Building simulation based optimization through design of experiments

Jay Dhariwal – Dept of Energy Sc and Engg, IIT Bombay, Mumbai, India – jaydhariwal@gmail.com
Rangan Banerjee – Dept of Energy Sc and Engg, IIT Bombay, Mumbai, India – rangan@iitb.ac.in

Abstract

Building thermal simulation based parametric methods are computationally intensive for optimizing the building design. This work uses experimental design techniques, i.e. fractional factorial design and response surface methodology, for sensitivity analysis and surrogate modeling respectively. These techniques find the solution in a reasonable time. Their application for building design optimization has not been found in the literature before.

Fractional factorial design has been used to identify the significant design variables. These variables are used to form a correlation for annual cooling load prediction, using response surface methodology. These methods are illustrated using two cases to minimize the life cycle cost of a single-storeyed, air-conditioned, solar powered, detached home, with 64 sq. m. floor area, for the warm and humid Mumbai climate.

For this climate, window solar heat gain coefficient, window to wall ratio, overhang depth and roof reflective coatings turn out to be the most important among the design variables used for this case study. The created response surface models show an error of less than 5% for more than 99% of the test data, which is comparable to other such models. Strategies are suggested to bring the error for the entire search space to less than 10%.

Life cycle cost minimization using the model for case 2 does 12 million iterations as opposed to 250 iterations using a parametric EnergyPlus simulation run at the same time. The solution is better and the design achieved is also different. The optimum design has a cooling load of 55 kWh m^{-2} yr^{-1}, while it varies from 46 to 118 kWh m^{-2} yr^{-1}.

This work adds an intuitive method for building design and opens up possibilities for optimization.

1. Introduction

Detailed thermal simulation programs like EnergyPlus, TRNSYS, etc. are needed to accurately estimate the cooling and heating loads in buildings. These programs can take from a few seconds to hours to find the building performance for each run. To find the optimal building designs for desired objective functions like minimum lifecycle cost (LCC), minimum discomfort hours, etc., simulation based parametric methods may not be viable as they are computationally intensive. The time taken may be in the order of years, for iterating over 15 design variables having 3 levels each, assuming 7 seconds per simulation run. A typical building design could easily have more than 30 design variables. Techniques like sensitivity analysis and optimization methods, coupled with simulation programs, are used to reduce the computational run time, while maintaining the accuracy of results.

1.1 Literature review

Sensitivity analysis (SA) can be used to find the design variables, having the most impact on the response variable, like the cooling load. Local SA involves changing one factor at a time over the base case. It is easy to use but does not consider the interactions between the variables. Global SA spans the entire input space and is preferred. Variance based methods give the most reliable results for global SA, as per the review paper by (Tian, 2013).

After eliminating the unimportant variables using SA, optimization methods can be used to find the optimal building design. Nguyen et al. (2014) and Machairas et al. (2014) provide a review of the building design optimization methods used. Genetic algorithm based direct search methods are

most popular (Machairas et al., 2014). Nguyen et al. (2014) suggest that improvements in surrogate or meta-models hold a lot of promise, however, there have only been limited studies using these models for building design.
Surrogate models (SMs) overcome the high run time limitation of simulation based optimization methods by creating an approximate model or a correlation, using a limited simulation run for predicting simulation outputs. Machine learning models like support vector machines and artificial neural networks have been mostly used to develop surrogate or meta-models (Machairas et al., 2014). Once the SMs are created and validated, they can be used parametrically or coupled with other optimization algorithms to optimize the design.
Nguyen et al. (2014) suggest that the performance of the optimization methods may be different for cold and warm climates.

1.2 Problem definition

This paper proposes fractional factorial design (Montgomery, 2007) for global SA, not used for building design before. This statistical method gives very reliable results, like other variance based global SA methods, but overcomes their limitation of high computational time by using only a fraction of parametric simulation runs to identify the most significant variables, while also analysing the key interactions between them. This technique can easily analyse 50 or more design variables in a short time. Gong et al. (2012) have used a factorial design technique called orthogonal method, for local SA only, to improve upon the base case.
Response surface methodology (Montgomery, 2007) has been used to create a surrogate model for cooling load prediction. Although response surface models (RSM) are the most popular surrogate models in general (Wikipedia, 2014), they have not been used for building design before. RSM is very intuitive to use and builds on fractional factorial design used for SA, to form a model for the significant variables, for further optimization. Using statistical methods like RSM can help avoid learning artificial intelligence techniques like artificial neural network or support vector machines, which require expert knowledge and have been the main reason for the limited use of SMs for building design so far (Machairas et al., 2014). Some commonly used RSM designs are Box-Behnken design and central composite designs (Montgomery, 2007).
The surrogate models developed in this paper are also tested to assess their accuracy and strategies are proposed for the regions with high error.
In addition, there is a dearth of studies on the application of building design optimization for warmer Indian climates. Hence, these methods are presented in an optimization framework to minimize LCC for an air-conditioned, single zone, solar powered, detached home, with a floor area of 64 m^2, based in the warm and humid Mumbai climate. It is inspired by the Indian team's house "h-naught" from Solar Decathlon Europe 2014 (SDE, 2014).

2. Methodology

Figure 1 explains the flowchart of the methodology used to optimize the building design. Other methodology related key points are addressed in this section. The simulation model for the base case is built in EnergyPlus. The design variables are estimated for their ranges, thermal properties and cost data. Local and global SA, surrogate modelling and optimization algorithm are applied as per the framework.
For global SA, the runs for the fractional factorial design are limited to 512 to achieve a target of finishing the simulations within an hour, with each simulation in EnergyPlus taking 7 seconds, on average, for the present case study. To create the fractional factorial design, Box-Behnken RSM design and ANOVA, Design-Expert software (Stat Ease, 2014) has been used. For running simulations in a batch mode, VBA programming language is used. If the simulated test cases for SM show an error of greater than 10%, then the region of high error should be identified and another SM should be fit for this region to reduce the error. For optimization, the best 5 solutions should be used to choose the most robust design among them.

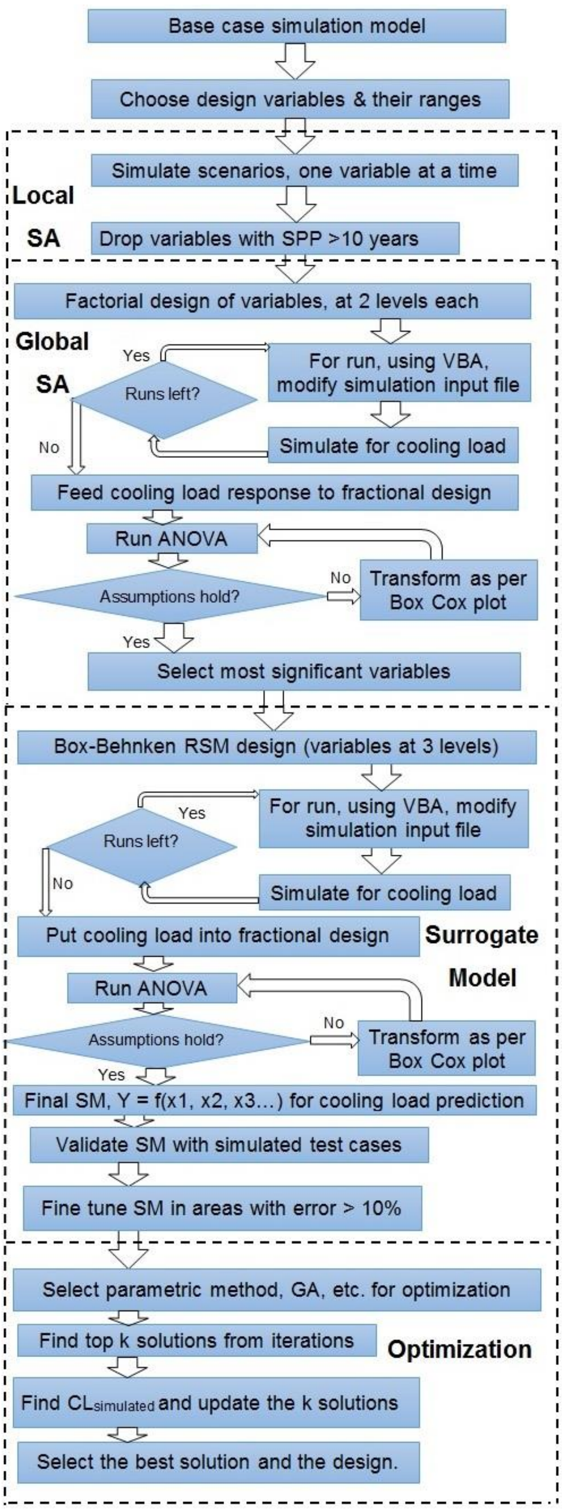

Fig. 1 – Flowchart for optimization methodology

3. Case study input data

The methodology is illustrated considering two cases for the Figure 2 building in Mumbai. The structural material for the base case is taken as concrete (Ramesh et al., 2012) for it to be representative of Indian buildings.

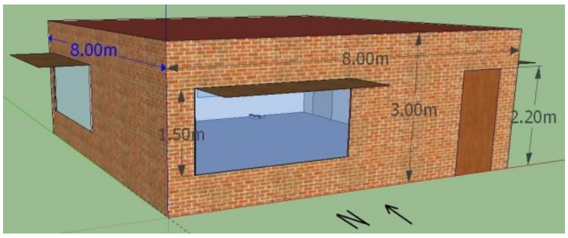

Fig. 2 – Detached home in Mumbai

Table 1 shows the important input parameters for the simulation. AC schedule was arrived at, from the survey of the load profile of the homes of middle class households of Gujarat, India (Garg et al., 2010). Other internal gains include the use of cooking gas, fridge and lighting.

Table 1 – Important input parameter assumptions

Input parameter	Data assumptions
Occupants	6 (for a middle class family)
Occupancy schedule	4 people on weekdays from 8 am to 6 pm; 6 people during rest of the time.
AC set points	T_{opt}: 22°C to 27°C (Manu et al., 2014)
AC schedule	April to June: 9 pm to 5 am, 2 pm to 4 pm for, July to October: 2 pm to 4 pm (Garg et al., 2010)
Infiltration rate	0.4 air changes per hour
Internal mass	40 m^2 (partitions), 20 m^2 (furniture)

Table 2 shows the 16 design variables along with their ranges to be used for the simulation runs. Insulation, thermal mass and PCM thickness have 3 variables each for wall, roof and floor. The WWR lower limit was fixed as per the requirements for LEED IEQ credit 8.1 (IGBC, 2011) using day-lighting simulation in EnergyPlus. Day-lighting controls in the simulation maintain the lighting levels in the house to greater than 100 lux. Curtains are used only when the solar radiation is greater than 500 W/m^2 to avoid the glare from the direct solar radiation. Glasswool is used for thermal insulation, having a density of 48 kg/m^3 (Twiga,

2014) and water is used for thermal mass. The base case run yields a cooling load of 100 kWh m^{-2} yr^{-1}.

Table 2 – Design variables and their range

Design Variable	Range
WWR	15% to 40% (IGBC, 2011)
Window U value	1 to 6
Window SHGC	0.2 to 0.8
Overhang depth	0 to 1 m (BIS, 2005)
Fin depth	0 to 0.5 m
Curtains SHGC	0.2 to 1.0
Cool roof ρ	0.25 to 0.85 (SSEF, 2014)
Insulation t	0-0.1 m (Wall), 0-0.2 m (Roof, Floor)
Thermal mass t	0-0.1 m (Wall), 0-0.2 m (Roof, Floor)
PCM t	0-0.1 m (Wall), 0-0.2 m (Roof, Floor)

4. Application of optimization method

Each step in the optimization method is applied as per Figure 1 for two cases, case 1 and case 2. The difference between the two cases is in the number of design variables. Case 1 has the 16 design variables, as per Table 2. For case 2, the design variables, namely, WWR, window U-value, overhang-depth, fin-depth and window SHGC have been made different for each façade. For example, WWR would be treated as four design variables, namely, WWR for north, east, west and south facades, respectively. This would result in a total of 31 design variables. Case 1 is easier to understand, whereas case 2 shows how SM could be used for designs with many design variables.

4.1 Local sensitivity analysis

As per the local sensitivity analysis and payback analysis of the costly items, shown in Table 3, it is seen that PCM would not be cost effective as its SPP is far greater than 10 years, as per scenarios run over the base case to find its cooling load reduction benefit, with utility rates of € 0.13 per kWh. So, it is eliminated at this step. This leaves 13 design variables for case 1 and 28 design variables for case 2.

Table 3 – Cost analysis for variables (1 € = 75 Indian rupees)

Variable	Cost (in €)	SPP
PCM	6000	100 years
Low solar gain windows	660	10 years
Cool roof	800	8 years
Glasswool insulation	260	4 years

4.2 Global sensitivity analysis

For case 1, with the range of 13 design variables as per Table 2, the fractional factorial design chosen is resolution VI 2^{13-4} design with a single replicate. This design can estimate the main effects and two factor interactions. It was created in Design-Expert software (Stat Ease, 2014). 512 simulation runs based on this design were automated using VBA, calling EnergyPlus batch file for each simulation and took a time of 55 minutes on an Intel® Core™ i7-2670QM processor. The cooling load varied from 43 to 143 kWh m^{-2} yr^{-1}, with a mean of 81 kWh m^{-2} yr^{-1}.

For case 2, for the 28 design variables, the factorial design has a resolution V. 408 simulation runs based on this design took a time of 45 minutes. The cooling load varied from 45 to 138 kWh m^{-2} yr^{-1}, with a mean of 83 kWh m^{-2} yr^{-1}.

The normal probability plot of residuals is well behaved for both the cases, suggesting that the normality assumptions hold.

Figure 3 shows that case 1 and case 2 have the same significant variables selected, albeit with different contribution levels. Window SHGC, WWR, cool roof ρ, overhangs and the interactions involving them explain more than 85% of the total variance in the model for both the cases. The most significant positive interaction is between the cool roof ρ and the roof insulation, which suggests that when the cool roof ρ is high, then the roof insulation has little effect. Figure 4 shows the breakup of facade-wise contribution for case 2. The west and east facade contribution for SHGC is higher than the north and south facade as the direct radiation is able to enter at a lower altitude angle for these facades.

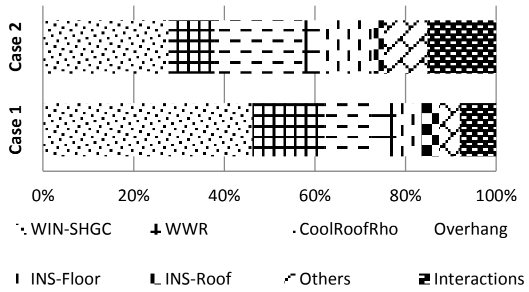

Fig. 3 – Percent contribution of design variables to cooling load

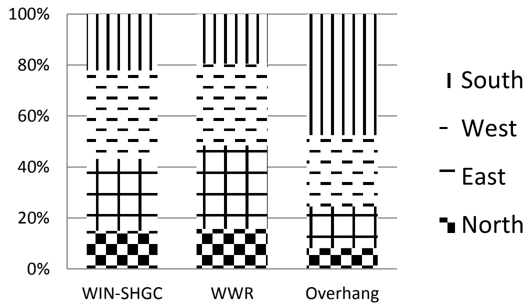

Fig. 4 – Breakup of variables for the façade wise contribution

The thermal mass, wall insulation and window U-value variables do not have a major effect as the temperature in Mumbai is fairly constant throughout the year. Curtains can be useful to avoid solar radiation but they also lead to blocking of daylight and views so they are only used to avoid harsh direct radiation, which happens only for a little time, when the AC is used. The variables, which are left out, are set at base case values in the simulation model. In most cases, it would mean removing these variables from further analysis. For example, thermal mass is removed but window is set at U-value of 6 W/m^2-K.

To create the surrogate model, the floor and the roof insulation are left out. The presence of floor insulation has a negative effect on the cooling load and it does not have any interaction with any other variables to be concerned about. For the roof, cool roof material is better suited than the roof insulation as per Figure 3. Thus, for case 1, four design variables, namely, cool roof ρ, window SHGC, overhang depth and WWR, and for case 2, the same variables for all four facades would be used to form the surrogate model, making it a total of 13 design variables.

4.3 Surrogate Model

4.3.1 SM for Case 1

Using RSM, a second order polynomial is fit to the 4 design variables to obtain the surrogate model, using the range from Table 2. Box-Behnken design is created in Design-Expert software using 3 levels for each factor, taking the lower and upper limit and the middle value of the range. For example, for the overhang depth, levels correspond to 0 m, 0.5 m and 1 m. The design leads to 29 simulation runs, which takes less than 3 minutes of computational time.

The fitted model has an adjusted R-squared of 0.99. The variables A, B, C and D have been defined in the nomenclature section.

$$AnnualCoolingLoad = 63.6 - 3.5A + 64.8B - 43.8C + 0.8D - 17.0\ AB - 11.7\ AC - 0.3\ AD + 11.8\ BC + 0.8\ BD + 0.5\ CD + 11.5\ A^2 - 44.5\ B^2 - 11.4\ C^2 \quad (1)$$

4.3.2 SM for Case 2

Like in section 4.3.1, a second order polynomial is fit for cooling load as the response, using RSM. Box-Behnken design leads to 220 simulation runs. The Box-Cox plot in Design-Expert software recommends a power transform to meet the normality assumptions for the residuals, as shown for the cooling load on the LHS of equation (2). The fitted model on the RHS of equation (2) has linear, quadratic and interaction terms between the variables for all façades. The fitted model achieves an adjusted R-squared of 0.99.

$$AnnualCoolingLoad^{2.1} = f(\text{overhang depth, window SHGC, WWR, cool roof } \rho) \quad (2)$$

4.3.3 Validation of SM

Table 4 shows the testing datasets created for the two cases. For case 1, the variables are varied parametrically. For example, WWR is varied from 15% to 40% in steps of 5%. For case 2, since the variables are large, so 4 fractional designs, with 512 runs each, have been created, to test for the entire range of variable values. WWR, window SHGC, overhang for all four facades and cool roof ρ would be varied at 0 & 100%, 10 & 90%, 25 & 75% and 40

& 60% of the range for the 4 fractional designs respectively. For example, WWR (%) for all facades would be 15 & 40, 17.5 & 37.5, 21.25 & 33.75, 25 & 30 respectively for the four fractional designs with 512 runs each. Table 4 shows a case 2 design.

Total test cases for case 1 and case 2 are 2058 and 512x4 = 2048 respectively. The error statistic MAPE for the difference between the simulated vs. SM predicted cooling load is less than 1% for both cases. The data point with minimum error for case 1 and case 2 are -12% and -36% respectively. The maximum error is less than 5% for each case. The areas with high error need to be analysed further.

Table 4 – Range of variables for testing

Variable	Case 1	Case 2 (design 1)
WWR (in %)	15-40 step 5	15, 40
Window SHGC	0.2-0.8 step 0.1	0.2, 0.8
Overhang (in m)	0-1 step 0.17	0, 1
Cool roof ρ	0.25-.85 step 0.1	0.25, 0.85

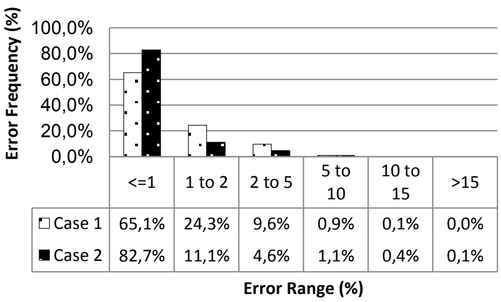

Fig. 5 – Error frequency for the surrogate model prediction

Figure 5 shows that about 98% of the test data have errors of less than 5%. These results are promising and validate the effectiveness of the surrogate model. However, for case 2, 0.5% of the test cases have errors higher than 10% and ways need to be found to reduce this error as well to make the SM valuable for the entire search space.

The analysis of high error points shown using Figure 6 suggests that the error has a tendency to increase on the negative side, with the reduction in the simulated cooling load, resulting in the underprediction of cooling load. This problem is fixed by using a reduced variable SM, having the same form as case 1, for the region of case 2 with CL less than 60 kWh m^{-2} yr^{-1} (as shown in Figure 6 with a vertical dotted line). This region accounts for 3.5 % of the total search space. The only difference between the reduced model and the case 1 model is the range for cool roof reflectivity, which is 0.65 to 0.85 for the reduced model. This strategy brings the error for all test data for case 2 to less than 10% but the detailed design ability of case 2 would be lost for this region.

To avoid compromising on the detail that case 2 offers, while keeping the error low, the optimization algorithm in the next section is made to search the space with CL less than 60 kWh m^{-2} yr^{-1}, for both the reduced case 1 model with less error and case 2 model with higher detail. The best 5 solutions are found from both the models and the optimum among them is chosen.

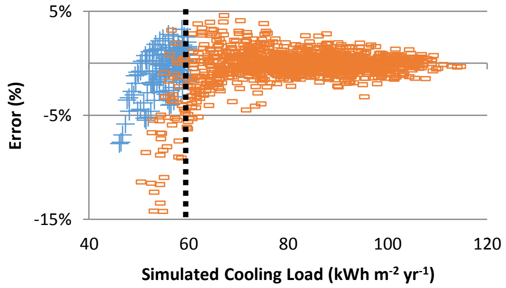

Fig. 6 – Error reduction for Case 2 ('-') with reduced model ('+')

4.4 Optimization for LCC

Using the surrogate model, LCC is optimized parametrically. As per the cost functions given in Table 5 and discount factor of 10%, the results for the minimum LCC design in constant euros are given in Table 6. The base case cost includes the construction cost of the house and the cost of the solar thermal hot water system. The life of the building components is given in Table 5.

As per Table 6, parametric EnergyPlus run method, case P is able to search 256 combinations of input parameters only as compared to 12 million iterations for case 2 in the same time. If the iterations for case 2 were done parametrically, it would have taken 2.7 years. In this case, SM is able to search 50000 times more solution space than parametric methods. This is where the real power of SM based methods lies.

Table 5 – Cost functions for design variables

Material	Cost function (in €)
Base cost of house	13.3 per sq. ft. (50 years life)
Solar PV	1066.7 per kWp (25 years life)
AC	400 per TR of capacity (10 years life)
Windows	$10000/75x(0.6-1.35B+0.85B^2)xWWR$
Overhangs	12 per m^2
Cool roof ρ	(12ρ/0.85) per m^2

Table 6 – Comparing results for optimized building design

	Case P	Case 1	Case 2
Method used	Parametric	SM	SM
Total iterations	256	0.3 mn	12 mn
Time taken (hrs)	0.5	0.05	0.5
$LCC_{optimum}$ (in €)	13973	13824	13757
Overhang depth	0.7 m	0.9 m	0, 1, 1, 1 m*
Window SHGC	0.6	0.375	0.2, 0.6, 0.6, 0.6*
WWR	15%	15%	15, 15, 15, 15%*
Cool roof ρ	0.65	0.85	0.85
CL for this design	69.3	51.8	55.1

*for N, E, W, S facades respectively

For case 2, since the cooling load for the minimum LCC lies in high error region as per Figure 6, so both the case 2 and reduced models were run in this region but the case 2 model gave a better solution in this instance. The minimum LCC for case 2 is 1.6% less than case P. The design is also different for all variables except WWR.

4.5 Applicability

For this paper, all variables chosen were continuous variables but this framework can be used for discrete variables in the same manner. In this case, for SM, for each combination of discrete variable values, there will be a separate equation.
SM was used for cooling load prediction only as heating is not required for the Mumbai climate but the methodology could be easily extended for cold climates as well.
SMs can be used as a correlation for mass housing projects for a given climate by architects or other building industry professionals, who do not have a background in building simulation.

5. Conclusion and next steps

A building simulation assisted optimization methodology has been presented for a single storeyed house in Mumbai. Fractional factorial design was used to find the significant variables for cooling load prediction. A correlation for cooling load prediction was fit for these variables and was used to optimize the life cycle cost. The entire process was completed in a couple of hours.
The surrogate model used to find the correlation was found to be thousands of times faster than the parametric method. The test data also showed that the prediction error could be kept to be less than 10% for the entire search space. This correlation, when used for optimization, also found a better solution with a different building design.
The case study revealed that for the Mumbai climate, window solar heat gain, window to wall ratio, roof reflective coatings and overhang depth are the most important design variables.
This work adds an intuitive optimization method into the building design kit. This approach opens up a lot of possibilities for building design optimization and could be used for the prediction of any simulation output. This method can also be extended to consider the effect of uncertainty in parameters like infiltration rate, occupancy schedule, etc. on building design.

6. Nomenclature

Symbols

ρ	reflectance
A, B, C, D	overhang depth to window height, window SHGC, cool roof ρ, WWR
AC	air conditioning
CL	cooling load (in kWh m^{-2} yr^{-1})
hrs	hours
INS, WIN	insulation, windows
k, U	thermal conductivity, overall heat transfer coefficient (in W m^{-2} K^{-1})
kWp	kilo-watt peak
LCC	life cycle cost
LHS, RHS	left hand side, right hand side
m	metre
mn	million
MAPE	mean absolute percentage error
min, max	minimum, maximum
N, E, W, S	north, east, west, south
PCM	phase change material
PV	photo-voltaic
RSM	response surface methodology
SA	sensitivity analysis
SHGC	solar heat gain coefficient
SM	surrogate model
SPP	simple payback period
sq. ft.	square feet
t	thickness
T_{opt}	operative temperature (Celsius)
VBA	Visual Basic for Applications
WWR	window to wall ratio

Subscripts/Superscripts

opt	operative temperature

References

BIS. 2005. *National Building Code of India 2005*. New Delhi: Bureau of Indian Standards.

Garg A., Maheshwari J., Upadhyay J. 2010. *Load Research for Commercial and Residential Establishments in Gujarat*. ECO-III-1024

Gong X., Akashi Y., Sumiyoshi D. 2012. "Optimization of passive design measures for residential buildings in different Chinese areas" *Building and Environment* 58: 46-57

IGBC. 2011. *Green Building Rating System.* LEED 2011 for India.

Machairas V., Tsangrassoulis A., Axarli K. 2014. "Algorithms for optimization of building design: A review" *Renewable and Sustainable Energy Reviews* 31: 101–112

Manu S., Shukla Y., Rawal R. 2014. *India Model for Adaptive (Thermal) Comfort.* Accessed Nov 15, 2014. www.cept.ac.in/carbse

Montgomery D.C. 2007. *Design and Analysis of Experiments*. New Delhi: Wiley India Edition.

Nguyen A., Reiter S., Rigo P. 2014. "A review on simulation-based optimization methods applied to building performance analysis" *Applied Energy* 113: 1043–1058

Ramesh T., Prakash R., Shukla K.K. 2012. "Life cycle energy analysis of a residential building with different envelopes and climates in Indian context" *Applied Energy* 89: 193–202

SDE. 2014. "Home page" Accessed Nov 15, 2014. http://www.solardecathlon2014.fr/en/

Stat-Ease, Inc., MN, USA. 2014. *Design-Expert® software, version 9*. www.statease.com

Tian Wei. 2013. "A review of sensitivity analysis methods in building energy analysis" *Renewable and Sustainable Energy Reviews* 20: 411-19.

Twiga. 2014. "Twiga Insul" Accessed Nov 15, 2014. http://www.twigafiber.com/twigainsul.php

Wikipedia. 2014. "Surrogate Model" Accessed Nov 15,2014.en.wikipedia.org/wiki/Surrogate_model

SSEF. 2014. "Cool Roofs for Cool Delhi" Accessed Nov 17, 2014. http://shaktifoundation.in/initiative/cool-roofs-for-cool-delhi/