Inter-Occupant Diversity in Occupant Behaviour Models:
Exploring Potential Benefits for Predicting Light Switch-on Actions

Farhang Tahmasebi[1], William O'Brien[2], Yan Wang[1]
Samuel Stamp[1], Dejan Mumovic[1], Ardeshir Mahdavi[3]
[1]UCL Institute for Environmental Design and Engineering, London, United Kingdom
[2]Department of Civil and Environmental Engineering, Carleton University, Ottawa, Canada
[3]Department of Building Physics and Building Ecology, TU Wien, Vienna, Austria

Abstract

The present paper reports on the development and cross-validation of a number of light switch-on behaviour models with a focus on the implications of inter-occupant and site-specific diversities. To this end, the study models light switch-on actions by occupants at three office buildings in London, Ottawa and Vienna. The models, which are developed at both single-occupant and office-level granularities are subjected to on-site and off-site cross-validations. Thus, the study can examine if the inclusion of behavioural diversity information enhances the predictive potential of occupant behaviour models in different settings. The results of the study suggest that light switch-on models with single-occupant granularity can better predict the occupants, based on the occupant for whom they are developed. This, however, does not translate into more accurate predictions of other occupants in different settings. Reflecting on the obtained results, the authors highlight other potential benefits of occupant behavioural diversity information for building performance simulation, which need to be further explored.

Introduction

The building simulation community has increased its efforts to reduce the gap between predicted and actual building energy use through probabilistic representations of occupants in buildings (Schweiker 2017). However, several studies have demonstrated that the use of existing occupant behaviour models involves considerable uncertainties and does not necessarily lead to a more reliable performance assessment (e.g., Tahmasebi & Mahdavi 2017; Gilani et al. 2018). Specifically, it is shown that without proper treatment of the diversity in occupants' behaviour, probabilistic occupancy-related models fail to provide representative ranges of occupant behaviour possibilities as intended (O'Brien et al. 2017; Tahmasebi & Mahdavi 2016). To address this issue, different approaches for inclusion of diversity in occupant behaviour modelling efforts have been examined (Reinhart 2004; Mahdavi 2017; O'Brien et al. 2017; Haldi et al. 2017). However, with regard to the usefulness of including diversity information in occupant behaviour models, previous studies have yielded somewhat contradictory results (Gilani et al. 2018; Tahmasebi & Mahdavi 2018). Consequently, the potential benefits of supplying occupant behaviour models with inter-occupant diversity information are not conclusively established. In this context, the current contribution revisits the problem to investigate three research questions:

1. Does modelling the inter-occupant diversity in light switch-on models contribute to better predictions of light use in the same setting?
2. To which extent do the data-driven light switch-on models offer a predictive potential in new settings across different countries?
3. Does inclusion of the inter-occupant diversity information in the modelling effort enhance the models' predictive performance in these new settings?

Method

Buildings and occupant data

The current study focuses on three office buildings in London, Ottawa and Vienna. The monitored areas in London and Ottawa offices each include 11 workstations and the Vienna office comprises of 8 workstations. In all three office areas the occupants have access to a light switch to control the lighting conditions of the workplace. The current study deploys streams of 15-min interval data on occupants' presence, state of the light and indoor illuminance from the offices. In the case of the London offices, workplane indoor illuminance has been measured, while at Ottawa and Vienna offices ceiling-mounted sensors have measured illuminance. The study uses one-year datasets from London and Vienna offices along with two years' worth of data from Ottawa offices (referred to as estimation period) to infer three sets of light switch-on models. A separate two-year-long dataset from Ottawa offices (referred to as validation period) serves to test the predictive potential of the developed models.

Table 1 includes basic information on the use of lights in the studied offices. Specifically, to provide an impression of the state of the lights during occupied and unoccupied intervals and the number of light switch-on actions, three indicators are shown in the table: Fraction of occupied intervals with switched-on lights, fraction of unoccupied intervals with switched-on lights and light switch-on rate at occupied intervals. The table includes the values of the metrics for individual occupants together with the mean and standard deviation of the metrics across occupants.

16th IBPSA
INTERNATIONAL
CONFERENCE
AND EXHIBITION

INTERNATIONAL
BUILDING
PERFORMANCE
SIMULATION
ASSOCIATION

Table 1: Fraction of occupied and unoccupied intervals with switched-on light and rate of light switch-on actions at the studied offices in London, Ottawa and Vienna.

Occupant	Fraction of occupied intervals with switched-on lights [-]			Fraction of unoccupied intervals with switched-on lights [-]			Light switch-on rate in occupied intervals [1/hour]		
	London	Ottawa	Vienna	London	Ottawa	Vienna	London	Ottawa	Vienna
L01	0.89	0.32	0.04	0.03	0.03	0.00	0.35	0.29	0.16
L02	0.73	0.93	0.04	0.23	0.08	0.00	0.44	0.41	0.15
L03	0.84	0.86	0.01	0.28	0.08	0.00	0.51	0.31	0.03
L04	0.84	0.28	0.01	0.28	0.04	0.00	0.51	0.34	0.03
L05	0.82	0.15	0.01	0.27	0.01	0.00	0.50	0.18	0.03
L06	0.82	0.79	0.01	0.27	0.09	0.00	0.50	0.62	0.04
L07	0.83	0.38	0.04	0.18	0.02	0.00	0.28	0.26	0.13
L08	0.83	0.13	0.01	0.18	0.04	0.00	0.28	0.12	0.04
L09	0.31	0.41	-	0.03	0.01	-	0.18	0.37	-
L10	0.33	0.78	-	0.02	0.01	-	0.18	0.73	-
L11	0.92	0.43	-	0.15	0.11	-	0.34	0.48	-
Mean	0.74	0.50	0.02	0.17	0.05	0.00	0.37	0.37	0.08
STD	0.20	0.28	0.02	0.10	0.04	0.00	0.12	0.17	0.06

Light switch-on models

To explore the research questions, the authors developed the following types of light switch-on models based on the monitored data obtained from the offices in the estimation period:

- Aggregate models (AGG), which were derived based on the data obtained from all occupants and light switches in each office area without any weighting factor;
- Individual models (L1-L11 in London and Ottawa offices, L1-L8 in Vienna office), which were derived based on the data obtained from each occupant and their associated light switch.

Each model is a logistic regression, which estimates the probability of turning on the lights based on indoor illuminance:

$$P = \frac{\exp(\beta_0 + \beta_1 E_{in})}{1 + \exp(\beta_0 + \beta_1 E_{in})}$$

Where P is the probability of a light switch-on action, E_{in} is the indoor illuminance, β_0 is the intercept and β_1 is the regression coefficient. Note that, without a distinctive treatment of different occupancy phases (arrival, intermediate and departure times) and by skipping a comprehensive variable selection procedure, the models are kept simple to focus the effort on the research questions. It should be also mentioned that, to determine the state of the lights in a building performance model throughout the simulation period, these models must be accompanied by switch-off models. This is, however, out of the scope of the current study.

Cross-validation scenarios

The present contribution follows two scenarios to examine the implications of inter-occupant and location-specific diversity for the explanatory power of light switch-on models.

Firstly, to answer the first research question in an ideal scenario, each single-occupant model from Ottawa dataset is used to predict the occupant, based on whom the model is developed. Thus, the study can explore the potential advantages of integrating inter-occupant diversity information in light switch-on models.

Secondly, to answer the second and third research questions, the models developed based on London and Vienna offices are tested to predict the occupants at Ottawa offices. Thereby, both aggregate and individual models are deployed to examine the explanatory power of the models in different settings with and without inclusion of inter-occupant diversity. In this scenario, the individual models are randomly assigned to the occupants in the new setting.

Model evaluation metrics

In order to measure the models' ability to discriminate between occurrence and non-occurrence of light switch-on, the authors deployed a rather novel metric in this field, namely the coefficient of discrimination. This metric is suggested by Tjur (2009) in the context of logistic regression as an analogue to coefficient of determination (R-squared) in ordinary regression models. Coefficient of discrimination (CoD) is calculated as follows:

16th IBPSA
INTERNATIONAL
CONFERENCE
AND EXHIBITION

INTERNATIONAL
BUILDING
PERFORMANCE
SIMULATION
ASSOCIATION

$$CoD = \bar{P}_1 - \bar{P}_0$$

Where $\bar{P}_1$ and $\bar{P}_o$ denote the average estimated probabilities of light switch-on, respectively, for the so-called successes (intervals, in which a switch-on has been observed) and failures (intervals, in which a switch-on has not been observed). Coefficient of discrimination takes values between 0 and 1. A value of 0 corresponds to "no explanatory power" (all fitted values are equal), and a value of 1 corresponds to "perfect fit" (the fitted values coincide with the observations). This indicator offers two advantages in evaluation of occupant behaviour models: Firstly, it is obtained without conducting Monte Carlo simulation of the model (as it deploys the estimated probabilities and not the randomly-sampled states). Secondly, it does not require the inclusion of models' feedback as it examines the models' predictions in a set of disconnected time intervals.

With this probability-based approach to model evaluation, it is also possible to estimate a number of widely-used metrics for occupant behaviour models (e.g. Schweiker et al. 2012) without relying on the state of devices through the Monte Carlo simulation of the model. These metrics, which in the present study serve to evaluate the models in terms of the agreement between predicted and monitored switch-on actions, are as follows:

- TPR (True Positive Rate), as the proportion of actual switch-on actions, which are correctly predicted. This metric is equivalent to $\bar{P}_1$.
- FPR (False Positive Rate), denoting the proportion of intervals without a switch-on action, for which occurrence of a switch-on action is predicted, as an equivalent of $\bar{P}_0$.
- TNR (True Negative Rate), denoting the proportion of intervals without a switch-on action, which are correctly predicted. This metric is equivalent to $1 - \bar{P}_0$.
- FNR (False Negative Rate) as the proportion of actual switch-on actions, which are wrongly predicted, as an equivalent of $1 - \bar{P}_1$.

In addition, the authors included another widely used metric in occupant behaviour studies referred to as Accuracy (ACC). While it has been argued that the way, in which this metric has promoted occupant behaviour models can be misleading, it can serve to put the models' predictive performance in perspective. In the present study, ACC is seen as an indicator for the models' potential in discriminating actions and nonactions weighted by their numbers, which is calculated as follows:

$$ACC = \frac{\bar{P}_1 s + (1 - \bar{P}_0)f}{s + f}$$

Where s denotes the number of successes (occurrence of light switch-on actions), and f denotes the number of failures (non-occurrence of light switch-on actions).

Results

Light switch-on models

Table 2 provides the estimated coefficients of logistic regression models for the individual and aggregate light switch-on behaviour at London, Ottawa and Vienna offices. Figure 1 to Figure 3 illustrate the models' response curves at indoor illuminance of 0 to 300 lux.

Note that the models based on the data from London offices use workplane illuminance, whereas the Ottawa and Vienna models take ceiling illuminance as input.

Table 2: Estimated coefficients for logistic regression models of light switch-on based on single occupant/light data (L1 – L11) and the aggregate data from all occupants and lights in each office area (AGG).

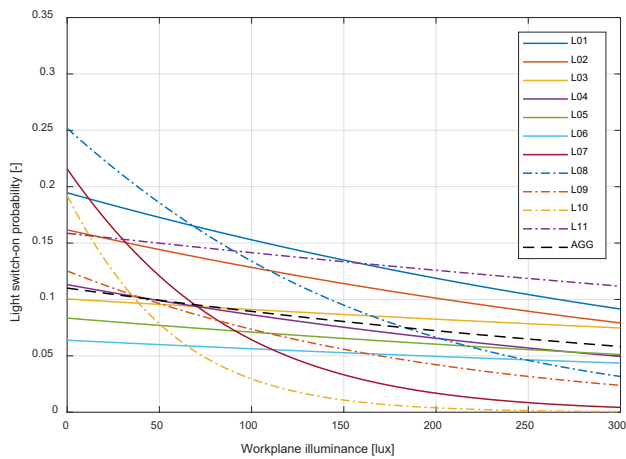| Model	Parameter	Estimate		
		London	**Ottawa**	**Vienna**
L01	β_0	-1.4203	-1.2938	-2.0994
	β_1	-0.0029	-0.0450	-0.0120
L02	β_0	-1.6457	1.2943	-2.2482
	β_1	-0.0027	-0.0172	-0.0098
L03	β_0	-2.1927	0.3499	-3.9995
	β_1	-0.0011	-0.0197	-0.0110
L04	β_0	-2.0588	-1.3910	-3.6027
	β_1	-0.0030	-0.0131	-0.0240
L05	β_0	-2.3964	-0.6994	-3.3241
	β_1	-0.0017	-0.0597	-0.0166
L06	β_0	-2.6852	0.7360	-3.6342
	β_1	-0.0013	-0.2909	-0.0048
L07	β_0	-1.2908	-0.1651	-1.8020
	β_1	-0.0139	-0.5580	-0.0315
L08	β_0	-1.0894	-1.6749	-2.8064
	β_1	-0.0078	-0.5726	-0.0140
L09	β_0	-1.9434	-1.2640	-
	β_1	-0.0059	-0.0844	-
L10	β_0	-1.4360	0.3004	-
	β_1	-0.0206	-0.1099	-
L11	β_0	-1.6678	-0.4881	-
	β_1	-0.0014	-0.1083	-
AGG	β_0	-2.1230	-1.0403	-2.9109
	β_1	-0.0020	-0.0193	-0.0107

Figure 1: Estimated single-occupant and aggregate light switch-on probabilities at London offices.
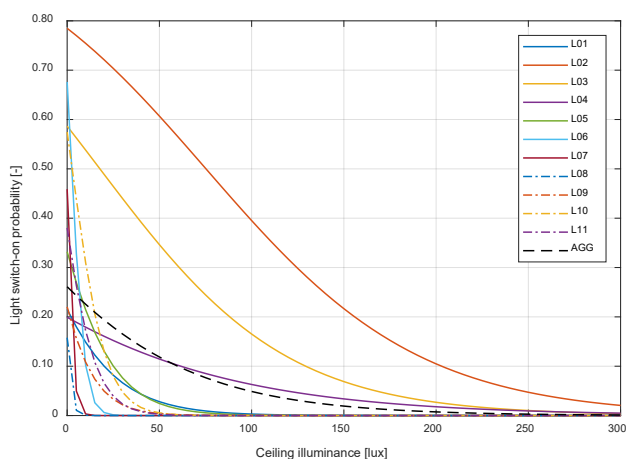

Figure 2: Estimated single-occupant and aggregate light switch-on probabilities at Ottawa offices.
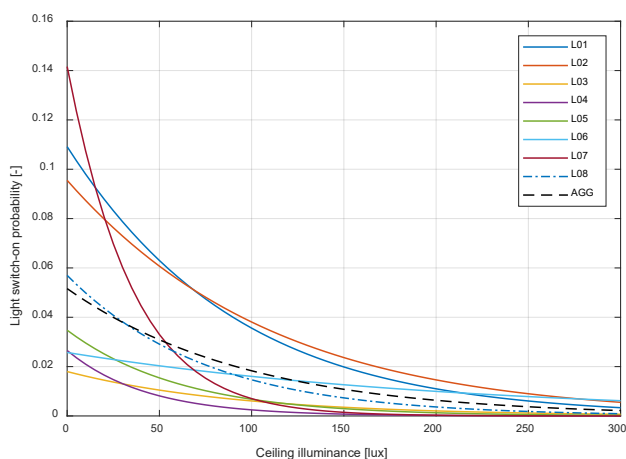

Figure 3: Estimated single-occupant and aggregate light switch-on probabilities at Vienna offices.

Models' predictive performance

Table 3 provides the values of $\bar{P}_1$, $\bar{P}_o$, Coefficient of Discrimination (CoD), TPR, FPR, TNR, FNR and ACC for the aggregate and individual models developed based on London, Ottawa and Vienna data as tested on the Ottawa offices in the validation period. The table also

includes a pseudo-model that does not return any switch-on actions (i.e. it assumes the lights are always switched off) to put the models explanatory power in context.

Moreover, to better compare the performance of Ottawa models in discriminating light switch-on actions, Figure 4 shows the predicted switch-on probabilities by aggregate and individual models for separate sets of intervals as follows: a) intervals with a switched-off light, where a switch-on actions is observed (i.e. the ideal prediction equals 1) on the left-hand side histograms; b) intervals with a switched-off light, where a switch-on action is not observed (i.e. the ideal prediction equals 0) on the right-hand side histograms.

Discussion and conclusion

Models' explanatory power in the same setting

Focusing on the predictive performance of Ottawa-based models to answer the first research question, the values of CoD and ACC provided in Table 3 indicate that the individual light switch-on models outperform the aggregate model (CoD of 0.1 versus 0.03, and ACC of 0.85 versus 0.79). The higher explanatory power of individual models can be also seen in Figure 4. Comparing the two histograms on the left-hand side shows that the individual models have, to some extent, pushed the histogram mass to the right endpoint (toward the ideal prediction of 1). More noticeably, contrasting the histograms on the right-hand side reveals that the individual models have pushed the histogram mass to the left end point (toward the ideal prediction of 0).

As mentioned at the outset, this validation procedure represents an ideal scenario, in that model estimation and validation happens at the same setting (Ottawa offices) and the single-occupant models predict the very occupants, based on whom they are developed. Nonetheless, the results underline the theoretical potential of inter-occupant diversity information toward enhancing the explanatory power of occupant behaviour models.

Models' explanatory power in different settings

In view of the applicability of the models in different settings (research questions 2 and 3), the results provided in Table 3 suggest the following: The models' predictive performance is far from satisfactory and the inclusion of inter-occupancy information, if anything, reduces the models' explanatory power. This can be seen, in particular, by the extremely low values of CoD (and by the lack of improvement due to the inclusion of inter-occupant diversity information) in testing London and Vienna models based on Ottawa data.

In this regard, one may argue that the high values of ACC, namely above 0.85 in four cross-validation tests, are promising. However, the ACC value of 0.98 resulting from the no-switch-on pseudo model (which does not require any modelling effort) put this in a better perspective. Note that, ACC is obtained by weighting the models' predictions based on the number of observed intervals with and without actions. Arguably, in any study on the use of lights, the intervals in which a switch-on

16th IBPSA
INTERNATIONAL
CONFERENCE
AND EXHIBITION

INTERNATIONAL
BUILDING
PERFORMANCE
SIMULATION
ASSOCIATION

action is observed comprise a very small portion of the intervals in which the occupant has the possibility to turn on the light. In the current case study, the proportion is as small as 1.7%, which results in an ACC value of 98.3% for a pseudo model that assumes that the occupants never turn on the lights. Therefore, since the occupant behaviour models are intended to capture the – perhaps infrequent but arguably influential – adaptive actions, an indicator such as CoD can better direct the researchers toward enhancing the explanatory power of the models.

Table 3: Values of the model evaluation metrics for aggregate and individual light switch-on models developed for London, Ottawa and Vienna offices as tested for Ottawa offices in the validation period.

| Model(s) type	Estimation location	Validation location	$\bar{P}_1$	$\bar{P}_0$	CoD	TPR	FPR	TNR	FNR	ACC
Aggregate	Ottawa	Ottawa	0.24	0.20	0.03	0.24	0.20	0.80	0.76	0.79
Individual			0.23	0.14	0.10	0.23	0.14	0.86	0.77	0.85
Aggregate	London		0.11	0.10	0.01	0.11	0.10	0.90	0.89	0.88
Individual			0.13	0.13	0.00	0.13	0.14	0.86	0.87	0.85
Aggregate	Vienna		0.03	0.02	0.01	0.03	0.02	0.98	0.97	0.97
Individual			0.10	0.09	0.01	0.10	0.09	0.91	0.90	0.90
No switch-on	-		0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.98
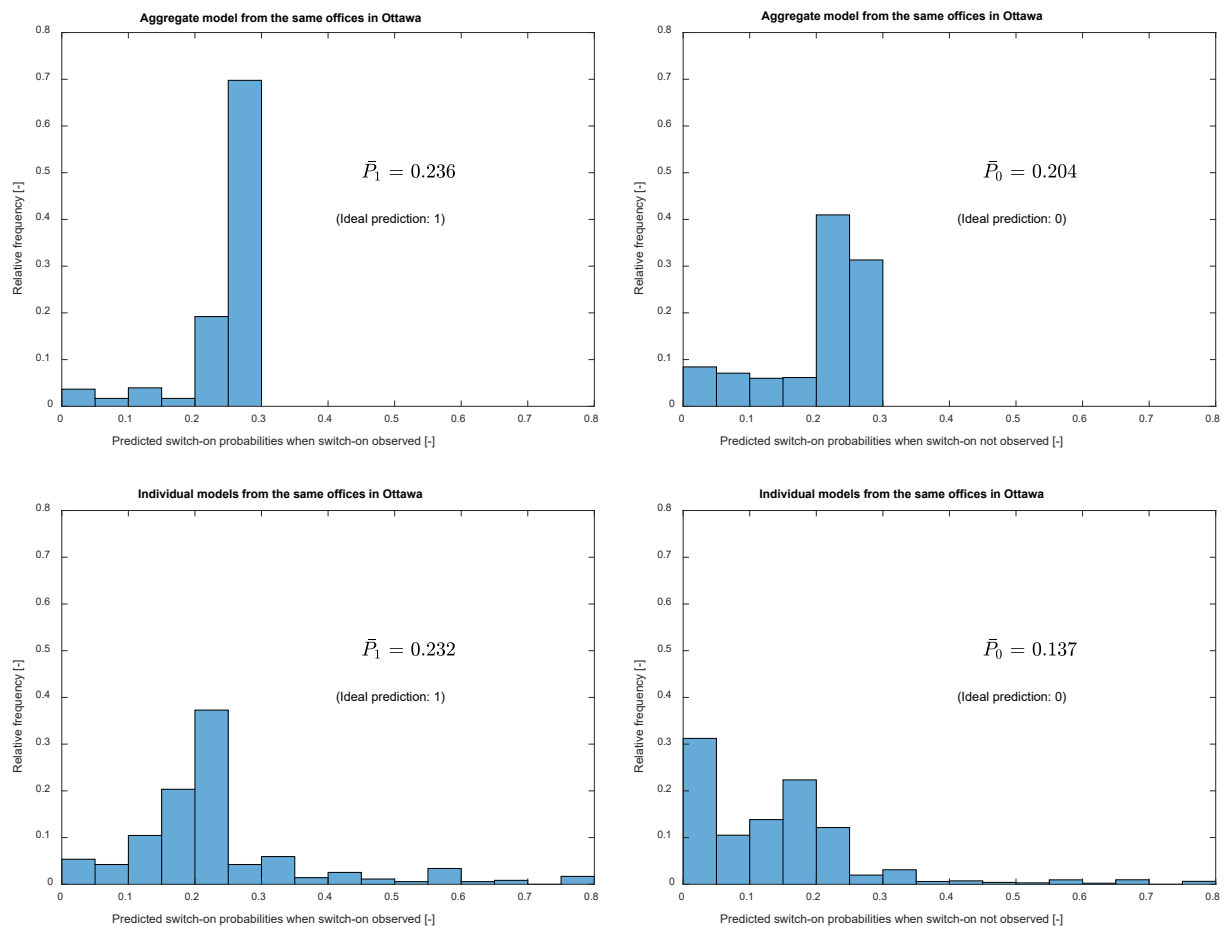
Figure 4: Histograms of predicted light switch-on probabilities for the Ottawa offices in the validation period provided by aggregate (top) and individual (bottom) models from the same offices in Ottawa. The histograms on the left-hand side show the predictions for the intervals in which a light switch-on is observed. The histograms on the right-hand side show the predictions for the intervals in which a light switch-on is not observed.

Reflections on the challenges and future research

Setting the findings on the three research questions side by side, the authors argue that, in principle, the inclusion of inter-occupant diversity in behaviour models can enhance their predictive power. However, this does not constitute an easy solution to extend the applicability and reliability of occupant behaviour models in different settings. In fact, given the very diverse nature of occupant behaviour in the three buildings studied here, relying on the off-site aggregate models to capture the average tendencies in occupants' use of lights seemed to a be more reasonable choice.

More specifically, as compared with other aspects of occupant adaptive behaviour in buildings, use of electric lighting seems to pose a greater challenge to develop occupant behaviour models for different settings. As it stands, indoor illuminance measurements are conducted differently from one place to another. Besides, given the evolving nature of everyday activities, setting an appropriate metric to represent the perceived adequacy of illumination does not seem to be a trivial task (see Raynham et al. 2019, as a recent effort in this area).

All in all, the authors believe that the motivation behind understanding and modelling inter-occupant behavioural diversity is not solely to enhance the accuracy of occupant behaviour predictions in different settings. Rather, to provide effective simulation-based design support, occupant models would have to be based on statistically representative diversity information. The necessity for inclusion of diversity depends on the design questions. For example, studies focused on robustness (e.g., O'Brien and Gunay, 2015) and other probabilistic results require that inter-occupant diversity be quantified. In contrast, for large building or community-scale projects, inter-occupant diversity tends to cancel out, with predictions converging on the equivalent of an average occupant (Gilani et al., 2018). However, there is little risk in including diversity; the main cost is the required model training and minor additional effort in interpretation of the model and simulation output.

The current results highlight the need for larger numbers of coordinated behavioural studies including diversity considerations. Moreover, further concerted efforts are necessary to formulate and implement practically feasible inclusion of diversity information in models of occupants' control-oriented actions in buildings.

Acknowledgements

References

Gilani, S., O'Brien, W., Gunay, B. 2018. Simulating occupants' impact on building energy performance at different spatial scales, Building and Environment 132, 327-337.

Haldi, F., Calì, D., Andersen, R.K., Wesseling, M., Müller D. 2017. Modelling diversity in building occupant behaviour: a novel statistical approach, Journal of Building Performance Simulation, 10:5-6, 527-544.

Mahdavi, A., 2015. The inter-individual variance of the defining markers of occupancy patterns in office buildings: a case study, Proceedings of BS2015, 2243-2247.

O'Brien, W. and H. B. Gunay, 2015. Mitigating office performance uncertainty of occupant use of window blinds and lighting using robust design, Building Simulation 8:6, 621-636.

O'Brien, W., H.B. Gunay, F. Tahmasebi, and A. Mahdavi. 2017. A preliminary study of representing the inter-occupant diversity in occupant modelling, Journal of Building Performance Simulation, 10:5-6, 509-526.

Raynham, P., Unwin, J. and Guan, L. 2019. A new metric to predict perceived adequacy of illumination, Lighting Research & Technology. doi: 10.1177/1477153519828416.

Reinhart, C.F. 2004. Lightswitch-2002: a model for manual and automated control of electric lighting and blinds, Solar Energy 77(1) 15-28.

Schweiker M., Haldi F., Shukuya M., Robinson D., 2012. Verification of stochastic models of window opening behavior for residential buildings, Journal of Building Performance Simulation, 5(1), 55-74.

Schweiker, M. 2017. Understanding occupants' behaviour for energy efficiency in buildings, Current Sustainable Renewable Energy Reports (2017) 4:8-14.

Tahmasebi, F., Mahdavi, A. 2016. An inquiry into the reliability of window operation models in building performance simulation, Building and Environment 105 (2016), 343–357.

Tahmasebi, F., Mahdavi, A. 2017. The sensitivity of building performance simulation results to the choice of occupants' presence models: a case study. Journal of Building Performance Simulation, 10:5-6, 625-635.

Tahmasebi, F., Mahdavi, A. 2018. On the utility of occupants' behavioural diversity information for building performance simulation: an exploratory case study, Energy and Buildings (2018), DOI: 10.1016/j.enbuild.2018.07.042.

Tjur, T., 2009. Coefficients of Determination in Logistic Regression Models − A New Proposal: The Coefficient of Discrimination, The American Statistician, 63:4, 366-372, DOI: 10.1198/tast.2009.08210.