Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 122 (2017) 337–342

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

Comparison of Different Classification Algorithms for the Detection of User's Interaction with Windows in Office Buildings

Romana Markovic[a,*], Sebastian Wolf[b], Jun Cao[a], Eric Spinnräker[a], Daniel Wölki[a], Jérôme Frisch[a], Christoph van Treeck[a]

[a]*RWTH Aachen University, Mathieustr. 30, 52074 Aachen, Germany*
[b]*Technical University of Denmark, Asmussens All, Building 303B, 2800 Lyngby, Denmark*

Abstract

Occupant behavior in terms of interactions with windows and heating systems is seen as one of the main sources of discrepancy between predicted and measured heating, ventilation and air conditioning (HVAC) building energy consumption. Thus, this work analyzes the performance of several classification algorithms for detecting occupant's interactions with windows, while taking the imbalanced properties of the available data set into account. The tested methods include support vector machines (SVM), random forests, and their combination with dynamic Bayesian networks (DBN). The results will show that random forests outperform all alternative approaches for identifying the window status in office buildings.

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: occupant behavior; window opening; office buildings, SVMs; Random forest;

1. Introduction

Occupant behavior is seen as the main source of discrepancy between predicted and measured energy consumption in buildings. Hence, understanding occupant behavior is crucial for achieving high performance and low-energy use, both in the commercial and the residential field [15].

Users' interactions with windows, in terms of window opening and closing are needed for modelling air exchange through ventilation. Nonetheless, they should be taken into account for controlling strategies of buildings' mechanical ventilation. Modelling window opening behavior is an important part of building performance simulation, in order to make reliable predictions of the buildings' energy consumption [24]. However, the conventional building simulation approaches still rely on synthetic window opening predictions, that do not lead to a realistic occupant's influence on the energy performance.

Occupant behavior and perceived thermal comfort in office buildings have been investigated in numerous studies [13], [8], [7], [11], [23], [9]. In addition, there are multiple studies that investigated window opening behavior in offices

* Corresponding author. Tel.: +49-241-80-25541 ; fax: +49-241-80-22030.
E-mail address: markovic@e3d.rwth-aachen.de

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.2017.07.333

[14], [10] and residential buildings [2], [20], [22]. Haldi and Robinson [14] showed that a Markov model provides higher accuracy compared to the logistic regression and agent based method. However, even though the model provided over 80 % of correct predictions in case of closed windows, the ability to predict an open window remained low. D'Oca and Hong [9] applied a data-mining approach to discover the patterns of window opening and closing in office buildings. They identified several behavioral patterns, including motivational and opening duration patterns. Similarly to Haldi and Robinson, they showed that indoor and outdoor air-temperature together with the time of a day and presence durations were the strongest factors leading to window opening and closing actions.
Machine learning (ML) and artificial intelligence (AI) techniques are widely used for predicting and evaluating occupant's actions in buildings [5], [16] as well as buildings' energy consumption [12], [1], [3]. However, there is little work that uses smart algorithms for modelling the occupants' interactions with windows in case of office buildings.
Furthermore, human behavior, including window opening and closing actions, cannot be modelled using analytical physical approaches. As a result, occupant actions have to be modelled using data-driven methods. For this purpose, machine learning methods offer a comprehensive alternative to modelling the occupant behavior in buildings and its' influence on the energy consumption.
This paper models occupant's actions and the resulting window status in office buildings by applying support vector machines (SVMs) and random forest. Based on monitoring data, the window status is defined as classification problem, where the status, open or closed, is identified. In addition, the temporal dependence of window actions is investigated by implementing a dynamic Bayesian network (DBN), with the aim of smoothing the classification results.

2. Method

2.1. Data Set

Data set includes monitoring data collected over two years in an office building in Frankfurt, Germany [17], [21]. The available data are collected on ten monitored offices in ten minutes time-steps. Due to the very low occupancy rate, one of the ten offices is excluded from the further evaluation. Measured data include indoor climate features (indoor air temperature) and outdoor climate features (outdoor air temperature, precipitation, wind velocity, wind direction, CO_2 concentration and relative humidity) as well as occupant's presence and actions (position sun protection, occupancy, time presence, occupancy state).
The window status is defined as a binary problem, where 0 and 1 refers to a closed and an open window, respectively. In addition, it is not distinguished between both windows in each office. As a result, all data points where at least one of the two windows is opened are labelled as class 1. In case of SVMs, features are scaled in range between 0 and 1 prior to data splitting into training and evaluation set. Since random forests does not require feature scaling, the random forest data remained in the original monitored range.

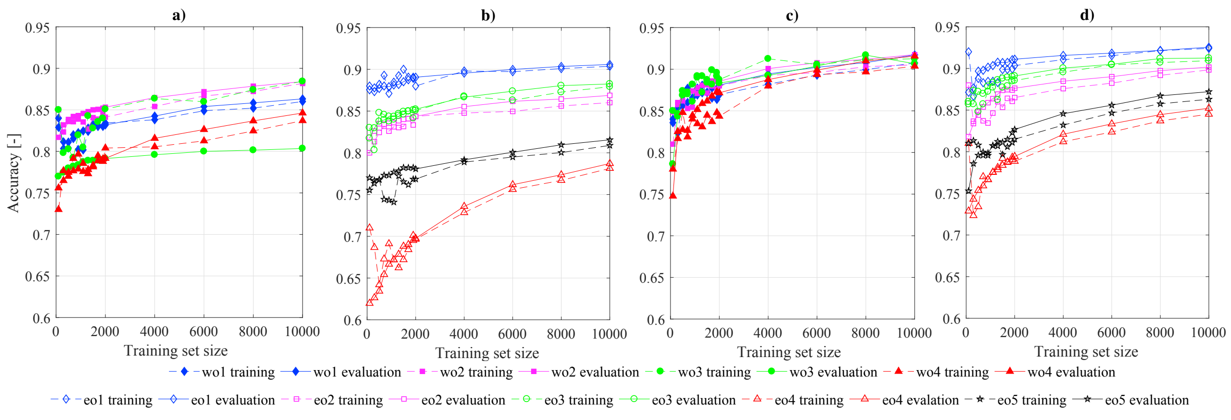

Fig. 1. Training and evaluation accuracy for varied training set size for SVMs ((a) and (b)) and random forest ((c) and (d)).