Using An Adaptive Meta-Model Evolutionary Algorithm For Mixed-Integer Type Building Design Optimization
for Building Simulation 2017 Conference

Weili Xu[1], Khee Poh Lam[1], Omer T. Karaguzel[1]
[1]School of Architecture, Carnegie Mellon University, Pittsburgh, U.S.

Abstract

The adaptive meta-model evolutionary algorithm has demonstrated its excellent convergence and faster speed than the conventional algorithms in building design evaluations (Xu et al., 2016). However, this algorithm is built on integer-type problems while most of the building designs have mixed integer-type parameters. In this study, numerical calculations are added in the operators of the algorithm. These updates allow the algorithm to work on mixed integer-type problems.

The updated algorithm is applied to a large office building with 350 thermal zones. The design of the daylight shelf and building-integrated photovoltaic (BiPV) systems are included in the design parameters. The results indicate that despite its high capital costs, BiPV system is favored by most of the optimal designs. On the contrary, the light shelf system is not favored in this integrated building design optimization for various reasons such as the low daylight harvesting potential from the high efficiency lighting technology and the objectives of the study.

Introduction

Design optimization has been one of the most popular approaches to achieve high performance buildings. In 2016, nearly 500 optimization papers have been published in the *Journal of Energy and Building* (ScienceDirect, 2016). These publications addressed the issue of optimizing various building performance related attributes, including capital cost (Wang et al., 2016; Xu et al., 2016), life cycle cost (Tokarik and Richman, 2016; Carreras et al., 2016) and energy (Du et al., 2016), with a large set of building design parameters. In industry, case studies (EcoGlobe, 2016) and design competitions (ASHRAE, 2016) also demonstrated the advantages of integrating optimization algorithms in the building design process.

Currently, most of the building design optimization problems are conducted by adding a design proposal evaluation process into the cost functions of optimization algorithms. The evaluation process can be performed with either a physical-based building simulation engine or a meta-model. Between these two methods, the physical-based building simulation engine appears more frequently in the optimization studies (Machairas et al., 2014). However, this method takes comparatively longer period to find the optimal design proposals. It is because the evaluation of the cost function involves of solving the differential algebra equations (DAE) with approximation techniques, which is not only computationally intensive but also introduces the discontinuity relation between cost functions and design parameters (Wetter, 2004). The discontinuity implies that the meta-heuristic algorithms, the algorithm that applies master strategies on searching through iterative heuristics, have a higher probability of finding better building design proposals than the direct search algorithms. Despite the advantages of meta-heuristic algorithms, they use stochastic processes to conduct heuristics that largely increases the required number of cost function evaluations, thus resulting longer calculation period.

To address this issue, researchers have proposed the use of meta-models, such as regression models, to replace the building simulation engines (Eisenhower et al., 2012; Costas et al., 2014). These meta-models are usually linear and smooth functions so that the time for evaluating cost functions is negligible. Also, the use of direct search algorithms can significantly reduce the number of evaluations. Therefore, the optimization process can be completed within a few seconds. However, a substantial amount of time is dedicated to forming a pre-simulated database for meta-model training. This database typically requires thousands of design options and simulation outputs that are generated from a physically-based building simulation engine.

In the past decade, an alternative algorithm, successive meta-modeling optimization, has evolved among the evolutionary algorithms (Deb and Nain, 2007). The core concept can be described as an integration of optimization algorithm with meta-models where the meta-models update themselves in every few generations along with design proposals moving towards the optimum region. Furthermore, this approach uses the data from a stochastic process in evolutionary algorithms, which eliminates the process of creating

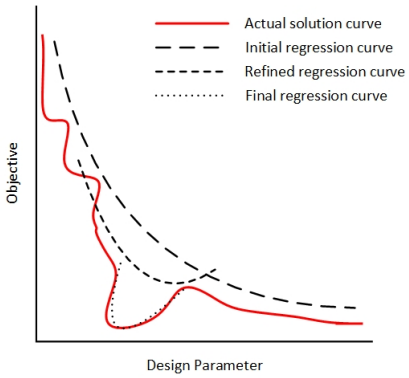

Figure 1: The coarse-to-fine adaptive meta-modeling theory.

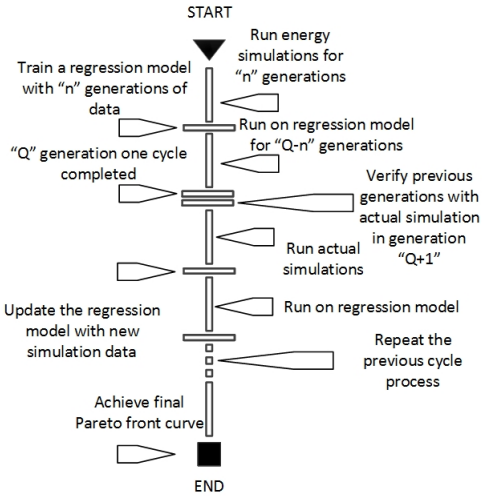

Figure 2: Proposed adaptive meta model optimization process

a pre-simulated database for meta-model training. This algorithm was studied and its performance attributes, including convergence, diversity, and speed, were compared with the non-dominated sorting genetic algorithm II (NSGAII) in an office building design problem (Xu et al., 2016). The results indicated that, although the diversity of design proposals was reduced slightly, implementing this algorithm could achieve 25% better convergence in 60% less time. However, this study focused on integer-type design parameters whereas most of the building design optimization problems had mixed integer and numeric type design parameters. This paper addresses the issue by reprogramming the data structure of the algorithm and implementing numerical calculations in the key operations. Therefore, the new approach can handle mixed-integer type optimization problems. A case study is presented with the updated algorithm and the results of this case study are discussed.

Method

The adaptive meta-model optimization algorithm proposes a novel approach that marries the machine learning technique with conventional search algorithms for convergence and speed improvements. The motivation behind the development of such a framework for building design optimization can be summarized into two key observations.

- The computationally intensive evaluation process of cost function dramatically limits the dimension of search spaces.
- The heuristic process in the conventional evolutionary algorithm generates some evaluations whose results are typically discarded after the algorithm reproduces their offspring.

The first observation has encouraged researchers to explore the possibility of replacing the building energy model with meta-models. These meta-models are typically regression models that are fine-tuned by learning patterns and relations between data and target values. In the building context, such data can be design parameters and the target values can be energy and life cycle cost. However, this method requires a pre-simulated database, which could be time-consuming to create. This issue urges researchers to find a new method that can eliminate the necessity of setting up the pre-simulated database in the optimization process. Consequently, the need leads to the second observation in the process of conventional optimization algorithms.

In the second observation, it implies that the evaluations in the heuristic process can be collected to form a database for training the meta-models. The accuracy of meta-models could be limited at the beginning stage of an optimization process due to the lack of sufficient data points for training. However, they can be refined in later stages when the search is conducted in a smaller but more focused region. Figure 1 demonstrates the concept of this "coarse to fine" meta-model theory. The initial regression curve in Figure 1 does not provide enough accuracy to predict the actual solution value, however, it provides the algorithm with a search direction to move to. As the optimization continues, the regression curve is refined to higher accuracy in predicting target values. Eventually, a final regression curve can successfully predict the values in the optimal region.

Adaptive meta model framework

Figure 2 shows the process of the adaptive meta-model optimization. In the first n generations, the cost functions are evaluated with EnergyPlus. At the same time, a database is created to record the choices for each design proposal and the outcomes. When the optimization reaches the n generation, the framework holds the process and switches to meta-model training and selection processes. Once a trained meta-model is produced, the optimization will then resume with all the evaluations of cost function done on the trained meta-model. After the Q generation, the op-

timization changes its cost functions back so that the $Q+1$ generation can use energy simulation to re-evaluate the best design proposals from the previous Q generations. Repeating this process for several cycles, the optimization can quickly find a set of optimal design proposals, which is called the Pareto Front.

Optimization algorithm

The core algorithm used in this framework is NSGAII, which is one of the most popular multi-objective genetic algorithms. The NSGAII is equipped with a non-dominated sorting algorithm, which allows it to perform multi-objective optimization. Tests have shown its superior performance on keeping the diversity of the optimal solutions while converging near the global pareto optimal front (Deb et al., 2002). NSGAII only handles integer or numeric types of design parameters. In order to equip the ammNSGAII algorithm with the ability to process the mixed data type, a genetic adaptive search scheme (GeneAS) is adopted (Deb, 1997). In GeneAS, each parameter is encoded based on its nature in the optimization process, and the crossover and mutation operators apply their operations parameter by parameter.

Crossover operator

Crossover is one of the operators used in producing offspring (i.e, candidate design solution alternatives). The essential idea of this operator is to enhance the exploration process by recombining the "high-performance" group of design parameters. In (Xu et al., 2016), the crossover was performed by a single point crossover algorithm, which only works with binary or integer data type. In the GeneAS framework, a simulated binary crossover (SBX) algorithm has been integrated with the single point crossover for processing numeric design parameters. This crossover algorithm is designed with respect to the properties of the single point crossover by ensuring the average design parameter values remain in the process, and the spread factor (β), which controls the spread of likely offspring, are similar to the single point crossover operation (Deb and Kumar, 1995). Equation 1 and 2 show the detail calculation method of SBX.

$$\beta = \begin{cases} (2\mu)^{\frac{1}{\eta_c+1}} & \text{for } \mu \leq 0.5 \\ (\frac{1}{2(1-\mu)})^{\frac{1}{\eta_c+1}} & \text{for } \mu > 0.5 \end{cases} \tag{1}$$

$$\begin{aligned} p_1' &= 0.5[(1+\beta)p_1 + (1-\beta)p_2] \\ p_2' &= 0.5[(1-\beta)p_1 + (1+\beta)p_2] \end{aligned} \tag{2}$$

In Equation 1, β represents the spread factor and μ is a number randomly drawn from the set of $[0, 1)$. In addition, the η_c is the distribution index which determines how close the offspring is to their parents. p' and p are the values for offspring and parents respectively. The mathematic formation of Equation 1 implies a high possibility of achieving a β that is close or equal to 1. The β calculated in Equation 1 is used in Equation 2 for calculating the values of design parameters in offspring. Equation 3 combines the two equations in Equation 2. It implies that the average design parameters values are the same for both parents and offsprings.

$$\frac{p_1' + p_2'}{2} = \frac{p_1 + p_2}{2} \tag{3}$$

Mutation operator

The mutation operator is primarily used to maintain the diversity of design solutions. A bit-flip algorithm is ordinarily used for parameters with integer data type. The algorithm simply replaces the original value of a design parameter by selecting a random integer value within the allowable range of the design parameter. For numeric design parameters, a polynomial mutation operator with a user-defined index parameter (η_m) is often used to perturb a design proposal in a parent's vicinity. The calculation consists of three steps.

1. Generate a random number μ between 0 and 1.
2. Calculate a parameter δ as follows:

$$\delta_q = \begin{cases} [2\mu + (1-2\mu)(1-\delta)^{\eta_m+1}]^{\frac{1}{\eta_m+1}} - 1 \\ \quad \text{for } \mu \leq 0.5, \\ 1 - [2(1-\mu) + 2(\mu-0.5)(1-\delta)^{\eta_m+1}]^{\frac{1}{\eta_m+1}} \\ \quad \text{for } \mu > 0.5. \end{cases} \tag{4}$$

   The $\delta = min[(p - p_l), (p_u - p)]/(p_u - p_l)$, where p_u and p_l are the upper and lower bounds of parameter p respectively. The η_m is the distribution index for mutation which controls the distances of the mutated value with its parent. A larger η_m shows a higher probability that the mutated value is closer to its parent. As demonstrated in (Deb and Agrawal, 1999), a value between $[20, 100]$ for η_m is sufficient for most of the problems.
3. Calculate the mutated offspring as follow:

$$p' = p + \delta_q(p_u - p_l) \tag{5}$$

In GeneAS, the polynomial mutation operator (Equation 4 and 5) has been integrated into the operator with a mutation clock scheme. The mutation clock scheme requires that the probability of a bit followed by a mutated bit should be determined by using an exponential probability distribution. (Deb and Deb, 2014).

Regression algorithms

Two regression algorithms, the Akaike Information Criterion (AIC) based linear regression (Hall et al., 2009) and Support Vector Machine (SVM) (Chang and Lin, 2011), are included in the optimization process. The linear regression and SVM are popular regression models that are often used in predicting building energy consumption. The hyper-parameters

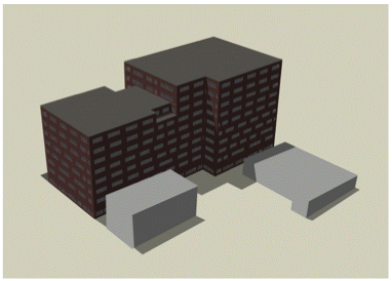

Figure 3: Rendered image of test building in DesignBuilder v4.7

of these two regression models are defined according to (Xu et al., 2016). Both algorithms can handle integer and numeric value concurrently. A 10-fold cross validation process is implemented for model selection to prevent under-fitting or over-fitting issues. The model selection criteria is based on the normalized root mean square error (NRMSE) (Equation 6).

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y)^2}{n}}$$
$$NRMSE = \frac{RMSE}{y_{max} - y_{min}} \tag{6}$$

Experiment

The building case

The test building is a multi-story office building located in Norristown, PA, U.S. Norristown has a heating dominant climate with a nearly 3000 heating degree day ($HDD18$). It consists of two towers, one has eight stories on the north side, and the other has ten stories on the south side. The total floor area of this building is 28,000 m^2 with 18,000 m^2 taking up the conditioned building area. The building energy model is built in DesignBuilder v4.7 (Figure 3).

Optimization formulation

The first cost and operation costs were selected to be the two objectives. The design problem can be formulated as:

$$min\left\{f_1(x), f_2(x)\right\} \tag{7}$$

Where $f_1(x)$ is building operation costs, which is stated:

$$f_1(x) = x_{rep} - x_{res} + x_e + x_{om\&r} \tag{8}$$

$f_2(x)$ is the capital costs (initial investment costs) and it can be formulated as:

$$f_2(x) = x_{ei} + x_{hi} + x_{li} + x_s + x_{pv} + x_{rest} \tag{9}$$

Where x_{ei} (envelope costs), x_{hi} (HVAC system costs), x_{li} (lighting system costs), x_s (light shelf costs), x_{pv} (BiPV costs), x_{rest} (other costs), x_{rep} (present value of capital replacement costs), x_{res} (present value of residual value minus the disposal costs), x_e (present value of energy costs), $x_{om\&r}$ (present value of non-fuel operating, maintenance and repair costs). Equation 9 is adapted from the NIST Handbook 135 with no water cost included (Fuller and Petersen, 1995). The constraint of this optimization process is:

$$X = \left\{x \in \mathbb{Z}^{sc} | x^{ij}, i \in \{1..c\}, j \in \{1..s\}\right\} \tag{10}$$

The $\mathbb{Z}^{sc}$ is the constraint set, which contains the predefined building system types s and their cost information c. The design parameter x is specified as discrete independent variables that can only take the values from $\mathbb{Z}^{sc}$.

Table 1: Window options

Window	Properties
Double Clear	U[1]-3.13, SHGC[2]-0.73, Vt[3]-0.8
Double Tinted	U-2.58, SHGC-0.37, Vt-0.53
Double Thick Clear	U-1.40, SHGC-0.41, Vt-0.61
Heat Reflective Clear	U-1.40, SHGC-0.25, Vt-0.45
Triple Glazing	U-0.81, SHGC-0.71, Vt-0.53
Quadruple	U-0.781, SHGC-0.46, Vt-0.62

1. U: U-value (W/m^2).
2. SHGC: solar heat gain coefficient.
3. Vt: visible transmittance.

Design parameters

The design parameters cover building envelope systems (wall and roof), window system, lighting systems and HVAC systems. The options for each design parameter are listed in (Xu et al., 2016). Table 2 shows the design parameters, their properties and the number of design options. For window assemblies, their detailed properties are listed in Table 1. Furthermore, three HVAC systems, variable air volume (VAV), variable refrigerant flow (VRF) and hybrid dedicated outdoor air system with variable refrigerant flow (DOAS+VRF), are included in the optimization. In addition, two new systems, daylight shelf and BiPV systems are added into the design parameter list.

Daylight shelf system

The daylight shelf system on the south facing windows is included in the optimization study. The daylight shelves are one of the many daylighting devices for bringing more daylight into a building. They are typically installed as an inside shelf, an outside shelf, or both on a window. The shelves reflect the exterior light onto the ceiling of a room to achieve extended daylight penetration distance and increase uniformity in daylight distribution levels.(LBNL, 2016).
In EnergyPlus, daylight shelves are modeled via the DaylightingDevice:Shelf object, where it requires a shelf host window, a heat transfer surface (inside light shelf), an attached shading surface (outside light shelf) and a shelf construction (LBNL, 2016). In order to add daylight shelves to an existing window, the window must be separated into upper window

Table 2: Design parameters

System	Design range	DP/O[1]
External wall assembly	Concrete, U: 0.27 - 1.06 W/m^2K	1/9
Roof assembly	Metal deck, U: 0.21 - 0.45 W/m^2K	1/5
Window assembly		1/6
Lights	$4.5, 7.5, 10.2\ W/m^2$	1/3
Daylight sensor	On / Off	1/2
WWR[2]	40%, 50%, 60%, 70%, 80%	4/5
HVAC	VAV, VRF, VRF + DOAS	1/3
Light shelf height	$10\% - 30\%$	1/NV[3]
Light shelf projection	Inside: $0.5 - 1.5\ m$ Outside: $0.5 - 1.5\ m$	2/NV
BiPV	Yes / No	4/2

1. DP/O: Number of design parameters / number of design options.
2. WWR: window-to-wall ratio.
3. NV: Numeric value.

and lower window sections. The height of the upper window section decides where the daylight shelves are mounted. In EnergyPlus, daylight shelves are simulated separately for daylighting and the zone heat balance; however, the calculation details vary between the inside and outside daylight shelves. The inside shelf is assumed to reflect all the transmitted light from the upper window onto the ceiling of the room as diffuse light, which means the upgoing flux is forced to be the total transmitted flux. On the other hand, the daylighting of the outside shelf is calculated by integrating over the sky and ground and summing the luminance contribution of each sky or ground element(**?**). Most of the daylighting interaction are assumed to be between shelves and the upper window, except the lower window receives shading from the outside shelf. For heat balance calculation, the inside shelf is defined as an inter-zone heat transfer surface (partition) and the outside shelf is simulated as an external shading device.

In this study, three numeric parameters are investigated for the daylight shelf design. They are the height (h) of the upper window, the projection (p_{in}) of the inside shelf and the projection (p_{out}) of the outside shelf. A threshold has been set for each of the parameters. For h, it can be any value between 10% to 30% of the original window height. Any value that is outside of this range is considered a no daylight shelf system. For p_{in} and p_{out}, the value is limited between $0.5m$ to $1.5m$. If the projection of a shelf is less than $0.5m$, this shelf will not be modeled in the design proposal. The cost of the shelf is quoted based on the construction material cost from RSMeans construction cost database (RSMeans, 2015). In this study, the construction consists of one lightweight concrete layer, which is identical to the daylighting shelf example file in EnergyPlus.

BiPV system

Design of on-site renewable energy technologies are included in this optimization study as building integrated solar photovoltaic (BiPV) systems in the form of exterior facade claddings for opaque envelope assemblies. The selected photovoltaic technology is applicable to opaque and mono-crystalline silicon (m-Si)-based solar cells encapsulated between multiple glass layers (i.e., glass-glass encapsulation). The objective here is to provide parametric evaluations of systems integrative energy performance of BiPV elements. Therefore, the PV heat transfer integration mode is selected as "Integrated Outside Face" within EnergyPlus. In such an integration mode, the solar cells are treated as integral elements of the building envelope assemblies with bi-directional thermal interactions, where solar cell temperatures are effected by conductive heat flux within the envelope assembly, while electrical energy generated by solar cells are taken as a heat sink term and removed from heat transfer surfaces. So as to achieve this thermal interaction, the PV integrated envelope assemblies are defined by a Construction: InternalSource in EnergyPlus (LBNL, 2016), which allows locating an internal heat source or sink element, such as a PV element, within the assembly. The front glass of the BiPV system selected as a clear uncoated monolithic glass with relatively high front-side solar and visible transmittances at normal incidence (0.903 and 0.912, respectively). The PV electrical performance model is selected as the equivalent one-diode type and necessary input parameters are retrieved from the manufacturer's specifications of an actual m-Si PV module with a peak power rating of 185Wp and a solar conversion efficiency of 16.32% under Standard Test Conditions (SolarWorld, 2016). A simple DC-to-AC inverter model with a constant operating efficiency of 0.94 is assumed for the entire BiPV system.

All the PV electricity generated by the studied BiPV system is assumed to be either instantaneously used by other building systems (after being converted to alternative current by the inverters) or surplus electricity is transferred to the utility grid using bi-directional net metering without a storage medium (such as battery banks). It should also be noted that due to current limitations in EnergyPlus, ancillary equipment of the BiPV system other than inverters are not modeled and it is assumed that the entire system is continuously operating at a maximum power point without any mismatch losses. The performance of the integrated BiPV-envelope system will be individually accessed for each orientation. The cost of this system is estimated based on an actual project, which is $4600/kW.

Table 3: Performance metrics of the regression models built at each cycle

Cycle	Objective	Model	NRMSE
1	Operation costs	SVM	8.3
	Capital cost	SVM	7.8
2	Operation costs	SVM	7.3
	Capital cost	SVM	6.6
3	Operation costs	SVM	5.2
	Capital cost	SVM	5.1
4	Operation costs	SVM	5.3
	Capital cost	SVM	5.1

Study assumptions

In total, 17 design parameters are defined for this study. Among the design parameters, 3 of them are the numeric data type. The cost for each design options is provided in RSMeans® Building Construction Cost Data Book (RSMeans, 2015). In addition to the total cost estimations, operation costs for 25 years are also estimated for this study. The discount rate (3%) is defined according to the annual supplement report for the NIST Handbook 135 (Rushing et al., 2016). As part of the operation costs, utility cost models are built using the EnergyPlus economic module based on the published tariff from local utility providers.

A desktop with a configuration of i7 quad-core 3.5 GHz CPU and 16 GB RAM is used for performing the case study. For a large building with 350 thermal zones, an annual EnergyPlus simulation takes about one hour depending on the selected design options. Besides these three design parameters in numeric data type, the total number of combinations of possible design proposal is close to 50 million. Using a brute-force search is nearly impossible to complete a full enumeration of the parameter space.

Discussion and result analysis

The optimization performed 83 generations in 50 hours. Each generation produced 30 design proposals, which means 450 actual simulations and 2040 meta-model predictions were conducted. The process included four cycles and in each cycle, the normalized root mean square error (NRMSE) was recorded to track the accuracy of the trained meta-models. Table 3 indicates the NRMSE for each meta-model. First, all the meta-models are SVM in this case study and all of them have shown acceptable accuracy with less than 10% NRMSEs. Second, the NRMSEs decreased as the number of cycles increased, which demonstrated the coarse-to-fine theory mentioned previously.

Algorithm performance

Another set of optimization is conducted in parallel using the conventional NSGAII algorithm. The same scheme, GeneAS, is also implemented in the

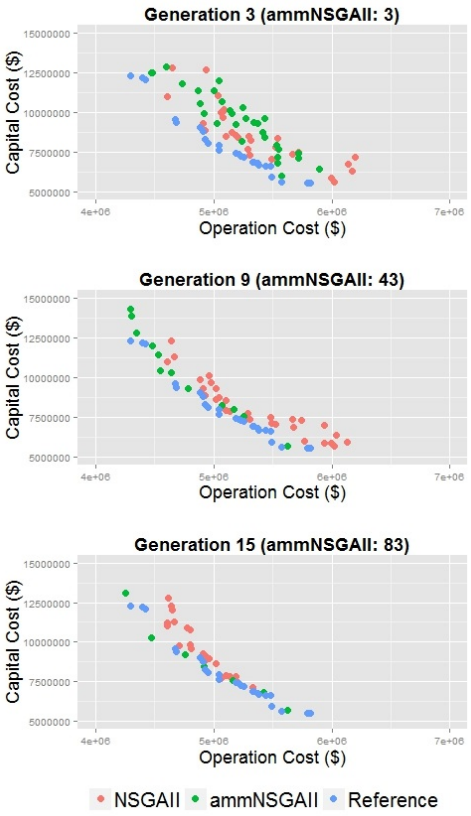

Figure 4: Optimization comparison between NSGAII and ammNSGAII

conventional NSGAII algorithm in JMetal package so that the algorithm is able to processing the mixed data types. The maximum number of evaluation (n_{max})is set to 450 which is equivalent to the number of actual simulations conducted in ammNSGAII algorithm. The constraint on the n_{max} ensures that the optimization can be completed in a relatively reasonable time frame. Besides n_{max}, the parameters of NSGAII and problem formulations are identical for both optimizations. A reference Pareto Front curve is formed by performing dominance and distance sorting on a jointed pool of optimization results. Figure 4 shows the comparison of the optimum design solutions at generation 3, 9 and 15. It should be noted that the 9^{th} generation in NSGAII is equivalent to the 43^{rd} generation in ammNSGAII because at this point, both of the algorithms have evaluated about 270 design solutions using EnergyPlus. Similarly, the 15^{th} generation in NSGAII is the same as the 83^{rd} in ammNSGAII. It can be observed that at generation 3, both of the algorithms perform similarly in terms of convergence and diversity. However, this changed significantly at generation 9. At generation 9, the ammNSGAII has completed two optimization cycles, and the results produced are quite close to the reference Pareto Front curve. On the contrary, the improvements observed in NSGAII are much slower. Similarly, in the last generation (15^{th}), the Pareto

Table 4: Performance metrics comparison between the NSGAII and ammNSGAII

Index	Model	Time (hr)	NCM	DM
3	NSGAII	14.5	0.22	0.43
	ammNSGAII	14.8	0.20	0.39
9	NSGAII	42.2	0.11	0.31
	ammNSGAII	39.6	0.06	0.37
15	NSGAII	59.4	0.08	0.28
	ammNSGAII	50.1	0.01	0.24

Front curve formed by ammNSGAII is almost identical to the reference curve while the NSGAII is still struggling to find better design solutions.

Table 4 lists the results of performance metrics for these three generations. The performance metrics include normalized convergence metric (NCM), diversity metric (DM) and speed (Time). The NCM normalizes the distances between the Pareto Front curve of the optimization and the reference curve (Equation 12). Therefore, a smaller distance indicates a better convergence. On the other hand DM measures the diversity of the generated Pareto Front curve against the reference curve (Equation 13). A smaller DM indicates a better diversity preservation of the generated Pareto Front curve.

$$d_i = \min_{j=1}^{|P^*|} \sqrt{\sum_{k=1}^{M} \left(\frac{f_k(i) - f_k(j)}{f_k^{max} - f_k^{min}}\right)^2}. \qquad C(P^{(t)}) = \frac{\sum_{i=1}^{|F^t|} d_i}{|F^t|}. \tag{11}$$

In equation 11, P^* is the reference optimal set and F is the generated optimal solution set, f_k^{max} and f_k^{min} are the maximum and the minimum function values of k^{th} objective function in P^* (Deb et al., 2002). $C(P^{(t)})$ is the convergence metric that averages the normalized distance d_i for all points in F^t. The normalized convergence metric is achieved in equation 12:

$$\bar{C}(P^{(t)}) = \frac{C(P^{(t)})}{C(P^{(max)})} \tag{12}$$

$$\Delta = \frac{d_f + d_l + \sum_{i=1}^{N-1} |d_i - \bar{d}|}{d_f + d_l + (N-1)\bar{d}} \tag{13}$$

In equation 13, d_f and d_l are the euclidean distances between the extreme solutions and the boundary solutions of the obtained non-dominated set. $\bar{d}$ is the average of all distance $d_i = 1, 2...(N-1)$ (Deb et al., 2002).

The NCM comparisons have shown that the adaptive meta-model process can effectively increase the search power. At generation 3, there is no significant difference between the ammNSGAII and NSGAII algorithm. However, after the completion of 2 optimization cycles, the ammNSGAII achieves 45% better convergence than NSGAII at generation 9. At the end, a 0.01 NCM indicates that ammNSGAII is almost converged to the reference curve, while the NCM of NSGAII is still 0.07, which is similar to ammNSGAII at generation 9.

Comparing the diversity, both of the algorithms seem to have a similar performance. The DM drops from 0.43 to 0.28 in NSGAII and from 0.39 to 0.24 in ammNSGAII. This indicates both of the algorithms can efficiently preserve the diversity in the optimal design solutions.

Lastly, a large difference is observed at the last generation. Since both of the optimization algorithms evaluated 450 design solutions with the EnergyPlus simulation, theoretically, the time elapse for optimization should be close. Although depending on the building systems chosen, an hour or two are expected in such a comparison. However, the proposed method has implemented a process that automatically skips the duplicated design solutions, which provides insight into the 10 hours difference at generation 15. The faster convergence secures most of the optimal design solutions in the ammNSGAII algorithm, thus, more duplicated cases are produced than the NSGAII algorithm. As these duplicated cases are skipped in the optimization process, they do not contribute time to the whole process.

Combined the results from (Xu et al., 2016), which used an office building with a 3700 m^2 floor area, both case studies demonstrated that the ammNSGAII outperforms the conventional NSGAII algorithm in convergence and speed performance regardless of the building size. However, it is interesting to note that the diversity preservation for ammNSGAII is improved in this case study. This is possible because there are less evaluations (450) conducted in this case study than that in (Xu et al., 2016) (900).

Design proposals with BiPV system

The high cost of the BiPV system has dramatically increased the capital cost. Depending on each orientation, the premium cost of installing BiPVs in similar designs can be somewhere between $900,000 to $5 million. However, by adding the system synergistic effects, these premium costs can be compensated by other building systems. In fact, the optimization results seemed to favor BiPV systems in this building case. Figure 5 shows the design proposals with the BiPV system at each orientation. It implies that installing the BiPV system on the south and west facades are more efficient. The south facade shows 10% higher average premium cost than the west facade due to its larger available wall areas; however, it also produces almost 9% lower operation costs.

In this study, the BiPV system interacts not only with building envelope systems, but also with the WWR. It is because larger a WWR could result in higher premium and lower operation costs. Figure 6 indicates that the optimal design proposals with the south BiPV system installed. The points in larger

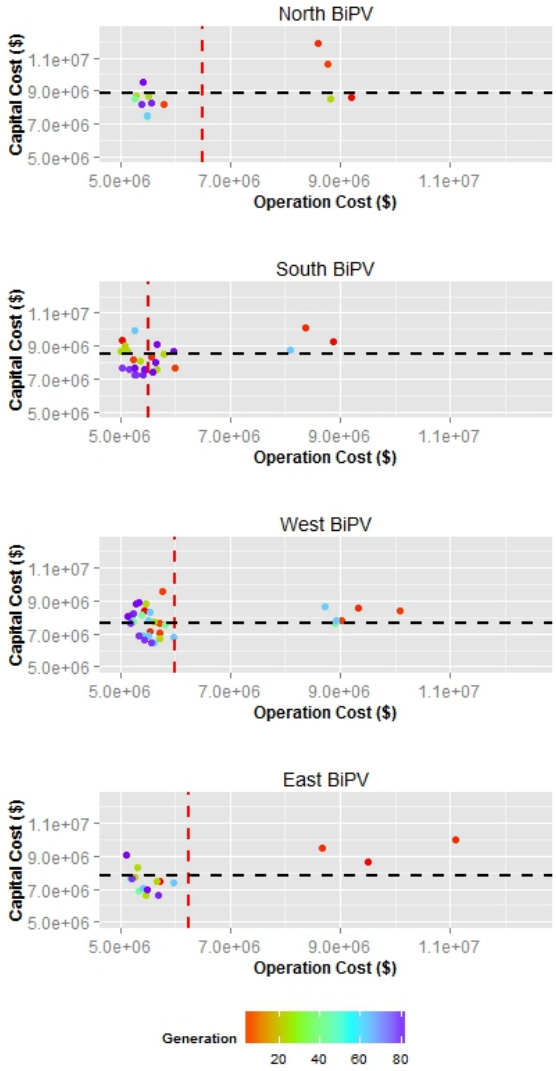

Figure 5: Design proposals with BiPV system at each orientation (red dashline: average operation costs, black dashline: average capital cost)

Figure 6: Optimal Design proposals with South BiPV system in different South WWR

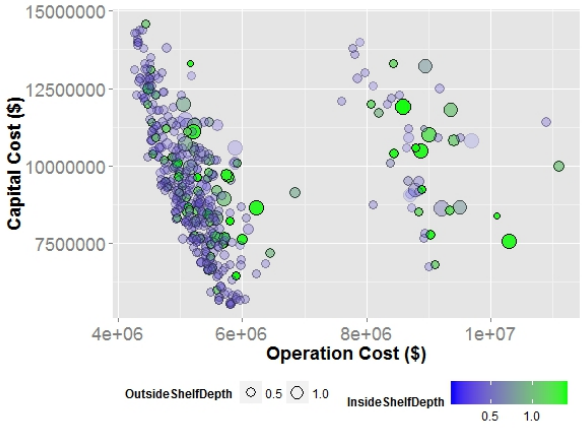

Figure 7: Design proposals with light shelf system

size imply large WWRs. It suggests that the BiPV system installed at the south facade can have a lower premium cost with WWRs from 60% to 80%. Although a large WWR may reduce the overall thermal insulation of the building, the benefits from PV generators and higher performance of daylighting sensors can compensate the loss.

Design proposals with daylight shelf system

The daylight shelf system is installed at the south facade. Figure 7 shows the generated design proposals with and without the shelves. Small circles indicate no outside shelves and blue shows no inside shelves. It can be observed that daylight shelf system is not favored by the optimization algorithm. The average operation costs for design proposals with daylight shelves are 4% higher than the rest. By comparing end-use energy consumption, the results suggests that the increase in heating and cooling energy are greater than the reduction in lighting energy. This is possibly because the daylight harvesting potential are not that effective with high efficiency lighting technologies and the sub-optimal daylighting sensor position. Furthermore, no visual comfort measures and their financial impacts are investigated in this study, which decreases the value of the daylight shelf system. Nevertheless, in Figure 7, inside shelves seems more energy efficient than the outside shelves because there are more small and green circles near the left bottom region. Also, $0.5 - 1.0m$ is a preferred depth range for both inside and outside shelves.

Design proposals

Figure 8 presents all the simulated design proposals. The color gradient represents the number of generations. In Figure 8, a substantial reduction regarding both capital and operation costs are achieved from the 1-3 generations to the 21-23 generations. The improvements show that the meta-model process in the first cycle has boosted the search power significantly. On the contrary, much smaller improvements

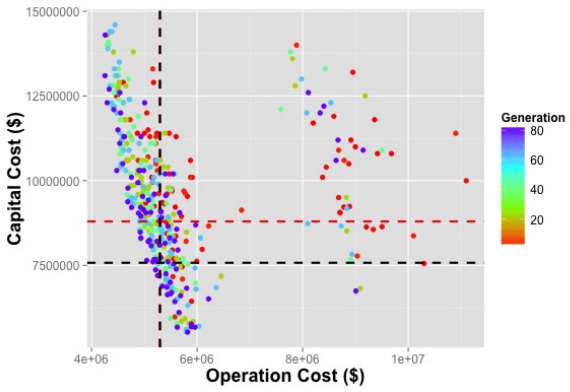

Figure 8: Design proposals generated in the optimization process

are observed between the 61-63 generations and the 81-83 generations. It suggests that the results should be very close to the global optimal design proposals.

Among the generated design proposals, 30 optimal designs formed a Pareto front curve based on the dominance theory in multi-objective optimization (Deb et al., 2002). Their capital cost ranges from \$5.5 to \$12.3 million and the operation costs are between \$4.2 to \$5.8 million. There are agreements on wall, window, roof and HVAC system selections among these design proposals. 28 out of 30 designs selected high wall constructions with a U-Value lower than $0.3W/m^2K$. Similarly, more than half of these designs choose the roof constructions with a U-Value of $0.21W/m^2$. Double thick clear window and DOAS+VRF HVAC systems have been used for 29 out of 30 design proposals. The double thick clear window has a U-Value of $1.4W/m^2K$, an SHGC of 0.41 and a visible transmittance (Vt) of 0.61. Although this window has similar SHGC and Vt but poorer insulation than a quadruple window, its lower price could be the key trade-off for the BiPV system. Similarly, half of the optimal cases choose T5 over LED lighting fixtures due to their lower capital costs.

Furthermore, 21 out of 30 design proposals decide to install BiPV systems. Among them, four design proposals installed BiPV systems on four orientations as well as having LED lighting fixtures to maximize the savings on operation costs. On average, these design proposals save 20% (\$740,000) on operations costs , but with a 40% (\$3 million) higher average capital cost compare to the others. In Figure 8, the red dashed lines map out the design proposal that has the lowest operation cost with no BiPV systems and the black dashed line map out a design proposal with a BiPV system. Both design proposals have very close operation costs (\$5.3 millions); however, the one with a BiPV system shows around 14% less capital cost due to its less expensive windows, HVAC and lighting system selections.

Conclusion

In this study, the application of an adaptive meta-model optimization framework has been extended to solve a large-scale office building design problem with mixed data type variables. Also, a renewable energy system design is included. A case study has demonstrated the effectiveness of the proposed framework in both its algorithm performance (speed and convergence) and the ability to explore the energy and costs interaction inside an integrated building system design.

The proposed algorithm, ammNSGAII, has successfully demonstrated its ability to find optimal design solutions in a large scale office building. With a maximum 450 evaluations, the ammNSGAII is able to achieve 800 times better convergences with similar diversity preservation in about 10 hour less than conventional NSGAII. The results of algorithm performances are similar to another case study conducted in (Xu et al., 2016).

In the case study, the results show that the trade-offs among the lighting fixture, window and BiPV systems could potentially minimize the operation costs up to 20% (\$740,000) in 25 years for a large size office building. On the other hand, the optimization algorithm also finds that the capital costs of these designs could be 40% (\$3 millions) higher. Additionally, the optimization algorithm thinks the daylight shelf system may not be especially efficient in reducing the building's operation costs. Therefore, the daylight shelf system does not appear in the Pareto front. Still, it demonstrates that the objectives of optimization study could be influential. If the human visual comfort benefits from the daylight shelf system were considered in this study, the results would be different. Nevertheless, the algorithm suggests an inside daylight shelf with $0.5 - 1m$ depth is preferred than other daylight shelf designs.

Future work will continue to focus on exploring the application of the adaptive meta-model optimization framework on HVAC system controls and a time-based maintenance plan for HVAC components. Furthermore, the development of a constraint handling system will also be conducted.

Acknowledgement

This research was supported by the Toshiba Research and Development Center. The author would like to thank the colleagues from Toshiba Corporation for their insight and expertise that greatly assisted the research.

References

ASHRAE (2016). Ashrae lowdown showdown modeling challenge teams recognized. `https://www.ashrae.org/news/2016/`.

Carreras, J., C. Pozo, D. Boer, G. Guillén-Gosálbez, J. A. Caballero, R. Ruiz-Femenia, and L. Jiménez

(2016). Systematic approach for the life cycle multi-objective optimization of buildings combining objective reduction and surrogate modeling. *Energy and Buildings 130*, 506–518.

Chang, C.-C. and C.-J. Lin (2011). Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST) 2*(3), 27.

Costas, M., J. Díaz, L. Romera, and S. Hernández (2014). A multi-objective surrogate-based optimization of the crashworthiness of a hybrid impact absorber. *International Journal of Mechanical Sciences 88*, 46–54.

Deb, K. (1997). Geneas: A robust optimal design technique for mechanical component design. In *Evolutionary algorithms in engineering applications*, pp. 497–514. Springer.

Deb, K. and S. Agrawal (1999). A niched-penalty approach for constraint handling in genetic algorithms. In *Artificial Neural Nets and Genetic Algorithms*, pp. 235–243. Springer.

Deb, K. and D. Deb (2014). Analysing mutation schemes for real-parameter genetic algorithms. *International Journal of Artificial Intelligence and Soft Computing 4*(1), 1–28.

Deb, K. and A. Kumar (1995). Real-coded genetic algorithms with simulated binary crossover: studies on multimodal and multiobjective problems. *Complex Systems 9*(6), 431–454.

Deb, K. and P. K. Nain (2007). An evolutionary multi-objective adaptive meta-modeling procedure using artificial neural networks. In *Evolutionary Computation in Dynamic and Uncertain Environments*, pp. 297–322. Springer.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *Evolutionary Computation, IEEE Transactions on 6*(2), 182–197.

DoE, U. (2016). Energyplus engineering reference. *The reference to energyplus calculations*.

Du, X., M. Song, and L. Li (2016). The energy efficiency assessment and optimization of collaborative working air-conditions. *Energy and Buildings 133*, 88–95.

EcoGlobe (2016). Case studies. `http://www.ecoglobe.de/case-studies/`.

Eisenhower, B., Z. ONeill, S. Narayanan, V. A. Fonoberov, and I. Mezić (2012). A methodology for meta-model based optimization in building energy models. *Energy and Buildings 47*, 292–301.

Fuller, S. K. and S. R. Petersen (1995). *NIST Handbook 135: Life-Cycle Costing Manual for the Federal Energy Management Program.* Gaithersburgh, MD: National Institute of Standards and Technology.

Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten (2009). The weka data mining software: an update. *ACM SIGKDD explorations newsletter 11*(1), 10–18.

LBNL (2016). Input output reference: The encyclopedic reference to energyplus input and output. Technical report, Lawrence Berkeley National Laboratory.

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

RSMeans (2015). *Building Construction Cost Data.* Norwell, MA: Construction Publishers & Consultants.

Rushing, A. S., J. D. Kneifel, and P. Lavappa (2016). Energy price indices and discount factors for life cycle cost analysis 2016. Technical report, National Institute of Standards and Technology.

ScienceDirect (2016). Energy and buildings. `http://www.sciencedirect.com/science/journal/03787788`.

SolarWorld (2016). Sunmodule solarworld module sw 185 mono. `http://pdf.wholesalesolar.com/module%20pdf%20folder/SW165175185.pdf`.

Tokarik, M. S. and R. C. Richman (2016). Life cycle cost optimization of passive energy efficiency improvements in a toronto house. *Energy and Buildings 118*, 160–169.

Wang, J., G. Huang, Y. Sun, and X. Liu (2016). Event-driven optimization of complex hvac systems. *Energy and Buildings 133*, 79–87.

Wetter, M. (2004). *Simulation-based building energy optimization.* Ph. D. thesis, UNIVERSITY OF CALIFORNIA, BERKELEY.

Xu, W., A. Chong, O. T. Karaguzel, and K. P. Lam (2016). Improving evolutionary algorithm performance for integer type multi-objective building system design optimization. *Energy and Buildings 127*, 714–729.