Similar to the previous step, this step is unique to exploratory modeling in relation to consolidative modeling. Taking into account uncertainties and generating a database of plausible futures in the previous step can easily result in an information overload. In order to deal with this potential information overload, machine learning or data mining techniques can be applied (Kwakkel and Pruyt, 2013).

Building on the EMA foundation as described in the introduction (Bankes, 1992, 1993), Bryant and Lempert (2010) developed a new participatory, computer assisted approach to scenario development which they call "scenario discovery". Scenario discovery is intended to aid decision makers in identifying policy or strategy relevant scenarios by interactively applying statistical and data mining algorithms to large databases of simulation model results (Bryant and Lempert, 2010). Bryant and Lempert (2010) developed this new approach to address the issue of problems with traditional discovery in case of systems where a large number of plausible futures can be identified and interests among stakeholders differ. Scenario discovery uses "statistical or data-mining algorithms to find easy-to-interpret, policy-relevant regions in the space of uncertain input parameters to computer simulation models" (Bryant and Lempert, 2010, p. 35). These simulations are run many different times with varying combinations of values for uncertain input variables. A threshold value for KPI's (i.e. outcomes of interest) can be identified to specific cases of interest. "Statistical or data-mining algorithms applied to the resulting multidimensional database then find simple descriptions of the input space that best predict these cases of interest. These regions of input space can then be usefully considered as scenarios for decision analytic applications, and the input parameters used to define the regions become the key driving forces for these scenarios" (Bryant and Lempert, 2010, p. 35).

8) Asses implications of exploration

The final step of the proposed approach in applying EMA on DES is to assess the implications of the exploratory analysis in light of business process performance at ING Arrears Management. The previous step will likely yield an overview of potential danger zones that might threaten the achievement of business process goals. These danger zones should not be perceived as predictions of the future but more as the provision of new information to help make an informed decision. This information can also form the basis for the development of adaptive policies (Hamarat et al., 2013). It is important to note that for this research adaptive policy design and testing is considered out of scope.

Assessing the results of the exploration stemming from different kinds of uncertainties can lead to the identification of alternative strategies, help prioritize research, and verify hypotheses that decision makers may have in regard to the target system (Bankes, 1993). Effective visualization and communication of the results is of crucial importance to a successful application in real world environments (Kwakkel and Pruyt, 2013). Explaining the danger zones as a summary of scenarios may help in effective communication as scenarios describe the future in a way that decision makers find easy to understand (Bryant and Lempert, 2010).

The approach to applying EMA on DES in the steps described above will be followed in this research as part of the design cycle in the overall research methodology, DSISR theoretical framework, introduced in chapter 1. It is important to note that the design cycle implicitly holds iterations, this means the stepwise approach presented in this paragraph will not simply be conducted once, but multiple times. This means analysis of results and implications discussed with decision makers and other stakeholders may form the basis for a new iteration starting with adjustments to the conceptual and/or specified model.

This paragraph provided an answer to the first scientific research sub-question: What does a methodology to apply an EMA approach on DES look like? The approach presented above will be tested in a case study at ING Arrears Management and evaluated based on the experiences gained from this case study on the one hand and other EMA documentation presented in literature on the other hand. The next three chapters are in service of answering the practical research questions. Chapter four will introduce the case study at ING Arrears Management which will be used to acquire a proof of concept for the proposed methodology presented above. In chapter 5, the actual application of DES from an EMA perspective on the case at ING Arrears Management is described. Chapter 5 aims to answer the practical research sub-question: What do simulation models for new processes in the ING Arrears Management department look like? In chapter 6, the analyses regarding based on model results will be presented for the purpose of answering the final practical sub-questions: What are the most important uncertainties in new processes at the ING Arrears Management department? And which combination of uncertainties lead to potential danger zones that could jeopardize chosen business objectives at ING Arrears Management?

Once the application of DES form an EMA perspective on the practical case study has been fully covered in chapter 4 through 6, chapter 7 and 8 will be dedicated to answering the scientific research questions.

4. Case study at ING Arrears Management

In the introductory chapter, the ING Arrears Management department was introduced in terms of the departments' main goal of reducing risk costs through implementing more efficient and effective business processes. In this chapter the environment at the ING arrears management department will be introduced in more detail in the form of a specific case study. This introduction will consist of reasons for change, the new processes in a nutshell, and uncertainties in the processes to be implemented.

4.1. Reasons for change: Introduction to the ING Arrears Management case

4.1.1. Background and context

Recent market developments force ING Domestic Bank to change in order to maintain its position within the Dutch banking field. External market influences can be translated to three major reasons for change within ING [redacted]. Firstly, benefits are under pressure due to a decrease in volumes and margins as a result of the current economic climate. Secondly, customer needs and demands are changing; customers have become more critical and expect more personal service. Thirdly, due to increasing demands and additional regulations introduced by supervising authorities costs are rising. In light of these developments, in 2011, ING decided to take measures which can be summarized in three categories [redacted]: Investments in process improvements, ease of use in banking services and personal and proactive advice, and more responsibility at the employee level. These general measures were taken into account while reorganizing the ING Arrears Management department. The next paragraph will zoom in on the current and future situation within ING Arrears Management.

4.1.2. Current and future situation: New processes at ING Arrears Management

Currently, customers with a mortgage arrear are approached in an accurate, but reactive fashion. One standardized process treats all customers equally. Once a customers has a mortgage arrear, letters are sent notifying the customer of their arrear. Only after a number of weeks, customers are approached actively. Furthermore, the number of instruments to help customers recover is currently limited. Besides a reactive approach, employees are forced to use multiple systems or do manual case management, in few cases technical support is available. It is both in the customers interest as well as in INGs interest to have a more proactive approach [redacted].

In the future situation, ING strives to help a growing amount of customers with financial problems. An important notion is that a substantial part of financial problems can be ascribed to behavioral problems (Jungman et al., 2012). For this reason differentiation of customers in terms of their willingness and ability to solve their financial problems will be key in the new processes [redacted]. Besides solving financial problems, the new approach aims at preventing financial problems through contacting customers as soon as possible once an arrear is identified. This way further increase in financial problems can be prevented.

The new processes at ING Arrears Management can largely be subdivided in three customer phases: 'prevention', 'maintenance', and 'parting' [redacted]. The focus of the case study presented in this research will be on the maintenance phase. During the prevention phase, based on experience and predictive models, customers can be approached proactively to investigate preventive solutions. [redacted] Contact with these customers will largely be attempted through SMS, E-mail or letters. In case the financial problems have not been solved after a certain period, [redacted] customers will be contacted by telephone. [redacted]

Once contact has been established, [redacted], a distinction is made between first and second line employees. The first line employees will attempt first contact and try to make a first payment agreement. In case this agreement is not successful, a more senior, second line employee will contact the respective customer to retry a payment arrangement after a more thorough screening of the customer. In case this second payment arrangement fails and all other measures to help customers solve their financial problems fail as well, the second line employee can decide to direct the customer towards the parting process. Before the parting process starts, a thorough check of the customer file will be performed to ensure nothing can be done to help this customer solve his/her financial problems. [redacted] The entire parting process is considered out of scope for this research. An aggregate depiction of the process in the current situation compared to the future situation is depicted in figure 4.1.

[redacted]

Figure 4.1: Current and new processes at the ING arrears management department, adapted from [redacted]

A more detailed process description including a more elaborate conceptual model will be discussed in chapter 5.

4.2. Uncertainties in new processes at ING mortgage arrears department

The previous paragraph provides a general overview of the new business processes at the ING arrears management department. Currently, these processes have not been implemented and the extent to which the process will operate as intended is uncertain. This uncertainty is present in different parts of the proposed business process and in all categories as described in the introduction: Structural uncertainty, parametric uncertainty and uncertainty in the desirability of outcomes. Furthermore, there are internal and external uncertain factors that play a significant role in the new processes at the ING arrears management department. For example, a largely external factor driven by economic developments is the number of customers with financial problems. An example of an uncertain internal factor is the processing time of customers in new parts of the process or the effectiveness of new measures. The aim of this case study is to identify uncertainties that have an impact on the new proposed processes at the ING arrears management department, to analyze this impact and to identify danger zones as a basis for developing steering measures to avoid undesired behavior in the new processes.

5. Design cycle: Simulation model design

The design cycle is at the heart of design science research and it forms the core of this research as well. The design cycle is continuously fed through the rigor cycle which connects design science research to the knowledge base and through the relevance cycle which connects design science research to the environment at the ING Arrears Management department discussed in chapter 4.

This chapter will aim to answer two sub questions related to the practical research question at ING:

- What are the most important uncertainties in new processes at the ING Arrears Management department?
- What do simulation models for new processes in the ING arrears management department look like?

These sub questions will be answered based on results from the conceptualization, (qualitative) uncertainty analysis, specification, and verification & validation phases as described in the detailed research approach in chapter 3.

Three simulation models were used for the purpose of EMA. The main difference between these simulation models lies in the chosen level of aggregation. In this thesis, only the most detailed model will be covered because all conclusions are based on this model and conclusions stemming from an earlier version of the simulation model can also be reproduced using the more detailed version.

The next paragraph will cover the conceptualization phase which that forms the basis for the simulation model used for exploration.

5.1. Conceptualization of ING Arrears Management business processes

In chapter 4, a brief description of the proposed new business processes at the ING Arrears Management department has been given. This description will be taken as a starting point for the development of a conceptual model. The conceptual model is depicted in figure 5.1.

Customers enter the process once they have a mortgage arrear. This means they have failed to pay their full monthly mortgage fee in a timely manner. [redacted] customers will receive a “Call For Action” (CFA). A CFA consists of an Email in case an Email address is available or a letter in case an Email address is not available and an additional SMS in case a cellphone number is available. Based on this CFA, customers can either repay the remaining mortgage debt (auto cure) or contact the Arrears Management department themselves (inbound calls) to make a payment arrangement [redacted] If after one month ING has not been able to contact a customer and the customer has failed to contact ING, then the customer will move on to the no contact process, this is true for all customer segments.

Outbound calling will be done through "smart calling" which means a customer will be called a maximum of [redacted] times [redacted] In case contact has been established, [redacted]

In case no contact can be established with a customer, administrative employees will do research for the purpose of finding a cellphone number in case it is not available or to verify the cellphone number that is available. Once a cellphone number is recovered, another outbound call (i.e. smart calling) attempt will be done by first line employees. In case contact information cannot be found or verified a "Kadaster check" will be done to verify the address of the customer and a BKR check will be done to check whether the customer has any debts at third parties. Letters will be sent to the customer depending on the outcomes of the respective checks. [redacted] In case contact can be established, a payment arrangement will be made, [redacted]

Once a payment arrangement made by a first line employee has failed (i.e. the customer did not pay sufficient and/or in time) [redacted], a second line employee will take over the customer case. First, the second line employee will review whether or not the first line employee has made a correct payment arrangement. If this is not the case the customer will be transferred back to the first line employee for another attempt to make a good payment arrangement. In case the arrangement was done properly, the second line employee will analyze the financial situation and customer behavior in more detail. Based on this analysis another payment arrangement can be made or other measure can be applied to help the customer recover from its financial problems.

Finally, all customers who have failed to honor their payment arrangements made during a face to face conversation or second line employees will be forwarded to the credit committee. The credit committee will review whether or not the employee has made a good payment arrangement and/or has taken sufficient measures to help the customer in question. In case the arrangement and or/measure is deemed sufficient by the credit committee, the customer will be forwarded to the parting process. If the efforts by second line or face to face employees have been insufficient the credit committee will send the customer file back to a second line employee for another attempt at a payment arrangement or other measures.

[redacted]

Figure 5.1: Conceptual model of new processes at ING Arrears Management

5.2. Qualitative uncertainties assessment

Based on management interviews, expert interviews, an expert data estimation workshop, business management searches (i.e. database queries) and construction of the conceptual model, uncertainties in the proposed new business process at ING Arrears Management are determined in a qualitative manner. The uncertainties used in the simulation model are depicted in table 5.1. The lay-out for this uncertainty table, including the column titles uncertainties, description, type and range or categories has been adapted from Hamarat et al. (2013).

Uncertainties	**Description**	**Min value**	**Max value**
Customer inflow and attributes			
Number of customers	Monthly inflow of customers in Arrears Management	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
First line			
Customer differentiation	Division of customers in ability and willingness categories	[redacted]	[redacted]
Success percentage of payment arrangements	Share of customers that honor their payment arrangement	[redacted]	[redacted]
Correctness percentage of payment arrangements	Share of correctly made payment arrangements	[redacted]	[redacted]
Process time customer call	Process time per (in- or outbound) call	[redacted]	[redacted]
Second line			
Success percentage of payment arrangements	Share of customers that honor their payment arrangement	[redacted]	[redacted]
Correctness percentage of payment arrangements	Share of correctly made payment arrangements	[redacted]	[redacted]
Max number of arrangements (first and second line)	Max attempts to make a good arrangement with a customer	[redacted]	[redacted]
No contact			
Effectiveness of research methods	Share of found cellphone numbers of total searches	[redacted]	[redacted]
Productivity			
Productivity	Number of productive hours per employee per day	[redacted]	[redacted]

Table 5.1: Uncertainties in business processes at ING Arrears Management, used in simulation model.

In addition to the uncertainties presented in table 5.1, the capacities for all employee types are considered uncertain in some experiments. [redacted]. Another reason for considering the capacity of different employee types is uncertain is that bottlenecks in the process may be present due to insufficient capacity. These bottlenecks have an impact on KPIs throughout the simulation model. Hence when a bottleneck is identified, the question arises what the impact on KPI's would be of a certain scenario in case the bottleneck was not present. If in a certain experiment the capacities are considered uncertain it will be mentioned explicitly in the experimental design step.

The next paragraph contains the specification of the conceptual model including the feeding of data to the model as well as the uncertainties and their corresponding range values.

5.3. Specification of the conceptual model

At this point there is an insight in the conceptual model reflecting new business processes at ING mortgage arrears. Furthermore, an overview of expected uncertainties has been given as well as plausible ranges and categories for these uncertainties. In this specification step the other model parameters will be fed with data, formulas and necessary variables to acquire usable output.

In this specification step the conceptual model depicted figure 5.1 will be translated to a DES simulation model in Arena. This means data will be fed to the model as well as formulas and structures to specify the model logic. A brief overview of the model specification including the location and expected impact of uncertainties on model behavior will be introduced in this paragraph. The detailed specified model can be found in Appendix D.

Figure 5.2 depicts a helicopter view of the entire simulation model.

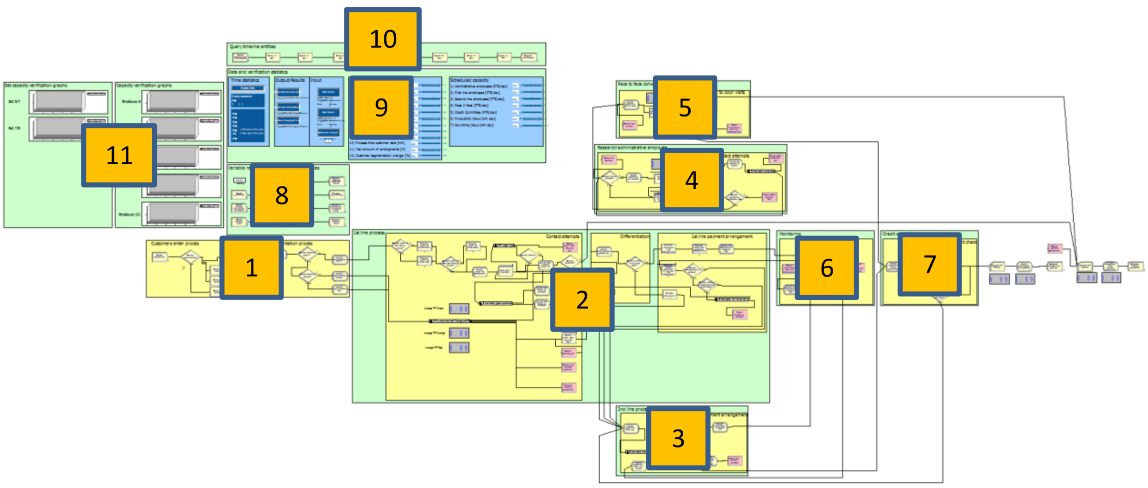

Figure 5.2: Helicopter view of ING Arrears Management processes simulation model

The numbers in figure 5.2 correspond to various sub processes and other model elements:

[redacted]

Not all sub processes enumerated above contain uncertainties. An overview of the uncertainties per sub process is given in table 5.2.

Sub process	**Uncertainties**
1) Arrival and segmentation process	Customer inflow, segmentation green, segmentation orange
2) First line process	Customer differentiation, success rate first line payment arrangements, correctness rate first line payment arrangements, process time customer calls, max number of arrangements
3) Second line process	success rate second line payment arrangements, correctness rate second line payment arrangements, max number of arrangements
4) No contact process	Effectiveness rate
Overall process	Productivity

Table 5.2: Location of uncertain parameters in simulation model

The customer inflow is expected to affect the entire process. If the customer inflow increases, average throughput times [redacted]

[redacted]

[redacted] The success and correctness rates of first line payment arrangements determine the amount of rework that has to be done by first line employees. In case both rates decrease, the amount of necessary rework increases which is expected to lead to higher throughput times for customers in the first line process. In case the success rate is low and the correctness rate high, it could lead to a significant extra inflow of customers in the second line process resulting in higher throughput times. Similar impacts are expected of the uncertainties in the second line process (success rate and correctness rate of second line payment arrangements.

The effectiveness rate of research refers to the fraction of times a research effort results in the finding of a usable 06-number. In case this effectiveness is very high, customers can be contacted by first line employees resulting in extra work for first line employees. In case the effectiveness rate is very low, the inflow at the face to face process will increase resulting in longer throughput times at that process. Besides the uncertain variables described in table 5.1, the model is fed with data for all other parameters. An overview of these variables including a description and their value is appendix E.

At this point the data input and uncertainties have been discussed. Next to these aspects, there are some model logic notions that should be mentioned. In contrast to what the swimming lanes in the conceptual model suggests, it is in fact possible (to a certain extent) for employees to perform tasks that are primarily assigned to other employee types. For this purpose, Resource sets are modeled instead of separate resources. The reason for doing so is that teams in the proposed business processes at Arrears Management will follow the “Super 7” principle in which smaller teams are formed and managed based on output steering. This means the entire team is responsible for the tasks of all employees in the specific team. Therefore, more senior employees can perform tasks of less senior employees in case this is required to achieve a certain goal. For the simulation model this means the following:

- Second line employees (Employee8) can perform first line employee (Employee7) tasks
- First line employees (Employee7) can perform research tasks (Employee6)

In order to ensure that employees primarily focus on their own respective tasks and only pick up tasks of other employees in case there are no available tasks for their primary responsibility, priorities have been assigned. As it is not possible in Arena to assign priorities per resource set, overall priorities have been assigned as follows (lower number means higher priority), the priority ranking can be found in table 5.3:

[redacted]	[redacted]
[redacted]	
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	
[redacted]	[redacted]
[redacted]	
[redacted]	[redacted]

Table 5.3: Priority values of employee tasks

Note that the only exception to the rule that employees primarily focus on their own tasks is receiving inbound calls. Concretely, this means second line employees will answer the phone in case all first line employees are busy. The final part of the specification step will include statistics collection for the purpose of comparing model performance to business goals.

The model has been specified in such a way that it can display relevant information regarding KPI's (Key Performance Indicators). Performance managers at ING have developed a thorough overview of relevant KPI's regarding the new processes at ING Arrears Management. A large portion of these KPI's are effectiveness related KPI's which cannot be measured in a business process simulation model. Rather, these effectiveness measures are used as (uncertain) input variables for the model. Despite the large amount of effectiveness KPI's, some KPI's remain relevant and measurable using the simulation model as specified in this paragraph and appendix D. The KPI's measured in the simulation model are:

[redacted]

The model behavior, results and analyses in terms of the KPI's discussed above will be discussed in chapter 6. However, before doing so, sufficient confidence in the simulation model for its intended purpose must be established. This will be achieved in the verification and validation step discussed in the next paragraph.

5.4. Verification, validation and simulation set-up

Now that the new business processes at ING Arrears Management have been translated to a model through conceptualization and fed with data and logic in the specification step, it is time to define the simulation set-up, verify and validate the model. It is important to note that even though verification and validation are described here as a "step", in practice it is an iterative process and verification and validation efforts are integrated in the life cycle of a simulation study (Balci, 1998). Furthermore, validation in the traditional sense is not possible considering the exploratory nature of this research. Hence, validation will be treated according to the discussion in paragraph 3.3.

5.4.1. Verification

Verification is meant to check whether the simulation model has been code correctly. According to Hoover and Perry (1989, p. 285), at least three questions should be answered:

- Are the input variables within the model processed correctly?
- Are the mathematical formulas and relationships in the model correct?
- Are the statistics and performance indicators calculated correctly?

Answering the first question requires looking at the simulation model itself for a check of normal input variables and to Scenario Navigator and the link between Scenario Navigator and the simulation model for correct translation of the uncertain input variables from Scenario Navigator input tables to the simulation model. In both cases, a structured walkthrough has been conducted based on a structured walkthrough as described by Balci (1998, p. 358). According to Balci (1998), the purpose of structured walkthrough is to discover errors and not to criticize the modeler as this will reduce the quality of error direction. Seven roles should be represented during a structured walkthrough (Balci, 1998, p. 358):

1. Coordinator: Organizes and follows up the walkthrough
2. Presenter: Presents the walkthrough
3. Scribe: Documents the walkthrough meeting(s)
4. Maintenance oracle: Considers long term implications
5. Standards bearer: Concerned with adherence to standards
6. Client representative: Represents the clients' interests
7. Other reviewers: Auditors, project manager, indirect stakeholders

Not all roles mentioned above were represented by different persons during the walkthroughs in this project. The coordinator, presenter and scribe roles were represented by Tim Markensteijn, the maintenance oracle and standards bearer role were included in the client representatives role which in turn was filled in by [redacted] and responsible for the overall process improvement project at ING Arrears Management. Other reviewers include [redacted] The actual contents of the walkthrough follow the description as discussed in the specification paragraph and in appendix D in more detail.

Based on the structured walkthrough several uncertainty range values have been slightly altered and a few uncertain variables were selected to be regular factors and vice versa. Besides a structured walkthrough, the data and verification dashboard depicted in Appendix D was used to verify if the uncertain variable values entered in the Scenario Navigator database were in fact translated correctly as input variables for the simulation model. This is the case. Furthermore, the graphs depicting the number of scheduled employees and the number of busy employees allows for easy verification of correct coding of schedule input variables.

The second question is also mainly answered through the result of a structured walkthrough. By checking all the model parts in detail as described in Appendix D, the model structure can be verified. Animations depicting scheduled and busy employees as well as several counters and statistics animations were used to verify the model structure and relationships within the model.

To answer the third question, two tests were done. The first test is one provided in the Arena software which is a syntax checks. Arena check the model for any errors and reports them if found. None were reported in this model. Besides a syntax check, 1, 100, and 1000 entities were sent through the model to verify whether the exact same amount would come out as came in. This is also the case. When 1 entity is sent through the model, several breakpoints where used to verify that variable and attribute calculations are done correctly. Now that the model has been verified, it is time to describe how sufficient confidence in the simulation model is acquired in the validation step in the next paragraph.

5.4.2. Validation

As discussed in chapter 3, the aim of validation in this research is to acquire sufficient confidence in the simulation model for the purpose of EMA. In the case study at ING Arrears Management this means that traditional validation is done where possible. I.e. in case data from the real world is available for some model parts, it will be used to compare with the data in simulation model. This will be done both for replicative and structural validation.

In case no data is available, iterative face validation is done so assumptions of both modeler and users at ING can be made explicit and can be validated for the purpose of exploration. In practice this could mean modeling different (contrasting) views of experts to make their assumptions explicit for the purpose of exploration. Furthermore, iterative presentation of model results in relation to hypotheses decision makers at ING might have contributed to the iterative validation effort.

5.4.3. Simulation set-up

The simulation set-up consist of four elements: Specification of exogenous data, collection of exogenous data, initialization conditions and run control conditions (Verbraeck and Valentin, 2006). The specification and collection of exogenous data, including the uncertain variables and their ranges has already been discussed in the specification step. Hence, this paragraph will only cover initialization conditions and run control conditions.

In light of the initialization conditions, the first important notion is that the business process in the simulation model is perceived as a non-ending system. [redacted]

[redacted] For this reason, the system is treated as a non-ending system. The definition of input data is twofold; a distinction can be made between factors and uncertainties. The initial values for all non-constant factors defined in the model depend on the seed value for a specific replication. The initial values for uncertainties are imported from an experimental design depending on a specific experiment. Hence, the exact input data value of uncertainties depends on a specific scenario. All replications for a specific scenario have the same values for uncertainties, but different values for other factors. The starting time and date of the simulation are set at Monday 30 September 2013 9.00 A.M. These settings are important for schedule verification purposes.

Run control conditions include the warm-up period, run length, statistics collection specifications and the amount of replications. Considering that we are interested in the performance of the system per month, a run length of one month would be an obvious choice. However, considering that payment arrangements could have an arrangement period of up to [redacted], rework could occur after [redacted]. For this reason, a warm-up period of [redacted] has been chosen as well as a run length of [redacted]. The total run length (i.e. replication length) amounts to [redacted] of which half is warm-up time. Statistics collection starts as soon as the warm-up period ends and the actual run time starts. For the sake of simplicity and technical database limitation, only one statistics value per outcome of interest (per replication and scenario) is written in a database. For details see the data collection module section of Appendix D.

Next, the amount of required replications has been determined. To determine the required amount of replications, an experiment has been done based on the same simulation model which is run in 1.024 scenarios using 1 replication and 10 replications. It is important to note that his experiment has been conducted on an earlier, less detailed version of the simulation model. The assumption is made that the conclusions are valid for the more detailed simulation model. First the results have been compared qualitatively by comparing PRIM boxes of both experiments. These results are shown in detail in Appendix F. Based on a qualitative comparison of PRIM boxes on the KPI for throughput times of red customers, it seems there is no significant difference between the experiments based on 1 and on 10 replications. To test this hypothesis, the KPI (All KPI's as stated in paragraph 5.3 have been tested) distribution results of both experiments have been compared using a nonparametric 2-sample Kolmogorov Smirnov (KS) test. A nonparametric method is chosen as normality cannot be assumed in the result distributions. Result distributions for the following KPI's have been tested:

1. Average throughput time for smart calling green customers
2. Average throughput time for smart calling red customers
3. Average throughput time for second line process
4. Average throughput time for no contact process
5. Average throughput time for face to face conversation customers
6. Average throughput time for credit committee process

Null hypothesis: There is no difference between the distributions of 1 and 10 replication results.

Table 5.4 shows the KS-Value results for the null-hypothesis as tested on all KPI's. This hypothesis is tested on a 0.05 confidence interval. Furthermore, table 5.4 provides a conclusion as to whether or not there is a significant difference between the result distributions of an experiment with 1 and 10 replications.

KPI	KS -Value	Significant difference?
1	0.011	No
2	0.01	No
3	0.009	No
4	0.015	No
5	0.076	Yes
6	0.022	No

Table 5.4: Two Sample KS test results

As the KS test statistic value for all KPI's is smaller than the 0.05 confidence interval, it can be concluded that the null hypothesis can be accepted for all KPI's and there is not sufficient evidence to conclude that the underlying distributions are different, with the exception of the average throughput time for face to face conversations. The significant difference of the average throughput time for face to face visits is likely caused by the bottleneck at the research process. As a result of this bottleneck, few customers are forwarded to the face to face process, this makes the throughput time more sensitive to the chosen distribution of face to face visit process time.

Considering the points made above, it can be concluded that 1 replication per scenario is sufficient for the purpose of this study. This could be caused by the sample size and/or by a sufficiently large number of entities flowing through the simulation model. The conclusion regarding the significant difference in average throughput time results for the face to face process seems to point towards the number of entities as a major cause for a reduced need of multiple replications per scenario. However, further research is required to accurately determine the relation between sample size, number of entities, and the required number of replications per scenario.

Finally, in addition to the run control conditions required in DES, the amount of scenarios has to be determined for the purpose of EMA. To this end, an extensive analysis of KPI results has been done through Kernel Density Estimate (KDE) plots. These plots show the probability density of a certain outcome value on the Y-axis and the corresponding outcome value of a certain KPI on the X-axis. The actual graphs represent the probability density function of a distribution based on a certain number of scenarios. The number of scenarios corresponding to a specific colored line is shown in the legend.

Interpretation of KDE plots:

Despite the fact that KDE plots show probability density functions, they are explicitly **not interpreted in a probabilistic manner** for the purpose of this research. Hence, the aim is not to study what the probability of a certain outcome related to certain KPI's is. Rather, the same outcomes based on the same KPI's for different experiments with varying amounts of scenarios are studied. Thus, the probability density functions corresponding to different amounts of scenarios are compared in a qualitative manner to determine whether or not increasing the amount of scenarios will lead to different results in the outcome distribution.

To determine the required number of scenario's a qualitative analysis of the KDE plots is made. The most suitable number of scenarios is the number of scenarios where the KDE graph deviates very little from the KDE graph depicting more scenarios. I.e. we are trying to minimize the amount of scenarios while coming as close as possible to the actual distribution of the results based on a certain KPI. Besides KDE plots, the mean and standard deviation of KPI distributions are studied. It is expected that the mean balances as the number of scenarios increases. A suitable number of scenarios should be chosen as the mean of the distributions starts to show very little fluctuations. Both KDE and mean/standard deviation plots are studied based on three different types of KPI's:

1. Throughput time
    a. Throughput time for customers from arrear to financially healthy
    b. Throughput time for red customers to be called (smart calls)
2. Scheduled utilization
    a. Scheduled utilization research employees
    b. Scheduled utilization first line employees
3. Number of customers in queues
    a. Inbound calls queue
    b. Second line outbound calls queue

Figure 5.3a depicts a KDE plot of throughput time for customers from arrear to financially healthy and figure 5.3b shows Mean and standard deviation plots of throughput time for customers from arrear to financially healthy. To be able to make a better qualitative comparison of distributions based on different scenario numbers, a closer look is taken at specific areas in figure 5.3a. Figure 5.3c zooms in on the top left part of figure 5.3a and figure 5.3d shows a closer look at the top right part of the figure. The complete overview of all KDE plots based on the KPI's described above can be found in appendix H.

[redacted]

[redacted]

Figure 5.3a: KDE plot full view

Figure 5.3b: Mean and standard deviation

[redacted]

Figure 5.3c: KDE plot of throughput time for customers from arrear to financially healthy top left zoom

[redacted]

Figure 5.3d: KDE plot of throughput time for customers from arrear to financially healthy top right zoom

When deciding what number of scenarios is appropriate for the purpose of this study, a few factors have to be taken into account. There is a trade-off between minimization of computing time on the one hand (i.e. minimize amount of scenarios) and maximizing reliability and maximizing the number of scenarios for the purpose of PRIM analyses on the other hand. Looking at figures 5.3c and d, it is very clear that 100, 500, and 1000 scenarios depicted in blue, green and red respectively, are too few to produce reliable results. Once the amount of scenarios surpasses 1.000, the KDE plots start to resemble each other pointing towards smaller differences in distribution of outcomes. Still, even though less obvious, 1.500 scenarios depicted in the light blue line seem to deviate from the other plots, this is especially visible in figure 5.3d. 2.000 scenarios seem to be reasonable amount based on figures 5.3a-d. However, considering the other KDE results shown in appendix H and the notion that more scenarios lead to more reliable results using PRIM analysis, 2.500 scenarios is chosen as the amount of scenarios to use in experiments.

The initialization conditions, run control settings and required number of scenarios for the purpose of exploratory modeling are summarized in table 5.5:

Run control setting	**Value**
Start time and date	Monday 30 September 2013
Warm-up period	[redacted]
Run length	[redacted]
Replication length	[redacted]
Number of replications per scenario	1
Number of scenario's per experiment	2.500

Table 5.5: Summary of initialization conditions and run control settings for simulation model experiments

6. Model results and analysis

The previous chapter showed what a simulation model of processes at ING Arrears Management looks like as well as the most important elements for experimental design in applying an EMA approach on DES. In this chapter, the final steps of the research approach are covered: Exploration and analysis of plausible scenarios and the assessment of exploration in light of danger zones and potential focal areas for adaptive policies. Hence, this chapter will aim to answer the sub research question: Which combination of uncertainties lead to potential danger zones that could jeopardize chosen business objectives at ING Arrears Management?

The first paragraph will briefly cover the model behavior. Afterwards, the second paragraph will cover how uncertainties influence this model behavior. This is done through scenario discovery as briefly introduced in chapter 3. Further explanation on scenario discovery and PRIM will be provided in this chapter.

6.1. Model behavior

To demonstrate the model behavior, a random scenario is chosen. The behavior will be illustrated based on various output graphs as a result of uncertain input variables as well as other elements as defined in the model specification step. The specific data for model parameters can be found in appendix D, the uncertainty values in the demonstrated scenario can be found in table 6.1:

Uncertainties	Description	Value
Number of customers	Monthly inflow of customers in Arrears Management	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
Customer differentiation	Division of customers in ability and willingness categories	[redacted]
Success percentage of payment arrangements	Share of customers that honor their payment arrangement	[redacted]
Correctness percentage of payment arrangements	Share of correctly made payment arrangements	[redacted]
Process time customer call	Process time per (in- or outbound) call	[redacted]
Success percentage of payment arrangements	Share of customers that honor their payment arrangement	[redacted]
Correctness percentage of payment arrangements	Share of correctly made payment arrangements	[redacted]
Max number of arrangements (first and second line)	Max attempts to make a good arrangement with a customer	[redacted]
Effectiveness of research methods	Share of found cellphone numbers of total searches	[redacted]
Productivity	Number of productive hours per employee per day	[redacted]

Table 6.1 Uncertainty values of demonstration scenario.

Essentially, the process at ING Arrears Management needs to be equipped to handle any customer inflow in terms of contacting customers with a mortgage arrear and helping them adequately. This means it is important that all different types of customers in all sub-processes are contacted and/or helped before the next major customer inflow, this part of the model behavior is demonstrated in figure 6.1. The X-axis shows time in hours and Y-axes number of customers in queues.

[redacted]

Figure 6.1. Customers in various call queues.

[redacted]

In this example scenario, the business process at ING Arrears Management seems sufficiently equipped to handle the customer inflow in a timely manner. To check this conclusion, the model behavior is also studied in terms of average throughput times and scheduled utilization of all employees. Figure 6.2 shows the scheduled utilization of all employees. [redacted]

[redacted]

Figure 6.2 Scheduled utilization statistics for all employee types

[redacted]

Figure 6.3 Average throughput time face to face conversation process

[redacted] To make sure no other

bottlenecks exist in the system we can take a look at the average throughput time developments of other sub processes, an overview of these graphs can be found in Appendix I. No other problematic areas can be identified based on an analysis of average throughput times.

Hence, except for the potential bottleneck in [redacted], this example scenario seems to be quite stable and the process dimensions seem sufficient in dealing with the customer inflow. However, this will not be the case in all scenarios. To illustrate this 100 scenarios have been plotted that show the sum of customers in outbound call queues. These 100 scenarios are shown in figure 6.4:

[redacted]

Figure 6.4: Illustration of sum of customers in call queues in 100 different scenarios

The colored lines in figure 6.4 correspond to specific scenarios (their specific scenario id/label is shown in the legend on the right hand side of the figure). The majority of scenarios resemble the behavior demonstrated in our example scenario in figure 6.1 where all monthly customer inflow is handled in a timely manner. However, in figure 6.4 there are six scenarios in which the business process at ING Arrears Management fails to cope with the customer inflow resulting in accumulations of waiting customers in outbound call queues. These are the type of scenarios that ING Arrears Management aims to avoid. In order to be able to avoid these scenarios, first the (common) causes for these scenarios must be identified. This will be done through scenario discovery and the application of PRIM. The next paragraph covers scenario discovery in the case at ING Arrears Management.

6.2. Scenario discovery and Patient Rule Induction Method (PRIM)

As mentioned in chapter 3, Scenario discovery is intended to aid decision makers in identifying policy or strategy relevant scenarios by interactively applying statistical and data mining algorithms to large databases of simulation model results (Bryant and Lempert, 2010). Specifically, for the purpose of this research, the Patient Rule Induction Method (PRIM) is applied to identify relevant scenarios.

PRIM was first described by Friedman and Fisher (1999), they intended PRIM as an addition to the toolbox that data analysts can use when the goal is optimization. PRIM is essentially a bump hunting algorithm that aims to find high density areas in a solution space. The basic idea of PRIM is to start with a large box including all observations (i.e. the solution space) and iteratively reduce that box by peeling the lower or upper side of the box (Chong and Jun, 2008). The peeling process stops when a certain value is reached of the fraction of observations that is excluded from the reduced box (Chong and Jun, 2008). This peeling process is repeated until the large box is completely divided into a number of smaller boxes (subsets). The last box may be improved through pasting, which is essentially inverse peeling (Friedman and Fisher, 1999). More recently, Bryant and Lempert (2010) have applied a form of PRIM for exploratory modeling purposes in their scenario discovery example. The identification of smaller boxes (containing high densities of outcomes of interest) from a large box with all observations through peeling is shown in figure 6.5., adopted from Bryant and Lempert (2010, p. 47). It is important to note that visualizations of PRIM through scatterplots as depicted in figure 6.5 become impossible as soon as the number of uncertainties increases.

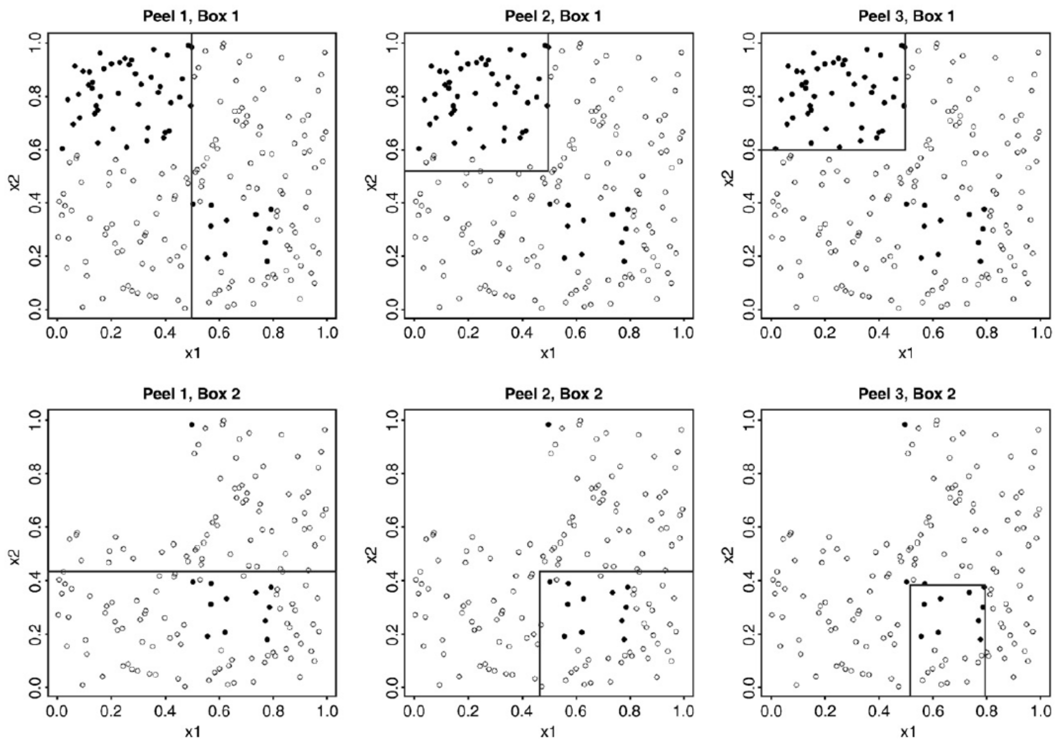

Figure 6.5: Visualization of PRIM peeling trajectory, adopted from Bryant and Lempert (2010, p. 47).

Bryant and Lempert (2010) use PRIM to determine which combinations of uncertain input variables best predicts certain outcomes of interest. So before applying PRIM, two elements need to be determined. First, the uncertain input variables (x1 and x2 in figure 6.5) and their ranges (normalized to a 0.0 – 1.0 scale in figure 6.5) have to be determined. Second, outcomes of interest have to be defined. This second step involves classifying outcomes in a binary way as an outcome of interest based on a selected threshold value for a chosen KPI. [redacted]

[redacted] In figure 6.5 cases of interest are depicted as black dots. Hence the aim is to find as much black dots as possible using the smallest possible amount of boxes.

Two uncertainties can be visualized in a scatterplot as demonstrated in figure 6.5 and three uncertainties can be visualized in a 3-D scatterplot. However, four or more uncertainties require a different visualization method. Boxes based on more than three uncertainties can be visualized through dimension plots (dim-plots). Figure 6.6 depicts an example of a dim-plot with nine uncertainties (x1 – x9). The blue lines corresponding to these uncertainties show the range of values (on a normalized scale) that result in a case of interest. The combination of blue lines in this dim-plot example shows a potential danger zone in regard to a threshold value for a chosen KPI.

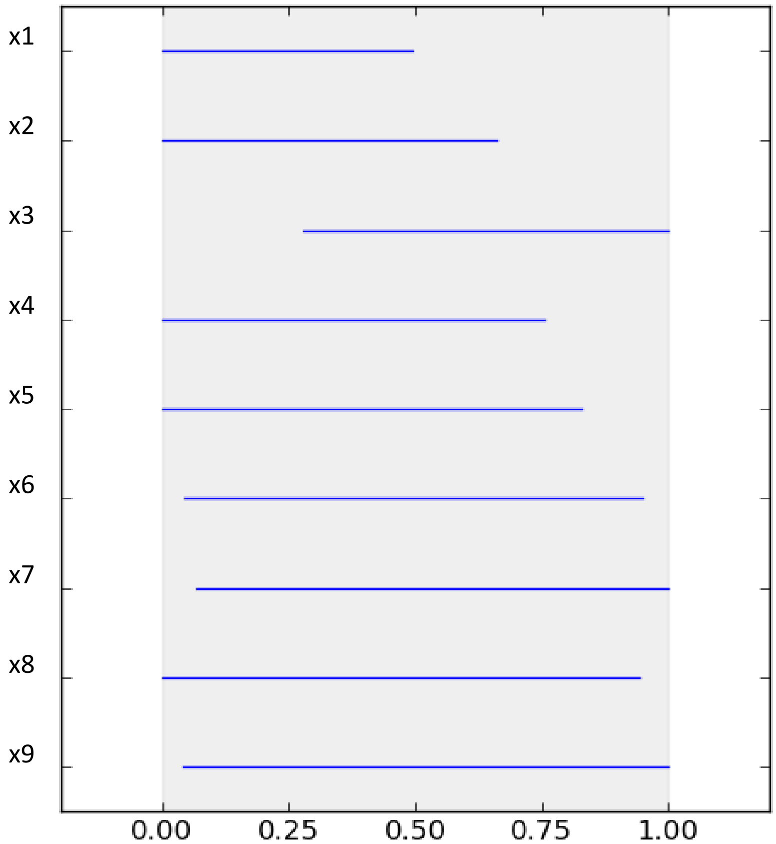

Figure 6.6: Dim-plot example based on a box with 9 uncertainties.

Bryant and Lempert (2010) describe the identification of usable boxes as the user's responsibility. It is up to the user to choose the box for their specific application that best balances coverage, density, and interpretability. High coverage means that a box set captures a high proportion of the total number of relevant cases to a certain KPI, high density means that a box set contains mostly relevant cases, and high interpretability means the outcomes are easy to understand. Ideally, a box set containing scenarios would have a high coverage, density, and interpretability. However, in practice these measures compete; increasing coverage often decreases density while higher interpretability can increase coverage but decrease density (Bryant and Lempert, 2010). Furthermore, even when boxes can be found with high coverage, high density and high interpretability, the user should warrant for over fitting. Over fitting means that outcomes of a certain experiment depend too much on the specific details of that experiment (Witten and Frank, 2005). Hence, the outcomes are likely to be caused by chance and may be different in another experiments, i.e. the outcomes are not robust (Belinga, 2013). Typically, in data mining over fitting is avoided by splitting historical data into a learning set and a test set (Chong and Jun, 2008). However, considering the fact that historical data is, for a large part, unavailable in this study, it is the responsibility of the user to verify whether or not boxes are formed due to chance. This verification is particularly important in experiments where many cases of interest can be found (more than half of the total scenarios) because boxes that are found in those cases tend to have misleading high densities. In experiments where fewer cases of interest are found, high density is a more reliable criterion for the purpose of preventing over fitting. For this reason, all experiments where more than half of the scenarios are classified as cases of interest, visual box verification is conducted through qualitative analysis of several box-plots (appendix J).

The next paragraph illustrates how scenario discovery has been applied in the case study at ING Arrears Management in terms of classification choices, trade-offs between coverage, density, and interpretability, and finally the visualization of the results. The final paragraphs of this chapter will be dedicated to a summary of all PRIM results and the assessment of implications of these results in the case study at ING Arrears Management.

6.3. Scenario discovery at ING Arrears Management

The previous paragraph provided a more detailed introduction of scenario discovery and PRIM. In this paragraph, many scenarios and their impact on chosen KPI's will be explored through the application of scenario discovery and PRIM. As mentioned in the previous paragraph, scenario discovery essentially consists of three steps:

1. Classification of cases of interest
2. Find boxes (i.e. identify danger zones)
3. Visualize boxes through dim plots

After these steps the (visualized) results must be interpreted and conclusions can be drawn in terms of danger zones towards business performance of the proposed new processes at ING Arrears Management. First, a classification is made based on different KPI's:

6.3.1. Classification

[redacted]

[redacted]

The black bullet points show the KPI's based on which the scenario discovery approach is conducted. The open bullet points show the corresponding threshold values of the KPI's. In case any of the open bullet conditions is false, an outcome will be classified as a case of interest. The next step involves finding boxes with high densities of cases of interest.

6.3.2. Finding boxes

For each of the KPI's stated above, based on each of the corresponding threshold values, an attempt is made to find a box with a high density of cases of interest based on common causes in terms of uncertain parameter values. For the purpose of this research the minimum required density is chosen to be at least 80% for any box. The reason for doing so is that a box with a lower density is considered to be unreliable as it contains too many cases of interest that cannot be explained by the common combination of uncertain parameter ranges. Hence, a resulting danger zone will not be useful as a basis for developing (adaptive) strategies.

While finding boxes, it is necessary to make trade-offs between density, coverage, and interpretability of any box. This trade-off can be made by selecting a specific amount of peelings based on the peeling trajectory. While making the trade-off, in general, the aim is to maximize density, coverage, and interpretability. However, in practice, a high score on all three factors is unlikely. For this reason, the choice is made to maximize the coverage while keeping a density of at least 80% and at the same time striving to minimize the amount of restricted dimensions as this increases interpretability. An example from the case at ING Arrears Management is given to illustrate the process of making trade-offs between density coverage, and interpretability. In this example, a box will be found based on the average throughput time for customers in the "no contact" process. I.e. scenarios where the average throughput time exceeds [redacted] will be classified as a case of interest. Based on this classification, [redacted] cases of interest are found out of 2.500 scenarios. The peeling trajectory for box 1 is shown in figure 6.7 (the complete peeling trajectory including numerical values per peeling is shown is appendix G).

Interpretation of peeling trajectory plots:

Figure 6.7 is an example of a plot depicting the peeling trajectory as applied through the use of the PRIM algorithm. The X-axis shows the number of peelings. The Y-axis shows mass, density and coverage as well as the number of restricted dimensions. Mass is depicted in green, coverage in red, density in blue and the number of restricted dimensions in purple. Mass depicts the fraction of cases in the box out of the total number of cases in the solution space. Coverage depicts the fraction of cases of interest in the box out of the total number of cases of interest in the solution space. Density depicts the fraction of cases of interest in a box out of the total number of cases in the box. Restricted dimensions correspond to the number of uncertainties that are peeled for the purpose of identifying boxes. With every peeling the total solution space is reduced in size in an effort to find a high density box, this is demonstrated by the mass that decreases as the number of peelings increase. The coverage also decreases as the number of peelings increases, because (inevitably) some cases of interest are omitted from the box during the peeling trajectory as well. Density increases as the number of peelings increase because more and more cases that are not of interest are omitted from the box. Finally, the number of restricted dimensions increases as all possible box dimensions are peeled in an effort to find a high density box.

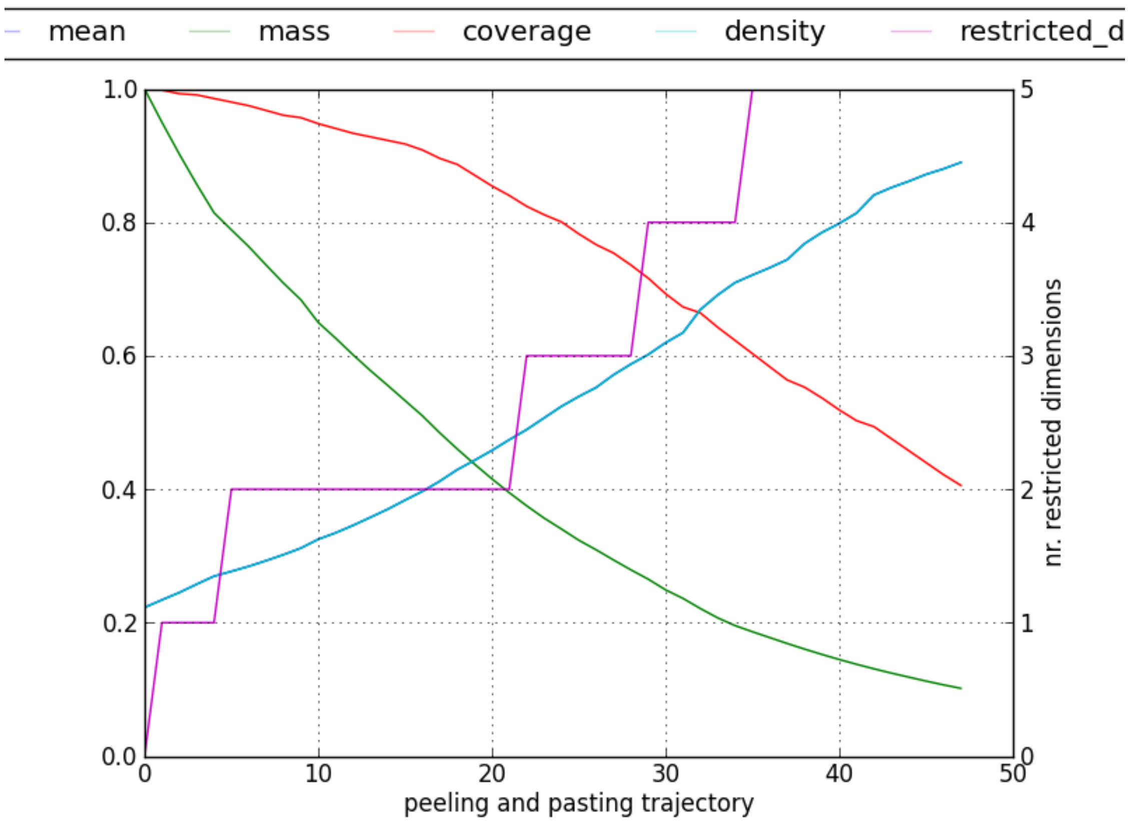

Figure 6.7: Peeling trajectory for box average throughput time no contact process

Before selecting a specific peeling in box 2, it is important to note that PRIM was unable to find a second box that meets the density threshold of 80%. Hence, at this point the choice is made to maximize coverage while maintaining a reasonable density (at least 80%) and minimizing the amount of restricted dimensions for the purpose of increasing interpretability. If a second box could be identified, the density would be maximized with less concern for high coverage, as the total coverage will increase through the identification of more boxes. Considering the aims mentioned, based on figure 6.7, 40 peelings are chosen for the identification of the box. Now that a box is found and the specific amount of desired peelings is selected, the box can be visualized for analysis.

6.3.3. Visualization of boxes

The remaining box resulting from the choices made during the peeling trajectory are visualized in a dimension plot (dim-plot). The box depicted in figure 6.8. The blue lines constitute the box and are plotted on a normalized range for the uncertain parameters [redacted]. This means a danger zone can be identified that may cause a failure in achieving the minimum desired threshold value of [redacted] for the average throughput time in the research process. [redacted] As figure 6.8 only shows the problematic uncertainty ranges on a normalized scale it is difficult to translate it into a concrete danger zone on which adaptive policies can be based. For this reason the exact numerical ranges per uncertainty are shown in table 6.2.

[redacted]

Figure 6.8: Dimplot of 2 box found based on classification for average throughput time of research process.

[redacted]

Table 6.2: Numerical representation of 2 boxes found based on classification for average throughput time second line process.

The approach described in this paragraph, including the detailed numerical peeling process described in appendix G is used to find all boxes that constitute danger zones in the case study at ING Arrears Management. The next paragraph contains the result analysis of four experiments analyzed according to the scenario discovery approach.

6.4. PRIM analysis: Results of experiments at ING Arrears Management

The previous paragraph elaborated on the approach towards scenario discovery in the case study at ING Arrears Management. This paragraph contains an analysis of PRIM results based on four experiments:

- Experiment 1: Uncertain input parameters and uncertain capacities
- Experiment 2: Uncertain input parameters
- Experiment 3: Fixed values for customer inflow, segmentation and differentiation
- Experiment 4: Priority of second line tasks over inbound calls for second line employees

All experiments are based on the same version of the simulation model and are subject to the simulation set-up and run control settings as described in chapter 5. Furthermore, the uncertainty ranges have been used as specified in paragraph 5.2. However, there are some differences as the experiment titles suggest. In the first experiment all uncertainties are taken into account and additionally, the capacities for all employee types have been treated as uncertain. Broad range estimations of the capacities have been made based on a simple spreadsheet estimation in which all sub processes were treated as independent processes. Based on this spreadsheet calculation, an estimation of all capacity ranges was made. The results of this estimation are depicted in table 6.3:

	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Table 6.3: capacity range based on min and max required capacity estimation of independent sub processes

The second experiment is exactly similar to the first experiment with the exception of capacities. Capacities are fixed in the second experiment [redacted] These values are depicted in table 6.4.

	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

Table 6.4: capacity as dimensioned in Arrears Management proposal

[redacted]

The fourth experiment has similar settings to the second experiment (i.e. fixed capacities as stated in table 6.4 and uncertainty ranges as specified in table 5.1 in paragraph 5.2). In the fourth experiment, inbound calls have been given a lower priority for second line employees. [redacted]

The results of all KPI's in all four experiments based on PRIM analysis will briefly be discussed per experiment. Only the detailed box visualizations and numerical box results can be found in Appendix J.

Results summary experiment 1: Uncertain input parameters and uncertain capacities

[redacted]

Results summary experiment 2: Uncertain input parameters

[redacted]

[redacted]

Results summary experiment 3: Fixed values for customer inflow, segmentation and differentiation

[redacted]

[redacted]

Results summary experiment 4: Priority of second line tasks over inbound calls for second line employees

[redacted]

[redacted]

This chapter aimed to answer the research sub-question: *Which combination of uncertainties lead to potential danger zones that could jeopardize chosen business objectives at ING Arrears Management*? "Business objectives" in this question are defined as the KPI's and their corresponding thresholds used for the purpose of scenario discovery. Overall, it can be concluded that [redacted]

6.5. Implications of exploration for ING Arrears Management

Even though the formulation of (adaptive) policies is explicitly out of scope in this research, the results discussed in this chapter can be useful as a starting point for the development of such policies as well as prioritization of monitoring and steering efforts at ING Arrears Management. Some of the uncertainties that are part of various danger zones identified in the previous paragraphs cannot be influenced. [redacted]

[redacted]

[redacted]

Based on the identified danger zones, the starting points for steering measures mentioned above may be effective. However, it is recommended to test and monitor it on a small scale once the process is implemented or in a simulation model for further research.

Chapter 4, 5 and 6 were dedicated to the introduction, execution and analysis of a case study at ING Arrears Management. The purpose was to provide valuable insights for ING Arrears Management through application of the proposed approach in chapter 3 as much as possible. The next chapter will focus on the discussion of EMA applied on DES as a suitable method for solving the case study at ING Arrears Management. Vice versa, as the case study was meant as a means to acquire a prove of concept for EMA applied in DES in deeply uncertain business process environments, the case study will be evaluated in terms of its suitability in providing such a proof of concept.

7. Discussion of EMA approach applied on DES at ING Arrears Management

Chapter 3 presented a research approach of applying EMA on DES which was applied in a case study at ING Arrears Management, described throughout chapter 4, 5 and 6. The aim of this chapter is twofold. First, the usefulness of the case study will be discussed for the purpose acquiring a proof of concept of the proposed approach. Second, the actual approach towards applying EMA on DES will be discussed by confronting theory and literature as presented in chapter 3 with lessons learned in practice during the case study at ING Arrears Management. The discussion presented in this chapter will form the basis for answering the two remaining scientific research questions:

- When is it desirable to apply an EMA approach on DES?
- What are the main strengths and weaknesses of applying an EMA approach on DES?

7.2. Case study discussion

The main purpose of the case study at ING Arrears Management in light of this research was to provide a proof of concept for the application of EMA on DES in deeply uncertain business process environment. In this paragraph an evaluation is made for the purpose of determining to what extent the case study was suitable as a tool in providing such a proof of concept.

Firstly, it is important to note that even if the case study is perfectly suitable, the results of this study cannot be deemed "proof" of the effectiveness of the proposed method as one case study is insufficient to provide such a proof in the academic sense. However, the results of this study given the limitations in line with only one case study can still be useful as a starting point for further research towards applying EMA on DES in business processes under deep uncertainty.

Secondly, as the aim is to show an application of EMA on DES can help decision makers in designing business processes under deep uncertainty, it is important that the case study does in fact meet the criteria for business processes in a deeply uncertain environment. As doing business for ING Arrears Management means minimizing risk costs through helping customers with their financial problems, the processes subject to (re)design in this case can be perceived as business processes. Next, the extent to which the case study is subject to deep uncertainty is evaluated based in the criteria for deep uncertainty as described by Lempert et al. (2003, p. 3):

"Deep uncertainty exists when analysts do not know, or the parties to a decision cannot agree on:

(1) the appropriate conceptual models that describe the relationships among the key driving forces that will shape the long-term future,
(2) The probability distributions used to represent uncertainty about key variables and parameters in the mathematical representations of these conceptual models, and/or
(3) How to value the desirability of alternative outcomes."

Condition 1 can be summarized as "structural uncertainty", condition 2 as "parametric uncertainty" and condition 3 as "valuation of outcomes uncertainty". Initially, the case study at ING Arrears Management seemed to meet all requirements as there was no final conceptual process design (structural uncertainty), the value of many factors was unknown because the process did not exist (parametric uncertainty), and the valuation of outcomes was debated upon in terms of which KPI's to use as valuation measures. Furthermore, the priority and desirability of values regarding these KPI's was unclear. In practice, only the first and third conditions were met. Even though the structure of the process was initially unknown it turned out be very difficult to identify alternative process structures that could be taken into account in parallel as categorical structural uncertainties. Multiple models were used during the case study, but these models differed in aggregation level and were used consecutively and never in parallel. Hence, despite the fact that the environment seemed to uphold all criteria for deep uncertainty, structural uncertainties were not taken into account during the case study. This means that if the level of confidence in the applicability of EMA on DES would be high in dealing with structural uncertainties, it can never be concluded based on this research.

This paragraph provided a limitation frame in regard to the extent to which conclusions based on the case study can be extrapolated towards business processes in general. Given these limitations, the next paragraph holds a discussion the proposed EMA approach applied on DES.

7.3. Discussion on EMA approach applied on DES

The proposed EMA approach applied on DES will be discussed briefly step by step according to the steps presented in chapter 3. Furthermore, a stepwise versus an iterative approach of EMA applied on DES will be discussed.

7.3.1. Stepwise discussion of EMA approach applied on DES

1. *Conceptualization of new processes*

During the conceptualization step it was argued that, in case of exploratory modeling, one or an ensemble of more simple models resembling business processes instead of one large complex model can be built that are all plausible (Bankes, 1993). Uncertainties in these business processes can find their origin in the lack of available information to construct a conceptual model or disagreement among stakeholders in regard to what the model structure looks like (Lempert et al., 2003). As mentioned in the previous paragraph, structural uncertainty was not taken into account during the execution of the case study at ING Arrears Management. This poses the questions whether or not this was due to the nature of the specific case study or as a result of the proposed approach. While performing the case study it became clear that there was quite a detailed view on what the processes should look like and identifiable disagreements in terms of process structure were not present. It is important to note that just because there were no identifiable disagreements in this case it does not mean that there will not be any in other (similar) cases. Hence, the omission of structural uncertainty cannot be ascribed to the proposed approach. Further research is required to determine to what extent an application of EMA on DES is suitable dealing with structural uncertainties.

2. *Uncertainty analysis*

During the analysis of uncertainties it is important to determine what exactly is known in regard to the target business processes, what is unknown and to what extent is it unknown. I.e. it is important to determine the amount, type, level and location of uncertainty in a target system. In the case study, determining what is known and what is unknown both proved to be quite challenging. Determining what is known in a large complex organization such as ING was difficult as central databases are not always present. Hence, it may even occur that someone is confident that certain information is available and at the same time it may be unknown where to find it. In terms of identifying what is unknown, the level of detail demanded by DES proposed a challenge in the case study. In order for DES to be of added value in comparison to other more simple analytical or mathematical models, a certain level of complexity is required. In deeply uncertain environments such as the case at ING, this level of complexity may not be achieved while maintaining a sufficient degree of confidence in the simulation model. This is a concern which can be investigated in future research. A good balance should be found between sufficient detail for DES to be of value and the lack of available details in terms of data in deeply uncertain environments resulting in the need for more simple models (Bankes, 1993). Furthermore, the case study shows that determining what is (un)known is not a onetime effort. What is (un)known in a business process may differ on a daily basis. Therefore, iterations are required in the uncertainty analysis phase. Rather, it can be perceived as an integral part of EMA throughout the modeling cycle.

3. *Specification of new processes: Development of simulation models*

Exploratory modeling may involve guessing ranges of data in case real world data is unavailable and the required variable data is considered uncertain (Bankes, 1993). Hence, even when specifying a single model to accurately reflect reality is not possible, but where relevant information exists, EMA can be a useful approach (Kwakkel and Pruyt, 2013). [redacted] By asking different experts to estimate (i.e. guess) the value of a certain parameter, the minimum and maximum guesses can be taken as a uniform range for the purpose of scenario discovery. Involving experts in this manner also proved to improve model confidence. Furthermore, the results of the case study as described in the previous chapter show that it is in fact possible to acquire valuable insights in case not all information is available.

4. *Verification and validation*

In short, validation for the purpose of EMA in this case study was defined as acquiring sufficient confidence in the simulation model for its intended purpose. This proved quite difficult as the simulation model was a black box for most stakeholders. One of the most important uses of EMA is to make (expert) assumptions explicit in a simulation model (Bankes, 1993). However, this was only achieved in terms of input validation and outcomes validation. Hence, a lot of time is necessary to explain how these input parameters were translated to the resulting outcomes. This is a risky exercise for the modeler, because failing to do so may result in the loss of confidence in the simulation model for the purpose of EMA. Still, based on experiences in the case study, this seems to be the most viable option as trying to transform the model into a white box requires time and expertise that decision makers seldom have.

The environment at ING Arrears Management is continuously evolving resulting in new developments and insights which calls for iterative validation. Hence, minor model adjustments must be made multiple times to ensure confidence in the simulation model remains at a high level. When applying the proposed approach to other cases, it is advisable to have short iterations where decision makers' assumptions are tested to slightly mitigate the use of black box models.

5. *Experiments for analysis dataset*

In chapter 3, the concern was formulated that it is important to realize that DES typically involves probabilistic information (Lempert et al., 2006) and inclusion of this probabilistic information might contaminate the scenario results regarding causality in an EMA approach (Bryant and Lempert, 2010). I.e. observed scenarios that are assumed to be caused by combinations of uncertain input variables may in fact be caused by the probabilistic nature of a DES model.

In the case study at ING Arrears Management a simulation model was used with many factors that are subject to probabilistic information (appendix D). Nevertheless, the observed scenarios in this case can be perceived as reliable because increasing the amount of replications per scenario does not result in significant differences in outcomes (see paragraph 5.4.3). Two explanations for the reliability of scenarios are given in paragraph 5.4.3. The first explanation is that the large amount of entities averages probabilistic information in the DES model used in this case. The second explanation is that running multiple scenarios averages probabilistic information in the outcome distributions. Either way, more research is required to accurately determine the relationship between the probabilistic nature of DES models, amount of replications, amount of scenarios and reliability of these scenarios in terms of cause and effect interpretation. However, in light of application of the proposed approach to other cases, it would be wise to be careful with models that use few entities as outcomes in these models are likely to be affected more by probabilistic information in comparison to the case presented in this research.

6. *Generate exploratory database*

The chosen approach involving 2.500 scenarios sampled using LHS works properly for the purpose of scenario discovery in this case. However, the practicality of generating 2.500 scenarios is questionable from a technical point of view as Arena is only able to use 1 pc core at the same time, computing times can be very long. The generation of a dataset containing 2.500 scenarios based on the case study at ING Arrears Management takes approximately 48 hours. Furthermore, as the amount of uncertainties increases, the required amount of scenarios to adequately represent the full solution space will also increase. Consequently, the required computing time will increase as well. More complex models also require longer computing times. Hence, in cases where larger models are required and/or more uncertainties are present, it is advisable to look for computing time reducing measures. Possibilities to this end are the use of more (virtual) pc's and/or to avoid demanding model constructs.

It is important to realize that long computing times are not only inconvenient for the analyst, it is in direct contrast with the argument made above: Short iterations are required for the proposed approach to be effective.

7. Explore and analyze plausible scenarios

The chosen analysis approach in the case study at ING Arrears Management is scenario discovery through PRIM analyses. Based on the analysis presented in chapter 6, it can be concluded that PRIM is a suitable analysis method in EMA applied on DES. Danger zones can be identified using PRIM which forms an important starting point for developing adaptive policies. However, some KPI's where boxes cannot be found leave cases of interest unexplained. Furthermore, black swan type of scenarios may go unnoticed which could have disastrous impacts on business goals in reality. PRIM only covers danger zones as a result of combined impacts of uncertainties on KPI's. Sometimes, the individual impact of uncertainties is also required. To a certain extent, manual examination of results can give an insight in possible black swan scenarios and the effect of individual uncertainties. However, more detailed research is required to identify a more suitable method.

Based on the results in experiment 1 presented in chapter 6, it seems PRIM may be used to determine the appropriate dimensions of the business processes at ING Arrears Management in terms of required capacities. However, more research is required to determine if this is in fact the case and if other more suitable methods are available to this end.

PRIM analyses prove very useful in bottleneck detection (see chapter 6). In case a very large amount of cases of interest is found as a result of the classification step, it may point to a bottleneck in the process as it does in this case. Visualization of scenarios may also help in the identification of bottlenecks as shown in figure 6.4.

8. Asses implications of exploration

In chapter 3, it was argued that assessing the results of EMA can lead to the identification of alternative strategies, help prioritize research, and verify hypotheses that decision makers may have in regard to the target system (Bankes, 1993). Effective visualization and communication of the results is of crucial importance to a successful application in real world environments (Kwakkel and Pruyt, 2013). Explaining the danger zones as a summary of scenarios may help in effective communication as scenarios describe the future in a way that decision makers find easy to understand (Bryant and Lempert, 2010).

Based on experiences in the case study at ING Arrears Management, it can be concluded that EMA does help in the identification of alternative strategies, help prioritize research (in this case prioritize IT implementation projects), and verify hypotheses. Dim-plots were mainly used to visualize the identified danger zones. These types of visualizations are not immediately clear and require some explanation. Furthermore, discussion of visualized results almost always leads to more questions requiring new experiments. It would be valuable to use more dynamic visualizations resulting in quick responses to questions rather than having to schedule follow-up meetings. Explaining the danger zones in a narrative manner as summaries of scenarios proves to be a useful method in communicating the identified dangers zones.

Lastly, the assessment of implications of exploratory efforts are largely limited to the identification of danger zones for the purpose this research. However, in future applications

EMA can be applied on DES for the purpose of developing adaptive policies along the lines of Adaptive Robust Design (Hamarat et al., 2013).

7.3.2. Discussion on stepwise vs. iterative approaches in EMA applied on DES

In chapter 3, it was noted that a traditional DES modeling approach is treated stepwise. Once a phase is completed, the next phase is initiated and the modeler moves on in the DES modeling study (provided that certain stage gates are passed successfully). Based on the previous paragraph, there are several observations that demand a different approach towards applying EMA on DES. It starts with the notion that the answer to the question "what is (un)known or what is (un)certain?" can never be given based on a snapshot. This question requires a dynamic answer that changes over time. Because what is (un)certain changes over time, the conceptual model and specified versions of that model have to change with it. As a result continuous iterative validation is required.

The changes in what is (un)known and (un)certain may result from a business process environment. However, the case study shows that it can also find its origin in the analysis and assessment of implications for experimental results. For this reason it can be concluded that the exploration of an exploratory database and the assessment of the results form the basis of a feedback loop towards re-evaluation of the uncertainty analysis and model conceptualization. Based on the case study, it can be concluded that the feedback loop that drives what is (un)known and (un)certain helps in identifying and prioritizing what should be known. As a result of this, decision makers are helped in identifying gaps, risks and weak spots in their proposed business processes. The feedback loops for iterative validation and iterations in analyses and re-evaluation of the conceptual model and uncertainty analyses are added (depicted in orange) to the proposed research approach in Appendix K.

The discussion presented in this chapter forms a basis to answering the remaining scientific research (sub-)questions: "When is it desirable to apply an EMA approach on DES?" and "What are the main strengths and weaknesses of applying an EMA approach on DES?" Besides answers to these questions, the following chapter holds the answers to the practical research sub-questions as well as answers to both main research questions.

8. Conclusions and recommendations

Firstly, the overall conclusion will be summarized based on the main practical and scientific research questions. Secondly, all sub-questions related to the main research questions will be answered. Finally, recommendations for future research will be made.

8.1. Conclusions based on main research questions

Two main research questions were formulated, one question from a practical point of view and another question from a scientific point of view. First, the practical research question will be answered:

How can Exploratory Modeling and Analysis, applied on DES models, help ING in designing efficient new business processes which are robust under uncertainty?

The aim in the case study for ING Arrears Management was to identify all relevant danger zones that might jeopardize the achievement of business objectives. Applying the proposed approach in the case study yielded the identification of several danger zones. Furthermore, bottlenecks were identified and potential capacity issues in various sub-processes. However, numerous potentially dangerous scenarios remain unexplained through the application of PRIM analysis. Therefore, based on the case study, it can be concluded that ING Arrears Management was partly helped in designing efficient new business processes in an uncertain business process environment. However, in order for the business processes to be considered robust under (deep) uncertainty further development and testing of adaptive steering measures is required.

Before an endeavor towards adaptive robust design is taken, it is recommended to improve the available tools from a technical perspective, as this would vastly increase the probability of successfully solving a case study in the future.

Even though the case study at ING Arrears Management was not completely solved through the application of the proposed research, it can be concluded that the approach shows great potential compared to a traditional DES approach. Not only in the appropriate application of tools and techniques for EMA, but also in the iterative approach as described in the discussion that resulted in helping decision makers at ING Arrears Management in identifying gaps, risks and weak spots in their proposed business processes.

Considering the conclusions mentioned above, the answer to the main research question can be summarized as follows. Exploratory Modeling and Analysis, applied on DES models can help ING in designing efficient new business processes which are robust under uncertainty when:

- An iterative approach towards execution of the model cycle is conducted;
- Exploratory tools and techniques are further developed;
- The identified danger zones are used to develop adaptive steering measures.

The main research question from a scientific point of view is:

How can an Exploratory Modeling and Analysis (EMA) approach be applied on Discrete Event Simulation (DES) in order to help decision makers design business processes and develop adaptive polices under uncertainty?

When answering this question, a few remarks must be made. First, the case study used to provide a proof of concept for the proposed approach towards applying EMA on DES has only been solved partly. Hence, it can only provide a partial proof of concept. Furthermore, the (partial) proof of concept is based on a single case study. For this reason, extrapolation of conclusions towards business processes under uncertainty in general must be made with great care.

Considering the limitations mentioned above, based on theory presented in this research and on the case study conducted at ING Arrears Management, the research question can be answered. An Exploratory Modeling and Analysis approach can be applied on Discrete Event Simulation in order to help decision makers design business processes and develop adaptive policies under uncertainty when the following methodological obstacles are overcome:

- Probabilistic information has to be dealt with in such a way that scenario interpretation is not contaminated. This can be achieved through appropriate use of replications or the inclusion of probabilistic information as an uncertainty (the latter solution may result in an information overload due to a high number of uncertainties)
- Based on the case study, it can be concluded that highly aggregate DES models do not yield very interesting results. Therefore, sufficient information must be available to make a detailed model to the extent that it is in fact more useful than a simple analytical/mathematical model
- Stakeholder involvement and commitment to the exploratory efforts has to be sufficient for the iterative modeling approach to be successful
- More data mining and machine learning techniques have to be available for the exploration of the full solution space (PRIM does not explain all scenarios)
- Computing time has to be limited as this will enable sufficiently short feedback loops for iterative modeling
- Technological tools have to become mature which will enable faster processing of results and easier communication of these results

8.2. Answers to sub-questions

Practical ING sub questions:

What are the most important uncertainties in new processes at the ING Arrears Management department?

During the conceptualization phase a qualitative analysis of uncertainties has been made. The resulting uncertainty list includes 12 uncertainties that are considered relevant by stakeholders at ING Arrears Management. [redacted]

[redacted]

What do simulation models for new processes in the ING Arrears Management department look like?

An important notion in answering this question is that any answer to this question is a snapshot of reality in an uncertain environment. Considering continuous developments at the ING Arrears Management department, a simulation model of the processes today may look different from the simulation model that should be used to represent the processes in two months. While keeping this in mind, a choice is made for the appropriate aggregation level which represents a trade-off between the desire for detail to accurately represent plausible process structures and the expiration date of the simulation model in light of the uncertain environment. What the resulting simulation model looks like can be found in chapter 5 and appendix D.

Which combination of uncertainties lead to potential danger zones that could jeopardize chosen business objectives at ING Arrears Management?

"Business objectives" in this question are defined as the KPI's and their corresponding thresholds used for the purpose of scenario discovery. [redacted] Hence, it can be concluded that these factors should be monitored carefully and to the extent to which it is possible, they should form the primary area of attention when developing adaptive policy measures. The next paragraph will briefly discuss the implications of the results presented in this paragraph (and in Appendix J in more detail).

Scientific sub questions:

What does a methodology to apply an EMA approach on DES look like?

The methodology that was used in this research is adapted based on a traditional DES approach from an EMA point of view. The application of this approach in the case study at ING Arrears Management shows that it can be used in practice for the purpose of applying EMA on DES. Stepwise differences can be found in all model phases (conceptualization, specification, validation, experimental design, analysis and assessment of results). Two steps are added that are unique to an application of EMA on DES which are uncertainty analysis and the generation of an exploratory database. Besides differences in specific steps there is as a difference in how the model cycle in a traditional DES approach is conducted and how it is recommended to be done while applying EMA on DES. A traditional DES approach involves execution of all steps in the model cycle consecutively. In contrast, the application of EMA requires iterative validation and iterations in analyses and re-evaluation of the conceptual model and uncertainty analyses. Furthermore, some steps should be perceived as dynamic. I.e. uncertainty analysis and validation are never completely finished in an EMA approach.

When is it desirable to apply an EMA approach on DES?

It is desirable to apply EMA on DES in business process environments when more simple analytical or mathematical techniques are insufficient. Based on theory discussed in this research, it can be concluded that other techniques are inadequate if the target business process is complex and deep uncertainty is present. From a practical point of view (i.e. based on the case study), it cannot be concluded that an application of EMA on DES is only desirable in cases of deep uncertainty considering the fact that structural uncertainty has not been taken into account in the case study. Furthermore, the case study does show an application of EMA on DES can yield interesting results even when structural uncertainties are not taken into account. However, other techniques such as sensitivity analysis may suffice as well in case structural uncertainty and uncertainty in the valuation of outcomes are not taken into account.

Besides requirements for the target business process in terms of complexity and uncertainty, the following methodological obstacles have to be overcome in order for an application of EMA on DES to be desirable:

- Probabilistic information can be dealt with in such a way that scenario interpretation is not contaminated;
- There is sufficient information to make a detailed model to the extent that it is in fact more useful than a simple analytical/mathematical model;
- Stakeholder involvement is sufficient for the iterative validation and discussion of hypothesis based on model iterations;
- Techniques are available for the exploration of the full solution space (PRIM does not explain all scenarios);
- Computing time can be limited, enabling sufficiently short feedback loops for iterative modeling;
- Technological maturity of tools is more advanced, enabling faster processing of results and easier communication of these results.

The answer to this sub-question stated above provides a good insight into what circumstances and under which conditions, an application EMA on DES is desirable. Once the decision is made to apply an EMA approach on DES, the question remains:

What are the main strengths and weaknesses of applying an EMA approach on DES?

The largest strength of applying an EMA approach on DES is that it allows decision makers to be supported through simulation models even in environments where (deep) uncertainty is present. Other strengths include the conclusion that an application of EMA on DES allows for:

- The use of partial information to gain insights in business processes;

- Confirmation of decision maker hypotheses and assumptions;
- Prioritization of process implementation steps and focal steering measures;
- High stakeholder involvement which results in model and analysis confidence;
- The identification of danger zones as a basis for (adaptive) measures.

Based on a confrontation with DES and EMA literature with the proposed method of applying DES on EMA as well as a confrontation with lessons learned from practical application of the proposed method in a case study, several weaknesses have been identified:

- The proposed approach is not suitable for dealing with models involving large amounts of probabilistic information because this contaminates the interpretation of scenarios
- There is a tension between the need for DES models to be relatively detailed for them to add value in a complex business process on the one hand and an uncertain environment requiring simple simulation models on the other hand
- The requirement for high stakeholder involvement forces stakeholders to invest time in the modeling cycle because they are involved in iterative uncertainty analysis, iterative validation and iterations of the entire model cycle as a result of exploratory analyses
- PRIM is insufficient as an exploratory technique to explain all potentially dangerous scenarios
- Visualization capabilities are limited at this point resulting in difficult communication of results
- Technical limitations result in long computing times as well as difficult interpretation of results

Based on the conclusions drawn in the first two paragraphs of this chapter, recommendations for future research that builds on this thesis will be given.

8.3. Recommendations for future research

In recommending future research a distinction is made between recommendations for future research at ING Arrears Management and from an academic and a technical point of view.

8.3.1. Future research at ING Arrears Management

As mentioned in the discussion above, the business processes at ING Arrears Management are continuously under development resulting in dynamic developments in what is (un)certain. As the implementation of the proposed new business processes continues, [redacted] Furthermore, it is recommended to re-evaluate the uncertainties in terms of determining whether an uncertainty has been deemed as such because data was not yet available (such as the productivity of employees) or because it will remain uncertain after the processes are completely implemented (for example the long term developments of customer inflow rates). By gradually improving the simulation model, confidence in the model and the model outcomes will increase. Furthermore, the model can be used to assess different process lay-outs in the second line process to determine what the impact of various second line measures will be on efficiency.

8.3.2. Future research from an academic and technical point of view

The recommendations for future research from an academic and technical point of view are largely based on the obstacles mentioned in answering the main scientific research question.

Considering the proof of (partial) proof of concept provided in this research is only valid for the case study presented in this report, the most important recommendations is to apply an EMA approach on DES on other cases where business processes under (deep) uncertainty can be identified. When choosing case studies for future research, I would recommend selecting case studies in which an attempt can be made to overcome the following three methodological obstacles:

- Inclusion of probabilistic information: Future research can focus on case studies that vary in the amount of probabilistic information. The usefulness of a scenario based approach can be investigated in light of probabilistic information that is expected to contaminate these scenarios. It would be valuable to know how (much) probabilistic information can be included in DES models such that scenario results acquired through EMA are not contaminated.
- Complexity vs. uncertainty: Future research can focus on case studies with different levels of uncertainty and complexity. The purpose of performing different case studies on these axes would be to try and identify under which conditions of complexity and uncertainty an application of EMA on DES is most valuable.
- Exploratory techniques: PRIM proved to be useful in identifying danger zones. However, not all scenarios can be explained through PRIM and not the entire solution space is explored properly through PRIM. Therefore, it is recommended to use different exploratory techniques in future research.

Besides overcoming methodological obstacles, it is recommended to dedicate future research to two technical aspects:

- Computing time: Complex DES models may take a long time to run. Long runtimes in combination with a large amount of required scenarios (and if necessary also replications) is likely to result in impractically long computing times for experiments. Future research should focus on reducing the computing time through the potential parallel application of (virtual) machines in combination with efficient data collections methods. Furthermore, research towards reducing runtimes in DES models could be done.
- Technical maturity: The technical approach applied for the purpose of this research is far from ideal. Different tools (Excel, @Risk, Scenario Navigator, Arena, Eclipse/Python, and SQL) had to be applied in order to be able to execute the required experiments. This approach makes applying EMA on DES a very time consuming exercise. Furthermore, the reliability of some tools in dealing with large amounts of scenarios and the resulting large quantity of data is questionable. For these reasons, I recommend to conduct future research to developing an integrated tool that allows for user friendly exploration, analysis and communication of results in business process models under deep uncertainty.

9. Reflection

9.1. Reflection on the research process

The main deliverable from a scientific point of view for this research was to identify the basic methodological principles for applying EMA on DES. However, to determine these basic principles, the methodology had to be applied in a real case study at ING Arrears Management. And in turn, to achieve this, a pragmatic technical construction which allows the execution of EMA experiments was required. Given the fact that such a study has never been performed and the notion that the required skills to perform these types of experiments are not part of the curriculum in the SEPAM MSc. program, a lot of time had to be invested in technical issues. This is of course in contrast with the main objective of this research which is to identify methodological principles. Hence, a difficult contradiction was present during the research process where methodological issues are expected to receive the most attention while technical issues demanded the most attention. A very real risk was to focus too much on the technical/practical requirements of this research and to lose track of what it is really about. I.e. there was the risk that the required simulation model and tools for analysis would evolve to become the goal itself instead of remaining a means to the end of identifying methodological principles.

In bi-weekly meetings at ING, an attempt was made to minimize these risks through intermediate result presentations which forced zooming out of the technical issues and thinking about the application in the case study at Arrears Management which in turn provided a foundations for methodological conclusions. I would recommend future graduate students to be very aware of the danger of focusing too much on the technical/practical side of a master thesis as this is unlikely to ever be the focal point of a graduation project at TPM.

On the other hand, the many technical discussions that were required during the research process (mainly at Systems Navigator for DES related issues and with [redacted] for EMA related issues) resulted in interesting methodological insights.

9.2. Reflection on the case study at ING Arrears Management

ING is of course a real company with real business goals and with limited available time at the decision making level. Considering the experimental nature of this research, it was very difficult to promise concrete results in terms of determining the most important focal areas for adaptive policies at ING Arrears Management in an early stage. Furthermore, there was a risk of not acquiring sufficient involvement in the EMA efforts resulting in lack of confidence in the simulation model among managers at ING Arrears Management. It was very helpful to have regular weekly or bi-weekly meetings to force the translation of technical efforts and issues into real (intermediate) advices. This helped the process of developing and testing an application of EMA on DES both from a technical as well as from a methodological point of view.

Unfortunately (though inevitably) there were many dead ends during the research as well as during the case study at ING Arrears Management. This can make the research quite demotivating. However, I have learned during the case study that dead ends are as important as conclusions and recommendations that will be used in the end because it helps to develop the methodological approach and it makes the results more reliable.

9.3. Reflection on EMA applied on DES

The proposed approach towards applying EMA on DES as demonstrated in the case study at ING Arrears Management seems to work and can even be perceived as a (partial) prove of concept. However, considering the experiences in applying this approach, both from a methodological point of view as well as from a technical and practical point of view there is a long way to go before this can be a widely applied approach.

From a methodological point of view it proved very challenging to involve all relevant stakeholders to the extent that sufficient confidence in the model and its outcomes could be created. However, in a future case at ING or anywhere else, the lessons learned from this research could be taken into account and this research can be shown to convince stakeholders of the potential benefits the approach could have to their business process. This would trigger more stakeholder interest which in turn could improve model confidence increasing the potential success of the approach.

From a technical point of view it is very desirable to develop an integrated tool that will make Exploratory Modeling and Analysis based on DES models easier and more accessible. By doing so the iteration cycles for experiments and feedback can be significantly shorter which again increases stakeholder involvement. Furthermore, exploratory potential can be improved as more models or specific areas in the solution space can be examined in the same time span. Until an integrated tool is developed for the purpose of applying EMA on DES, I would not recommend using this approach to anyone in practice as it is very time consuming, technically unreliable, and not user friendly.

10. References

Balci, O. (1998). Verification, validation, and testing. In J. Banks (Ed.), *Handbook of simulation* (pp. 335-393). New York: John Wiley & sons, inc.

Bankes, S. (1992). Exploratory Modeling and the use of Simulation for Policy Analysis *N-3093-A*. Santa Monica: RAND.

Bankes, S. (1993). Exploratory Modeling for Policy Analysis. *Operations Research, 41*(3), 435-449.

Banks, J. (1998). *Handbook of Simulation*. New York: John Wiley & sons, inc.

BCBS. (2009). Strengthening the resilience of the banking sector. Basel: Bank for International Settlements.

BCBS. (2013). Basel Committee on Banking Supervision Charter. Basel: Bank for International Settlements.

Belinga, P. B. (2013). *Economic and Environmental Assessment of Extended Producer Responsibility and No Reuse Policies in Mobile Phone End of Life Networks.* (MSc.), Delft University of Technology, Delft.

Ben-Tal, A. & Nemirovski, A. (2002). Robust optimization–methodology and applications. *Mathematical Programming, 92*(3), 453-480.

Bishop, P., Hines, A. & Collins, T. (2007). The current state of scenario development: an overview of techniques. *Foresight-The journal of future studies, strategic thinking and policy, 9*(1), 5-25.

Blundell-Wignall, A. & Atkinson, P. (2010). Thinking beyond basel iii: Necessary solutions for capital and liquidity. *OECD Journal: Financial Market Trends, 2010*(1), 5-6.

Boersma, S. K. T. & Hoenderkamp, T. (1981). *Simulatie: een moderne methode van onderzoek*. Schoonhoven: Academic Service.

Bryant, B. P. & Lempert, R. J. (2010). Thinking inside the box: A participatory, computer-assisted approach to scenario discovery. *Technological Forecasting and Social Change, 77*(1), 34-49.

Chong, I.-G. & Jun, C.-H. (2008). Flexible patient rule induction method for optimizing process variables in discrete type. *Expert Systems with Applications, 34*(4), 3014-3020.

[redacted]

de Vreede, G.-J., Verbraeck, A. & van Eijck, D. T. (2003). Integrating the conceptualization and simulation of business processes: A modeling method and arena template. *Simulation, 79*(1), 43-55.

Epstein, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation, 11*(4), 12.

Friedman, J. H. & Fisher, N. I. (1999). Bump hunting in high-dimensional data. *Statistics and Computing, 9*(2), 123-143.

Hamarat, C., Kwakkel, J. H. & Pruyt, E. (2013). Adaptive Robust Design under deep uncertainty. *Technological Forecasting and Social Change, 80*(3), 408-418.

Hevner, A. R. (2007). The three cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(2), 87 - 92.

Hevner, A. R., March, S. T., Park, J. & Ram, S. (2004). Design science in Information Systems research. *Mis Quarterly, 28*(1), 75-105.

Hoover, S. V. & Perry, R. F. (1989). *Simulation: a problem-solving approach*. Reading, Massachusetts: Addison-Wesley Longman Publishing Co., Inc.

Jahangirian, M., Eldabi, T., Naseer, A., Stergioulas, L. K. & Young, T. (2010). Simulation in manufacturing and business: A review. *European Journal of Operational Research, 203*(1), 1-13.

Jansen-Vullers, M. & Netjes, M. (2006). *Business process simulation–a tool survey.* Paper presented at the Workshop and Tutorial on Practical Use of Coloured Petri Nets and the CPN Tools, Aarhus, Denmark.

Jungman, N., Van Geuns, R., Klaver, J., Wesdorp, P. & van der Wolk, J. (2012). Op weg naar effectieve schuldhulp, Preventie: Voorkomen is beter dan genezen. Utrecht: Hogeschool Utrecht.

Kuechler, B. & Vaishnavi, V. (2008). On theory development in design science research: anatomy of a research project. *European Journal of Information Systems, 17*(5), 489-504.

Kwakkel, J. H. (2011). Exploratory Modelling and Analysis (EMA) Workbench. Retrieved 27-5-2013, 2013, from http://simulation.tbm.tudelft.nl/ema-workbench/contents.html

Kwakkel, J. H. & Pruyt, E. (2013). Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change, 80*(3), 419-431.

Law, A. M. & Kelton, D. W. (1982). *Simulation modeling and analysis*. New York: McGraw-Hill.

Lempert, R. J., Groves, D. G., Popper, S. W. & Bankes, S. C. (2006). A general, analytic method for generating robust strategies and narrative scenarios. *Management science, 52*(4), 514-528.

Lempert, R. J., Popper, S. W. & Bankes, S. C. (2003). *Shaping the next one hundred years: new methods for quantitative, long-term policy analysis*. Santa Monica: RAND.

Li, Z. & Ierapetritou, M. (2008). Process scheduling under uncertainty: review and challenges. *Computers & Chemical Engineering, 32*(4), 715-727.

Mulvey, J. M., Vanderbei, R. J. & Zenios, S. A. (1995). Robust optimization of large-scale systems. *Operations Research, 43*(2), 264-281.

Neelamkavil, F. (1987). *Computer simulation and modelling*. Chichester: John Wiley & Sons Ltd.

Palisade Corporation. (2010). *Guide to Using @Risk*. Ithaca, NY USA: Palisade Corporation.

Pruyt, E. & Hamarat, C. (2010). The influenza A (H1N1) v pandemic: an exploratory system dynamics approach *18th International Conference of the System Dynamics Society*. Seoul.

Pruyt, E. & Kwakkel, J. (2012). *A Bright Future for System Dynamics: From Art to Computational Science and Beyond.* Paper presented at the System Dynamics Conference, at St. Gallen, Switzerland.

Pruyt, E., Kwakkel, J., Yucel, G. & Hamarat, C. (2011). *Energy transitions towards sustainability: A staged exploration of complexity and deep uncertainty.* Paper presented at the Proceedings of the 29th International Conference of the System Dynamics Society, Washington DC.

Sage, A. & Armstrong, J. (2000). *Introduction to Systems Engineering.* New York: John Wiley & Sons.

Sargent, R. G. (1996). *Verifying and validating simulation models.* Paper presented at the Proceedings of the 28th conference on Winter simulation.

Schlesinger, S., Crosbie, R. E., Gagne, R. E., Innis, G. S., Lalwani, C., Loch, J., . . . Bartos, D. (1979). Terminology for model credibility. *Simulation, 32*(3), 103-104.

Shannon, R. E. (1975). *Systems simulation: the art and science*. New Jersey: Prentice-Hall Englewood Cliffs.

Shannon, R. E. (1981). *Tests for the verification and validation of computer simulation models*. Paper presented at the Proceedings of the 13th conference on Winter simulation-Volume 2.

Verbraeck, A. & Valentin, E. C. (2006). *Discrete modellen Deel II Discrete Simulatie*. Delft: TU Delft.

Witten, I. H. & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. San Francisco: Morgan Kaufmann Publishers.

Zouaoui, F. & Wilson, J. R. (2003). Accounting for parameter uncertainty in simulation input modeling. *Iie Transactions, 35*(9), 781-792.

11. Appendices

Appendix A: ING mortgage arrears risk costs calculation

[redacted]

[redacted]

Figure A1: Risk costs based on provisions and losses, adapted from [redacted].

Appendix B: Monte Carlo Sampling vs. Latin Hypercube Sampling

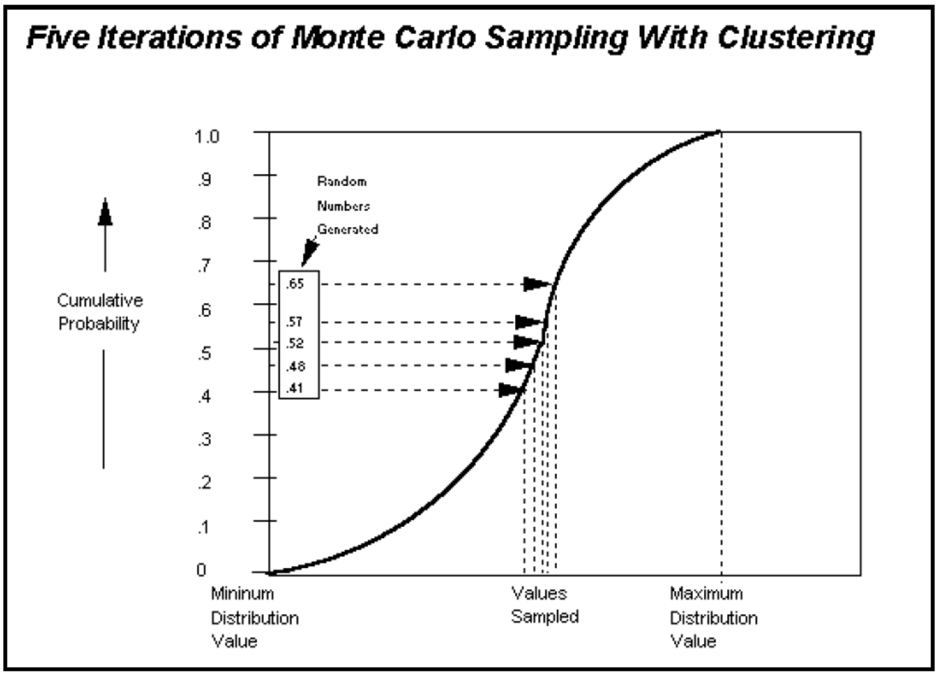

Figure B1: Cumulative probability distribution example with five samples based on MCS, adopted from Palisade Corporation (2010, p. 650)

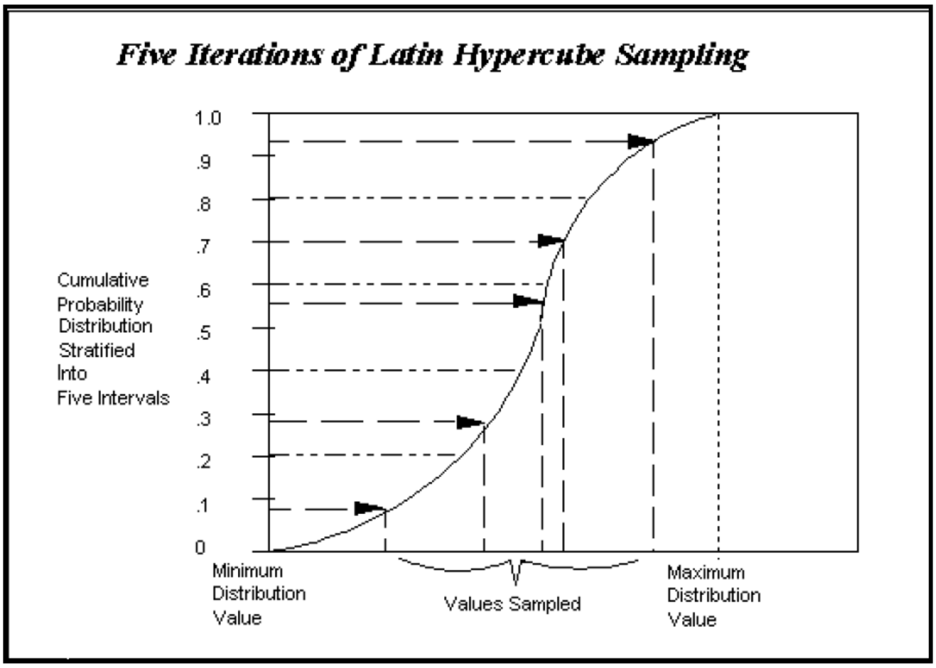

Figure B2: Cumulative probability distribution example with five samples based on LHS, adopted from Palisade Corporation (2010, p. 651)

Appendix C: Detailed description of differentiation groups

	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Table C1: Detailed description of customer type and process aim for all willingness/ability groups.

Appendix D: Specification of ING Arrears Management simulation model

In paragraph 5.2, 5.3, and appendix E the data for uncertain variables and other variables values are discussed. This appendix adds more detailed information to the specification chapter including model structure and model logic regarding the simulation model. Figure D1 depicts a helicopter view of the entire simulation model.

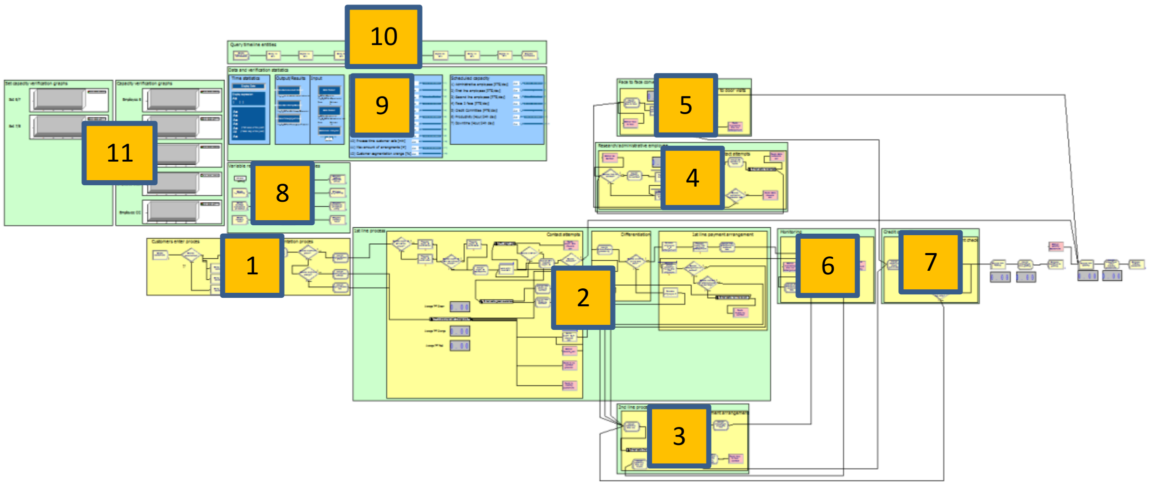

Figure D1: Complete overview of detailed aggregation simulation model

The numbers in figure D1 correspond to various sub processes and other model elements:

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Appendix E: Specification of input data

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

Appendix F: PRIM results comparison 1 vs. 10 replications

The PRIM results in this appendix are based on average throughput time for red customer smart calls. All experimental design elements were similar except for the number of replications.

[redacted]

Figure F1: Dim plot of average throughput time for red customer smart calls based on 10 replications

[redacted]

[redacted]

Table F1: Table of average throughput time for red customer smart calls based on 10 replications

[redacted]

Figure F2: Dim plot of average throughput time for red customer smart calls based on 1 replication

[redacted]

[redacted]

Table F1: Table of average throughput time for red customer smart calls based on 10 replications

It is important to note that the results in this appendix are based on an earlier, more aggregate version of the simulation model and that the assumption is made that we can apply the conclusions on the final, more detailed simulation model.

Based on a qualitative comparison of both PRIM analysis results presented in this appendix, it can be concluded that the same danger zones are found [redacted]. Furthermore, the detailed ranges corresponding to the danger zone seems largely similar.

Appendix G: Detailed PRIM peeling trajectory analysis of average throughput time for no contact process

This appendix shows the detailed box finding approach for average throughput time for no contact process including the peeling trajectories (figure G1) and the specific choices for number of peelings (highlighted in yellow in table G1).

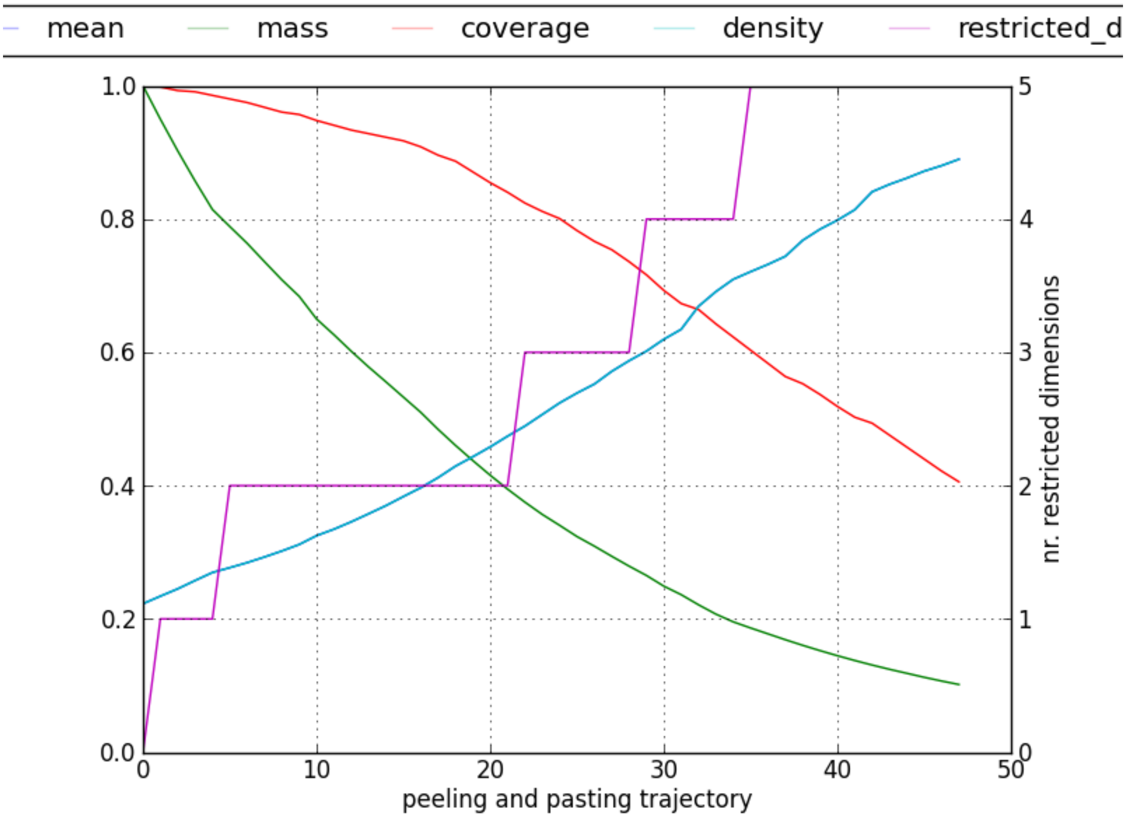

Figure G1: Visualization of peeling trajectory box 1 average throughput time no contact process.

box	mean	mass	coverage	density	res dim
0	0.22	1	1	0.22	0
1	0.23	0.95	1	0.23	1
2	0.25	0.9	0.99	0.25	1
3	0.26	0.86	0.99	0.26	1
4	0.27	0.81	0.99	0.27	1
5	0.28	0.79	0.98	0.28	2
6	0.28	0.76	0.97	0.28	2
7	0.29	0.74	0.97	0.29	2
8	0.3	0.71	0.96	0.3	2
9	0.31	0.68	0.96	0.31	2
10	0.32	0.65	0.95	0.32	2
11	0.33	0.63	0.94	0.33	2
12	0.35	0.6	0.93	0.35	2
13	0.36	0.58	0.93	0.36	2
14	0.37	0.56	0.92	0.37	2
15	0.38	0.53	0.92	0.38	2
16	0.4	0.51	0.91	0.4	2
17	0.41	0.48	0.9	0.41	2

18	0.43	0.46	0.89	0.43	2
19	0.44	0.44	0.87	0.44	2
20	0.46	0.42	0.85	0.46	2
21	0.47	0.39	0.84	0.47	2
22	0.49	0.38	0.82	0.49	3
23	0.51	0.36	0.81	0.51	3
24	0.52	0.34	0.8	0.52	3
25	0.54	0.32	0.78	0.54	3
26	0.55	0.31	0.77	0.55	3
27	0.57	0.29	0.75	0.57	3
28	0.59	0.28	0.74	0.59	3
29	0.6	0.27	0.72	0.6	4
30	0.62	0.25	0.69	0.62	4
31	0.63	0.24	0.67	0.63	4
32	0.67	0.22	0.66	0.67	4
33	0.69	0.21	0.64	0.69	4
34	0.71	0.2	0.62	0.71	4
35	0.72	0.19	0.6	0.72	5
36	0.73	0.18	0.58	0.73	5
37	0.74	0.17	0.56	0.74	5
38	0.77	0.16	0.55	0.77	5
39	0.78	0.15	0.54	0.78	5
40	0.8	0.14	0.52	0.8	5
41	0.81	0.14	0.5	0.81	5
42	0.84	0.13	0.49	0.84	5
43	0.85	0.12	0.48	0.85	5
44	0.86	0.12	0.46	0.86	5
45	0.87	0.11	0.44	0.87	5
46	0.88	0.11	0.42	0.88	5
47	0.89	0.1	0.41	0.89	5

Table G1: Numerical peeling trajectory for box found in average throughput time no contact process.

The peeling choices made in the trajectory depicted above results in the following box identification shown in figure G2 and table G2.

Figure G2: Dimplot of 2 box found based on classification for average throughput time of no contact process.

[redacted]

Table G2: Numerical representation of 2 boxes found based on classification for average throughput time no contact process.

The demonstration of peeling trajectory and choices made in this appendix have been done for all identified PRIM boxes.

Appendix H: Complete results KDE experiment for required number of scenarios based on 5.000 scenario

This appendix contains the results of all KDE plots and mean/standard deviation plots based on six KPI's (two KPI's per KPI type):

1. Throughput time
    a. Throughput time for customers from arrear to financially healthy
    b. Throughput time for red customers to be called (smart calls)
2. Scheduled utilization
    a. Scheduled utilization research employees
    b. Scheduled utilization first line employees
3. Number of customers in queues
    a. Inbound calls queue
    b. Second line outbound calls queue

Considering the results shown in this appendix, trade-off between minimization of computing time on the one hand and maximizing reliability and maximizing the number of scenarios for the purpose of PRIM analyses on the other hand, 2.500 scenarios seems to be a good amount of scenario runs to use in experiments.

Figure H1a depicts a KDE plot of throughput time for customers from arrear to financially healthy and figure H1b shows Mean and standard deviation plot of throughput time for smart calling red customers.

Figure H1a

Figure H1b

[redacted]

Figure H1c: KDE plot of throughput time for smart calling red customers top zoom

[redacted]

Figure H1c: KDE plot of throughput time for smart calling red customers bottom zoom

Figure H2a depicts a KDE plot of throughput time for customers from arrear to financially healthy and figure H2b shows Mean and standard deviation plot of scheduled utilization of research employees.

Figure H2a

Figure H2b

Figure H2c: KDE plot of scheduled utilization of research employees top zoom

[redacted]

Figure H2d: KDE plot of scheduled utilization of research employees bottom zoom

Figure H3a depicts a KDE plot of throughput time for customers from arrear to financially healthy and figure H3b shows Mean and standard deviation plot of scheduled utilization of first line employees.

[redacted] [redacted]

Figure H3a Figure H3b

Figure H3c: KDE plot of scheduled utilization of first line employees top zoom

Figure H3d: KDE plot of scheduled utilization of first line employees bottom zoom

Figure H4a depicts a KDE plot of throughput time for customers from arrear to financially healthy and figure H4b shows Mean and standard deviation plot of customers in inbound calls queue.

Figure H4a

Figure H4b

Figure H4c: KDE plot of customers in inbound calls queue top zoom

Figure H4d: KDE plot of customers in inbound calls queue bottom zoom

Figure H5a depicts a KDE plot of throughput time for customers from arrear to financially healthy and figure H5b shows Mean and standard deviation plot of customers in second line outbound calls queue.

Figure H5a

Figure H5b

Figure H5c: KDE plot of customers in second line outbound calls queue top zoom

Figure H5d: KDE plot of customers in second line outbound calls queue bottom zoom

Appendix I: Model behavior graphs of throughput times for processes

[redacted]

[redacted]

Figure I1: Average throughput time for second line process

[redacted]

Figure I2: Average throughput time for credit committee process

[redacted]

Figure I3: Average throughput time for complete process. [redacted]

[redacted]

Figure I4: Average throughput time for no contact / research process

[redacted]

Figure I5: Average throughput time for smart calling green, red, and orange customers (depicted in green, red and orange lines respectively).

Appendix J: Detailed PRIM Analysis results

This appendix holds a detailed PRIM analysis including identified boxes for all experiments. The identified boxes are visualized in dim-plots. For more details, please refer to the numerical values for all boxes and uncertainty ranges as presented per PRIM analysis. Furthermore, if necessary verification of boxes has been conducted through scatter-plot analysis for the purpose of avoiding conclusions based on over fitting. All boxes have been identified based on the following classification of KPI's

[redacted]

All cases that meet the threshold values stated above are classified as a case of interest. The results of the following experiments will be discussed in order:

- Experiment 1: Uncertain input parameters and uncertain capacities
- Experiment 2: Uncertain input parameters
- Experiment 3: Fixed values for customer inflow, segmentation and differentiation
- Experiment 4: Priority of second line tasks over inbound calls for second line employees

Appendix J.1. Results experiment 1: Uncertain input parameters and uncertain capacities

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Appendix J.2. Results experiment 2: Uncertain input parameters

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Appendix J.3. Results experiment 3: Fixed values for customer inflow, segmentation and differentiation

[redacted]

Appendix J.4. Results experiment 4: Priority of second line tasks over inbound calls for second line employees

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Appendix K: Addition of feedback loops to proposed approach

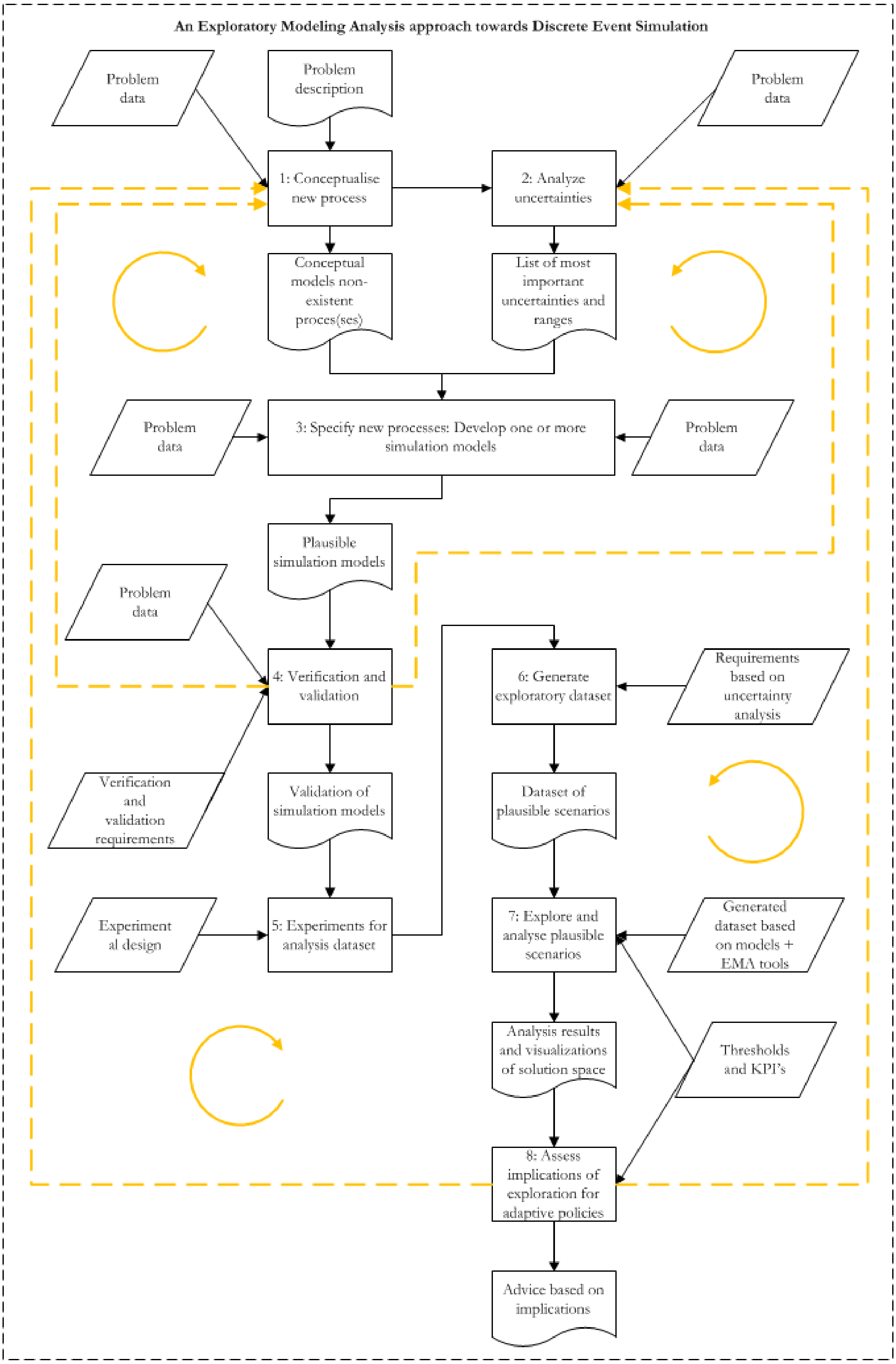

Figure K1: Detailed research design visualization including feedback loops adapted from Verbraeck and Valentin (2006).