Energy and Buildings 96 (2015) 30–39

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

# Four-state domestic building occupancy model for energy demand simulations

Eoghan McKenna*, Michal Krawczynski, Murray Thomson

*Centre for Renewable Energy Systems Technology (CREST), School of Electronic, Electrical and Systems Engineering, Loughborough University, LE11 3TU, UK*

ARTICLE INFO

*Article history:*
Received 8 September 2014
Received in revised form 2 March 2015
Accepted 4 March 2015
Available online 12 March 2015

*Keywords:*
Domestic occupancy
Markov chain
Energy modelling

ABSTRACT

Stochastic building occupancy models are increasingly used to underpin building energy demand models, especially those providing high-resolution electricity demand profiles. This paper describes the development of an established two-state active-occupancy model into a four-state model in which the absent/present state and the active/inactive state are treated separately. This provides a distinction between sleeping and absence and so offers an improved basis for demand modelling, particularly high-resolution thermal modelling. The model uses a first-order Markov chain technique and the paper illustrates the value of this approach in duly representing the naturally occurring correlation of occupancy states in multiply occupied dwellings. The paper also describes how the model has been enhanced to avoid under-representation of dwellings with 24 h occupancy. The model has been implemented in Excel VBA and made available to download for free. The model is constructed from and verified against UK time-use survey data but could readily be adapted to use similar data from elsewhere.



## 1. Introduction

The transition to a low-carbon economy may be expected to require high penetrations of low-carbon technologies such as heat pumps, electric vehicles and photovoltaics [1,2]. These large and potentially undiversified loads and generation could present a considerable challenge to the operation of electricity distribution networks, potentially necessitating significant network reinforcement at high cost [3]. Furthermore, the accurate determination of exactly where and when such reinforcement is required is not straightforward – conventional low-voltage network design procedures typically use rather simple representations of the varying demand and rely heavily on experience – experience which is not available for high penetrations of low-carbon technologies [4].

To address this, and for other applications, 'bottom-up' models of domestic electricity demand that use probabilistic methods to provide stochastic high-resolution data for individual dwellings are currently being developed. A common feature of these models is a core representation of the occupancy of individuals within dwellings, which is used as the basis for subsequent modelling of end-use demands. The high-resolution model of domestic electricity demand developed by Loughborough University [5–8] is constructed in this way. It uses a two-state active-occupancy model that feeds into determining stochastic switch-on events for individual lighting and domestic appliances. The published model has been used widely within academia and industry for electricity network modelling [9–11]. It does not, however, include any detailed representation of thermal demands and, therefore, cannot yet be used to properly investigate the effects of the electrification of heating or CHP. Work is now underway at Loughborough to construct an integrated thermal–electrical demand model that can provide a convenient basis for future network studies, including the electrification of heating. This paper describes the first stage of that development.

A requirement for the thermal modelling is to account for casual gains associated with heat produced by lighting, appliances, and occupants. While the first two can be readily derived from the existing lighting and appliance models, the latter requires knowledge of when occupants are present within the dwelling, including when they are sleeping. The existing occupancy model [5], however, does not differentiate between occupants who are asleep and those who are not at home. The model has therefore been extended from a 'two-state' model to a 'four-state' model, where Table 1 describes the various occupancy states. The first aim of this paper is to describe the development and verification of the new four-state occupancy model. The model has been developed as a Microsoft Excel workbook and has been made available to download for free [12].

* Corresponding author. Tel.: +44 07958 531 842.
*E-mail address:* e.j.mckenna@lboro.ac.uk (E. McKenna).

http://dx.doi.org/10.1016/j.enbuild.2015.03.013

**Table 1**
comparison of different states of occupancy for two-state and four-state models.

| | Description of state |
|---|---|
| *Two-state occupancy model* | |
| 1 | Active occupant – at home and active |
| 0 | Not an active occupant |
| *Four-state occupancy model* | |
| 00 | Not at home, and not active |
| 01 | Not at home, and active |
| 10 | At home, and not active |
| 11 | At home, and active |

In addition, this paper presents a comprehensive review of the occupancy modelling literature and identifies two modelling assumptions that require further scrutiny. The first is that occupancy can be adequately modelled as a first-order Markov process, and the second is that occupancy patterns within dwellings of more than one occupant can be adequately modelled by assuming the occupants are independent of each other. The second aim of the paper is therefore is to evaluate the discrepancies between real occupancy data and the synthetic occupancy data generated by models with these assumptions.

## 2. Literature review

### 2.1. Early 'bottom-up' models of demand

Swan and Ugursal describe the requirements for comprehensive models of residential energy consumption to assess the impacts of technology and behaviour change, and critically assess the strengths and weaknesses of different modelling approaches [13]. A distinction is made between 'top-down' and 'bottom-up' approaches. Top-down models take a whole-system view of demand, and are based on identifying correlations between 'macro' variables (e.g. price, climate, etc.) and aggregated demand data. Bottom-up approaches model individual end-uses at the building level, based on detailed data from samples of the population, which are then aggregated together to produce a wider view of demand. Top-down models tend to be deterministic, while bottom-up models tend to use probabilistic methods to account for demand diversity. Of the two approaches, a bottom-up approach is the more appropriate choice for the requirement of modelling of individual buildings for future network studies as it can represent the diversity of demand at this level.

Early work on developing bottom-up models of domestic electricity demand emphasised the importance of including occupancy as a core variable. Capasso et al. developed a complex residential load model that used a probabilistic 'availability at home' input for each member of the household [14]. The high-resolution domestic lighting model developed by Stokes et al. takes into account the number of occupants but not variations in daily occupancy patterns, with the authors noting that "taking account of these patterns would improve the modelling of diversity" [15]. Jardine used household occupancy derived from measured household demand data as a key input parameter to generate disaggregated high-resolution demand profiles [16]. Yao and Steemers developed a method for generating household load profiles based on five pre-determined occupancy patterns and noted that "the load profile depends very much on the occupancy pattern" [17].

### 2.2. Markov-chain technique and use of time-use surveys

Following these early models, important advances were made by the Richardson–Thomson occupancy model [5], and the Page occupancy model [18]. These are the earliest published models that use a first-order Markov-chain technique to generate stochastic synthetic occupancy patterns. The concept of a first-order Markov-chain technique is that the probability of being in any state in a given time step depends only on the state in the previous time step (an $n$th-order Markov-chain would base these probabilities on the previous $n$ time-steps). The probabilities of changing from one state to another ('transition probabilities') are held in 'transition probability matrices' which are derived from observed occupancy states. Both models use 'time-inhomogeneous' Markov-chains, which means the transition probability matrices vary in time.

The models are differentiated by the type of occupancy data used to calibrate the model. The Page model was based on occupancy data from five single-occupancy office rooms. By contrast, an important feature of the Richardson–Thomson model was that it addressed the issue of the lack of availability of occupancy data with which to calibrate models. This was done by inferring occupancy from the UK time-use survey. Time-use surveys are large nationally representative surveys of how people use their time, which typically contain many thousands of 24-h diary entries. The UK time-use diaries detail participant location and activity at 10-min resolution, and allowed the introduction of the concept of 'active occupancy' – defined as when an occupant is at home and not asleep.

The technique of using a first-order Markov chains combined with national time-use surveys has been widely adopted in the literature. Widén developed a similar model based on the Swedish time-use survey [19]. López-Rodríguez et al. implemented the technique to develop an active occupancy model based on the Spanish time-use survey [20]. Muratori et al. used it to develop an activity model based on the American time-use survey [21].

The first-order Markov chain technique has also been widely adopted to develop models of occupancy in office buildings [22–24]. There are, nonetheless, important ways in which the literature can be distinguished. Broadly, four trends can be identified, as described below.

### 2.3. Techniques to reduce data input requirements

Given that one of the main drivers for the development of occupancy models was to address the lack of availability of real occupancy data, it is perhaps unsurprising that research has focussed on further techniques to produce synthetic occupancy data given less input data. This has been of particular interest in the field of non-domestic occupancy models.

Page's technique rearranges the Markov-chain formulae such that they depend on the probability of presence and a 'parameter of mobility' [18] – a measure of the likelihood that occupants change state. Wang's technique estimates transition probabilities based on 'expected sojourn times', or the mean time occupants spent in a state [22]. Both techniques are proposed to simplify data input requirements. Page argues that the 'probability of presence is a rather standard input' for simulations which 'should be available to the user'. Wang's technique is proposed to 'further simplify the specifications for [transition probability] matrices'.

Both techniques, however, still require a basis for estimating the model parameters. Page's 'parameter of mobility' is derived from the transition probabilities, while Wang's 'expected sojourn time' should be based on a representative sample of occupancy data. To be accurate therefore, both techniques still have a requirement for actual occupancy data, in which case it is arguably simpler to use this data to calculate and use the transition probabilities directly. In the absence of detailed occupancy data, however, simplifying assumptions about occupancy are required, and these two