Proceedings of Building Simulation 2011:
12th Conference of International Building Performance Simulation Association, Sydney, 14-16 November.

HOW USEFUL ARE BUILDING ENERGY MODELS FOR POLICY? A UK PERSPECTIVE

Alex J Summerfield[1,2], Rokia Raslan[1], Robert J Lowe[1] and Tadj Oreszczyn[1]
[1]UCL Energy Institute, University College London
[2]email: a.summerfield@ucl.ac.uk

ABSTRACT

Energy demand models are central to the efforts of many governments to reduce carbon emissions from buildings. The lack of empirical research to ensure the appropriate use of predictions from the models has implications for building regulations and evaluating policy initiatives. We present three recent examples from the UK that highlight challenges: the discovery of a heat by-pass in party walls, trends in household gas consumption and the impact of condensing boilers, and inter-model variation in the non-domestic sector. We emphasise and contrast the approach of health sciences to support policy, and suggest that a far more systematic and integrated approach between empirical research, model development, and policy evaluation is needed.

INTRODUCTION

"*Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful.*" (Box & Draper, 1987)

Energy performance models lie at the core of policies of many governments, aimed at reducing carbon emissions from the built environment, such as via building regulations. The Standard Assessment Procedure (SAP), that underpins the Building Research Establishment Domestic Energy Model (BREDEM), is the UK government's approved methodology for assessing the energy ratings of dwellings. Designs for new domestic buildings and those with major renovations are evaluated according to SAP for their estimated carbon emissions, in order to satisfy the standards of Part L (that addresses building energy performance) of the Building Regulations. Non-domestic buildings, though often more complicated than residential buildings, follow an equivalent regulatory process based on satisfying criteria for modelled or rated energy performance.

Numerous other countries operate similar procedures whereby certification for energy rating is part of the regulatory requirements or incentivised at the national, state, or local level. For instance, Europe is covered by the Energy Performance of Buildings Directive (EPBD), where various energy models (with standard calculation methodologies) are used for rating energy consumption of new and existing buildings. In the US there are certification initiatives, such as the National Green Building Standard (NGBS) for the domestic sector. In Australia, states have adopted various rating schemes, such as BASIX in NSW, that use predicted annual energy to help determine if the design has reached a prescribed minimum rating level or category.

This represents a major shift in policy over the last decade or so, yet there is scarce empirical evidence in peer-reviewed research on the reliability and accuracy of the underlying energy models either in terms of quantifying efficacy of individual energy efficiency measures or in energy reductions at the stock level. There are many papers on the energy performance of exemplar buildings, and of the energy modelling of innovative technologies. These are not representative samples, and it is often not clear if agreement between predicted and measured performance occurred only after re-modelling (to account for the specifications as constructed) or intervention in the building post-construction (to ensure the building operated as specified).

This has implications for predicting energy at the single building level, as well as for policymakers and utilities at the building stock level. For instance, there are no empirically based findings on the extent that the energy consumption of UK dwellings built since 2001 and 2005, when Part L of the Building Regulations were progressively tightened, differs from the older stock. Nor are there peer-reviewed findings for impact of different types of condensing boilers on energy consumption, one of the key aspects of the new regulations in the UK. The empirical basis of parameters used in energy performance models, assumptions in the calculation methods such as average indoor temperature settings, and the uncertainty implied in their predictions or simulations, also remains far from clear.

This position paper argues for the need to refocus on the guiding axiom for the use of models of Box and Draper (1987). Its aim is to clarify the implications of this approach for energy performance modelling and policy evaluation in the built environment, and the consequences for the research field. We present three examples from the UK that highlight some of the deficiencies in our current approach. It ends with a discussion of the contrast with the approach used in the health sciences and epidemiology.

METHODOLOGICAL ISSUES

The differences between predicted and actual energy consumption arise from a range of sources for uncertainty, including:

- Input data both in terms of the accuracy of the operator's entry of the building and material specifications, and the empirical basis for the performance values explicitly associated with different building components.
- Limitations in the calculation engine, for instance in not sufficiently accounting for pathways for thermal transmission, such as can occur thermal bridging; also inaccuracies in the values implicitly assumed in the calculation for the performance characteristics of elements.
- Differences between the building as modelled and the specification relevant as constructed, varying from replaced components to incorrect installation of equipment.
- Differences in the occupancy patterns or building operation to that specified, such as extended office hours.
- Issues with post-occupancy surveying and monitoring of the building, for instance in accessing inner wall construction or obtaining accurate annual energy data.
- If the results are to be generalised to the national stock, or some section of it, then issues of representativeness of the sample become important, including socioeconomic factors for housing.

As is the case for standard energy models in many countries, the calculations and assumptions contained in SAP are laid out in freely available documentation. However the implementation of SAP in approved software such as BREDEM is a 'black box', since it is impossible to inspect directly the implementation. Thus it becomes problematic to determine exactly the source of the discrepancy if two approved energy models generate disparate predictions for carbon emissions. Second, the empirical basis and the associated uncertainty for many of the assumptions used in SAP, such as for average indoor temperature with different heating systems, is not provided.

Model users and uses

Given the lack of verification of each energy performance rating, it can be reasonably expected that the main focus of many SAP users is whether the specification of their design can produce a SAP prediction for emissions that satisfies building regulations without incurring too many extra costs. There are a multitude of specifications available to be assigned, though for SAP the scope for potential 'adjustments' having a major impact have reduced substantially with the tightening of building regulations in recent years. Nevertheless, it may still be determined that rather than change construction method to reach external wall U-values, it is more cost effective to enlarge the floor area of dwelling slightly and hence lower the emissions per unit floor area. From the developers perspective that can be a rational decision even though it may result in higher overall emissions. Or they could choose to include have internal doors on a highly glazed section, so that it can be treated as a conservatory by SAP. If so, then it should be noted that survey results have indicated that occupants typically leave conservatory doors open through winter – if not remove them (Summerfield et al. 2008). For the interest of these users of SAP to extend to issues beyond design approval requires model predictions to be routinely compared with the monitored energy performance of the constructed building.

In this regulatory use, the predictions generated by SAP for design approval operate conceptually as a maximum limiting value whereby any higher values may mean the design fails to meet the criteria. Implicit in this notion is that the prediction is absolutely precise. Theoretically then, approval can hinge on the difference due to sizing of one window. When compared to the as-built design, however, the same SAP predictions might be expected to reflect the minimum expected energy consumption, since the performance of building components are more likely to be below specifications than above. There is simply more scope through construction variability for the energy consumption to be worse than expected. The recent requirement for air pressure testing and thermographic imagery of a sample of dwellings from each development, at least provides a first step in testing if the constructed building meets the standard expected given the specified design performance.

Moreover as an asset rating, SAP predictions are based on 'standard' climatic conditions and assumptions about 'standard' rather than necessarily average occupant behaviour, such as heating patterns for a given type of heating system. Such occupant influences on aspects of demand may have changed with demographics, lifestyle and consequent occupancy patterns of UK households over the last two decades, from when SAP was first developed. In the absence of empirical evidence for indoor temperature, one recent UK energy model has provided improved predictions for building stock consumption using occupancy patterns for heating that are derived from the household employment status data (Cheng & Steemers, 2011).

Policymakers at the national and local level are also key users of building stock energy models to help to formulate and evaluate policy measures. Developing models for the non-domestic sector is far more complicated than for the domestic sector, for instance due to the range of energy demand activities, from restaurants to gyms to offices, within a single building envelope. For the domestic sector the UK has used BREHOMES, which is also based on

BREDEM but essentially takes weighted averages of energy predictions according to the proportion of various dwelling typologies in the stock. This has been used to predict average energy consumption as the stock grows and also to estimate the impact of specific measures, such as cavity insulation (Utley & Shorrock, 2009). Unfortunately although it relies on extensive survey data of the stock, the disaggregated data used for the model are not published – so for instance we do not know the individual contributions from each typology that lead to a stock average heating heat loss. The extent that the average indoor temperatures are based on empirical data is also unclear, instead they are adjusted to help 'harmonise' the predicted stock energy with actual consumption (Shorrock & Dunster 1997).

Much of the data regarding detailed composition of the housing stock in BREHOMES is based on the English Housing Survey (EHS), which includes an extensive social survey as well as a building survey by qualified surveyors for around 8000 households per year on an on-going basis, describes housing conditions. For instance, the details of where – in building stock terms – wall cavity insulation has been installed. Unfortunately for the ongoing evaluation of BREDEM, BREHOMES, and other UK energy models the last occasion energy data was obtained for even a limited sub-sample was more than a decade ago.

RESULTS: THREE UK EXAMPLES

In the following, we explore results from recent work in the UK to illustrate how our current approach manifests as a lack of uncertainty in in the way energy models can currently be used to support and evaluate policy.

Heat loss and party walls

Lowe et al (2008) have described their action research approach to evaluating Stamford Brook, which was designed as a low-energy development with conventional appearance for UK dwellings. Their study included close observation of the construction process to identify and remedy potential issues as they arose, such as thermal bridging around windows. However co-heating tests of dwellings undertaken post-construction still found energy consumption exceeded that expected from the models of the design, and particularly for those dwellings with party walls. After considerable investigation, high temperatures were recorded in the roof-space at the top of the party wall. This confirmed that the heat loss was due to a lack of insulation or a cap for the cavity in the party wall. So instead of the assumption made in energy models that the party wall represented no heat loss (U-value =0), it had worse performance than the external walls. Simply, it was acting as a thermal chimney, bypassing roof insulation, and convecting heat directly into the upper roof space.

The significance of this result goes beyond its role as an interesting example about the need to fully understand the details of construction in determining fabric heat loss. Terraces and semi-detached dwellings are common in the UK stock, but it is not clear what proportion are cavity rather than solid party walls. The issue is of most significance for well-insulated dwellings built in the last decade, and for other dwellings as they are retrofitted. In 2010 the building regulations were amended to address the issue. But the real questions is, given evaluation and testing of energy models for more than two decades, why was it not detected sooner? The study highlights the importance of having an organised programme of action research in working through the complete cycle from design performance to measured performance, diagnosis and remedy of the disparities, and then feeding the resultant information back into the regulatory and modelling process.

Gas consumption and condensing boilers

Since 2005 UK building regulations have specified that new and replacement gas boilers should be high efficiency condensing boilers, which pass hot flue gases through a heat exchanger in order to preheat the water that is then heated directly from gas combustion. Boiler thermal efficiency is specified according to a SEDBUK rating which provides a seasonal efficiency and claims to provide a more accurate assessment of *in situ* performance with standard heating patterns and controls than that obtained under laboratory testing conditions. Seasonal efficiency of domestic Boilers in the UK (SEDBUK) ratings for winter thermal efficiency are typically around 90%.

Since 2005 around 1-1.2 million condensing boilers, or ~5% of gas connected households, are installed in the UK each year. With ~20% gain in efficiency over the standard boiler being replaced, this might be expected to result in ~1% reduction in total gas demand annually. So after 5 years of implementation we should expect to see a ~5% decline in annual household gas demand. Clearly the UK government would like to know if this is an effective policy initiative. Specifically: how many tons of carbon is this measure saving, amidst a raft of other policy initiatives including installation of cavity wall and roof insulation? And if it is not saving as much as expected, then how can the policy be revised?

UK quarterly delivered energy data since 1998 for household gas use shows a significant decline in recent years (figure 1). In previous work, Summerfield et. al. (2010) constructed a simple top-down model for household total delivered energy, though the results for gas consumption were equivalent. Their findings indicated that until 2005, variation in gas demand could be explained by temperature variation, but that from 2006 it departed from this model. This juncture marks not only the introduction of the relevant changes in building regulations, but it was also a period where gas prices began a steep rise, by 80% in real terms in 3 years to 2008. For data up to 2008, gas prices were a better explanatory variable for the changes in demand than the number of condensing boilers. In other words, no evidence was found for changes in demand beyond that explained by variations in temperature and gas price. The high gas prices could have led to behavioural change, such as turning thermostats down, or ‘structural’ change such as the installation of insulation and condensing boilers.

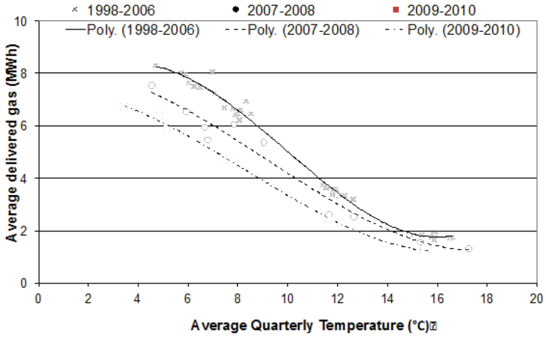

Figure 1: Household quarterly gas consumption in the UK vs. Temperature (1998-2010)

However, since 2008 gas prices have declined back to ~40% above their 2005 level, while gas consumption has continued to decline. As a result, the number of condensing boilers in the building stock has now replaced gas prices in the regression model as an explanatory variable for the change in gas consumption. The effect size corresponds to a ~5% *annual* decline and is considerably larger than the estimated ~1-2% annual decline due to introducing condensing boilers identified above. Even the summer consumption data, which mainly reflects hot water demand and is not influenced by levels of building fabric insulation, show a ~20% reduction over 5 years.

While this evidence supports the effect of condensing boilers, there are clearly other factors at work that have impacted at the same time as this energy policy measure. Overall we would expect the decline to reflect further ‘structural’ effects in the cooler seasons, such as due to increased wall and roof insulation. The installation of condensing boilers may also be highly correlated to other measures, such as improvements to hot water tank insulation. There has also been a shift to combi-boilers, which supply hot water on demand, and therefore do not require external cylinders. Furthermore, there may be other ‘additional’ effects, i.e. not structural, where occupant behaviour is an important factor. Since 2008 the UK has also been greatly affected by the global financial crisis and economic austerity measures. So one other explanation is that although gas prices declined recently, this variable does not reflect the household perception of prices as much as disposable income or other wealth related measures. So although we can say there has been a significant and incremental reduction in gas consumption since 2005, we are unable to identify from this data the contributions to savings from specific policy measures.

After five years of policy implementation, there are still no peer reviewed and empirically based estimates for contributions to reductions in gas consumption in UK dwellings from the raft of energy efficiency measures. Various studies or investigations are under way in the UK, including modelling within the Department for Energy and Climate Change (DECC) using household annualised energy data, however while these are useful they are typically have issues such as limited occupant reported building information and social data. They are typically not based around a standard study design with population based studies. Nor can the small scale studies that focus on a specific development or technology, and are often the topic of academic papers, serve the broader purpose for stock estimates.

Yet the uncertainty could be easily addressed at little extra cost through, for instance an augmented version of the existing EHS, that monitored gas and electricity use, especially as smart (high frequency electronic) metering is currently being installed. The impact of condensing boilers could be addressed through a randomly selected sample of households, and monitoring of their gas and electricity consumption and indoor temperatures before and after the boiler installation.

Inter-model variation

While we have focussed thus far on the domestic sector, building energy performance models play an essentially equivalent role in the non-domestic sector. Again to meet Part L of the UK building regulations and for an Energy Performance Certificate, designers need to achieve a prescribed standard by entering the building specifications into one of a number of approved versions software. Detailed accreditation and verification procedures have been established for both the quasi-steady state models, such as simple Energy Building Model (SBEM), and dynamic simulation models (DSM).

While the methodology for testing models has been described in some detail (Judkoff & Neymark, 2006), scarce recent findings are available that compare predicted with actual energy consumption in large samples of buildings. One exception is the report on the energy performance of LEED rating scheme for new construction in the USA, that used data from 121 mainly ‘medium energy use’ buildings (Turner & Frankel 2008). They found that while the LEED predictions were a relatively reliable predictor of the average performance of the sample, measured energy use for over half the projects deviated by more than 25% from the generated design projections. Among other issues they note that there was systematic underestimation for ‘high energy use’ buildings, which used on average 2.5 times the predicted energy use at the design phase. Turner and Frankel (2008) conclude that the energy performance of these buildings is not well understood. It is not known if the same issues arise with the software used in the UK or Australia as no comparable studies have been published.

It has been recognised for some time that a considerable range in the methods or algorithms employed *between* the various energy simulation tools used for calculating basic building physics lead to a divergence in the predicted energy performance (Judkoff & Neymark, 2006, Neymark & Judkoff, 2002). From their industry survey, Raslan and Davies (2006) also reported a similar concern by operators and other professionals regarding discrepancies between results from simulation software. They subsequently evaluated 12 simulation tools, SBEM, nine FI-SBEM and two DSM, using specifications for three single-zone models (shallow plan office, deep plan office and retail shed) that are considered to be representative of the main typologies that cover much of the UK non-domestic stock (Raslan & Davies 2009). They controlled for extraneous differences, for instance using the same experienced energy modeller throughout, but the findings indicated a disturbingly wide range in the predicted benchmark carbon emissions, with results varying by 50% or more and the two DSM tools consistently generating markedly lower estimates than other tools. Variations were also found in the calculated external areas and calculated U-values for building components. Although there were noted benefits in their relative simplicity to use, some of the FI-SBEMs were constrained by the inability of the current SBEM calculation engine to model a range of ventilation strategies, HVAC systems, and energy efficient lighting systems. They also recommended that testing procedures should be more consistent and rigorous and that the model be operated under test conditions that are the same as it is used in practice. Last, they point to the need to standardise the process of data entry and building specification. There has been no equivalent peer-reviewed research testing across designs for the numerous implementations of SAP software for domestic sector in the UK, but an Italian study has also found a high degree of intra-model variability for asset ratings for two simple dwellings (Milone et al., 2009).

This work raises the more fundamental question of why is energy performance software, which is crucial to effective policy implementation in this area, not under much closer scientific testing and scrutiny? The lack of clarity in the current situation may be advantageous for those not genuinely interested in energy performance of their buildings, but it undermines support the policy objective of lower carbon emissions.

DISCUSSION

We have provided a brief overview of some of the issues that currently permeate the use of building energy models both as part of the implementation of energy policy and in policy evaluation. We have used three examples to illustrate how the current situation can be remedied: the role for a systematic programme of active research following up discrepancies in designed performance; the need for large scale studies to unravel changes already occurring in energy demand at the national level; and need for a greater scientific scrutiny of models and their operation.

The disparity between predicted and actual energy demand may reflect everything from errors in data entry to software implementation, to construction compliance issues, to energy monitoring issues. What is needed is a far more integrated approach at all levels. We currently simply do not have the data or research programme in place to evaluate any of the uncertainty in a systematic way.

It is worthwhile to reframe the current predicament from the perspective of health sciences and epidemiology. Suppose that governments have embarked on the national eradication of an endemic disease, which in this case we might call *excessive energy consumption*; obesity levels represent an appropriate parallel. Our understanding of the technical mechanism of this disease has convinced the government to prescribe an array of *treatments* to dwellings (i.e. the population), such as insulation, condensing boilers, campaigns for behavioural change, and so on, both to the existing population and the 'new build'. In that medical and population health context, we might have expected the government and the research community to have *already* established:

- Population based longitudinal studies, i.e. using representative samples and gathering time series data, to provide an assessment of the prevalence and incidence of the energy disease as well as the influence of social factors, such as demographic changes. This identifies target groups or hard-to-treat sectors of the population or building stock. It provides a platform for on-going evaluation of the national policy effectiveness.
- Specific studies aimed at the evaluation of individual treatment measures and testing of intervention programmes for efficiency measures, such as via randomised control trials. This would involve pre and post intervention monitoring.
- A formal 'disease' reporting procedure and data collection, with standard protocols and methods, whereby professionals involved in administering treatment measures report their observations and concerns about their effectiveness, with systematic follow-up of cases where outcomes

need further investigation (such as when they disagree with the prediction of models).

- Findings from this publicly funded research are subject to peer-reviewed (rather than consultancy reports), access is given to disaggregated datasets, and subsequent models are open to scrutiny.

This comparison may seem wildly optimistic, yet it is essentially standard international practice in health sciences. The reason why it is not the case in building science may concern broader issues of culture and the value we place as a society on meeting energy performance targets, compared with improving population health. For instance, Williamson (2010) has been one of the few in this field to discuss the role of ethics in reliable design simulation.

In practical terms a great deal could be achieved at relatively little cost. For instance, we have suggested existing large-scale studies on housing can easily be augmented with the collection of energy consumption data. Verification procedures with follow-up of discrepancies should be a standard part of an on-going evaluation process. Lastly there should be far greater emphasis on and recognition of the statistical uncertainty in energy prediction results.

CONCLUSION

In summary, as policy requires the delivery of increasing reductions in energy demand from buildings, a far greater emphasis on empirical evidence is needed to for robust model development to support policy formulation. Overall a more integrated and scientific approach should be adopted, much as is standard practice in the health sciences. We need a clearer sense of the uncertainty in the results at the individual building and stock level; a better understanding of the extent and conditions where energy models are 'wrong', so we can determine how 'useful' they can be for improving energy performance.

REFERENCES

Box, G.E.P., Draper N.R. 1987. Empirical Model-Building and Response Surfaces, p. 424, Wiley. ISBN 0471810339

CEBR 2011 British Gas Home Energy Report 2011: An assessment of the drivers of domestic natural gas consumption. Centre for Economics and Business Research Limited, London.

Cheng, Vicky, and Koen Steemers. 2011. Modelling domestic energy consumption at district scale: A tool to support national and local energy policies. Environmental Modelling & Software 26 (10): 1186-1198.

Judkoff, R., Neymark, J. 2006. Model Validation and Testing: The Methodological Foundation of ASHRAE Standard 140. In ASHRAE 2006 Annual Meeting, 24–29June 2006, Quebec City, Canada: ASHRAE.

Neymark J. & Judkoff, R. 2002. Building Energy Simulation Test and Diagnostic Method for Heating, Ventilating, and Air-Conditioning Equipment Models (HVAC BESTEST) Volume 1: Cases E100–E200. International Energy Agency

Lowe, R., Wingfield, J., Bell, M. & Bell, J. 2007. Evidence for heat losses via party wall cavities in masonry construction. Building Service Engineering Research and Technology, 28(2), 161–181.

Milone A., Milone D. & Pitruzzella S. Asset rating: disagreement between the results obtained from software for energy certification. In Proceedings of the Eleventh International IBPSA Conference (BS'09), 569-575. Glasgow. Scotland.

Raslan, R., & Davies, M. 2006. The Implementation of a Simulation-Based Compliance Methodology for the UK Building Regulations 2006: An Industry Survey. In Proceedings of the European Modelling Symposium 2006, 149-153. University College London, London: Simulation Society & the Graduate School, UCL.

Raslan, R. & Davies, M. 2009. Results variability in accredited building energy performance compliance demonstration software in the UK: an inter-model comparative study. Journal of Building Performance Simulation. 3 (1) 63 – 85.

Shorrock, LD, & Dunster JE. 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 (12): 1027-1037.

Summerfield A.J., Pathan A., Hong S.H., & Oreszczyn T. 2008. Energy, Ventilation, and Occupant Behaviour, In D. Mumovic & M. Santamouris (Eds.) A Handbook of Sustainable Building Design and Engineering. London: Earthscan

Summerfield A.J., Lowe R.J., & Oreszczyn T. 2010. Two models for benchmarking UK domestic delivered energy. Building Research & Information, 38:12-24.

Utley, J., Shorrock, L. 2009. BRE: Domestic Energy Fact File 2008, Building Research Establishment, Watford.

Williamson, T. J. 2010. Predicting building performance: the ethics of computer simulation. Building Research & Information 38(4):401-410