

Multi-Linear Regression Modelling for Prediction of Net Energy Consumption of a HVAC System

Dr. Arvind A R¹, Nagamanikkam N², Priyadharshan S M³, Gaytheswar P R⁴, Vishal S⁵

Deputy General Manager, Project Planning, Ashok Leyland, India¹

Manager, Project Planning, Ashok Leyland, India²

Department of Mechanical Engineering, PSG College of Technology, India^{3,4,5}

ara.arvind@gmail.com, nagamanikkam.n@hotmail.com

priyadharshansm@gmail.com, reach.gaytheswar.pr@gmail.com, vishalsankar115@gmail.com

Abstract: Prediction of cooling load requirements is an important step [1], [2], [6] in designing HVAC (Heating, Ventilation & Air Conditioning) systems for maintaining correct set-point temperatures and humidity levels. Our study aims to predict the net energy for a HVAC system in an industrial unit using sensitivity analysis data from DesignBuilder. Two HVAC systems — Unitary Heat Cool and VAV Reheat DX Cooling with Dehumidification — were compared for cooling performance and energy consumption across seven Indian cities. With the data, a multi – linear regression model is built in Microsoft Excel with variables like plant area, cooling set-point temperature, internal gains, and climate factors to estimate the net energy consumption of the HVAC system. This simple model has potential to be built into a more extensive model which can predict the net energy consumption of whole infrastructures with minimal input requirements.

Keywords: HVAC, Prediction modelling, Design builder, Simulation

I. INTRODUCTION

The need of an effective HVAC system design has increased in order to optimize energy resources and cost. Simulation tools like DesignBuilder enable performance analysis [3], [9], [40] of different HVAC systems across climates before implementation.

Investigation has been carried out on performance variations of different HVAC systems [2], [3], [5] in a closed room of different areas (4000 and 6000 m²) in a part of a manufacturing plant under varying weather conditions and internal loads. By simulating cooling energy usage for both simple and complex HVAC systems across seven Indian cities, factors affect energy & performance were identified and system capacity estimates projected using annual energy outputs.

OBJECTIVES

With the above understanding, the set of objectives are framed.

- To compare two different HVAC system templates in DesignBuilder across multiple climates
- To determine the sensitivity of cooling energy consumption with respect to input parameters like set-point temperature, internal gains, and location
- To estimate system sizing (in Tonnage TR) using energy consumption outputs
- To derive regression models for predicting the cooling energy in future scenarios



II. LITERATURE REVIEW

Multiple nonlinear regression (MNR) models have effectively predicted short-term cooling loads in public buildings by using same-day calibration. They significantly outperform traditional regression methods in accuracy and practicality for real-world HVAC operation [2].

Validation studies with DesignBuilder/EnergyPlus have shown how important model inputs are. User-defined internal load schedules greatly improved annual electricity prediction accuracy, reducing error from 18% to 0.2%. In contrast, customized weather data had little effect [3]. Other studies used DesignBuilder to assess energy conservation measures (ECMs) like VAV systems, photoelectric dimming, and double-glazed windows, achieving savings up to 26.5% without sacrificing comfort [4], [18].

Comparative analyses of residential HVAC systems with tools like EnergyPlus and OpenStudio show that unitary systems usually result in the lowest annual energy use. However, simplified components such as “ideal air loads” often overestimate consumption, highlighting the need for accurate COP modeling [5], [13], [16].

A basic classification of HVAC systems based on configurations such as all-air, all-water, and unitary is seen as crucial for system selection, tailored to space, load, and cost limits [1]. Using simulation-based optimization further improves energy design when combined with inputs from climatology and solar engineering, especially during the early design stages [9], [26], [27].

Machine learning methods like support vector regression and neural networks have shown high accuracy in predicting building energy use [11], [22], [32]. However, they often require large datasets and careful tuning, making them less practical for early-stage or small-scale projects. In such cases, simpler methods like multi-linear regression can provide reliable predictions with minimal inputs when key variables are clearly defined [2], [10].

For model calibration, recent advancements have focused on matching simulations with measured performance through automated or semi-automated calibration techniques [12], [21], [24]. Calibrated models also serve as a basis for predictive control, including model predictive control (MPC) strategies to optimize HVAC operation under uncertainty [14], [29], [33], [34].

Thermal comfort modeling and control, another key design aspect, relies on understanding occupant behavior, internal heat gains, and passive cooling strategies [23], [28], [31]. Additionally, lifecycle assessments help in selecting efficient and sustainable systems from both performance and environmental perspectives [15], [19], [36].

Broader reviews have looked into building stock models [35], uncertainty analysis in simulation workflows [39], and the integration of energy modeling in net-zero and climate-responsive buildings [25], [37], [38].

To minimize the need for repeated simulations, a regression model was developed in Microsoft Excel. Statistical methods detailed in established regression modeling literature [10] were used, including dummy variables for categorical inputs like city data, along with a trial-and-error process to refine predictive performance.

III. METHODOLOGY

The modelling and analysis were performed using **DesignBuilder**. As a widely used software in the building industry for energy analysis, especially for HVAC design. DesignBuilder provides a visual interface that simplifies many of EnergyPlus’s complex features [40] and better suited for the study.

Initial Setup and Building Modelling

Modelling of two closed rooms of size 6000 m² and 4000 m² assumed to be a part of a manufacturing unit. The building shape, wall thickness, height, and roof type were input using DesignBuilder’s block-building tool. Built-in construction templates were used with good insulation properties for walls and roofs varying the thermal resistances (R-values) of every layer.

A single thermal zone was used to represent the main working area, and an upper zone was added to place additional components in the building like Thermocol false ceiling and PUF roof insulation. The internal loads (like equipment and lighting) were added using estimated power densities based on common industrial values (15–45 W/m²) [7], [8].



HVAC Systems and City Selection

Two HVAC system templates were used:

Unitary Heat Cool (a simple DX system commonly used in small factories or commercial spaces)

VAV Reheat DX Cooling with Dehumidification (a more advanced, multi-zone system capable of better humidity control).

Both systems were selected from DesignBuilder's inbuilt HVAC templates without making deep modifications to their internal components (like exact fan or coil specifications). Set-point managers were left as default for the most part, though cooling set-points were modified in the activity templates to study their impact.

The designed systems were test-simulated under **seven Indian cities viz.**, Chennai, Kochi, Bangalore, Lucknow, Alwar, Nagpur, and Pune. The idea was to cover different climate zones — hot and humid (like Chennai), composite (like Nagpur), and moderate (like Pune). The weather files (EPW) were selected from DesignBuilder's own library and from the ISHRAE website.

Simulation and Sensitivity Analysis

After the basic model was built and the HVAC templates were assigned, simulations and sensitivity analysis were run for each city based scenario with variations in three key inputs:

Cooling Set-point Temperature (from 19°C to 23°C).

Power Density (15 to 45 W/m²).

Occupancy Density (0.05 to 0.15 people/m²).

A sensitivity analysis module [17] inside DesignBuilder was used, which runs the same model many times with different combinations of these inputs. There are a total of 11 input variables (cooling set-point temperature, power density, occupancy density, the two types of HVAC templates, and the seven cities). A total of 500 data-points was collected from the simulation for one area value, thus making it 1000 data-points for the two areas modelled.

Based on simulated trial runs on the model outputs were collected:

Cooling (Electric) [kWh] – It is the input load given to the HVAC system

Sensible Cooling Energy [kWh] – It is the total sensible cooling energy produced by the HVAC system

Total Cooling Energy [kWh] – It is the total cooling energy produced by the HVAC system

These outputs were exported into Excel for regression models testing using LINEST to predict cooling energy based on the inputs. One city (Pune) was left out of the dummy variables to avoid multi-collinearity.

IV. RESULTS AND DISCUSSION

HVAC System Comparison

The VAV reheat system offered better humidity control but significantly higher cooling energy use due to continuous fan operation and reheat loads. Unitary Heat Cool, while simpler, performed more efficiently in low humidity zones.

Sensitivity Analysis

The following are few statistical data from the sensitivity analysis report generated by DesignBuilder for one model (even though twomodels of different areas are being used, but the statistics are similar hence details of only one model is present):

Sampling and Simulation Summary

This section of the report lists the settings used for the uncertainty propagation and subsequent simulations runs undertaken. The failed iterations, if any, have been excluded from the results presented here and have not been used in the analysis.

Sampling Method:	LHS
Samples Requested:	500
Samples Created:	498
Failed Iterations:	0



Successful Iterations:	498
-------------------------------	-----

Table 1 – Sampling data

Input Sample Details

This section of the report lists the statistics for the samples generated, and used in the analysis, for all the inputs. The spread of sampled space is captured in the corresponding graphs.

Input variable: Hourly weather data

Option Name	Frequency
Bangalore : 01	71
Alwar : 02	71
Pune : 03	72
Nagpur : 04	72
Lucknow : 05	71
Kochi : 06	72
Chennai : 07	69

Table 2 – Input Options Frequency: Hourly weather data

Input variable: Cooling set-point temperature

Mean	SD	Min	Q1	Median	Q3	Max
21	1.1	19	20	21	22	23

Table 3 – Summary Statistics: Cooling Set-point Temperature

Input variable: Occupancy density

Mean	SD	Min	Q1	Median	Q3	Max
0.1	0.0269	0.05	0.078	0.1	0.1	0.2

Table 4 – Summary Statistics: Occupancy Density

Input variable: Equipment power density

Mean	SD	Min	Q1	Median	Q3	Max
30	8.1	15	23	30	37	45

Table 5 – Summary Statistics: Equipment power density

Input variable: HVAC template (Detailed HVAC)

Option Name	Frequency
UNITARY HEAT COOL : 01	248
VAV REHEAT DX COOLING DEHUMIDIFICATION: 02	250

Table 6 – Input Options Frequency: HVAC Template Type

Output Uncertainty

This section of the report presents the uncertainty analysis results for all the outputs requested. Summary statistics and corresponding graphs captures the spread of the outputs due to the variation in the inputs.



Output: Cooling (Electricity)

Summary Statistics:

Mean	SD	Min	Q1	Median	Q3	Max
1049008	381132	441500	733098	982037	1327961	2273545

Table 7. Summary Statistics - Cooling (Electricity)

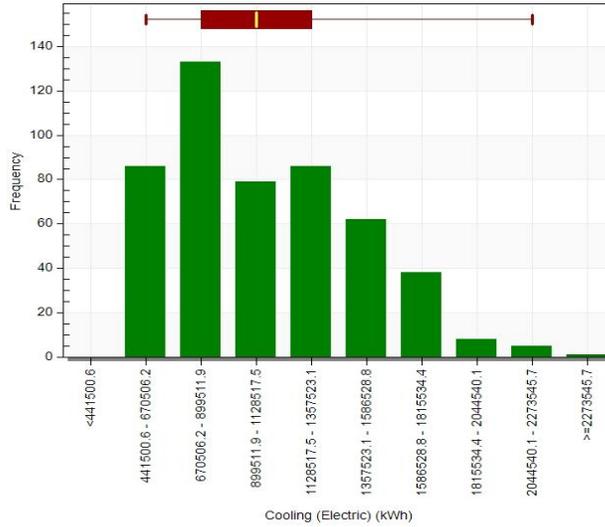


Figure 1 – Uncertainty Histogram: Cooling (Electric)

Output: Cooling energy

Summary Statistics:

Mean	SD	Min	Q1	Median	Q3	Max
3163716	1091533	1422094	2276278	2992100	3979350	6743693

Table 8. Summary Statistics - Cooling energy

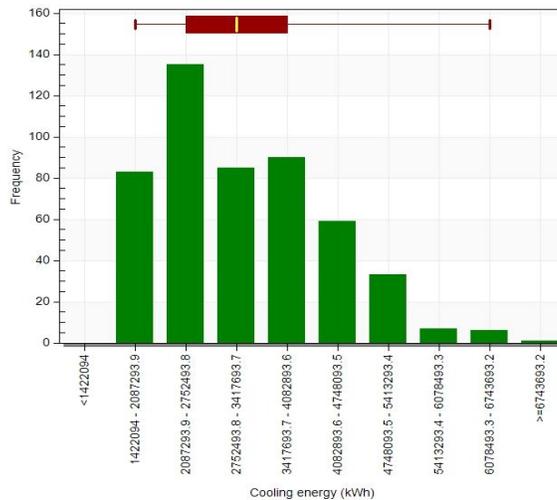


Figure 2 – Uncertainty Histogram: Cooling energy



Output: Sens Cooling energy

Summary Statistics:

Mean	SD	Min	Q1	Median	Q3	Max
-1929063	551540	-3741607	-2317773	-1819829	-1525858	-974998

Table 9. Summary Statistics – Sens Cooling energy

Output Uncertainty Histogram:

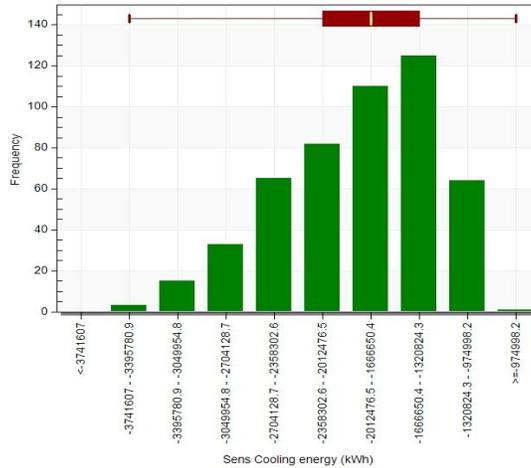


Fig 3. Uncertainty Histogram: Sensible Cooling energy

Regression Model

All the data points are put in Microsoft Excel and using its LINEST function, we are able to get the regression coefficients and standard error of each of the variables. The various regression coefficients are as follows:

Regression Coefficients

Variable	Cooling(Electricity)		Cooling energy		Sens Cooling energy	
	Reg. Coef.(a _i)	Std. Error	Reg. Coef.	Std. Error	Reg. Coef.	Std. Error
Nagpur	78527	8440.55	125656	24309.6	70343	11899.3
Lucknow	80401.9	8454.23	197263	24349	23237.2	11918.6
Kochi	234827	8448.21	729824	24311.7	127328	11910.1
Chennai	243735	8428.09	661745	24273.7	168264	11881.8
Bangalore	-46098	8458.69	-82467	24361.8	-73529	11924.9
Alwar	25379.8	8460.86	629.904	24368.1	12258.1	11928
HVAC Template (1 or 2)	600844.3	4530.2	1673747.8	13047.4	871860.9	6386.6
Power density (W/m ²)	9583.9	280.6	27970.1	808.1	23002.2	395.5
Occupancy Density(people/m ²)	3191653.1	84413.6	10009265.7	243119.2	4045629.8	119004
Cooling Set-point Temperature(°C)	-64582.3	2101.2	-177235.3	6051.7	-145264.3	2962.2
Area	133.656	2.25	429.7	6.5	235.727	3.17
Intercept	6153.86	48776	-283587	140480	1106895	68763.5

Table 10: Regression Coefficients



With these values we can formulate a regression equation for these output variables. Since we have 11 variables, the regression equation is:

$$Y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_{11}x_{11}$$

Where x_1, x_2, \dots, x_{11} are the input variables and a_1, a_2, \dots, a_{11} are its corresponding regression coefficients. a_0 is the intercept value.

With the formulated equation, prediction for the values for the three outputs with desired inputs is feasible. An example is illustrated below,

City: Chennai

HVAC system: VAV Reheat with DX Cooling Dehumidification

Power density: 15 W/m²

Occupancy Density: 0.05 people/m²

Cooling set-point: 22°C

Area: 5000 m²

Regression coefficients can be referred from the table and the corresponding input entered in the simulation model. For the variables associated with cities, the input value is in binary format. For the Chennai City, enter the value 1 for x_4 (the variable corresponding to Chennai) and 0 for the other city variables. Pune is being taken as the baseline in this regression model so it is excluded as a variable.

$$Y = 2895345 \text{ kWh (Estimated value of Cooling energy)}$$

To find the average annual Tonnage TR, the facility is assumed to be operated for 16 hours a day and 6 days a week, for a total of 52 weeks,

$$\text{Annual number of working hours} = 16 * 6 * 52 = 4992 \text{ hours}$$

$$\text{And } 1 \text{ TR} = 3.5 \text{ kW}$$

$$Y = 2895345 / (4992 * 3.5) = 165.7 \text{ TR}$$

d) Limitations

The limitations of the study are listed below,

Weather File Accuracy: The simulations used EPW weather files downloaded from the internet, which may not reflect recent climate shifts. Some cities may be represented by outdated data, slightly affecting load predictions.

DesignBuilder Learning Curve: The software requires careful input configuration. Parameters like internal gains, humidity set-points, and system schedules can produce unexpected results if not thoroughly understood.

Building Usage Assumptions: The plant was modelled with ideal usage patterns (e.g., continuous 16-hour operation), which does not acknowledge part-load behaviour.

No Cost or Carbon Analysis: The study focused on energy performance, but did not extend to cost, life-cycle impact, or carbon footprint — which could affect system selection in practice.

Tool-Specific Behaviour: The results are tightly coupled to EnergyPlus via DesignBuilder. Using another engine (e.g., TRACE, eQuest, or IES VE) may yield slightly different results.

V. CONCLUSION

This study shows that HVAC system type, climate, and internal load factors significantly affect cooling energy consumption in large industrial spaces. The regression model offers a reliable way to estimate energy demand in new locations which have similar climate profiles to the seven cities used in this regression model. The VAV system, though more advanced, may not always be energy efficient in low-latent-load regions.

REFERENCES

[1] Prof.S.M. Gheji, "Basic Classification of HVAC Systems for Selection Guide," International Journal of Innovative Research in Science, Engineering and Technology, vol. 5, no. 4, p. 10, 2016.



- [2] Fan Chengliang, Ding Yunfei, "Cooling load prediction and optimal operation of HVAC systems using a multiple nonlinear regression model," *Energy and Buildings*, vol. 197, p. 10, 2019.
- [3] Holly A. Wasilowski, Christoph F. Reinhart, "MODELLING AN EXISTING BUILDING IN DESIGNBUILDER/ENERGYPLUS: CUSTOM VERSUS DEFAULT INPUTS," *Proceeding of Building Simulation*, p. 8, 2009.
- [4] M. R. M. Rahman, "Energy conservation measures in an institutional building in sub-tropical climate in Australia," *Applied Energy*, vol. 87, no. 10, p. 10, 2010.
- [5] S. K. Alghoul, "A Comparative Study of Energy Consumption for Residential HVAC Systems Using EnergyPlus," *American Journal of Mechanical and Industrial Engineering*, vol. 2, no. 2, p. 6, 2017.
- [6] G. H. Zhengwei Li, "Re-evaluation of building cooling load prediction models for use in humid subtropical area," *Energy and Buildings*, vol. 62, p. 8, 2013.
- [7] ASHRAE (. American Society of Heating, 2024 ASHRAE Handbook)—HVAC Systems and Equipment, Atlanta, GA: ASHRAE, Inc., 2024.
- [8] ASHRAE (. American Society of Heating, 2024 ASHRAE Handbook)—HVAC Applications, Atlanta, GA: ASHRAE, Inc., 2024.
- [9] G. Bhagyesh S. Pawar, "Energy Optimization of Building Using Design Builder Software," *International Journal of New Technology and Research*, vol. 4, no. 1, 2018.
- [10] J. Montgomery, E. Peck, and G. Vining, *Introduction to Linear Regression Analysis*, 5th ed., Wiley, 2012.
- [11] C. Yan, Y. Ouyang, "Energy performance prediction of commercial buildings using support vector regression," *Energy Procedia*, vol. 142, pp. 2073–2078, 2017.
- [12] T. Hong, S. Taylor-Lange, "Advances in calibrated building energy modeling: A review," *Energy and Buildings*, vol. 97, pp. 234–245, 2015.
- [13] S. Wetter, C. H. Keeling, "Simulation-based energy performance assessment of HVAC systems using Modelica," *Journal of Building Performance Simulation*, vol. 7, no. 3, pp. 191–202, 2014.
- [14] Y. Ma, F. Borrelli, "Model predictive control of thermal energy storage in building cooling systems," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 796–803, 2012.
- [15] L. Perez-Lombard, J. Ortiz, "A review on buildings energy consumption information," *Energy and Buildings*, vol. 40, no. 3, pp. 394–398, 2008.
- [16] M. Wetter, "Modelica-based modelling and simulation to support research and development in building energy and control systems," *Journal of Building Performance Simulation*, vol. 2, no. 2, pp. 143–161, 2009.
- [17] D. P. Jenkins, Y. Fu, "Sensitivity analysis in building energy modeling: a review of studies using DesignBuilder," *Building Simulation*, vol. 11, pp. 885–903, 2018.
- [18] K. F. Fong, V. I. Hanby, "HVAC system optimization for energy management by evolutionary programming," *Energy and Buildings*, vol. 38, no. 3, pp. 220–231, 2006.
- [19] Y. Huang, R. J. Fuller, "Life cycle assessment of energy performance of residential buildings," *Energy and Buildings*, vol. 40, no. 3, pp. 415–423, 2008.
- [20] Z. Wang, S. Wang, "Energy consumption prediction for office buildings based on support vector machine," *Energy and Buildings*, vol. 37, no. 3, pp. 287–297, 2005.
- [21] T. A. Reddy, "Literature review on calibration of building energy simulation programs: uses, problems, procedures, uncertainty, and tools," *ASHRAE Transactions*, vol. 112, no. 1, pp. 226–240, 2006.
- [22] M. Aydinalp-Koksal, V. Ismet Ugursal, "Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector," *Applied Energy*, vol. 85, no. 4, pp. 271–296, 2008.
- [23] T. Catalina, V. Iordache, "Thermal comfort models and indicators for building energy performance analysis and design," *Energy and Buildings*, vol. 47, pp. 392–402, 2012.



- [24] E. Fabrizio, V. Monetti, "Methodologies and advancements in the calibration of building energy models," *Energies*, vol. 8, no. 4, pp. 2548–2574, 2015.
- [25] M. Santamouris, "Innovating to zero the building sector in Europe: Minimizing the energy consumption, eradication of the energy poverty and mitigating the local climate change," *Solar Energy*, vol. 128, pp. 61–94, 2016.
- [26] A. Beagon, M. Hajdukiewicz, "Influence of construction materials on the energy performance of buildings," *Building Simulation*, vol. 13, pp. 1173–1185, 2020.
- [27] L. D. Danny Harvey, "Reducing energy use in buildings by the use of natural ventilation," *Renewable and Sustainable Energy Reviews*, vol. 5, no. 2, pp. 97–121, 2001.
- [28] S. P. Corgnati, E. Fabrizio, "Free-running buildings: Climate, passive solutions and comfort," *Energy and Buildings*, vol. 40, no. 3, pp. 281–288, 2008.
- [29] B. Dong, K. P. Lam, "A real-time model predictive control for building heating and cooling systems based on the occupancy behavior pattern detection and local weather forecasting," *Building Simulation*, vol. 7, no. 1, pp. 89–106, 2014.
- [30] K. Liu, B. Pan, "Predicting hourly building energy consumption using machine learning models with a hybrid feature selection algorithm," *Energy*, vol. 202, p. 117754, 2020.
- [31] J. V. Caicedo, M. A. Pandharipande, "Energy-efficient lighting control for buildings based on occupant presence and ambient light level," *Energy and Buildings*, vol. 48, pp. 56–62, 2012.
- [32] P. Xu, M. J. Hu, "Forecasting the energy consumption of HVAC systems using hybrid statistical and machine learning techniques," *Energy Reports*, vol. 6, pp. 389–398, 2020.
- [33] A. Afram, F. Janabi-Sharifi, "Theory and applications of HVAC control systems – A review of model predictive control (MPC)," *Building and Environment*, vol. 72, pp. 343–355, 2014.
- [34] K. Fong, H. Lee, "Improving the energy performance of HVAC systems in commercial buildings using predictive control strategies," *Applied Energy*, vol. 88, no. 1, pp. 475–484, 2011.
- [35] M. Kavgic, A. Mavrogianni, "A review of bottom-up building stock models for energy consumption in the residential sector," *Building and Environment*, vol. 45, no. 7, pp. 1683–1697, 2010.
- [36] H. Yoshino, A. Bamba, "A review of energy performance in residential buildings using data from field studies," *Energy and Buildings*, vol. 40, no. 3, pp. 345–352, 2008.
- [37] M. Attia, W. W. Braham, "Post-occupancy evaluation of indoor environmental quality in net-zero energy buildings," *Energy and Buildings*, vol. 84, pp. 276–288, 2014.
- [38] P. Heiselberg, "Integrated design of low energy buildings," *Renewable and Sustainable Energy Reviews*, vol. 4, no. 2, pp. 79–98, 2002.
- [39] C. Hopfe, J. Hensen, "Uncertainty analysis in building performance simulation for design support," *Energy and Buildings*, vol. 43, no. 10, pp. 2798–2805, 2011.
- [40] J. Crawley, L. Lawrie, "EnergyPlus: Creating a new-generation building energy simulation program," *Energy and Buildings*, vol. 33, no. 4, pp. 319–331, 2001.

