A PROBABILISTIC MODEL TO PREDICT BUILDING OCCUPANTS' DIVERSITY TOWARDS THEIR INTERACTIONS WITH THE BUILDING ENVELOPE

Frédéric Haldi[1]

[1] Gartenmann Engineering SA, Avenue d'Ouchy 4, 1006 Lausanne, Switzerland

ABSTRACT

Based on observations conducted in an office building, we apply advanced statistical analysis methods, leading to the formulation of stochastic models for the prediction of buildings occupants' actions on window openings and shading devices.

The statistical analysis method – based on generalised linear mixed models – enables a correct treatment of the longitudinal nature of the datasets, an accurate estimation of the calibration parameters' uncertainty and a detailed study of the differences between the occupants surveyed. This analysis results in the formulation of stochastic models for the prediction of occupants' interactions with the key elements of the building envelope, which include explicit in-built probabilistic terms to account for occupants' diversity if required.

Furthermore, we show that the properties of these probabilistic terms can be used to infer a statistical distribution of the model's calibration parameters, which comprehensively represent the diversity of observed behaviours between building occupants, and can be applied to simulate their behavioural properties.

INTRODUCTION

The issue of the ability of building simulation programs to correctly represent reality is regularly discussed, in particular regarding the incomplete representation of occupants' presence and interactions with environmental controls, which undermines the accuracy of predictions. Based on field survey monitoring of heating or electricity consumption, these factors are known to be of great importance, as the performance of identical building is generally estimated to vary by a factor of two, a spread that is entirely induced by the differences in behaviour between occupants.

In order to account for the impact of occupants' interactions with the components of the building envelope which determine building energy flows (Figure 1), deterministic and stochastic models were developed to predict interactions with window openings (reviewed by Roetzel et al. (2010) and Fabi et al. (2012)) and shading devices (reviewed by O'Brien et al. (2013)). The data supporting their development, the proposed modelling approaches, the robustness of their validation and their scope of application are variable, which advocates for the strengthening of published work into a robust formulation, adaptable to a sufficiently wide range of situations.

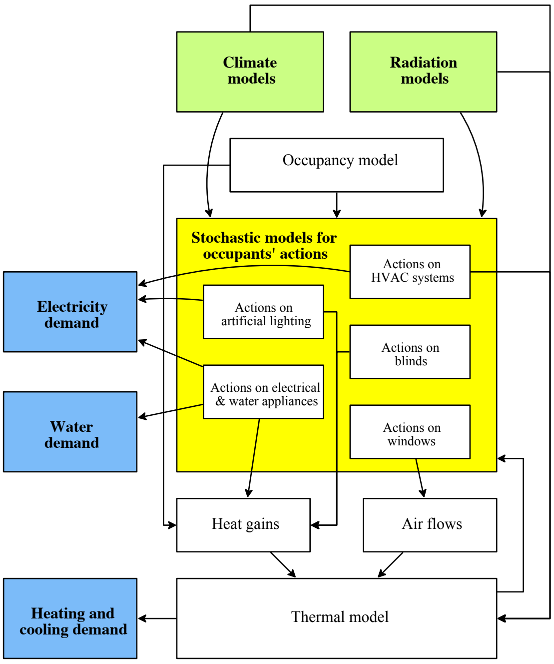

Figure 1. An overview of the principal occupant-induced processes in building energy simulation

On the other hand, the validation of these models is still to be completed, particularly regarding their ability to predict behaviour in buildings that are different from their original calibration basis. Validation studies of Haldi et al. (2010) and Schweiker et al. (2012) suggest that similar patterns occur in residential and office environments of temperate climates, but that they strongly differ in humid climates and within buildings equipped with cooling devices.

Preliminary attempts to implement stochastic models representing interactions with windows (Haldi and Robinson, 2009) and shading devices (Haldi and Robinson, 2010) into building simulation tools have led to encouraging results which are in coherence with experimental evidence. Based on an implementation of this approach, Haldi and Robinson (2011) have shown that in the context of a single zone model, the simulated behaviour of different behavioural patterns lead to heating and cooling demands that vary by a factor of two. In a different simulated context of an office building with sealed windows, Parys et al. (2011, 2012) have shown that the demand at the level of a whole building varies by a magnitude close to 20%.

These results are encouraging, but the quality of the predictions of such models depends heavily on the quality of calibration parameters, particularly with respect to the representation of the diversity between occupants. A correct methodology is of central importance, particularly towards a correct prediction of extreme values of energy use, or to assess the robustness of a given building design towards particular behavioural patterns.

With regard to this important issue, previously published models have either:

a) ignored the representation of behavioural diversity and only presented aggregated results (Rijal et al. (2007), Herkel et al. (2008)), or
b) proposed simplified classification mechanisms based on fixed thresholds (eg. active versus passive occupants, Reinhart et al. (2002)), or
c) presented occupant-specific results from their surveys, but without the integration of this diversity into a modelling framework (Haldi and Robinson (2009, 2010), Yun and Steemers (2008), Mahdavi et al. (2008)).

In order to strengthen existing models, we propose to develop an original approach which integrates the issue of behavioural diversity into the existing modelling approaches of Haldi and Robinson (2009,2010) and extends their results for predicting actions on windows and shading devices.

EXPERIMENTAL DATA

The data that support the development of these models were collected from the Solar Energy and Building Physics Laboratory (LESO-PB) experimental building, located in the suburb of Lausanne, Switzerland (46°31'17"N, 6°34'02"E, alt. 396 m).

In every office, occupants have the possibility to tilt or open up to any angle each of the two windows (height 90 cm, width 70 cm) and to control two external blinds (width 350 cm): a lower blind potentially covering the totality of the vision window (height 100-185 cm) and an upper blind covering an anidolic system (height 210-270 cm). These blinds are controlled by switches (one to start and one to stop lowering/raising) allowing occupants to shade any desired fraction. Occupants may also close and tilt internal vertical slat blinds at the upper window to reduce glare whilst benefiting from direct solar gain during the heating season.

Six offices are each occupied by two persons, who can both individually access their own window, while eight offices accommodate single occupants also able to act on the two windows. It is safe to leave windows open (eg. for night ventilation) during periods of absence, except on the ground floor.

All 14 south-facing cellular offices of this building have been equipped with sensors whose real-time measurements were archived by a centralised EIB data acquisition system. For a period covering 19 December 2001 (1 January 2004 for blinds) to 8 September 2009 (with the exception of a few short interruptions caused by maintenance and technical reasons), measurements of local indoor and outdoor temperature and illuminance, occupancy, window openings and closings, blinds lowering and raising with their unshaded fractions, actions on electrical lighting were continuously recorded (Figure 6).

The reader is referred to Haldi and Robinson (2009, 2010) for a detailed description of the surveyed building and exploratory data analysis.

THE MATHEMATICAL MODEL

Following the approach previously developed by Haldi and Robinson (2009, 2010), occupants' interactions with windows and shading devices are modelled as a discrete-time Markov process (Figures 2 and 3). This dynamic method can account for the real adaptive processes of occupants, based on the relevant set of time-evolving physical predictors as well as special patterns such as occupancy events.

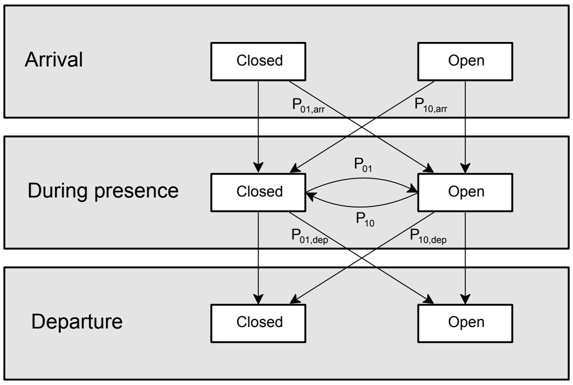

Figure 2. General scheme of the modelling approach for predicting actions on windows.

Formulation of action probabilities

In this approach action probabilities p are formulated as logistic models which include a set of n predictors $x_1, \ldots, x_n$:

$$\text{logit}(p) = \log(p/(1-p)) = \beta_0 + \Sigma_{k=1,\ldots,n}\, \beta_k\, x_k \qquad (1)$$

where β_k are the regression parameters (see the nomenclature for other definitions). The relevant predictors x_k (possibly including any environmental parameter such as temperature or occupancy status) are then selected by forward selection on the basis of their statistical significance.

In the case of windows, action probabilities refer to the opening and closing probabilities, depending on the current status of the window. These two probabilities are separately estimated in the situations of occupants' arrival, departure and during their presence (Figure 2).

The model for the prediction of actions on shading devices (Figure 3) defines lowering and raising probabilities, which refer to events that may both occur at any simulated time step, except when the shaded fraction is zero or one. A specific sub-model then determines the occupant chosen shaded fraction. The analysis shows that occupants' behaviour is specific on arrival, but that there is no significant difference between actions during presence and at departure.

Both models predict occupants' actions on 5-minute intervals, but the approach can be generalised to any other desired interval.

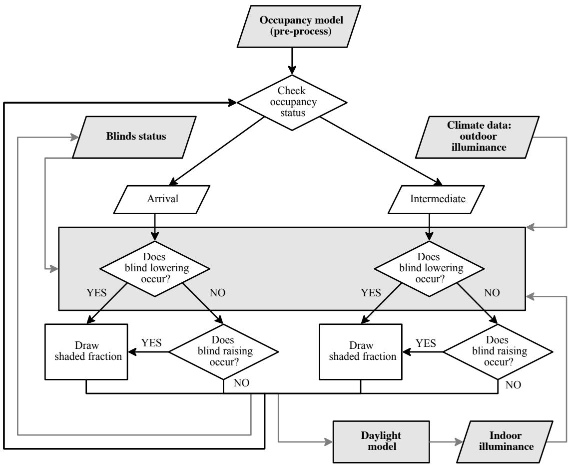

Figure 3. General scheme of the modelling approach for predicting actions on shading devices and the resulting shaded fraction.

Extension to generalised linear mixed models

An extension of the model for action probabilities in order to integrate the effect of individuals' diversity requires a particular treatment, as this latter is associated with the characteristics of individuals that are drawn at random from a population; unlike the design or environmental variables such as temperatures and occupancy status, which are definite values of experimental conditions. The former type of variable is treated as a random effect, while these latter (such as x_k in Equ. 1) are fixed effects. Considering together these factors, a mixed-effects logistic model is defined for action probabilities:

$$\text{logit}(p) = \beta_0 + b_0 + \Sigma_{k=1,\dots,n} (\beta_k x_k + b_k x_k) \qquad (2)$$

where b_k denotes a random variable representing the deviation from the population mean of the mean logit(p) for the ith individual. With this convention, the vector $\mathbf{b} = (b_0,\dots,b_n)$ is assumed to be distributed as a multivariate normal distribution $\mathcal{N}_{n+1}(\boldsymbol{\mu},\boldsymbol{\Sigma}^2)$, whose density function f is defined as:

$$f(b_0,\dots,b_n) = (1/(2\pi)^{k/2}|\boldsymbol{\Sigma}|^{1/2}) \cdot \exp\left(-(1/2) \cdot (\mathbf{b}-\boldsymbol{\mu})^T \cdot \boldsymbol{\Sigma}^{-1} \cdot (\mathbf{b}-\boldsymbol{\mu})\right), \qquad (3)$$

where $\boldsymbol{\mu}$ is the expectancy of the vector random variable associated with $\mathbf{b}$ ($\boldsymbol{\mu}$ is set by definition to 0 in the context of regression within Equ. 2), $\boldsymbol{\Sigma}$ the associated positive-definite covariance matrix and $|\boldsymbol{\Sigma}|$ its determinant.

With these definitions, generic models for the action probabilities are deduced by regression with Equ. (2), if the fitted values of the random effects $\mathbf{b}$ are ignored, but with an estimation of the standard errors of the regression parameters β_k that is coherent with the longitudinal nature of the surveyed data.

Furthermore, for each predictor x_k, the obtained distribution of each b_k measures the variability around the associated regression parameter β_k, induced by the behavioural differences among occupants.

RESULTS

Based on a dataset similar as Haldi and Robinson (2009, 2010), the results regarding action probabilities on windows and shading devices obtained by this methodology are presented in this section.

Estimated parameters and uncertainties

As expected, we observe that this methodology generally selects the same significant variables (with a few exceptions) with slightly different values of their associated regression parameters (Table 1) than those obtained by Haldi and Robinson (2009, 2010), where the distinction between fixed and random effects was ignored.

This difference is caused by a different treatment of the weight put to each surveyed occupant in the computation of the regression parameters for the fixed effects, which solves the problem of previously overweighing observations from occupants that were surveyed for a long duration.

An important change lies in the estimation of standard errors of the parameters for the fixed effects, which typically increase by a factor of three for most predictors (Table 1). The correct integration of the part of the variance in the data which is due to inter-occupant variability results in this larger uncertainty in the instrinsic effect of the retained predictors (Figure 4).

On the other hand, a new result produced by this mixed model approach is the covariance matrix $\boldsymbol{\Sigma}$ which comprehensively describes the correlations between all the random effects b_k which are linked with the predictors x_k. A comprehensive presentation of the elements of the matrices $\boldsymbol{\Sigma}$ is out of the scope of this article. However, the square root of their diagonal elements corresponds by definition to the standard deviation of the random effects retained in the models. Their values are also presented in Table 1.

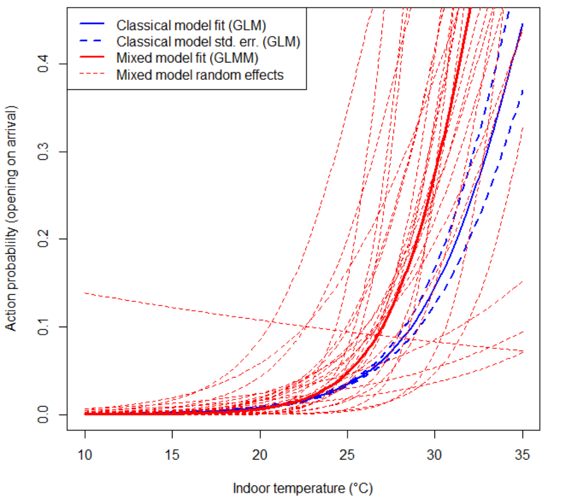

Figure 4. Fitted window opening probability on arrival (for the case where θ_{out} = 20°C, $f_{abs,prev}$ = 1, f_R = 0)

These latter provide an estimation of the magnitude of the deviations b_k from the 'average' behaviour expressed by β_k with respect to environmental stimuli x_k that are caused by the behavioural diversity among individual occupants. It is of particular interest to point out that these deviations are generally of a magnitude three to four times larger than the statistical errors on the fixed effects β_k. This fact points out how behavioural diversity with respect to the use of environmental controls is expected to heavily impact the associated energy flows through the building envelope.

Predicting individual behavioural patterns

The approach presented here is particularly constructive as it also produces further valuable results. Indeed, with this regression method the distinct sources of uncertainty due to statistical spread and the diversity among occupants are correctly identified and estimated using a detailed statistical treatment. However an important added value lies in the possibility to simulate individual behavioural patterns on a sound statistical basis. Indeed, by knowing the parameters β_k of the included fixed effects it is straightforward to generate any desired number of individual behavioural profiles as this information is entirely encapsulated within the statistical distribution of the random vector $\mathbf{b}$. Therefore, by knowing its covariance matrix $\mathbf{\Sigma}$, any behavioural profile based on a number of parameters n can be sampled from its multivariate normal distribution $\mathcal{N}_{n+1}(\mathbf{0},\mathbf{\Sigma}^2)$.

An example of this procedure is provided in Figure 3, where 1000 individual behavioural profiles are displayed for the case of opening actions on windows on occupant arrival, together with their statistical univariate (histograms) and bivariate (scatterplots) graphical representations.

DISCUSSION

Statistical models

The results obtained using generalised linear mixed models offer: an appropriate treatment of the longitudinal nature of the dataset; an accurate estimation of the calibration parameters' uncertainty; a detailed study of the differences between the occupants surveyed; and a specific quantification of the inter-individual variability with respect to interactions with windows and shading devices in buildings.

The underlying assumptions related to the use of generalised linear mixed models were verified (particularly the assumed normal distribution of random effects) without observing violations of this important requirement; although a database including a larger number of occupants would be desirable to thoroughly investigate this issue. This assumption of normality allows a coherent and accessible representation of the diversity between occupants which can be easily programmed for simulation purposes.

Behavioural diversity and its representation

This approach implicitly considers the issue of active and passive occupants within buildings. This distinction is closer here to that which is often observed in practice meaning that there is a continuity rather than a categorical separation within the observed behavioural patterns.

One strength of this approach is to treat behavioural diversity as a continuum, which is achieved by inferring a comprehensive statistical distribution of the parameters that describe individual interactions between environmental stimuli and actions in the building.

After integration within building simulation tools this approach facilitates the testing of the impact of a realistic set of behaviours on the energy demand of a building, and to identify which of these result in the particular outcomes under investigation eg. thresholds in energy consumption, overheating risk, moisture induced damage due to insufficient air renewal, etc. Such studies may also inform the added value of integrating automatic control systems within the building such as the control of shading devices or ventilation openings.

Statistical properties of diversity indicators

Based on Figure 7, it is possible to classify categories of behaviours. Occupants that interact only at high temperatures with windows tend to occupy an upper right position in the first chart of the second row of

Figure 7, as this is the domain of high characteristic temperatures $\theta_{in,50}$ (for which action probability reaches 0.5). These latter behaviours can also be simulated with the resulting regression parameters and statistical properties of the covariance matrix, as shown in Figure 5 for the indoor characteristic temperatures for opening actions at arrival.

In this latter case, the statistical properties of such derived indicators require particular care. As the characteristic temperature is computed by the ratio of two quantities (β_k + b_k) which are both normally distributed, the result theoretically converges to a Cauchy distribution; (in the case where these quantities are independent, approached in principle when the correlation is zero. This distribution is heavy-tailed implying that predicted extreme values may strongly diverge, which is not consistent with what is observed in practice. However, although the Cauchy distribution has undefined mean and variance, its median and quantiles can be defined, which allows a rigorous classification of the 'activity' of occupants towards controls based on statistical domains (eg. characteristic temperature situated between the median and the 3rd quartile).

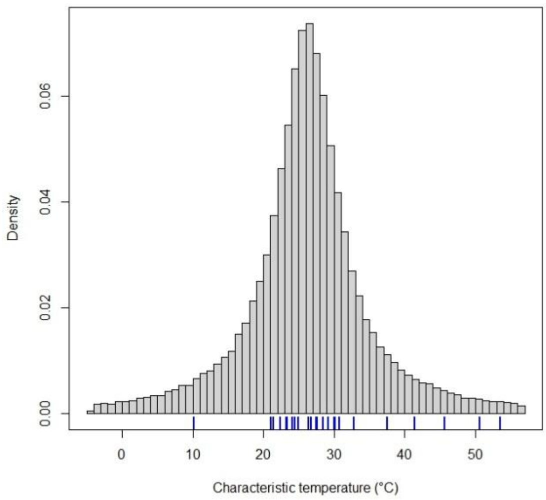

Figure 5. Observed (rugs) and simulated (histogram) indoor characteristic temperatures for opening actions at arrival (for the case where θ_{out} = 20°C, $f_{abs,prev}$ = 1, f_R = 0)

In the context of the simulation of individual profiles, this observation illustrates the importance of integrating the in-built correlations between the random effects, which are frequently significant (see for instance Fig. 6), and to perform simulations on the basis of a statistical distribution which is inferred on the basis of a sufficiently large number of observed occupants.

CONCLUSION

Main findings

This analysis results in the formulation of stochastic models for the prediction of occupants' interactions with the building envelope, which include explicit in-built probabilistic terms to account for occupants' diversity.

The method implemented allows to correctly extract from longitudinal databases the standard errors of calibration parameters related to environmental stimuli, as well as the effect and dispersion of individual occupants surveyed.

Limitations and further research

It would be of interest to strengthen the calibration basis of this model by including data from other field surveys or other climates. If relevant, specificities may also be accounted for by a specific random effect.

This inclusion of further data can also allow for the performing of detailed statistical tests on a larger database regarding the validity of the assumption of normality for the random effects. Such examinations are particularly needed for the correct treatment of extreme values which needs a wider calibration basis.

A promising application of this method probably lies in the analysis of risks in buildings, as it allows an estimation of the prevalence of behavioural patterns that can lead to problems in buildings. Once this model is implemented, a detailed assessment at the design phase whether buildings are robust to the influence of occupants' behaviour becomes – in principle – possible.

ACKNOWLEDGMENTS

The staff of the Solar Energy and Building Physics Laboratory who contributed to the installation and maintenance of the data acquisition sensors is gratefully acknowledged, particularly Antoine Guillemin, David Lindelöf and Laurent Deschamps.

NOMENCLATURE

θ_{in}	Indoor temperature (°C)
$\theta_{in,50}$	Indoor temperature at which action probability equals 0.5 (°C)
θ_{out}	Outdoor temperature (°C)
$\theta_{out,dm}$	Daily mean outdoor temperature (°C)
f_R	Rainfall (binary variable)
T_{pres}	Ongoing presence duration (min)
$f_{abs,prev}$	Preceding absence longer than 8 hours (binary)
$f_{abs,next}$	Following absence longer than 8 hours (binary)
f_{GF}	Window higher than ground floor (binary)
B	Unshaded fraction of a shading device
W	Window status (binary - 0: closed, 1:open)
E_{in}	Indoor workplane illuminance (lux)
$E_{gl,hor}$	Outdoor global horizontal illuminance (lux)
x_k	Fixed effect predictor in a regression model
β_k	Regression parameter for a fixed effect predictor x_k

b_k Esimated random effect associated to a predictor x_k

Σ Covariance matrix, whose $(i,j)^{th}$ element is the covariance between the $(i,j)^{th}$ elements (here b_i,b_j) of the vector **b** of random effects

REFERENCES

Fabi V., Andersen, R. V., Corgnati, S., Olesen, B. W., 2012. Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models, Building and Environment 58, 188-198.

Haldi F., Robinson, D. 2009. Interactions with window openings by office occupants, Building and Environment, 44(12), 2378-2395.

Haldi F., Robinson, D. 2010. Adaptive actions on shading devices in response to local visual stimuli, Journal of Building Performance Simulation, 3(2), 135-153.

Haldi F., Robinson, D., Pröglhöf C., Mahdavi, A. 2010. A partial double blind evaluation of a comprehensive window opening model, Proc. BauSIM 2010 Conference, 331-336, Vienna.

Haldi F., Robinson, D. 2011, The impact of occupants' behaviour on building energy demand, Journal of Building Performance Simulation, 4(4), 323-338.

Herkel, S., Knapp, U., Pfafferott, J., 2008. Towards a model of user behaviour regarding the manual control of windows in office buildings. Building and Environment, 43(4):588–600.

Mahdavi, A., et al., 2008. Occupants' operation of lighting and shading systems in office buildings. Journal of Building Performance Simulation, 1 (1), 57–65.

O'Brien, W., Kapsis, K., Athienitis, A. K., 2013. Manually-operated window shade patterns in office buildings: A critical review, Building and Environment 60, 319-338.

Parys, W., Saelens, D., Hens, H., 2011. Coupling of dynamic building simulation with stochastic modelling of occupant behaviour in offices – a review-based integrated methodology, Journal of Building Performance Simulation, 4(4), 339-358

Parys, W., Breesch, H., Hens, H., Saelens, D., 2012. Feasibility assessment of passive cooling for office buildings in a temperate climate through uncertainty analysis, Building and Environment 56, 95-107.

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

Roetzel, 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.

Rijal, H. B., et al., 2007. Using results from field surveys to predict the effect of open windows on thermal comfort and energy use in buildings. Energy and Buildings, 39(7):823–836.

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

Yun, G. Y., Steemers, K., 2008. Time-dependent occupant behaviour models of window control in summer. Building and Environment, 43(9):1471– 1482.

Table 1
Overview of the regression parameters for action probabilities on windows (upper part) and shading devices (lower part) with standard deviation of the associated random effects.
*Significance: no sign: $p<0.001$, ** $0.001<p<0.01$, * $0.01<p<0.05$*

CONTROL & OCCUPANCY	PREDICTOR	OPENING / LOWERING PROBABILITY		CLOSING / RAISING PROBABILITY	
		Estimate fixed effect	**Std. dev. random effect**	**Estimate fixed effect**	**Std. dev. random effect**
Windows Arrival	Intercept	-16.0 ± 1.2	4.7	6.0 ± 1.2	4.1
	θ_{in}	0.407 ± 0.054	0.21	-0.361 ± 0.053	0.151
	θ_{out}	0.058 ± 0.012	0.045	-0.058 ± 0.022**	0.087
	$f_{abs,prev}$	1.65 ± 0.16	0.66		
	f_R	-0.43 ± 0.13**	0.43		
Windows Intermediate	Intercept	-14.7 ± 1.3	0.002	-0.2 ± 1.1	5.1
	θ_{in}	0.371 ± 0.056	0.065	-0.100 ± 0.045	0.194
	θ_{out}	0.038 ± 0.012**	0.070	-0.093 ± 0.018	0.088
	T_{pres}		0.0043	-1.41 ± 0.18 E-03	0.60 E-03
	f_R	-0.34 ± 0.15*	0.0225		
Windows Departure	Intercept	-15.1 ± 1.9	6.35	-5.66 ± 0.70	1.03
	θ_{in}	0.317 ± 0.091	0.308	0.078 ± 0.032*	0.043
	$\theta_{out,dm}$	0.096 ± 0.027	0.094	-0.079 ± 0.029	0.120
	$f_{abs,next}$			1.88 ± 0.16	0.57
	f_{GF}	0.76 ± 0.46*	1.64	-1.20 ± 0.31	0.83
Lower blinds Arrival	Intercept	-8.50 ± 0.59	2.39	-0.016 ± 0.452	2.09
	E_{in}	1.263 ± 0.076 10^{-3}	0.30 10^{-3}	-1.90 ± 0.34 10^{-3}	1.51 10^{-3}
	B_L	2.39 ± 0.51	1.98	-4.67 ± 0.40	1.77
Lower blinds Intermediate & departure	Intercept	-8.25 ± 0.10	0.008	-2.48 ± 0.39	1.83
	E_{in}	4.88 ± 0.70 10^{-3}**	0.33 10^{-3}	-0.91 ± 0.16 10^{-3}	0.74 10^{-3}
	B_L	1.43 ± 0.10	1.17	-3.88 ± 0.36	1.66
Upper blinds Arrival	Intercept	-8.11 ± 0.34	1.36	-1.35 ± 0.28	1.25
	E_{in}	1.307 ± 0.078 10^{-3}	0.31 10^{-3}	-2.74 ± 0.42 10^{-3}	1.80 10^{-3}
	B_U	2.11 ± 0.16	0.42	-4.38 ± 0.23	0.78
Upper blinds Intermediate & departure	Intercept	-8.92 ± 0.30	1.30	-3.59 ± 0.27	1.26
	E_{in}	1.071 ± 0.062 10^{-3}	0.27 10^{-3}	-1.30 ± 0.25 10^{-3}	1.18 10^{-3}
	B_U	1.84 ± 0.16	0.64	-3.53 ± 0.21	0.90
Any blind Arrival	Intercept	-7.36 ± 0.37	1.55	-0.25 ± 0.35	1.61
	E_{in}	1.224 ± 0.076 10^{-3}	0.32 10^{-3}	-2.09 ± 0.33 10^{-3}	1.45 10^{-3}
	B_{tot}	1.90 ± 0.21	0.75	-4.87 ± 0.33	1.40
Any blind Intermediate & departure	Intercept	-8.51 ± 0.47	2.16	-3.08 ± 0.33	1.55
	E_{in}	1.074 ± 0.077 10^{-3}	0.35 10^{-3}	-0.87 ± 0.17 10^{-3}	0.79 10^{-3}
	B_{tot}	1.74 ± 0.33	1.49	-3.30 ± 0.31	1.40

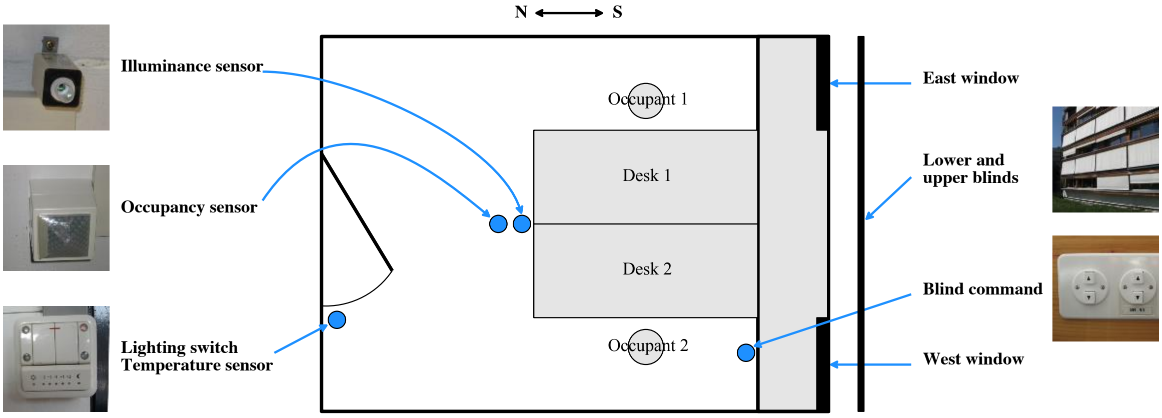

Figure 6. General scheme of the devices used for the field survey

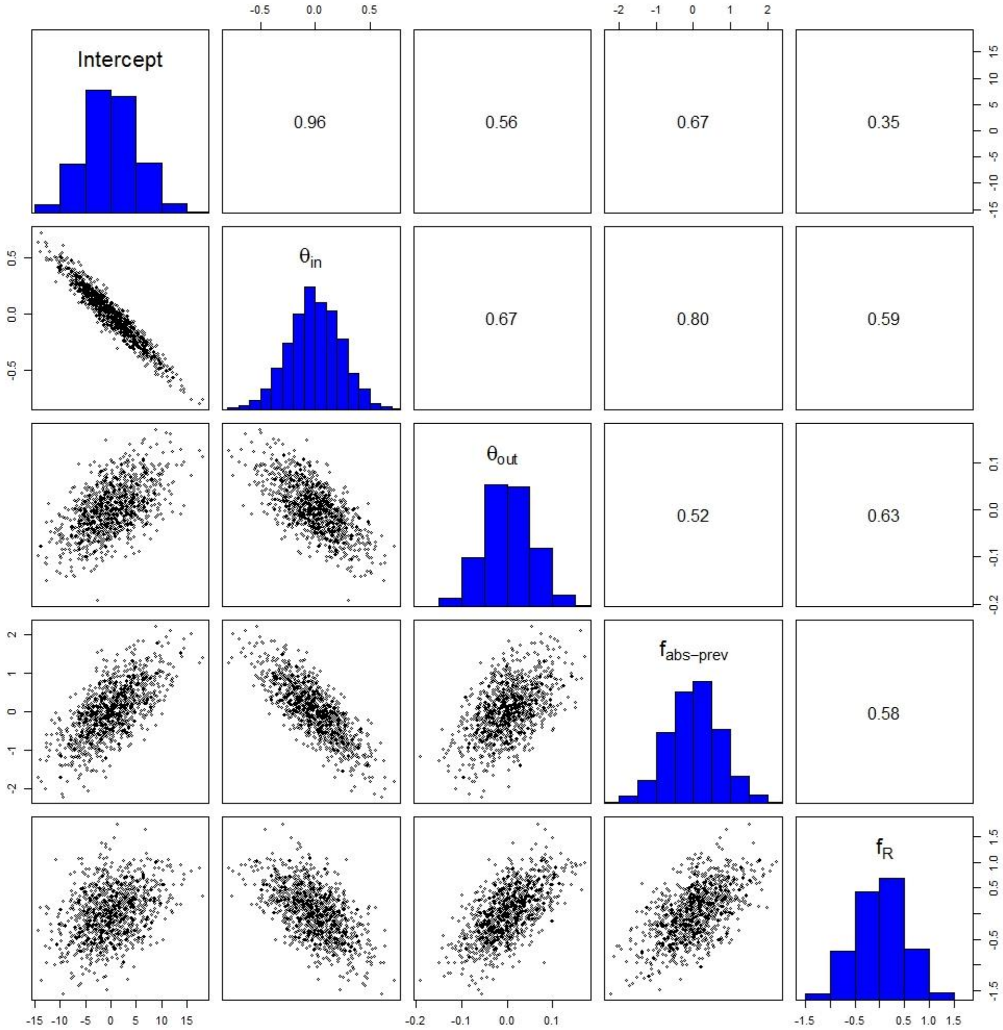

*Figure 7. Multivariate distribution of the vector of random effects **b** for window opening probability on arrival, superposed with a thousand simulated random effects, presented as bivariate plots with frequency histograms and correlations*