Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 122 (2017) 277–282

Energy Procedia

www.elsevier.com/locate/procedia

CISBAT 2017 International Conference – Future Buildings & Districts – Energy Efficiency from Nano to Urban Scale, CISBAT 2017 6-8 September 2017, Lausanne, Switzerland

Bayesian calibration of building energy models: Comparison of predictive accuracy using metered utility data of different temporal resolution

Martin Heine Kristensen[a,*], Ruchi Choudhary[b], Steffen Petersen[a]

[a]*Department of Engineering, Aarhus University, Inge Lehmanns Gade 10, 8000 Aarhus, Denmark*
[b]*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom*

Abstract

Modern smart meters in heating systems offer building energy data of high temporal resolution. Compared to the annually aggregated readings used for conventional billing, the continuous information flow from these smart meters can be made available as time series data containing monthly, daily or even hourly aggregated values. In this paper, the effect of different temporal aggregation levels of commercial smart meter data on building energy model (BEM) calibration is investigated. Four different aggregation levels of a training data set were applied for calibration of six BEM input parameters to set up a Gaussian process emulator of the physical system. The performance of the emulator was subsequently tested on an unseen validation data set. Results reveal a systematic pattern of increasing predictive accuracy as a function of increasing training data resolution.

Peer-review under responsibility of the scientific committee of the CISBAT 2017 International Conference – Future Buildings & Districts – Energy Efficiency from Nano to Urban Scale

Keywords: Bayesian calibration; building energy modeling; temporal data resolution; smart meter; detached single-family house

* Corresponding author.
E-mail address: mhk@eng.au.dk

1876-6102 © 2017 The Authors. Published by Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the CISBAT 2017 International Conference – Future Buildings & Districts – Energy Efficiency from Nano to Urban Scale
10.1016/j.egypro.2017.07.322

1. Introduction

Setting up a valid building energy model (BEM) is often a difficult task, e.g. in the case of modeling an already existing building stock where one, to some extent, have to resort to conjecturing about construction details, type of materials and their state of condition. The uncertainty embedded in such BEMs may be dealt with by means of calibration [1, 2] where parameters are fitted to metered energy use data; however, depending on the level of uncertainty in the model, it can be difficult to find a deterministic *best fit* of model parameters. Probabilistic calibration is suitable for incorporating and quantifying this error, as one does not have to rely on single-value estimates, but can allow noisy data and the uncertainty of unidentifiable parameters to be propagated through the model.

In recent years, a plethora of Bayesian-based calibration techniques have been proposed and demonstrated, e.g. [3, 4, 5, 6, 7, 8, 9]. These references all offspring from the original emulator-based framework proposed by Kennedy and O'Hagan [10] and Higdon et al. [11] utilizing Gaussian process (GP) regression to match BEM evaluations with observed data by fitting calibration parameters. The emulator-based Bayesian approach enables probabilistic calibration of uncertain inputs for a given BEM using a limited number of evaluations from the building model. Implementing prior information about uncertain input parameters further enables the modeler to bias or even constrain the posterior inference – an option that can be reasonable to use when a limited amount of observed data is available for calibration. The calibration efficacy of the Bayesian framework has previously been studied under different levels of uncertainty [6] and different training set sizes [8]. However, it remains unclear how the temporal data resolution of the observed training set affects the posterior parameter inference and overall model accuracy.

In this paper, we investigate the issue of training data resolution by presenting a study on how different temporal resolution of metered district heating (DH) data affects the predictive accuracy of a BEM of a detached single-family house. We compare the posterior estimates of the calibration parameters and the predictive capabilities of the posterior model as measures of this effect. As such, this paper advances our understanding of how the temporal resolution of currently available DH smart meter read data affects calibration quality. This knowledge is valuable in many situations, for instance when modeling existing buildings for retrofit decision making under uncertainty, and for the future design of building energy management systems.

2. Methods

In the following subsection, we shortly outline the Kennedy and O'Hagan calibration formulation [10] in the context of BEM and point out changes made for the purpose of this study.

2.1. Emulator-based Bayesian calibration framework

The building-physical system used to generate $i = 1,2,\dots,n$ observations of building energy use y_i at observed settings $\boldsymbol{x}_i$ and unknown observation error $\varepsilon_{obs,i}$ is represented as

$$y_i = \zeta(\boldsymbol{x}_i) + \varepsilon_{obs,i} \qquad i = 1, \dots, n, \tag{1}$$

where $\zeta(\boldsymbol{x}_i)$ denotes the true energy-consuming process. The observable setting $\boldsymbol{x}_i$ consists of a p-dimensional vector of explanatory design points $\boldsymbol{x} \in \mathbb{R}^p$. In this study, we took $p = 2$ by letting x_1 index the outdoor air temperature and x_2 index the insolation. The inclusion of a Gaussian distributed noise-term $\varepsilon_{obs,i} \sim N(0, \sigma^2_{obs})$ allowed for different observations of y at identical settings of $\boldsymbol{x}$, hereby accounting for the very stochastic nature of the energy-consuming process, e.g. occupant behaviour.

Using a BEM to represent the energy-consuming process, the observations were statistically modeled as

$$y_i = \eta(\boldsymbol{x}_i, \boldsymbol{\theta}) + \delta(\boldsymbol{x}_i) + \varepsilon_{obs,i} \qquad i = 1, \dots, n, \tag{2}$$

where $\eta(\boldsymbol{x}_i, \boldsymbol{\theta})$ denotes evaluations of the BEM at the p+q-dimensional input vector $(\boldsymbol{x}_i, \boldsymbol{\theta})$ comprising observed design points $\boldsymbol{x}_i$ and additional calibrated parameters $\boldsymbol{\theta} \in \mathbb{R}^q$. In this study, we selected $q = 6$ BEM input parameters for calibration based on a Sobol sensitivity analysis on the model as demonstrated by Kristensen and Petersen [12]. They were *U-value (windows)*, *Infiltration@50Pa*, *Thermal mass*, *Heating temperature set point*, *Occupant density*