DOMESTIC BUILDING ENERGY PREDICTION IN DESIGN STAGE
UTILIZING LARGE-SCALE CONSUMPTION DATA
FROM REALIZED PROJECTS

Ji-Hyun Kim\textsuperscript{1}, Godfried Augenbroe\textsuperscript{1}, Hye-Soo Suh\textsuperscript{2} and Qinpeng Wang\textsuperscript{1}

\textsuperscript{1}College of Architecture, Georgia Institute of Technology, Atlanta, GA, U.S.
\textsuperscript{2}POSCO Engineering & Construction, South Korea

ABSTRACT
This paper suggests the use of consumption data of realized apartment buildings for estimation of energy use in future projects. This is relevant in the case of project developers that want to estimate average consumption of projects under development. A major unknown in the design prediction is the role of occupants in the energy use of the apartments. Previous research has shown that the variability of occupancy in terms of schedule, set points and appliance use is very large. We propose to capture this variability from an analysis of the consumption data of similar apartments. The prediction method uses a simple yet accurate building energy model in which the occupancy variability is added based on the observed variability in the monitored buildings. The cumulative effects of schedule, household composition, and issues of occupant behavior such as the tendency to rely on natural ventilation versus the use of the mechanical cooling systems, are rolled up into one factor, referred to as the lifestyle factor. This factor is a stochastic variable representing the variability of lifestyle of households for a given demographics. The hypothesis of our approach is that by using the lifestyle factor as the input parameter in the apartment energy prediction model we capture all aspects of occupancy variability in an apartment complex. Cross-validation results show that the proposed method accurately estimates the full range of building energy outcomes for a proposed apartment design.

INTRODUCTION
Building energy measurement data have been utilized for a variety of purposes, especially for the evaluation of building energy performance, the calibration of simulation models, and the development and calibration of estimation models. To construct a building energy estimation model based on measurement data from multiple buildings, architectural engineers have widely adopted (multiple) regression analysis (Catalina et al., 2008 and Zhao et al., 2012) since this statistical method is relatively easy to construct by processing large-scale energy data. However, regression models typically use a limited number of input parameters that are identified as key variables that have a strong influence on energy consumption in all known cases. In most instances, the regression parameters have no direct link with design parameters. Therefore regression models have limited use to test alternative building designs because the link between design parameters and regression parameters is lost in translation. This paper presents a building energy estimation method that provides better support for building design variables by adopting a simple (normative) energy calculation tool that is first principles based, and therefore uses design parameters as model input parameters, contrary to regression models. The energy calculation tool used in this study is a lightweight, Excel-based calculator based on CEN-ISO standard 13790:2008 (ISO, 2008).

Recent studies have devoted a great deal of effort to developing better occupancy models for building energy simulation, such as stochastic and agent-based approaches. The common goal of these studies is to produce estimates of energy use that are close to measurements by accounting for more realistic occupant behavior in building simulation (Azar et al., 2010, Wang et al., 2011, Widén et al., 2012, and Rysanek et al., 2014). However, without massive data gathering about occupancy variables such as presence and actions in different types of buildings, it is still hard to generate general and comprehensive occupancy models that accurately represent the role of occupancy variability in current building energy models. In this paper, we suggest a way of utilizing building energy consumption data to predict energy use in future projects without knowing the detailed occupancy variables in buildings.

While using commonly known design parameters as primary inputs in the simple energy calculation tool, we introduce a “lifestyle factor” into our calculation as an additional input in order to capture the combined effect of occupancy variables, such as presence, the operation of set-point temperatures, lighting schedules, and appliance use. This lifestyle factor is thus a rolled-up macro version of a fully descriptive occupancy model, representing cumulative impacts of occupant behaviour as related to energy consumption. This factor is particularly relevant in cases where we are unable to adopt a detailed occupancy model. We argue that especially
in residential buildings, the variability over households, attitude, affordability, and many special personal circumstances usually results in a very wide range of stochastic occupancy models that requires extensive data gathering and analysis to model. For that reason, we postulate a factor that is easier to derive and may capture all the influences sufficiently for the purpose of a project developer. Along the lifestyle factor, we will introduce other parameters to test which ones have the power to explain the variability over the observed consumption. This leads to a set of calibration parameters (including lifestyle factor) that we treat as stochastic variables that will enable us to estimate not only the means, but also the range (variability) of energy consumption over apartment households in future projects.

METHODOLOGY

The method used in this study consists of two steps (Figure 1). In Step 1, we construct 50 distinct apartment unit models, out of total 54 types of floor plans, with known inputs such as the dimensions of a building and the thermal properties of the envelope. The calculated results from each apartment model are compared with the measurement data from identical units. We use the comparison to calibrate five selected input parameters that we regard as uncertain or unknown: cooling coefficient of performance (COP), set-point temperature, internal gain, infiltration, and lifestyle factors. The calibration method is as follows: we compare the results for each apartment from the calculator with the measurements and run simulations to minimize the root-mean-square error (RMSE). Once we find the optimum values for the parameters (Step 1) for each apartment with an acceptable resulting average of the RMSE (20.1% on average), we generate cumulative distribution functions (CDFs) for each unknown parameter by using the obtained values for the apartments. In Step 2, we construct and simulate other six apartment models (out of 54 types) for cross-validation, but this time using the five calibrated parameters with the CDFs from Step 1. We then compare the results of these simulations with measurement data to check if the estimation results from Step 2 are adequate and close enough to the measured energy use data. The following sections provide more details on energy data, apartment units, and building design parameters.

Measured energy consumption data

With the support of POSCO Engineering and Construction (POSCO E&C, 2014), a Korean construction company, this study utilizes cooling energy consumption data of 2,182 POSCO E&C apartment units located in Asan, Incheon, Gyeonggi in one climate zone, and Busan, Geoje in another climate zone in South Korea. These five apartment communities have unique floor plans according to the size of their units. Each community consists of at least 8 units with the same floor plan and as many as 143 units. That is, the units share the same floor plan and thermal properties of the building envelope. Therefore, these data are useful for analysing the isolated effect of occupant behaviour on energy consumption in each unit.
Figure 2 illustrates the patterns of cooling energy use intensities during the cooling months (from June to September) for units with the same floor plan and the same orientation. The floor areas of Types 24, 25, and 26 are 108 m² and that of Type 26 is 112 m², and the numbers of the units for Types 24, 25, and 26 are 143, 139, and 137 units, respectively. All of these types are located in Gyeonggi, South Korea. As captured in this figure, the cooling energy consumption is widely dispersed from 0 kWh/m²/month to 3.19 kWh/m²/month. Even though the units have the same floor plan and envelope thermal properties, they exhibit a large discrepancy between their actual and average (red-dotted line) uses (Figure 2), which originates from the unique occupant behavior in each unit such as occupant schedules, temperature set point control, appliance and lighting uses, and natural ventilation use. Figure 2 also shows that the deviation from the average is wider in the hottest months (July and August) than in the intermittent months (June and September). For example, the standard deviations in August for
Types 26, 27 and 28 are 35.61%, 56.06%, and 37.08% and in June, 15.38%, 21.23%, and 18.15% respectively. When we introduce a lifestyle factor in the energy calculation in a later section, we expect that this factor will be able to capture the unknown impact of occupant behavior and its wide variability on energy consumption captured in these data.

Table 1 shows the number of units of each floor plan, its location, floor areas, and in step 1 or step 2. The gray-colored units in Table 1 are utilized for the validation steps (Step 2).

**Normative calculation tool**

This study uses a normative building energy calculation tool developed by the Georgia Institute of Technology and based on CEN-ISO standard 13790:2008, the Energy Performance Coefficient (EPC) calculator (ISO, 2008). This lightweight and Excel-based calculator is widely adopted and recognized as an adequate tool especially for large-scale building performance analysis (Lee et al., 2011 and Heo, 2012) and building rating purposes (Kim et al., 2013) since normatively-defined modeling assumptions and parameters constrain the number of building inputs and thus circumvent modeler’s bias and reduce the chance of modeling errors. Another benefit is the limited computational effort, which is a benefit since we have to perform 50 calibrations that requires optimization with on average of 1,500 evaluations of yearly energy consumption before full convergence. In this study, we use an EPC calculator based on the monthly quasi-steady-state calculation method (ISO, 2008).

**Building design parameters**

Table 2 shows the categories of the modeling parameters of the apartment units, separated into known and unknown parameters based on the data that was available in this study. Known parameters are treated as deterministic in all models because either we are able to collect the information or make good guesses throughout the construction data. On the other hand, five unknown parameters including building air leakage level, cooling system COP, internal gain, occupant/appliance/lighting schedules, and cooling set-point temperature are treated as stochastic parameters since those values are potentially different in each apartment unit due to the lack of knowledge and/or data of occupants’ behavior such as their household particulars. Among these we identify the five calibration parameters, since internal gain, schedules, and set-point temperature largely depend on occupants’ presence and control actions in real life, we decide to calibrate them to make the closest match with the measurement, instead of using typical occupancy schedules of residential housing given by standards or used in current simulation tools. In the case of occupants, appliances, and lighting schedules, which are defined based on the time of the day and the day of the week, it contains more than 140 unknown input parameters. Instead of calibrating all of these schedule parameters inputs, we use standard schedules for residential buildings defined in commercial reference buildings (mid-rise apartment) of the U.S. Department of Energy (DOE, 2012) and estimate the total internal gain alternately. Because of this simplification, we only have one unknown input parameter accounting for the internal gain.

Unlike typical residential buildings in the United States, cooling systems in South Korea housing are usually not installed in all units by default; heating systems such as a radiant floor heating system with a boiler, however, are typically installed when the units are built. For example, the data set of this study include only one apartment community in Incheon, South Korea, with built-in air-conditioning systems inside the units. Consequently, without a housing inspection, we do not know if a unit has a cooling system. Furthermore, even if a unit has a cooling device such as a floor-standing air-conditioning system, its actual use may be occasional (Bae et al., 2009). In Korean housing, natural ventilation through open windows is a more common way of creating an acceptable level of indoor thermal comfort (Cho et al., 2011) and as a result, the use of the cooling system will be significantly reduced. In this study, we consider the cooling system COP unknown as different systems may be installed in the units, and the use of the systems at different part load fractions may vary significantly between different households. In addition, we consider the average outside air infiltration as an additional calibration parameter to account for a random use of natural ventilation in the units, which obviously may differ significantly across the apartment units.

**Lifestyle factor**

In addition to the calibration parameters explained above, we introduce a lifestyle factor associated with all possible occupant actions in an apartment unit that are almost impossible to predict because of their randomness in nature and possible constraints and intrinsic factors such as demographic, health, age, gender, household size that are not recorded for the monitored apartments. One additional factor is the accessibility to cooling systems (built-in or optional post-installation) inside units. Figures 3 shows the energy use intensities (EUIs) of the apartment units in Incheon and Gyeonggi. Other than the size of their floor plans, one major difference between these units is the cooling system option. In other words, the units in Incheon have a built-in air-conditioning system and those in Gyeonggi do not, so occupants in Gyeonggi have less access to cooling. Figure 3, which plots the EUIs of all units in Incheon and Gyeo-
Figure 3 EUIs of Units in Incheon and Gyeonggi, Korea

Figure 4 Measured and Calculated EUIs of One Unit in Gyeonggi, Korea

onggi, shows a stark difference between the two communities: higher cooling energy consumption in Incheon resulting from the accessibility to cooling systems. From the large-scale data set of 2,182 apartment units in this study, determining actual building use, such as window-opening behavior, seems virtually impossible without a detailed occupant survey and/or monitoring. As seen in Figure 4 for one selected apartment, when we apply standard occupant inputs in the energy model, the results of monthly cooling energy consumption (dark-blue-colored) differ markedly from those of the measurements (light blue). Once we manually set the input conditions, the results (dotted) closely resemble those of the measurements with the RMSE of 14.09%. In this case, we assumed the conditions of cooling system use based on the fact that natural ventilation is a common way of cooling in Korean housing. We set several conditions of natural ventilation use by using outdoor temperature, wind speed, and occupant schedule. Figure 4 shows the closeness of the result to measurement, which is based on the confirmation that occupants utilize natural ventilation through open windows if it is relatively cool outside (less than 25.8 °C), the wind speed is sufficient for natural ventilation, and if only a few people are in the house. These results could indicate that we need a detailed occupant survey and monitoring when targeting the estimation of actual energy use of individual apartments. For our study the individual prediction is less relevant as the developer is primarily interested in the spread over the apartments in a new targeted development.

In the validation (Step 2), we found that a single summer long lifestyle factor for four cooling months from June to September was not able to capture the occupant diversity adequately for different months. As shown in Figure 5, 30% to 40% of units of Type 22 in Gyeonggi have zero cooling energy use in June and September, and one lifestyle factor could not account for this tendency. Therefore, this study introduces two separate lifestyle factors for June and September (the intermittent months) and July and August (the hot summer months) based on the different cooling energy use patterns shown in the measurement data.

RESULTS AND DISCUSSION

Step 1: Calibration of unknown parameters

Among the 54 types of floor plans, 50 units were modeled and calculated in Step 1 and five unknown parameters were calibrated for each unit separately to make the calculated results close to those of the measurement data using the EPC monthly calculator and MATLAB coding (MathWorks, Inc., 2005). These 50 types of floor plans account for 1,966 units among 2,182 units in the data set. Table 3 shows the lower and upper bounds of unknown parameters, including cooling COP, infiltration rate, internal gain, set-point temperature and lifestyle factors, which we calibrated in MATLAB coding to minimize the RMSE (20.1% on average) for each apartment unit.

Step 2: Estimation with pre-defined unknown parameters

For a cross-validation test, we estimate the cooling energy use of Type 2 in Incheon, Types 28 and 33 in Geoje, and Type 45 in Busan (Table 1) and compare the results of the estimations with those of the measurements (Figures 6 to 9). By utilizing the CDFs from Step 1, we sample the values of the five unknown parameters from these distributions to create 1,500 samples that are subjected to an individual calculation run. The results are used to make a probability distribution of the energy consumption. This is done for all different floor plans. Figures 6 through 9 show both the estimated and measured cooling energy uses with their probabilities in percentages for each unit listed above. The figures also present the deterministic EPC calculation results, calculated without pre-defined parameters, in green. This deterministic prediction did not include a lifestyle factor, whereas the other calibration parameters (cooling COP, infiltration rate, internal gain, and cooling set-point temperature) were for all apartments identically set to their typical values used in current simulation practice based on commercial reference buildings (mid-rise apartment) of the U.S. Department of Energy (DOE, 2012). Overall, Figures 6 through 9 show that the estimated variation of the energy use of individual floor types corresponds well to the
measured energy use pattern across the data set. In the case of Type 45 in Busan (Figure 9), its probability distribution and mean of cooling energy use are relatively similar with an average 12% difference. However, when a smaller number of units was available for the energy data, a weaker resemblance of the probability distributions occurred, for example, in Type 2 in Incheon (Figure 6). In the same context, the estimated results of Type 33 (Figure 8) in Geoje with 59 units show closer agreement with the measured cooling energy use pattern than Type 28 (Figure 7) in Geoje with only 33 units. Despite the relatively poorer match for some floor types, all of the estimated results, by considering occupant impact, show a much more realistic indication of the cooling energy use than the deterministic EPC results. When the number of measurement data is large enough to provide the variability of occupancy, this method, unlike typical simulation tools, is relatively simple and efficient in terms of modeling and computational effort that must be devoted to estimating more realistic energy use of apartments. Moreover, the outcomes allow the developer to make better estimates of the aggregated consumption of the apartment complex.

CONCLUSION

This paper proposes a way of utilizing energy consumption data of apartment units to predict energy use in future projects by estimating the cumulative effects of occupant behavior from the data of realized projects. This study introduced a new parameter, lifestyle factor, to capture the cumulative effects of occupant behavior and possible but unknown operational constraints, such as accessibility to cooling systems (i.e., built-in or optional post-installation of a cooling system). When we estimated the cooling energy consumption of new designs using the CDFs of the five calibrated parameters in Step 1, we found that this method provides good estimates of not only the means, but also the range (variability) of energy consumption, i.e. far more realistically than typical simulation results based on standard assumptions. The results are significant for project developers of new apartment complexes, as they want to not only guarantee a certain mean energy consumption of the total building but also anticipate the wide range of actual consumption of individual apartment, as this has implications for their tenant leasing agreements as well as energy code compliance and potential electric utility contracting. By examining a larger amount of measurement data including cultural, spatial, and demographical differences of buildings, future research could yield findings that would prompt engineers and architects to apply this method more universally.

<table>
<thead>
<tr>
<th>Optimized Parameters</th>
<th>Cooling COP</th>
<th>Infiltration [ACH]</th>
<th>Internal Gain [W/m²]</th>
<th>Cooling Set-point Temperature [°C]</th>
<th>Life-style factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower bound</td>
<td>3</td>
<td>0.1</td>
<td>15</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>Upper bound</td>
<td>4</td>
<td>9</td>
<td>60</td>
<td>28</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 6 Step 2: Result Comparison (Type 2, Incheon)

Figure 7 Step 2: Result Comparison (Type 28, Geoje)

Figure 8 Step 2: Result Comparison (Type 33, Geoje)
ACKNOWLEDGEMENT

The authors would like to acknowledge the support of POSCO Engineering and Construction, South Korea.

REFERENCES

Mitchell, J.W., Beckman, W.A. 1995. Instructions for IBPSA Manuscripts, SEL, University of Wisconsin, Madison USA.


POSCO E&C, POSCO ENGINEERING & CONSTRUCTION., LTD. 2014


Zhao, H. X., Magoulès, F. 2012. A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, 16(6), 3586-3592.