

Paper submission

Modelling Uncertainty in District Energy Simulations by Stochastic Residential Occupant Behaviour

R. Baetens^{a,b*} and D. Saelens^{a,b}

^a*Building Physics Section, Department of Civil Engineering, KU Leuven, Leuven (Belgium);*

^b*EnergyVille, Dennenstraat 7, 3600 Genk (Belgium)*

(v0.1 released April 9, 2015)

Occupant behaviour has since long been of main interest in the domain of building energy savings and indoor air quality; and its importance is recognized by its wide coverage in literature. In the recent developments of detailed transient building energy simulations, including the occupant behaviour as boundary condition for the thermal comfort and system efficiency calculations has been a major research topic given its significant impact. Simultaneous growing interest in district energy simulations raises similar questions at the aggregate level, where upscaling from the building to an aggregate neighborhood level at the spatial scale of a low-voltage feeder results in a natural regression to the mean lowering uncertainty compared to the level of the household.

The presented work starts with the description of **StROBe**, a stochastic residential occupant behaviour for district energy simulations; integrating the modelling of receptacle loads, internal heat gains, thermostat settings and hot water tapings based on occupancy and activity prerequisites. Given this model, the uncertainty for district energy simulations is addressed. The epistemic uncertainties are elaborated first comparing model results with reference values and denoting local disaggregation of demographic statistics as possible main hiatus of general modelling methods for building energy occupant behaviour used at the neighborhood level. To conclude, the aleatory uncertainty caused by stochastic residential occupant behaviour in integrated district energy simulations are quantified. Here, the expected value of the objective functions have to a large extend the same minimizers as the measures of the proposed robustness. As such, optimizing an objective value for its expected value generally seems to result in an optimum *near* the optimum of robustness. However, 95 percent of the observed objectives lay between 0.81 and 1.6 times the expected value for a feeder larger than 10 houses and between 0.88 and 1.3 times the expected value for more than 20 houses denoting an overall ‘rather small’ uncertainty on the possible objective functions caused by user behaviour. Furthermore, we show that the design of the building energy system has its impact on the robustness of the objective criteria and it could thus be minimized as part of an optimization exercise.

Keywords: District simulation; Stochastic modelling; Occupant behaviour; Uncertainty.

1. Introduction

Occupant behaviour has since long been of main interest in the domain of building energy savings and indoor air quality; and its importance is recognized by its wide coverage in Annex 5 (1979), Annex 8 (1984) and Annex 20 (1988) of the International Energy Agency (IEA) Energy in Buildings and Communities (EBC) Programme focussing on ventilation habits, and in Task 26 (1998) of the IEA Solar Heating and Cooling (SHC) Programme focussing on hot water consumption. In the recent developments of detailed transient building energy simulations, including the - possibly irrational - occupant behaviour as boundary condition for the thermal comfort and system efficiency calculations has been a major research topic given its significant impact. Here, the recent Annex 66 (2014) of the IEA-EBC Programme aims to harmonize research on stochastic modelling of energy related occupant behaviour in buildings at over 20 research

*Corresponding author. Email: ruben.baetens@bwk.kuleuven.be

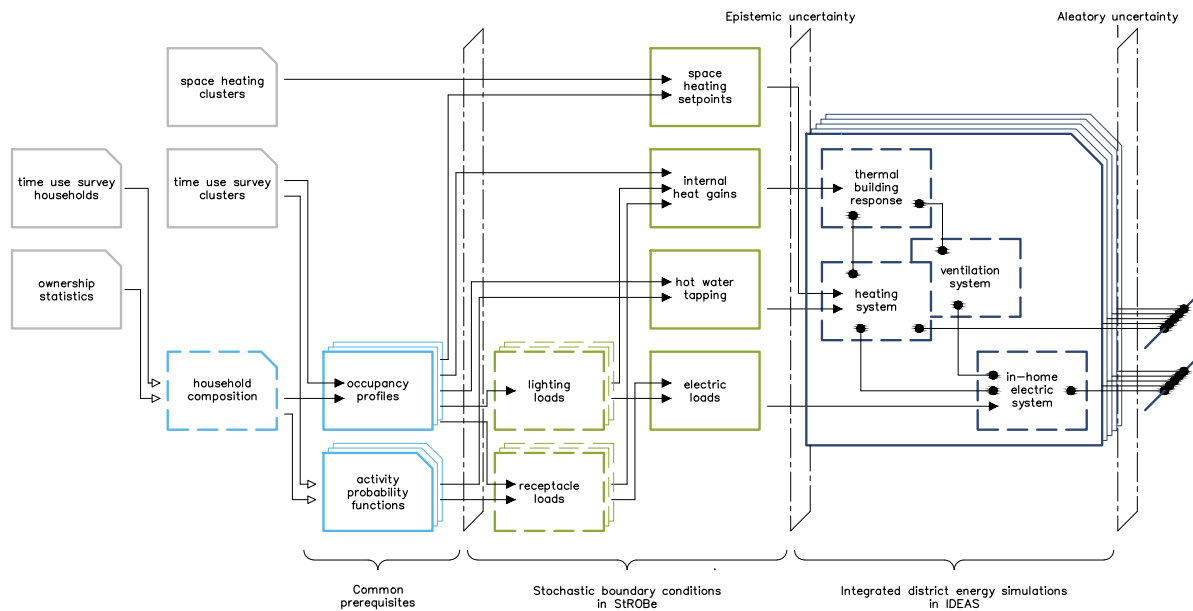


Figure 1. General overview of the implemented algorithms in *StROBe*. The framework is found on the stochastic determination of occupancy and activity chains as prerequisites; based on the clustered time use survey data. Given the generated occupancy chains and activity probabilities, the profiles for the receptacle load of appliances and lighting, for the hot water tapplings, the internal heat gains, and the space heating set points are determined.

groups. Simultaneous with these developments in building energy simulations, the interest grows in district energy simulations; due to the possible externalities of rooftop-mounted photovoltaic systems and electricity-driven heat pumps for space heating and domestic hot water.

Given the impact of thermostat adjustments and hot water tapping on the simultaneity of heat pump load profiles and the additional aggregation of receptacle loads at neighbourhood level, the presented work aims to estimate the uncertainty in district energy simulations regarding occupant behaviour within the presented context. Here, up-scaling from the building to an aggregate neighbourhood level at the spatial scale of a low-voltage feeder results in a natural regression to the mean lowering uncertainty compared to the level of the household.

To address this, the presented work starts by (a) the description of a developed stochastic residential occupant behaviour model *StROBe* (short for ‘Stochastic Residential Occupant Behaviour’) for district energy simulations; integrating the modelling of receptacle loads, internal heat gains, thermostat settings and hot water tapping based on occupancy and activity prerequisites. Given this model, the uncertainty for district energy simulations is addressed: epistemic uncertainties (b) are elaborated first, comparing model results with reference values and denoting possible hiatus of general modelling methods for building energy occupant behaviour used at the neighbourhood level. At last, the aleatory uncertainty (c) is estimated describing the uncertainty in district energy simulations inherent to the non-deterministic occupant behaviour.

An overall overview of the followed methodology is given in Figure 1; making distinction of the used data sources used to model occupancy profiles and activity probabilities as prior common prerequisites for the modelling of the stochastic phenomena in occupant behaviour with physical impact on (building and) district energy simulations.

2. Prior considerations

Without elaborating on a full-length literature review of related work on occupant behaviour modelling for building energy simulations, we will precede the description of developed occupant model by a set of literature-based modelling assumptions.

Modelling approach. Review of the extensive related literature on (a) receptacle load modelling, on (b) modelling the interaction between occupants and building systems providing indoor environmental quality, and on (c) hot water tap load modelling allows us to draw the first methodology outlines for a correct representation and modelling of the pervasive system in residential district energy simulations:

- A bottom-up model using occupancies and activities of the household individuals as prerequisites for modelling the heating settings, plug loads, internal heat gains and domestic hot water tapping time series is preferred. The former is introduced in early work of by C.WALKER (1982) and A.CAPASSO (1993) *et al.* (Walker, 1982; Walker & Pokoski, 1985; Capasso *et al.*, 1993, 1994), and is adopted by a majority of load models (Jordan & Vajen, 2001b,a; Spur *et al.*, 2006; Widén *et al.*, 2009a,b; Widén & Wäckelgård, 2010; Widén *et al.*, 2012; Stokes *et al.*, 2004; Richardson *et al.*, 2008, 2009, 2010; Richardson, 2010; Wilke *et al.*, 2013; ?; Aerts, 2015).
- Demographic and socio-economic survey data are used in preference of load data as basis for modelling. As such, the model allows cross-sectional and longitudinal analyses of behavioural aspects. The latter approach is adopted in all previously reference models, while load-based comprehensive models are scarce (Mansouri *et al.*, 1996; Paatero & Lund, 2006; Armstrong *et al.*, 2009).

Additionally, though not indisputably proven in latest model developments regarding occupant behaviour, it is the author's opinion that:

- Existing comprehensive stochastic models for window opening and solar shading control in a residential environment are not sufficiently described or validated, in contrast to the models for offices (Warren & Parkins, 1984; Fritsch *et al.*, 1990; Johnson & Long, 2005; Rijal *et al.*, 2008, 2011; Yun & Steemers, 2008; Yun *et al.*, 2009; Yun & Steemers, 2010; Haldi & Robinson, 2011; Schweiker *et al.*, 2012; Andersen *et al.*, 2013; Yun *et al.*, 2009; Guerra-santin & Itard, 2010; Haldi & Robinson, 2011).
- Clustering and survival analysis are preferably combined for modelling the variables which' state values duration sequences influence most the system efficiencies (Wilke *et al.*, 2013; Aerts, 2015).
- First-order Markov chains are preferably only used for modelling the variables which' duration sequences have little influence on system efficiencies.
- A time resolution of ten minutes or less is recommended (Widén *et al.*, 2010; Baetens *et al.*, 2011; Groscurth *et al.*, 1995).

Remarkably, the earliest model of C.WALKER (1982) yet includes majority of the quoted preferred principles, of which many of them were abandoned in successive research (Walker, 1982). We will adopt his model philosophy, updated with the latest developments and available data. As such, the Python StROBe Package (short for 'Stochastic Residential Occupant Behaviour') is developed as a modular framework in Python 2.7 and describes the occupant behaviour for residential (building and) district energy simulations based on the cited related work and national surveys.

Additional assumptions. Prior to the modelling of the stochastic nature of occupant behaviour, we will state all assumptions shared by two or more of the related sub-systems. As such, prior, we will assume that:

- Occupant behaviour is spatially homogeneous with respect to the national demography; and the technical systems and occupant behaviour are spatially homogeneous with respect to the weather.
- Structural demographic changes can be neglected during a single simulation.
- Households and household members behave independently of each other, and indepen-

dently of the building and district energy systems' state.

The first premise means that 'people living in a single neighbourhood do not predominantly tend to fall in one or two consumer types', and that appliance ownerships, demography and dwelling type can be described independently of each other, but studies on this assumption are not unequivocal (Gardiner & Manson, 1994; Vanneste *et al.*, 2001). The premise additionally means that system interactions on environmental data do not depend on the location within the system. This may introduce an overestimation of the simultaneity of climate-induced loads, *e.g.* on the lighting load or photovoltaic output on cloudy days. The true spatial inhomogeneity can however not be modelled due to a lack of data, but its effect is assumed to be *negligible* due to the small spatial scale of interest in the research. Considering the last two premises; the stated research question is generally addressed by performing energy simulations representing the duration of one year. The structural changes of occupant behaviour in such a period are rather small, and negligible for energy system evaluations. The changes in a long-term perspective are more pronounced, *e.g.* the evolution to more energy efficient appliances up to changing demographics in time.

3. The StROBe Package; a residential human behaviour model

Given these assumptions, the effective implementation of the stochastic nature of occupant behaviour is a research question on its own and will therefore be elaborated on in the following sections.

An overview of StROBe is given in Figure 1 on page 2. In a first stage, available survey data is clustered and used for household composition. Following, the household proclivities are determined, and occupancy and activity chains are generated based on survival models. In a third and last stage, the effective occupant behaviour concerning receptacle loads, thermostat settings and domestic hot water tapping is modelled.

3.1 Subspace clustering

The majority of the used statistics in this work are based on two surveys, *i.e.* the decennial Belgian Time-Use Survey and Household Budget Survey (from now on referenced as BeTUS'05 and BeHBS'05) collected in 2005 by the Directorate-general Statistics and Economic Information. (Glorieux *et al.*, 2008b,a; Glorieux & Minnen, 2008; Heylen *et al.*, 2007) The BeTUS'05 and BeHBS'05 datasets relate to a population of 6400 individuals from 3474 households, who completed questionnaires describing the chronological course of activities in their diaries in 10 min increments throughout 24 hours, starting and ending at 4:00 AM. These activities are assigned from an extensive list of 272 different activity types.

The physical quantities of interest in BeTUS'05 are the probabilities to perform different activities during the day, and the possible cross-correlations with all demographic parameters. At the resolution of the individual, observed occupancy and activity chains are respectively characterised by the discrete-time data series $o^{(x)}[n]$ and $a^{(x)}[n]$ of the individual x over time bins n . Here, three possible occupancy states Ω_o are determined for each time step, *i.e.* being *awake at home*, being *asleep at home*, and being *absent*.

As depicted by D.AERTS *et al.*, subgroup population clustering is required for model validity allocated at the single building scale (Aerts, 2015). In order to predict occupancy and activity chains so that individual behaviour or possible sequential activity occurrence can be captured, the aggregated BeTUS'05 data is clustered in seven partitions β_j by D.AERTS *et al.* by means of agglomerative hierarchical clustering $C(\cdot)$ based on the Levenshtein distance as metric between

Table 1. BeTUS'05 coverage of individual employment types for all partitions β_j if applied for weekdays, representing household members behaving according to a typical occupancy profile j . The bold coverages denote an over-representation in β_j in comparison to the overall population, as such being dominated by these clusters.

	β_1	β_2	β_3	β_4	β_5	β_6	β_7	$U(\beta)$
Full-time employment	0.657	0.413	0.323	0.177	0.142	0.138	0.158	0.294
Part-time employment	.085	.130	.097	.098	.066	.107	.084	.079
Being retired or unemployed	.067	.150	.368	.599	.642	.679	.712	.296
Minor	.192	.308	.213	.084	.194	.076	.046	.145

all observed occupancy state sequences $o^{(x)}[n]$ (Aerts, 2015)

$$\{\beta_j\}_{1:7} \triangleq C_{\text{lev}} \left(\left\{ o^{(x)}[n] \right\}_s \right) : \text{BeTUS}'05 \quad (1)$$

with $\beta_j \subseteq \text{BeTUS}'05$ Each partition β_j represent household members behaving according to a typical occupancy profile j . All partitions together cover 81.4 percent of the original dataset for weekdays, 76.6 percent for Saturdays and 80.0 percent for Sundays as shown in Table 1 on page 5.

3.2 Household composition & cluster allocation

The household composition must be determined before generating the household behavioural variables. With respect to StROBe, it is required to know how many people live in the dwelling, to which cluster β_j they can be allocated, and which appliances they own.

Household composition is determined based on the compositions found in BeTUS'05, *i.e.* a random household composition is each time taken from the BeTUS'05 database determined by the professional life stages of the individuals who are part of it. Here, distinction is made between 'minors', 'full-time' employed individuals, 'part-time' employed individuals and individuals being 'idle'. The latter is later on divided into 'unemployed' and 'retired' individual as will be required for determining thermostat settings. As such, the household is at this point determined as *e.g.* $\{full-time, unemployed, minor\}$. Given these professional life stages of the individuals, each individual is allocated to a cluster β_j based on the known representations of a certain employment as shown in Table 1 on page 5. As such, the household is at this point determined as *e.g.* $\{\beta_1, \beta_6, \beta_2\}$.

The owned appliances are determined by the average ownerships found in BeHBS'05, *i.e.* the ownership is evaluated for each household based on the observed probabilities.

3.3 Occupancy and activity prerequisites

Based on the clustered datasets, the occupancy chains and the activities of all household members will be modelled as a common prerequisite for the modelling of the variables with physical impact. This improves their resulting auto- and cross-correlations.

3.3.1 Member occupancy

The occupancy $o^{(x)}[n]$ of each household member x is explicitly modelled as a common prerequisite for the physical behavioural variables. The generation of the behavioural profile assumes members to behave independently of each other¹ and consists of three steps, *i.e.* defining the start state, and successively evaluating a *survival time density* and an *event density* to determine the time n at which the next event will take place and to which state the occupancy will change.

¹Independent behaviour is an assumption made in a majority of behavioural research, although correlated behaviour can be suspected (Tanimoto *et al.*, 2008b; Widén & Wäckelgård, 2010; Richardson, 2010). The use of clusters slightly improves the current approaches, while combining individual probabilities with household proclivities is an active research domain which might improve the generation of $o^{(x)}[n]$ in the near future (Wilke *et al.*, 2013; Aerts, 2015).

As such, the resulting occupancy chain $o^{(x)}[n]$ is described as

$$o^{(x)}[n] = \left\{ \left(o_{n_0}^{(x)}, l_1 \right), \left(o_{n_0+l_1}^{(x)}, l_2 \right), \dots, \left(o_N^{(x)} \right) \right\} \quad (2)$$

where $o_n^{(x)} \in \Omega_o$ is the occupancy state of individual x at time n , Ω_o is the occupancy state space and l_n is the survival time for which the state $o_n^{(x)}$ remains equal to its previous value $o_{n-1}^{(x)}$.

The occupancy state $o_{n_0}^{(x)}$ at time n_0 is directly determined for each x as

$$o_{n_0}^{(x)} = X \left(\Pi \left(\{ o_{n_0}^{(x')} \}^{\beta^{(x)}} \right) \right) : x \propto \beta^{(x)} \subseteq \text{BeTUS}'05 \quad (3)$$

where $\{ o_{n_0}^{(x')} \}^{\beta^{(x)}}$ is the set of observed occupancy states at n_0 in $\beta^{(x)} \subseteq \text{BeTUS}'05$ to which the individual x was allocated, $\Pi(\cdot)$ describes the distribution of the given variable and $X(\cdot)$ denotes a random value from the given distribution function.

Once $o_{n_0}^{(x)}$ defined, the stochastic process is ruled by a heterogeneous discrete-time Markov chain with alternate transitions in the state space and in time space. These transitions are determined by an *event density* and a *survival time density*. We define the event density $O_n^{(x)}$ and lead time density $L_n^{(x)}$ at time n as

$$O_n^{(x)} \triangleq \Pi \left(o_n^{(x)} | o_{n-1}^{(x)}, n \right) : \forall n \in \mathbb{N}, \forall o_n^{(x)} \in \Omega_o \setminus o_{n-1}^{(x)} \quad (4)$$

$$L_n^{(x)} \triangleq \Pi \left(o_{n+k}^{(x)} = o_n^{(x)} | o_n^{(x)}, n \right) : \forall n, k \in \mathbb{N} \quad (5)$$

whose image $O_n^{(x)}$ is a complete description of the cumulative distribution function of $o_n^{(x)}$ conditional to $o_{n-1}^{(x)}$ and the presupposition that an event occurs at time n ; and whose image $L_n^{(x)}$ is a complete description of the cumulative distribution function of the lead time k conditional to the presupposition that $o_{n+k}^{(x)} = o_n^{(x)}$. Both event and survival time densities are defined at clock-time τ_{30} but will be evaluated at τ_{10} in the algorithm when profiles are generated, and the parametrization set of day-types is restricted to *Weekday*, *Saturday* and *Sunday*. An example day-profile of the observed $O^{(x)}[n_d]$ and $L_{n_d}^{(x)}$ is given in Figure 2 for two arbitrary clusters β_j .

The resulting final occupancy chain $o^{(x)}[n]$ for x as defined in Equation 2 is derived by successively determining the start state $o_{n_0}^{(x)}$ as in Equation 3, obtaining its survival time $l_0^{(x)} = X(L_{n_0}^{(x)})$ denoting the time after which a change in occupancy state will occur, proceed to $n_0^{(x)} + l_0$, determine the new state $o_{n_0+l_0}^{(x)} = X(O_{n_0+l_0}^{(x)})$ and repeat the last three steps until the end of the simulation period. The generation of a one-year occupancy profile for IDEAS 0.3 usually consists of a one-week profile repeated 52 times; as creating a full-year stochastic profile shows an increased regression to the mean.

3.3.2 Activity proclivities

The exerted activities are not modelled explicitly but only described by their statistics, in contrast to the occupancies. As such, we recall C.WALKERS' proclivity function often referred to as I.RICHARDSONS' static activity function and formally defined as

$$g_{j,n}^{(x)} \triangleq p \left(a_n^{(x)} = j | o_n^{(x)} \right) : \forall n \in \mathbb{N}, \forall j \in \Omega_a \quad (6)$$

whose image $g_{j,n}^{(x)}$ denotes the likelihood that this individual x is engaged in the depicted activity j at time n conditional to $o_n^{(x)}$ (Walker, 1982; Richardson *et al.*, 2010). Similar to the event density and survival time density, $g_{j,n}^{(x)}$ is defined at clock-time τ_{30} but evaluated at τ_{10} in the

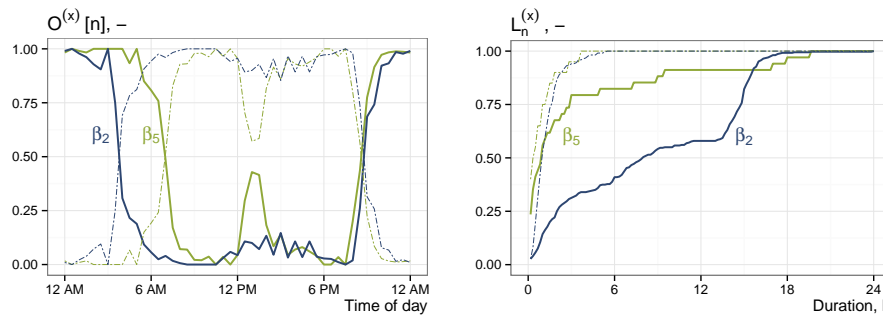


Figure 2. Example profile of event density $\hat{O}^{(x)}[n_d]$ (left) and survival time density $\hat{L}_{n_d}^{(x)}$ (right) for clusters β_2 and β_5 . The given event density $\hat{L}_{n_d}^{(x)}$ denotes the probability that o_{n+k} is ‘home and awake’ (bold) or ‘away’ (dashed) given that o_n equals ‘home and sleeping’. The given $\hat{L}_{n_d}^{(x)}$ denotes the survival time density at 6:00 AM for a transition from ‘home and sleeping’ to ‘home and awake’ (bold) or ‘away’ (dashed).

algorithm when profiles are generated, and the parametrization set of day-types \mathcal{S} is restricted to *weekday*, *Saturday* and *Sunday*.

3.4 Temporal pervasive space of human behaviour

The defined approach on household proclivities result in occupancy chains $o^{(x)}[n]$ and the related static activity likelihoods $\hat{g}_j^{(x)}[n]$ for each household member x and all time bins n . Both variables have no first-order impact on the energy simulations, but are prerequisites to determine the correlated space heating settings, the plug loads for appliances and lighting, internal heat gains, and domestic hot water tapping flows which influence all commodity flows.

3.4.1 Receptacle loads

The first conclusive model in **StROBe** models the receptacle loads and is divided in two sections, *i.e.* modelling of the residential appliances loads $P_\alpha[n]$ and the lighting loads $P_\lambda[n]$ respectively.

Residential appliance loads The implemented appliance load model adopts the approach of I. RICHARDSON *et al.* (Richardson *et al.*, 2010), determining the load profile $P_\alpha[n]$ for appliances α based on the generated occupancy chains $\{o^{(x)}[n]\}_h$, appliance power characteristics, appliance-activity assignment, static activity likelihoods which correspond to the proclivity functions $g_j^{(x)}[n]$ and the determination of a calibration scalar. In general, domestic appliances have two states as they may either be *standby* or *on* reducing the calculation of $P_\alpha[n]$ to defining the switching-on and -off events.

For each appliance α and member x , the probability of switch-on $p_\alpha^{(x)}(t)$ is defined as

$$p_{\alpha,n}^{(x)} = \begin{cases} g_{a_\alpha,n}^{(x)} c_\alpha & \text{if } o_n^{(x)} = 1 \\ c_\alpha & \text{if } (a_\alpha = o_n^{(x)}) \text{ or } (a_\alpha = \emptyset) \end{cases} \quad (7)$$

with $c_\alpha \leq 1$ a calibration scalar, $a_\alpha \in \Omega_a$ the activity to which the use of appliance α is related, and $g_{a_\alpha,n}^{(x)}$ the likelihood of activity a_α related to appliance α for $x \in \beta^{(x)}$; $g_{a_\alpha}^{(x)}$ is determined at τ_{10} being the time resolution of the generated $o^{(x)}[n]_h$ but evaluated at τ_1 in the algorithm. The calibration scalar c_α determines the average number of times the respective appliance α is used in a year, and is determined iteratively after hundred of algorithm runs per iteration. Here, c_α is required as there is no one-to-one link between the use of appliances and the activities. Once a switch-on event is detected as $X(p_{\alpha,n}^{(x)})$ for appliance α by x , the duration δ_α of the *on*-state

ν_α is determined as

$$\delta_\alpha = X(\mathcal{N}(\mu_\alpha, \sigma_\alpha^2)) : \sigma_\alpha \triangleq \mu_\alpha/10 \quad (8)$$

with μ_α the appliance' mean duration of the *on*-state, and where-after the procedure is repeated for the entire occupancy chain.

Once all events are defined and all durations determined, the resulting load profile $P_\alpha[n]$ of α can be easily determined as $\nu_\alpha[n]P_{\nu_\alpha}$ with $\nu_\alpha[n]$ the generated state of α and P_{ν_α} the characteristics power of α . Given the load profile, the resulting internal convective and longwave heat gains are straightforwardly determined as $\dot{Q}_{a,g}^{(\alpha)} \triangleq f_{a,\alpha}P_\alpha[n]$ and $\dot{Q}_{lw,g}^{(\alpha)} \triangleq f_{lw,\alpha}P_\alpha[n]$ respectively based on the simplification that appliances do not have a thermal capacity. Here, $f_{a,\alpha}$ and $f_{lw,\alpha}$ are the convective and radiative fraction of the heat gain for appliance α . Both factors do not necessarily sum up to unity to include heat generation in appliances which are not transferred to the building zone, *e.g.* appliances for cooking and laundry.

Residential lighting loads A distinct model is used for defining the receptacle lighting loads $P_\lambda[n]$ for the lighting fixtures λ . Here, the approach of J.WIDEN *et al.* (2009) for simulating the use of lighting for residential buildings is chosen in favour of the model of I.RICHARDSON *et al.* as it does not require the nominal power of all fixtures separately (Widén *et al.*, 2009a; Richardson *et al.*, 2010). The approach starts by defining a theoretical *desired* power level $P_{\lambda,id}[n]$ as

$$P_{\lambda,id}[n] \triangleq u_{ab}[n]P_{\lambda,ab} + u_{ac}[n]P_{\lambda,ac}[n] + u_{ia}[n]P_{\lambda,ia} \quad (9)$$

where $P_{\lambda,ab}$ and $P_{\lambda,ia}$ are default electric loads for lighting in case the most-active occupant is *absent* or *inactive* respectively and where the functions $u_i[n]$ are step functions stating $o^{(x)}[n] = i$; while $P_{\lambda,ac}[n]$ is an ideal load required to meet global comfort when *active* given by

$$P_{\lambda,ac}[n] = \begin{cases} P_{\min}\psi(t) + P_{\max}(1 - \psi(t)) & \text{if } E_e^{(h)}(t) \leq E_{\lim} \\ P_{\min} & \text{if } E_e^{(h)}(t) > E_{\lim} \end{cases} \quad (10)$$

where $\psi(t) \triangleq E_e^{(h)}(t)/E_{\lim}$ as function of the global irradiation levels $E_e^{(h)}(t)$ on the horizontal surface based on the assumption that the illuminance on a surface is linearly dependent on the irradiance and where E_{\lim} is a threshold value. The effective lighting load $P_\lambda[n]$ is defined by and incremental power ΔP_λ with a constant probability p_λ reducing the difference between the effective load for lighting $P_{\lambda,n}$ and the instantaneous ideal load $P_{\lambda,id,n}$ at evaluation in each time step for $o_n^{(x)} = 1$ denoting the slow response of occupants to changing light conditions. Given the load profile, the resulting internal heat gains can be straightforwardly determined as $\dot{Q}_{a,g}^{(\lambda)} \triangleq f_{a,\lambda}P_\lambda[n]$ and $\dot{Q}_{lw,g}^{(\lambda)} \triangleq f_{lw,\lambda}P_\lambda[n]$ respectively; with $f_{a,\lambda}$ and $f_{lw,\lambda}$ the convective and radiative fraction of heat gain for appliance λ .

3.4.2 Internal heat gains

Once the occupancy chains $o^{(x)}[n]$ of all household individuals x and the receptacle load chains $P_\alpha[n]$ and $P_\lambda[n]$ of all appliances and lighting fixtures are determined, the overall convective and longwave internal heat gain profiles $\dot{Q}_{a,g}[n]$ and $\dot{Q}_{lw,g}[n]$ can be defined straightforwardly as the sums $\dot{Q}_{a,g}^{(\alpha)}[n] + \dot{Q}_{a,g}^{(\lambda)}[n] + \sum_x f_a^{(x)}Q_M(o^{(x)}[n])$ and $\dot{Q}_{lw,g}^{(\alpha)}[n] + \dot{Q}_{lw,g}^{(\lambda)}[n] + \sum_x f_{lw}^{(x)}Q_M(o^{(x)}[n])$ respectively, with $f_{a,x}$ and $f_{lw,x}$ respectively the convective and radiative fraction of heat gain for humans and $Q_M(o^{(x)})$ the metabolic heat production of x in state $o^{(x)}$.

3.4.3 Space heating settings

There are no *comprehensive* stochastic models known focussing on the space heating settings based on prerequisite occupancy and activity chains. The recent survey data analysis method of K.LEIDELMEIJER *et al.* denoting clustering of the space heating comfort desires, patterns and heated spaces, and the resulting findings closely match a $o^{(x)}[n]$ -prerequisite approach when reformulated for modelling and will be adopted here (Leidelmeijer & Van Grieken, 2005). As such, the used statistics in this section are based on the Dutch Qualitative House Registration (from now on referenced as DuQHR'00) collected in 2000 by the Directorate-general Housing of the Ministry of Housing, Spatial Planning and the Environment VRO/DG Wonen (2000). The dataset relates to a population of 15000 dwellings, who completed questionnaires describing their heating and ventilation behaviour.

To analyse the temporal behavioural patterns, DuQHR'00 is clustered in seven partitions φ_j by K.LEIDELMEIJER *et al.* based on the observed space heating state chains $T_{sh}^{(s_d)}[n]$ of the living area s_d (Leidelmeijer & Van Grieken, 2005)

$$\{\varphi_j\}_{1:7} \triangleq C \left(\left\{ T_{sh}^{(s_d)}[n] \right\}_s \right) : \text{DuQHR}'00 \quad (11)$$

with $\varphi_j \subseteq \text{DuQHR}'00$. Each partition φ_j represent households heating according to a typical pattern j , *e.g.* partition φ_3 covers the households in DuQHR'00 with a typical space heating profile heating the day-zone just below 20°C all day long, increasing their set-point in the evening while reducing it at night to 15°C. All partitions together cover 98 percent of the original dataset as shown in Table 2 on page 10; though partition φ_1 covering 4 percent of DuQHR'00 is removed for modelling new neighbourhoods denoting the households which never heat their dwelling. The set $\{\varphi_j\}_{1:7} \subseteq \text{DuQHR}'00$ covers two distinct aspects at the same time, *i.e.* the characterisation of the heating pattern in time and of the desired comfort levels. If we only consider the temporal behaviour in φ_j replacing $T_{sh}^{(s_d)}[n]$ by a function $u_{sh}^{(s_d)}[n]$ with state space *high*, *medium* and *low*, the partitioning φ_j matches the earlier found partitioning $\{\beta_j\}_{1:7} \subseteq \text{BeTUS}'05$ on occupancy with state space *awake at home*, 'asleep at home' and *absent*. We therefore conclude that the space heating set-point model can start from the generated occupancy chains $\{o^{(x)}[n]\}_h$ of all members x in household h . The model is as such defined based on the instantaneous *most-active* occupancy state of all members x . Given the stated $o^{(x)}$ -prerequisite, the probabilistic allocation of household h to the cluster $h \propto \varphi^{(h)}$ is based on the given un-parametrized coverages in DuQHR'00 and translated to the parameter set $T_{sh,u_1}^{(s_d)}(\Omega_o, \varphi^{(h)})$ denoting the space heating set-points for the living area as function of the occupancy state of the most-active member.

In order to analyse the spatial behavioural patterns, DuQHR'00 was again clustered by LEIDELMEIJER & VAN GRIEKEN based on the observed heated spaces $\{s'_{sh}\}_h$. (Leidelmeijer & Van Grieken, 2005) As such, DuQHR'00 is clustered in six partitions σ_j

$$\{\sigma_j\}_{1:6} \triangleq C \left(\{s'_{sh}\}_h \right) : \text{DuQHR}'00 \quad (12)$$

with $\sigma_j \subseteq \text{DuQHR}'00$. Each partition σ_j represent households heating according to a typical pattern j , *e.g.* partition φ_3 covers the households in DuQHR'00 which heat the day-zone, kitchen and bathroom according to the pattern observed in the living area s_d . For convenient handling of multi-zone building energy simulations, the state space of DuQHR'00 is converted to the smaller state space with *day-zone*, *night-zone* and *bathroom* by dealing the living area, kitchen and scullery equal. The coverages of the subgroups on heated spaces $\{\sigma_j\}_{1:6}$ are not quantified but only expressed qualitatively in relation to the subgroups on heating patterns $\{\varphi_j\}_{1:7}$, *i.e.* the most common combination for each cluster σ_j with φ_j is given. As such we can only deterministically allocate $h \propto \sigma^{(h)}$ given the allocation $h \propto \varphi^{(h)}$ as shown in Table 2 on page 10.

Table 2. DuQHR'00 coverage of household heating behaviour patterns for all partitions φ_j representing household members behaving according to a typical settings j with set point T_{sh} as function of the occupancy state for the heated spaces $\{s_{sh}\}_h$. The bold coverages denote an over-representation in φ_j in comparison to the overall population, as such being dominated by these clusters. The possible heated spaces are the dayzone (d), bathroom (b) and nightzone (n).

	φ_2	φ_3	φ_4	φ_5	φ_6	φ_7	$U(\varphi)$
Coverage	0.35	0.16	0.08	0.11	0.05	0.20	
T_{sh} if home & awake	20.0°C	18.5°C	20.0°C	20.0°C	21.0°C	21.5°C	19.11°C
T_{sh} if home & asleep	19.5°C	18.5°C	19.5°C	15.0°C	21.0°C	21.5°C	18.35°C
T_{sh} if absent	15.0°C	15.0°C	11.0°C	14.0°C	20.5°C	15.5°C	13.13°C
Heated spaces $\{s_{sh}\}_h$	{d,b,n}	{d,b}	{d,n}	{d}	{d,b,n}	{d,b}	

3.4.4 Hot water redraws

The implemented hot water tap model adopts the exact same approach as the receptacle loads, determining the load profile $\dot{m}_\gamma[n]$ for taps γ based on the generated occupancy chains $\{o^{(x)}[n]\}_h$, tap flow characteristics, tap-activity assignment, static activity likelihoods which correspond to the proclivity functions $g_j^{(x)}[n]$ and the determination of a calibration scalar. In general, domestic taps have two states as they may either be *closed* or *open* reducing the calculation of $\dot{m}_\gamma[n]$ to defining the opening and closing events of all taps.

For each tap γ and member x , the probability of switch-open $p_\gamma^{(x)}[n]$ is defined

$$p_{\gamma,n}^{(x)} = \begin{cases} \widehat{g}_{a_\gamma,n}^{(x)} c_\gamma & \text{if } o_n^{(x)} = 1 \\ c_\gamma & \text{if } (a_\gamma = o_n^{(x)}) \text{ or } (a_\gamma = \emptyset) \end{cases} \quad (13)$$

with $c_\gamma \leq 1$ a calibration scalar, $a_\gamma \in \Omega_a$ the activity to which the use of tap γ is related, and $\widehat{g}_{a_\gamma,n}^{(x)}$ the likelihood of activity a_γ related to tap γ for $x \in \beta^{(x)}$. The calibration scalar c_γ determines the average number of times the respective tap γ is used in a year, and is determined iteratively after hundred of algorithm runs per iteration. Here, c_γ is required as there is no one-on-one link between the use of taps and the activities. Once a switch-open event is detected as $X(p_{\gamma,n}^{(x)})$ for tap γ by x , the duration δ_γ of the *open*-state ν_γ is determined as

$$\delta_\gamma = X(\mathcal{N}(\mu_\gamma, \sigma_\gamma^2)) : \sigma_\gamma \triangleq \mu_\gamma/10 \quad (14)$$

with μ_γ the tap' mean duration of the *open*-state, and where after the procedure is repeated for the entire occupancy chain.

Once all events are defined and all durations determined, the resulting load profile $\dot{m}_\gamma[n]$ of γ can be easily determined as $\nu_\alpha[n] \dot{m}_{\nu_\gamma}$ with $\nu_\gamma[n]$ the generated state of γ and \dot{m}_{ν_γ} the characteristics tap flow of γ .

3.4.5 Ventilation and solar shading control

As the focus lies on (new) low-energy dwellings in the proposed research questions, we will state that the required ventilation is achieved by a mechanical ventilation system and its control is, together with the control of the solar shading, performed by a building energy management system.

4. Epistemic uncertainty: evaluating methods in StROBe for IDEAS

The Python StROBe Package has been developed to include stochastic residential occupant behaviour as a main boundary conditions in integrated (building and) district energy simulations with the Modelica IDEAS Library. The main improvements achieved in StROBe compared to the state-of-the-art is the use of clustered BeTUS'05 data for occupancy and activity modelling, and combining all different commodities in a single model found on the same prerequisites; as such improving the representation of their cross- and autocorrelations which are important for district

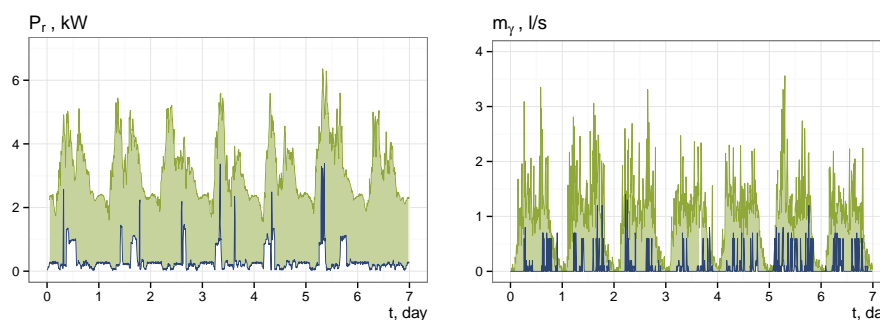


Figure 3. Example week profile of the receptacle load $P_r[n]$ (left) and hot water redraws $\dot{m}_\gamma[n]$ (right) for a single dwelling (blue) compared to the instantaneous aggregate profile (green) derived by summing hundred profiles scaled with a factor .10 for readability.

energy simulations. Resulting example profiles of $P_r[n]$ and $\dot{m}_\gamma[n]$ at dwelling and aggregated level are given in Figure 3.

The Python StROBe Package is one of the main implicit sets of equality constraints for the parameter set \mathbf{x} in integrated district energy simulations. The set of assumptions made in the previous sections and which are specific to the development of a bottom-up building-energy occupant behaviour model such as StROBe cause *epistemic uncertainties*, *i.e.* uncertainties attributable to incomplete knowledge about a phenomenon that affects our ability to model it. As epistemic uncertainties are due to ‘things we could in principle know but don’t in practice’, they are evidently hard to quantify. As such we will try to estimate them on three aspects, *i.e.* by verifying the StROBe model outcome with known reference values on the annual loads, the simultaneity of loads and on the autocorrelation of loads.

Annual loads. An overview is given of the histograms for hundred annual simulations in Figure 4 regarding the total annual receptacle loads, the total annual hot water tap and the average space heating set point temperature for the day-zone.

The observed annual average of the modelled electricity loads is 3.1 MWh. This annual consumption can be compared to the values found in the Belgian Energy Consumption Survey (from now on referenced as BeECS’10) collected in 2010 under the authority of EUROSTAT (Jespers *et al.*, 2010). The BeECS’10 dataset relates to a population 3396 households who completed questionnaires related to their energy demand, and denotes a median annual electricity consumption of 3.6 MWh. The observed underestimation can be explained with two arguments:

- An underestimation of the electricity loads is often found in bottom-up load models based on occupancy, activity and ownership; and is generally addressed by the notion that the considered list of appliances is never exhaustive and the notion that occupant behaviour is not always argumentative (Aerts, 2015).
- Residential energy consumption statistics contain more than just the receptacle loads, *e.g.* the loads of pumps and fans. Additionally, for Belgium, twenty percent of households provide hot water based on an electric boiler, while ten percent of the households heat their house based on electricity and an additional ten percent sporadically heat an additional space with an electric heater.

As such, we would expect an underestimation compared to energy consumption statistics, though we can not directly estimate it. The latter does not apply for hot water tapping. The observed average of the modelled hot water load is 135 litre per household per day normalized at a temperature of 60°C. This number is hard to compare with Belgian statistics, but matches well with the rule-of-thumb denoting 60 litre per person per day which results in an average hot water load of 140 litre per household per day given the average household size of 2.3 persons. The observed average of the modelled space heating set-point in the day-zone is 18.3°C. Also

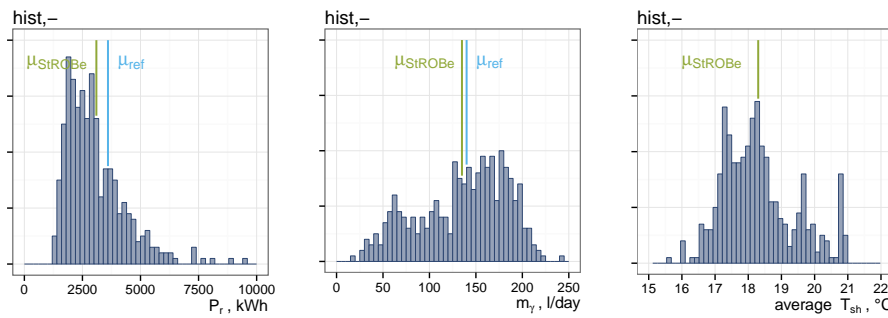


Figure 4. Comparison of the resulting epistemic histograms for the annual receptacle load P_r , the average daily hot water consumption m_γ and the average space heating setpoint T_{sh} with (average) reference values given in literature.

this number is hard to compare as it is based on the supposed set-point temperature given when the space heating is considered to be switched off, but matches well with the considered average building temperature of 18°C in the Belgian energy directive on the energy performance of buildings.

Factors of simultaneity. When modelling the pervasive space at a neighbourhood level, the *simultaneity* of loads at the aggregate level is one of the main factors influencing the resulting system performance and the design assessment. To address this, the factor of simultaneity $k_s \leq 1$ or its inverse diversity factor is defined in both electrical and hydronic design standards. The factor of simultaneity is the ratio of the peak load of a group of loads to the sum of peak powers for each load in this group and can thus be formally determined as:

$$k_s \left(\dot{m}_\gamma^{(\mathcal{H})} \right) = \frac{\max \left(\sum_h^{\mathcal{H}} \dot{m}_\gamma^{(h)} [n] \right)}{\sum_h^{\mathcal{H}} \left(\max \dot{m}_\gamma^{(h)} [n] \right)} \quad \text{s.t. } h \in \mathcal{H} \quad (15)$$

where \mathcal{H} is the set of households h in the depicted feeder or neighbourhood. The same definition is applied to $P_r[n]$ for the design of electricity systems.

The resulting factors of simultaneity k_s for the modelled $P_r[n]$ and $\dot{m}_\gamma[n]$ at τ_{15} are shown in Figure 5 as function of the number of dwellings n considered in the feeder. To get a correct representation of k_s , each neighbourhood with n dwellings is simulated a hundred times.¹ The average k_s of the modelled data is $0.94/n + 0.23$ for $P_r[n]$ and $1.12/n + 0.09$ for $\dot{m}_\gamma[n]$ for $n \in [2, 30]$. Given the simulated set of hundred cases for each feeder size $n \in [2, 30]$, we can compare the found factors of simultaneity with the known design standards; which values should indicate an upper limit of k_s as function of the system size, bound by a confidence interval or safety factor. Before comparing the simulated values with the standard, we will state that ‘the factors of simultaneity k_s may be treated as being normally distributed for a single n ’; which allows us to describe k_s based on standard deviations. For both $k_s(m_\gamma[n])$ and $k_s(P_r[n])$, the null hypothesis H_0 against $\mathcal{N}(\mu, \sigma^2)$ is refuted by SHAPIROS’ W -test for normality for an α -level of .05 and $W \geq .95$.² As such, we can calculate the coefficient of variation of k_s , equal to .2 $n - 14$ percent for the receptacle loads $P_r[n]$ and .07 $n - 12$ percent for the water redraws $\dot{m}_\gamma[n]$.

Having defined the standard deviation under the assumption of normality, we can compare an $\mu + n \cdot \sigma$ upper limit for $k_s(P_r[n])$ and $k_s(m_w[n])$ at τ_{10} with the reference values given in their respective system design standards IEC Std. 60439 and Std. EN 806-3 on low-voltage switchgear

¹The size of hundred cases is set based on the suggestion of K.LOMAS *et al.* and I.MACDONALD *et al.* stating that the accuracy does not improve much above a hundred runs (Lomas & Eppel, 1992; Macdonald, 2009); though this can be questioned (Janssen, 2013).

²To be complete, also the null hypothesis H_0 against $\ln \mathcal{N}(\mu, \sigma^2)$ is refuted by SHAPIROS’ W -test for (log-)normality for the same α -level of .05 and $W \geq .95$, which may allow us to treat k_s as being a log-normal distributed with a longer upper-tail.

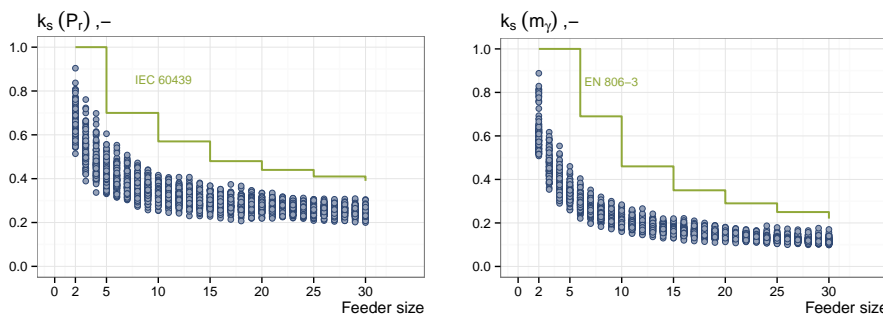


Figure 5. Verification of the modelled factors of simultaneity k_s for τ_{15} receptacle loads $P_r[n]$ (left) and hot water redraws $m_\gamma[n]$ (right) denoted for 100 cases per feeder size in comparison to its design standards.

and in-home water piping respectively; which 'values should describe an upper limit of observed values, bound by a confidence interval or safety factor.

The comparison of the calculated $k_s(P_\alpha[n])$ with IEC Std. 60439 and $k_s(\dot{m}_\gamma[n])$ with Std. EN 806-3 as shown in Figure 5 gives an indication of the correctness of the modelled data. The modelled upper 3σ limit for $k_s(P_\alpha[n])$ of $1.84/n + 0.26$ corresponds visually well with the bottom limit of $3.77/n + 0.26$ in the design standard IEC Std. 60439, and also the bottom limit of $1.81/n + 0.34$ in the design standard Std. EN 806-3 matches visually well to the calculated upper 3σ limit for $k_s(\dot{m}_\gamma[n])$ of $1.61/n + 0.11$. From this we may draw the implication that $k_s(\dot{m}_\gamma)$ and thus the simultaneity of the modelled \dot{m}_γ at feeder level is satisfyingly verified, though the difference between both remains significant. Two explanations can be given for the underestimated k_s of the modelled time series;

- A first reason may be the premise that the pervasive system is spatially homogeneous with respect to the national demography. As stated earlier, A.GARDINER *et al.* and A.RIDDELL repelled this assumption by stating based on measured household and transformer load profiles that 'time series $P_r[n]$ found at individual-house level are apparent at the feeder-transformer' implying that people living in a given neighbourhood predominantly tend to fall in just one or two clusters $\beta_j \in \beta$ (Gardiner & Manson, 1994; Riddell, 1996). Such a local concentration of clusters β_j would amplify the log-normality of $k_s(P_r[n])$ compared to its normality, and as such raise the upper 3σ limit.
- On the other hand, Std. EN 806-3 includes domestic hot water as well as cold water tapping which gives a distorted comparison with the modelled hot water loads.

Unfortunately, there is not enough data available in BeTUS'05 or BeHBS'05 to allow a proper modelling of the local inhomogeneity of the pervasive space. Despite the observed underestimation of the factors of simultaneity compared to their design values in standards, we will say that the simultaneity is modelled well enough for its purpose in district energy simulations.

Receptacles' autocorrelation. When modelling the pervasive space at a building level, the *autocorrelation* of loads is one of the indicators addressing the correct representation of the load profile. To address this, the autocorrelation function (ACF) of the modelled $P_r[n]$ loads at τ_1 for hundred households is compared to an equal number of random profiles measured in a single district at τ_{15} as showed in Figure 5 for a one-week period (Labeeuw, 2013).

Visual comparison of the modelled and measured autocorrelation functions shows an underestimation for the modelled $P_r[n]$ -profile compared to the measured profiles, while the lag is correctly represented. Autocorrelation functions however depend on the sample size and first-order lag, and due to the difference in time resolution of both datasets we have to compare them in relation to their respective threshold levels of correlation needed to reject the null hypothesis H_0 of zero population correlation with .95 significance for an α -level of .05. The resulting $r_{.95}$ are .062 for the modelled load profiles and .094 for the measured load profiles (Andersen, 1941).

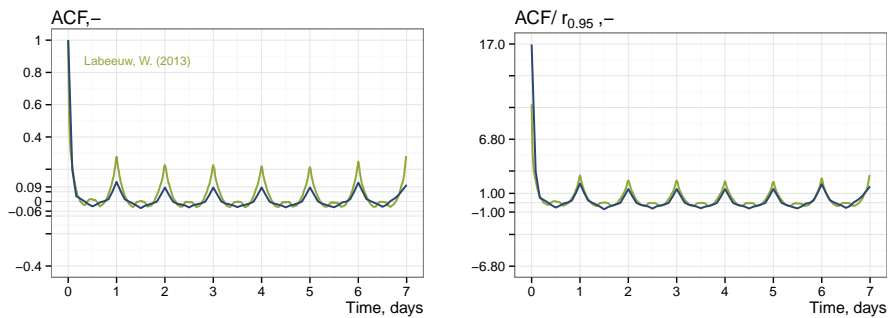


Figure 6. Verification of the modelled autocorrelation function (ACF) for τ_1 for receptacle loads $P_r[n]$ (blue) in comparison to the average ACF for 100 random profiles measured by W.LABEEEUW *et al.* (green) as such and (right) corrected for to their respective threshold levels of correlation $r_{.95}$ needed to reject the null hypothesis H_0 of zero population correlation with .95 significance for an α -level of .05.(Labeeuw, 2013)

As showed in Figure 6 (right), the $r_{.95}$ threshold level is exceeded at lags which are a plurality of a day similar to the observed data, and reaches $2.05 \cdot r_{.95}$ after 24 hours and $1.47 \cdot r_{.95}$ at its pluralities for the modelled load profiles in comparison of $2.98 \cdot r_{.95}$ and $2.41 \cdot r_{.95}$ respectively in the observations. As such, as observed before, a slight underestimation of the auto-correlation function of the measured load profiles remains even when corrected for $r_{.95}$. Two explanations can be given;

- On the one hand, *repeatability* of behaviour is only accounted for in the explicit modelling of $o^{(x)}[n]$. The activity modelling could include a correction for the recorded repetition of daily behaviour by adopting a Markov-chain method as proposed by U.WILKE *et al.* , J.TANIMOTO *et al.* or J.WIDÉN *et al.* who focus on rectifying $a_{x,n}$ (Wilke, 2013; Tanimoto *et al.* , 2008a; Widén *et al.* , 2012). The latter however requires behavioural data of consecutive days and a rejection of the assumption that household individuals behave individually from each other.
- On the other hand and similar to the underestimation of k_s , stating that the pervasive system is spatially homogeneous with respect to the national demography may underestimate the autocorrelation function. As stated earlier, A.GARDINER *et al.* and A.RIDDELL repelled this assumption (Gardiner & Manson, 1994; Riddell, 1996). Such a local concentration of clusters β_j would lower the observed spread and raise the mean observed autocorrelation.

Despite the small observed underestimation of the autocorrelation functions compared to their design values in standards, we will say that the autocorrelation is modelled well enough for its purpose in district energy simulations.

5. Aleatory uncertainties; the inherent spread on results by occupant behaviour

As stated earlier, the Python StROBE Package is one of the main implicit sets of equality constraints for the parameter set \mathbf{x} in integrated district energy simulations. When used for simulation and optimization, the objective function $z_i(\mathbf{x})$ is defined by $Z : \mathbb{R}^m \rightarrow \mathbb{R}$ as $z_i(\mathbf{x}) \triangleq Z(\mathbf{y}_t(\mathbf{x}, 1))$, $\forall i \in \{1, \dots, n\}$ where $\mathbf{y}_t(\mathbf{x}, 1) \in \mathbb{R}^m$ is the simulation solution for all state variables \mathbf{y}_t in time as implemented in the Modelica IDEAS Library or another district energy simulation environment. The stochastic occupant behaviour as modelled in StROBE cause *aleatory uncertainties*, *i.e.* uncertainties inherent in a non-deterministic phenomena. As such, and contrary to epistemic uncertainties, aleatory uncertainties cannot be reduced by further study, as it expresses the inherent variability of a phenomenon. The latter requires us to rewrite the previous definition of the objective function z_i as:

$$z_i(\mathbf{x}, \alpha) \triangleq Z(\mathbf{y}_t(\mathbf{x}, \alpha, 1)) , \forall i \in \{1, \dots, n\} \quad (16)$$

with \mathbf{y}_t the state variables and α the input quantities provided by the environment the system is embedded in, including the pervasive system environment.

Different methods are available to incorporate uncertainties, which thrive to define a robust counterpart \mathbb{Z}_i of the original objective function z_i ; The probabilistic nature of aleatory uncertainties makes probability distributions an adequate mean for the mathematical description of these uncertainties, which is generally translated to ‘an expectancy measure’ and ‘a probabilistic threshold measure’ of robustness in literature.

We therefore simulate the states of $y_t(x, \alpha, 1)$ a hundred times for a radial feeder of variable length n with given energy system designs, only differing in its pervasive space as α -variable.

Case description. The aleatory uncertainty of district energy simulations is defined for low-voltage distribution feeder with 5, 10, 15, 20, 25 and 30 residential detached dwellings connected to a EXAVB 4G150 cable, all with a heated area of 123 m². For each feeder size, a version is simulated with 0, 20, 40, 60, 80 and 100 percent of the buildings equipped with a modulating air-to-water heat pump influenced by the occupant behaviour and influencing the feeder dynamics; and this combined without or with a photovoltaic system of 2 or 4 kVA. As such, 108 different neighbourhood designs have been simulated 100 times to define the aleatory uncertainty caused by occupant behaviour.

The architectural types is determined earlier as representative for the Belgian building stock and is modelled as a 2-zone model with the day zone (*e.g.* living area, kitchen) and night zone (*e.g.* bedrooms) of the dwelling respectively. The distribution of internal gains by occupant behaviour is based on a volume-weighted ratio. All dwellings are designed to result in a low-energy standard. Heat losses by conduction are reduced to a minimum by applying thermal insulation to obtain an overall mean heat transfer coefficient of 0.11 W/m²K, 0.13 W/m²K, 0.10 W/m²K and 0.8 W/m²K for the cavity walls, concrete foundation floors, timber roof constructions and windows respectively; while the dwellings are as airtight as possible with a natural infiltration rate of 0.03 ACH. All dwellings are equipped with mechanically balanced, air-to-air heat-recovery ventilation with an air change rate of 0.5 h⁻¹ and a recovery efficiency of 0.84. These measures result in a design heat load of 26 W/m² for the moderate climate of Uccle, Belgium.

The modulating air-to-water heat pump connected to low-temperature radiator in each thermal zone. The heat pump model is based on interpolation in a performance map retrieved from manufacturer data. The interpolation defines the heating power and electricity use as a function of condenser outlet temperature, the ambient temperature and can modulate to 30 percent. The coefficient of performance based on manufacturer data is 3.17 at 2/35°C test conditions (*i.e.* air/water temperature) and 2.44 at 2/45°C test conditions for full load operation. The heatpump is controlled based on the measured and set-point values for the thermal comfort. The HP control set-points are based on a heating curve for space heating (*i.e.* 55°C at an outdoor temperature of -8°C and 20°C at an outdoor temperature of 15°C).

Objective functions. Four objective functions will be focused on in the uncertainty assessment, *i.e.* the annual Ohmic losses in the grid E_Ω , the annual net electricity demand E_γ , the absolute peak transformer load P_{tra} and the characteristic voltage deviation U_ϕ^{rms} defined as

$$U_\phi^{(k)} = \frac{1}{U_0^{(k)}} \left[\frac{1}{n_d} \sum_{d=1}^{n_d} \max_{\tau_{10}} \left(\left| U_{\phi,d}^{(k)}(t) \right| - U_0^{(k)} \right)^2 \right]^{0.5} \quad (17)$$

where $U_{\phi,d}^{(k)}$ is the line voltage of phase ϕ in feeder k at day d and $U_0^{(k)}$ be the reference voltage, *i.e.* 230 V. Before quantifying the aleatory uncertainty, we will state that ‘all four objectives $z_i(\mathbf{x}, \alpha)$ may be treated as being log-normally distributed for α given a single n ’; which allows us to describe z_i based on a ‘coefficient of variation’. For all four objectives, the null hypothesis H_0

against $\ln\mathcal{N}(\mu, \sigma)$ is refuted by SHAPIROS' W -test for log-normality with a test statistic $W \geq .90$ and an α -level of .001 for feeder sizes n larger than 2, and with $W \geq .95$ and an α -level of .05 for feeder sizes larger than 10. The observed test statistics for log-normality are, hereby, much higher than the test statistics of H_0 against $\mathcal{N}(\mu, \sigma^2)$, where $W \geq .90$ is only observed for feeder sizes higher than 10 and an α -level of .05 for feeder sizes of 25 or higher. As such, we may formulate all criteria as being log-normally distributed for the pervasive space α denoting a tail in the probability density function at one side of the mean.

5.1 Expectancy measures of robustness

A first approach to deal with aleatory uncertainties is defining one or more integral measures of robustness, *i.e.* 'expectancy measures' and 'dispersion measures'; The expectancy measure of an objective $z_i(\mathbf{x}, \alpha)$ describes the long-run average value of repetitions of the variable it represents, whereas the dispersion measure of $z_i(\mathbf{x}, \alpha)$ describes the possible values of the objective around the expected value.

Within this context, the robust counterpart $\mathbb{Z}_i(\mathbf{x}, \alpha)$ of $z_i(\mathbf{x}, \alpha)$ can be defined as the expectation of $Z(y_t(\mathbf{x}, \alpha, 1))$ denoted as $E[z_i(\mathbf{x}, \alpha)]$; which has to be traded off to its variance denoted as $\text{Var}[z_i(\mathbf{x}, \alpha)]$. The multi-objective optimization task then becomes

$$(E[z_i(\mathbf{x}, \alpha)], \text{Var}[z_i(\mathbf{x}, \alpha)]) , \forall i \in \{1, \dots, n\} , \mathbf{x} \in S \quad (18)$$

$E[z_i(\mathbf{x}, \alpha)]$ and $\text{Var}[z_i(\mathbf{x}, \alpha)]$ may have different minimizers $\hat{\mathbf{x}}$ representing conflicting objectives. Therefore, we might need to aggregate both objectives in a single objective function using a weighted sum of both functions or consider the Pareto-optimal solutions.¹ The stated trade-off is given graphically for four $z_i(\mathbf{x}, \alpha)$ in Figure 7 for the given three different building energy solutions in a radial feeder and for a variable degree of heat pump and photovoltaic system implementation. No general relation can be observed between the variance and expectation of a random objective $z_i(\mathbf{x}, \alpha)$. However, following remarks may be stated based on the observed trade-offs.

$E[z_i(\mathbf{x}, \alpha)]$ and $\text{Var}[z_i(\mathbf{x}, \alpha)]$ have different minimizers $\hat{\mathbf{x}}$ when regarding the ohmic losses E_Ω , characteristic voltage deviation U_ϕ^{rms} as defined in Equation 17 on page 15, net energy off-take E_γ and transformer peak load P_{tra} as to be minimized objective functions; therefore representing conflicting goals. All observed $\text{Var}[z_i(\mathbf{x}, \alpha)]$ are, however, lower than or equal to the observed variance for the annotated reference case only existing of receptacle loads; except for $x \in \mathcal{S}_1$ considering $E_\Omega \rightarrow \min$. The latter allows us to state that $E[z_i(\mathbf{x}, \alpha)]$ may be seen as the dominating objective function, as a reduction in $\text{Var}[z_i(\mathbf{x}, \alpha)]$ is almost always achieved though not to a minimum. The denoted increase of the conditional variance $\text{Var}[z_i(\mathbf{x}, \alpha)]$ for a reduction in conditional expectation $E[z_i(\mathbf{x}, \alpha)]$, may however induce a flaw in using $E[z_i(\mathbf{x}, \alpha)]$ as objective function for minimisation; If the increase of $\text{Var}[z_i(\mathbf{x}, \alpha)]$ is more pronounced than the decrease in $E[z_i(\mathbf{x}, \alpha)]$, an absolute increase of $z_i(\mathbf{x}, \alpha)$ given \mathbf{x} could be noticed for certain α .

As shown in Figure 8, besides E_Ω , all stated objectives functions $z_i(\mathbf{x}, \alpha)$ show a general low $\text{Var}[z_i(\mathbf{x}, \alpha)]$ with respect to $E[z_i(\mathbf{x}, \alpha)]$; generally denoting a factor of variation below 15 percent. Taking into account the log-normality and leaving aside E_Ω , we may say that 95.4 percent of the observed objectives $z_i(x)$ lay in the (.81, 1.6) $E[z_i|x]$ interval for $n \leq 10$ with a p -value of .92, and in the (.88, 1.3) $E[z_i|x]$ interval for $n \leq 20$ with a p -value of .80.

For the considered systems \mathbf{x} , we observe that integrating (or increase) photovoltaic systems always lowers the variance of the denoted $z_i(\mathbf{x}, \alpha)$. This may be explained logically, as the integration of deterministic loads (*i.e.* equal in all simulated cases) influencing $z_i(\mathbf{x}, \alpha)$ lowers

¹The former approach using a weighted sum of both functions is not explored in the consecutive paragraphs as there is no rational method to defining the weight factor.

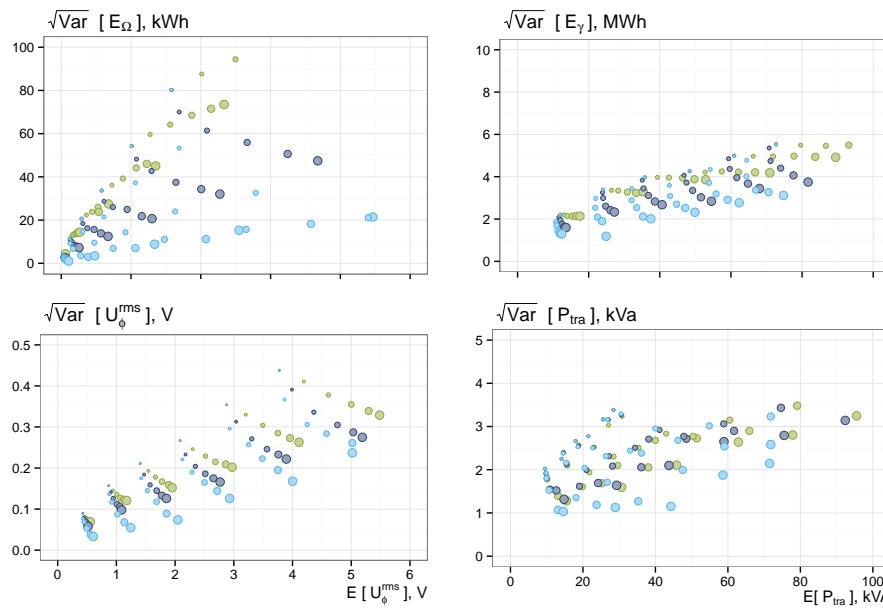


Figure 7. Trade-off between the variance $\text{Var}[z_i(\mathbf{x}, \alpha)]$ induced by human behaviour and the expectation $E[z_i(\mathbf{x}, \alpha)]$ of the Ohmic losses E_Ω , voltage quality U_ϕ^{rms} , net energy offtake E_γ and peak transformer load P_{tra} for a feeder of variable length, building system design and degree of implementation of low-energy dwellings.

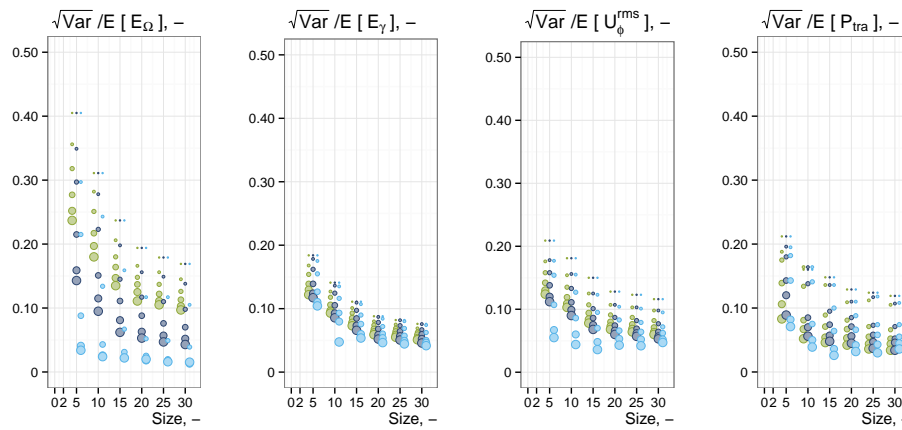


Figure 8. Variance $\text{Var}[z_i(\mathbf{x}, \alpha)]$ induced by human behaviour as function [of the feeder size as shown in Figure 7 on page 17, normalized to the expected value $E[z_i(\mathbf{x}, \alpha)]$ for a feeder of variable building system design and degree of implementation of low-energy dwellings.

the dependency of $z_i(\mathbf{x}, \alpha)$ on the behaviour-related α -variables; thus lowering $\text{Var}[z_i(\mathbf{x}, \alpha)]$. The latter can not be formulated for the integration of the heat pump system: we observe that adding heat pumps may increase the absolute value of the coefficient of variation, while lowering it when normalized to the expected value. Furthermore, the impact of the integration of heat pumps on $z_i(\mathbf{x}, \alpha)$ depends on the behaviour-induced control of the system. However, both statements can not be generalized for all possible solutions based on the limited considered number of technical systems in this section.

5.2 Threshold measure of robustness

A second approach to deal with aleatory uncertainties is proposed based on defining a probabilistic threshold measure of robustness. Within this context, we do not consider the expected value of an objective function $z_i(\mathbf{x}, \alpha)$ but the distribution of the variate directly. That is, in the case of minimization, we look for the threshold $\{z_i(\mathbf{x}, \alpha)\}_{p_0}$ below which we expect $p_0 n$ number

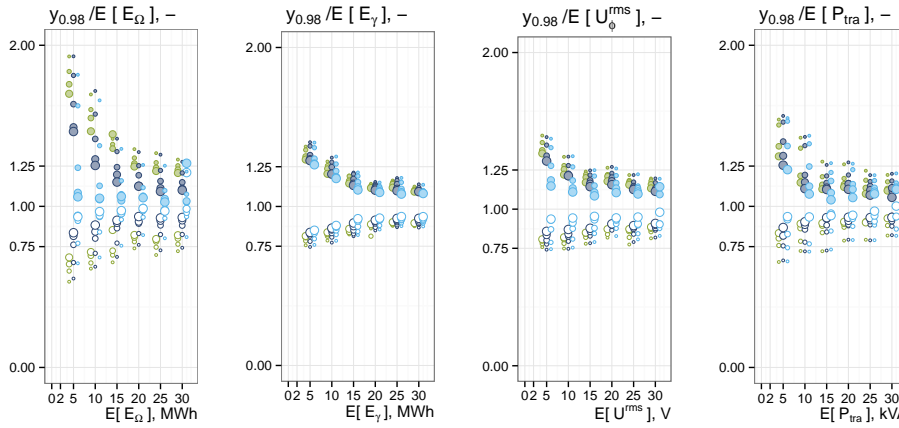


Figure 9. Normalized 2nd and 98th percentile $\{z_i(\mathbf{x}, \alpha)\}_j$ induced by human behaviour of the Ohmic losses E_Ω , voltage quality U_ϕ^{rms} , net energy offtake E_γ and peak transformer load P_{tra} for a feeder of variable building system design and degree of implementation of low-energy dwellings.

of samples fulfilling $z_i(z_i(\mathbf{x}, \alpha)) \leq \{z_i(\mathbf{x}, \alpha)\}_{p_0}$ for a fixed number of samples n . We therefore yield the threshold dependent criterion $\{z_i(\mathbf{x}, \alpha)\}_{p_0}$ for robustness of $z_i(\mathbf{x}, \alpha)$ as

$$p(z_i(\mathbf{x}, \alpha) \leq \{z_i(\mathbf{x}, \alpha)\}_{p_0}) \geq p_0 \quad (19)$$

with p_0 defines the desired conditional relative frequency of the $z_i(\mathbf{x}, \alpha)$ in the case of minimization.

In relation to the earlier mentioned 95.4 percent confidence interval based on the 3σ -rule, we will focus on $\{z_i(\mathbf{x}, \alpha)\}_{.02}$ and $\{z_i(\mathbf{x}, \alpha)\}_{.98}$, closely representing the expected infimum and supremum in simulations. The threshold values of objectives are given graphically for four $z_i(\mathbf{x}, \alpha)$ in Figure 9 in relation to the feeder size, for the given three different building energy solutions for a variable degree of implementation. These results confirm the observations found in the previous section based on $\text{Var}[z_i(\mathbf{x}, \alpha)]$: the statement that 95.4 percent of the observed objectives $z_i(x)$ lay in the $(.81, 1.6) E[z_i|x]$ interval for $n \leq 10$ and in the $(.88, 1.3) E[z_i|x]$ interval for $n \leq 20$ corresponds well with the found range of $z_{.02}(\mathbf{x}, \alpha)$ to $z_{.98}(\mathbf{x}, \alpha)$ leaving aside E_Ω . Similarly to the observations for $\text{Var}[E_\Omega(\mathbf{x}, \alpha)]$, the 2nd and 98th percentiles for E_Ω are show a much larger range compared to the other objective functions showing an uncertainty which is roughly twice as large.

Furthermore, we noticed that $E[z_i(\mathbf{x}, \alpha)]$ and $\{z_i(\mathbf{x}, \alpha)\}_{.98}$ have the same minimizers \hat{x} therefore representing harmonized goals. This observation is however not absolute, *i.e.* different points can be observed which have alternatives north-west in $E[z_i(\mathbf{x}, \alpha)] \mapsto \{z_i(\mathbf{x}, \alpha)\}_{.98}$ or south-east in $E[z_i(\mathbf{x}, \alpha)] \mapsto \{z_i(\mathbf{x}, \alpha)\}_{.02}$ when regarding $z_i(\mathbf{x}, \alpha)$ as to be minimized objective functions. The latter means that two options can be weighted differently against each other, depending on assuming the conditional expectancy or the 98th percentile of the observed objective functions $z_i(\mathbf{x}, \alpha)$, and thus depending on the occupant behaviour.

6. Conclusions

The presented work started by (a) the description of StROBe, a stochastic residential occupant behaviour for district energy simulations; integrating the modelling of receptacle loads, internal heat gains, thermostat settings and hot water tapping based on occupancy and activity prerequisites. Given this model, the uncertainty for district energy simulations is addressed: epistemic uncertainties (b) are elaborated first comparing model results with reference values and denoting possible hiatus of general modelling methods for building energy occupant behaviour used at the neighbourhood level, followed by the aleatory uncertainty (c) describing the uncertainty in

district energy simulations inherent to the non-deterministic occupant behaviour.

The resulting outcome of **StROBe** is verified with known standards and measurement indicating its fit for purpose and main shortcomings. At the neighbourhood level, the simultaneity of electric and hot water tap loads is compared to existing design standards which show a good agreement with the 5σ -limit of the modelled data. As the statistics behind the reference standards is unknown, the distance from the modelled factors of simultaneity to the design values is assigned to the assumption that ‘occupant behaviour is spatially homogeneous with respect to the national demography’. Here, local disaggregation of demographic statistics might be of high importance in future work on modelling of occupant behaviour for district energy simulations.

To conclude, we quantified the aleatory uncertainty caused by stochastic residential occupant behaviour in integrated district energy simulations. For the simulated cases, we have shown that the design of the building energy system has its impact on the robustness of the objective criteria and it could thus be minimized as part of an optimisation exercise. Furthermore, the expected value of the objective functions have to a large extend the same minimizers as the measures of the proposed robustness. As such, optimizing an objective value for its expected value generally seems to result in a optimum *near* the optimum of robustness. However, 95 percent of the observed objectives lay between 0.81 and 1.6 times the expected value for a feeder larger than 10 houses and between 0.88 and 1.3 times the expected value for more than 20 houses denoting an overall ‘rather small’ uncertainty on the possible objective functions caused by occupant behaviour.

In future work, the presented occupant behaviour model can be used to properly assess the impact of *e.g.* residential heat pump based space heating and domestic hot water systems on the electricity distribution grid, and estimate the potential of thermal demand side management with respect to electric grid constraints.

7. Nomenclature

Physical symbols

$a^{(x)}[n]$	Activity chain in discrete time of individual x
α	Input parameters causing uncertainty
α	Statistical significance, -
α	Appliance
β_j	Subset $j \subseteq \text{BeTUS}'05$
$\beta^{(x)}$	Cluster to which x is appointed to
$C(\cdot)$	Clustering of dataset \cdot
d_{pdf}	Density distance function, -
E_Ω	Annual Ohmic feeder losses, W
E_γ	Net energy off-take, W
$E(\cdot)$	Expected value of the depicted set of variables
$g_{j,n}^{(x)}$	Proclivity function for activity j of individual x , -
$\hat{g}_j^{(x)}$	Approximate of $g_{j,n}^{(x)}$ based on observed data, -
$g_{j,n_d}^{\beta^{(x)}}$	Observed $g_{j,n}^{(x)}$ in $\beta^{(x)} \subseteq \text{BeTUS}'05$ at time n_d
H_0	Null hypothesis
$i_{d_{pdf}}$	Distance, -
k_s	Factor of simultaneity, -
l_n	Lead time at time n , s
$L_n^{(x)}$	Occupancy lead time density at time n , -
$\hat{L}_n^{(x)}$	Approximate of $L_n^{(x)}$ based on observed data
$L_{n_d}^{\beta^{(x)}}$	Observed $L_n^{(x)}$ in $\beta^{(x)} \subseteq \text{BeTUS}'05$ at time n_d

- $\dot{m}_w[n]$ Hot water tapping profile, L/s
 n_d Clock-time
 n, k Time in discrete notation
 n Number of dwellings in depicted neighbourhood, -
 $o^{(x)}[n]$ Occupancy chain in discrete time of individual x
 $o_n^{(x)}$ Occupancy state of x at time n
 $o'^{(x)}[n]$ Observed occupancy chain in discrete time of individual x
 $O_n^{(x)}$ Occupancy event density at time n
 $\hat{O}_n^{(x)}$ Approximate of $O_n^{(x)}$ based on observed data, -
 $O_{n_d}^{(\beta^{(x)})}$ Observed $O_n^{(x)}$ in $\beta^{(x)} \subseteq \text{BeTUS}'05$ at time n_d
 Ω_o Occupancy state space
 Ω_a Activity state space
 $P_{hp}(t)$ Heat pump load, W
 $P_{net}(t)$ Net power exchange, W
 P_{tra} Peak transformer load, VA
 $P_\alpha[n]$ Receptacle load profile for appliance α , W
 $P_\lambda[n]$ Lighting load profile for appliance α , W
 $\dot{Q}_{a,g}[n]$ Convective internal heat load profile, W
 $\dot{Q}_{lw,g}[n]$ Longwave internal heat load profile, W
 $r_{.98}$ Threshold value at the 98th percentile
 s Day-type
 \mathcal{S} Day-type state space
 $T_{op}(t)$ Operative building zone temperature, K
 τ_i Time resolution of i minutes
 $T_{sh}^s[n]$ Space heating set point temperatures for zone s , K
 U_ϕ Phase feeder voltage, Vor pu
 U_0 Reference phase voltage, *i.e.* 230 V
 U_ϕ^{rms} Voltage quality, V
 $\text{Var}(\cdot)$ Variation of the depicted set of variables
 \mathbf{x} Set of parameters
 \mathbf{y}_t State variables
 $\tilde{\mathbf{y}}_t$ Uncertain approximation of the state variables \mathbf{y}_t
 z_i Evaluation criteria i

Sub- or superscripts

- (x) Variable related to individual x
 (h) Variable related to household h
 n Defined at time n

8. Acknowledgements

The authors gratefully acknowledge the The Research Foundation - Flanders and the KU Leuven Energy Institute (EI) for funding this research through granting the project entitled Optimized energy networks for buildings.

References

- Aerts, Dorien. 2015. *Occupancy and activity modelling for building energy demand simulations, comparative feedback and residential electricity demand characterisation*. PhD, Vrije Universiteit Brussel.

- Andersen, R L. 1941. Distribution of the serial correlation coefficients. *Annals of Mathematical Statistics*, **8**(1), 1–13.
- Andersen, Rune Vinther, Fabi, Valentina, Toftum, Jø rn, Corgnati, Stefano Paolo, & Olesen, Bjarne W. 2013. Window opening behaviour modelled from measurements in Danish dwellings. *Building and Environment*, **69**, 101–113.
- Armstrong, Marianne, Swinton, Mike, Ribberink, Hajo, Beausoleil-Morrison, Ian, & Millette, Jocelyn. 2009. Synthetically derived profiles for representing occupant-driven electric loads in Canadian housing. *Journal of Building Performance Simulation*, **2**(1), 15–30.
- Baetens, Ruben, De Coninck, Roel, Helsen, Lieve, & Saelens, Dirk. 2011. Integrated dynamic electric and thermal simulations for a residential neighborhood: Sensitivity to time resolution of boundary conditions. *Pages 1745–1752 of: Proc. of the 12th Conf. of The International Building Performance Simulation Association*.
- Capasso, A, Grattieri, W, Insinga, F, Invernizzi, A, Lamedica, R, & Prudenzi, A. 1993. Validation tests and applications of a model for demand-side management studies in residential load areas. *Pages 1–5 of: 12th International Conference on Electrivty Distribution*.
- Capasso, A, Grattieri, W, Lamedica, R, & Prudenzi, A. 1994. A bottom-up approach to residential load modeling. *IEEE Transactions on Power Systems*, **9**(2), 957–964.
- Fritsch, R, Kohler, A, Nygard-Ferguson, M, & Scartezzini, J.-L. 1990. A Stochastic Model of User Behaviour Regarding Ventilation. *Building and Environment*, **25**(2), 173–181.
- Gardiner, A, & Manson, K. 1994. The relationship between individual domestic appliance load profiles and aggregate distribution loading. *In: Electrical Supply Engineers Association Conference*.
- Glorieux, Ignace, & Minnen, J. 2008. *Website 'Belgisch tijdsbudgetonderzoek' (www.time-use.be) - Onderzoeksgroep TOR Vrije Universiteit Brussel & Algemene Directie Statistiek en Economische Informatie, Brussel*.
- Glorieux, Ignace, Minnen, J, & Van Tienoven, J-P. 2008a. *Het collectieve ritme van België. Evoluties in het levensritme van de Belgen op basis van tijdsbestedingsonderzoek uit 1966, 1999 en 2005*. Tech. rept. Onderzoeksgroep TOR, Vakgroep Sociologie, Brussel.
- Glorieux, Ignace, Mestdag, I, & Minnen, J. 2008b. *Technisch verslag Time & Budget*. Tech. rept. Onderzoeksgroep TOR, Vakgroep Sociologie, Brussel.
- Groscurth, H.-M., Bruckner, Th., & Kümmel, R. 1995. Modeling of energy-services supply systems. *Energy*, **20**(9), 941–958.
- Guerra-santin, Olivia, & Itard, Laure. 2010. Occupants behaviour : determinants and effects on residential heating consumption. *Building Research & Information*, **38**, 318–338.
- Haldi, Frédéric, & Robinson, Darren. 2011. The impact of occupants' behaviour on building energy demand. *Journal of Building Performance Simulation*, **4**(4), 323–338.
- Heylen, Kristof, Le Roy, Marie, Vanden Broucke, Steven, Vandekerckhove, Brecht, & Winters, Sien. 2007. *Wonen in Vlaanderen - De resultaten van de Woonsurvey 2005 en de Uitwendige Woonschouwing 2005*. Tech. rept. Vlaams Kenniscentrum voor Duurzaam Woonbeleid.
- Janssen, Hans. 2013. Monte-Carlo based uncertainty analysis: Sampling efficiency and sampling convergence. *Reliability Engineering & System Safety*, **109**, 123–132.
- Jespers, Kaat, Dams, Yoko, Aernouts, Kristien, Simus, Pascal, Jacquemin, Frederic, Delaite, Laurent, & Vanderhoeft, Camille. 2010. *Energy consumption survey for Belgian households*. Tech. rept.
- Johnson, Ted, & Long, Tom. 2005. Determining the frequency of open windows in residences: a pilot study in Durham, North Carolina during varying temperature conditions. *Journal of exposure analysis and environmental epidemiology*, **15**(4), 329–49.
- Jordan, U, & Vajen, K. 2001a. Influence of the DHW load profile on the fractional energy savings: A case study of a solar combi-system with TRNSYS simulations. *Solar Energy*, **69**(July), 197–208.
- Jordan, Ulrike, & Vajen, Klaus. 2001b. *Realistic Domestic Hot-Water Profiles in Different Time Scales*. Tech. rept. Universität Marburg, Marburg.
- Labeeuw, Wouter. 2013. *Characterisation and modelling of residential electricity demand*. PhD, KU Leuven.
- Leidelmeijer, K, & Van Grieken, P. 2005. *Wonen en energie - stook- en ventilatiegedrag van huishoudens*. Tech. rept. RIGO Research en Advies BV.
- Lomas, Kevin J, & Eppel, Herbert. 1992. Sensitivity analysis techniques for building thermal simulation programs. *Energy and Buildings*, **19**, 21–44.
- Macdonald, Iain A. 2009. Comparison of sampling techniques on the performance of Monte-Carlo based sensitivity analysis. *Pages 992–999 of: Building simulation 2009*.
- Mansouri, I, Newborough, M, & Probert, D. 1996. Energy consumption in UK households: Impact of domestic electrical appliances. *Applied Energy*, **54**(3), 211–285.
- Paatero, J V, & Lund, P D. 2006. A model for generating household electricity load profiles. *International Journal of Energy Research*, **30**(5), 273–290.
- Richardson, I, Thomson, M, & Infield, D. 2008. A high-resolution domestic building occupancy model for energy demand simulations. *Energy and Buildings*, **40**(8), 1560–1566.
- Richardson, I, Thomson, M, Infield, D, & Delahunty, A. 2009. Domestic lighting: A high-resolution energy demand model. *Energy and Buildings*, **41**(7), 781–789.
- Richardson, I, Thomson, M, Infield, D, & Clifford, C. 2010. Domestic electricity use: A high-resolution energy demand model. *Energy and Buildings*, **42**(10), 1878–1887.
- Richardson, Ian. 2010. *Integrated high-resolution modelling of domestic electricity demand and low voltage electricity distribution networks*. Ph.D. thesis, Loughborough University.
- Riddell, a. 1996. Parametrisation of domestic load profiles. *Applied Energy*, **54**(3), 199–210.
- Rijal, H. B., Tuohy, P., Nicol, F., Humphreys, M. a., Samuel, a., & Clarke, J. 2008. Development of an adaptive window-opening algorithm to predict the thermal comfort, energy use and overheating in buildings. *Journal of Building Performance Simulation*, **1**(1), 17–30.
- Rijal, Hom B., Tuohy, Paul, Humphreys, Michael a., Nicol, J. Fergus, & Samuel, Aizaz. 2011. An algorithm to represent occupant use of windows and fans including situation-specific motivations and constraints. *Building Simulation*, **4**(2), 117–134.
- Schweiker, Marcel, Haldi, Frédéric, Shukuya, Masanori, & Robinson, Darren. 2012. Verification of stochastic models of window opening behaviour for residential buildings. *Journal of Building Performance Simulation*, **5**(1), 55–74.
- Spur, R, Fiala, D, Nevala, D, & Probert, D. 2006. Influence of the domestic hot-water daily draw-off profile on the performance of a hot-water store. *Applied Energy*, **83**(7), 749–773.
- Stokes, M, Rylatt, M, & Lomas, K. 2004. A simple model of domestic lighting demand. *Energy and Buildings*, **36**(2),

- 103–116.
- Tanimoto, J, Hagishima, A, & Sagara, H. 2008a. Validation of methodology for utility demand prediction considering actual variations in inhabitant behaviour schedules. *Journal of Building Performance Simulation*, **1**(1), 31–42.
- Tanimoto, Jun, Hagishima, Aya, & Sagara, Hiroki. 2008b. Validation of probabilistic methodology for generating actual inhabitants behavior schedules for accurate prediction of maximum energy requirements. *Energy and Buildings*, **40**(3), 316–322.
- Vanneste, D, De Decker, P, & Laureyssen, I. 2001. *Sociaal-Economische Enquête 2001. Monografie. Woning en woonomgeving in België*. Tech. rept.
- VRO/DG Wonen. 2000. *Kwalitatieve Woningregistratie*.
- Walker, C. 1982. *A residential electrical load model*. PhD, University of New Hampshire.
- Walker, C F, & Pokoski, J L. 1985. Residential load shape modeling based on customer behavior. *IEEE Transactions on Power Apparatus and Systems*, **PAS-104**(7), 1703–1711.
- Warren, P, & Parkins, L. 1984. Window-opening behavior in office buildings. *ASHRAE Transactions*, **90**, 1056–1076.
- Widén, J, & Wäckelgå rd, E. 2010. A high-resolution stochastic model of domestic activity patterns and electricity demand. *Applied Energy*, **87**(6), 1880–1892.
- Widén, Joakim, Nilsson, Annica M., & Wäckelgå rd, Ewa. 2009a. A combined Markov-chain and bottom-up approach to modelling of domestic lighting demand. *Energy and Buildings*, **41**(10), 1001–1012.
- Widén, Joakim, Lundh, Magdalena, Vassileva, Iana, Dahlquist, Erik, Ellegå rd, Kajsa, & Wäckelgå rd, Ewa. 2009b. Constructing load profiles for household electricity and hot water from time-use data - Modelling approach and validation. *Energy and Buildings*, **41**(7), 753–768.
- Widén, Joakim, Wäckelgå rd, Ewa, Paatero, Jukka, & Lund, Peter. 2010. Impacts of different data averaging times on statistical analysis of distributed domestic photovoltaic systems. *Solar Energy*, **84**(3), 492–500.
- Widén, Joakim, Molin, Andreas, & Ellegå rd, Kajsa. 2012. Models of domestic occupancy, activities and energy use based on time-use data: deterministic and stochastic approaches with application to various building-related simulations. *Journal of Building Performance Simulation*, **5**(1), 27–44.
- Wilke, Urs. 2013. *Probabilistic bottom-up modeling of occupancy and activities to predict electricity demand in residential buildings*. PhD, Ecole polytechnique de fédérale de Lausanne.
- Wilke, Urs, Haldi, Frédéric, Scartezzini, Jean-Louis, & Robinson, Darren. 2013. A bottom-up stochastic model to predict building occupants' time-dependent activities. *Building and Environment*, **60**(Feb.), 254–264.
- Yun, G, & Steemers, K. 2010. Night-time naturally ventilated offices: statistical simulations of window-use patterns from field monitoring. *Solar Energy*, **84**, 1216–1231.
- Yun, G, Tuohy, P., & Steemers, K. 2009. Thermal performance of a naturally ventilated building using a combined algorithm of probabilistic occupant behaviour and deterministic heat and mass balance models. *Energy and Buildings*, **41**, 489–499.
- Yun, Geun Young, & Steemers, Koen. 2008. Time-dependent occupant behaviour models of window control in summer. *Building and Environment*, **43**(9), 1471–1482.