The vernacular architecture of household energy models

Abstract

Energy use in buildings is driven by a socio-technical system providing energy services to building occupants. Despite the irreducible presence of people in the system, they are currently poorly represented in mainstream building energy models. This poor representation stems at least partly from the dominant 'folk ontology' of the predominantly physical science trained modelling community with their strongly visual, physical and causal mental image of the modelling system they are representing in their model descriptions. The introduction of new methods of embedding more sophisticated models of occupants within these models using Bayesian Networks shows promise, but presents its own set of challenges. Such models endogenise uncertainty, making correct interpretation of modelled results difficult, and create model structures that are theory agnostic to the fields the model's variables are draw from. Such models do, however, have advantages in multidisciplinary modelling environments where such theory agnosticism can provide a neutral territory for debate, and their graphical representation makes them useful vehicles for negotiating understandings between disciplines.

Introduction

"People use energy – not buildings" is an accepted truism of energy in buildings research. Energy use is driven by occupants' needs and desires for energy services (heat, light, hot water, 'infotainment') and the technologies through which they are supplied. In this complex socio-technical system, houses built to identical technical specifications will vary in their energy consumption by a factor of three once occupied. Despite, and in part because of this, occupants are only crudely represented in building models. This paper explores a range of these representations, from new approaches to established methods, and through them the entailed vernacular epistemologies of the modellers who construct them.

The paper proceeds as follows. An example of modelling occupant influences on energy use in homes using Bayesian Networks is given. This focuses on modelling the factors influencing internal temperatures in homes, as internal temperature is the single largest determining factor of UK domestic energy use. This will include a brief introduction to Bayesian Networks, their methods of construction, the data underpinning their construction, their central assumptions and their graphical representation. The paper then discusses this example in the context of the definition of modelling, the practice of model construction in the wider building energy modelling field, the use and function of models in research and policy making, and the dominant epistemologies entailed in these modelling practices.

Modelling home internal temperatures using Bayesian Networks

Introduction

Energy is used in buildings to provide services for occupants. Despite this, the influence of occupants on home energy use is usually very poorly represented in building energy models. The current dominant approach is through the use of what are called 'occupancy schedules'. These are used to standardise some of the known occupancy-related influences on energy use in homes. One of the most important of these is internal temperature. Internal temperature is the parameter to which BREDEM class models are most sensitive. BREDEM class models are important in the UK, as they form the basis of the most widely used building stock models, and underpin the UK Standard Assessment Procedure (SAP) for Energy Performance Certificate rating of homes. As Firth *et al* (2010, p.33) note in their analysis of the sensitivity of their BREDEM based Community Domestic Energy Model *"The heating demand temperature (which in most cases is the thermostat set-point temperature used in the dwelling to control the heating system) results in the most sensitivity… [This] suggests that heating demand temperature is the key determinant of CO_2 emissions in housing.*"
Currently these models vary mean internal temperature as a function of purely technical variables. For example, the SAP rating system varies mean internal temperature of living areas as a function of the Heat Loss Parameter (a measure of the thermal efficiency of the envelope), heating type, heat gains and heating controls (BRE, 2008b). It is therefore both an important parameter, and one believed to be influenced by both technical and non-technical (socio-demographic and behavioural) variables. This makes a natural target for extending the representation of people within building energy models. One way to do this is through construction of statistical models of the factors influencing domestic energy use. The approach illustrated here treats physical and occupant related influences equivalently, by building statistical models from empirical data as measured in the CaRB Home Energy Use survey of English homes (Shipworth, et al., 2010). The models are built from data using Bayesian Networks. This internal temperature statistical model is designed to replace the exogenous default set-point temperatures used in BREDEM models. It allows for prediction of living room temperature in two degree bands ($< 17 \leftrightarrow 19 \leftrightarrow 21 \leftrightarrow 23 \leftrightarrow 25 >$) in English homes based on a relatively small set of variables.

Bayesian networks

Bayesian Network models consist of a set of variables called ‘nodes’, and a set of links joining related variables called 'edges'. A generic example of such a model is shown in Figure 1. In Figure 1, each circle represents a variable, each arrow represents a relationship between variables, each variable contains within it a conditional probability table determining the nature of the probabilistic relationship between each variable and its 'parents', i.e. those nodes linked to it where the links are pointing towards it. Mathematically, such an object is called a ‘graph’, hence Bayesian Networks are referred to as ‘graphical’ models. In order for the algorithms calculating the probabilistic interrelationships between variables within Bayesian Networks to work, the network cannot have any cycles. The links are therefore directed ('parent' to 'child') and the networks are termed Directed Acyclic Graphs (DAGs).

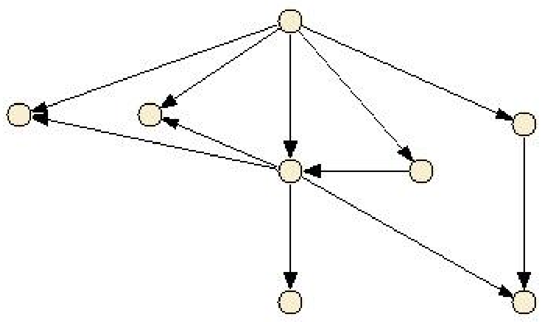

Figure 1: Bayesian Network showing variables (circles), directed relationships (edges), and no cycles.

Bayesian Networks are one type of statistical graphical model. As Jordan (1999) notes in the introduction to 'Learning in Graphical Models':

> *'Graphical models are a marriage between probability theory and graph theory. ... Probability theory provides the glue whereby the parts are combined, ensuring that the system as a whole is consistent, and providing ways to interface models to data. The graph theoretic side of graphical models provides both an intuitively appealing interface by which humans can model highly-interacting sets of variables as well as a data structure that lends itself naturally to the design of efficient general-purpose algorithms.' (p.1)*

A Bayesian Network defines a joint probability distribution specified over a set of variables and their relationships as defined by the structure of the graph. The structure of the graph explicitly encodes conditional interdependencies between the variables. A graph-theoretic definition is provided by Bottcher & Dethlefsen (2003, p.2):

"Let D = (V,E) be a Directed Acyclic Graph (DAG), where V is a finite set of nodes and E is a finite set of directed edges (arrows) between the nodes. The DAG defines the structure of the Bayesian network. Each node $v \in V$ in the graph corresponds to a random variable X_v. The set of variables associated with the graph D is then $X = (X_v)_{v \in V}$. Often, we do not distinguish between a variable X_v and the corresponding node v. To each node v with parents pa(v) a local probability distribution, $p(x_v|x_{pa(v)})$, is attached. The set of local probability distributions for all variables in the network is P. A Bayesian network for a set of random variables X is the pair (D,P)."

Because not all nodes in the network *D* are directly linked, some nodes are conditionally independent of other nodes. This allows the joint probability distribution over all the nodes to be factorised into the product of a series of conditional dependencies between nodes (Equation 1). It is this capacity of the DAG to structure the joint probability distribution over the set of variables *V* that makes calculating the joint probability distribution tractable.

$$p(X_1, X_2, ..., X_V) = \prod_{v \in V} p\left(x_v \middle| x_{pa(v)}\right)$$

Equation 1: Factorization of the joint probability distribution

Constructing Bayesian networks

There are two epistemically distinct approaches to the construction of Bayesian Networks: elicitation and learning. Elicitation is the process of eliciting the structure and probabilities for networks from domain experts and is widely used in constructing applied Bayesian Networks in environmental and other fields (O'Hagan, 1998). Learning is the application of algorithms to extract the structure and probabilities from datasets. Learning approaches have grown from developments in the data-mining and artificial intelligence fields (Mackinnon and Glick, 1999). These approaches can be combined, and where this is done elicitation is often used to determine the structure of the networks, and learning to extract the probabilities within the models from datasets. In multidisciplinary environments however, elicitation of either network structure or probabilities can be problematic, as there is no single natural domain of experts from which to elicit. In the domain of energy use in homes, building physicists, building services engineers, economists, sociologists and social psychologists all hold different beliefs about which factors to measure and model, how these factors are linked, and the strength of the relationships between them. Pilot work conducted on elicitation of network structure showed so little agreement between researchers within and between fields that this approach was rejected in favour of the use of learning algorithms for both determining the network structure and probabilities.

Bayesian network construction

There are five main steps in the construction of Bayesian Networks. The first, variable selection, encodes existing findings from a range of disparate cognate disciplines into the model through the process of variable choice. The second, instrument development, encodes a range of different epistemologies and methods into the model through instrument development. The third, variable measurement, similarly encodes domain specific methods and research designs into the data. Statistical models are only as good as the data they are built on so variable selection, instrument development, and variable measurement is as integral a component of Bayesian Network construction as choices of modelling algorithms.

Variable selection

The CaRB team of social scientists, building physicists and building technologists reviewed literature from many fields including building science, sociology, psychology and economics. Variables were selected on the basis of evidence from prior studies of the strength of their influence on home energy use. An implicit epistemic filter was therefore applied through the requirement to satisfy largely unexplicated notions of 'evidence' which themselves would have varied between fields. Variables were identified across a wide range of topics including: energy use; internal temperature; fuel types; built form; heating type and its usage; heating system controls and their usage; ventilation; occupancy patterns; bathing technology and practices; household appliances and practices; household socio-demographics; attitudinal measures; health measures; and comfort practices.

Variable instrument development

Instruments were developed to measure each variable to meet a range of pragmatic constraints. Because Computer Assisted Personal Interview (CAPI) was to be the dominant method, instruments for many variables were drawn from the UK Office of National Statistics harmonised methods. Additionally, some instruments were designed to harmonise with other current or historical surveys to which the findings were to be compared. In many cases these instruments were known to be imperfect, but were used to allow comparability across datasets for analysis purposes. Given the breadth of variables measured, instruments adopted and adapted from disparate fields themselves embedded these fields' theories and methods into the data.

Variable measurement: The CaRB Home Energy Survey

Sample size calculations for Bayesian Network analysis is best done using subject to item ratios. These are determined by the size of the local conditional probability table within each variable, not by the joint probability distribution of all the variables. Using this method, and the widely accepted 10:1 subject to item ratio used in PCA and EFA (Costello and Osborne, 2005) & (Osborne and Costello, 2004), a sample size of 500 homes was established.

For a sample of this size, face to face social surveys and inexpensive physical instrumentation were selected as the methods best able to minimise uncertainty in measurements within the logistical and financial constraints.

The Survey was based on a representative sample of English households selected by stratified random sample drawn from the Postcode Address File sampling frame. To ensure a good geographic and socio-demographic spread, postcode sectors were stratified by Government Office Region and by the percentage of households where the Census Household Reference Person was in NS-SeC categories 1 or 2. Fifty-four postcode sectors were selected at random in proportion to the number of addresses they covered, and 21 addresses were sampled in each selected postcode sector. Out of 1134 addresses, 427 households were interviewed – a response rate of 44%. For additional detail of the CaRB HES survey see (Shipworth, et al., 2010).

Raw data preparation and interpretation was conducted and SPSS syntax for construction of computed variables written. At time of model building over 100 variables (~10% of raw variables) were developed and assessed for inclusion in the temperature classifier model.

Temperatures within the CaRB HES sample were monitored continuously from mid July 2007 to early February 2008 and the average over each 45 minute period logged. The temperature measurements used as the dependent (target) variable in the network was the average winter temperature in the main living room at 20:15. These were measured with Hobo UA-001-08 temperature loggers with a manufacturer reported accuracy of ±0.47°C at 25°C which were calibrated prior to instillation. These were placed during the interviews by occupants with trained interviewer guidance between knee and head height away from direct heat or sunlight.

Network construction part 2: model selection and conditional probability learning

Variable model selection (network structure learning)

Determining the structure of a Bayesian Network is at least as important as determining the conditional probabilities linking the variables. Druzdzel and van der Gaag (2000) note that '*Experience with constructing probabilistic networks for various domains of application has established a consensus that the graphical structure of a network is its most important part...*' p.483. This importance arises for two reasons. Firstly, the output of the network is more sensitive to changes in the structure than to changes in conditional probabilities. Secondly, the amount of data required to construct a network is heavily dependant on the number of conditional probabilities required, which is in turn heavily dependent on the structure (Laskey and Mahoney, 2000).

Learning network structure from data however is computationally intensive and involves heuristic search over the space of possible networks, i.e. all possible ways to connect the variables in the models. The number of all possible networks, the size of the 'model-space', grows super-exponentially with the number of variables. Consequently, where there are more than around eight nodes, heuristics are needed to search for the best model within the model-space.

Structure learning, i.e. selecting the best possible way to link the variables within the model space, is a heuristic search of the model space for models which maximise the Bayesian Information Criterion (BIC)(Akaike, 1979). BIC is a widely used measure of model fitness and has been adapted for use with incomplete, case based datasets by O. Francois (François and Leray, 2006, François, 2008).

Variable conditional probabilities (parameter learning)

Within each variable in the network is a Conditional Probability Table (CPT). The probabilities in the CPT determine the strength of the relationships between variables in the model. The size of the CPT is the Cartesian product of the number of states of the variable and all its parents, for example, a variable with 3 states, and three 2-state (binary) parents, would have a CPT with 3*2*2*2= 24 conditional probabilities to be learnt from the data. The data requirements for parameter learning of Bayesian networks are therefore strongly related to both how the variables are discretised, and the structure of the network.

Representation of results

Presented is an example of a temperature classifier model constructed using all available variables in the CaRB Home Energy Use survey. The representation shows the variables (as outlined in Table 1) and the graphical structure of the model with the strength of relationships between variables indicated using Blind Average Link Strength percentage. This means that for any two linked variables, knowing what state one of the pair is in, increases what we know of the other, by reducing our uncertainty of its state by the percentage shown.

Living Room Temperature classifier model 1

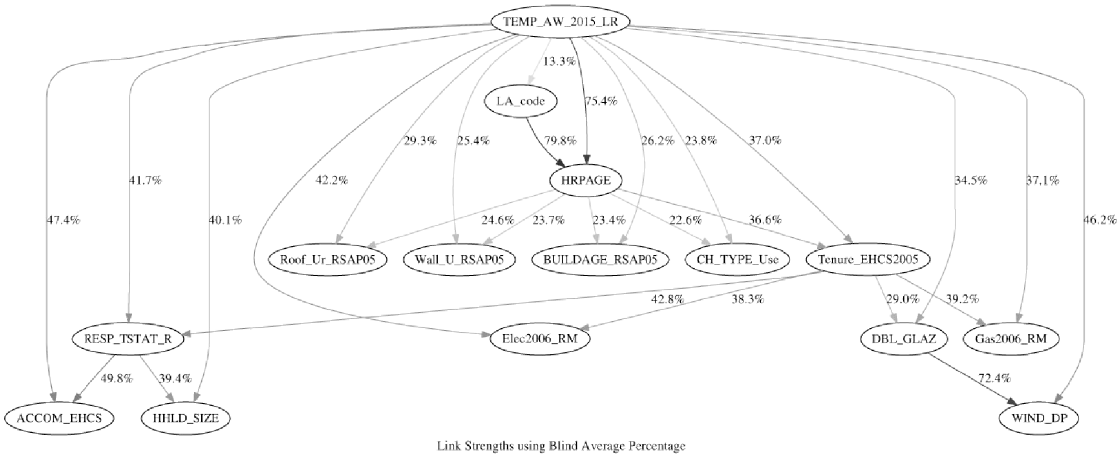

Figure 1: CaRB Home Energy Use Survey room temperature classifier model.
Target variable TEMP_AW_2015_LR with states: < 17 ↔ 19 ↔ 21 ↔ 23 ↔ 25 > (°C).
Correct classification rate: 85%
Expected value classification rate: 75% of cases within 10% of their observed value.

Table 1: Variables used in Living Room Temperature classifier model 1

Variable Name	Variable Description	Blind average link strength percentage
TEMP_AW_2015_LR	Avg. main living room temp. in Winter at 20:15 (measured)	Not Applicable
HRPAGE	Age of Household Reference Person	75.4%
ACCOM_EHCS	Accommodation Type (EHCS 2001 harmonised)	47.4%
WIND_DP	Proportion of windows draught-proofed	46.2%
Elec2006_RM	Electricity consumption (2006 BERR data)	42.2%
RESP_TSTAT_R	Thermostat setting (respondent reported)	41.7%
HHLD_SIZE	Number of occupants in household	40.1%
Gas2006_RM	Gas consumption (2006 BERR data)	37.1%
Tenure_EHCS2005	Tenure (EHCS 2005 harmonised)	37%
DBL_GLAZ	Proportion of windows Double Glazed	34.5%
Roof_Ur_RSAP05	Roof U-value (Respondent reported - RD SAP 2005 harmonised)	29.3%
BUILDAGE_RSAP05	Building Age (RD SAP 2005 harmonised)	26.2%
Wall_U_RSAP05	Wall U-value (RD SAP 2005 harmonised)	25.4%
CH_TYPE_Use	Central Heating type used in winter	23.8%
LA_Code	Local Authority	13.3%

Note: RD SAP = Reduced Data Standard Assessment Procedure

Discussion of the Bayesian Network approach

The approach discussed above differs markedly from that used in the rest of the buildings energy modelling field. Unlike conventional modelling approaches, the relationships between the variables are not dictated by theory directly. They are, at least ostensibly, empirical relationships learnt from data. Theory does play a role in the selection of variables to measure, but as these are drawn from a variety of different fields with different theories, this process is quite indirect. Information theory does play a role in determining the relationships between variables, but this is likewise agnostic to any correspondence relationship between the model and what it represents. In part because of this, and because they contain sets of variables not included in current domain specific models, the models seem to have a greater capacity for surprise. This capacity for surprise seems also to stem from the fact that the relationships between the variables are also not simply an encoding of existing domain specific theories into the model, and therefore frequently challenge existing beliefs.

As a guide to further empirical work the models have countervailing effects. Sensitivity analysis conducted on the models highlights those areas where improved data is likely to yield the greatest benefits. In contrast to this however, statistical models endogenise instrument error, making it indistinguishable from the underlying aleatory uncertainty of the variable, thus lack of strength or sensitivity to a variable in the model may result from measurement error. This is particularly problematic where different variables were measured with different instruments with highly variable but unquantifiable levels of instrument error. Reaching robust conclusions about the relative importance of different variables is therefore problematic and a matter of both methodologically informed judgement and model testing.

The Bayesian Network approach does however have a range of benefits. Its instrumental benefit lies in creating a classifier model of internal temperature. Its conceptual benefit lies in having a clear model structure that is easily to interpret and discuss. Its multidisciplinary benefit it lies in bringing together a range of variables from across disciplines in a 'theory agnostic' environment which can form the basis for sparking healthy debate within and between disciplinary communities. Such models can therefore form a neutral territory where disciplines can find a common graphical representation of model variables and structure they can understand, engage with and debate over, in order to elucidate methodological and theoretical differences.

Discussion of broader modelling issues

This section is a 'reflective practitioners' view on the process of model definition and construction in the broader field of building energy and energy systems modelling.

John Casti (1992, p.2) defines mathematical modelling as follows:

“Our viewpoint is that the study of natural systems begins and ends with the specification of observables describing such a system and a characterization of the manner in which these observables are linked.”

In many ways, this is an appealing and succinct definition of the modelling process, and it maps neatly onto the processes of Bayesian Network construction outlined by Druzdzel & van der Gaag (2000) of:

- Identification of the domain variables;

- Identification of the relationships between these variables and;
- Identification of the probabilities describing these relationships

While this process describes the mechanics of modelling, it does beg a range of more fundamental questions:

- How are we to define our 'natural system' and its boundaries?
- What forms of 'observation' constitute valid knowledge?
- What forms of knowledge form a valid basis for determining the relationships between variables?

Casti's definition however largely severs modelling from the rest of the scientific process, particularly from questions of theory and measurement which so dominate statistical modelling.

Ruttkamp (2002 p.17), with reference to the process of science per say, observes:

“The only way in which we can have scientific contact with the world…is through actions involving selection, abstraction, and generalisation, which are always executed within some theoretical framework or disciplinary matrix….”

This serves to set Casti's initial process of 'the specification of observables' into its necessary theoretical and disciplinary context. In the context of energy use in buildings generally, or the internal temperature classifier model in particular, this leads to two disciplinary questions: Through which theoretical framework(s) are we observing the role of people's influence on energy in buildings? and; 'through which disciplinary matrices do we look at the role of people's influence on energy in buildings and how do these colour both how and what we choose to observe?

To answer these questions we need to step away from the construction of individual models, around which these definitions and questions appear to be framed, and look at the practice of building energy modelling. The temperature classifier model presented above is designed to endogenise one currently exogenous variable, household internal temperature, within BREDEM class models. As discussed above, BREDEM class models are the dominant building energy models used in the UK. These models also form the core of BREHOMES, the dominant national building stock models, which are based on the BREDEM individual building energy modelling engine. However these models are themselves not monolithic. They are conglomerations of smaller models (e.g. physics-based models of heat-flow through walls), empirical observations (laboratory based measurements of the thermal conductivity of materials), parametric models from data (hot water use as a function of floor area) compliance ratings of pieces of technology (efficiency ratings of boilers), normative accounting standards (the carbon intensity of energy generated by different technologies), simplifying assumptions (steady-state heat-flow), known omissions (ventilative heat loss as a function of external wind-speed) and deliberate exclusions (occupant window opening behaviours).

This picture resonates with Winsberg's (2009 p.837) description of the epistemology of simulations. He argues that:

“[T]he knowledge produced by computer simulations is the result of inferences that are downward, motley, and autonomous. They are downward in…that [they] are drawn (in part) from high theory, down to particular features of phenomena… They are motley in that they draw on a wide variety of sources. These include theory, but also physical insight, extensive approximations, idealizations, outright fictions, auxiliary information, and the blood, sweat, and tears of much trial and error. … Finally, they are autonomous

in the sense that the knowledge produced by simulation cannot be sanctioned entirely by comparison with observation."

Building energy modelling exhibits all three of these characteristics. They are downward in that much of the design and behaviour of these models stems from their basis in steady-state heat balance equations from thermodynamics. The system is known to be dynamic, with the temperature gradient across the envelope of the building known to vary spatially and temporally, and the thermal resistance and inertia of the envelope being similarly dynamic, yet applied building physics theory and measurement ignores these effects.

Building energy modelling is "motley". There are currently no modelling methodologies within building energy modelling (at individual building or stock level) that distinguish between data of different types and quality, and maintain that distinction through the model in a form transparent to either model builders or users. Building models are epistemic 'sausage machines' – combining inputs of all qualities and types into outputs of homogenous and indeterminate quality and type.

Building energy modelling is autonomous. Many energy models, including BREHOMES, the UK national building stock model, are calibrated by top-constraint against data from the Digest of UK Energy Statistics (DUKES) on energy use in the UK housing stock (Shorrock and Dunster, 1997). Top calibration is done by adjusting the variable within the model to which output is most sensitive, internal temperature. This endogenises all model error within one variable. This has led, through a process of back-calibration of the model (built in 1990) against yearly data back to 1970, to the conclusion that average household internal temperatures in the UK have risen from an average of 12° in 1970, to an average of 18° in 2006 (BRE, 2008a), a figure widely cited in government (CCC, 2009). Against a backdrop of little observed reduction in energy use in the UK housing stock, this has lead to blaming occupants for 'taking-back', through improved comfort, the energy efficiency improvements assumed to follow from decades of Government regulatory and programme initiatives in the area. This has, in its own way, created a strong and largely negative 'model' of occupant behaviour in the minds of policy makers.

This example of model underdetermination is analogous to that highlighted by Betz (2009) as applying to modelling scenarios, but one applicable to the more immediate problem of model calibration. Betz cites Quine's famous statement on undetermination of logical linguistic statements:

"[T]he total field [of logically connected statements] is so underdetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to re-evaluate in the light of any single contrary experience. (Quine, 1953 p. 42f)"

In the context of model calibration, this translates into that fact that if our models don't match our observations, there are usually a wide variety of things we can adjust to calibrate them, and too little data to decide which the is the right element to adjust. Following Betz arguments, this makes the values of internal variables in such models calibrated by top-constraint epistemically equivalent to scenarios – i.e. it renders them descriptions of what is *possible* – not what is *factual*. The wide-spread application of top-constraint model calibration in energy modelling creates a problem for deterministic models comparable to that of multicoliniarity in statistical models. That is, while the gross output may be true, the accuracy of individual variables within the model is not, and of necessity, it is at the level of individual variables that policy interventions are targeted. In addition, as Winsberg argues, simulation is partially epistemically autonomous and is undertaken in areas where data is sparse and insufficient to support decision making, with again the effect that energy stock modelling constructs representations only of possible futures between which it is not empirically possible to distinguish. For both these reasons, models' epistemic authority rests as much or more on the veracity of their construction, as on methods of validation, suggesting far more care needs to be taken to track data type and quality through our models to allow us to understand how and where the impacts become manifest in our results.

Extending Winsberg's arguments, Godfrey-Smith argues for the emergence of what he calls 'model-based science' (2006). As he puts it *"…[This involves] a rejection of the idea that modelling is a mere heuristic adjunct to the real business of theory-construction*" (2006 p.730). The central components of Godfrey-Smith's model-based science seem well reflected in the practice of energy systems modelling generally, and energy in buildings modelling in particular. Godfrey-Smith's work builds on Giere's which distinguishes between the specification of the *model system*, and (frequently multiple) *similarity* relations between the model system and the *target system* (the world). Godfrey-Smith distinguishes between the *model system* (some abstracted and simplified 'imagined concrete thing'), the *model description* (a description of the model system in some language – mathematical symbols, graphical representation, words), and the *similarity relation* (that aspect of the model system that a modeller regards as similar to the real-world target system). Godfrey-Smith's treatment of model systems as 'imagined concrete things' is particularly relevant for energy in buildings research. He notes: "*It is important to the practice of model-based science, at least some of the time, that model systems can be conceived and treated in a more concrete way. Roughly, we might say that model systems are often treated as "imagined concrete things" – things that are imaginary or hypothetical, but which would be concrete if they were real*". (2006 p.735). It is clear from personal experience that modellers working in building energy modelling suffer from a form of physical literalism and an over-familiarity with their subject matter of their models. We spend over 90% of our lives in buildings, and this immediacy feeds into what may be called a completionist mentality, particularly in physical science trained building energy modellers. They can walk around buildings pointing to every energy consuming appliance and model its energy demand as a function of its rated power consumption, its load, and the duration of its use. They therefore feel that their 'imagined concrete thing' must represent everything they can see. This frequently leads to an insistence on modelling even trivial energy consuming technologies, irrespective of the quality of the available data. It has a complementary consequence in that things that cannot be literally seen, or are without explicit physical causal mechanisms, tend not to be modelled. Statistical relationships without plausible causal mechanisms visualisable by the modellers within their model systems are discounted. Similarly, issues of the diversity of buildings, technologies, and patterns of occupancy are not easily accommodated in model systems which are so instantiated and cemented in specific individual buildings within the modellers' experience. This is reflected in the structure of national building stock models like BREHOMES (and virtually all others) which are constructed by modelling a relatively small number of individual 'archetypal' buildings, and multiplying them by their frequency in the national stock. Again these 'archetypal'

buildings are defined by overt physical characteristics (terrace house; semi-detached house; flat, etc) rather than through any identification of characteristics which may cluster the households into statistically distinct groups in terms of their energy consumption. The dominance of the visual resemblance relation (the mental image) of the 'imagined concrete thing' that is the energy use in buildings' *model system* is so powerful as to largely determine the structure and content of models in the field. Godfrey-Smith refers to these 'imagined concrete things' as a form of 'folk ontology' of a field. In the building energy modelling field that which exists is the visual, physical and causal – that which doesn't is the invisible, variable and correlational. This is the 'vernacular architecture' of building energy models.
There remain yet other characteristics of building energy modelling which seem unexplained by the model-based science approach. There is an 'agenda power' or 'narrative power' that stems from the developers of specific energy efficiency technologies, and then flows into the energy modelling community. The Energy modelling community adopts the predominantly technical agenda of the renewable energy and energy efficiency technology development community. These technology communities are themselves driven in large measure by government and social agendas on climate change, energy security and fuel poverty. This agenda and the accompanying narrative, frame the way the modelling community look at, and integrate, each new technology or intervention into their models. Industry established product ratings for efficiency or generation are often calculated for ideal conditions in laboratories, for perfectly installed demonstration equipment, operating in isolation from other elements of building technology, building fabric, and occupancy related factors. This agenda focuses on the scope for potential benefits, and not on the scope for average, or underperformance from real-world conditions or systems-level conflicts. This creates a kind of technological optimism in the models that overestimates performance by factors of 50% to 500%.

Conclusion

Energy use in buildings is driven by a socio-technical system providing energy services to building occupants. Despite the irreducible presence of people in the system, they are currently poorly represented in mainstream building energy models. This poor representation stems at least partly from the dominant 'folk ontology' of the predominantly physical science trained modelling community with their strongly visual, physical and causal mental image of the modelling system they are representing in their model descriptions. The introduction of new methods of embedding more sophisticated models of occupants within these models using Bayesian Networks shows promise, but presents its own set of challenges. Such models endogenise uncertainty making correct interpretation of modelled results difficult, and create model structures that are theory agnostic to the fields the model's variables are draw from. Such models do, however have advantages in multidisciplinary modelling environments where theory agnosticism can provide a neutral territory for debate, and their graphical representation makes them useful vehicles for negotiating understandings between disciplines.

References

Akaike, H. (1979) A bayesian Extention of the Minimum AIC procedure of Autoregressive Model. *Biometrika,* 66**,** 237-242.

Betz, G. (2009) Underdetermination, Model-ensembles and Surprises: On the Epistemology of Scenario-analysis in Climatology. *Journal for General Philosophy of Science,* 40**,** 3-21.

Bottcher, S. and Dethlefsen, C. (2003) Deal - A package for Learning Bayesian Networks. *Journal of Statistical Software,* 8.

Bre (2008a) Domestic Energy Fact File. Watford, Building Research Establishment.

Bre (2008b) SAP 2005 specification with 9.82 update. 2005 Revision 2 ed. Watford, Building Research Establishment.

Casti, J. L. (1992) *Reality rules: 1 - Picturing the world in mathematics - The fundamentals,* New York, John Wiley and Sons.

Ccc (2009) Energy use in buildings and industry: Technical appendix. London, Committee on Climate Change.

Costello, A. and Osborne, J. (2005) Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research and Evaluation,* 10

Druzdzel, M. and Van Der Gaag, L. (2000) Building Probabilistic Networks: 'Where Do the Numbers Come From?' *IEEE Transactions on Knowledge and Data Engineering,* 12**,** 481-486.

Firth, S. K., Lomas, K. J. and Wright, A. J. (2010) Targeting household energy-efficiency measures using sensitivity analysis. *Building Research & Information,* 38**,** 25 - 41.

François, O. (2008) Efficient Bayesian Network Learning Using EM or Pairwise Deletion. *The Fourth European Workshop on Probabilistic Graphical Models - PGM'08.* Hirtshals, Denmark.

François, O. and Leray, P. (2006) Learning the Tree Augmented Naive Bayes Classifier from incomplete datasets. *Third European Workshop on Probabilistic Graphical Models (PGM'06).* Prague.

Godfrey-Smith, P. (2006) The strategy of model-based science. *Biology and philosophy,* 21**,** 725-740.

Jordan, M. (Ed.) (1999) *Learning in Graphical Models*, MIT Press.

Laskey, K. and Mahoney, S. (2000) Network engineering for agile belief network models. *IEEE Transactions on Knowledge and Data Engineering,* 12**,** 487-497.

Mackinnon, M. J. and Glick, N. (1999) Data Mining and Knowledge Discovery in Databases - An Overview. *Australian & New Zealand Journal of Statistics,* 41**,** 255-275.

O'hagan, A. (1998) Eliciting Expert Beliefs in Substantial Practical Applications. *The Statistician,* 47**,** 21-35.

Osborne, J. and Costello, A. (2004) Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research & Evaluation,* 9**,** 11.

Quine, W. V. O. (1953) Two dogmas of empiricism. IN Quine, W. V. O. (Ed.) *From a logical point of view.* Cambridge, Harvard University Press.

Shipworth, M., Firth, S., Gentry, M., Wright, A., Shipworth, D. and Lomas, K. (2010) Central heating thermostat settings and timing: building demographics. *Building Research & Information,* 38**,** 50-69.

Shorrock, L. D. and Dunster, J. E. (1997) The physically-based model BREHOMES and its use in deriving scenarios for the energy use and carbon dioxide emissions of the UK housing stock. *Energy Policy,* 25**,** 1027-1037.
Winsberg, E. (2009) Computer Simulation and the Philosophy of Science. *Philosophy Compass,* 4**,** 835-845.