GAUSSIAN PROCESS EMULATOR FOR OPTIMAL OPERATION OF A HIGH RISE OFFICE BUILDING

Ji-Eun Kang[1], Young-Jin Kim[2], Ki-Uhn Ahn[3], and Cheol-Soo Park[1*]
[1]Department of U-City design and Engineering, SungKyunKwan University, Suwon, South Korea
[2]School of Civil & Architectural Engineering, SungKyunKwan University, Suwon, South Korea
[3]Korea Infrastructure Safety & Technology Corporation, South Korea
*E-mail: cheolspark@skku.ac.kr

ABSTRACT

This paper reports the development of a Gaussian Process (GP) emulator to make an energy efficient operational decision by emulating complex and dynamic physical behavior of a building. The GP emulator demands much less time and effort as well as produces almost identical outputs compared to the whole-building simulation tools such as EnergyPlus. This study started from a request by an office-building owner who wanted to investigate whether EnergyPlus can be applicable for real time operation and management of a high-rise office building. The EnergyPlus model of the building developed by the authors required significant computation time for producing stochastic outputs. Under this context, the GP emulator was developed for operation strategies such as answering what-if scenarios in real time. This paper reports construction process and performance of the GP emulator. It is shown in this paper that the GP emulator is good enough to be used for real time optimal operation.

INTRODUCTION

High rise office buildings consum significant cooling energy due to high internal heat gain and trasnparent envelopes. It is highly required to provide energy efficient operation strategy (e.g. optimal chilled water temperature, water tempeature of cooling tower, etc.) in real time for better buildng energy management.

The whole building simulation tools (e.g. EnergyPlus, ESP-r, TRNSYS, etc.) have been widely used to answer such what-if scenarios. The whole building simulation tools can provide stochastic prediction using Monte Carlo simulation. However, it takes significant time and efforts to develop a building energy model using the whole building simulation tools. In addition, computation time of the tools is not negligible. In other words, building operation managers or building owners cannot derive stochastic prediction results considering the uncertainty of the building in real time.

The Gaussian Process (GP) emulator has recently received much attention due to its remarkable post-processing capabilities in building simulation. The GP emulator, based on Bayesian approach, is a surrogate model, which can mimic complex and dynamic physical behaviors of a building (systems) while being computationally cheap to run (Fricker et al, 2011). Brown et al (2012) proposed a kernel regression technique for emulating a building energy model. Heo & Zavala (2012) showed an applicability of her GP model for Measurement and Verification (M&V) practices. Eisenhower et al. (2012) suggested that the GP emulator be used to perform a stochastic optimal design of building systems. In general, it has been reported that the GP emulator produces reliable stochastic prediction without demanding significant time and efforts compared to the whole-building simulation tools (Kennedy & O'Hagan, 2001; Oakley & O'Hagan, 2004; Goldstein & Rougier, 2006; Rasmussen & Williams, 2006; Liu & West, 2009; Brown et al, 2012; Heo & Zavala, 2012; Eisenhower et al, 2012).

This paper addresses the following: development of the GP emulator for a large-scale office building, search for optimal operation strategies to reduce electricity consumption of chilled water systems. The GP emulator provides very fast stochastic results for building operation and energy management. The results from the GP emulator is approximately identical to those of dynamic simulation tools.

BUILDING DESCRIPTION

The client asked the authors to develop a simulation model to investigate applicability of EnergyPlus in their daily operation and management of a real high rise office building. The client wanted to know whether EnergyPlus could be a right choice for the aforementioned purposes. One of the request by the client is to develop the EnergyPlus simulation model to be as close to the reality as possible. The authors had to report the client our assumptions and professional judgment during the modeling process.

The building is a telecommunication company's headquarter building in Seoul, Korea (Figure 1). The building has 33 stories above ground and 6 underground levels, and the total floor area is 91,898 m^2. The main exterior materials are low-e double pane glazing, and the window-wall ratio is approximately 70%.

Primary system components include three steam boilers, one centrifugal turbo chiller for AHU, and two centrifugal chillers for an ice storage system, two absorption chillers for backup, and seven heat exchangers. Secondary system components placed between heating and cooling plants and each zone of the building are composed of Constant Air Volume (CAV) for lobby, Variable Air Volume (VAV) for office spaces, Fan Power Unit (FPU), and Fan Coil Unit (FCU) for extreme control zones for high level of comfort condition. The distribution systems of the HVAC include variable or constant fans and pumps. In particular, the Building Energy Management System (BEMS) includes 1,692 on-line sensor points.

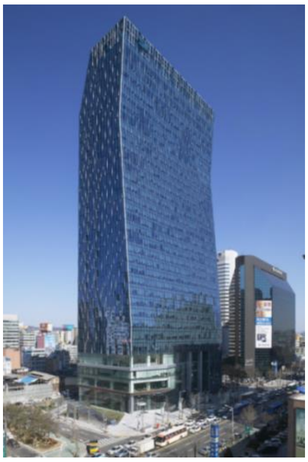


Figure 1 Target building

Firstly, the authors developed an EnergyPlus model as shown in Figure 2. The size of the Input Data File (IDF) is 25MB and the number of the classes in the IDF is 125. The number of the zones is 785 and the number of the surfaces is 3,802. It took almost two hours to run one month simulation. (used PC: Intel i7-870 (2.93 GHz) CPU and 6 GB of memory).
As inferred from the simulation runtime, the EnergyPlus simulation model was not qualified for real time decision making and energy management. The following section will describe the development of the GP emulator based on the EnergyPlus simulation model of the building.

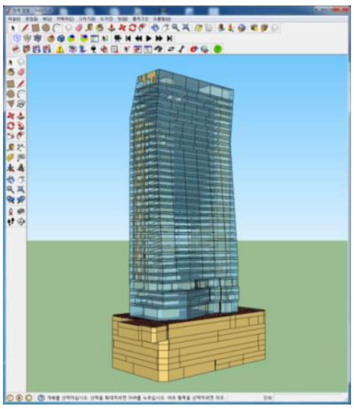


Figure 2 EnergyPlus simulation model

GP EMULATOR

In general, the construction of the GP emulator takes four steps as follows: (1) selection of training dataset, (2) construction of Gaussian regression model, (3) Bayesian approach, and (4) uncertainty or sensitivity analysis (Figure 3).

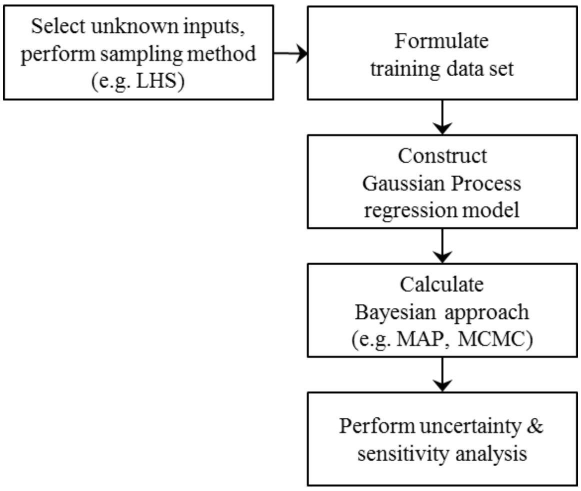


Figure 3 Development of GP emulator

The training dataset is composed of unknown inputs and simulation outputs. To generate the dataset, Latin Hypercube Sampling (LHS) method has been widely used to generate simulation cases based on probability distribution (min, base, max) of unknown inputs. The training dataset propagated by LHS method is used to construct a linear regression model with Gaussian noise. The GP is based on Gaussian probability distribution and can give a very general treatment of a GP regression function (O'Hagan, 1978). However, it should be noted that the GP regression function has three unknown parameters such as scaling parameter, length-scale, and variance of Gaussian noise, and the posterior distribution of unknown parameters needs to be estimated according to Bayesian approach (Busby, 2009; Vanhatalo et al, 2011).
The Bayesian approach has been widely used by specifying the individual belief (or prior knowledge) into probability distribution, and minimizing the unintended subjective effect. In general, a maximum a posterior (MAP) and Markov Chain Monte Carlo (MCMC) estimate can be used for determining the posterior distribution of those.

Chilled water system of the study

For this study, three chilled water systems were chosen: two chillers for the ice thermal storage system and one turbo chiller for cooling. The ice thermal storage system is used for cooling of the entire building while one turbo chiller is used for 24 hours continuous cooling of an IT center located in 22F in the building.

As mentioned above, the GP emulator was developed to find optimal operation strategies in real time decision making. The scenarios of our interests are to find optimum T_{chws} and T_{ctws}.

- T_{chws}: chilled water temperature of supply-side in the chilled water loop
- T_{ctws}: cooling water temperature of supply-side in the cooling tower loop

Figure 4 shows a schematic diagram of the given chilled water system. The system consumes significant electricity to generate the chilled water. In such large office facilities, one of the dominant energy consumers is cooling energy. Therefore, it is important to optimally control the chilled water system based on the accurate simulation model (in this study, a meta-model or the GP emulator) that predicts the system's characteristics (Monfet & Zmeureanu, 2011).

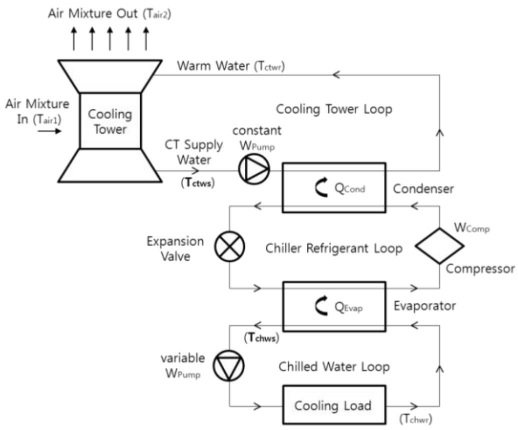


Figure 4 Chilled water system of the study

The chilled water system has two identically-sized ice thermal storage chillers and one trubo chiller. Two absorption chillers were excluded from the GP emulator since they are installed as backup and rarely used. The two ice thermal storage chillers have a rated capacity and COP of 1,531 (*kW*) and 4.31 (*W/W*). The two cooling towers for the chillers have a rated capacity, fan power and water flow of 879 (*kW*), 7.46 (*kW*), and 0.42 (m^3/min) respectively. The turbo chiller has a rated capacity and COP of 704 (*kW*) and 3.35 (*W/W*). The cooling tower for the turbo chiller has a rated capacity, fan power and water flow of 879 (*kW*), 7.46 (*kW*) and 0.42 (m^3/min).

Simple feasibility test of energy saving by EnergyPlus

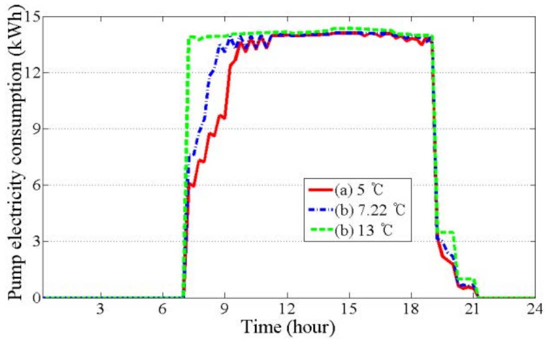
As mentioned above, the given problem is to find optimal T_{chws} and T_{ctws} as control variables. If we increase T_{chws} (ranging from 5.0°C to 13.0°C) and T_{ctws} (ranging from 28.0°C to 32.0°C) with increment of 0.1°C, the possible number of combination becomes 3,321 (81 × 41 = 3,321).

Rather than attempting to simulate all possible combination by EnergyPlus, the authors firstly tested the following three cases to investigate importance of finding optimal T_{chws} and T_{ctws}.

- Case 1: T_{chws} of 5.0°C (minimum T_{chws}) (Lee & Cheng, 2012)
- Case 2: T_{chws} of 7.22°C (default T_{chws} in EnergyPlus)
- Case 3: T_{chws} of 13.0°C (maximum T_{chws} in EnergyPlus) (Lee & Cheng, 2012)

For the three cases, T_{ctws} was set to 29.4 °C or a default value of T_{ctws} in EnergyPlus.

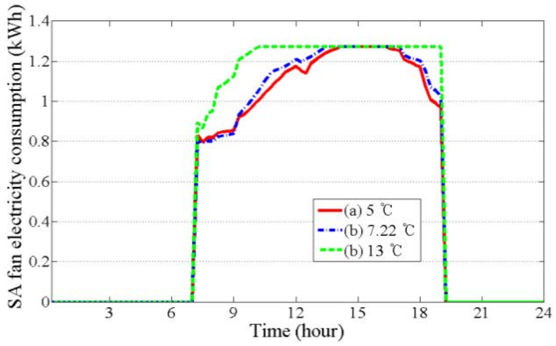
As shown in Figure 5, the pump and fan electricity consumptions are proportional to T_{chws} (Figure 5(a), 5(b)) and the cooling electricity consumption is inversely proportional to T_{chws} (Figure 5(c)). The total energy saving per day by using optimal T_{chws} is 10.1%, which is significant (Table 1).



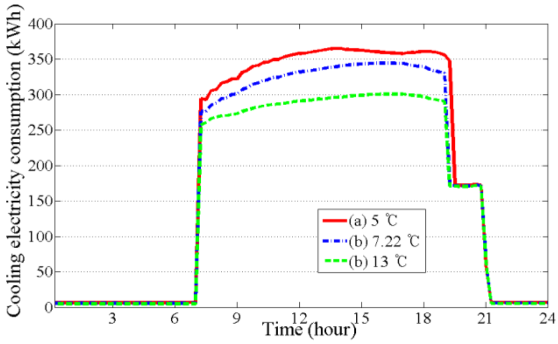
(a) Pump electricity consumption

Table 1 Electricity energy consumptions for 1st August

T_{CHWS} (°C]	ELECTRICITY CONSUMPTION (KWH)				SAVING RATE (%)			
	PUMP	SA FAN	COOLING	TOTAL	PUMP	SA FAN	COOLING	TOTAL
5.0	627	53	18,446	19,126	-	-	-	-
7.22	655	54	17,173	17,882	-4.5	-1.9	6.9	6.5
13.0	695	59	15,322	16,076	-6.1	-9.3	10.8	10.1



(b) Supply air fan electricity consumption



(c) Cooling electricity consumption

Figure 5 Three sample cases on 1st August

Generation of training set

Table 2 shows five unknown inputs which are assumed to have a normal distribution. The normal distribution maximizes the information entropy among all distributions with a known mean (μ) and standard deviation (σ) (Simlab, 2011). For uncertainty propagation, LHS method was employed. LHS method provides good coverage of the input space with relatively few samples compared to the standard brute force random sampling (Saltelli et al., 2004). A total of simulation runs was set to 70. The number of generated samples was well above the value of 10 × k (where, k is five unknown inputs as shown in Table 3) (MUCM, 2013). The confidence interval was set to 95%.

Table 2
Unknown inputs for uncertainty propagation

UNKNOWN INPUTS		MEAN	STANDARD DEVIATION
Ice thermal storage chiller1	COP	4.31	0.862
Ice thermal storage chiller2	COP	4.31	0.862
Turbo chiller	COP	3.35	0.670
Temperature set-points(°C)	T_{chws}	7.22	0.361
	T_{ctws}	29.4	1.470

As shown in Table 3 and Figure 6, cooling electricity consumption has greater variance than pump and fan electricity consumption. In other words, the cooling electricity consumption is more sensitive to the probability ranges of the selected unknown inputs than the others do.

With regard to computation time, the simulation runs (70 cases) took about 2 days for one month simulation (from 1st August to 31th August 2012).

Table 3
Uncertainty results of pump, fan, cooling, and total electricity consumption (kWh/m^2)

ELECT-RICITY CONSUMP-TION	MEAN	STANDARD DEVIATION	CONFIDENCE INTERVAL	
			2.5%	97.5%
Pump	1.27	0.0047	1.26	1.28
Fan	0.91	0.0028	0.91	0.92
Cooling	4.42	0.2189	3.99	4.85
Total	6.60	0.2172	6.18	7.03

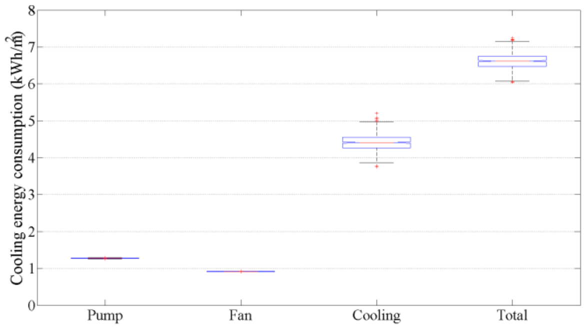


Figure 6 Uncertainty results using boxplot

Formulation of GP emulator

The GP emulator was constructed using GPstuff software. GPstuff is in-house M-scripts coded in MATLAB developed by Vanhatalo et al (2011). The GPstuff is a collection of MATLAB functions to build and analyse Bayesian models built over Gaussian Processes. Three unknown parameters (scaling parameter, length-scale, and variance of Gaussian noise) of the GP regression function were calculated using MAP estimate, which it is relatively easy and fast to evaluate (Vanhatalo et al, 2011).

Validation of GP emulator

To verify the GP emulator, the exhaustive method was adopted. The exhaustive method is to test fall possible cases. In this study, another set of 70 simulation cases were newly made for validation purpose and cross-compared to each other (EnergyPlus and GP emulator).

Table 4 shows comparison results. The differences in mean values of pump, fan, cooling and total energy between EnergyPlus and the GP emulator are 0.00, 0.01, 0.02, and 0.01, respectively, which is significantly close to each other. The GP emulator developed in this study can be used as a surrogate model of EnergyPlus.

Figure 7 shows total electricity consumption using Cumulative Distribution Function (CDF). The developed GP emulator accurately mimics dynamic simulation.

Table 4
Comparison of stochastic prediction results between EnergyPlus and GP emulator

STOCHASTIC RESULTS		MEAN	STANDARD DEVIATION
Pump (*kWh/m²*)	EnergyPlus	1.26	0.0043
	GP emulator	1.26	0.0046
Fan (*kWh/m²*)	EnergyPlus	0.90	0.0010
	GP emulator	0.89	0.0023
Cooling (*kWh/m²*)	EnergyPlus	4.65	0.2301
	GP emulator	4.67	0.0023
Total (*kWh/m²*)	EnergyPlus	6.81	0.2302
	GP emulator	6.82	0.2731

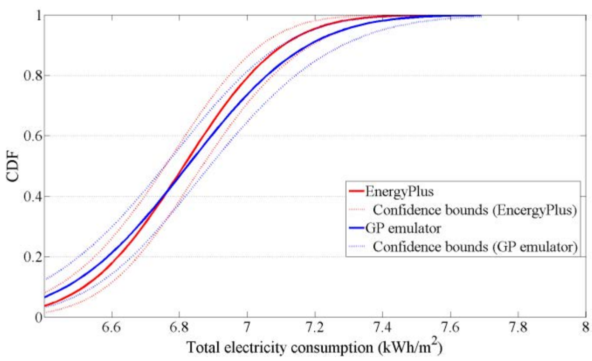


Figure 7 Uncertainty results of total electricity consumption

RESULTS

For the search of optimal T_{chws} and T_{ctws}, a Pareto dominance criterion was used. The Pareto dominance criterion, which is one of the principles of mean-standard deviation, identifies alternatives having both low mean and low standard deviation values. The remaining alternatives, which are not dominated by any other alternatives, are called "Pareto optimal set" or "efficient frontier" (Charnes et al, 1985). The GP emulator generated uncertainty results for 3,321 cases. It took about 10 minutes.

Table 5 shows non-dominated Pareto solutions in terms of energy consumption. In terms of pump and fan electricity consumption, the non-dominated Pareto solutions are T_{chws} of 5.0°C and are independent of T_{ctws}. Since pumps in the chilled water system is a variable speed, the mass flow rate through the pump is reduced when T_{chws} has lower termpature.

In terms of cooling electricity consumption, the non-dominated Pareto solution is T_{chws} of 13.0°C (maximum T_{chws}) and T_{ctws} of 28.0°C (minimum T_{ctws}).

The non-dominated Pareto solution in terms of the total electricity consumption is equal to that of cooling electricity consumption. In other words, the cooling electricity consumption is dominant in total energy consumption. The optimal solution of the pump and fan electricity consumptions have lower variance than that of the cooling electricity consumption.

Table 5
Non-dominated Pareto solutions

OPTIMAL SOLUTION	CHILLED WATER SYSTEM		ELETRICITY CONSUMPTION (*kWh/m²*)	
	T_{chws} (°C)	T_{ctws} (°C)	MEAN	STANDARD DEVIATION
Pump	5.0	28.0~32.0	1.2609	0.0045
Fan	5.0	28.0~32.0	0.893	0.0022
Cooling	13.0	28.0	3.989	0.1565
Total	13.0	28.0	6.217	0.1548

Figure 8 shows a non-dominated Pareto solution in terms of the total electricity consumption. The chosen Pareto optimal solution has the least mean and standard deviation.

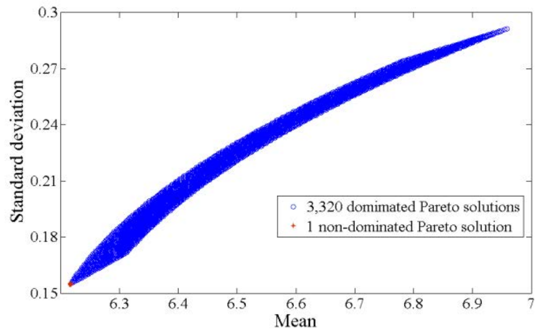


Figure 8 Results of Pareto dominance (non-dominated vs. dominated solution)

To validate the non-dominated Pareto solution, the authors compared the optimal solution in terms of the total electricity consumption with randomly chosen eight dominated Pareto solutions (Table 6). The non-dominated Pareto solution is superior to the dominated Pareto solutions.

Table 6
Validation of Pareto optimal solution

DESIGN	CHILLED WATER SYSTEM		TOTAL ELETRICITY CONSUMPTION (kWh/m^2)	
	T_{chws} (°C)	T_{ctws} (°C)	MEAN	STANDARD DEVIATION
Non-dominated	***13.0***	***28.0***	***6.217***	***0.1548***
Dominated #1	5.0	28.0	6.777	0.2736
Dominated #2	7.0	28.0	6.577	0.2440
Dominated #3	9.0	28.0	6.417	0.2141
Dominated #4	11.0	28.0	6.296	0.1842
Dominated #5	13.0	29.0	6.241	0.1591
Dominated #6	13.0	30.0	6.265	0.1633
Dominated #7	13.0	31.0	6.289	0.1676
Dominated #8	13.0	32.0	6.313	0.1718

CONCLUSIONS AND FUTURE WORK

In this study, the GP emulator was developed to replace the EnergyPlus simulation model for real time optimal decision making in building energy management. The GP emulator deveoped in this study is accurate enough compared to the whole building simulation tool (e.g. EnergyPlus) and is very fast for energy prediction.

The GP emulator was used for stochastic simulation to determine optimal chilled water temperatures from the chillers and cooling tower. The GP emulator could successfully find a Pareto-optimal solution set. The GP emulator may replace the use of the whole building simulation tools for real time energy management and optimal decision making. Future study may include embedding the GP emulator to real Building Energy Management System with sensor network. The online self-calibration of the model may be of need in near future for prediction accuracy and easy application.

ACKNOWLEDGEMENT

This work is financially supported by the Korea Minister of Ministry of Land, Transport and Maritime Affairs (MLTM) as “U-City Master and Doctor Course Grant Program”.

REFERENCES

Brown, M., Barrington-Leigh, C. and Brown, Z. 2012. Kernel regression for real-time building energy analysis, Journal of Building Performance Simulation Available Online, In Press.

Busby, D. 2009. Hierarchical adaptive experimental design for Gaussian process emulators, Reliability Engineering and System Safety, Vol. 94, pp. 1183-1193.

Charnes, A., Cooper, W.W., Golany, B. and Seiford, L. 1985. Foundations of data envelopment anaylsis for Pareto-koopmans efficient emprical production functions. Journal of Econometrices, Vol. 30, pp. 91-107.

Eisenhower, B., O’Neil, Z., Narayanan, S., Fonoberov, V.A. and Mezic, I. 2012. A methodology for meta-model based optimization in building energy models, Energy and Buildings, Vol. 47, pp. 292-301.

Fricker, T.E., Oakley, J.E., Sims, N.D. and Worden, K. 2011. Probabilistic uncertainty analysis of an FRF of a structure using a Gaussian process emulator, Mechanical System and Signal Processing, Vol. 25, pp. 2962-2975.

Goldstein, M. and Rougier, J.C. 2006. Bayes linear calibrated prediction for complex systems, Journal of the American Statistical Association, 101, pp. 1132-1143.

Heo, Y.S. and Zavala, V.M. 2012. Gaussian process modeling for measurement and verification of building energy savings, Energy and Buildings, Vol. 53, pp. 7-18.

Kennedy, M.C. and O’Hagan, A. 2001. Bayesian calibration of computer models, Journal of the Royal Statistical Society, Series B, 63, Part 3, pp. 425-464.

Lee, K.P. and Cheng, T.A. 2012. A simulation-optimization approach for energy efficiency of chilled water system, Energy and Buildings, Vol.54, pp.290-296.

Liu, F. and West, M. 2009. A Dynamic Modelling Strategy for Bayesian Computer Model Emulation, Bayesian Analysis, pp. 393-412.

Managing Uncertainty in Complex Models. 2013. http://mucm.aston.ac.uk/MUCM/MUCMToolkit/index.php?page=MetaHomePage.html, last accessed on 30 Jan 2013.

Monfet, D. and Zmeureanu, R. 2011. Identification of the electric chiller model for the EnergyPlus program using monitored data in an existing cooling plant, Proceedings of Building Simulation 2011, Syydney, 14-16 November.

Oakley, J. and O’Hagan, A. 2004. Probabilistic sensitivity analysis of complex models: a bayesian approach, J. R. Statist. Soc. Ser. B 66, pp. 751–769.

O’Hagan, A. 1978. Curve fitting and optimal design for prediction, Journal of the Royal Statistical Society, Series B (Methodological), Vol. 40, No. 1, pp. 1-42.

Rasmussen, C.E. and Williams, C.K.I. 2006. Gaussian Processes for Machine Learning, the MIT Press, 2006, ISBN 026218253X.

Saltelli, A., Tarantola, S., Campolongo, F. and Ratto, M. 2004. Sensitivity Analysis in Practice, Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England.

Simlab. 2011. http://simlab.jrc.ec.europa.eu/, version 3.2, last accessed on 30 Jan 2011

Vanhatalo, J., Riihimaki, J., Hartikainen, J. and Vehtari, A. 2011. Bayesian modeling with Gaussian processes using the MATLAB toolbox GPstuff. submitted.