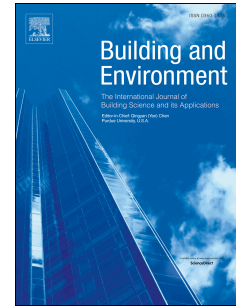


Accepted Manuscript

Reconstruction of the indoor temperature dataset of a house using data driven models for performance evaluation

Luis M. Candanedo, Véronique Feldheim, Dominique Deramaix



PII: S0360-1323(18)30247-6

DOI: [10.1016/j.buildenv.2018.04.035](https://doi.org/10.1016/j.buildenv.2018.04.035)

Reference: BAE 5431

To appear in: *Building and Environment*

Received Date: 11 January 2018

Revised Date: 16 April 2018

Accepted Date: 26 April 2018

Please cite this article as: Candanedo LM, Feldheim Vé, Deramaix D, Reconstruction of the indoor temperature dataset of a house using data driven models for performance evaluation, *Building and Environment* (2018), doi: 10.1016/j.buildenv.2018.04.035.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 **Reconstruction of the indoor temperature dataset of a house using**
2 **data driven models for performance evaluation**

3 Luis M. Candanedo, Véronique Feldheim, Dominique Deramaix

4 Luismiguel.candanedoibarra@umons.ac.be

5 Tel +32(0)65 37 4471 Fax +32(0)65 37 4400

6 Thermal Engineering and Combustion Laboratory, Rue de l' Epargne 56, 7000 Mons.

7 University of Mons, Belgium

8 **Abstract:**

9 Whenever the long term monitoring of a building is attempted it is likely that specific sensors or the
10 whole monitoring system used may experience long-term failure therefore creating important gaps
11 in one or more variables of special interest. These long gaps may not be addressed using simple
12 linear interpolation. The option of only using the available data for descriptive statistics would
13 produce results that are biased towards the season of measurement. In addition discarding the
14 incomplete data represents a significant waste of time and effort in the research study. A work
15 around to reduce the bias problem is to predict the missing data from other measured variables
16 using machine-learning techniques. Some questions that follow are: How much data is necessary to
17 be able to train a regression model? What is the expected error of such prediction? What is the best
18 model for such a task? This paper addresses the problem of completing a data set for the interior
19 temperatures inside a passive house using different monitored predictors such as exterior
20 temperature, humidity, wind speed, visibility, pressure and electrical energy use inside the building.
21 Two regression models, multiple linear regression and random forest are compared using learning
22 curves for the training and testing sets for visualizing the so-called bias-variance trade off. The
23 learning curves help to answer the question of optimal sample size for training, model selection and
24 expected error. Finally, descriptive statistics such as median, maximum, minimum, and room
25 temperature averages are presented before and after completing the data sets.

26 *Keywords: Learning curves, multiple linear regression, random forest, passive house, temperatures,*
27 *sample size*

Download English Version:

<https://daneshyari.com/en/article/6697237>

Download Persian Version:

<https://daneshyari.com/article/6697237>

[Daneshyari.com](https://daneshyari.com)