Mind Reading Building Operation Behaviour

Neves, L. O.[1], Hopes, A. P.[2], Chung, W. J.[2], Natarajan, S.[2]

[1]*School of Civil Engineering, Architecture and Urban Design, University of Campinas, Brazil*

[2]*Department of Architecture and Civil Engineering, University of Bath, United Kingdom*

Abstract

Modern buildings are operationally classified as fully naturally ventilated (free running, FR), fully mechanically ventilated and conditioned (air-conditioned, AC) or intermediary (mixed mode, MM). Each strategy implies (i) the selection of different internal comfort models and controls, e.g. adaptive thermal comfort + operable windows in FR and static thermal comfort (Predicted Mean Vote - PMV) + non-operable windows in AC, and (ii) different energy consumption and carbon emissions with FR being the least (zero) and AC being the most. In each instance, the choice of strategy is dictated by the designer with little choice for the occupants. Here we ask the question: what operational mode would occupants select, given complete freedom? We examine this question using real concurrent data from three offices of similar size and use. In each, occupants have complete freedom in adopting FR, AC or a series of intervening MM strategies. We use monitored data on window and air-conditioning operation combined with internal and external thermal conditions to create validated computer models for each office. We discover the true operational mode of each office by comparing the computer model against 23 different scenarios using Dynamic Time Warping for binary (window open/ close, AC on/ off) and Euclidean distance for continuous operative temperature time series. Strikingly, results demonstrate that while each office used divergent strategies across the seasons, the indoor conditions were very similar and attainable through NV alone. This suggests that while a purely NV strategy is likely to deliver indoor thermal comfort, understanding occupant motivation and educating them on the impact of AC operation is needed to minimise energy use.

Keywords: occupant behaviour; mixed-mode ventilation; office building; thermal comfort.

1. Introduction

Occupant behaviour is acknowledged to have a major, often negative, impact on the energy consumption of office buildings (Gunay et al., 2013). The primary hypothesis is that occupants' actions are often at odds with the design intent of the spaces or actions taken by facility management to minimise energy consumption. Indeed, it has been suggested that the difference in overall consumption between energy profligate and energy conscious occupants could be as much as 330% (Andersen et al. 2007). The behaviour of occupants in these instances can be driven by both conscious and unconscious actions. Conscious actions might be driven by a desire for personal control over their working environment in an effort to maximise comfort. A good example of this is the opening of windows when the heating or cooling system is operational, thus rejecting or admitting heat to/from the external environment, resulting in an increase in energy consumption (Sorgato et al., 2016). Unconscious actions such as leaving systems and appliances on when they are not needed also result in unnecessary energy consumption (Kingsland and Hunter, 2011). Some studies have hence suggested that automated systems can reduce up to 80% wasted energy in comparison to spontaneous occupant control, since they can optimize operation and reduce the occupants' energy impact (Gunay et al., 2013; Chen et al., 2019).

Naturally, the level of automation of the building's systems directly influences the users' degree of freedom to manipulate the occupied space. Indeed, it is well known that the *perception* of control is positively correlated with actual thermal acceptability of the indoor environment (Hoes et al., 2009; Vellei et al., 2016). This suggests that there is a direct conflict within conditioned environments[1] between occupant desire on the one hand, and low energy consumption on the other. Our hypothesis is that this conflict is driven by the imposition of an operational regime on occupants by designers and facility managers, possibly through a lack of understanding of the diversity in user behaviour and preference. Classically, this emerges from the definition

[1] We define conditioned environments here as those provisioned with and actively using a heating or cooling system.

of an "average occupant" towards whom the design is orientated. The pitfalls of "designing to the average" are well known in other fields such as aviation and automotive design. For example, in the 1940s, US Airforce cockpits designed to the dimensions of an average pilot were shown to not fit a single pilot from a sample of 4,063 (Rose, 2015). This led to significant redesign that provided greater flexibility and adaptability to individual pilot needs. In building design, the emergence of the adaptive thermal comfort standard (McCartney and Nicol, 2002) was driven in large part by the idea that occupants should be able to exercise greater control over their buildings through, for example, being able to open and close windows – a simple adaptive opportunity that is often removed in fully conditioned buildings to prevent heat or coolth loss (Rijal et al., 2009).

In climates that feature external temperatures that are either below or above the adaptive thermal comfort threshold (17.4 °C to 31.5 °C over a mean outdoor temperature range of 10 °C to 33.5 °C, in the ASHRAE 55-2017 standard, considering metabolic rates ranging from 1.0 to 1.3 met, clothing range from 0.5 to 1.0 clo and air speed up to 0.3 m/s), some heating or cooling may be inevitable. Hence, depending on climate and other requirements, designers have a choice of three operational modes:

- Fully naturally ventilated (free running, FR) → zero operational energy/ carbon.
- Fully mechanically ventilated and conditioned (air-conditioned, AC) → high operational energy/ carbon.
- Intermediary (mixed mode, MM), where the building usually operates in either FR or AC mode at different times or different parts of the building, to provide comfort while minimising energy use (Omer, 2009).

Each mode of operation implies the selection of different internal comfort models and controls. For example, FR buildings will normally be designed to the adaptive thermal comfort standard and include operable windows whereas AC buildings will be designed to the "fixed" Predicted Mean Vote (PMV) standard (ASHRAE, 2017), i.e. full mechanical ventilation and, in many instances, non-operable windows. These choices impose or create completely different indoor thermal environments. For example, in typical office environments, the PMV standard translates to an indoor operative temperature of 23 ±2 °C, whereas the adaptive model allows a range of temperatures between ~18 °C to ~31 °C, depending on seasonal weather. Hence, understanding occupant choice can dramatically affect the design of the building. Yet, little is known about what mode of operation occupants might choose, if they had choice *a priori*.

Hence, our research question is: what operational mode would occupants select, given complete freedom? While FR or AC driven design essentially represents a binary choice for designers, this is not true for MM driven design where three key modes of operation can be defined. These are: (i) zoned, where different parts of a building operate in either FR or AC; (ii) changeover, where the mode of operation is dependent on time of day or year; and (iii) concurrent, where both modes can operate simultaneously. Appendix A summarizes previous studies that have investigated modes of operation of MM office buildings focused on occupant behaviour analysis and prediction, used as a general reference for defining a series of theoretical modes of operation in this paper (see Section 2.3.4).

A key advantage of the concurrent mode is that it provides maximum flexibility (compared to the other two modes) since it could, theoretically, be set up to operate in any of the other modes of operation or indeed fully FR or AC. In fact, concurrent is the most common mode in MM office buildings, especially in the developing world (Neves et al., 2017). Hence, concurrent mode buildings provide an ideal platform to investigate our research question.

2. Methods

Our overall approach is summarised in Figure 1, covering the following steps:

Step 1. We undertake field monitoring of three office rooms in two concurrent mixed-mode office buildings located in the city of São Paulo, Brazil.

Step 2. We infer occupant behaviour through binary observations of window (open/ close) and air-conditioning (on/ off) operation and relate these to the co-incident indoor and outdoor thermal environment.

Step 3. These are fed into calibrated thermal models of each office to produce an indoor operative temperature time-series for a 'spontaneously controlled' base-case.

Step 4. We define a range of theoretical modes of MM operation, driven by our literature review (Appendix A), capped by FR and AC at either end. These are fed into the calibrated models to create time series for 23 different 'imagined' operational scenarios.
Step 5. Finally, we use a series of mathematical techniques to discover which of the created imagined temperature time series best matches reality.

We discuss each of these steps below.

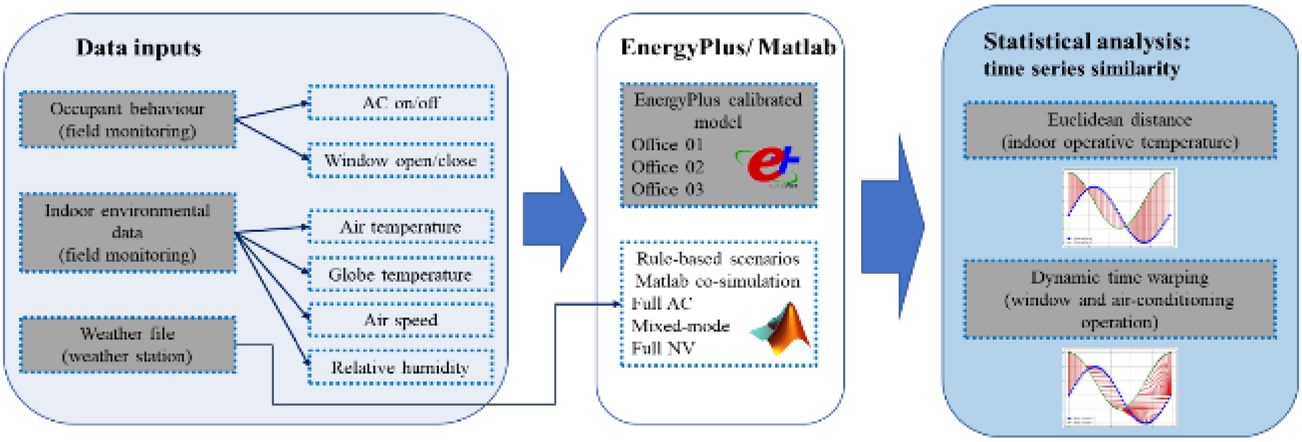
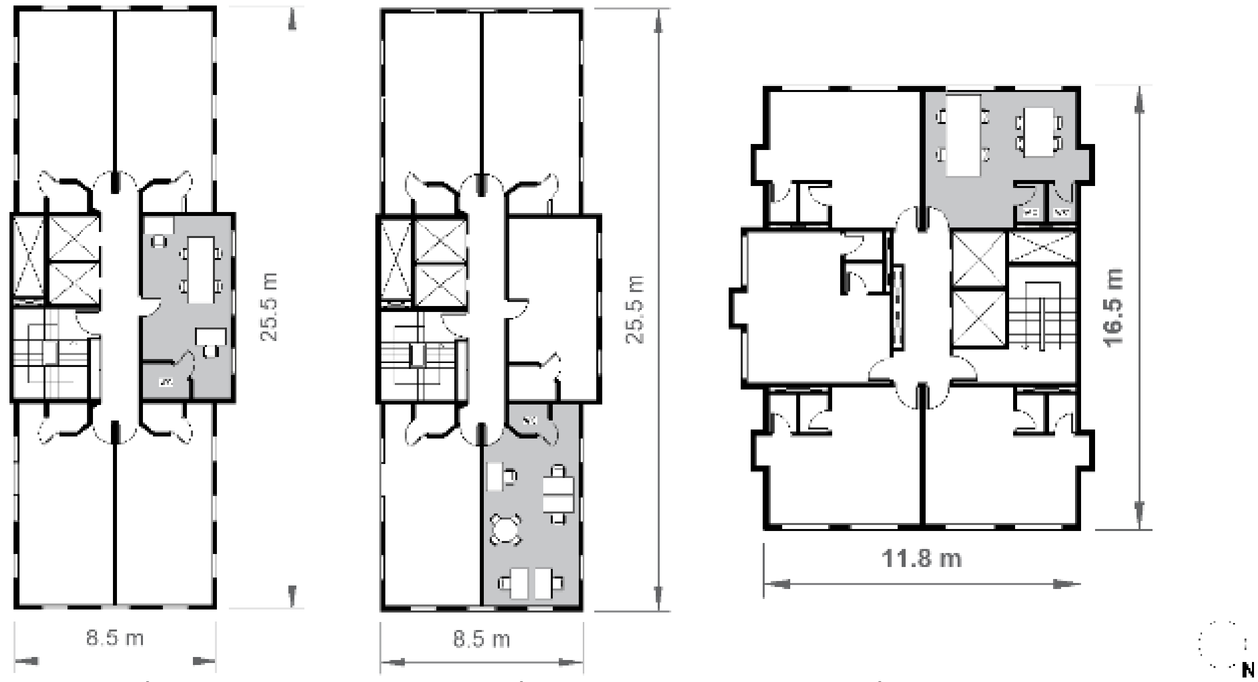


Figure 1. Analysis workflow.

2.1 On-site measurements: characteristics of the selected offices

Figures 2 and 3 show floor plans and views of the monitored offices, and Table 1 their geometry, envelope thermal properties and internal loads information. Offices 1 and 2 share a Northwest facing (i.e. equatorial) aspect whereas Office 3 faces Southeast.



(a) Office 1 (4th floor) (b) Office 2 (6th floor) (c) Office 3 (4th floor)

Figure 2. Dimensions, floor location and orientation of the monitored office rooms in São Paulo, Brazil.

(a) Office building – rooms 01 and 02 (Northwest façade)

(b) Office building – room 03 (Southeast façade) (Google Earth, 2019)

Figure 3. External view of the monitored buildings from ground level. The monitored offices are highlighted in red.

Table 1. Office rooms input data

Parameter	*Office 1*	*Office 2*	*Office 3*
Office floor area	25 m²	33 m²	29 m²
Office floor-to-ceiling height	2.6 m	2.6 m	2.75 m
Glazing thermal transmittance	5.67 W/(m².K) (single pane)		
Glazing solar heat gain coefficient	0.87 (clear glass)		
Slab thermal capacity	295 kJ/(m².K) (10 cm concrete slab)		
Wall thermal transmittance	2.75 W/(m².K) (mortar 2 cm + concrete block 14 cm + mortar 2 cm)		2.0 W/(m².K) (mortar 2 cm + concrete block 19 cm + mortar 2 cm)
Wall thermal capacity	238 kJ/(m².K)		291 kJ/(m².K)
Natural ventilation strategy	Single sided ventilation	Cross ventilation (adjacent facades)	Single sided ventilation
Glazed area that opens	74% (fixed glazing at the bottom)		79% (fixed glazing at the bottom)
Free area of opening	64% (OT hung window frame)		19% (OT hung window frame)
Internal loads – occupancy	8.3 m²/person		7.3 m²/person
Internal loads – lights	7.2 W/m²	4 W/m²	8.3 W/m²
Internal loads – equipment	16 W/m²	14.5 W/m²	6.9 W/m²
Occupancy schedule	Weekdays from 0900 to 1800		
Clothing insulation	winter = 0.97 CLO, spring = 0.55 CLO, summer = 0.45 CLO, autumn = 0.57 CLO		

Indoor environmental variables monitored *in situ* comprise: air temperature (T_a), globe temperature (T_g), air speed (V_a) and relative humidity (RH). User control variables were also monitored, comprising the manual operation of the air-conditioning and natural ventilation systems (i.e. operable windows). Instruments used to perform the field monitoring are listed in Table 2. T_a + RH sensors were placed at the air-conditioning outlet to infer air-conditioning on/ off state. The other equipment was placed on a tripod shielded from direct sources of radiation, at the height correspondent to a standing person (1.2 m). The tripods were located at the corner of the rooms to minimise occupant disturbance.

Table 2. Technical specification of instruments used during field monitoring.

Instrument	*Range*	*Accuracy*
Datalogger air temperature/ relative humidity	-20 °C to +55 °C 0 to 100%	± 0.4 °C ± 2%
Datalogger globe temperature (black globe)	-35 °C to +55 °C	±0.5 °C
Datalogger air temperature/ air speed	1 to + 20 m/s	± (0.03 m/s + 5%)
Air speed probe	0 to + 10 m/s	± (0.03 m/s + 5%)
Air temperature probe	-50 °C to + 125 °C	± 0.2 °C

Datalogger air temperature/ relative humidity (for air-conditioning triggering monitoring)	-20 °C to + 85 °C 0 to + 100 %	±0.5 °C 0.6 %
State measure datalogger (for window position monitoring)	Maximum frequency 1 Hz	± 1 minute per month at 25 °C

Data acquisition occurred every 15 minutes for indoor environmental variables, every 5 minutes for air-conditioning triggering monitoring and every time a state change was detected for window operation. Data collection occurred for each season of the year, as follows:

- Winter: June 19th to July 3rd 2017 = 15 days.
- Spring: October 3rd to 26th 2017 = 24 days.
- Summer: January 19th to February 5th 2018 = 18 days.
- Autumn: April 4th to 18th 2018 = 15 days.

2.2 Weather data

Outdoor environmental data from June 2017 to May 2018 were taken from a meteorological station located 9 km distant from the buildings (INMET, 2018). Acquired information comprise hourly data of atmospheric pressure, air temperature, relative humidity, air speed and direction, global solar radiation and rainfall index. Weather Converter v8.8 was used to calculate direct normal and diffuse horizontal solar radiation, by using Perez model (EnergyPlus, 2017), and to create an EnergyPlus weather file (epw). Figure 4 presents the weather file profile for dry bulb temperature (DBT) and relative humidity (RH).

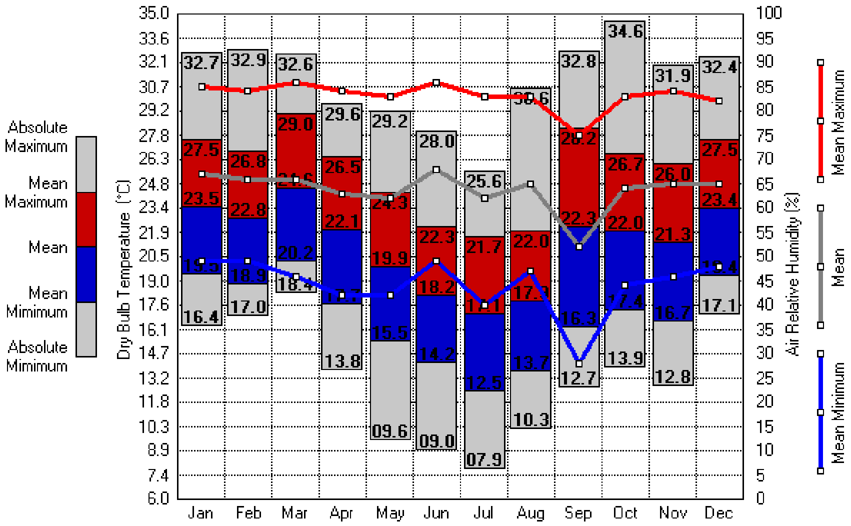


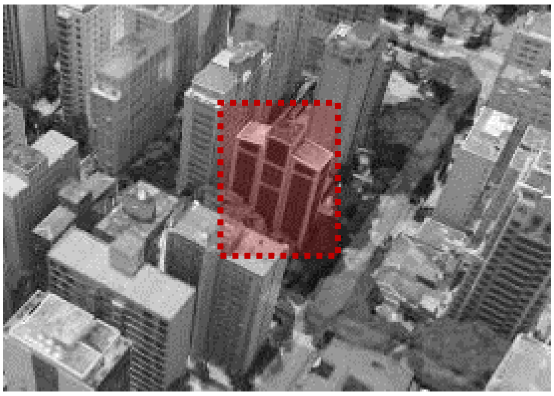
Figure 4. Monthly mean minimum and maximum dry bulb temperature (DBT) and relative humidity (RH) for São Paulo. January to May corresponds to year 2018 and June to December corresponds to year 2017.

2.3 Thermal modelling

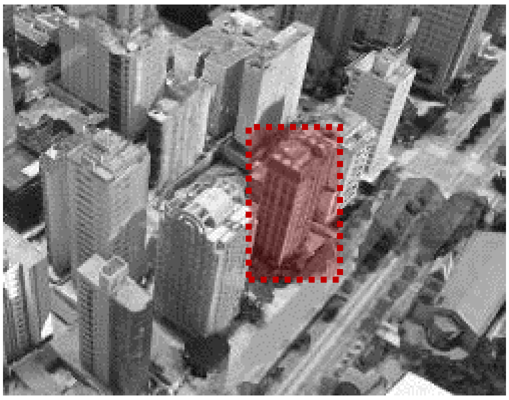
Our approach is to discover the true mode of operation of the offices, through comparing observed and simulated data. However, uncertainty in inputs such as operation schedule, building physics (i.e. precise thermal characteristics of building materials), site characteristics and solar radiation preclude a direct comparison. Instead, our approach is to undertake a *model-to-model* comparison where the "observed" data is represented by a base case simulation. We begin with a thermal model calibrated against observed data when gains were minimal and no occupant behaviour was present (i.e. window state is unchanged and AC is off), i.e. weekend evenings. This model is then injected with the real observed operation behaviour to create the base case. The comparators are then created using the same calibrated thermal model but with algorithmically derived operation behaviour. We discuss this approach in detail below.

2.3.1 Building energy simulation

We use the well-known EnergyPlus (v8.8) software for our modelling. Natural ventilation was modelled with the multizone Airflow Network (AFN) model. One limitation of the AFN model is its inability to control the window opening area. This was mitigated using an EnergyPlus Runtime Language (ERL) code injected via the Energy Management System (EMS) to override the window control modulation. This produces a binary (open/ closed) control state. EnergyPlus uses a surface average calculation for estimating wind pressure coefficients. However, as our offices are in a densely built urban area (Figure 5) and at low to middle floor levels, we use input values from Tokyo Polytechnic University's aerodynamic database for non-isolated buildings (TPU, 2019) as a first approximation to account for the surrounding obstructions. The discharge coefficient (C_d) was set to the standard value of 0.6 (Florentzou et al., 1998).



(a) Office building – rooms 01 and 02



(b) Office building – room 03

Figure 5. Aerial view of selected buildings. The monitored buildings are highlighted in red. (Google Earth, 2019).

The cooling system was modelled using an ideal air conditioning system, with an outdoor airflow rate of 0.5 air changes per hour and a cooling setpoint of 25 °C. Since the monitored offices did not have a heating system, an artificially low setpoint (5 °C) was set for heating to ensure the heating system would never turn on. In order to justify disregarding unmet hours for heating, we analysed the number of hours below the thermal comfort limit within the monitored data, according to the ASHRAE 55 (2017) adaptive model. The result was zero unmet hours in office 1, four unmet hours in office 2 (which corresponds to 0.5% of the monitored period) and fifteen unmet hours in office 3 (which corresponds to 2.2% of the monitored period). Therefore, we assumed acceptable that the indoor temperature conditions during the occupied period never falls below the lower threshold.

2.3.2 Model calibration

The main parameters that influence indoor conditions are outdoor air temperature, solar radiation and internal heat gain. To minimise uncertainty, we utilised weekend night-time data – i.e. when the building conditions were not influenced by either solar radiation or occupant-related heat gains – and hence only the temperature time series for calibration. Changes in outdoor air temperatures could influence the indoor conditions with conduction heat transfer through the wall, the window and the infiltration. The influences from infiltration in three offices were accounted by inputting the minimum values of AFN. For selecting the most accurate minimum AFN values, simulations with 0.01 increment of opening factor up to opening factor of fully opened window (1) for each office were run and selected the opening factor value with minimum Mean Absolute Error (MAE) of the simulated and measured indoor operative temperatures.

After calibrating the building physical model, indoor conditions during the day were calibrated by extracting the setpoint temperatures and internal loads of each office building. Setpoint temperatures could be estimated by observing the indoor air temperature while the air conditioning was on. With calibration on building physics and setpoint temperatures, and maximum internal gains estimated based on counting fixtures, appliances and workers at each office, multiple parametric simulations were run to find the closest operative temperatures to measured values.

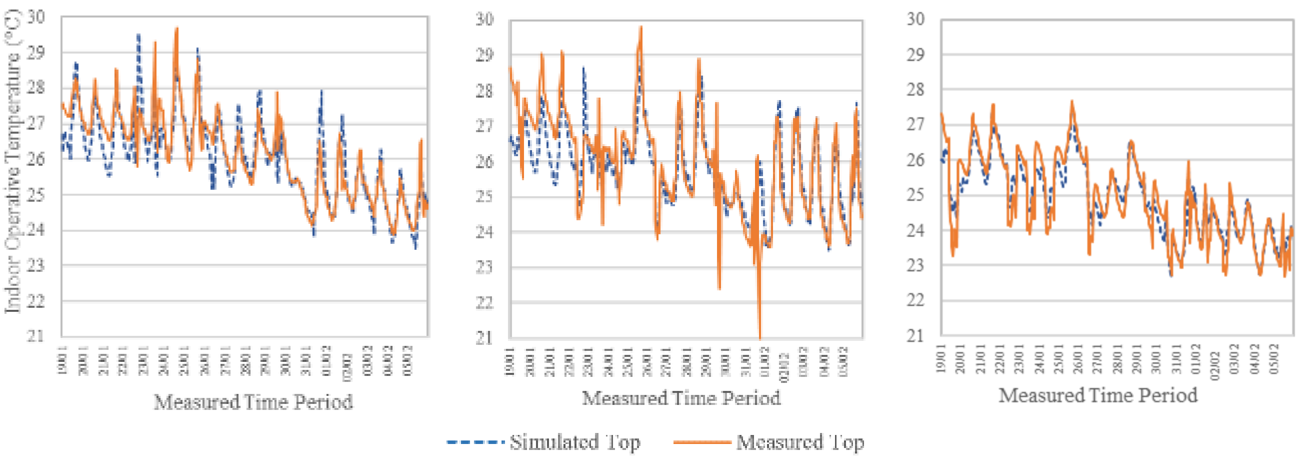
To select the simulation that most closely matched measured operative temperatures, we used MAE to observe overall difference and Normalised Mean Bias Error (NMBE) and Coefficient Variation of Root Mean Square Error (CV RMSE) were used for accuracy. MAE in each office was maintained below 1 ℃ and NMBE and CV RMSE were significantly lower than the threshold in standards (10% and 30% respectively in ASHRAE guideline 14 (ASHRAE, 2002)) as demonstrated in Table 3.

Table 3. Accuracy of calibration

Office	*MAE [℃]*	*NMBE [%]*	*CV RMSE [%]*

Office 1	0.9	-2.8	4.0
Office 2	0.5	-0.1	-3.1
Office 3	0.3	-0.1	1.8

Based on the calibration of building physics, setpoint temperatures, and occupants' schedules, the measured and simulated operative temperatures (T_{op}) of each office are shown in Figure 6 (excerpt for summer period).

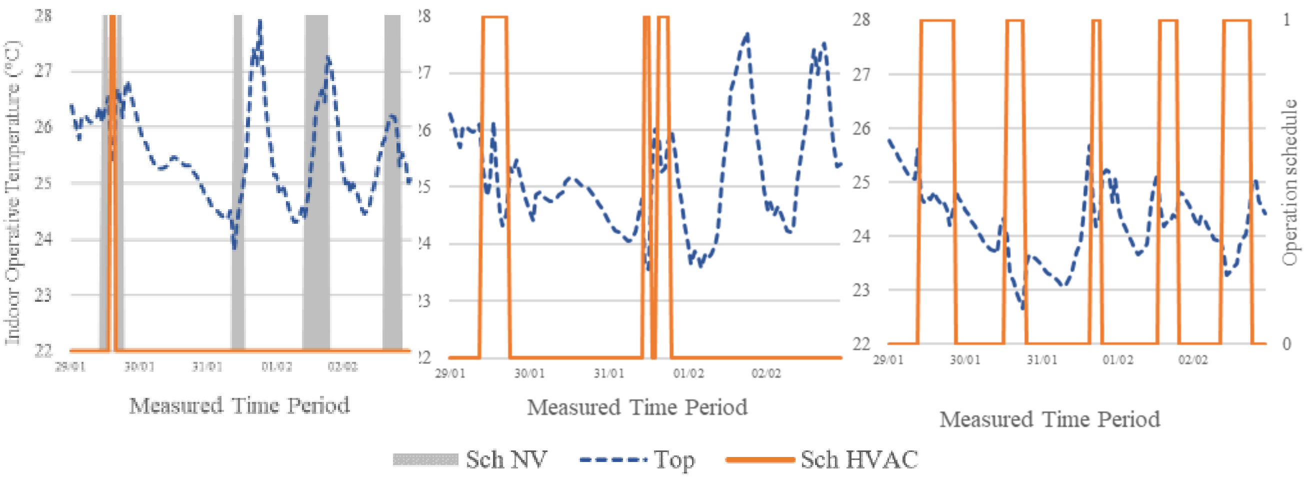


(a) Office 1 (b) Office 2 (c) Office 3

Figure 6. Calibrated operative temperature examples for each office, excerpted over the summer period.

2.3.3 'Spontaneously controlled' base-case

We injected real observed window and AC operation data into the calibrated models obtained above to create a 'spontaneously controlled' base case for each office. An illustrative week for each office in summer is shown in Figure 7, with the resultant operative temperature plotted alongside the injected AC and window operation binary data.



(a) Office 1 (b) Office 2 (c) Office 3

Figure 7. Illustrative summer week of the spontaneously controlled model – window operation (Sch NV), operative temperature (Top) and air-conditioning operation (Sch HVAC).

2.3.4 Operational scenarios

In order to "read" the mode of operation chosen by the occupants, a series of rule-based scenarios were created (Table 4). These were based on a realisation of the simplest set of design rules driven by the static and adaptive models of ASHRAE 55 standard (ASHRAE, 2017).

There are five scenarios for AC operation based on 1 °C increments in indoor setpoints from 22 °C to 26 °C. These were obtained from the suggested central estimates of operative temperature in the static model of ASHRAE 55 (2017) for typical CLO values gathered in this study (0.45 < CLO < 1 and assuming RH = 50%

and Va < 0.1 m/s). NV operation has three scenarios. Two are derived from the ASHRAE 55 adaptive comfort acceptability thresholds of 80% and 90%, (ASHRAE, 2017) and the third is ‘always on’. MM operation has fifteen scenarios based on the above modes of operation: 5 AC x 3 NV modes. In this case, the ASHRAE 55 static comfort model (Predicted Percentage of Dissatisfied < 10%) (ASHRAE, 2017) was added as an option to determine upper limit comfort temperatures, since the standard does not specify a specific thermal comfort model for MM buildings. This results in 23 distinct scenarios, each of which with an operational ruleset defined by the mode of operation (AC, NV or MM) (Table 4). For example, under AC operation, the rules can be understood as follows (NV is always OFF):

AC Rule 1. If PPD < 10% on arrival then set AC to OFF.
AC Rule 2. If PPD > 10% OR previous timestep AC = 1, set AC = ON.
AC Rule 3. At departure, set AC = OFF.

For the MM and NV scenarios, we used a deadband of 2 °C above T_{out} as the excursion threshold for indoor operative temperature, as suggested by Rijal et al. (2007). This is necessary to avoid model instability, such as windows opening and closing at the same temperature.

These scenarios were fed into the calibrated thermal models to produce time series of operative temperature and window/ AC operation that could be compared against the three base cases. Simulations were performed on a 1 h timestep.

Table 4: Rule-based scenarios used in BES.

Mode of operation	*Scenarios*	*Rule*	*Actuation*	
			AC	*Window*
Fully air-conditioned	AC setpoint: 22, 23, 24, 25, 26 °C = (5 scenarios)	Arrival: PPD < 10%	0	0
		PPD ≥ 10% OR previous timestep AC = 1	1	0
		Departure	0	0
Mixed-mode	AC setpoint: 22, 23, 24, 25, 26 °C x NV setpoint: T_{adp} 80%, T_{adp} 90%, PPD < 10% = (15 scenarios)	Arrival: $T_{op} < T_{adp}$/PPD &/or $T_{op} > T_{out} + \Delta = 2$ °C	0	1
		Arrival: $T_{op} < T_{adp}$/PPD & $T_{op} \leq T_{out} + \Delta = 2$ °C	0	0
		$T_{op} < T_{adp}$/PPD & previous timestep AC = 0; NV = 0	0	0
		$T_{op} < T_{adp}$/PPD & previous timestep AC = 0; NV = 1	0	1
		$T_{op} \geq T_{adp}$/PPD & $T_{out} > T_{adp}$/PPD & $T_{op} > T_{out} + \Delta = 2$ °C	1	0
		$T_{op} \geq T_{adp}$/PPD & $T_{out} \leq T_{adp}$/PPD & $T_{op} > T_{out} + \Delta = 2$ °C	0	1
		$T_{op} \geq T_{adp}$/PPD & $T_{op} \leq T_{out} + \Delta = 2$ °C	1	0
		Departure	0	0
Fully naturally ventilated	NV setpoint: T_{adp} 80%, T_{adp} 90%, = (2 scenarios)	Arrival: $T_{op} < T_{adp}$ &/or $T_{op} \geq T_{out} + \Delta = 2$ °C	0	1
		Arrival: $T_{op} < T_{adp}$ & $T_{op} < T_{out} + \Delta = 2$ °C	0	0
		$T_{op} \geq T_{adp}$ & $T_{op} < T_{out} + \Delta = 2$ °C	0	0
		$T_{op} \geq T_{adp}$ & $T_{op} \geq T_{out} + \Delta = 2$ °C	0	1
		Departure	0	0
	Always on (1 scenario)	Arrival and during the day	0	1
		Departure	0	0

Where:
PPD = predicted percentage of dissatisfied (Static thermal comfort model) (ASHRAE, 2017)
T_{op} = indoor operative temperature (°C)
T_{adp} = adaptive model upper limit operative temperature (models for 80% and 90% limits) (ASHRAE, 2017)
T_{out} = outdoor dry bulb temperature (°C)
0 and 1 indicate the binary actuation states of CLOSED/OFF and OPEN/ON, for window and AC operation respectively.
Rules marked as Arrival or Departure are evaluated only under those conditions.

2.4 Comparison between ‘spontaneous control’ base model and rule-based models

To answer the research question formulated earlier we aim to compare the following features of the 'spontaneously controlled' base thermal model against the rule-based models:

- Indoor operative temperatures (T_{op}).
- Window and air-conditioning (WAC) operation.

All of our temporal data can be represented as hourly time series. The time periods for monitored and generated data are the same so time series to be compared will have the same length. For the first case we have a univariate time series with hourly measurement of indoor operative temperatures. For the second case the time series is multivariate as it contains hourly observed binary modes for both window and air-conditioning operation (window open/ close, AC on/ off). Hence the comparison problem can be reformulated as finding similarity measures between two sequences of time series data.

Determining time series similarity measure i.e. the degree to which one time series resembles another is central to many data mining, pattern recognition and classification tasks (Keogh and Kasetty, 2003; Wang et al., 2013). However, the choice of similarity function heavily depends on the type of similarity sought (pattern matching, trend similarity or pairwise distance), time series dimensionality and types of data. We describe our approach to analysing both sets data below.

2.4.1 T_{op} time series

Let us consider two time series of the same length:

$X = \{x_t , x_t \in R^1\}$ and

$Y=\{y_t, y_t \in R^1\}$,

where t=1..n; n is a positive integer number representing the length of time series.

x_t denotes T_{op} for 'spontaneous control' base model at time t and

y_t denotes T_{op} for selected rule-based model at time t.

Here, the problem comprises of a simple pair (i.e. 'spontaneous' base vs rule driven) of temporally co-incident time series of continuous data. Hence a method that can compare the spatial proximity of samples that are at exactly the same temporal location is needed. The Euclidean distance (ED) approach is a widely applied time series similarity measure designed for such problems. The main advantage of Euclidean distance compared with other similarity measures such as autocorrelation and cepstrum is its computational simplicity and indexing capabilities[2] (Xing et al., 2011). The results from several studies conducted to compare similarity measures on different test sets of continuous data have shown that the accuracy of ED approach is approximately 85% and its error rates are lower than those obtained from applying other similarity measures (Keogh and Kasetty, 2003; Wei and Keogh, 2006; Xing et al., 2011). Also, ED is not sensitive to alignment of values (shifting) (Cassisi et al., 2012).

The ED measure sums the Euclidean distance computed for each sample at the same temporal location (Equation 1):

$$d_E(X,Y) = \sqrt{\sum_{t=1}^{n}(x_t - y_t)^2} \tag{1}$$

The use of a difference measure in the definition above makes it clear that this method will work well for continuous data but is unsuited to binary or categorical data. Hence, a different method is needed for the WAC time series.

2.4.2 WAC operation time series

For this case let us consider two time series of the same length:

$X = \{x_{1t}, x_{2t}, (x_{1t}, x_{2t}) \in \{0, 1\}\}$ and

[2] Indexing is the process of comparing one time series to another to find similarity/dissimilarity.

$Y = \{y_{1t}, y_{2t}, (y_{1t}, y_{2t}) \in \{0, 1\}\}$,

where t=1.. n; n is a positive integer number representing the length of time series.

(x_{1t}, x_{2t}) pair denotes window operation (open/close) and AC on/ off states correspondingly for 'spontaneous control' base model at time t.

(y_{1t}, y_{2t}) pair denotes window operation (open/ close) and AC on/ off states correspondingly for selected rule-based model at time t.

Since the window and air-conditioning operation profiles are in binary form (window open/ close, AC on/ off) and ED is unsuited to binary multivariate data, a different approach is needed for pattern matching the operation profiles. Essentially, the chosen approach will need to account for (i) the need to match two time series pairs simultaneously, one for window and the other for AC operation and (ii) temporal differences in the search for an overall pattern. Temporal differences exist over a continuum defined by differences in event duration and whether the event centres are temporally co-incident. Apart from the trivial cases of a perfect match (durations equal *and* centres temporally co-incident) or perfect mismatch (no overlap in events), for all cases with some overlap in events, the following scenarios are possible (middle three cases in Figure 8):

- Equal duration but time shifted.
- Differing duration but temporally co-incident.
- Differing duration and not temporally co-incident, but overlapping.

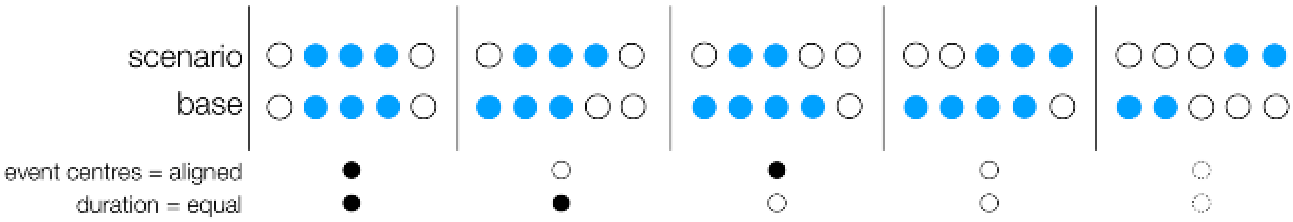


Figure 8. Possible matching scenarios.

Our interest is to ascertain whether events in the base and scenario pairs are *broadly*, but not necessarily *exactly*, temporally co-incident. Due to the complex nature of occupant ↔ building interaction, noted elsewhere (Haldi and Robinson, 2009), we are also agnostic towards event lengths. However, a method is needed that can appropriately grade performance such that the situations shown in Figure 8 decrease in score from left to right (with situations (2) and (3) being co-equal). These requirements are met through Dynamic Time Warping (DTW) (Sakoe and Chiba, 1978; Berndt and Clifford, 1994). DTW is a classic approach for computing similarity between two univariate or multivariate time series. It has been widely applied in areas such as speech recognition (Myers et al., 1980) and behavioural perception (Corradini, 2001). DTW applies an elastic transformation of time series data to recognize similar phases between different patterns along time hence minimizing distortion effects that could arise as a result of time-dependency.

DTW approach arranges time series X and Y into a matrix *D* of size *n x n*, where each element represents distance between elements $x_i = (x_{1i}, x_{2i})$ and $y_j = (y_{1j}, y_{2j})$. Note that in general *D* does not have to be square (i.e. it can contain *n* rows and *m* columns). In case of time series of binary data, we choose magnitude of the difference (Manhattan distance) (Craw, 2017) as the most suitable measure (Equation 2).

$$d(i,j) = |x_{1i} - y_{1j}| + |x_{2i} - y_{2j}| \tag{2}$$

A warping path $W = w_1 \ldots w_k$ is a sequence of points $w_k = (i_k, j_k)$, where each w_k corresponds to a mapping between x_i and y_j, such that the following conditions are satisfied:

- Boundary: $w_1 = (1, 1)$ and $w_k = (n, n)$ where *k* is the length of the warping path.
- Continuity: if $w_k = (i_k, j_k)$ and $w_{k-1} = (i_{k-1}, j_{k-1})$ then $i_k - i_{k-1} \leq 1$ and $j_k - j_{k-1} \leq 1$.
- Monotonicity: if $w_k = (i_k, j_k)$ and $w_{k-1} = (i_{k-1}, j_{k-1})$ then $i_{k-1} \leq i_k$ and $j_{k-1} \leq j_k$.

DTW similarity measure between time series X and Y can be calculated by solving the following minimization problem (Equation 3):

$$DTW(i,j) = \min_{W} \sum_{l=1}^{k} d(w_k) \tag{3}$$

The calculation of DTW relies on a dynamic programming algorithm and has a complexity of O (n^2). The difference between ED and DTW is visually illustrated in Figure 9.

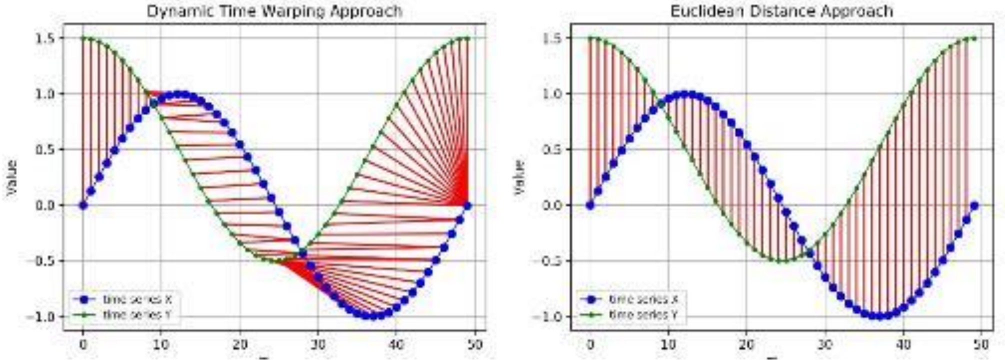
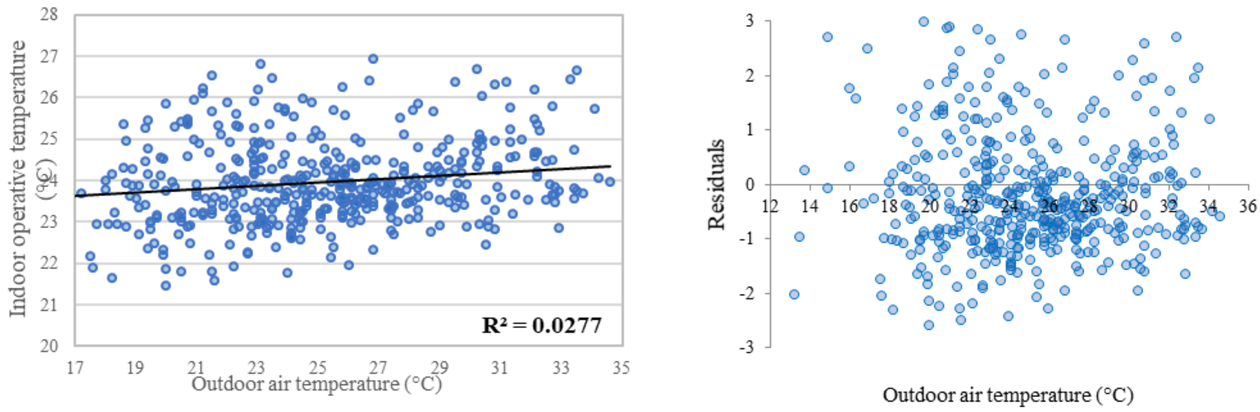


Figure 9. Dynamic Time Warping vs. Euclidean Distance approach.

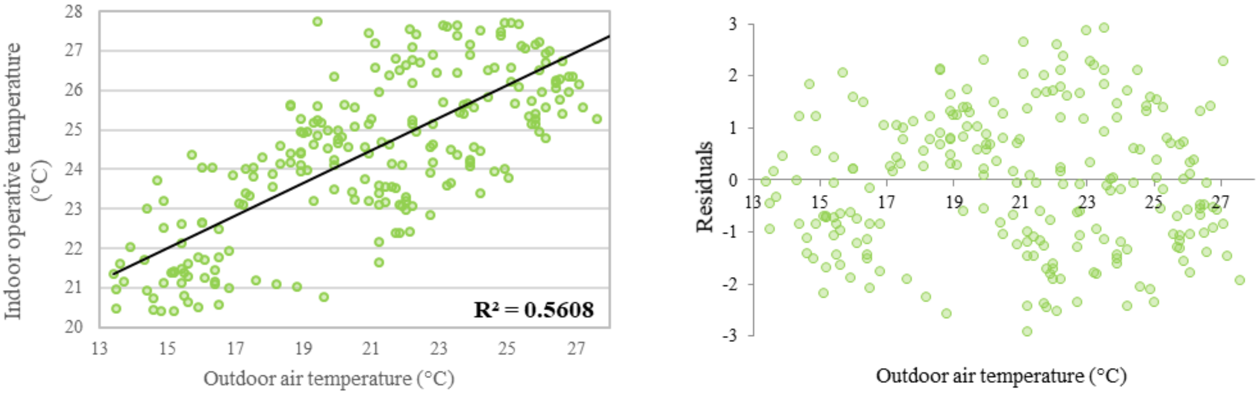
3. Results and discussion

3.1 General observations

In this section, we make some general observations about our data as a broad “sanity check”. It is well-known that air-conditioning operation effectively disconnects the indoor and outdoor environments, whereas natural ventilation creates a close connection between the two. This is verifiable using a regression analysis comparing outdoor air temperature (T_{out}) and indoor operative temperature (T_{op}). As expected, Figure 10 indicates a very poor correlation between T_{out} and T_{op} for all hours when the air-conditioner was known to be turned on, but a good correlation for all periods when the windows were known to be open (R^2 = 0.56). Interestingly, we observe significantly more hours of AC operation (485) than NV operation (236). These observations are supported by box-plots of the seasonal differences between the AC and NV periods (Figure 11); where we observe that the NV mode shows a larger range of temperatures than the corresponding AC mode whilst also showing clear seasonal drift in median temperatures.



(a) Air-conditioning ON (N=485)



(b) Window OPEN (N=236)

Figure 10. Outdoor air temperature plotted against indoor operative temperature for all three offices (N = number of hours).

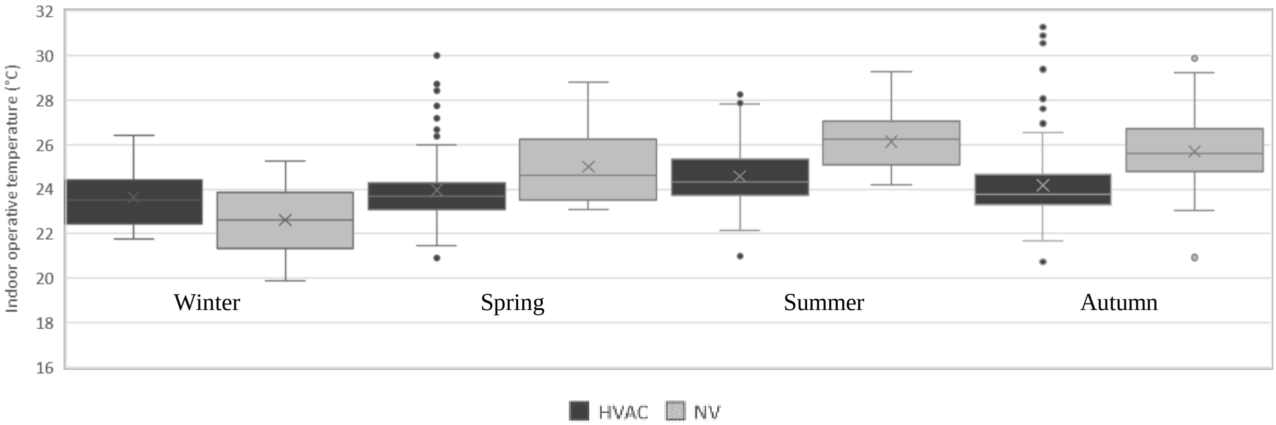


Figure 11. Box plot of indoor operative temperature for all three offices. The line within each box is the median, the cross is the mean, the edges of the box are the 25th and 75th percentiles, the dots are the outliers.

Indoor air velocity (V_a) was low ($\bar{x}$ = 0.04 m/s, s = 0.02 m/s), even when windows were opened. No correlations could be found between indoor relative humidity and outdoor air temperature during natural ventilation and air-conditioning triggering. A slight correlation could be found between indoor and outdoor relative humidity for the naturally ventilated period (R^2 = 0.28), but not for the air-conditioned period (R^2 = 0.05) (Figure 12).

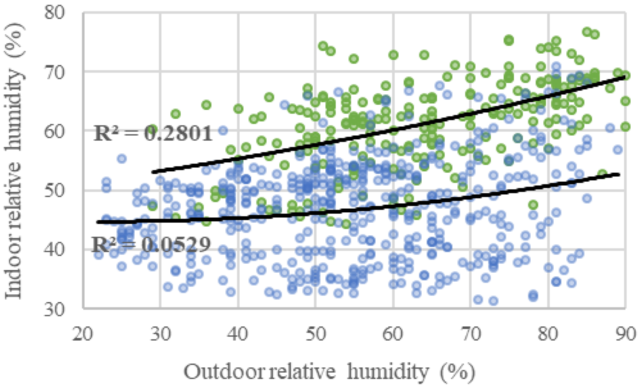


Figure 12. Indoor relative temperature plotted against outdoor relative humidity for all three offices (NV represented in green; AC represented in blue).

3.2 Comparison between 'spontaneous control' base model and rule-based models

Monitored data on indoor operative temperatures and on window and air-conditioning operation were pre-processed using the following rules:

- Time intervals outside of known working hours (1900 to 0900) were removed.
- Since there was no monitoring of occupants' actual times of arrival and departure, the days when no WAC operation events were observed were also removed from the sample.
- Due to a technical problem with the state measure datalogger, spring measurements for office 1 resulted in only one valid monitoring day, and was hence excluded from the analysis.

When taken with the fact that mean events per day were 1.04 (with 89% of days with only one event), the analysis of observed data suggests low interaction with the indoor environment. This is expressed mathematically in Equations 4 and 5 where we estimate the likelihood of having window open or air-conditioning turned on for all three offices throughout all seasons using logistic models and Markov Chain Monte Carlo (MCMC) sampling (Gill, 2008; Ross, 2017). As expected, the probability distributions generated for window and air-conditioning actuation for all the offices are similar and suggest that a strategy, once selected, generally tended to remain unchanged for the day (Figure 13). Window actuation is concentrated in the mornings (67% of events occur before 1200 and 62% between 0900 and 1000), whereas air-conditioning is mainly triggered in the afternoon (59% of the events happen after 1200 and 35% between 1400 and 1500), when outdoor air temperatures are usually higher.

$$p(NV_{on}|t) = \frac{1}{1 + e^{\beta * t + \beta_0}} \tag{4}$$

Equation 4 presents logistic probability distribution for window actuation where t denotes the time and β and β_0 are unknown parameters estimated using MCMC sampling.

$$p(AC_{on}|t) = \frac{1}{1 + e^{\gamma * t + \gamma_0}} \tag{5}$$

Equation 5 presents logistic probability distribution for air-conditioning actuation where t denotes the time and γ and γ_0 are unknown parameters estimated using MCMC sampling.

Similarly, seasonal between-day analysis for all offices suggests that strategies did not change between days. The exceptions were Office 1 in summer and Office 2 in winter. The between-day transition probability matrix for office 1 can be seen in Equation 6. For office 2, there were not enough transitions to construct a matrix.

$$\begin{array}{cc} & \begin{array}{cc} AC & NV \end{array} \\ \begin{array}{c} AC \\ NV \end{array} & \begin{pmatrix} 3/4 & 1/4 \\ 2/3 & 1/3 \end{pmatrix} \end{array} \tag{6}$$

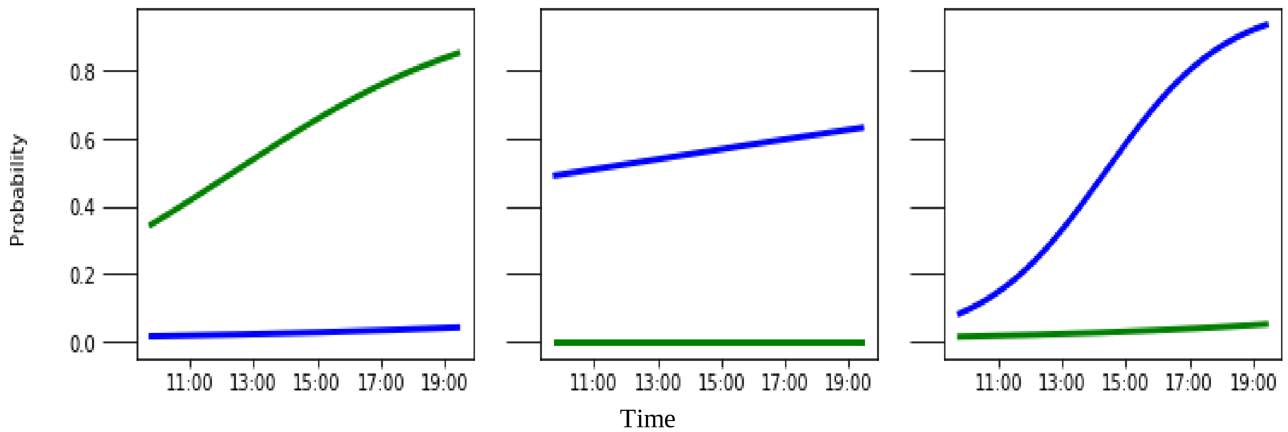


Figure 13. Probability for window and/or air-conditioning actuation for all three offices.

Results of the comparison between the spontaneously controlled base model and the rule-based scenarios were obtained through the application of ED on T_{op} and DTW on WAC operation. The theoretical range for ED in our dataset was between 6.23 and 31.50 whereas it was between 24 and 198 for DTW. Since, for both methods, the more similar the time series, the smaller the distance between them (see Sections 2.4.1 and 2.4.2), we normalise scores using the following expression (Equation 7), to facilitate comparison. This converts scores from both methods to a value between 0 (best fit) and 1 (worst fit). Detailed scores can be seen in Appendix B.

$$\text{score}_{norm} = \frac{\text{score}_{actual} - \text{range}_{min}}{\text{range}_{max} - \text{range}_{min}} \tag{7}$$

Table 7 summarizes the best fit strategy for each office by season. It is clear that while identical matches for both methods rarely occur (the exception being summer in Office 1), there is considerable agreement between them. In several instances ED identifies a single setpoint temperature whereas DTW is unable to identify the temperature but correctly identifies the operational mode. This is unsurprising given that DTW is purposely expansive in its definition of a match (per Figure 9). It is noteworthy that the NV scenarios can be considered as being in hierarchical sets where a match at the top level of the hierarchy will automatically result in matches for the lower levels. This is because all instances of 90% adaptive acceptability will be captured within 80% acceptability. Similarly, all instances of 80% acceptability can be expected to be contained within "always on"[3]. Hence, Table 7 only shows the highest matched level of the hierarchy. The same logic also applies for NV regimes within MM operation.

Table 7: Best fit strategy for each office by season, considering indoor operative temperature (Euclidean Distance – ED) and window and air-conditioning operation (Dynamic Time Warping – DTW).

	Winter		**Spring**		**Summer**		**Autumn**	
	ED	DTW	ED	DTW	ED	DTW	ED	DTW
Office 1	NV Always on	NV Always on	*	*	MM Adaptive 90%	MM Adaptive 90%	NV Always on	MM Adaptive 80%
Office 2	AC 26 °C	AC 22—26 °C	AC 25 °C	AC 22—26 °C	MM PPD < 10% 22 °C	AC 22—26 °C	AC 26 °C	AC 22—26 °C
Office 3	MM Adaptive 90%	**	MM PPD < 10% 22—26 °C	AC 22—26 °C	AC 25 °C	AC 22—26 °C	MM PPD < 10% 22—26 °C	AC 22—26 °C

* Spring measurements for office 1 resulted in only one valid monitoring day and was excluded from the analysis.
** The best match for WAC operation was AC 22-26°C, but the AC never turned on, which means that no rule-based scenario was considered a best match, for this case.

For office 1, we observe that NV was the best match for indoor operative temperature in winter and autumn. This is confirmed through the T_{op} time series in Figures 14 and 16, which show the match between the spontaneously-controlled base model and the best fit scenario. For the summer period, MM was the best match (Figure 18). For window and air-conditioning operation, NV was the best match in winter, since air-conditioning did not turn on during the winter (Figure 15), while in autumn and summer the MM scenario showed a better correlation (Figures 17 and 19). In autumn, the air-conditioning system turned on during only three hours in the autumn (Figure 17b). In summer, even though the air-conditioning system was turned on during three afternoons, natural ventilation was in use 50 out of 80 occupied hours during the measurement period (Figure 19a).

[3] Though this may not happen if external temperatures regularly go above or below the upper and lower thresholds of the adaptive criteria.

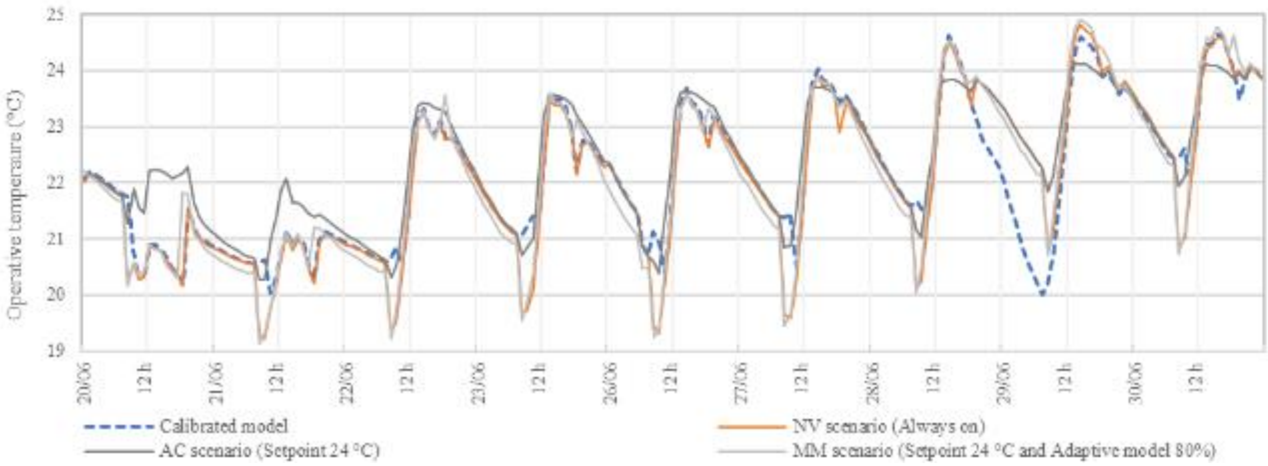
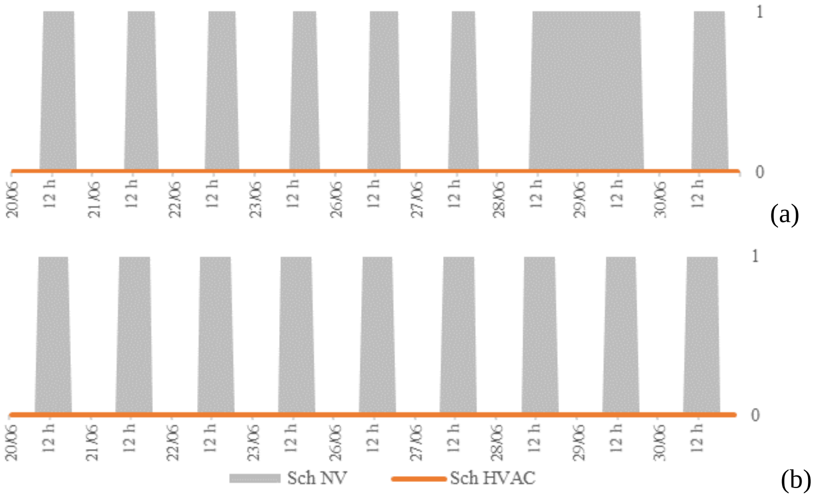


Figure 14. Indoor operative temperature comparison for the winter period – Office 1.



(a) Spontaneous control base case
(b) Rule-based scenario (NV – Always on)

Figure 15. Comparison between natural ventilation (Sch NV) and air-conditioning (Sch HVAC) schedules of operation for the winter period – Office 1.

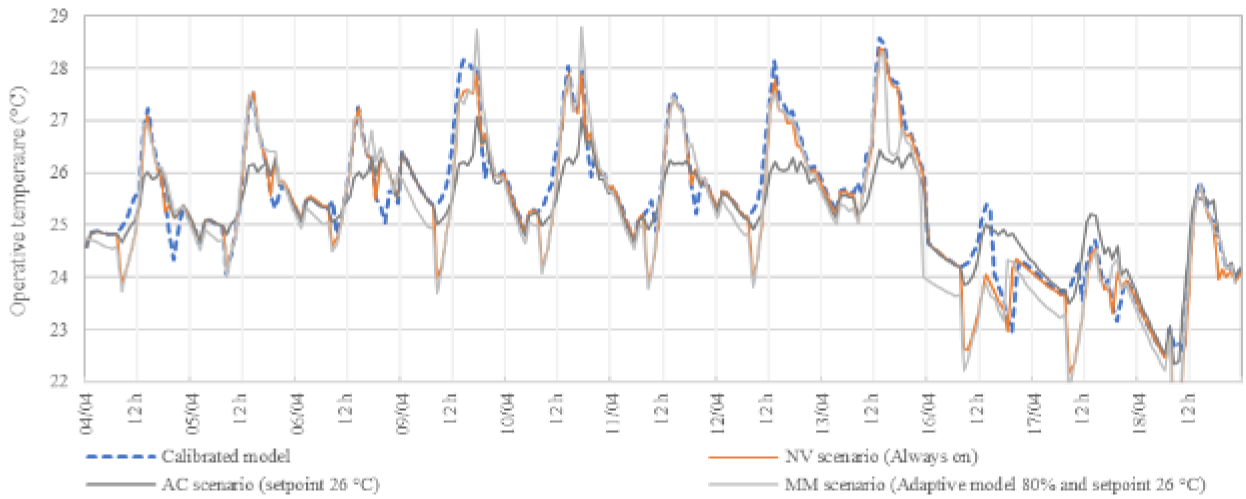
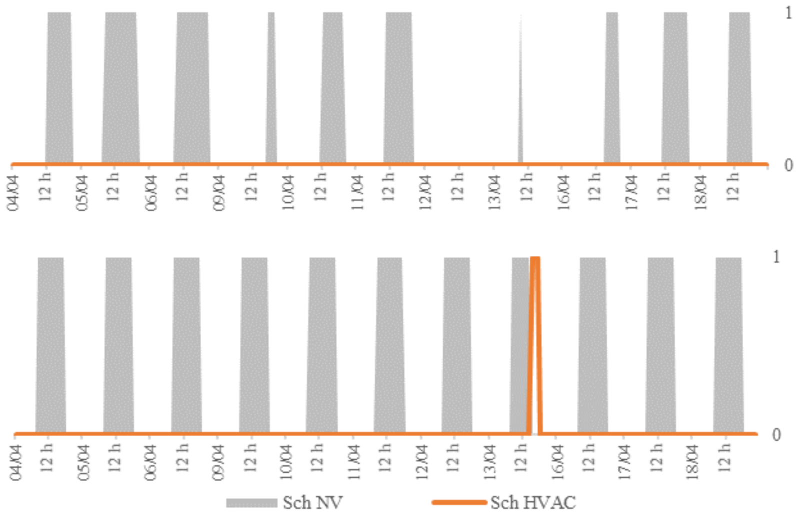


Figure 16. Indoor operative temperature comparison for the autumn period – Office 1.



(a)

(b)

(a) Calibrated model (spontaneous control)
(b) Rule-based scenario (MM – Adaptive model 80%, AC setpoint 24 °C)

Figure 17. Comparison between natural ventilation (Sch NV) and air-conditioning (Sch HVAC) schedules of operation for the autumn period – Office 1.

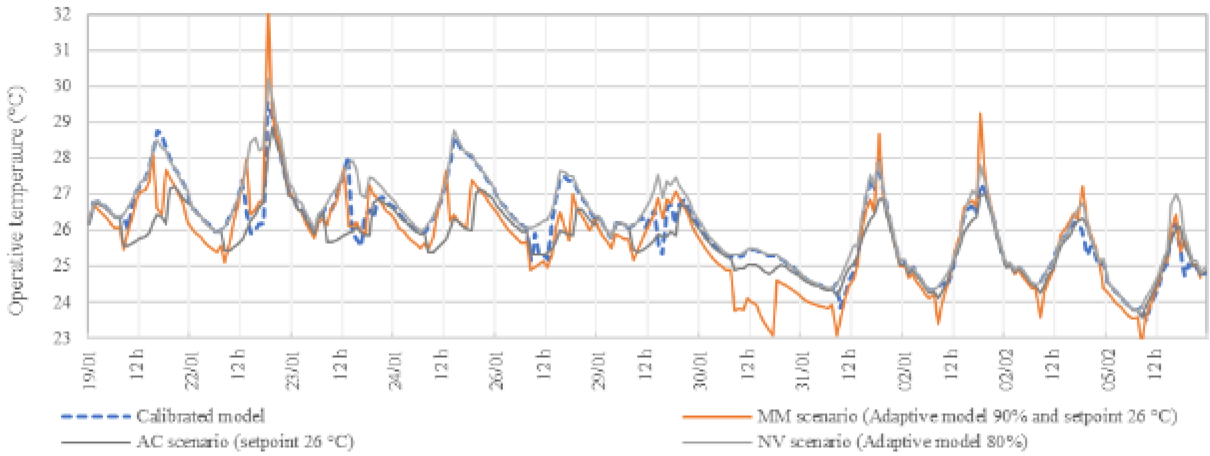
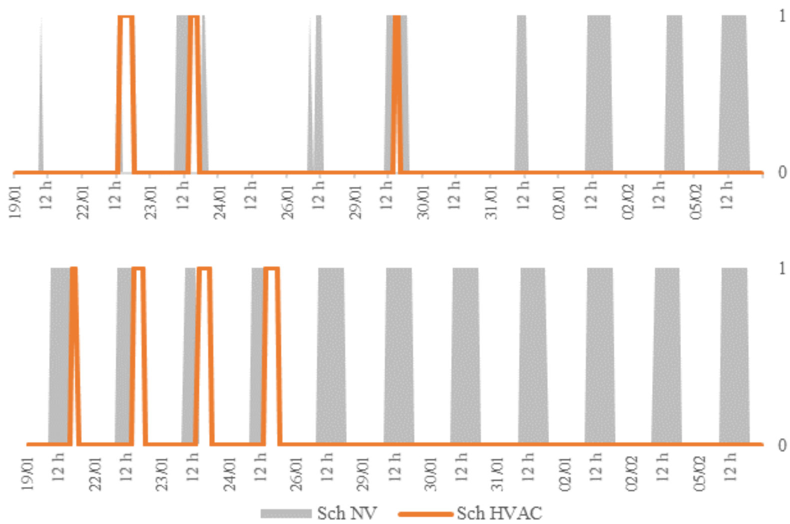


Figure 18. Indoor operative temperature comparison for the summer period – Office 1.



(a)

(b)

(a) Calibrated model (spontaneous control)
(b) Rule-based scenario (MM – Adaptive model 90%, AC setpoint 26 °C)

Figure 19. Comparison between natural ventilation (Sch NV) and air-conditioning (Sch HVAC) schedules of operation for the summer period – Office 1.

Results for offices 2 and 3 differ significantly from office 1 since the air-conditioning mode was the best match for WAC operation, for all seasons. Also, we can notice that the best match for T_{op} was either the air-conditioning system with a high setpoint temperature (25 °C or 26 °C) or the mixed-mode system. These results imply that more energy efficient operating modes could be used in order to obtain the same indoor thermal conditions as the ones obtained from real occupancy data. The results of the comparison between the monitored data and the rule-based scenarios for offices 2 and 3 can be seen in Appendix B.

3.3 Discussion

Results obtained through running multiple scenarios for different operating modes such as air-conditioning mode, mixed-mode and natural ventilation mode enabled us to perform an inter-office comparison for both T_{op} and WAC operation schedules. We observed that offices had different WAC operation for similar climatic conditions, indicating the existence of variable occupant behaviours. Whilst occupant choices from office 1 are mainly driven by the NV system along all seasons, occupant choices from offices 2 and 3 are mainly driven by the AC system (see Appendix B). This is unlikely to be driven by solar orientation as offices 1 and 2 share the same building and same orientation, though other aspects such as noise and air pollution could be at play. These findings support the idea that the imposition of a single behavioural pattern at design stage will not take into account the diversity of choices people exhibit in real buildings even under very similar conditions.

In some senses, the diversity of behaviour is unsurprising given the significant freedom available to occupants in these three offices. What is more interesting is the net result of these behaviours on the indoor temperature. In the previous section, we observed that despite the diversity of practices, the resulting indoor conditions are broadly the same. Indeed, the seasonally separated box plots (Figure 20) show that mean temperatures in the three offices in spring, summer and autumn are within 1.2 to 1.7 °C of each other. It is only in winter that the lower temperatures in Office 3 push this range to 2.5 °C. Indeed, the difference in means in summer, when one would expect the greatest demand for air-conditioning, is only 1.3 °C. This would suggest that the lowest energy-cost means of obtaining comfort (i.e. natural ventilation) would be practical, provided outdoor conditions such as temperature, noise and pollution are not unfavourable. This is in contrast to the assumption of much lower design indoor temperature in air-conditioned buildings. For example, using ISO 7730 (2005), assuming low air speeds and business attire results in the specification of indoor operative temperature of 23 ±2 °C.

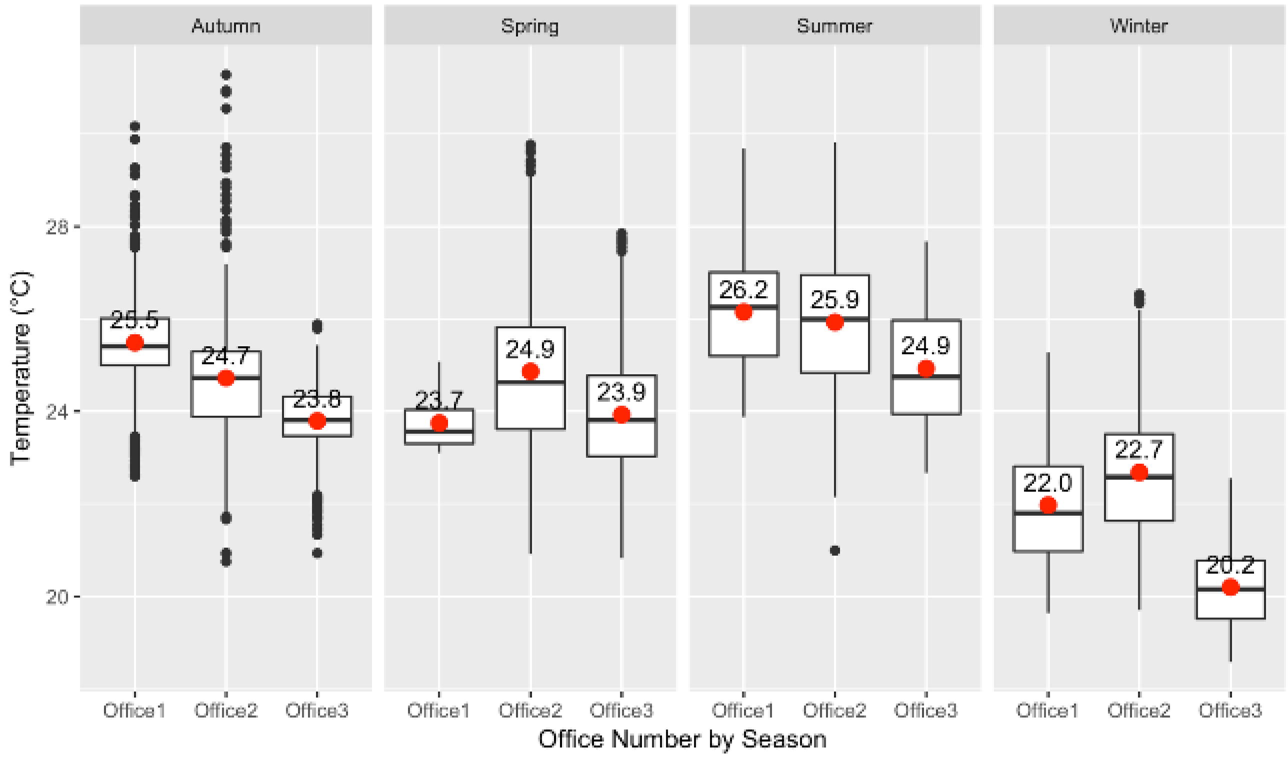
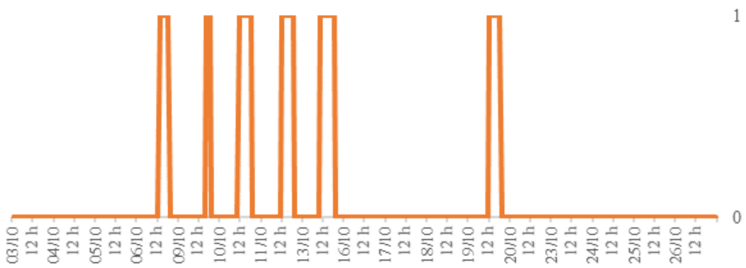
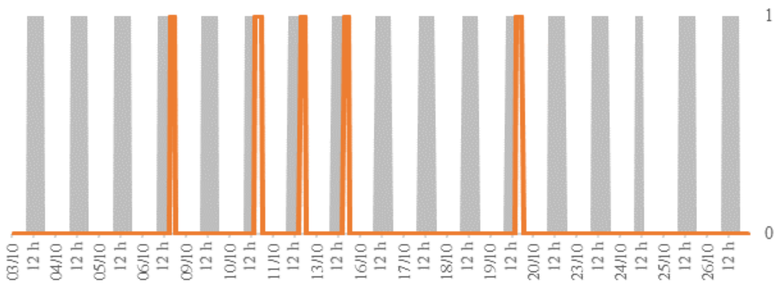


Figure 20: Box and whisker plots for indoor operative temperature in each office split by season. The boxes show the inter-quartile range with the median as a black bar. The red dots are the means.

The impact of this is illustrated using data for office 3. Here, the rule-based model that presented the best match with spontaneous control for WAC operation was an HVAC scenario (Figure 21a) and the rule-based model that presented the best match with spontaneous control for T_{op} (Figure 21b) was a mixed-mode scenario. The reduction in hours of operation of the air-conditioning system for the second case was 23 hours (-58%) over the measured period, whilst the indoor operative temperature maintained close to what occupants chose in the spontaneous control base model (the average temperature during occupied hours was 23.9 °C for the rule-based model and 23.8 °C for the spontaneous control based model). Therefore, similar operative temperature conditions as the ones found in the spontaneous control model could be obtained by making better use of the natural ventilation system and less use of the air-conditioning system. This means that a more energy efficient mode of operation could guarantee the same indoor thermal conditions, as expected by the occupants. It is important to highlight that both models have different operation premises: whilst the best match for WAC operation used the static thermal comfort model and an air-conditioning setpoint of 24 °C, the best match for T_{op} used the adaptive thermal comfort model and an air-conditioning setpoint of 26 °C.



(a)



(b)

(a) Rule-based scenario (Air-conditioning system – static model PPD < 10% and AC setpoint of 24 °C)
(b) Rule-based scenario (Mixed-mode system – adaptive model 90% acceptability and AC setpoint of 26 °C)

Figure 21. Comparison between natural ventilation (Sch NV) and air-conditioning (Sch HVAC) schedules of operation for the spring period – Office 3.

4. Conclusion

In this paper, we have investigated what operational mode occupants would select in a mixed-mode office room, considering that the air-conditioning and the natural ventilation systems can operate simultaneously and occupants have complete freedom of operation. We have undertaken field monitoring of three office rooms in order to infer occupant behaviour through binary observations of window (open/ close) and air-conditioning (on/ off) and calibrate a computer simulation model. Then, we compared the 'spontaneously controlled' base-case simulation model with rule-based models containing different operational scenarios and used a series of mathematical techniques to discover which of the created imagined temperature time series best matches reality.

We observed that the three offices had different window and air-conditioning operational habits but very similar indoor climatic conditions, indicating the existence of variable occupant behaviours that lead to similar thermal comfort conditions. Resulting indoor conditions were very alike and attainable through natural ventilation alone, suggesting that the free running mode or the mixed-mode could deliver indoor thermal comfort most of the time. Also, the air-conditioning system was mainly used with a high setpoint temperature (25 °C or 26 °C), which imply that more energy efficient operating modes could be used to obtain the same indoor thermal conditions as the ones obtained from real occupancy data.

These findings support the idea that the imposition of a single behavioural pattern at design stage does not take into account the diversity of choices people exhibit in real buildings. Also, despite the diversity of practices, the resulting indoor conditions are broadly the same, which suggests that a more energy efficient mode of operation could guarantee the same indoor thermal conditions, as expected by the occupants. Therefore, understanding occupant motivation and educating them on the impact of AC operation is needed to minimise energy use.

One possible reason that could prevent occupants from being more energy efficient is the low occupant interaction with the indoor environment, which characterizes them as "passive" users, according to the terminology used by Haldi and Robinson, 2009 and Gunay et al., 2013.

Research findings indicate that setting design goals considering occupant priority could change how buildings are being designed in practice and taking into account real occupant behaviour could set better directions on how to improve building design, in order to meet a greater diversity of users.

5. Acknowledgments

We wish to gratefully acknowledge funding from São Paulo Research Foundation (FAPESP), Brazil (grant number 2016/02734-0) / FAEPEX (grant number 2875/2018) as well as the Engineering and Physical Sciences Research Council (EPSRC) for the Zero Peak Energy Building Design for India (ZED-i, EP/R008612/1). Neither funder was involved in the design or execution of this study.

6. References

Andersen, R. K., Olesen, B. W. and Toftum, J. (2007) 'Simulation of the Effects of Occupant Behaviour on Indoor Climate and Energy Consumption Simulation of the Effects of Occupant Behaviour on Indoor Climate and Energy Consumption', in *Proceedings of Clima 2007: 9th REHVA world congress: WellBeing Indoors*.

ASHRAE. (2002) Guideline 14: Measurement of energy and demand savings. Technical Report ISSN 1049-894X, ASHRAE Standards Committee.

ASHRAE. (2017) ANSI/ASHRAE 55-2017 Thermal environmental conditions for human occupancy. Atlanta: American Society of Heating, Refrigerating and Air Conditioning Engineers, 2017.

Berndt, D.J., Clifford, J. (1994) Using dynamic time warping to find patterns in time series, in: Proc. of the AAAI Workshop on Knowledge Discovery in Databases, pp. 359–370.

Cassisi C, Montalto P, Pulvirenti A. (2012) Similarity measures and dimensionality reduction techniques for time series data mining. In: Advances in Data Mining, Knowledge Discovery and Applications, INTECH, chap 3.

Chen, Y. *et al.* (2019) 'Achieving natural ventilation potential in practice : Control schemes and levels of automation', *Applied Energy*. Elsevier, 235(October 2018), pp. 1141–1152. doi: 10.1016/j.apenergy.2018.11.016.

Corradini, A. (2001) Dynamic time warping for off-line recognition of a small gesture vocabulary. In Proceedings of the Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, Vancouver, BC, Canada, 13 July; pp. 82–89.

Craw S. (2017) Manhattan Distance. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning and Data Mining. Springer, Boston, MA.EnergyPlus. (2017) Auxiliary programs. U. S. Department of Energy.

Flourentzou, F., Van der Maas, J. and Roulet, C. A. (1998) 'Natural ventilation for passive cooling: measurement of discharge coefficients', *Energy and Buildings*, 27(3), pp. 283–292. doi: 10.1016/S0378-7788(97)00043-1.

Gill, J. (2008). Bayesian methods: a social and behavioral sciences approach (Second Edition.). Chapman and Hall/CRC.

Gunay, H. B., Brien, W. O. and Beausoleil-morrison, I. (2013) 'A critical review of observation studies, modeling , and simulation of adaptive occupant behaviors in offices', *Building and Environment*. Elsevier Ltd, 70, pp. 31–47. doi: 10.1016/j.buildenv.2013.07.020.

Haldi, F. and Robinson, D. (2009) 'Interactions with window openings by office occupants', *Building and Environment*. Elsevier Ltd, 44(12), pp. 2378–2395. doi: 10.1016/j.buildenv.2009.03.025.

Hoes, P. *et al.* (2009) 'User behavior in whole building simulation', 41, pp. 295–302. doi: 10.1016/j.enbuild.2008.09.008.

Hu, J. and Karava, P. (2014a) 'Model predictive control strategies for buildings with mixed-mode cooling', *Building and Environment*. Elsevier Ltd, 71, pp. 233–244. doi: 10.1016/j.buildenv.2013.09.005.

Hu, J. and Karava, P. (2014b) 'A state-space modeling approach and multi-level optimization algorithm for predictive control of multi-zone buildings with mixed-mode cooling', *Building and Environment*. Elsevier Ltd, 80, pp. 259–273. doi: 10.1016/j.buildenv.2014.05.003.

INMET. (2018) Instituto Nacional de Meteorologia. 'Weather station data'. Available at: http://www.inmet.gov.br/portal/index.php?r=estacoes/estacoesAutomaticas (Accessed: December 20th 2018).

ISO. (2005) International Organization for Standardization. Moderate thermal environments-determination of the PMV and PPD indices and specification of the conditions for the thermal comfort. ISO 7730. Switzerland.

Keogh, E.J., Kasetty, S. (2003) On the need for time series data mining benchmarks: a survey and empirical demonstration. Data Mining and Knowledge Discovery 7, 349–371.

Kingsland, J. and Hunter, B. (2011) ‘Using existing meters for web based real time energy monitoring’, in *CIBSE Technical Symposium*. Leicester UK, pp. 1–12.

Mccartney, K. J. and Nicol, J. F. (2002) ‘Developing an adaptive control algorithm for Europe’, *Energy and Buildings*, 34, pp. 623–635.

Myers, C.; Rabiner, L.; Rosenberg, A. (1980) Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. IEEE Trans. Acoust. Speech Signal Process. 28, 623–635.

Neves, L. O. et al. (2017). Envelope design of mixed-mode office buildings: theory versus practice. Proceedings… In: Passive and Low Energy Architecture. PLEA: Edimburgh.

Omer, A. M. (2009) ‘Energy use and environmental impacts : A general review’, *Journal of Renewable and Sustainable Energy*, 053101(September). doi: 10.1063/1.3220701.

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’, 39, pp. 823–836. doi: 10.1016/j.enbuild.2007.02.003.

Rijal, H. B., Humphreys, M. A. and Nicol, J. F. (2009) ‘Understanding occupant behaviour: The use of controls in mixed-mode office buildings’, *Building Research and Information*, 37(4), pp. 381–396. doi: 10.1080/09613210902904221.

Roetzel, A., Tsangrassoulis, A. and Dietrich, U. (2014) ‘Impact of building design and occupancy on office comfort and energy performance in different climates’, *Building and Environment*. Elsevier Ltd, 71, pp. 165–175. doi: 10.1016/j.buildenv.2013.10.001.

Rose, T. (2015) *The end of average*. Harper Collins.

Ross, S. (2017) ‘Logistic Regression’, Introductory Statistics (Fourth Edition).

Sakoe, H., Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. on Acoustics, Speech, and Language Processing 26, 43–50.

Sorgato, M. J., Melo, A. P. and Lamberts, R. (2016) ‘The effect of window opening ventilation control on residential building energy consumption’, *Energy & Buildings*. Elsevier B.V., 133, pp. 1–13. doi: 10.1016/j.enbuild.2016.09.059.

TPU – Tokyo Polytechnic University. (2018). TPU aerodynamic database. The School of Architecture and Wind Engineering. Available at: < http://wind.arch.t-kougei.ac.jp/system/eng/contents/code/tpu> (Accessed: January 29th 2019.

Vellei, M. *et al.* (2016) ‘The effect of real-time context-aware feedback on occupants ’ heating behaviour and thermal adaptation’, *Energy & Buildings*. Elsevier B.V., 123, pp. 179–191. doi: 10.1016/j.enbuild.2016.03.045.

Wang, L. and Greenberg, S. (2015) ‘Window operation and impacts on building energy consumption’, *Energy and Buildings*. Elsevier B.V., 92, pp. 313–321. doi: 10.1016/j.enbuild.2015.01.060.

Wang, X., Mueen, A., Ding, H., Trajcevski, G., Scheuermann, P., Keogh, E.J. (2012). Experimental comparison of representation methods and distance measures for time series data. Data Mining and Knowledge Discovery 26, 2, 275-309.

Wei L., Keogh E.J. (2006). Semi-supervised time series classification. In: KDD’06: proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, New York, NY, USA, 748–753.

Xing, Z., Pei, J., Yu, P.S. (2011). Early classification on time series. Knowledge and Information Systems 31, 105-127.

Zhao, J. *et al.* (2016) ‘Occupant-oriented mixed-mode EnergyPlus predictive control simulation’, *Energy and Buildings*. Elsevier B.V., 117, pp. 362–371. doi: 10.1016/j.enbuild.2015.09.027.