SEVENTH FRAMEWORK PROGRAMME OF THE EUROPEAN COMMUNITY
(EC GRANT AGREEMENT N° 288819)

INTELLIGENT SERVICES FOR ENERGY-EFFICIENT DESIGN AND LIFE CYCLE SIMULATION

Deliverable D2.1: Overall Stochastic Approach for the Virtual Energy Lab Platform

Responsible Authors:

Amin Zahedi Khameneh, Raimar J. Scherer, Gudni Gudnason

Due date: 00.12.2012
Issue date: 15.01.2013
Nature: public

Start date of project: **01.12.2011** Duration: **36 months**

Organisation name of lead contractor for this deliverable:

TRIMO INZENIRING IN PROIZVODNJA MONTAZNIH OBJEKTOV, D.D.

History

Version	Description	Lead Author	Date
	Initial report structure	TUD-CIB	7.09.2012
0.1	First Draft	TUD-CIB, NMI	20.11.2012
0.2	Second Draft	All involved partners	14.12.2012
0.9	Pre-final released	TUD-CIB	10.1.2013
1.0	**Checked and approved final version**	Coordinator	15.01.2013

Copyright

Citation

Zahedi Khameneh A. ,Scherer R. J., Gudnasson G.(2012): ISES Deliverable D2.1, Overall Stochastic Approach for the Virtual Energy Lab Platform, © ISES Consortium, Brussels.

Acknowledgements

The work presented in this document has been conducted in the context of the seventh framework programme of the European community project ISES (n° 288819). ISES is a 36-month project that started in December 2011 and is funded by the European Commission as well as by the industrial partners. Their support is gratefully appreciated.

The partners in the project are TECHNISCHE UNIVERSITÄT DRESDEN (Germany, Coordinator), OLOF GRANLUND OY (Finland), UNIVERZA V LJUBLJANI (Slovenia), SOFISTIK HELLAS A.E. (Greece), NYSKOPUNARMIDSTOD ISLANDS(Iceland), NATIONAL OBSERVATORY OF ATHENS (Greece), LEONHARDT ANDRÄ UND PARTNER BERATENDE INGENIEURE VBI GMBH (Germany) and TRIMO INZENIRING IN PROIZVODNJA MONTAZNIH OBJEKTOV, D.D. (Slovenia). This report was created through the joint effort of the above organisations.

Project of the SEVENTH FRAMEWORK PROGRAMME OF THE EUROPEAN COMMUNITY		
Dissemination Level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

TABLE OF CONTENTS

Executive Summary

The **objective** of WP2 of ISES is to (1) set up the overall cloud-based ISES platform architecture, including data preparation via local CAD and FM systems as well as distributed and local databases for climate data, user profiles and product catalogues, (2) specify the component services and their interrelations to prepare for the needed harmonized APIs that will be developed in WPs 3-7, (3) define the prerequisites for multiple parallel simulation runs on the cloud, thereby setting the basis for the RTD work in WP7, specifically dedicated to the cloud environment, (4) develop the overall stochastic approach for the Virtual Energy Lab platform in alignment with the objectives of the project and the technical platform architecture. The platform will take into account both remote web services, especially the services for life cycle energy, CO2 and cost simulations executed on a cloud, and the local CAD, FM and product catalogue systems that will be used in ISES.

This deliverable covers the overall work performed in WP2 within task T2.2development and specification of the overall stochastic approach. During this task the baseline for the detailed stochastic approach in WPs 4, 5 and 6 is set up. The stochastic processes involved in the envisaged energy, emissions and cost simulations are analyzed from systemic point of view in order to develop a pragmatic, manageable treatment of the stochasticity of the product life-cycle. The developed stochastic approaches are described in concise form, highlighting its innate features. Special attention is put on the use of material properties, climate models and energy consumption profiles.

The **involved partners** are:

- TUD-CIB – lead, development of the overall stochastic approach for the Virtual Energy Lab, structure and editing of the deliverable report
- OG – usage / occupancy profiles
- SOF – stochastic input
- NMI - overall stochastic approach with regard to the coordination and interoperability with WP4 (information framework)

1. Introduction

Regarding the ISES outlook to evaluate, simulate and optimize the energy efficiency of facilities and building energy performance, considering the stochastic approach in a realistic assessment through the product life-cycle becomes necessary. Due to the largely random nature of the material properties, climate/weather data and the usage profile over the lifecycle, standard worst case evaluations of life cycle cost not only for the component product but for the interaction with the host systems has to be carried out during design. The stochastic nature of the overall life-cycle has to be approximated by a stochastic discrete process of possible sequences of characteristic energy patterns and profiles. This can be achieved during simulation of a large number of combinations of possible stochastic energy patterns and profiles. Such simulation tasks may require hundreds of individual simulations run in parallel in a cloud environment, with target-oriented feedbacks between evaluated and further simulations that cannot be anymore configured by hand but have to be managed highly automatically, with only general control interaction by the user.

The stochastic simulation process and its fundamental components as well as sampling methods are discussed in the following chapter. In chapter three, several stochastic variable types, the stochastic variables which are considered in stochastic simulation as well as providing the proper samples size and treatment with the simulation outcomes will be elaborated.

2. Stochastic Simulation

2.1 Uncertainties Encountered in Buildings Energy during Life-Cycle

Forecasting the building energy demand is a complicated task. In addition to primary models, which characterize building systems and components, detailed information about the **building material**, **HVAC systems**, **weather** and **user behavior** must be taken into consideration. The interaction between the weather conditions and building operations on the one side, and the impact of multiple building characteristics on the other side call for the use of sophisticated simulations – to facilitate design and operation for better building performance. Moreover, significant deviations in terms of building energy consumption between measured performance and model-predicted results at design stage are reported for low-energy buildings [Turner and Frankel 2008]. The current practice for energy load calculations have traditionally focused on determining energy consumption of buildings by a prescriptive approach as detailed in national standards and regulations and thereby complying with set requirements. Similarly, design tools for calculation of energy consumption of buildings have been more focused on calculating energy loads for dimensioning of heating, cooling and air conditioning systems [Hopfe 2009]. Usually these tools are based on static (deterministic) calculation methods, applied at the later design stages of the building and leaving little opportunity for design optimisation in the energy efficiency of the design.

Despite these efforts, simulation results are obtained based on a number of basic assumptions about the simulation model and the influencing factors e.g. climate, building properties and occupant behaviour that cannot be realistically replicated and the associated uncertainties quantified. Many of the input parameters are depended on discreetness, non-linearity, uncertainty or variability [Hopfe 2009] and depend on many varying factors both dependent and independent of one another. In such cases average values can only give a reasonable estimation of the actual values especially given the complexity of the context in which the object is being simulated. Uncertainty, inaccuracy and errors in input parameters have raised concerns in the literature as these propagate through the simulation model resulting in inaccuracy and uncertainty in the simulation output [Fabiet al.2011] which is a well known fact in computational engineering, namely “rubbish in-rubbish out” and means numerical approach must not be more sophisticated than the input knowledge about the input parameters and hence the input model is. In practice, the uncertainty analysis has also the benefit that by changing the input of the parameters and showing the effect on the outcome of a model, it provides a “what-if analysis”. It is for instance used in multiple decision support tools [Gokhale2009].

An estimate of the degree of uncertainties contributed from each factor is of importance to improve the robustness of simulation models and help the modeler and customer have a better understanding of building simulation results. Several important research efforts focused on the investigation of uncertainties in input parameters for building design support. However, a review of the literature shows there are limited data available describing uncertainties for design parameters in building simulation. [Wang et al. 2012] have categorized the major uncertainties in building energy in two fundamental groups:

- Uncertainties in annual energy use due to weather variation
- Uncertainties in annual energy use due to building operations

Besides the two major uncertainty sources in building energy simulation,

- Uncertainties in material should also be considered.

[MacDonald and Strachan2001] applied a Monte Carlo uncertainty analysis for thermal properties of construction materials, weather, internal heat gains, and infiltration rate to evaluate the variation of energy consumption using assumed uncertainty distribution patterns.[Holm and Kuenzel2002] evaluated the impacts of materials properties and surface coefficients on hydrothermal building simulation using a Monte Carlo analysis.

As it was previously elaborated during D1.2 [Manspergeret al. 2012], the solutions provided by ISES are focused on all possible scenarios in the life-cycle of buildings and facilities:

(1) Development of new building components and products
(2) Design and engineering of new buildings and facilities
(3) Refurbishment and retrofitting of existing buildings and facilities

The simulation outputs resulting from the different scenarios in the life-cycle of buildings and facilities are the probabilistic components. In building component product development and design of facilities, almost all design parameters are subject to uncertainty. The stochastic approach addresses all the categories of the parameters space, climate profiles, solar gain, usage profiles, and energy related building material properties.

Combining building operations scenarios can yield significantly different building energy consumption results should any of these scenarios change. In fact, there is not a unique or exact answer; we are faced with a wide range of possible performance outcomes. Applying the stochastic modelling of several uncertain input components involved in energy consumption, will lead to a wide range of results, from which the favour statistical indices can be extracted.

2.2 Stochastic Simulation Process

Stochastic simulation has been studied in numerous researches, recognizing the necessity of applying stochastic methods in Building Performance Simulation (BPS) [MacDonald 2002, De Wit 2001, Hopfe 2009, Jacobs 2011]. General findings are in agreement that stochastic methods generate different results from traditional deterministic analyses methods and deliver more valuable design information and support a more robust decision-making process in design.

ISES provides a platform in which different existing simulation software can be integrated to run in a grid-based (cloud based) environment. In this case, classical deterministic simulation models are used, but treating the input parameter space as a separate stochastic model defining the building physical properties, system solutions as well as the initial and boundary conditions of the buildings context. Monte Carlo techniques have successfully and widely been deployed in the area of BPS [Kim 2011]. Although Monte Carlo is refer to a specific sampling method, it always be used as a generalization of stochastic simulation (In this deliverable, we will also use the term “Monte Carlo” generally as stochastic analysis approach, except where it is called as specific sampling method). The technique has proven suitable for integrating both uncertainty and sensitivity analyses with current deterministic simulation tools such as EnergyPlus, DOE 2x, ESP-r etc. However, the efficiency and robustness of the stochastic simulation can be improved by applying the reduced variance sampling approaches like Latin Hypercube Sampling.

The Monte Carlo simulation is generally described as systematically "inverts" simulation, treating deterministic problems by first finding a probabilistic analogy. Common methods of simulation and statistical sampling generally did the opposite: using simulation to test a previously understood deterministic problem. Though examples of an "inverted" approach do exist historically, they were not considered a general method until the popularity of the Monte Carlo method spread.

Monte Carlo Simulation is a technique used to determine the probabilistic distribution of an outcome that relies on all probable scenarios and it produces not only one answer, but rather a series of answers or a range over which the results vary as a function of probability of occurrence and also a most expected result. The answer may fall anywhere within the range of the results produced.

In classic Monte Carlo analysis it is assumed that all inputs are independent of each other. Thus, when the variables are correlated a correction to the Monte Carlo Simulation is required. To account the correlation, a correlation matrix in the Monte Carlo Simulation is used. The correlation value for each set of variables is entered in the matrix, and when the simulation is run, correlation among the variables is accounted for.

This process is repeated many times so as to provide enough information to construct a probability distribution of the model output. The stochastic simulation process can be divided into the (1) Pre-processing, (2) Simulation and (3) Post-processing (Figure 1). Following, the simulation steps are discussed briefly.

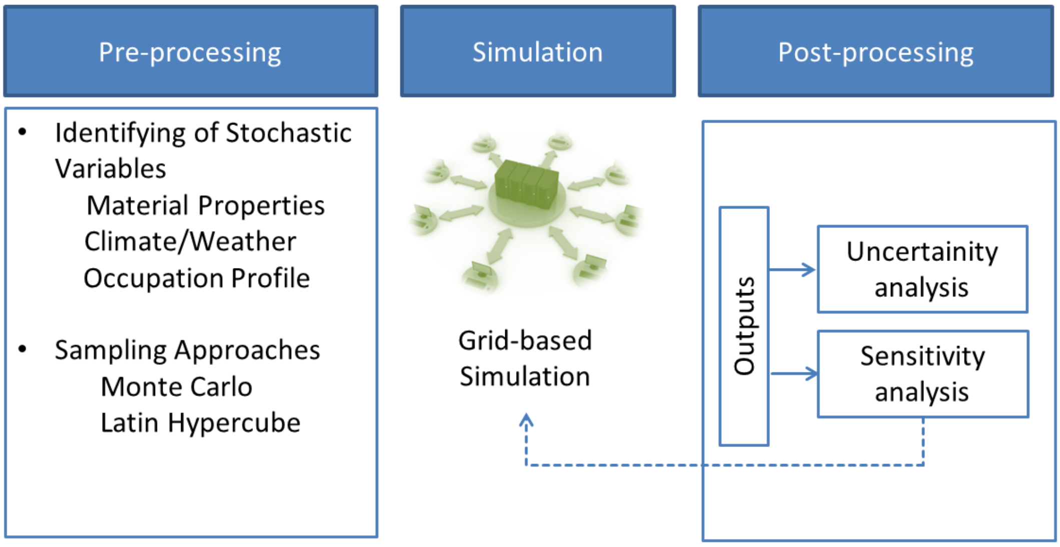


Figure 1: Illustration of the stochastic simulation process divided into the three steps pre-processing, simulation, and post-processing

Pre-Simulation

Identifying the critical aspects of the design and design alternatives and decision on performance metrics and design objectives are performed during pre-processing step. The first step of uncertainty analysis is to determine the uncertainty in estimates of design variables of importance. This entails two major steps;

(1) Identification of stochastic variables, their ranges and scope, appropriate probability distribution function (PDF) (mean and standard deviation) are selected for each stochastic variable. These can include variables with missing or incomplete design information or specification, uncertainty relating to accuracy of parameter values, parameters subject to discrete random events etc. The stochastic model types and stochastic variables which will be considered in ISES are mentioned and discussed in chapter 3.

(2) Identification appropriate sampling methods. The purpose is to get an unbiased estimate for the design variable from the PDF representing the variable population mean. The most popular sampling approaches which are commonly used in stochastic simulation process are described in detail in section 2.3.

Since the correlation or dependency of the variables on each other may influence the reliability of the simulation results, the dependency of stochastic input variables should also be studied.

Typically simulation software uses several input files to collect and combine related relevant input values representing the simulation model. These need to be constructed from the samples obtained during the uncertainty analyses step. For each sample vector X=[xi_1, xi_2…xi_k] (i= 1,..,N, N is sample size and k is number of independent variables) one set of input files are generated e.g. if design variables where sampled for sample size of 100 elements. 100 sets of input files need to be generated and simulated.

Simulation

To perform simulation with the defined uncertainties, the creation of a multiple stochastic simulation controller has required, which affects the necessary changes in the data model for each parametric variation and initiates the simulations. Simulation in Energy Enhanced Building (eeB) usually is performed with hourly time steps extending over a period of **one year** for each simulation. The simulation algorithm is therefore executed 8760 times (24 times365). In Monte Carlo simulation the number of executions of the algorithm increases in direct proportional to the number of samples selected for variables i.e. variable sample size times 8760. Given a sample size of 100 each simulation will require 876000 executions. If we want to investigate extreme values and/or correlation, we will fastly reach or exceed 1000 samples. Minimizing the sample size and number of stochastic parameters to be analysed is therefore a priority to save computer effort.

[Macdonald and Strachan 2001] suggest that the controller should read the input data model into memory and all subsequent changes to the model are made there. In order to avoid the corruption of the input data model this can be referenced between each simulation, prior to data manipulation. Before the simulations are commissioned, the total number required is calculated. This is straightforward after the analysis method has been chosen. After running the simulation the simulation controller references the model uncertainty file and, using information held there, changes one or more parameters in the model (depending on the analysis method chosen).

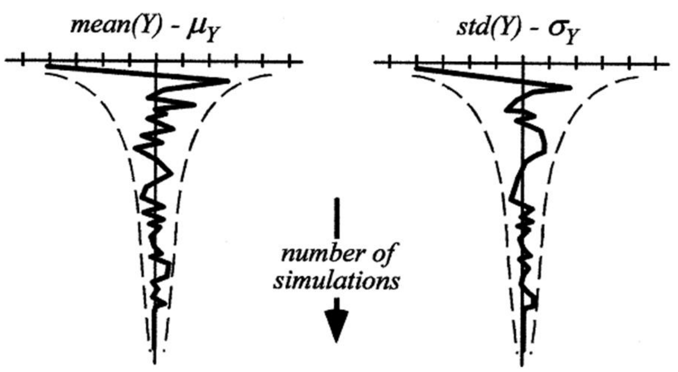
Post-Simulation

Once the model evaluations have been performed a post-simulation step collects all the results from the multiple simulations and therefore requested uncertainty and sensitivity analysis can be performed. In this context, the uncertainty analyses typically come before Sensitivity Analyses. Uncertainty Analyses determine the uncertainty in simulation outputs derived from uncertainties of the input variables. The results are collected for each measured output variable and summaries are usually presented in the form of means, variances and probability distribution functions. Based on the results, if uncertainties lie outside acceptable tolerances or some variables are seen to have greater influences on the model outputs. The stochastic simulation has to be partially repeated with accordingly adjusted values. Further on a sensitivity analyses can be performed. Sensitivity analyses will determine how sensitive model outcome is to changes in the model inputs, by analysing the mappings between every output variable and input variables. It can be concluded that the uncertainty arising from different sources are propagated through the model for uncertainty analysis and their relative importance is quantified via sensitivity analysis.

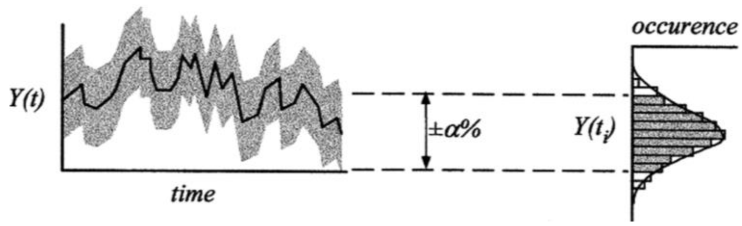
Uncertainty Analyses

In Energy simulation we generally assume that the composition of the model being analysed is usually well known and understood in a deterministic way, i.e. the same input data will produce the same outcome for any consecutive runs of the simulation model. For example we assume that the geometry of the building has been verified during the design process using appropriate design tools and that building component types, their composition and properties are known for a particular design solution. However many of the building systems are numerical and algorithmic approximations when employed in the simulation models that can have significant variation and uncertainties embedded compared to the real system.

With the uncertainty analyses the objective is to assess and quantify the uncertainty in the model outcomes that derives from uncertainty in the input parameters whereby the method gives information on how reliable and confident the simulation outcome is [Hopfe 2009, Eisenhower et al.2010, Fabiet al.2011]. Uncertainty analyses explores the mapping between uncertain output results $Y(X) = [y_1(X), y_2(X), \ldots ,Y_n(X)]$ as a function of uncertain input parameters $X=[x_1, x_2, \ldots , X_n]$ exploring what is the uncertainty in Y(X) given the uncertainty in X [Helton 2006]. The uncertainty of input parameters X is expressed in terms of probability distribution functions (e.g. Normal, lnNormal and etc.), but can also be specified by samples of measured values, i.e. empirical probability distributions (Figure 2).



a) convergence of the mean and standard deviation within confidence interval



b) time series resulting from successive Monte-Carlo simulations

c) histogram of one simulation corresponding the one time step

Figure 2: Stochastic simulation, In (a) the convergence of the mean and standard deviation is checked by plotting the statistics obtain after each simulation (b) the confidence interval around the time-series of an output Y(t) is estimating using Monte Carlo technique, to each time step corresponds a set of results which can be analysed with histogram (c)[from Roulet, 1999]

With many input parameters generating the probability distributions can present a major effort and costs. As part of the analyses strategy is to do initial exploratory analyses using fairly basic definitions of probability distributions and identify the most important input parameters. Further definition on the identified input parameters distribution can then be done with a second analyses step using refined probability distribution.

In simple Monte Carlo analyses generally it is assumed that input parameters are not related and independent of one another that is, there exist no correlation between them. Since this is not the case for many input parameters e.g. moisture and temperature, outdoor and indoor temperature. There may be a need to induce a desired correlation structure onto the samples being generated [Ekström 2005]. Specifically correlated variables should have correlation closed to their specified value and uncorrelated variables should have correlation close to zero [Helton 2006]. The Rank Order correlation method by Iman and Conover [Ekström 2005, Helton 2006] is available in many statistical software applications to impose rank correlation on samples. The Rank order correlation method is a non-parametric statistic for quantifying the correlation between two variables where the correlation statistics is not affected by the type of mathematical relationship between the variables. Furthermore the method is applicable where complex correlation structures and where multiple input parameters are involved in the analyses and the method is independent of sample distributions and works well with different sampling methods.

Sensitivity Analysis

By systematically changing the input variables, sensitivity analysis is used to investigate the relationship between input variables and output variables and address which input variable has larger contribution on output variables. Thus, the sensitivity analysis can determine the following questions [Rackwitz and Fiessler 1978)]: which input variable is more critical compared to others and need additional knowledge on them? With this answer prior to modelling the precise of the significant variables can be improved to reduce the output uncertainty; which input variable has little contribution to the model? With this answer the insignificant variable can be fixed in the nominal value to simplify the model. Is the model performed in the proper way? If the model is sensitive to some non-influential variables judged by expert experience the chosen range of the variable or the model structure need to be further examined. The application of sensitivity analyses depends on the context and nature of the investigation. Sensitivity analyses can provide a general evaluation of the model precision when used for evaluating model performance indicators in alternative simulation scenarios or for detailed study in the significance and interaction of individual input parameters.

There are several approaches available to perform the sensitivity analysis. Each approach has its own capability and applicability. To facilitate the decision-maker to select the most appropriate method, the sensitivity analysis approaches are broadly classified into three categories: (1) mathematical methods, (2) statistical (or probabilistic) methods and (3) graphical methods [Frey and Patil 2002].

(1) Mathematical methods evaluate the impact of the range of variation of input variable on the output variable [Morgan and Henrion 1990; Frey and Patil 2002]. Mathematical methods include nominal range sensitivity, differential sensitivity analysis, etc. The method typically assesses the sensitivity of the output to a few values of the input variable and mostly valid for the linear model [Frey and Patil 2002]. When applied to the non-linear model the result of the evaluation could be misleading.

Differential Sensitivity Analysis (DSA) is one type of mathematical methods, which only varies one input variable in each simulation while keeping others fixed in their expected values. Therefore, it is also known as a local sensitivity analysis method. The method is structured on the behaviour of the model for a base-case scenario, which is resulted from the set of input variables in their expected values [Hamby 1994]. The sensitivity coefficient can be computed from first-order partial derivative of the output variable with respect to the input variable in the Taylor series approximation of the model [Saltelli et al. 2008]. In case nonlinearities are neglected the first-order partial derivative can be approximated as the ratio of the corresponding variation in the output to the variation in the input [Hamby 1994].

(2) Statistical methods incorporate the influence of both the range and distribution of the input variables by repeatedly implementing the model. This method evaluates the sensitivity of individual input with varying the other input variables in the same time. Therefore, it considers the interaction effect among the multiple input variables [Frey and Patil 2002]. Statistical method includes regression based sensitivity analysis, partial correlation, Fourier Amplitude Sensitivity Test (FAST), Sobol's method, etc.

In the statistical sensitivity analysis, one of the important steps which should not be ignored is the definition of the distribution of the input variable. The choice of the distribution of input variable determines the uncertainty of the output variable, as well as the relative importance of input variable in the model. The inappropriate distribution of input could lead to large influence on the output variable, even draw the wrong conclusion. The reasonable choice on the range and distribution of input variables might come from the measurement, experienced expert opinion and rational estimation.

(3) Graphical methods detect the relation between input variable and output variable by the graphs and charts. The graphical methods provide a more intuitive way for the analyst to explore the model behaviour. It provides a complement to the mathematical and statistical methods [Frey and Patil 2002]. The common used graphical methods include scatter plots, histograms and cobweb plots.

2.3 Sampling Approaches

Sampling is the process by which values are randomly drawn from input probability distributions. Sampling in a simulation is done, repetitively, with one sample drawn every iteration from each input probability distribution. With enough iterations, the sampled values for a probability distribution become distributed in a manner which approximates the known input probability distribution. The statistics of the sampled distribution (mean, standard deviation and higher moments) approximate the true statistics input for the distribution. Samples can be obtained by various sampling methods, Simple random sampling, stratified sampling, Latin Hypercube sampling and Sobol sequences for example. These methods vary in computational cost and consequently in number of elements required to obtain convergence to the population mean. Generating stochastic samples for building stochastic simulations will always be a compromise between precision of the estimate of the population mean and cost of doing simulation runs. Precision increases as the number of elements in the sample grows, but at the same way, the cost of simulation increases with the sample size (adjusting the sample size will be discussed in 3.3).

In this section the most popular sampling method, Monte Carlo will be represented and its advantages and shortages will be denoted. In addition the variance reduction paradigm, which is applied to reduce the calculation effort during stochastic simulation, will be discussed briefly. The overall aim is to determine, appropriate sample sizes with as few elements as possible and an estimated sample mean with variance of insignificant value as it converges to the actual population mean. One of the improved sampling methods, namely Latin Hypercube Sampling method is also presented in this section. The Latin Hypercube sampling, forces the samples drawn to correspond more closely with the input distribution, and thus converges faster on the true statistics of the input distribution.

Monte Carlo Sampling Method (MCS)

Monte Carlo sampling refers to the traditional technique for using random or pseudo-random numbers to sample from a probability distribution. A wide variety of algorithms are available for generating random samples from different types of probability distributions. Monte Carlo sampling techniques are entirely random — that is, any given sample may fall anywhere within the range of the input distribution. Samples, of course, are more likely to be drawn in areas of the distribution which have higher probabilities of occurrence. In the cumulative distribution shown earlier, each Monte Carlo sample uses a new random number between 0 and 1. With enough iterations, Monte Carlo sampling “recreates” the input distributions through sampling. A problem of clustering, however, arises when a small number of iterations are performed. In the Figure 3 shown here each of the 5 samples drawn falls in the middle of the distribution. The values in the outer ranges of the distribution are not represented in the samples, and thus their impact on your results is not included in the simulation output.

Clustering becomes especially pronounced when a distribution includes low probability outcomes, which could have a major impact on your results. It is important to include the effects of these low probability outcomes. To do this, these outcomes must be sampled, but if their probability is low enough, a small number of Monte Carlo iterations may not sample sufficient quantities of these outcomes to accurately represent their probability [Guide to Using @RISK, 2010].In stochastic building simulation which is an evaluation of a multi-dimensional parameter space, that normally involves a very large number of stochastic variables, can require hundreds of thousands of simulation runs with large sample sizes. For most parts, sample sizes determine the computational cost of the simulation since that number of sample elements equals the required number of simulation runs.

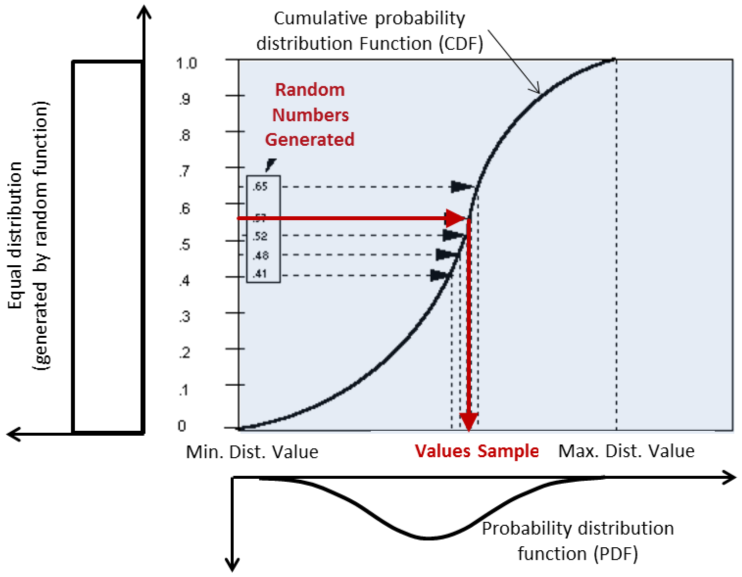


Figure 3: Five iteration of Monte Carlo Sampling[modified from Guide to Using @RISK, 2010]

In mathematics, more specifically in the Monte Carlo methods, variance reduction is a procedure used to increase the precision of the estimators that can be obtained for a given number of simulations. The possibility of variance reduction is what separates Monte Carlo from direct simulation. Simple variance reduction methods often are remarkably effective and easy to implement. It is good to think about them as ways to reduce the burden of the Monte Carlo simulation.

Several sampling techniques, also called variance reduction techniques, have been developed in order to improve the computational efficiency of the method by reducing the statistical error that is inherent in Monte Carlo simulation and keeping the sample size to the minimum possible. Furthermore, advanced solution methods and parallel processing have been recently implemented having a beneficial effect on the efficiency of Monte Carlo simulation [Papadrakakis and Lagaros 2002].Importance sampling (IS) is generally recognized as one of the efficient reduction technique [Schüller 1981, Frangopol 1984, Bucher 1988 and Hurtado and Barbat 1997]. The most popular alternative for Monte Carlo simulation is the stratified sampling techniques such as the Latin Hypercube sampling which is represented in the following section.

Important Sampling Method (IS)

To estimate extreme or rare probabilities, the tails of the distribution are more important than the average values. Rare or extreme events can be associated with dramatic costs, like in finance or because of reasons of safety in environment. Importance sampling (IS) tunes Monte Carlo to the area in parameter space from where the rare events are generated. IS is based on the idea to make the occurrence of rare events more frequent, or in other words, to speed up the simulation. Technically, IS aims to select a probability distribution that minimizes the variance of the IS estimate. A suitable distribution would be one that has higher probabilities in its tails than a Gaussian distribution. Figure 4 represents the general idea of IS. Suppose the target of the inference is the mean of a statistic m(X), i.e. E(m(X)), where m(x) only depends on the sample values x greater than a constant c. Then a proper choice of a new density h(x) of X will make the sample mean of m(X)f(X)/h(X) an unbiased estimator of the target with smaller variance. The role of the new density h(x) is to produce more samples on the area (greater than c) that affects the values of m(x) [Lu et al. 2010].

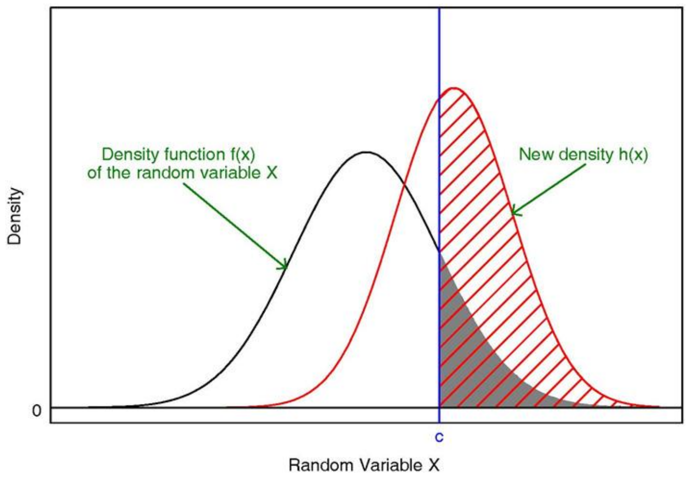


Figure 4: Illustration of Importance Sampling [Lu et al. 2010]

The most difficult aspect to importance sampling (IS) is in choosing a good sampling density, g. In general, one needs to be very careful for it is possible to choose h(x) according to some good heuristic such as the maximum principle, but to then end that h(x) results in a variance increase.

Latin Hypercube Sampling Method (LHS)

Latin Hypercube sampling is one of the most recently developed sampling approach, designed to accurately recreate the input distribution through sampling in fewer iterations when compared with the Monte Carlo method. **Stratification** of the input probability distributions is the key to Latin Hypercube sampling. It is performed by dividing the cumulative curve into equal intervals on the cumulative probability scale. A sample is then randomly taken from each interval or "stratification" of the input distribution. Sampling is forced to represent values in each interval, and thus, is forced to recreate the input probability distribution.

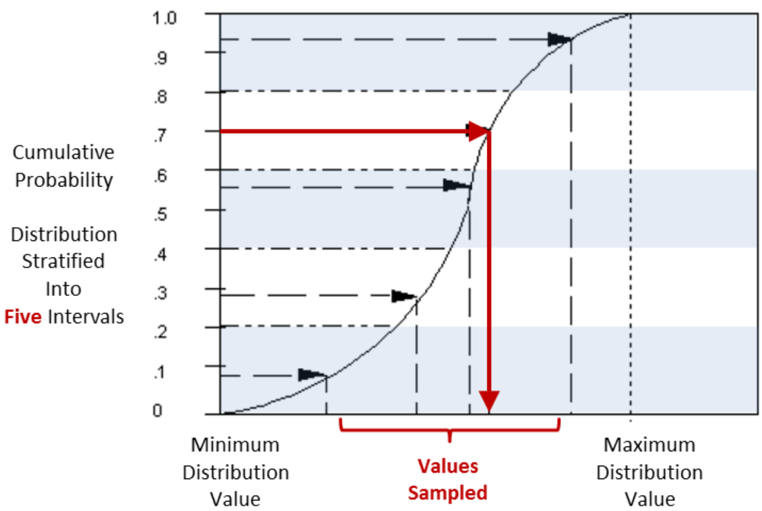


Figure 5: Five iteration of Latin Hypercube Sampling[Guide to Using @RISK, 2010]

In the illustration above (Figure 5), the cumulative curve has been divided into 5intervals. During sampling, a sample is drawn from each interval, whereas comparing to the Monte Carlo method, the 5 samples can be distributed randomly over the whole range and even clustered as shown in Figure 3. With Latin Hypercube, the samples more accurately reflect the distribution of values in the input probability distribution. Already for a few samples, if the sample size is large enough, there is no difference between MCS and LHS. The values sampled, for one variable, need to be independent of those sampled for another. This independence is maintained by randomly selecting the interval to draw a sample from for each variable. In a given iteration, Variable #1 may be sampled from stratification #4, Variable#2 may be sampled from stratification #22, and so on. This preserves randomness and independence, and avoids unwanted correlation between variables.[Guide to Using @RISK, 2010]

As a more efficient sampling method, Latin Hypercube offers great benefits in terms of increased sampling efficiency and faster runtimes (due to fewer iterations). Latin Hypercube also aids the analysis of situations, where low probability outcomes are represented in input probability distributions. By forcing the sampling of the simulation, to include the outlying events, Latin Hypercube sampling assures they are accurately represented in the simulation outputs.

Several researchers have compared the performance of different sampling methods and their efficiency pertaining to building simulation. [Matala 2008] compared simple random sampling and Latin Hypercube sampling in order to finding optimized sample sizes for Monte Carlo simulation for fire problems using complex non-linear models for fire, [MacDonald 2009] examined the performance of several sampling methods; simple random (SRS), stratified (SS) and Latin Hypercube sampling (LHS) in application an integrated natural ventilation problem. It was concluded that for the same number of simulations the LHS method produces a more robust result compared to the stratified method, which in turn produces a more robust result compared to the simple method.

Screening Methods

In such cases in which the number of parameters is very large, applying the screening methods can be very beneficial. Screening methods simplify the models and reduce the number of uncertain input parameters propagating through the model [de Wit, 1997]. Screening methods consider the global sensitivity meaning the input parameters are varied over the whole range of their possible values. A well-established representative is the Morris analysis. For assessing global sensitivity measure, a design composed of individual randomized one factor at a time is built in order to determine, for each factor X_j, the elementary effects $d_j(y)$

$$d_j = \frac{y\left(x_1, \dots, x_{j-1}, x_j + \Delta_j, x_{j+1}, \dots, x_j\right) - y(x)}{\Delta_j}$$

where Δ_j is a value in {1/(p−1),..., 1−1/(p−1)}, with p as the number of levels.

Considering L different trajectories, a statistical analysis of these elementary effects provides the mean $\mu_j(y)$ which assesses the global influence of the factor X_j

$$\mu_j(y) = \frac{1}{L}\sum_{l=1}^{L} d_j^l\,(y)$$

The standard deviation σ_j (y) which indicates the presence of higher order effects and measures the non-linearities or the interactions of the j^{th} factor with others factors is

$$\sigma_j(y) = \sqrt{\frac{1}{L}\sum_{l=1}^{L}(d_j^l(y) - \mu_j(y))^2}$$

In Morris analysis, the uncertainty of the output is characterized by a value called "effect". By varying the input parameter set, the "effect" is calculated several times [Zador et al., 2006]. It allows the selection of important input parameters, by evaluating the model with different inputs. The results of the Morris analysis consist of one graph where the averaging coefficient for each parameter μ_j is compared against the dispersion σ_j from this coefficient per parameter (Figure 6). According to the values of μ_j and σ_j, Morris shows that studied factors can be classed into three groups as follows: factors having (1) negligible effects, (2) linear and additive effects or (3) nonlinear or interaction effects[Santiago et al., 2010].

A drawback of the Morris analysis is that it does not allow uncertainty analysis due to the fact that it does not take the shape of the probability density function of the parameters into account [De Wit et al., 2001].

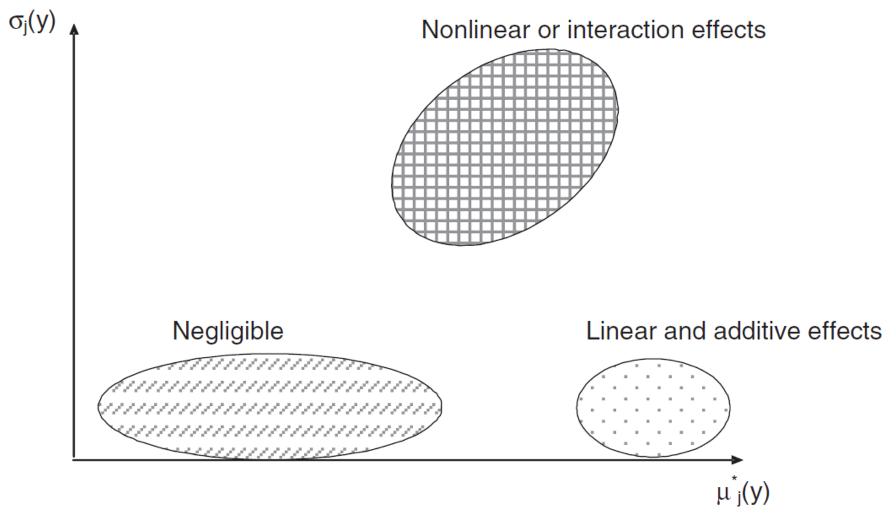


Figure 6: Theoretical disposition of means μj(y) and standard deviations σj(y) of the effects distribution [Santiago et al., 2010]

3. Stochastic Model Types

3.1 Stochastic Model Data Types

The ISES Virtual Energy Laboratory will support product manufacturers, architects, HVAC designers and energy experts in component development (ISES simulation cycle 1, see Deliverable D1.2), new building design (ISES simulation cycle 2a, 2b) and retrofitting/refurbishment of existing buildings (ISES simulation cycle 3a, 3b) by enabling comprehensive simulation of energy efficiency and evaluation of performance and comfort, taking into account probabilistic input values and semi-stochastic computational methods (see table 1). Regarding the simulation intent several data types (domain) have to be applied to the simulation engine. In building component product development and design of facilities, almost all design parameters subject to uncertainty. The stochastic approach addresses all the categories of the parameters space, climate profiles, usage profiles, and energy related building material properties.

Table 1: Data requirements and model data types involving several simulation cycles (from ISES D1.2)

Simulation Cycle	Data Requirements	Data Domain
Component Development Simulation Cycle 1	• Properties of Materials	Stochastic Parameter
	• Climate • Usage Profile (Occupancy)	Stochastic Process
Early Design Simulation Cycle 2a	• Definition of Material	Stochastic Parameter
	• Cubature • Number of Stories • Windows size and orientation…	Deterministic Parameter
	• Solar Gains	Stochastic Process
	• Space Division	Uncertain Parameter (Fuzzy or bandwidth)
Early Design Simulation Cycle 2b	• Definition of Material	Stochastic Parameter
	• Space division • Cubature • number of Stories	Deterministic Parameters
	• Climate • Usage Profile (Occupancy)	Stochastic Process
	• HVAC Type(s)	Uncertain Parameter (Fuzzy or bandwidth)

Table 1 (continued): Data requirements and model data types involving several simulation cycles (from ISES D1.2)

Simulation Cycle	Data Requirements	Data Domain
Retrofitting &Refurbishment Simulation Cycle 3a	• Current HVAC Type current Façade	*Deterministic Parameters*
	• Climate • Solar Gains	*Stochastic Process*
	• Usage Profile (Occupancy)Windows size & Orientation	*Uncertain Parameter (Fuzzy or bandwidth)*
	• Floor Material	*Stochastic Parameter*
Retrofitting & Refurbishment Simulation Cycle 3b	• Floor Material • Windows size & Orientation	*Deterministic Parameters*
	• Climate • Usage Profile (Occupancy) • Solar Gains	*Stochastic Process*
	• Current HVAC Type • Current Façade	*Uncertain Parameter (Fuzzy or bandwidth)*

Regarding the Simulation Cycles in ISES the stochastic model data can be categorized into the six parameter types; (1) deterministic parameter, (2) stochastic parameter, (3) time-invariant stochastic processes, (4) time-variant stochastic processes, (5) time-invariant stochastic field, (6) time-variant stochastic field.

(1) Deterministic parameter

The deterministic parameters are the physical parameters which are not allowed to variate during simulation. One characteristic of the simulation cycles are their different the deterministic parameters. For instance, during the simulation cycle 2a the parameters cubature, number of stories and the windows size and orientation are considered as deterministic parameters.

(2) Stochastic parameter

In order to consider the uncertainty in input parameter characteristics, the stochastic parameters are assigned. The stochastic parameter is defined via a probability distribution function (PDF), which shows the relative likelihood for this random variable to take a given value. The thermal properties of material, for examples, are candidates of stochastic parameters.

(3) Time-invariant and (4) Time-variant stochastic processes

The stochastic process is characterized as a family of random variables which can be discrete or continuous in time. The time-invariant or stationary stochastic processes exhibit statistical properties concerning the invariant time. Thus, for example, second-order stationarity implies that the statistical properties of the pairs $\{X(t_1), X(t_2)\}$ and $\{X(t_1+c), X(t_2+c)\}$ are the same for any c. The solar radiation and weather temperature are the data which are modeled by using the stochastic processes.

(5) Time-invariant and (6) Time-variant stochastic field

A random field is a generalization of a stochastic process such that the underlying parameter need no longer be a simple real or integer valued "time", but can instead take values that are multidimensional vectors, or points on a topological space. At its most basic, discrete case, a random field is a list of random variables whose indices are mapped onto a space (of n dimensions). Values in a random field are usually spatially correlated in one way or another. For instance, the occupancy should be modeled as a stochastic filed.

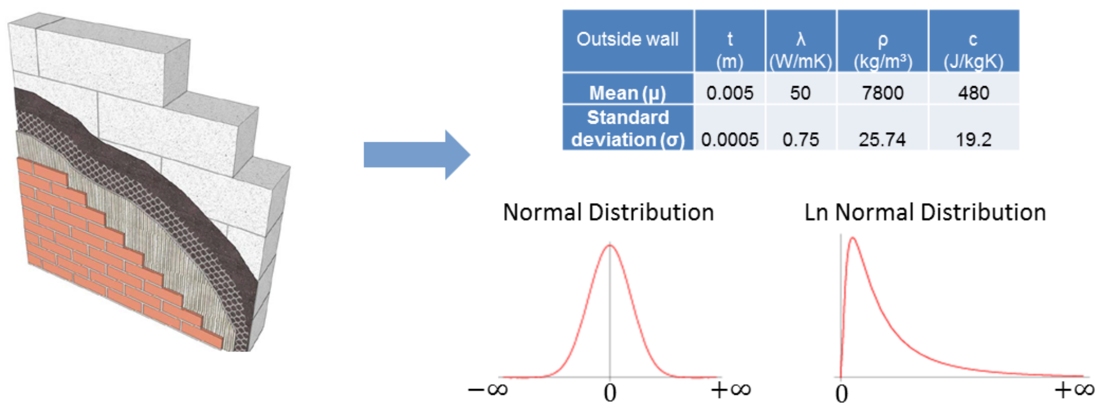
3.2 Stochastic Variables

Material Properties

The uncertainty in material properties is caused by (1) material inherent inhomogeneity and manipulation during the production: the natural variability of the physical properties for a specific material, being an intrinsic property of natural materials(2) measurement: errors caused by experimental setup, evaluation and interpretation of experiments[1]and (3) modeling methodology: errors due to the fact that the material functions (e.g., liquid water conductivity) are generated by using simplified models which could not represent the real properties perfectly.

During the stochastic simulation the uncertainty in material properties is considered by modeling the thermal conduction (λ), density (ρ), specific heat capacity (c) and thickness (t). The material modeling requires a series of experiments to acquire either the basic parameters or the data for further functionalization. Some of the experiments have a long test period, which may extend the analysis process or affect the precision of the results, i.e. when taking the stochastic values from similar material. Some statistical methodologies, which can group the similar parameters (cluster analysis) and reveal the relations between in-group parameters, further predict unmeasured variables (regression analysis), can be used to simplify the test procedure. [Zhao, 2012]

Finally, the uncertainty in material properties may be represented in random distribution form. Applying the Normal Distribution or a ln-Normal Distribution the mean (μ) and standard deviation (μ) are sufficient to describe the uncertainty relies in material properties. (Figure 7)



Outside wall	t (m)	λ (W/mK)	ρ (kg/m³)	c (J/kgK)
Mean (μ)	0.005	50	7800	480
Standard deviation (σ)	0.0005	0.75	25.74	19.2

Figure 7: Representing the uncertainty relies in material properties by applying random distribution

[1] Measurement uncertainty is to be modeled with Fuzzy methods. However, for simplification in our concern it is as well modeled as stochastic uncertainty.

Climate/Weather

The stochastic simulation of building energy demand requires climate/weather data to calculate energy balance of buildings and HVAC systems. Future weather prediction is critical for the success in predicting energy requirements for building heating and cooling. Climatologists and building scientists have been working together actively to improve methodologies for future weather data preparation [Crawley 2007 and Guan 2009]. The values of climate elements are based on time-series of measurements/predictions which are modelled by applying time-variant (non-stationary) stochastic process. Development of **stochastic weather data** provides an opportunity to produce synthetic data representative of future climatic conditions that may influence more comprehensive lifecycle energy consumption analyses [de Wilde & Coley 2012]. [Aguiar et al. 2002] adopted a stochastic method which constructs future meteorological test reference years by matching historical records of the same location with predicted mean monthly air temperature.

Predictions published by the Intergovernmental Panel on Climate Change (IPCC) indicate an increase in global average surface temperature in different scenario ranges of 1.1–2.9 °C to 2.4–6.4 °C from a 1990s baseline towards the end of the 21st century [IPCC, Climate Change 2007]. The main cause of the climate change trend is the emissions from buildings, business, agriculture and transport [Pachauri 2005]. There are several methods to construct climate change weather data for building simulation from results of global or regional circulation models. One of them is adjustment of present-day weather data with regional climate change model predictions, generally termed "morphing" [Belcher et al. 2005]. Based on a probabilistic modelling [Schölzel and Hense 2011] have developed a method to assess the regional climate change in Germany and performed the estimation of its uncertainty during the first 3–4 decades of the twenty-first century. Their statistical model extracts information from an ensemble of regional climate simulations to estimate probability distributions of future temperature change in Southwest Germany (figure 8). The method was used is related to kernel dressing which has been extended to a multivariate approach in order to estimate the temporal auto covariance in the ensemble system.

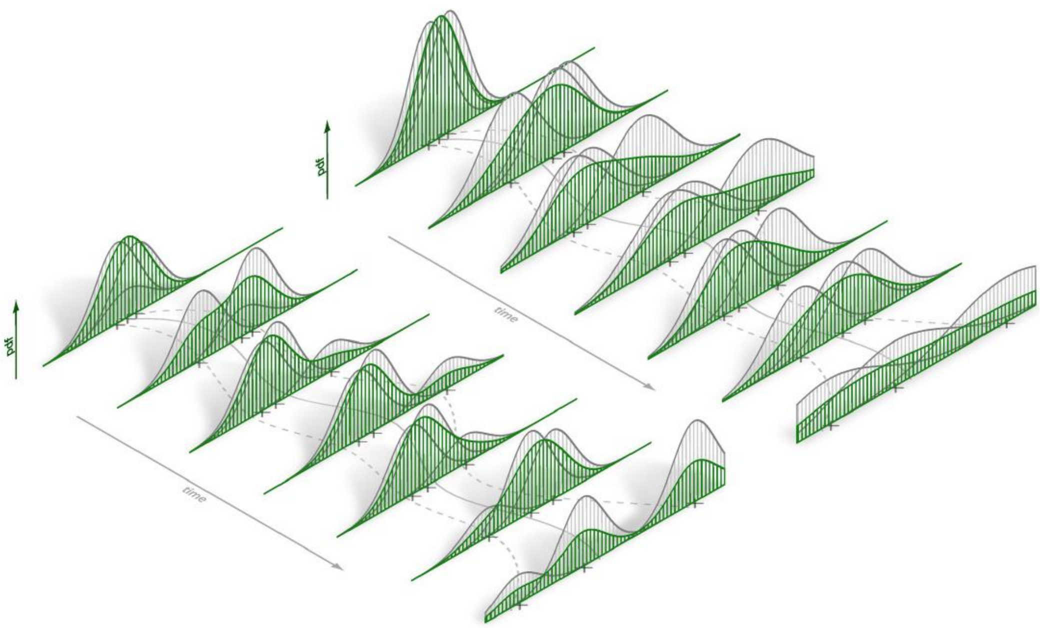


Figure 8: Illustration of different kernel dressing methods: Bayesian model averaging (left) and affine kernel dressing (right) based on an example of three ensemble members (plus symbol). The lines indicate ensemble mean and spread, the gray hatching the (weighted) kernel functions, and the green hatching the resulting probability density [Schölzel and Hense 2011]

So far, there is still no standard methodology available. As it was briefly reviewed several approaches have been used by various research groups. The required climate elements can be characterized as the outdoor air temperature, the relative outdoor air humidity, the overall solar radiation on a horizontal plane, the direct/diffuse solar radiation on a horizontal plane, the wind direction and wind velocity, the precipitation and cloudiness etc.

Occupancy

Various studies have observed the impact of consumer behavior on energy demand [e.g. Santin 2011, Sardianou 2008]. These studies highlight various factors related to occupants behaviors that influence building energy use but they present some difficulties of assessing the specific contribution of each factor. Because of the direct influence of occupancy (or activity) patterns on the energy consumption profile, it is preferred to develop the occupancy profile which is used to model the usage profile. In the domestic sector, it depends not only on the number of people who live at a property but also on whether they are at home and active (i.e. not asleep). From Occupancy the heating schedule, ventilation rate, shading, cooling schedule, equipment scheduler, etc. Are dependent (figure 9). Among several occupancy modeling studies, applying the Markov chain to model the occupant presence was used most frequently. [Page et al. 2008] have modeled the occupancy by considering occupant presence as an inhomogeneous Markov chain interrupted by occasional periods of long absence. The model generates a time series of the state of presence (absent or present) of each occupant of a zone, for each zone of any number of buildings. In the following, the occupancy pattern modeling in residential building will be illustrated in detail. Since the occupancy modeling of residential building is more complicated than the office or commercial buildings, the following approach can be used in a simplified manner for the office or commercial buildings.

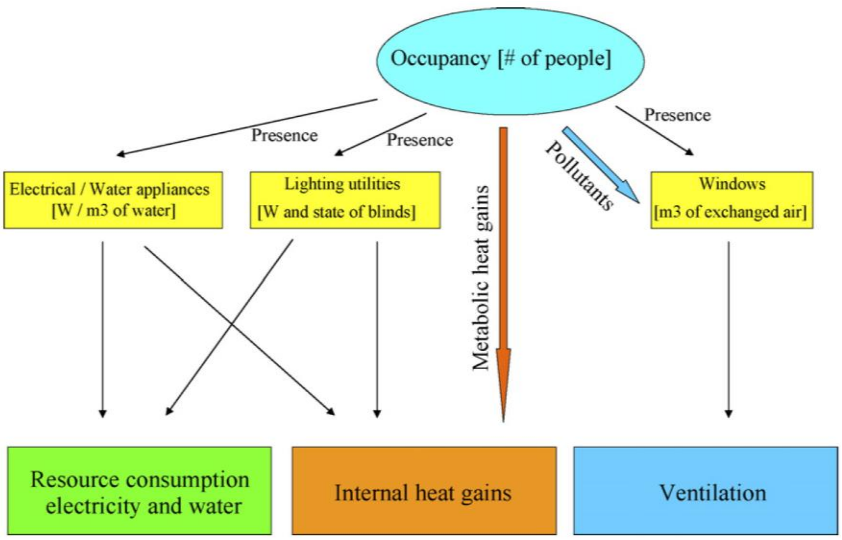
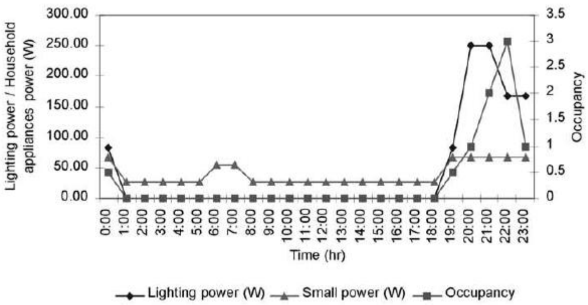


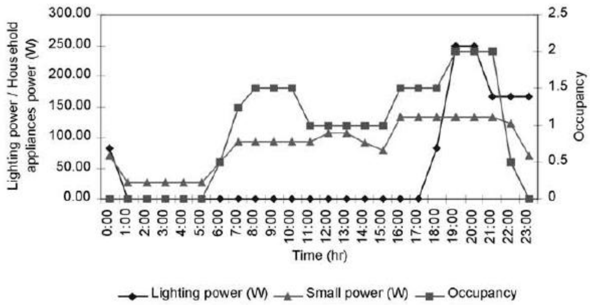
Figure 9: Outputs of the occupancy model and their later use by stochastic models of occupants behaviour [Page et. Al. 2008]

The occupancy pattern in residential building itself depends on several social and demographic parameters. An extensive study performed in Germany [Daily Life in Germany, analysis of the time use, 2004] show that the families with and without poverty risk follows different time consumption patterns. The sexuality and the age of children can also affect the occupancy pattern. For instance, the 14-18 years old girls spend 67% more their time in household works comparing the boys in the same age, while the ratio for the 10-14 years old children limited to 33%.The investigations of [Wan and Yik 2004] shows that the correlation between the occupancy profile and energy consumption might be extremely different for several family types and several separated zones in a building unit. In figure 10 the power energy consumption versus occupancy is illustrated for two separated zones for two family types; namely a family in which all four members are working full time and a family in which only two persons have a fulltime activity and two are unemployed. As it is expected during the evening the lighting power consumption in living and dining room follow the occupancy pattern while during the daylight the small power consumption shows the better correlation with the occupancy profile.



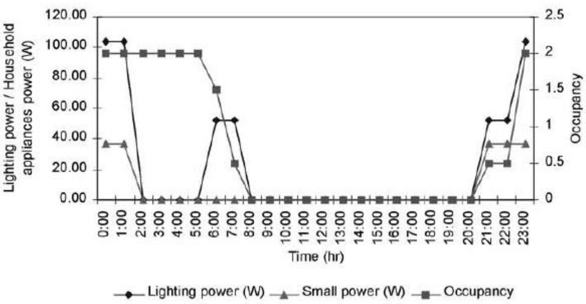
living and dining room, family type F1

(All 4 family members working fulltime)



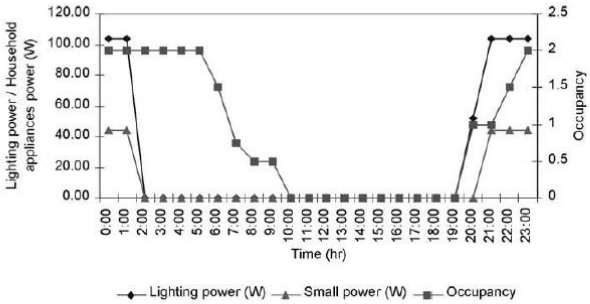
living and dining room, family type F2

(2 family members working fulltime and 2 are unemployed)



Bedroom, family type F1

(All 4 family members working fulltime)



Bedroom, family type F2

(2 family members working fulltime and 2 are unemployed)

Figure 10: Daily occupation-patterns and lighting and appliances power-profiles for a living and bedroom for two family types[from Wan and Yik 2004]

Another study from [Meesteret al. 2012] shows the following picture. Considering an active couple works outside the house during the day while their three children go to school. The family lives in a two stories house with 182 m^2 living area (Figure 11). The building is divided to six separated zones; namely living, kitchen, office, couples' room, 3 children's' room and bathroom.

Figure 11: Plans of the ground floor and the first floor of the studied house [from Meester et al. 2012]

The occupancy profiles and their schedules based on surveys realized in Belgium about Dwellings, Households and families, and education and employment [Meesteret al. 2012] are illustrated in figure 12.To normalize occupancy with the surface area, the occupancy density which is defined as the number of occupants in area unit (e.g. cubic meter) can be used (in this case, the occupancy density for every resident is 1/182=0.0055). The total occupancy density profile for the building in a working day is got by summing up the occupancy density profiles for several buildings zones.

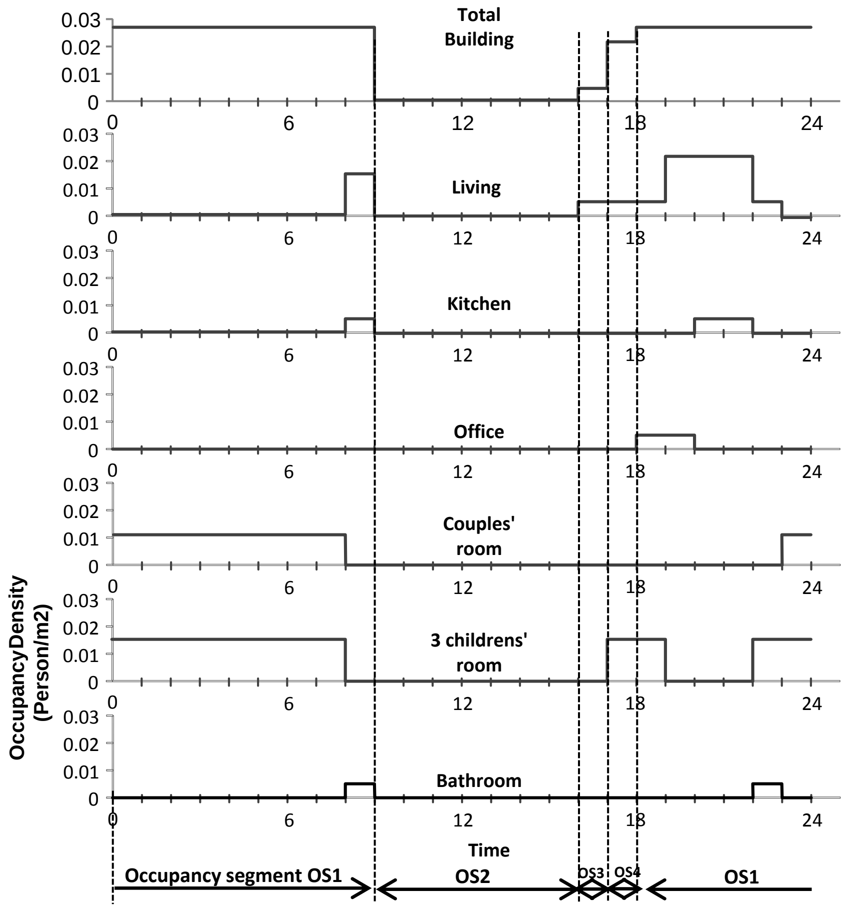


Figure 12: Occupancy profiles per building zone

Modeling the stochastic occupancy profile is performed by dividing the occupancyprofile in its dominant segments. According to the clustered statistical data regarding several building and family types, the probability distribution for every occupancy profile segment is driven and the probability distribution relevant parameters are extracted (in the case normal and ln-normal distribution, mean and standard deviation). Mean values dedicatedtheboundary of occupancy segmentsare represented on figure 13 as t_1, t_2, t_3 and t_4. Correspondingto the occupancy profile segments, the value of occupancy density is represented inprobability form as well, where the O_1, O_2, O_3 and O_4 are the mean values regarding the occupancy density of every occupancy segment (figure 12).

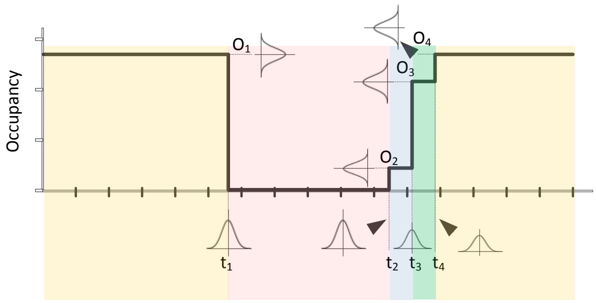


Figure 13: Propability distribution of the stochastic occupancy profile segmentats

In fact, the occupancy profiles of separated building zones are not independent from each other. Presence of a specified resident(s) in a specified zone during the specified time instant,means the absence of the specified resident(s) in the rest zones of the building. Spatial Correlation between the different zones in a building unit can be modeled by applying the random field paradigm. The random field is constrained by the sum of the occupancies (Figure 14). A random field is also called a random (or stochastic) process, although the term "field" indicates that the parameter space is multidimensional. A random field X(t) is a collection of random variables at point with coordinates t=(t_1, …, t_n) in an n-dimensional "parameter space". Second-order information about point-to-point variation is contained in the covariance function B(t,t'), the covariance between values of the random filed at two location t and t'[Vanmarcke 1983]. The variation of n-dimensional random field x(t) at two location t and t' is characterized by the covariance function

$$B(t,\acute{t}) \equiv B_x(t,\acute{t}) = cov[X(t),V(\acute{t})] = E[X(t)X(\acute{t})] - m(t)m(\acute{t})$$

Or by the correlation function

$$\rho(t,\acute{t}) \equiv \rho_x(t,\acute{t}) = \frac{B(t,\acute{t})}{\sigma(t)\sigma(\acute{t})}$$

Where m(t) and m(t') are the means and σ^2(t) and σ^2(t') the variances of X(t) and X(t'), respectively.

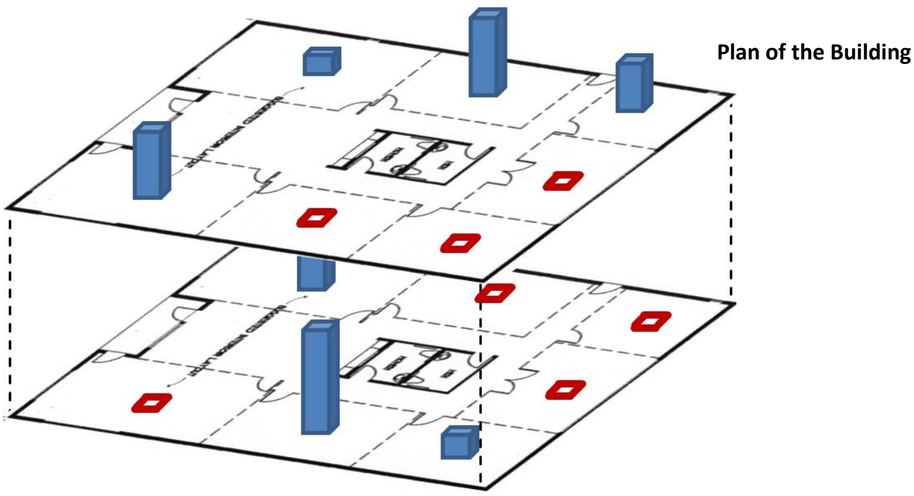


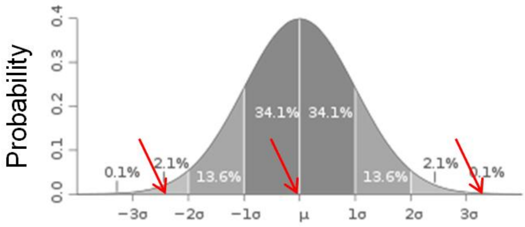
Figure 14: Modeling the spatial correlation between several zones in a buildingunit.The empty zones are represented by the red squers while the occupated zones are marked by the blue cuboid

3.3 Sample Size

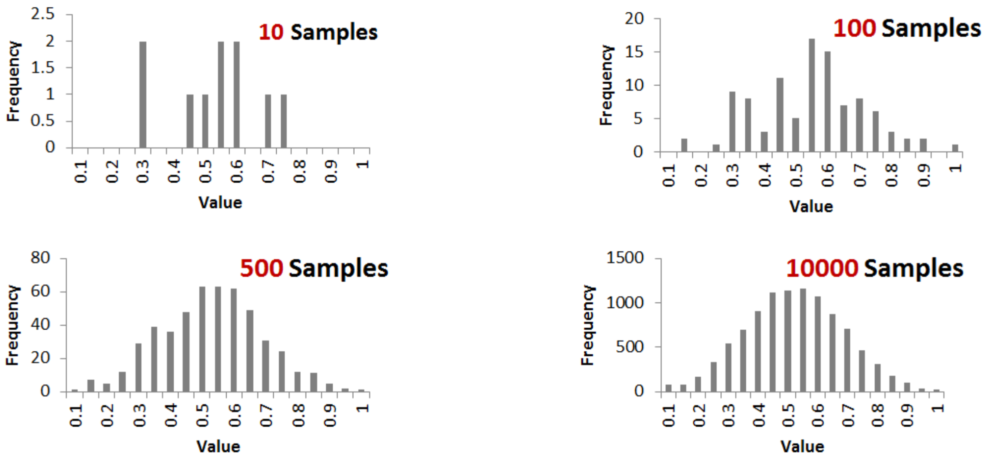
A common question asked during Monte Carlo sampling (or the other sampling approaches like LHS) is, "What sample size is enough for the stochastic simulation?" The sample size required for a particular analysisdepends on various factors such as type of model, the random number generator used, type of distributions, and the output probabilistic measure and cannotbe universally defined. The sample size is also strongly depended on the precision which is expected [Schüller, 1981]. The required number of sample (N) can be calculated by

$$N = \frac{1}{\varepsilon^2}\frac{\sigma^2}{\mu^2}$$

Where σ is the standard deviation, μ is the mean and relative error is represented by ε.The general tendency is to reducethe samples as much as possible without realizingthe effect on decisions. For example, the mean of theoutput requires a number of samples that is an order ofmagnitude less compared to the variance. Therefore, itis desirable to use a sampling technique that can predictthe output probabilistic measure accurately withthe minimum number of samples. Over the years, several rules of thumb have been proposed such as 5-10 observations perparameter, no less than 100, and so on. At least one thing is clear, the larger the sampling size is, the better coverage of the designed distribution of the input can be obtained. However, due to the computational cost, the numbers of the sampling should be controlled in a reasonable limit. In reality there is norule of thumb that applies to all situations. The sample size needed for a study dependson many factors including the size of the model, distribution of the variables, amount ofmissing data, reliability of the variables, and strength of the relationships among thevariables[Muthén and Muthén 2002].



(a)



(b)

Figure 15: (a) Small and large fractiles in probability distribution, (b) The effect of sample size in regenerating a random distribution (mean=0.5, standard deviation=0.1667)

Considering the normal distribution in figure 15 (a).A numerical descriptor for the span of frequency distributions is the fractile. A p-fractile is defined as the x-value of the distribution which includes p*N observations, with 0<p<1 and N being the number of observations. An example may clarify this: the 2.1 % fractile of the distribution shown below is almost -2.4σ, as it includes 2.1 % of all observations (starting from the left). Accordingly, the median (μ) is the 50% fractile and the large fractile 99.9 % is 3.4σ.

It means to reach the occurrence probability of 99.99 %, 10000 samples are required (Figure 15, b), that means at least one sample lies above the 99.99 % fractile. [Lomas and Eppel 1992] applied Monte Carlo analysis on building thermal modeling filed and found after 60-80 simulations the accuracy of confidence interval on standard deviation of outputs marginally increase regardless of the number of input parameters. Nevertheless, having a relatively small sample size (N<100 independent of the number of variables) it is possible to determine the means and standard deviations of the output parameters with 10% accuracy [Roulet 1999].

3.4 Stochastic Model Approach

The stochastic simulation in ISES will be done for the whole life-cycle of buildings and facilities. The simulation task may require hundreds of individual simulation runs, performed in parallel, which should be managed and configured applying highly automated approach, with only general control interaction by the user.The virtual energy lab is structured havingtwo feedback cycles, one for handling the stochastic life-cycle nature of climatic, usage conditions, material properties and one for the optimization of the design. ISES covers a wide spectrum comprising the selection the most energy efficient product, designing and retrofitting scenarios. Since the stochastic variables involved in the simulation approach vary for every ISES use case scenario (simulation cycle 1-3), the preliminary principle stochastic model approach for every use case scenario will be developed separately in this section. Afterward, a method to generate the random sample dataset regarding stochastic model data types is presented and obtaining the energy demand for the building life-cycle will be characterized.

Simulation cycle 1; Component Development

The efficient design of energy related building components and their optimal control in operation need several feedback cycles to reach an optimal design and operation control process. As it was represented in table 1 the major stochastic variables involving in component development are the climatic and usage profiles. While the climate strongly depends on the location, the user requirements depends on the end use of a building which is only partially influenced by the climate. Regarding the important role of façadeelements in energy efficiency and energy saving design, the main focus of this simulation cycle is dedicated to designing the optimal façade element for each location and end user type.

The model approach for the simulation cycle 1 is represented in figure 16. The stochastic simulation is feed with the stochastic values generate by applying several stochastic parameter and process models types can be given in Figure 16. The obtained probabilistic results of uncertainty analysis will be used to estimate the energy demand in the building life-cycle. The feedback of the sensitivity analysis may be used to estimate the role of every engaging stochastic parameter and every stochastic process in the stochastic modelling.

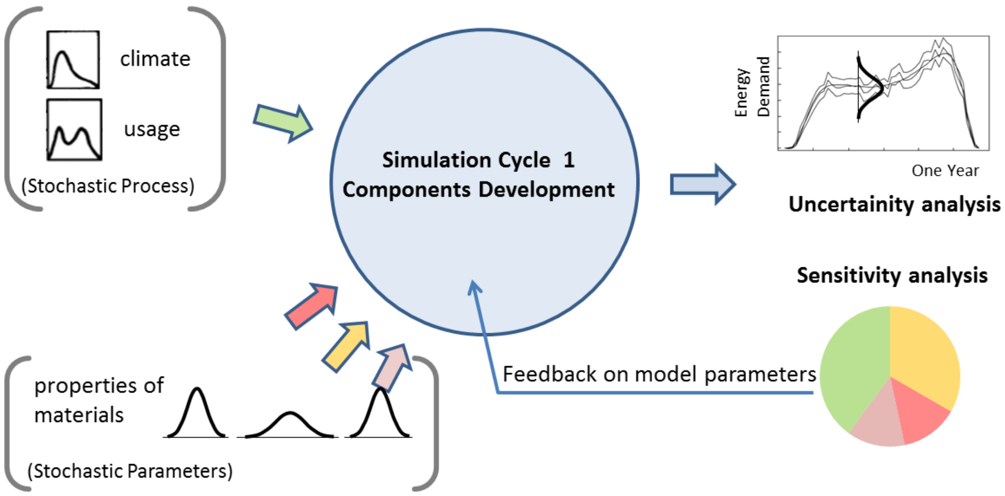


Figure 16: Illustration of the stochastic simulation approach regarding the simulation cycle 1; components development

Simulation cycle 2; Early Design

The main and principle behaviour and layout of the building will be determined during early design. In the early phase the architectural alternatives are designed concentrating on the spatial layouts and functions with attention paid to both energy efficiency and life cycle costs. In this phase the results from the life-cycle analysis and the information from different spatial layout alternatives the actual energy consumption is calculated to determine the operational costs. During ISES the simulation cycle 2 is divided into two stages (detailed information can be found in ISES D1.2).

The stochastic simulation regarding Cycle 2b is shown in figure 17.During this stage the early design of building energy with already defined heating, ventilation and air-conditioning (HVAC) types will be performed. The required input parameters/variables are stochastic process, deterministic parameters, stochastic (and uncertain) parameters. The obtained uncertainty analysis results will be represented in probabilistic manner and applied into the decision making process.

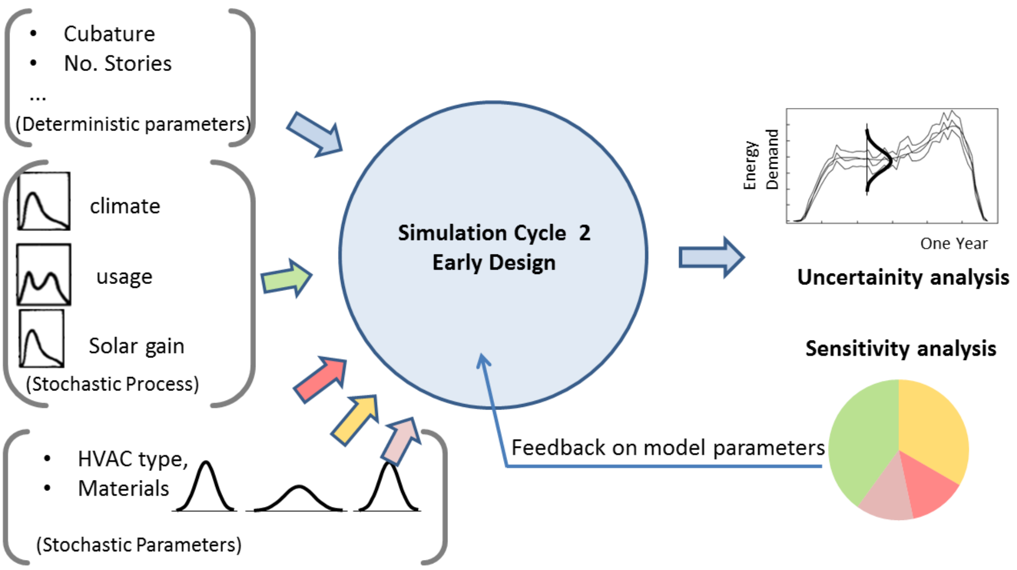


Figure 17: Illustration of the stochastic simulation approach regarding the simulation cycle 2b; early design

Simulation cycle 3; Refurbishment/Retrofitting

If building components or technical installations of existing buildings do not longer meet sufficiently current technical, economical, ecological or regulatory requirements, the client or facility manager can choose between different alternatives of retrofitting or refurbishment measures on the basis of the results from energy simulations. Two probabilistic simulations cycles can be applied within the developed retrofitting/refurbishment process; namely 3a and 3b (see ISES D1.2 for more information).

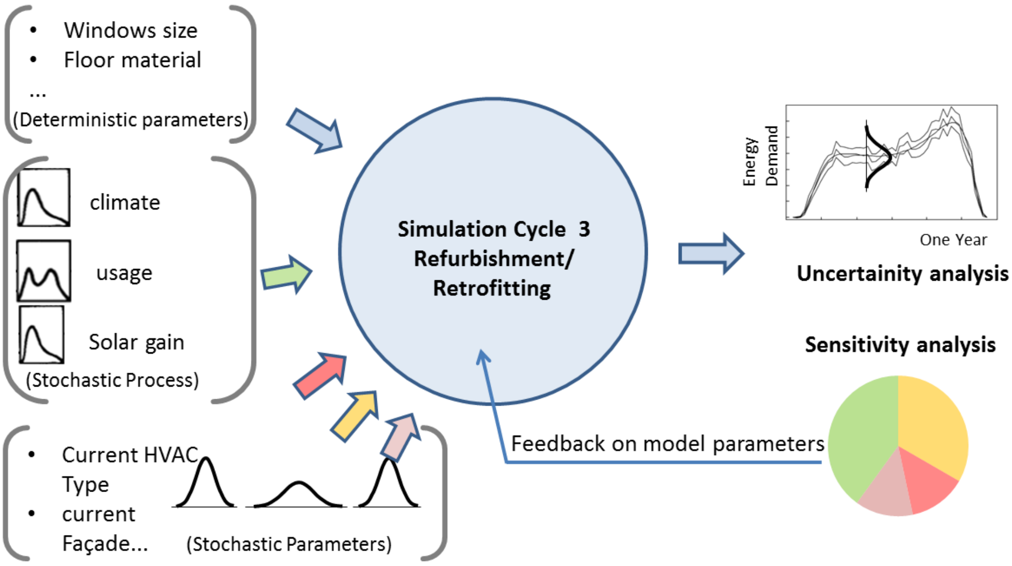
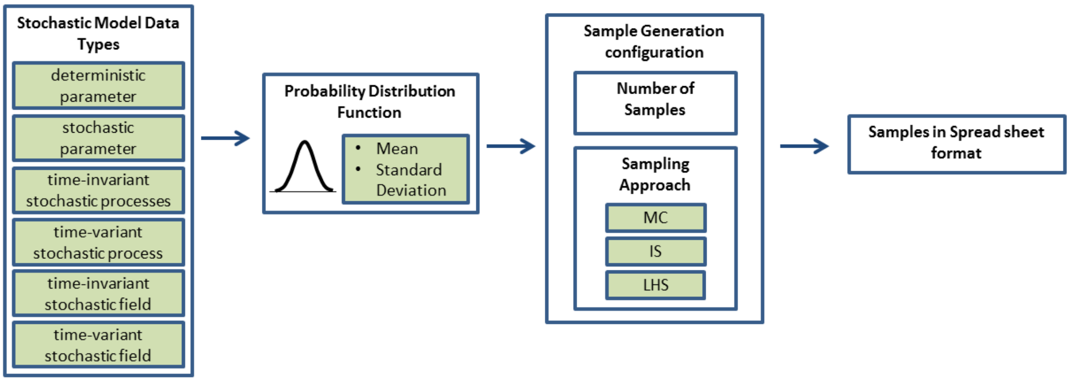


Figure 18: Illustration of the stochastic simulation approach regarding the simulation cycle 3b; Refurbishment/Retrofitting

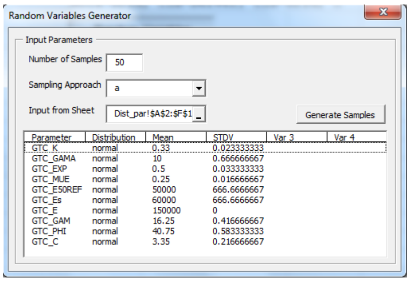
During Simulation Cycle 3b the impact of different alternative building elements on the overall building energy performance will be studied (see Figure 18). The required input parameters/variables are again categorized under deterministic parameters, stochastic processes and stochastic (and uncertain) parameters. The obtained results of the uncertainty analysis will be represented in probabilistic manner and applied into the cost estimation regarding several demand scenarios.

Generation of Sample Datasets

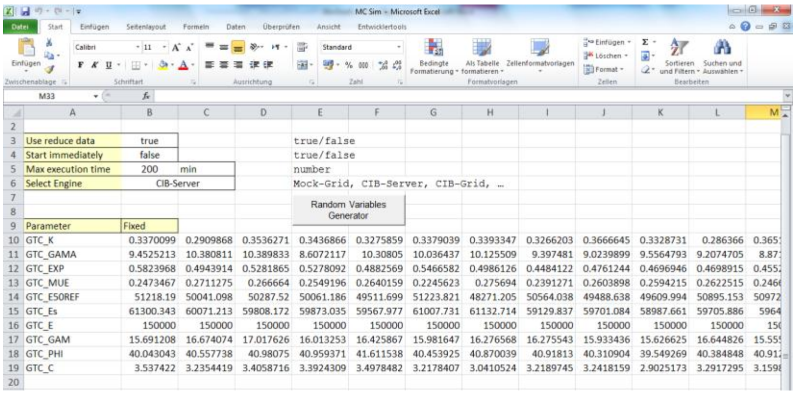
Regarding the stochastic model data types, which will be applied in stochastic simulation, samples dataset can be generated. The process of generating sample datasets for the data type stochastic parameters is illustrated in figure 19 a. In this case the parameter of a normal probability distribution function (PDF), mean and standard deviation have to be given. Applying the PDF parameters and selecting the sampling approach as well as number of samples, the sample datasets can be generated in a sample file, i.e in a Excel spreadsheets in a specified format directed applying Excel built-in random generator engines and Excel-VBA (Figure 19 c), in order to be readable by the Grid-based simulation engine.



a) Generating random sample dataset procedure



b) Configuration of the sample datasets regarding several stochastic parameters



c) Generated sample datasets in excel spread sheet

Figure 19: Generating sample datasets

The highly non-stationary stochastic process concerning usage and climate models have to be approximated e.g. evolutionary stochastic process or by a stochastic process of partly stationary processes. We will use the approximation of the partially stationary processes to model the non-stationary stochastic process as a macro discrete stochastic process where the discrete events are the micro stationary stochastic process modules. A further approximation will be introduced as described below, because the stochastic nature in the modules is of minor importance and the design objective is either the average values of the life-cycle or the extreme values or the extreme change values over a short time (e.g. hours).

Energy Demand Estimation

The stochastic approach of ISES is focused on lifecycle energy performance. Envisaged energy, CO2 emissions and costs are the parameters which will be derived from the stochastic simulations.These design objectives are set in order to obtain

(1) Overall mean energy demand and CO2 emissions

(2) The worst-cases of changes in heating and cooling to design the energy providing systems (99% fractile of energy demand)

Extreme values and extreme change values are the result of the respective worst-case sequences of the patterns. This means the task of non-stationary stochastic processes can be reduced and approximated by many discrete deterministic sequence processes of the sequence of deterministic characteristic patterns.Once the stochastic simulations were performed, to examine the obtained results and derive from that the resulting stochastic distribution outputs presentation, the stochastic process analysing models are deployed.

In order to estimate the building energy demand over the building life-cycle (e.g. 30 years), the annual energy demand is integrated over the time. Since the observations *X(u)* are made continuously, the purely random process becomes an „ideal white noise " and integral is known as the Wiener Process :

$$I(t) = \int_0^t X(u)\,du$$

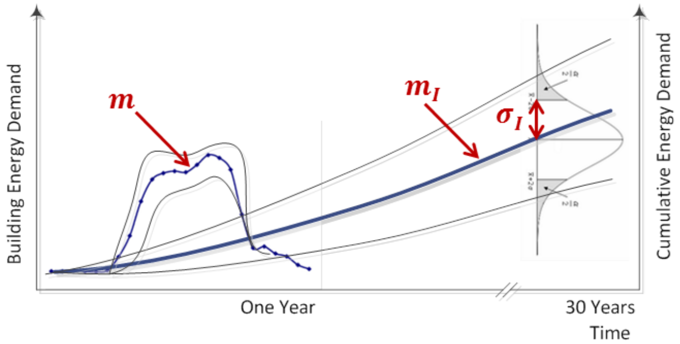


Figure 20: Mean (m) and standard deviation (σ) of the energy demand

If *m* is the mean of *X(u)*, the mean of the Wiener process *I(t)* is

$$m_I = mt$$

Its variance is also proportional to *t*

$$\sigma_I^2 = wt$$

where, *w* measures the intensity of white-noise ($w = \sigma^2 \Delta t$). Therefore, the mean of the integral of the stochastic energy demand over the time is calculated by multiplying the mean of the stochastic process to the time span. The standard deviation of the integral is calculated by rooting the multiplication of the intensity of white noise (*w*) to the time span (Figure 20).

Several studies show growing in energy consumption demand over the time. During the only last fifty years, the demand was approximately doubled [Eddy and Marton, 2011] and it seems the trend will continue in the future as well (Figure 21).The gradually increasement should also be considered during calculation of the building life-cycle energy demand. The trend pattern can be applied as a linear function over the time.It should be remembered that the slope of the trend is strongly depends on the region, for which the study is performed.

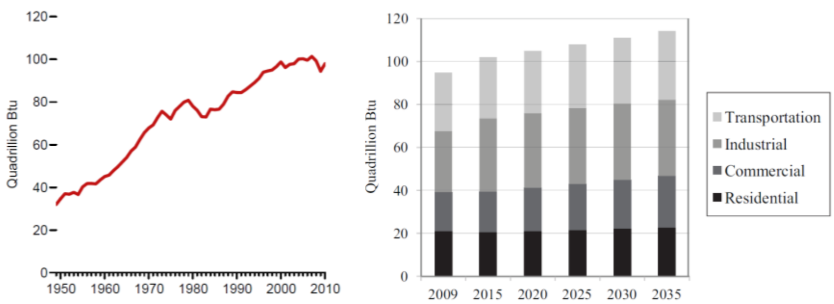


Figure 21: Energy demand by end use sector, 1950-2009 (left) and 2009-2035 (right)[from Eddy and Marton, 2011]

Presentation of the Simulation Results

One of the most challenging aspects in stochastic simulation of energy demand is the presentation of the simulation results in a transparent and meaningful way. Visualization of the building energy data allows for the comparison of the measured data with simple rule models and therefore the finding of dominant trends as well as detection of possible faults.Numerous graphical presentations have been presented in the literature. However, for designers to be able to make informed decisions based on multi-variant simulation results and to evaluate different design alternatives results need to be presented in terms of clear performance metrics which have meaning for the designers involved in the design process and not just the simulation expert. The overall idea is to use scatter, carpet and box plots for the analysis of the data as they provide clear characteristic patterns. When unusual behavior is identified, time series plots are then used to explore in more detail single data streams for a given period of time.

4. Conclusions

In this deliverable, the three phases of the stochastic simulation for ISES are described, namely the pre-processing, simulation and post-processing phases. Identification of the stochastic variables, their ranges and scope and identifying the appropriate sampling methods are conducted. Once the model evaluations have been performed, during the post-simulation step all results from the multiple simulations are collected and uncertainty and sensitivity analysis are performed. In order to minimize the number of stochastic variables, before the uncertainty analysis the sensitivity analysis is carried out. Since the simulation itself is an expensive computing process, finding a proper sampling approach (reducing the variance) and the optimal sample numbers are of enormous importance for the stochastic simulation.

During this deliverable a special attention was put on defining the stochastic variables: (1) Material properties, (2) Climate/Weather and (3) Energy consumption pattern in form of occupancy. A model based on statistical analysis was suggested to establish the stochastic occupancy density profile. In this regard, the stochastic occupancy profile is formed through several separated profile segments which are characterized by probability density functions. In fact, the occupancy profiles of separated building zones are not independent from each other. We have suggested applying the random field paradigm to model the Spatial Correlation between different zones in a building unit.

The stochastic simulation in ISES will be done for the whole life-cycle of buildings and facilities. Since the stochastic variables involved in the simulation approach vary for every ISES use-case scenario (simulation cycle 1, 2 and 3), the preliminary principle stochastic model approach for every use case scenario was developed separately. Furthermore, a method to generate random sample dataset regarding stochastic model data types is presented in this deliverable. Using Excel-VBA and Excel built-in functions, generation of random values for stochastic parameters on Excel Spreadsheets was illustrated.

Envisaged energy, CO_2 emissions and costs are the parameters which will be derived from the stochastic simulations. Overall mean energy demand and the worst-cases demand (99% fractile of energy demand) are described as the indicators which will represent the energy demand for the building life-cycle in ISES. Along these lines, the Deliverable 2.1 achieves three goals:

(1) The vital features of the stochastic simulation have been described concisely

(2) The stochastic model data types and the stochastic variable types have been characterized

(3) The stochastic model approaches for several simulation cycles in building life-cycle are developed.

References

Aguiar, R., Oliveira, M. and Goncalves, H. (2002): Climate change impacts on the thermal performance of Portuguese buildings, Results of the SIAM study, *Building Services Engineering Research and Technology* 23 (4) , 223–231.

Belcher, S.E. , Hacker, J.N. and Powell, D.S. (2005): Constructing design weather data for future climates, *Building Services Engineering Research and Technology* 26 (1) , 49–61.

Bucher, C.G. (1988): Adaptive sampling – an iterative fast Monte Carlo procedure, *Structural Safety* 5 , 119–126.

Crawley, D.B. (2007): Creating weather files for climate change and urbanization impacts analysis, *Proc.: Building Simulation 2007*, 1075–1082.

Daily Life in Germany (2004): Analysis of the time use (Alltag in Deutschland: Analysen zur Zeitverwendung), *Beiträge zur Ergebniskonferenz der Zeitbudgeterhebung 2001/02 am 16./17. , Wiesbaden, Statistisches Bundesamt*, Band 43.

De Wilde, P., Coley, D. (2012): The implications of a changing climate for buildings, *Building and Environment* 55, 1–7.

De Wit, M. S. (1997): Identification of the important parameters in thermal buildingsimulation models, *Journal of Statistical Computation and Simulation* 57, no.1: 305-320.

De Wit, M.S. (2001): *Uncertainty in Predictions of Thermal Comfort in Buildings*, Phd Theses, TU Delft.

Ekström, P.-A. (2005): *Eikos A Simulation Toolbox for Sensitivity Analyses*,-UPTEC IT05015, degree project, Uppsala Universitet.

Eddy, W.F. and Marton, K. (2011): *Effective Tracking of Building Energy Use: Improving the Commercial Buildings and Residential Energy Consumption Surveys*, National Research Council, NAS Press, ISBN: 0309254019.

Eisenhower, B., O'Neill, Z., Fobonorov, V. and Mezic, I. (2011): Uncertainty and Sensitivity Decomposition of Building Energy Models, *Journal of Building Performance Simulation*, 11(200):1-18.

Fabi, V., Andersen, R.V., Corgnati, S.P., Olesen, B.W. and Filippi, M. (2011): Description of occupant behavior in Building energy simulation: state-of-art and concept for improvements , *Proc. Building Simulation 2011*, Sydney.

Frey, H. C., and Patil, S. R. (2002): Identification and review of sensitivity analysis methods, *Risk Analysis*, 22(3), 553–578.

Frangopol, D. (1984): Interactive reliability based structural optimization. *Comp. Struct* 19 (4) , 559–563.

Gokhale, Swapna S. (2009): Model-based performance analysis using block coverage measurements, *Journal of Systems and Software* 82, no. 1 (January): 121-130.

Guan, L. (2009): Preparation of future weather data to study the impact of climate change on buildings, *Building and Environment* 44, 793–800.

Guide to Using @RISK (2010), Version 5.7.

Hamby, D.M. (1994): A review of techniques for parameter sensitivity analysis, *Environmental Monitoring and Assessment*, vol. 32, no. 2, pp. 135-154.

Helton, J.C. (2006): Uncertainty and Sensitivity Analysis for Models of Complex Systems, *Proc. IMAC-XXIV: Conference & Exposition on Structural Dynamics - Looking Forward: Technologies for IMAC*.

Holm, A. and Kuenzel, H. (2002): Practical application of an uncertainty approach forhygrothermal building simulations—drying of an AAC flat roof, *Building andEnvironment* 37 (8–9) 883–889.

Hopfe, C. J. (2009) : *Uncertainty and sensitivity analysis in building performance simulation for decision support and design optimization*, PhD thesis.

Hurtado, J.E. and Barbat, A.H. (1997): Simulation methods in stochastic mechanics, in: J. Marczyk (Ed.), *Computational Stochastic,Mechanics in a Meta-computing Perspective*, CIMNE, Barcelona, pp. 93–116.

IPCC(2007): *Climate Change : The Physical Science Basis*, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA.

Jacob, D., Burhenne, S., Herkel, S., Wagner, A., Dodier,R.H. and Henze, G.P. (2011): Comparing Two Methods of Stochastic Modeling for Buildings , *Proc. Building Simulation 2011,* Sydney.

Kavcic, M., Balaras, C., Gudnason, G., Guruz, R., Kaiser, J., Katranuschkov, P., Laine, T., Mansperger, T., Pappou, T., Černe B. and Zahedi, A. (2012): *ISES Deliverable D1.1: Gap Analysis*, © ISES Consortium, Brussels.

Kim, Y.-J., Oh, S.-M., Park, C.-S., Kim, I.-H. and Kim, D.-W. (2011): Self-Activating Uncertainty Analysis for Bim-Based Building Energy Performance Simulations, *Proc. Building Simulation 2011,* Sydney.

Lomas, K.J., and Eppel, H. (1992): Sensitivity analysis techniques for building thermal simulation programs, *Energy and Buildings* 19, no. 1: 21-44.

Lu, J.-C., Shuen-Lin Jeng, Wang, K. (2010): A Review of Statistical Methods for Quality Improvement and Control in Nanotechnology, *Journal of Quality Technology*, Vol. 41, No. 2, 148-164.

MacDonald, I. and Strachan, P. (2001):Practical application of uncertainty analysis, *Energy and Buildings* 33 (3) 219–227.

MacDonald, I. (2002): *Quantifying the Effects of Uncertainty in Building Simulation*, Phd Theses, University of Strathclyde, July 2002.

MacDonald, I. (2009): Comparison of Sampling Techniques on the Performance of Monte-Carlo Based Sensitivity Analysis, in Building Simulation, *11th International IBPSA Conference Glasgow*, Scotland, July 27-30.

Mansperger T., Guruz R., Guðnason G., Kavcic M., Laine T., Forns-Samso F., Zahedi A., Zingel U., Zschippang S., Katranuschkov P., (2012): *ISES Deliverable D1.2: Use Case Scenarios and Requirements Specification*, © ISES Consortium, Brussels.

Matala, A. (2008): *Sample Size Requirements for Monte Carlo simulations using Latin Hypercube sampling*, Helsinki university.

Meester, T. D., Marique, A.-F., Herde, A. D. and Reiter, S. (2010): Impactsof occupant behaviours on residential heating consumption for detached houses ina temperate climate of the northern part of Europe, *Energy and Buildings*, doi:10.1016/j.enbuild.2012.11.005.

Morgan, M.G. and Henrion, M. (1990): *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*,Cambridge University Press, NY.

Muthén, L.K. and Muthén, B.O. (2002): How to use a Monte Carlo study to decide on sample size and determine power, *Structural Equation Modeling*, 4, 599-620.

Page, J. , Robinson, D., Morel, N. and Scartezzini, J.-L. (2008): A generalised stochastic model for the simulation of occupant presence, *Energy and Buildings* 40, 83–98.

Papadrakakis, M. and Lagaros, N.D. (2002): Reliability-based structural optimization using neural networks and Monte Carlo simulation, *Methods Appl. Mech. Engrg.* 191, 3491–3507.

Pachauri, R.K. (2005) in: *Proceedings of the 11th Conference of the Parties to the United Nations Framework Convention on Climate Change and 1st Conference of the Parties serving as Meeting of the Parties to the Kyoto Protocol*, Montreal, Canada, http://www.ipcc.ch.

Papoulis, A. and Pillai, S. U. (2002): *Probability, Random Variables, and Stochastic Processes*, Mcgraw-Hill Publ.Comp.; 4th edition, ISBN: 978-0071226615.

Rackwitz, R. and Fiessler, B. (1978): Structural Reliability under Combined Random Load Sequences, in: *Comp. & Struct.* 9, pp. 484-494.

Roulet, C.-A. (1999): Confidence of simulation results: put a sensitivity analysis module inyour MODEL. The IEA-ECBCS Annex 23 experience of model evaluation, *Energy and Buildings* 30, 61–71.

Saltelli , A., Ratto, S., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008): *Global sensitivity analysis*,Wiley & Sons.

Santiago, J., Corre, B., Claeys-Bruno, M. and Sergent (2012): Improved sensitivity through Morris extension, *Chemometrics and Intelligent Laboratory Systems* 113 52–57.

Santin, O. G. (2011): Behavioural Patterns and User Profiles related to energy consumption for heating, *Energy and Buildings* 43 2662-2672.

Sardianou, E. (2008): Estimating space heating determinants : An analysis of Greek households, *Energy and Buildings* 40 1084 – 1093.

Schölzel, C. and Hense, A. (2011): Probabilistic assessment of regional climate change in Southwest Germany by ensemble dressing, *Clim Dyn* 36:2003–2014.

Schüller, G. I. (1981): *Introduction to Safety and Reliability of Structures (Einführung in die Sicherheit und Zuverlässigkeit von Tragwerken)*, Ernst & Sohn, Germany, ISBN: 3-433-00874-4.

Turner, C. and Frankel, M. (2008): *Energy Performance of LEED for New Construction Buildings*,NBI, www.newbuildings.org.

Vanmarcke, E. (1983): *Random Fields, Analysis and Synthesis*, MIT Press, ISBN: 0-262-22026-1.

Wang, L., Mathew, P. and Pang, X. (2012): Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building, *Energy and Buildings* 53 , 152–158.

Wan, K.S.Y. and Yik, F.H.W. (2004): Representative building design and internal loadpatterns for modelling energy use in residential buildings in Hong Kong, *Applied Energy* 77, 69–85.

Zador, J., Zsely, I.Gy. and Turanyi, T. (2006): Local and global uncertainty analysis ofcomplex chemical kinetic systems. *Reliability Engineering & System Safety* 91, no. 10-11: 1232-1240.

Zhao, J. (2012): *A systematic method for material characterization and a stochastic approach for simulating hygrothermal performance of building enclosure assemblies*, PHD thesis.

Appendix I: Acronyms

BIM	Building Information Modelling
BPS	Building Performance Simulation
CAD	Computer Aided Design
DSA	Differential Sensitivity Analysis
eeB	Energy Enhanced Building
FAST	Fourier Amplitude Sensitivity Test
FM	Facilities Management
HESMOS	EU Project No 260088 "ICT Platform for Holistic Energy Efficiency Simulation and Lifecycle Management Of Public Use FacilitieS"
HVAC	Heating, Ventilation, Air Conditioning
ICT	Information and Communication Technology
IS	Importance Sampling
LHS	Latin Hypercube Sampling
MCS	Monte Carlo Sampling
PDF	Probability Distribution Function
RTD	Research and Technology Development
VBA	Visual Basic Application