Ybl Journal of Built Environment, Volume 10. Issue 1. (2025)
10.12700/jbe-2025-0014

BEHAVIORAL ASPECTS OF REAL AND SIMULATED ENERGY CONSUMPTION IN ARTIFICIAL LIGHTING SYSTEMS OF OFFICE BUILDINGS

Johannes Weninger[1], Sascha Hammes[2]

[1]Bartenbach GmbH, research & development, Aldrans, Austria
johannes.weninger@bartenbach.com
[2]Unit of Energy Efficient Building, University of Innsbruck, Innsbruck, Austria
sascha.hammes@uibk.ac.at

Abstract: *Given that a considerable portion of the global energy demand is directly attributable to artificial lighting systems in buildings, an understanding of the determinants affecting their energy consumption is imperative to align current building planning practices with environmental policy objectives. Although it is already recognized that individual behaviours of building occupants exert a noteworthy impact on the energy performance of artificial lighting systems, a thorough quantification of this influence remains largely deficient. Based on a one-year, minute-by-minute monitoring of workplace occupancy, environmental conditions and energy consumption in an open-plan office, a synthetic dataset was generated to represent diverse usage scenarios. Leveraging advanced machine learning techniques, this dataset facilitated a comprehensive quantification of the multidimensional factors influencing both real and simulated energy consumption of the system. The results emphasize the critical need for an enhanced incorporation of behavioural aspects in the strategic planning of artificial lighting systems to optimize energy efficiency.*

Keywords: *artificial lighting, user behaviour, energy efficiency, energy performance gap, machine learning.*

1. INTRODUCTION

Despite the increasing maturity of advanced modelling techniques and energy-efficient technologies, buildings frequently fall short of achieving anticipated energy savings. Studies indicate that actual energy consumption in buildings can be up to three times greater than the estimates made during the design phase [1-4]. In addition, since the efficiency of control implementations is seldom evaluated post-commissioning, the attainment of environmental targets remains uncertain. Therefore, it is essential to understand and quantify the causes of these discrepancies to improve the accuracy of building performance planning and simulation, thereby ensuring the achievement of current energy policy objectives.

Considering that the building sector, despite numerous efforts aimed at enhancing energy efficiency, is still responsible for approximately 30% of greenhouse gas emissions and about 40% of global energy demand [5], with artificial lighting being a significant electrical consumer [6,7], the importance of reducing performance gaps [1,2,8,9] is increasingly critical [10].

To date, numerous research activities have been undertaken to identify and mitigate the causes of the energy performance gaps (EPG, e.g. [1,4,8,11–13]). In general, the discrepancy between actual energy consumption and simulated forecasts described by the EPG is strongly linked to concepts of optimal energy consumption, which must be seen in relation to the best possible function, structural realization and design of buildings and at the same time ensure that the requirements of the building users are met [4,9]. The requirements profile for estimating the energy consumption of buildings turns out to be correspondingly complex and the causes of deviations in real operation are often the result of an interaction of several influencing factors [13].

Potential influences on the operational deviations range, for example, from inadequate fine-tuning of control systems, suboptimal settings of technical systems when implementing new technologies for the first time [11], inaccuracies in measurement technology to execution errors and improper use of devices by building users [9,14]. In this context, it is important to note that building occupants do not always understand the energy impacts of system interventions. Rather, they make their decisions regarding interaction with control systems primarily on the basis of satisfying their immediate personal needs [15], which do not necessarily have to coincide with long-term strategic and energy-saving control concepts.

In addition, specification uncertainties in building modelling as well as inaccuracies in geometric representation and stored material properties prove to be major causes of planning-related estimation errors [3]. Finally, unrealistic assumptions and forecast errors in climate data [16] and occupancy models during the design phase [12,13] contribute to the manifestation of the EPG. In particular, the uncertainty of occupancy behaviour is currently assumed to have a critical impact on the accuracy of performance predictions for energy demands of buildings made during the design phase [8,12,17].

The decisive role of user behaviour [17] for the occurring performance gap is based on the lack of detailed information during the planning phase about the organizational and socio-cultural factors that influence user behaviour during the usage phase of the building [1,14]. In this context, planning and simulation assumptions about occupancy behaviour are currently primarily based on empirically validated and standardized models that are formulated as generally as possible to ensure broad applicability [18]. However, the dynamics of occupancy behaviour at workplaces are significantly influenced by inter-individual factors such as work activities or the professional position in the company as well as social framework conditions that can vary considerably between organizations.

As a result, the energetic influences of occupancy profiles usually prove to be stochastic and therefore do not correspond to the static occupancy models [19], which are assumed to be generally valid. The potential failure of today's model assumptions becomes obvious, for example, in contexts with flexible social structures, such as flexitime regulations and home offices (cf. [20,21]). In addition, work-specific dynamics, such as the different proportions of meetings depending on the position (cf. [22]), represent an obvious application-relevant challenge that can only be inadequately covered by current assumptions.

The improvement of simulation-relevant model assumptions therefore represents a key requirement to avoid inefficient building operation and ensure the achievement of energy objectives through better estimation in planning processes. For this reason, some research efforts have already been carried out on this in the past.

1.1. *Related work*

According to the review of Cozza et al. [9] EPG reduction strategies can be divided into two groups: (1) measures to improve the building design phase, including improved planning and simulation methods, and (2) optimization of building systems in operation after commissioning by monitoring and deriving appropriate measures.

In relation to these so-called post-occupancy evaluations (POE), Fedoruk et al. [23] emphasise the relevance of (1) appropriate energy monitoring methods (e.g. via building management systems and/or meters), (2) building occupant satisfaction surveys, (3) a better understanding of the limits of the energy system and (4) feedback processes over the different phases of a building life cycle to identify causes of energy performance gaps and derive adequate measures. In this context, point (4) of the study results primarily from the fact that the obstacles in the construction project under consideration were of an institutional nature and that there was a lack of information availability and information procurement [23]. In general, BIM (Building Information Modelling) can support this as collaboration and information platform in the future [24]. Through POE, not only can a better understanding of the sources of the EPG be created and subsequently reduced, but the knowledge gained can also be fed back into the planning process in order to reduce performance gaps in advance [4].

In order to obtain a better picture of potential influencing factors as early as possible in the planning phase, behavioural approaches based on virtual reality (VR) technology in conjunction with digital building models and the Design-with-Intent (DwI) method can potentially be used [8]. For example, building planners can use user-oriented VR experiments to analyse behavioural patterns, such as human-environment interaction, in the building model in advance and check whether these correspond to the specified design conditions and then adapt design strategies. In this context, the authors coin the term pre-occupancy evaluation [8,12].

In addition, there are currently several approaches to modelling the interaction behaviour of building users with the artificial and daylight system. Discrete-time Markov processes based on predictors, whose statistical significance is selected via forward and backward selection [25], for example, offer the added value of representing individual behaviour on a statistical basis. Machine learning methods have also been increasingly used in recent years (for example [10]). The derivation of improved methods for modelling user behaviour and their implementation in simulation environments was and is also the subject of previous and current research activities within the framework of the Energy in Buildings and Communities Program of the International Energy Agency (IEA EBC, see [17,26,27]). However, to date there is still no comprehensive and suitable quantification of the multidimensional factors influencing energy demand and the resulting performance gap.

1.2. *Objective of the current study*

Understanding the extent and underlying causes of EPGs is essential to be able to forecast and understand energy consumption more reliably and, consequently, to contribute to increasing energy efficiency and thus to reducing CO2 [4,28] emissions. The literature states that the uncertainty of occupancy behaviour determines the failure or success of the statements made in the building planning phase regarding the subsequent energy performance of the system. Although previous studies have recognised the existence of this gap, the precise measurement of its extent remains a complex task due to various influencing factors. A breakdown of the characteristics that influence energy demand and their weighting can help to counteract performance deficits in a targeted manner.

The aim of this study is therefore to comprehensively quantify the energy performance gap by investigating the effects of individual occupancy behaviour at the workplace in relation to other relevant influencing factors such as daylight availability, time of year and time of day, and building-related conditions in an open-plan office.

As part of a post-occupancy evaluation, numerous user-specific and building-related data were collected in high resolution over a period of one year. The collected data was then used to generate a synthetic dataset that represents a variety of different space usage scenarios. For this purpose, energy consumption was systematically calculated for different user combinations in relation to the seating configuration of the office. The importance of the influencing factors in the derived dataset was then quantified using both the real data and simulation-based assumptions using a histogram-based gradient boosting regression tree and a subsequent SHAP (SHapley Additive exPlanations) analysis. The results illustrate the significant influence of individual behaviour in the open-plan office on both energy consumption and the energy performance gap. No previous study was found that uses a comparable methodology to quantify the EPG.

2. METHODOLOGY

2.1. *Study object and raw data collection*

The research and development building of Bartenbach GmbH in Aldrans, Austria, includes a 160 m² open-plan office, accommodating a maximum of 28 workstations. To ensure optimum operation and comfort, it is primarily operated for 18 people, who are distributed across nine workstation zones. Four zones for two people each are located on the north side along the skylight and five further workstation zones, each for two people as standard and a maximum of four people, along the south façade (see Figure 1 left).

Figure 1 Interior (left) and exterior view (right) of the Bartenbach R&D building in Aldrans, Austria. In the right part of the interior area, the skylights of the north façade can be seen, the exterior view shows the static daylight system on the south façade.

Both the daylight and artificial lighting systems in the office space have been optimized over several years. As a result, the daylight and artificial lighting systems in the study object can be controlled separately for each workplace zone in order to meet individual lighting preferences [29–33], prevent associated conflict situations (cf. [34]) and significantly reduce the overall energy consumption of the system (cf. [35]). The artificial lighting system features two colour temperatures, ranging from 5,000 K in the morning to 2,200 K in the evening, to promote the circadian rhythm of the users and is controlled in relation to occupancy by ceiling-mounted passive infrared sensors (PIR; Thermokon, RDI). The implemented switch-off delays used are adapted to an industrial standard of 15 minutes (cf. [36,37]) to prevent false system switch-offs. In addition, the necessary artificial light is reduced using table-mounted horizontal illuminance sensors (Thermokon, LDF 1000A) based on the amount of daylight available. In this respect, a normative standard of 500 lx [38] is assumed as the target value.

The office is characterized by a high proportion of glazing on the southern façade, which ensures a high level of daylight integration. On an annual average, horizontal illuminance levels of over 500 lx occur at workplaces between 9:00 a.m. and 4:00 p.m., resulting in a daylight autonomy (DA) of 81.56% (Figure 2) and limiting the necessary use of artificial light primarily to the morning and evening hours (cf. [20,21,39]). To avoid glare and overheating, there are automatically controlled shading systems mounted on the outside of the southern façade and on the inside of the northern skylights as well as an external static daylight system (see Figure 1 right), which has been adapted to the specific situation and geographical location of the object. The automated control logic of both artificial and daylight can be overridden at any time by users within a work zone by means of switches to ensure a high level of end-user acceptance [30].

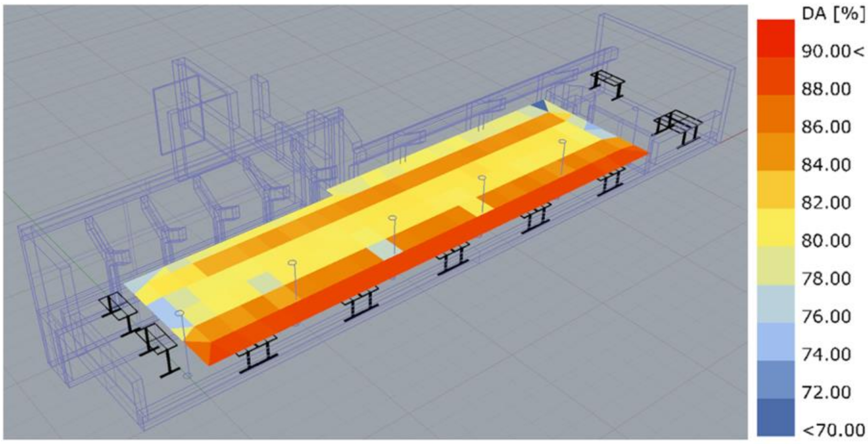

Figure 2 Daylight simulation of the study object, implemented with Radiance; simulation related to the normative minimum illuminance of 500 lx according to [38]; reference time: 8:00 to 18:00, daylight savings time not considered, calculated with necessary glare protection.

The occupancy structure in the building is proving to be highly dynamic. The core working hours in the object are Monday to Thursday 9:00 a.m. - 12:00 p.m. and 2:00 p.m. - 5:00 p.m. and Fridays from 9:00 a.m. - 12:00 p.m. Furthermore, as an organizational framework, there is both the possibility of home office and a flexitime arrangement between 6:00 a.m. - 08:00 p.m. To record the occupancy behaviour on an individual level, there are PIR sensors (NodOn, PIR 2 1 01) mounted under each individual work desk, whose detection ranges are limited to the respective workstation.

The building is controlled centrally via a programmable logic controller (PLC, BECKHOFF, CX5140-0141), which is also used for comprehensive logging of all sensor data and system statuses of the actuators. With over 100 sensors in the R&D building, a complete monitoring of the indoor and outdoor climate as well as the presence and absence of the occupants at the workstations is created in compliance with data protection aspects.

The data collection of the present study covered a period of exactly one year (February 2022 to January 2023). The continuous data from the illuminance sensors were collected every minute during this period. Attendance at the workstation was recorded on an individual level when the status changed. All data was collected pseudonymously in a machine-processable data format (.csv).

During the study period, 18 people were employed in the R&D department, 2 women and 16 men. Two people were employed part-time. To avoid conflicts with data protection regulations, a declaration of consent was obtained on a voluntary basis from all users at whose workplaces the occupancy profile was recorded and evaluated. The study was conducted outside of COVID-19 influences. Accordingly, there were no occupancy restrictions in this regard.

2.2. *Data synthesizing procedure*

The energy consumption of the artificial lighting system in the building depends on two key factors: (1) the zonal presence, which defines the basic switch-on times of the system, and (2) the amount of daylight available at the times of presence, which can vary between the individual zones at a specific time due to the structural situation of the building. Since the individual workstation zones can be used by two users at the same time, the switch-on times of the system are dependent on the coupled individual presence profiles of the two zone users (see also Figure 3). If the two profiles are as similar as possible, the energy is used much more efficiently than with different usage structures within a zone, as operating times in which only one user is present are minimized. In this regard, an earlier study [20] was already able to show the significant influence of both user pairing and seating organization on the overall energy consumption of the office building.

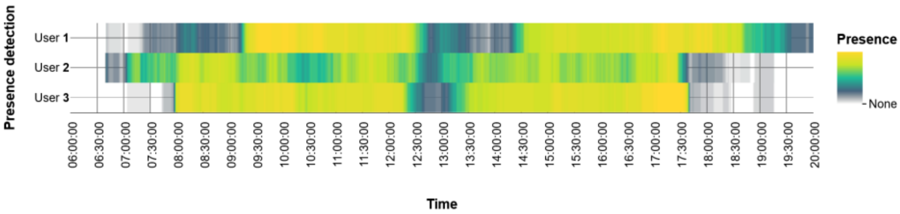

Figure 3 Exemplary representation of three individual occupancy profiles in Bartenbach's open-plan office to visualize the different workplace usage; mean values for the period from September 2, 2020 to November 3, 2020; color coding: yellow-green representation - high occupancy, blue to gray representation - low occupancy, transparent areas - periods without occupancy

In this regard, to exclude a potentially excessive influence of individual user pairings occurring in the real usage scenario on the results and to increase the general validity and significance of the analyses, the data was synthesized for the further analysis process. For this purpose, all possible combinations of individual users were considered, resulting in a total of 153 coupled user scenarios (number of possible combinations of 2 users each from a total of 18 users). The 153 individual user couplings were also considered for each of the existing 9 workplace zones in order to achieve a complete representation of all possible presence and daylight availability combinations. All features calculated for the further analysis were then derived for each of the resulting 1,377 zonal usage cases.

2.3. *Feature modelling*

As part of the feature modelling procedure, a comprehensive representation of all potentially relevant influencing parameters for the energy consumption of the artificial lighting system was carried out. In addition to general indicators for the time of day and year as well as for the location of the zone under consideration within the building structure, this included meaningful characteristics for the coupled presence in the workplace zone and the availability of daylight. Both all continuous input features and the associated energy consumption as a target criterion were calculated based on real data.

In addition, comparable features and target values were derived using dedicated building simulations or related assumptions and feature-related gaps were calculated. All calculations were carried out as hourly averages for each month to achieve a resolution of the real data that matches the current simulation assumptions. Data were normalized based on the working days in each month (holidays and weekends were excluded, but personal vacations were not taken into account). The time ranges for the calculation were limited to the period between 06:00 a.m. and 08:00 p.m. in order to do justice to the flexitime regulations in place in the building.

Regarding the continuous input features, the presence in the workplace zone was calculated as a percentage using the linked individual profiles of the individual users (logical OR connection of the individual profiles). As a simulation-relevant assumption, the presence profiles currently used primarily in simulations were used, which also indicate an hourly percentage value [40].

Regarding daylight availability, two input features were constructed, (1) the average horizontal illuminance at the workplace and (2) the period of time in which the normative requirement of 500 lx horizontal illuminance [38] cannot be sufficiently achieved by daylight. Both values could be derived for real operation from the available table-mounted sensors for each zone. The simulation-relevant comparison values were determined using a building simulation.

For the simulation, the study object was modelled in a high level of detail using Rhinoceros 3D (version: Rhino 6). Honeybee[+], a Grasshopper plug-in in Rhinoceros, with RADIANCE as the daylight simulation engine was used as simulation platform. Illuminance levels depending on the room position were generated in RADIANCE with EnergyPlus weather files (resolution: 1h, location: Innsbruck, Austria, approx. 5 km of air gap from Aldrans, Austria). To use bidirectional scattering distribution functions (BSDF; generated for the study object via WINDOW7), an annual daylight simulation was carried out according to the three-phase method (see [41]).

The energy consumption resulting from the presence and daylight availability for both real operation and the simulation assumptions was calculated using dedicated consumption measurements of the luminaires. A descriptive breakdown of the continuous features can also be found in Table 1.

The additionally constructed indicator features to consider both daily and seasonal effects as well as the dependence of the zonal location within the building were numerically coded to avoid categorical modelling processes in the analysis. Since daylight availability is subject to periodic fluctuations in relation to both seasons and times of day, the relevant input features were constructed as distances from the respective middle. For the month-related feature, the middle of the year (months of June and July) was assigned a distance value of 0. The values then increased with temporal distance from mid-year, reaching a maximum of 5 for January and December. The same coding procedure was used for the hour-related feature, which successively increased the feature value from the middle of the day (12:00 a.m. and 01:00 p.m.) to the edge of the day (06:00 a.m. and 07:00 p.m.) by 1 up to a maximum of 6.

With regard to the zonal location in the building, an indicator for the location of the zone in relation to west-east and south-north was modelled. The south-north coding was carried out on a binary level (0 for south, 1 for north zones). The west-east coding was carried out in 5 levels, with the westernmost zones being coded with 0 and the feature value being increased by 1 with increasing distance to the east up to a maximum of 4. All 9 existing zones in the building could thus be clearly described in their location by a value of the resulting 10 combination options (there is no workplace zone at one combination option due to the entrance situation into the building).

Feature	***Real situation***	***Simulation***	***Feature Gap***
Occupancy *[min]*	21.03 ± 13.91	25.71 ± 18.06	4.69 ± 15.61
Illuminance *[lx]*	597.86 ± 560.40	569.90 ± 399.36	-27.97 ± 310.57
Time < 500 lx *[min]*	31.88 ± 21.75	36.15 ± 18.48	4.27 ± 16.09
Energy demand *[Wh]*	8.05 ± 10.68	7.23 ± 9.48	-0.82 ± 7.69

Table 1 Descriptive listing of the continuous features calculated as part of the modelling procedure; values are stated as respective means and standard deviations

In total, 16 features were derived (4 continuous values, each in relation to the real situation, simulation assumption and the occurring feature gap, plus 4 indicator features), which were calculated for all user combinations and workplace zoning. Taking into account the monthly and hourly resolution, this resulted in a 16 x 231,336 matrix (approx. 4 million dedicated individual values) which was fed into the analysis process.

2.4. *Data analysis*

Regression analyses were carried out to analyse the influence of the individual input features on the resulting energy consumption. A total of 4 different models were calculated: (1) modelling of the real consumption based on the real feature values to estimate the influence in real operation, (2) modelling of the simulated consumption based on the simulation assumptions to estimate the influence in relation to the associated building simulation, (3) modelling of the real consumption based on the simulation assumptions to estimate the accuracy of the simulation on real consumption, and (4) modelling of the deviating consumption based on the feature gaps to determine the influence on the occurring energy performance gap.

Each model was implemented using a histogram-based gradient boosting regression tree (Python 3.9, scikit-learn 1.3.2), an ensemble method that avoids overfitting and was specially designed for large amounts of data. For modelling, the data set was split into a training and a test data set in a ratio of 80:20. To avoid confounding effects, the data split was stratified in relation to the user pairings. The overall fit of the derived models was determined using the determination coefficient R^2, which describes the proportion of the variation in the test data that can be explained by the derived regression model.

The created models were then interpreted using SHAP analyses (Python 3.9, shap 0.44.0), a game-theoretic method that explains machine learning predictions by combining optimal credit allocation with local interpretability based on Shapley values. On the one hand, the individual feature importances were derived, which describe an impression of the average explanatory power of the individual feature on the target variable. An additional interpretation of the influence of the individual features was achieved using summary plots. A more detailed description of SHAP analyses can also be found in [42]. Additional statistical analyses were carried out using JASP (version 0.18.3.0).

3. RESULTS

3.1. *Explanatory value of real features on real energy consumption*

With an R^2 of 0.979, the modelling of energy consumption based on the calculated input features for real operation proved to be extremely accurate. A more detailed analysis of the contributions of the individual features in the model (see Figure 4) showed that the occupancy time at the workplaces had the greatest influence on the resulting energy consumption, with the influence being positively correlated, i.e. longer average occupancy times also resulted in an increase in energy consumption. Other features of similar importance were the south-north location of the zone in the building (while workplaces located to the south contributed to an increase in energy consumption, northern zones resulted in a reduction) and the length of time with insufficient lighting at the workplace (positive correlation, i.e. increases in the duration resulted in an increase in energy consumption).

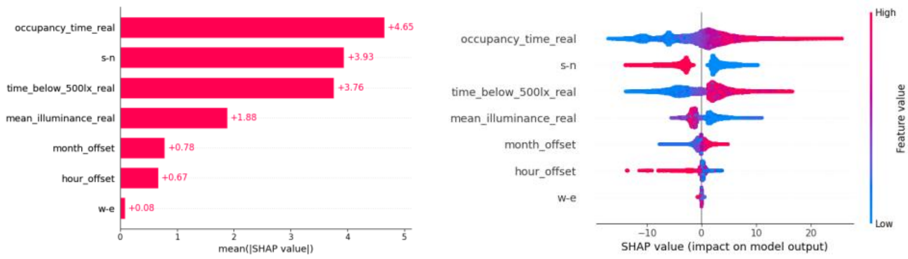

Figure 4 SHAP analysis for the explanatory value of real features on real energy consumption; left: derived feature importances, right: summary plot

Of lesser importance was the average illuminance at the workplace, which showed a negative correlation as expected, with low average illuminance increasing energy consumption. The monthly and hourly influencing factors proved to be negligible, similar to the west-north location of the zone in the building.

3.2. *Explanatory value of simulated features on simulated energy consumption*

The modelling of the simulated energy consumption based on the simulation assumptions also proved to be extremely accurate with an R^2 of 0.986. A more detailed analysis of the contributions of the individual features in the model (see Figure 5) showed that the two most important input features were just as pronounced as in the real model. The occupancy time at the workplaces proved to have the greatest influence on the resulting energy consumption and correlated positively, i.e. longer average occupancy times also resulted in an increase in energy consumption. The south-north location of the zone in the building also proved to be of comparable importance in the simulation model, with the contributions of the individual locations being similar to those in the modelling of the real situation (south-located workplaces contributed to an increase in energy consumption, northern zones to a reduction).

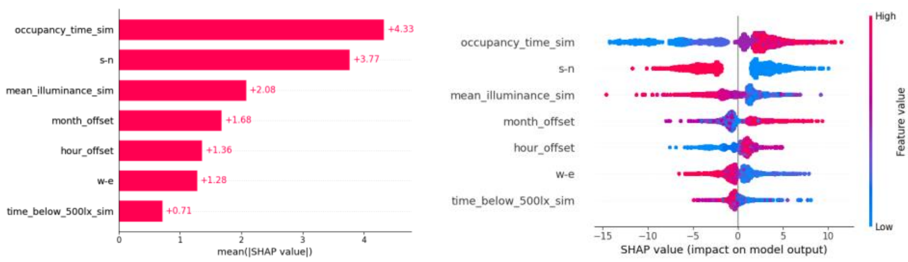

Figure 5 SHAP analysis for the explanatory value of simulated features on simulated energy consumption; left: derived feature importances, right: summary plot

With regard to the remaining input features, a very similar structure was also found compared to the modelling of the real situation. The gradations of importance in the derived model of the features average illuminance at the workplace and monthly and hourly offset were basically the same, although rated much higher. In contrast, the period of time with insufficient illuminance at the workplace, which was assumed to be very important in the real model, was rated significantly lower and as having the least influence on the simulated energy consumption. Due to this deviation at a higher level, a more detailed analysis of the derived contributions of the individual features to the model result showed a significant difference between the models for the real situation and the simulation (Chi-Square Goodness of Fit Test, X2(6, N = 100) = 1457.25, p < .001).

3.3. *Explanatory value of simulated features on real energy consumption*

The modelling of the explanatory value of the simulation assumptions with regard to the real energy consumption showed a significant reduction in terms of accuracy with an R^2 of 0.715. A more detailed analysis of the contributions of the individual features in the model (see Figure 6) showed that the two most important input features of the model of the real situation were assigned the least explanatory power. Both the occupancy time at the workstations and the time periods in which the standard requirement of 500 lx was not met appear in the model with only very little importance. The ranking of the remaining features, however, remained unchanged, even if their influence was rated much higher.

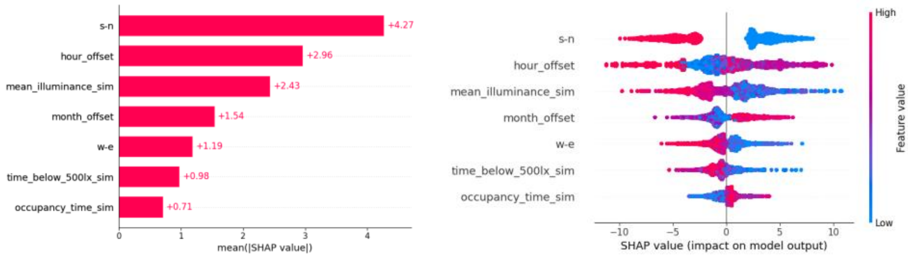

Figure 6 SHAP analysis for the explanatory value of simulated features on real energy consumption; left: derived feature importances, right: summary plot

The correlation of the individual features remained fundamentally unchanged compared to the previous models. Zones facing south also contributed to an increase in energy consumption, as do low average illuminance levels, high duration of occupancy and time periods in which the illuminance target value was not met. However, due to the strongly different evaluations of the importance of the individual features, the results of the explanatory values of real input features and simulation assumptions for estimating the real energy consumption turned out to be significantly different (Chi-Square Goodness of Fit Test, X2(6, N = 100) = 2939.90, p < .001).

3.4. *Explanatory value of the feature gaps on the energy performance gap*

The final modelling of the energy performance gap based on the individual feature gaps showed an R^2 of 0.952, which indicates that the deviations between the simulation and the real situation can be explained almost completely. A more detailed analysis of the contributions of the individual features (see Figure 7) shows that the differing consumption is largely due to differing occupancy assumptions. The differing occupancy duration was assigned a feature importance that was more than five times higher than that of the input feature, which was ranked as the second most important.

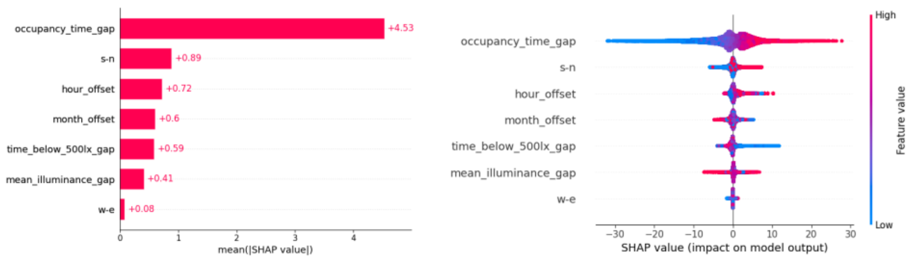

Figure 7 SHAP analysis for the explanatory value of the individual feature gaps on the energy performance gap; left: derived feature importances, right: summary plot

Interestingly, a closer look at the individual influence of the occupancy duration revealed a clear trend. Overestimated occupancy at the workplace proved to be clearly responsible for increasing the resulting EPG, while underestimated durations significantly reduced the EPG. All other individual feature gaps can be neglected with a high degree of probability due to the comparatively low explanatory value in the model.

4. DISCUSSION

The actual energy consumption always turns out to be the result of a causal relationship that results from different influencing factors and their control-related implementation. In the context of energy consumption derived by simulation, a similar causal relationship between the influencing factors is considered to be given on the basis of the mathematical calculation. Both models derived in this regard had a determination coefficient of more than 0.95 and accordingly an extremely high degree of explanatory accuracy of the target values. It can therefore be assumed in principle that the chosen regression analysis approach is very well suited to the analysis of consumption-related influencing factors and that underlying causal connections can be mapped. Histogram-based gradient boosting regression trees can therefore also be suitable for extended analyses of other buildings in the future and thus provide a solid basis for better comparable results.

The individual influencing factors in the two models of the real situation and the simulation approximation were generally given comparable explanatory values. The associated correlations also prove to be comparable. It can therefore be assumed that the causal relationships of the real situation were generally correctly assumed in the simulation environment. However, the significantly lower explanatory value of the assumed occupancy times for the real consumption shows that the deviations within the individual influencing factor causes significant changes in the results achieved. The intrinsic consistency that arises within the simulation-related model due to the mathematical relationship is therefore no longer justifiable in comparison to the real situation.

This result is reinforced by the fact that the energy performance gap that occurs is almost exclusively due to incorrect assumptions within the occupancy behaviour, which demonstrates the strong dependence of the simulation results on individual factors. Even though significant improvements have been made to simulation methods in recent years that allow a better estimation of the building-related key performance indicators (e.g. [41]), individual incorrect assumptions in other areas are proving to be decisive, meaning that the progress achieved cannot be transferred in a targeted manner. Further development of all relevant influencing factors is therefore of great importance in order to achieve environmental policy objectives.

4.1. *Limitations of the current study*

Even though the present study attempted to use a synthetic dataset to strengthen the general significance of the results and increase their validity, it still proves to be a case study. Personal influences on energy consumption are fundamentally tied to individual behaviour. Other building users, different usage scenarios or even a different organizational use of the building can therefore lead to different results. Even though the present study identified incorrect assumptions about individual behaviour as a key influencing factor for different energy consumption in real situations and simulations, it must be assumed that further case studies with different uses are required to make a general statement.

It must also be noted that the evaluated building is characterized by an above-average availability of daylight. Accordingly, large parts of the day do not require the use of artificial light in view of the normative requirements, which leads to a higher variability of energy consumption at the edges of the day. Since there is generally a higher fluctuation in occupancy times during these periods (e.g. due to different start and end times for daily work), it is potentially possible that the influence of occupancy behaviour is overestimated in the results. However, whether this overestimation really exists and in what dimension it manifests itself would require comparative studies, which are not currently available.

Finally, it must be noted that there are additional influences, particularly in relation to occupancy times, which make it difficult to generalize the results. In principle, the analytically derived performance gaps can be traced back to deviations from the presence models for the simulation, which are usually resolved hourly. Due to their hourly resolution, the models have only a very limited ability to correctly depict dynamic processes in user behaviour. The control strategy included in the building used a follow-up time of 15 minutes for absence detection using PIR sensors, based on the current industrial standards [36,37]. In principle, however, it is possible to adjust the follow-up time more closely to real presence patterns, which can also have a positive effect on energy consumption (cf. [39]). However, an adjustment in this regard would further increase the existing discrepancies between real conditions and simulation-relevant assumptions and accordingly potentially contribute to increasing existing gaps. A complete quantification of the mutual influence of user-specific behaviour and control-related implementations was not the subject of this study and should be examined more closely in future research in order to achieve a complete picture of potential influencing factors.

5. CONCLUSION

This study examines the influence of different factors on energy consumption in relation to both a real situation and an associated simulation model. Regression analyses were used to determine, among other things, the explanatory value of the individual simulation assumptions in relation to the real energy consumption and to derive specific insights into the causes of the resulting EPG.

Not only did user behaviour prove to be largely responsible for energy consumption in the real context, but the simulation-relevant and generalized assumptions related to it also caused the related deviations in the simulation results, primarily due to the lack of an opportunity to adequately depict individual behaviour. An improvement in the currently valid user model for building simulations is therefore absolutely necessary. In the future, these should not only be reduced as much as possible in terms of their temporal resolution in order to better depict basic dynamics, but above all they should also be implemented at an individual level.

The present results serve to create a better understanding of EPG in the lighting sector and to quantify the influence of the relevant factors. Since the object of investigation presented has some special features, such as a high level of daylight autonomy and a high level of occupancy dynamics, it is not easy to transfer them to other office applications. In order to be able to transfer the findings to the application in the long term and to increase the general validity of the statements, further research is necessary.

ACKNOWLEDGMENTS

This work is the result of a long-term cooperation between the University of Innsbruck and Bartenbach GmbH. The authors received no financial support from any public or private institution in developing the content. The authors declare no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the authors upon request.

REFERENCES

[1] **LIANG, J., QIU, Y., HU, M.,** Mind the energy performance gap: Evidence from green commercial buildings. *Resources, Conservation and Recycling, Volume 141 (2019), 364-377.*

[2] **GALVIN, R.,** Making the 'rebound effect' more useful for performance evaluation of thermal retrofits of existing homes: Defining the 'energy savings deficit' and the 'energy performance gap'. *Energy and buildings, Volume 69 (2014), 515-524.*

[3] **CALÌ, D., OSTERHAGE, T., STREBLOW, R., MÜLLER, D.,** Energy performance gap in refurbished German dwellings: Lesson learned from a field test. *Energy and buildings, Volume 127 (2016), 1146-1158.*

[4] **ZOU, P. X., ALAM, M.,** Closing the building energy performance gap through component level analysis and stakeholder collaborations. *Energy and buildings, Volume 224 (2020), 110276.*

[5] https://wedocs.unep.org/20.500.11822/32152

[6] **DUBOIS, M. C., BLOMSTERBERG, Å.,** Energy saving potential and strategies for electric lighting in future North European, low energy office buildings: A literature review. *Energy and buildings, Volume 43 (2011), No 10, 2572-2582.*

[7] **MADIAS, E. N. D., KONTAXIS, P. A., TOPALIS, F. V.,** Application of multi-objective genetic algorithms to interior lighting optimization. *Energy and Buildings, Volume 125 (2016), 66-74.*

[8] **NIU, S., PAN, W., ZHAO, Y.,** A virtual reality supported approach to occupancy engagement in building energy design for closing the energy performance gap. *Procedia engineering, Volume 118 (2015), 573-580.*

[9] **COZZA, S., CHAMBERS, J., BRAMBILLA, A., PATEL, M. K.,** In search of optimal consumption: A review of causes and solutions to the Energy Performance Gap in residential buildings. *Energy and Buildings, Volume 249 (2021), 111253.*

[10] **YILMAZ, D., TANYER, A. M., TOKER, İ. D.,** A data-driven energy performance gap prediction model using machine learning. *Renewable and Sustainable Energy Reviews, Volume 181 (2023), 113318.*

[11] **ZOU, P. X., WAGLE, D., ALAM, M.,** Strategies for minimizing building energy performance gaps between the design intend and the reality. *Energy and Buildings, Volume 191 (2019), 31-41.*

[12] **NIU, S., PAN, W., ZHAO, Y.,** A virtual reality integrated design approach to improving occupancy information integrity for closing the building energy performance gap. *Sustainable cities and society, Volume 27 (2016), 275-286.*

[13] **DE WILDE, P.,** The gap between predicted and measured energy performance of buildings: A framework for investigation. *Automation in construction, Volume 41 (2014), 40-49.*

[14] **MENEZES, A. C., CRIPPS, A., BOUCHLAGHEM, D., BUSWELL, R.,** Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap. *Applied energy, Volume 97 (2012), 355-364.*

[15] **BARTHELMES, V. M., BECCHIO, C., CORGNATI, S. P.,** Occupant behavior lifestyles in a residential nearly zero energy building: Effect on energy use and thermal comfort. *Science and Technology for the Built Environment, Volume 22 (2016), No 7, 960-975.*

[16] **ERBA, S., CAUSONE, F., ARMANI, R.,** The effect of weather datasets on building energy simulation outputs. *Energy Procedia, 134 (2017), 545-554.*

[17] **YOSHINO, H., HONG, T., NORD, N.,** IEA EBC annex 53: Total energy use in buildings—Analysis and evaluation methods. *Energy and Buildings, Volume 152 (2017), 124-136.*

[18] **WANG, C., YAN, D., REN, X.,** Modeling individual's light switching behavior to understand lighting energy use of office building. *Energy Procedia, Volume 88 (2016), 781-787.*

[19] **ZHOU, X., YAN, D., HONG, T., REN, X.,** Data analysis and stochastic modeling of lighting energy use in large office buildings in China. *Energy and Buildings, Volume 86 (2015), 275-287.*

[20] **Hammes, S., Weninger, J., Pfluger, R., Pohl, W.,** Take the right seat: the influence of occupancy schemes on performance indicators of lighting in open plan offices. *Energies, Volume 15 (2022), No 9, 3378.*

[21] **Hammes, S., Hauer, M., Geisler-Moroder, D., Weninger, J., Pfluger, R., Pohl, W.,** The impact of occupancy patterns on artificial light energy demand-simulation and post-occupancy-evaluation. In *Building Simulation 2021, Volume 17 (2021), 3536-3543.* IBPSA.

[22] **Panko, R. R., Kinney, S. T.,** Meeting profiles: Size, duration, and location. In *Proceedings of the Twenty-Eighth Annual Hawaii International Conference on System Sciences, Volume 4 (1995), 1002-1011.* IEEE.

[23] **Fedoruk, L. E., Cole, R. J., Robinson, J. B., Cayuela, A.,** Learning from failure: understanding the anticipated–achieved building energy performance gap. *Building Research & Information, Volume 43 (2015), No 6, 750-763.*

[24] **Villa, V., Naticchia, B., Bruno, G., Aliev, K., Piantanida, P., Antonelli, D.,** Iot open-source architecture for the maintenance of building facilities. *Applied Sciences, Volume 11 (2021), No 12, 5374.*

[25] **Haldi, F., Calì, D., Andersen, R. K., Wesseling, M., Müller, D.,** Modelling diversity in building occupant behaviour: a novel statistical approach. *Journal of Building Performance Simulation, Volume 10 (2017), No 5-6, 527-544.*

[26] https://www.annex66.org/

[27] https://www.iea-ebc.org/projects/project?AnnexID=79

[28] **Van Dronkelaar, C., Dowson, M., Burman, E., Spataru, C., Mumovic, D.,** A review of the energy performance gap and its underlying causes in non-domestic buildings. *Frontiers in Mechanical Engineering, Volume 1 (2016), 17.*

[29] **Boyce, P. R., Eklund, N. H., Simpson, S. N.,** Individual lighting control: task performance, mood, and illuminance. *Journal of the illuminating Engineering Society, Volume 29 (2000), No 1, 131-142.*

[30] **Despenic, M., Chraibi, S., Lashina, T., Rosemann, A.,** Lighting preference profiles of users in an open office environment. *Building and Environment, Volume 116 (2017), 89-107.*

[31] **Newsham, G. R., Aries, M. B., Mancini, S., Faye, G.,** Individual control of electric lighting in a daylit space. *Lighting Research & Technology, Volume 40 (2008), No 1, 25-41.*

[32] **Veitch, J. A., Newsham, G. R.,** Preferred luminous conditions in open-plan offices: research and practice recommendations. *International Journal of Lighting Research and Technology, Volume 32 (2000), No 4, 199-212.*

[33] **Juslén, H. T., Wouters, M. C. H. M., Tenner, A. D.,** Preferred task-lighting levels in an industrial work area without daylight. *Lighting Research & Technology, Volume 37 (2005), No 3, 219-231.*

[34] **Chraibi, S., Lashina, T., Shrubsole, P., Aries, M., Van Loenen, E., Rosemann, A.,** Satisfying light conditions: A field study on perception of consensus light in Dutch open office environments. *Building and Environment, Volume 105 (2016)., 116-127.*

[35] **Hammes, S., Weninger, J., Canazei, M., Pfluger, R., Pohl, W.,** Die Bedeutung von Nutzerzentrierung in automatisierten Beleuchtungssystemen. *Bauphysik, Volume 42 (2020), No 5, 209-217.*

[36] **Nagy, Z., Yong, F. Y., Schlueter, A.,** Occupant centered lighting control: A user study on balancing comfort, acceptance, and energy consumption. *Energy and buildings, Volume 126 (2016), 310-322.*

[37] **Garg, V., Bansal, N. K.,** Smart occupancy sensors to reduce energy consumption. *Energy and Buildings, Volume 32 (2000), No 1, 81-87.*

[38] **European Committee for Standardization,** European Standard EN 12464-1: Light and lighting - Lighting of work places - Part 1: Indoor work places.

[39] **Hammes, S., Weninger, J., Geisler-Moroder, D., Pfluger, R., Pohl, W.,** Reduzierung des Kunstlichteinsatzes durch Anpassung der Nachlaufzeit an individuelle Anwesenheitsmuster. *Bauphysik, Volume 43 (2021), No 1, 50-64.*

[40] **Schweizerischer Ingenieur- und Architektenverein,** Standard-Nutzungsbedingungen für die Energie- und Gebäudetechnik

[41] **McNeil, A., Lee, E. S.,** A validation of the Radiance three-phase simulation method for modelling annual daylight performance of optically complex fenestration systems. *Journal of Building Performance Simulation, Volume 6 (2013), No 1, 24-37.*

[42] https://shap.readthedocs.io/en/latest/