Advanced Engineering Informatics xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Application of clustering for the development of retrofit strategies for large building stocks

Philipp Geyer [a,*], Arno Schlüter [b], Sasha Cisar [b]

[a] KU Leuven, Belgium
[b] ETH Zürich, Switzerland

A R T I C L E I N F O

Article history:
Received 23 December 2014
Received in revised form 26 November 2015
Accepted 3 February 2016
Available online xxxx

Keywords:
Algorithmic clustering
Building retrofitting
Strategic building stock management
Data mining

A B S T R A C T

In order to reduce energy consumption and emissions from the built environment, it is vital to transform the existing building stock and develop retrofit strategies to achieve energy efficiency and building-integrated renewable energy supply. Compared to developing cost-optimal retrofit strategies for one building, the development of strategies for 100 to up to 10,000 buildings remains a major challenge. This paper presents a method to cluster buildings based on their sensitivity to different retrofit measures, focusing on the cost-effectiveness. Derived from algorithmic clustering and combined with time and cost data, a tailored development of retrofit strategies for large building stocks becomes possible. Improved identification of retrofit measures and strategies, in contrast to the conventional classification based on building type and age, is demonstrated. The method is illustrated, using the data from the case study project 'Zernez Energia 2020', which aims to achieve carbon neutrality of a Swiss alpine village.

1. Introduction

Achieving the objective of energy efficiency and emission reduction in urban structures requires the management of retrofitting the building stock. A central need is the identification of the energy reduction potential for heating, cooling and related emissions and to locally produce renewable energy within the constraints of a limited financial budget. To achieve an optimal retrofit it would be beneficial to assess each single building in an urban structure for the applicability and effect of measures for efficiency improvement and decentralized energy production and then to draw an overall conclusion about the potential. Furthermore, as building owners need to finance the largest portion of the measures themselves, it is essential to identify strategies about how to best retrofit and improve the energetic behavior of the buildings in relation to the available means of investment.

The currently most applied approach to deal with building stocks of 100 up to 10,000 buildings is a type-age classification. However, for the determination of measures for improving energy efficiency and for reducing CO_2 emissions, many other factors than just age and building type have an impact on the effectiveness of retrofit measures. Therefore, the type-age classification provides only a rough estimation to develop retrofit strategies. Using modern databases and GIS, comprehensive information of the building stock is increasingly available and this paves the way to develop more sophisticated approaches that utilize this information allowing retrofit strategies to be implemented in a timelier, efficient and cost effective manner.

1.1. Objectives

The objective of this work is to utilize available rich building data for simulation, analysis and identification of cost-optimal retrofits measures for groups of individual buildings. The resulting matching of measures to building groups can then be utilized to define overarching strategies in order to allow faster and more cost-effective retrofits. The development of strategies for clusters instead of individual buildings facilitates more effective strategy development, compared to dealing with each building individually, and makes it thus more feasible to address the complete building stock. Such information can lead to a change in policies for retrofit subsidies and a more effective utilization of public funds. Furthermore, a coordinated retrofit opens up possibilities of the economy of scale.

We present the method of performance-based clustering in Section 2. This includes a comparison of different clustering techniques. The case study of Zernez, which is introduced in Section 3, is used to demonstrate that the proposed method improves the impact of retrofit measures while minimizing the

* Corresponding author.
E-mail address: p.geyer@kuleuven.be (P. Geyer).

http://dx.doi.org/10.1016/j.aei.2016.02.001
1474-0346/

Please cite this article in press as: P. Geyer et al., Application of clustering for the development of retrofit strategies for large building stocks, Adv. Eng. Informat. (2016), http://dx.doi.org/10.1016/j.aei.2016.02.001

resource consumption (material and costs). The method is applied on the data collected during the case study and related available information. Section 4 reports on the result of the application of the method on the case study. This section also shows how to derive similar reacting groups of buildings, how to interpret them and how to develop retrofit strategies.

1.2. Background

For predicting the future energy use and emissions and for evaluating scenarios and strategies, building stock models are important. Swan and Ugursal [18] and Kavgic et al. [9] review the state of the art techniques of building stock modelling in research, existing models and approaches and their use in policy making. The typical parameters that these models build on are categories of building age, type of building, heat distribution type, energy source, construction or refurbishment year and dwelling type. In particular, bottom-up models that are based on building physics engineering serve to assess the reduction potential of energy efficiency measures and technologies for the building stock. For this purpose, existing approaches usually define scenarios and strategies using standard combinations of measures and estimation for the whole building stock. However, these approaches do not derive measures or strategies from the analysis of the actual building stock and do not derive individual retrofit strategies.

Furthermore, classification methods are applied to building stock data. Usually, the parameters in the data serve to derive groups and to assign measures for energy efficiency. A typical example is the type-age classification that has been developed for Germany [10] and recently extended to Europe [6]. Many examples of classifying are based on type, age and other similar parameters in building stock. For instance, Kohler and Yang [11] address long term management of building stocks by using the German model of Building Stocks. Uihlein and Eder [19] use four age categories in the development of a broad strategy for EU-27 residential building stock. Boardman [1] examines low- and zero-carbon technologies for residential building stock in UK. He uses Oxford's UK Domestic Carbon Model that is based on age class, dwelling type, tenure type, number of floors and construction type.

Moreover, approaches that use algorithmic clustering for building stocks based on energy consumption and other parameters exist. Santamouris et al. [15] apply clustering to a database of 320 schools in Greece and build groups based on the energy consumption with climatic normalization. Gaitani et al. [3] identify typical building properties and parameters of the schools by k-means clustering. Jones et al. [8] cluster building stock by building properties, such as heated ground floor area, façade, window to wall ratio. Yamaguchi et al. [21] identify district types and provide typical energy performance by simulating buildings in a representative district.

Furthermore, in engineering fields other than building stock management, de Oliveira et al. [14] use density-based spatial clustering of failures in a water network. Jazizadeh et al. [7] apply heuristic unsupervised hierarchical clustering for autonomous partitioning of appliances signature space in non-intrusive load monitoring (NILM). Motamedi et al. [13] use spatial clustering combined with other criteria during localization of RFID-equipped assets. Hung and Kang [4] develop a method for grouping objects for collision detection in real-time construction simulation using k-means clustering. Hyun et al. [5] use hierarchical clustering for similarity analysis of car designs in order to identify car manufacturers' design strategies.

All mentioned approaches related specifically to building stocks use either the descriptive parameters (age, type, etc.) or the current performance to set up groups of buildings. Only a very few, such as Gaitani et al. [3] explicitly use algorithmic clustering. However none of the existing methods apply algorithmic clustering based on effectiveness of measures to develop retrofitting strategies for a building stock. This is of major interest for the development of strategies, as it is much easier to make decisions and develop strategies for groups of buildings that react similarly to energy efficiency measures.

2. The method of performance-based clustering

In this section, we describe the methodology for the fundamental shift from clustering based on describing parameters to clustering based on performance-based indicators of measure effects. This change allows identification of groups that react similarly to energy efficiency measures and therefore form the basis for group strategies. The main innovation of the method consists of selecting appropriate performance indicators for clustering that individually describe the reaction of retrofit measures instead of parameters such as type and age or general energy consumption. This is the basis for deriving groups that similarly react on a planning strategy developed for the group. Fig. 1 provides an overview of the method. The four major phases of the process are: (1) data preparation, (2) pre-processing, (3) clustering and (4) post-processing.

2.1. Data preparation

The aim of the data preparation is to develop a database that includes all buildings with the describing parameters that are required to estimate the effect of energy efficiency measures. Important parameters include the total and heated floor area and the use of the buildings as well as external surface of façade, windows and doors, roof, walls, etc. Orientation, geometry and available roof area play an important role for solar collectors (photovoltaic, PV and solar thermal, ST). Moreover, the existing heating system and the consumed fuel or electric power per year are valuable information. Finally, the setup of a GIS model was a helpful means in the Zernez project for interpreting results, for spatial analysis and for proposing district networks.

2.2. Pre-processing

2.2.1. Defining performance-based feature space for clustering

The first step is the definition of the dimensions of the feature space for clustering. This definition leads to j features characterizing a building instance. These features consist either of the effect of a simple energy efficiency measure (EEM), for instance, insulation of the building envelope or the effect of a measure set, for instance, insulation of the envelope and new heating system. For each feature, to describe the cost effectiveness, the effect e of the measures or the measure set is determined as the quotient of emission reduction ΔV_{CO_2} and the investment costs c_{invest}:

$$e_j = \frac{\Delta V_{CO_2,j}}{c_{invest,j}} \text{ in g } CO_2/\text{a per CHF}. \quad (1)$$

Alternatively, clustering using only CO_2 reductions was performed. This delivers information insensitive to costs for cases where only the reduction of emission are considered without analysis of any investment costs.

2.2.2. Determining the effect matrix for clustering

The basis for clustering is the response of each building to each measure, which defines the matrix **E**. This matrix locates every building in the virtual feature space defined in the previous step. For the building i and the measure j, the effect matrix composes as follows:

$$\mathbf{E} = e_{i,j}. \quad (2)$$