User’s behaviours in non-residential mixed-mode buildings: a case study in a tropical climate

M Payet¹ ², M David³, P Lauret¹ and F Garde¹

¹ Laboratory PIMENT (Physics and Mathematical Engineering for Building, Energy and Environment), University of La Réunion, Campus universitaire de Terre Sainte, 97410 Saint Pierre, La Réunion
² LEU Reunion (Urban Ecology Laboratory), 97410 Saint Pierre, La Réunion, France

Abstract. In this paper, we present the results obtained by modelling the users’ behaviours in a mixed mode office building in a tropical climate, more exactly in La Réunion. Few specific research studies on comfort in tropical climates have been published, and there is little feedback on the users' behaviour in these buildings.

In order to improve users' assumptions in the design phase, users' actions on ceiling fans and windows have been measured and analysed. These data have then been modelled by machine learning methods, according to hygrothermal comfort and occupancy. The F1 scores eventually obtained for predicting fan use by random forests, decision trees and Bayesian networks are 99%, 98% and 95% respectively. For windows use, the F1 scores obtained are 92%, 91% and 70%, which demonstrates the ability of the models tested to predict the users’ behaviours.

1. Introduction

To reduce energy needs in the intertropical areas is an important challenge, due to their continuous development. Currently, air conditioning is one of the most energy-consuming items. However, the mild all-year temperatures and the potential of the trade winds for natural ventilation are both advantages of the tropical climate, which offer the possibility of employing passive systems to achieve comfort in buildings.

Therefore, thanks to these climatic conditions, mixed mode buildings are developing. They allow users to choose between passive systems or active systems for the hottest days. However, when indoor temperatures and humidity levels rise, it is more difficult to maintain comfortable conditions using only passive solutions, and users are more likely to add active cooling systems, and to adapt their behaviours.

In the publications by Liu, Y et al. and Yu, Z. et al., occupants' behaviour is mentioned as one of the elements that affect the energy consumption of buildings, as well as climatic conditions, the building envelope or the efficiency of the systems[1] [2]. For example, users will open or close windows to regulate the indoor environment by adjusting their perceived ambient temperature. This possibility is part of the building controls, as explained by Turner, W. et al. [3] and Raja, et al. [4]. Users will both adapt to the building and behave to adapt the building to their individual needs [5].

The common controls available to users of non-residential mixed mode buildings in tropical climates are:

- Natural cross ventilation, made possible by numerous windows on opposite facades. In fact, effective window size and operability are more important factors than weather protection in hot climates, according to the study conducted by Roetzel, A. et al. [6].
- Active systems for hot days, such as variable speed ceiling fans and air conditioning.
In order to reduce the gaps that may exist in the design phase of a building, the modelling of users' behaviours is becoming more and more common in the literature. However, according to the state of the art of Carlucci, et al., tropical climates are poorly represented, and the behaviours in buildings using natural ventilation only represent 14% of the articles studied, while air-conditioned buildings are in top position [7]. This has led us to ask the following question: Will users of mixed mode buildings in tropical climates use the building as planned in the design phase? Several models have been used in the literature to model these behaviours, but few in a tropical context.

After having monitored a case study in La Réunion, we have modelled the behaviours regarding the opening of windows and the use of ceiling fans, to predict future actions.

2. Research methodology

2.1. Presentation of the case study

The 310m² office building we have investigated is part of a residential building, located in the south of La Réunion. It is composed of 2 individual offices, a meeting room, a server room and open-space areas, where architects, urban planners and building engineers work.

It is a mixed mode bioclimatic building with large solar protections and lots of vegetation, where users can operate the many manually adjustable windows, called louveres, to naturally ventilate. They can also use ceiling fans, which have a variable rotation speed. There is no air-conditioning system, except for the server room and the meeting room. From an architectural point of view, this building is very representative of the current building trend.

2.2. Field measurements

We have monitored the building for approximately 1 year, from November 2019 to the beginning of 2021. The actual users' behaviours have been followed. The hourly database includes the specific electrical consumption of the ceiling fans and sockets, the indoor air conditions in terms of temperature and relative humidity, the state of opening or closing of the louveres thanks to magnetic contacts, and the outdoor conditions thanks to a weather station.

The magnetic contacts give binary signals. 0 = louver is open, 1 = louver is closed. The electrical consumptions are in fact the average power demand over 10 minutes, which are measured in Watts. An overview of the installed equipment is provided in Table 1.

| Instrument Denomination | Number |
|-------------------------|--------|
| Air temperature and humidity sensor | Testo 174H |
| Magnetic contact | NODON SDO-2-1-05 |
| Multichannel concentrator + submeters modules | OMEGAWATT |
| Data acquisition system | JEEDOM Pro v2 JEEDOM_PR |

Figure 1 Vertical solar protections in stretched canvas material protecting the South façade of the offices.

Figure 2 Green patio with large solar protections on the North façade.
2.3. Occupancy
We have not directly measured the number of people present in the building. Indeed, the solution of CO2 sensors in naturally ventilated rooms is not appropriate, and the installation of movement detectors is difficult in open spaces. However, we know that occupancy is a precondition for action to take place. We therefore estimated the number of people present from the power consumption of the sockets, which includes the consumption of each user's computer. Since there are only fixed workstations, when a user is present at his or her workplace, the computer consumes energy.

2.4. Modeling implementation
To make predictions and to classify the users' behaviours from our data, we have chosen to compare the deterministic machine learning models of decision tree (DT) and random forests (RF), and the probabilistic model of Bayesian networks (BN).

The models have been built from a training database, which allows to learn the relationships between the variables. These relationships have been applied to a testing database to predict the values of the variable we want to model and the bootstrap method has been applied. This involves the creation of new statistical 'databases', by random sampling with replacement, from the original database. The training base is the same for each model.

Decision trees and random forests have been performed using mixed qualitative and quantitative variables (classification models), while the Bayesian network model uses only qualitative data. The unsupervised clustering algorithm K-means can determine qualitative classes from quantitative data. The objective of this type of algorithm is to divide n points of a database into k distinct groups, called clusters, according to the similarities existing between these points. A point can only be in one cluster at a time. It is associated with a particular cluster when its distance from the centroid of that cluster is the smallest. The centroid represents the centre of the cluster. See [8] for more details on K-means.

The occupancy qualitative variable have taken the values Absent, Medium or High. The qualitative variables of louvers opening and ceiling fan power have been classified as follows: Low, Medium or High. Table 2 summarises the type and possible values for each variable.

| Table 2 Summary of models inputs |
|----------------------------------|
| Variable                        | Type       | Value                                                                 |
| --------------------------------|------------|-----------------------------------------------------------------------|
| Air temperature [°C]            | Quantitative | [19.95 : 31.51]                                                      |
| Humidity [%]                    | Quantitative | [46.5 : 93.14]                                                        |
| Level of comfort                | Qualitative | Discomfort, Comfort at 0 m/s, Comfort at 0.5 m/s, Comfort at 1 m/s   |
| Occupancy                       | Qualitative | Absent, Medium, High                                                  |
|                                 |             | (Absent for Occupancy ≤ 2 and High for Occupancy ≥ 19)               |
| Level of use for louvers        | Qualitative | Low, Medium, High                                                     |
|                                 |             | (Low for louvers opened ≤ 7 and High for louvers opened ≥ 20)        |
| Level of use for Ceiling fans   | Qualitative | Low, Medium, High                                                     |
|                                 |             | (Low for ceiling fans power ≤ 152 W and High for ceiling fans power ≥ 425 W) |

For the decision trees and the random forest, thermal comfort is directly expressed in terms of indoor air temperature and indoor relative humidity. For the Bayesian networks, the qualitative form of the thermal comfort level has been defined using the comfort levels described on the GIVONI diagram, a widely accepted tool in tropical climate design offices, which identifies comfort zones based on air temperature and relative humidity. A comfort level of 0.5 m/s means that an air speed of 0.5 m/s is required to be in a comfortable situation. It should be noted that an air speed of 1 m/s can be achieved with ceiling fans. For the comfort zone at 0 m/s, no wind is needed to be in a comfortable situation [9].

2.5. Algorithms and performance evaluation
Decision trees are used for regression or classification problems. From the database, their objective is to provide decisions, i.e. the possible values that the variable to be predicted will take, according to interconnected "If, then" rules. These rules describe the values taken by the input variables. In a tree,
final decisions are taken at the final leaves and rules are obtained at each node. The algorithm tests the input data at each node, and partitions it into groups of data so that they are as similar as possible and explain the values of the variable to predict. When we apply the rules to a new dataset to predict a value of the output variable, the results obtained at the different tests will determine the path to follow. They offer an explicit and visual prediction.

A random forest works on the same principle as decision trees but is actually a combination of several trees. It will fit decision trees to various subsamples of the database and take the average to improve the accuracy of predictions. The strengths and weaknesses of each tree are aggregated. The results obtained by random forest are not easily readable.

Finally, the last model we have tried is the Bayesian network, built on probabilities. Like trees, they provide easily interpretable results, in the form of a directed graph where the cause and effect relationships between variables are explicit. In this type of graph, nodes are variables and are associated with a probability function that takes a particular set of values. The arcs connect the nodes and show the cause/effect relationship. Bayesian networks are machines for calculating probabilities for one variable, knowing specific observations on the other connected variables [10]. The structures of the models used for this work are summarised in Figure 3.

![Diagram of models](image)

**Figure 3** Implementation of the models

To evaluate the performance of the models, the F1 score has been calculated. This is a way of combining model precision and recall into a single number. An F1 score reaches its best value at 1 and its worst at 0. Its formula is as follows:

$$F1 \text{ score} = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

In our study, the variables are multi-class. The weighted average F1 score of each class has been calculated and then, they have been combined into a single number, the overall F1 score of the classifier. As our classes are not balanced, we have weighted the F1 score of each class by the number of samples in that class. The F1 score is a good method for comparing different models in a synthetic way, although it is sometimes contested for giving equal weight to recall and precision.

3. Results and discussion

As illustrated in Figure 4, the F1 scores eventually obtained for predicting fan use by random forests, decision trees and Bayesian networks are 99%, 98% and 95% respectively. For windows use, the F1 scores obtained are 92%, 91% and 70%, which demonstrates the ability of the models tested to predict the users' behaviours. The modelling of fans is more efficient than the modelling of windows opening.
The random forests has obtained the best scores. All the models have obtained close scores on the low class and on the weighted average. The high and medium classes are less well represented. These are also the classes with the least amount of data, which could explain why the algorithms have not acquired enough knowledge.

![Figure 4: Results for the prediction of the level of use for ceiling fans](image)

Even if they are satisfactory, bayesian networks have shown the lowest results. However, we have improved them by extracting predictive knowledge from random forests to classify the variables, instead of using kmeans. It can therefore be useful to link models together instead of simply comparing them individually. The prediction of the level of use for fans has also increased with the incorporation of windows opening data as input in the models. These controls are therefore linked.

![Figure 5: Comparison between prediction by RF and current value of ceiling fans power](image)

The plots in Figure 5 represent the prediction of the fans’ usage level by the random forests in green, and their actual usage in blue. The prediction have properly taken into account the switch on and switch off times of the systems, however we notice some discrepancies on the maximum power demand, sometimes underestimated.
4. Conclusion
Data on hygrothermal comfort and on the users' behaviours towards ceiling fans and windows have been collected in a mixed mode office building in a tropical climate. From this database, we have modelled the users' actions, using decision tree, random forest and Bayesian network algorithms. Random forests have obtained the best prediction score, even though all models have produced satisfactory results, reflecting the relationship between occupancy and behaviours, as well as between comfort level and behaviours. Improving the assumptions made in the building design phase requires an understanding of how users acts. Buildings employing as much as possible passive solutions such as natural ventilation and ceiling fans to achieve users' comfort should be developed and could be extended to other climates, which are likely to become warmer in the coming years. A fuller paper on our modelling work is in progress.

5. Acknowledgments
We are grateful to the team of the GSCOP laboratory in Grenoble, especially Stéphane Ploix and Manar Amayri, for their experience sharing and their precious advices.

6. References
[1] Liu, Y., Yang, L., Zheng, W., Liu, T., Zhang, X., & Liu, J. 2018. A novel building energy efficiency evaluation index: Establishment of calculation model and application. Energy Conversion and Management 166, 522-533.
[2] Yu, Z., Fung, B. C., Haghighat, F., Yoshino, H., & Mrofšky, E. 2011. A systematic procedure to study the influence of occupant behavior on building energy consumption. Energy and buildings 43(6), 1409-1417.
[3] Turner, W., & Hong, T. 2013. A technical framework to describe occupant behavior for building energy simulations (No. LBNL-6671E). Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).
[4] Raja, I. A., Nicol, J. F., McCartney, K. J., & Humphreys, M. A. 2001. Thermal comfort: use of controls in naturally ventilated buildings. Energy and buildings 33(3), 235-244.
[5] Mahdavi, A. 2011, November. The human dimension of building performance simulation. In 12th International IBPSA Conference: Building Simulation, 14-16.
[6] Roetzler, A., Tsangrassoulis, A., Dietrich, U., & Busching, S. 2010. A review of occupant control on natural ventilation. Renewable and Sustainable Energy Reviews 14(3), 1001-1013.
[7] Carlucci, S., De Simone, M., Firth, S. K., Kjærgaard, M. B., Markovic, R., Rahaman, M. S., & van Treeck, C. 2020. Modeling occupant behavior in buildings. Building and Environment 174, 106768.
[8] Gao, X., & Malkawi, A. 2014. A new methodology for building energy performance benchmarking: An approach based on intelligent clustering algorithm. Energy and Buildings 84, 607-616.
[9] Givoni, B. 1992. Comfort, climate analysis and building design guidelines. Energy and buildings 18(1), 11-23.
[10] Naïm, P., Wuillemin, P. H., Leray, P., Pourret, O., & Becker, A. 1999. Réseaux bayésiens. Eyrolles, Paris, 3, 120.