Accepted Manuscript

Title: A new feature selection method to improve the document clustering using particle swarm optimization algorithm

Author: Laith Mohammad Abualigah Ahamad Tajudin Khader Essam Said Hanandeh

PII:	S1877-7503(16)30500-2
DOI:	http://dx.doi.org/doi:10.1016/j.jocs.2017.07.018
Reference:	JOCS 733

To appear in:

Received date:	18-12-2016
Revised date:	23-6-2017
Accepted date:	30-7-2017

Please cite this article as: Laith Mohammad Abualigah, Ahamad Tajudin Khader, Essam Said Hanandeh, A new feature selection method to improve the document clustering using particle swarm optimization algorithm, *<![CDATA[Journal of Computational Science]]>* (2017), http://dx.doi.org/10.1016/j.jocs.2017.07.018

Bayesian calibration of building energy models with large datasets

Adrian. Chong[a,*], Khee Poh. Lam[a,b], Matteo. Pozzi[c], Junjing. Yang[b]

[a] *Department of Building, National University of Singapore, Singapore*
[b] *Center for Building Performance and Diagnostics, Carnegie Mellon University, USA*
[c] *Civil and Environmental Engineering, Carnegie Mellon University, USA*

Abstract

Bayesian calibration as proposed by Kennedy and O'Hagan (2001) has been increasingly applied to building energy models due to its ability to account for the discrepancy between observed values and model predictions. However, its application has been limited to calibration using monthly aggregated data because it is computationally inefficient when the dataset is large. This study focuses on improvements to the current implementation of Bayesian calibration to building energy simulation. This is achieved by: 1) using information theory to select a representative subset of the entire dataset for the calibration, and 2) using a more effective Markov chain Monte Carlo (MCMC) algorithm, the No-U-Turn Sampler (NUTS), which is an extension of Hamiltonian Monte Carlo (HMC) to explore the posterior distribution. The calibrated model was assessed by evaluating both accuracy and convergence.

Application of the proposed method is demonstrated using two cases studies: 1) a TRNSYS model of a water-cooled chiller in a mixed-use building in Singapore, and 2) an EnergyPlus model of the cooling system of an office building in Pennsylvania, U.S.A. In both case studies, convergence was achieved for all parameters of the posterior distribution, with Gelman-Rubin statistics $\hat{R}$ within 1 ± 0.1. The coefficient of variation of the root mean squared error (CVRMSE)

☆Fully documented templates are available in the elsarticle package on CTAN.

*Corresponding author
Email address: bdgczma@nus.edu.sg (Adrian. Chong)

Preprint submitted to Energy and Buildings *August 29, 2017*