Predicting Heating Energy Needs for Residential Building Clusters Using a Non-Linear Data-Driven Approach

Andreas Koch, Sebastien Cajot
European Institute for Energy Research, Karlsruhe, Germany
koch@eifer.org

Abstract

To assist cities in reaching their environmental and climate protection targets, a number of modelling approaches are currently developed to support urban energy planning. Especially deterministic models have gained importance in recent years. For simulation of the existing building stock, these models are difficult to correctly parametrize given a general lack of precise building information. For such application cases, a data-driven model is tested at the scale of a residential building cluster. The comparison with measured energy use data shows good results and confirms the underlying hypothesis of increasing performance of the model with increasing scale (i.e. number of buildings). The paper thus argues to develop a differentiated discussion on selecting specific modelling approaches for the various use-cases in urban energy planning.

Introduction

In the context of an increasing importance of cities in the global debate on climate change (UNEP 2011), opportunities for urban mitigation actions emerge at the scale of building clusters and neighborhoods as a distinct target scale for urban energy planning.

For a small number of buildings, advanced building simulations can deliver accurate and reliable results. The quality and reliability of the results is highly dependent on detailed knowledge of the buildings' properties. For large areas such models "require more input than the available data can support" (Coakley, Raftery et al. 2014). This is especially the case in existing urban areas. Mendes, Ioakimidis et al. (2011) provide a detailed assessment of a number of applied models for urban energy planning. In their analysis of quantitative energy assessment methods for existing buildings Wang, Yan et al. (2012) point out the difficulties of acquiring reliable input data for deterministic models. At the scale of urban neighborhoods, where detailed building descriptions are lacking, such approaches are in many cases over-parameterized for the planning tasks (Coakley, Raftery et al. 2014). Here, data-driven models can be used instead of detailed forward modelling approaches, as the former can easily be applied based on a limited amount of information.

In order to support local energy planning already in early stages of urban (re)development projects, energy models are required that are robust in their application and capable of dealing with limited information and a high degree of uncertainty. In later phases of urban development projects as well as following their implementation, continuous benchmarks are necessary to assess the performance of the implemented scheme. In this way continuous simulation for urban neighborhoods can help to scale up monitoring practices often used at the scale of single buildings to detect inefficient operations for example in district heating systems within short time delays.

To deliver robust simulation results that can serve as continuous benchmarks, the presented case study investigates the applicability of a data-driven modelling approach, based on a limited amount of input data at the scale of a building cluster. In comparison to measured energy use data, the limits in terms of scale are evaluated.

Data-driven modelling of heating energy needs

At the scale of buildings, data-driven models are widely used in the context of monitoring (Kissock, Haberl et al. 2003, Raffio, Isambert et al. 2007) or for the assessment of the performance of HVAC equipment. In these cases, measurements are used to build specific models for a given case. Approaches that consider the dependency of heating energy needs and outdoor temperature as well as further environmental variables are referred to as energy signatures (Bauer and Scartezzini 1998; Rabl and Rialhe 1992). Jacobsen (1985) was among the first to use energy signatures to assess building performance data. Day (2006) provides an overview of this category of models, a reference to design such models is described in ASHRAE's inverse modelling toolkit (Kissock, Haberl et al. 2003). Raffio, Isambert et al. (2007) as well as Mazzarella, Liziero et al. (2009) used energy signatures to assess the performance of larger building stocks.

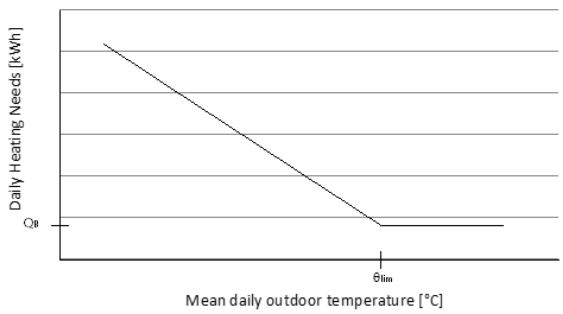

Figure 1: Example of a linear energy signature model

In its simplest form an energy signature is expressed as a linear single-variant regression model based on e.g. the correlation of heating energy use and mean outdoor temperature (Figure 1). Girardin, Marechal et al. (2010) used such linear regression models to calculate the thermal needs of urban areas. Ali, Mokhtar et al. (2011) used this class of models for predicting electricity demand at urban scale for cooling and heating applications.

Figure 1 shows a linear single change point model with the heating limit temperature θ_{lim} and a non-temperature dependent load Q_B. Further developments included models with multiple change points to represent phenomena such as limited heating capacity in winter (Mazzarella, Liziero et al. 2009). Dotzauer (2002) developed a linear model with four change points for simulating energy demand in district heating systems.

Simulation

Sigmoid Energy Signature Model

Based on a review of a number of applied methods, Koch (2010) proposed a sigmoid model for the simulation of heating needs at neighborhood scale. The author successfully applied the model at the scale of a district heating system (Woods 2012). The sigmoid function was initially developed by Geiger and Hellwig (2002) for the application in gas distribution networks. Such calculations are carried out at the regional scale to estimate day-ahead gas consumption for a given regulatory zone (Eichlseder 2008, BDEW 2010, BDEW, VKU et al. 2014). The non-linear data-driven model was adapted to current building standards and was improved regarding the performance at low outdoor temperatures (Koch 2016) by minimizing the coefficient of variation of the root mean square error when compared to measured data. Kissock, Haberl et al. (2003) proposed a least square regression analysis to calibrate data driven models to specific application cases.

The non-linear model (1) provides a distribution function for heating energy needs over the simulated period resulting in the normalized daily energy needs (h_a). The three parameters “A”, “B” and “C” are defined in relation to the type of building use and determine the slope as well as the change point of the function. Parameter “D” shifts the curve along the vertical axis and thus describes the non-temperature dependent part of the energy needs (e.g. domestic hot water).

$$h_a = \frac{A}{1 + \left(\frac{B}{\vartheta_a - \vartheta_0}\right)^C} + D \qquad (1)$$

The sigmoid model is based on a calculated daily outdoor temperature (ϑ_a) as its single regressor variable, which is represented by a geometric sequence with four elements (2). The calculated temperature considers the mean outdoor temperature of the actual day (ϑ_t) and a weighted value of the three previous days (ϑ_{t-n}). The temperature ϑ_0 describes the point of discontinuity with T equals 40°C. The geometric sequence can be seen as the representation of the buildings’ inertia. Therefore, the resulting curves represented in Figure 2 and 3 are not exactly linearly correlated to the measured outdoor temperature.

$$\vartheta_a = \frac{\vartheta_t + \frac{1}{2} \times \vartheta_{t-1} + \frac{1}{4} \times \vartheta_{t-2} + \frac{1}{8} \times \vartheta_{t-3}}{1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8}} \qquad (2)$$

Case Study

Data-driven models inherently perform better with increased sample size. Therefore, the model was applied to different scales between single buildings and a small building cluster to assess the limits of its applicability. The selected case study is based on monitoring data from eight individual buildings in the South West of Germany. The buildings are situated within a 20 km radius, so that the same weather conditions were assumed for the simualtion. The buildings, which were all recently renovated, were of similar size and each equipped with a new central heating system for provision of space heating and domestic hot water. Data was anonymised so that no specific information could be correlated to individual house owners. The buildings are referred to as “Building A” to “Building H”. Measured energy use as well as temperature data was available in hourly time steps. It was recorded between the 1st of August 2008 and the 31st of July 2010. Eventually, five buildings were selected which provided a continuous year of measurements for the same period.

Based on the described model, daily heating energy needs were calculated. The demand calculation is based on the measured total annual heating needs and the calculated daily outdoor temperature. Individual simulation was carried out for each building. In addition, the buildings were added in random order to a larger sample, adding one building in each consequent step. Simulation results were then evaluated against measured data at different aggregation scales. For the simulation a parameter set for existing buildings (“Single Building (old)”) was selected which is provided in Table 1 (Koch 2016).

Table 1: Parameter (A, B, C, D) for simulation of single dwelling units (Koch 2016)

	A	B	C	D
Single Building (old)	3,13	-37,19	5,75	0.00
Single Building (new)	2.84	-36.89	6.57	0.00

Selected Statistic Indicators

Important indicators to assess the fitness of a given simulation model are the mean biased error (MBE) that delivers a non-dimensional description of the bias of the model as well as the coefficient of variation of the root mean square error (CV RMSE) (Coakley, Raftery et al. 2014). In the case of the energy signature model the former indicator does not provide useful results when comparing measured data against simulation results as the sigmoid function is used to distribute the total heating energy needs over the days within the simulation period.

Therefore, the MBE of both series is equal. A number of sources provide minimal performance criteria for the MBE and the CV RMSE for monthly and hourly simulations. Table 2 includes acceptance criteria proposed by the American Society of Heating, Refrigerating and Air-Conditioning Engineers - ASHRAE (ASHRAE 2002), the International Performance Measurement and Verification Protocol - IPMVP (EVO 2007), and the American Federal Energy Management Program - FEMP (US DOE and Nexant 2008).

Table 2: Acceptance criteria for building energy performance simulation

Standard/guideline	Monthly criteria		Hourly criteria	
	MBE	CV RMSE	MBE	CV RMSE
ASHRAE Guideline 14	5%	15%	10%	30%
IPMVP	20%	–	5%	20%
FEMP	5%	15%	10%	30%

As further indicators, the coefficient of determination and the Bravais-Pearson correlation coefficient are used to evaluate the results (Kühlmeyer 2001). Grohmann (2000) proposed using a combination of indicators. In the same line of thought, the correlation coefficient is regarded alongside the variance to include the assessment of the spread of the compared series. The variance is described by the ratio of the standard deviation of simulated and measured values (σ_s/σ_m). Finally, the regression analysis with the coefficient of determination is used as an important indicator combined with the line of equality (Bland and Altman 1986), the former alone would only provide for a valid evaluation of the correlation, while for building simulation also the match of the absolute values is an important aspect.

Analysis of results

At the scale of a building cluster, the simulation shows good results even though, no calibration of the model's parameters was performed.

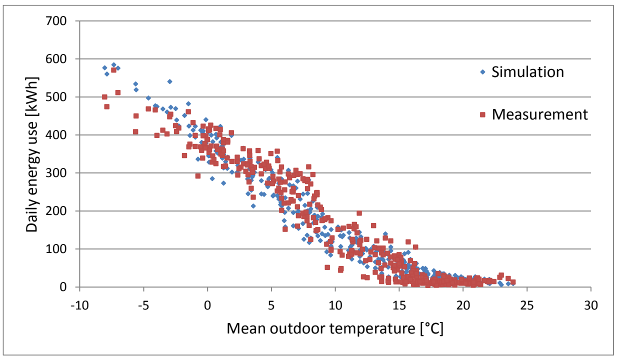

Figure 2: Simulated and measured daily heating energy needs vs. outdoor temperature, aggregated for five single dwelling units, (Koch 2016)

Figure 2 shows a good match of the slope of the heating curve aggregated for five buildings and the simulation of the daily heating energy needs. The heating limit temperature, which corresponds to the change point in linear signature models, is also well represented. In contrast, the simulation for individual buildings showed the expected high values for the coefficient of variation of the root mean square error (CV RSME) of up to 44% (Table 3). By aggregating all buildings the CV RSME was reduced to a good value of 19% for the whole sample, which falls in the range of acceptability proposed by ASHRAE (Table 1). In Figure 2 this relates to a comparable width of both scatter plots which is not the case for the individual building (Figure 3). The coefficient of determination (R^2) improved from 0.83 to a value of 0.95 for the complete building cluster (Table 3).

Table 3: Statistic analysis of the simulation results at different aggregation scales

Daily Series	**CV RMSE**	R^2	**ρ**	σ_s/σ_m
Building A	35 %	0.87	0.93	1.10
Building B	25 %	0.92	0.96	1.02
Building C	24 %	0.96	0.98	0.85
Building D	30 %	0.88	0.94	1.00
Building G	44 %	0.83	0.91	0.78
All Buildings	19 %	0.95	0.97	0.98

The Bravais-Pearson correlation coefficient (ρ) showed an increase to 0.97 from an already high value of 0.91. The ratio of the standard deviations of simulated and measured data series (σ_s/σ_m) improved to a value of 0.98 indicating a very good match of the spread of both data series. The coefficient of determination (R^2) was above 0.83 for all individual cases, with a value of 0.95 for the building cluster. Finally, the correlation coefficient was above 0.91 for all simulation cases and for most cases improved with the larger sample size. Among the monitored cases studies, “Building G” was least well represented when regarded individually. The results are shown in Figure 3 to contrast the well-represented scatter plot for all buildings shown in Figure 2.

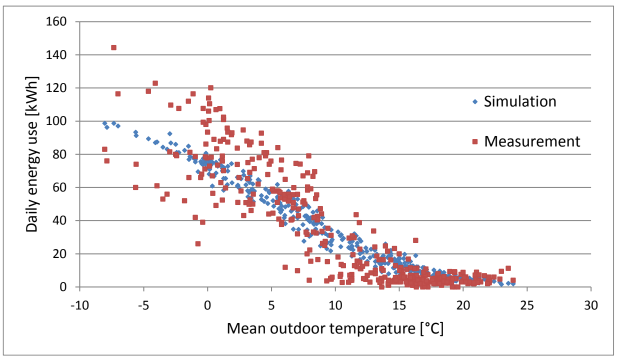

Figure 3: Simulated and measured daily heating energy needs vs. outdoor temperature for a single dwelling unit (Building G), (Koch 2016)

The CV RMSE showed a high error of 44%, the ration of the standard deviations resulted in 0.78, which relates to a less strict correlation of heating energy needs and outdoor temperature. Still measured and simulated values are highly correlated with a value of 0.91. The model represents well the trend but fails to depict correctly the spread at the scale of individual buildings. This latter result further stresses the need to not only rely on one indicator such as the coefficient of determination in the assessment of the performance of energy simulation models (Bland and Altman 1986). Measured and simulated heating energy needs for all buildings are furthermore compared in Figure 4, including the trend line as well as the line of equality necessary to measure not only the high correlation but also the good representation of absolute values of the heating energy needs.

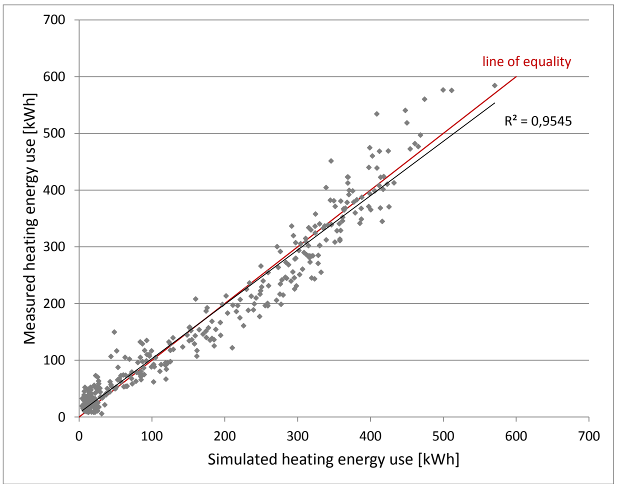

Figure 4: Scatter plot of the measured and simulated daily energy needs for all buildings with trend line and line of equality, (Koch 2016)

Conclusion

Comparing the simulation results to measured data of single-dwelling residential units enabled the investigation of the importance of scale in the application of the model. At the scale of regions the energy signature models is used to predict aggregated demand (BDEW, VKU and GEODE 2014). The presented work investigated the lower limits of scale. The statistical analysis shows an improvement of all considered indicators with increasing sample size. Already at the relatively small scale of a building cluster, the results indicate the model's capability to explain and correctly predict the trend of heating energy needs.

At the scale of building clusters, the approach proved to deliver good results within the range of acceptability suggested for building simulation by ASHRAE Guideline 14-2002 (ASHRAE 2002). In nearly all considered cases, individual buildings are less well represented by the data-driven model than the full sample. The work supports the research hypothesis of a sensitivity to scale of the energy signature model and indicates an applicability at the larger scale of neighborhoods or cities. Yet, further tests are required, to better identify the limits of scale for the application of the energy signature model. In previous case studies, the model was successfully used for the demand simulation of districted heating systems (Woods 2012).

The shown application used default model parameters. In future works the calibration of the energy signature model for specific sites is likely to further improve the results. Such a calibration is feasible by minimizing the CV RMSE in comparison to measured data. In addition to providing a fast yet robust model for predicting heating energy needs, the approach could thus be applied in the context of continuous commissioning for the existing building stock. Based on the work, parameter sets will be developed for specific building types for the application in comparable conditions. Important factors ensuring a minimal transferability are central heating systems with a heating based operation mode and comparable climatic conditions.

In addition to the direct assessment of the performance of the model, the results show, that in neighborhood or city scale simulation, aggregating effects are important factors to consider in the choice of the modelling approach. Application cases for which data-driven simulation can provide a good solution include demand estimation for district heating systems in order to provide annual load duration curves. In addition, the simulation of daily and hourly heating needs for advanced local energy planning in the existing building stock at the scale of communities or cities are further possible use cases.

Acknowledgement

The authors would like to express their thanks to Niklas Grießbaum for his kind support in the treatment of the monitoring data as well as to Andreas Wagner of the Building Science Group at the KIT for his advice during the supervision of the author's thesis.

References

Ali, M. T., M. Mokhtar, M. Chiesa and P. Armstrong (2011). A cooling change-point model of community-aggregate electrical load. Energy and Buildings 43(1): 28-37.

ASHRAE (2002). Guideline 14-2002, Measurement of Energy and Demand Savings. Atlanta, Georgia, American Society of Heating, Ventilating, and Air Conditioning Engineers.

ASHRAE (2005). Energy Estimating and Modelling Methods. 2005 ASHRAE Handbook - Fundamentals. Atlanta, GA, USA, American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Bauer, M. and J. L. Scartezzini (1998). A simplified correlation method accounting for heating and cooling loads in energy-efficient buildings. Energy and Buildings 27(2): 147-154.

BDEW, VKU and GEODE (2014). BDEW/VKU/GEODE Leitfaden - Abwicklung von Standardlastprofilen Gas. Berlin, Bundesverband der

Energie- und Wasserwirtschaft e.V.; Verband kommunaler Unternehmen e.V.; Groupement Européen des entreprises et Organismes de Distribution d'Énergie.

BDEW. (2010). Endenergieverbrauch der privaten Haushalte 2008 nach Anwendungsarten. Arbeitsgemeinschaft Energiebilanzen e.V., Berlin.

Bland, J. M. and D. G. Altman (1986). Statistical methods for assessing agreement between two methods of clinical measurement. The lancet 327(8476): 307-310.

Coakley, D., P. Raftery and M. Keane (2014). A review of methods to match building energy simulation models to measured data. Renewable and Sustainable Energy Reviews 37: 123-141.

Day, T. (2006). Degree-days: theory and application. London, The Chartered Institution of Building Services Engineers (CIBSE). TM41.

Dotzauer, E. (2002). Simple model for prediction of loads in district-heating systems. Applied Energy 73(3-4): 277-284.

Eichlseder, H. (2008). Lastprofile nicht-leistungsgemessener Kunden (HE, HM, HG, PG, PK und PW) der Gasnetzbetreiber Österreichs – Überarbeitung 2008. Graz, Technische Universität Graz.

EVO (2007). International performance measurement & verification protocol. Concepts and Practices for Determining Energy Savings in New Construction Volume III Part I, Efficiency Valuation Organisation.

Geiger, B. and M. Hellwig (2002). Entwicklung von Lastprofilen für die Gaswirtschaft im Auftrag des Bundesverbandes der deutschen Gas- und Wasserwirtschaft e.V. und des Verbandes kommunaler Unternehmen e.V. - Haushalte. München, Lehrstuhl für Energiewirtschaft und Anwendungstechnik im Institut für Energietechnik, TU München.

Girardin, L., F. Marechal, M. Dubuis, N. Calame-Darbellay and D. Favrat (2010). EnerGis: A geographical information based system for the evaluation of integrated energy conversion systems in urban areas. Energy 35(2): 830-840.

Grohmann, J. (2000). Lastanalyse und -synthese zur Optimierung leitungsgebundener Energiesysteme. Herrsching, E & M Energie und Management VerlagsGmbH.

Hellwig, M. (2003). Entwicklung und Anwendung parametrisierter Standard-Lastprofile. PhD, Technische Universität München.

Jacobsen, F. R. (1985). Energy signature and energy monitoring in building energy management systems. Proceeding of CLIMA 2000 World Congress, Copenhagen.

Kissock, J. K., J. S. Haberl and D. E. Claridge (2003). Inverse Modeling Toolkit: Numerical Algorithms. ASHRAE Transactions-Research 109 (Part 2): 425-434.

Koch, A. (2010). Heat Demand Modeling at the Scale of Urban Neighborhoods, Next Generation Infrastructure Systems for Eco-cities Conference, 11-13 November 2010, Shenzhen, China.

Koch, A. (2016). Continuous Simulation for Urban Energy Planning Based on a Non-Linear Data-Driven Modelling Approach, thesis, KIT.

Kühlmeyer, M. (2001). Statistische Auswertungsmethoden für Ingenieure. Berlin, Springer.

Masuda, H. and D. E. Claridge (2014). Statistical modeling of the building energy balance variable for screening of metered energy use in large commercial buildings. Energy and Buildings 77: 292-303.

Mazzarella, L., M. Liziero, C. Neumann and D. Jacob (2009). Description of European Prototype Tool for Evaluation of Building Performance and the national tools. Milano, Politecnico di Milano, Fraunhofer Institut für Solare Energieysteme.

Mendes, G., C. Ioakimidis and P. Ferrão (2011). On the planning and analysis of Integrated Community Energy Systems: A review and survey of available tools. Renewable and Sustainable Energy Reviews 15(9): 4836-4854.

Rabl, A. and A. Rialhe (1992). Energy signature models for commercial buildings: test with measured data and interpretation. Energy and Buildings 19(2): 143-154.

Raffio, G., O. Isambert, G. Mertz, C. Schreier and K. Kissock (2007). Targeting residential energy assistance. Energy Sustainability, Long Beach.

UNEP (2011). Towards a Green Economy: Pathways to Sustainable Development and Poverty Eradication. Geneva, United Nations Environment Programme.

US DOE and Nexant (2008). M&V guidelines: measurement and verification for federal energy projects version 3.0. Boulder, Colorado, US Department of Energy.

Wang, S., C. Yan and F. Xiao (2012). Quantitative energy performance assessment methods for existing buildings. Energy and Buildings 55, 8

Woods, P. (2012). ETI Macro Distributed Energy Project - Summary Report. Loughborough, Energy Technologies Institute.