Renewable Energy 129 (2018) 570–582

Contents lists available at ScienceDirect

Renewable Energy

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

Comparison of clustering algorithms for the selection of typical demand days for energy system synthesis

Thomas Schütz*, Markus Hans Schraven, Marcus Fuchs, Peter Remmen, Dirk Müller

RWTH Aachen University, E.ON Energy Research Center, Institute for Energy Efficient Buildings and Indoor Climate, Mathieustr. 10, Aachen, Germany

ARTICLE INFO

Article history:
Received 14 February 2018
Received in revised form
27 May 2018
Accepted 7 June 2018
Available online 14 June 2018

Keywords:
Building energy system
Clustering
Mixed integer linear programming
Renewable energies
Typical demand days

ABSTRACT

The optimal design, sizing and operation of building energy systems is a complex problem due to the variety of available generation and storage devices as well as the high-resolution input data required for considering seasonal and intraday fluctuations in the thermal and electrical loads as well as renewable supply. A common measure to reduce the problem's size and complexity is to cluster the demands into representative periods. There exist many different algorithms for the clustering, but to the best of our knowledge, no comparison has been made that illustrates which algorithms are the most appropriate for such problems.

Therefore, this paper compares six aggregation methods for reducing full year input data to typical demand days for energy system synthesis. We consider seasonal and monthly classification as well as sophisticated clustering methods such as k-centers, k-means, k-medians and k-medoids for aggregating the heat and electricity demand as well as solar irradiation onto the roof of a single-family house and an apartment building.

The results show that all clustering methods are able to determine energy systems that are close to the optimal system, however their demand related costs are approximated best and most reliably with k-medoids.

1. Introduction

The optimal design of building energy systems is a highly complex problem that addresses the usage, generation and storage of multiple forms of energy. An appropriate model simultaneously has to consider a wide range of generation and storage technologies as well as high-resolution input data for accurately modeling the seasonal and intraday fluctuations of thermal and electrical loads.

The combination of these aspects – high-resolution inputs, the variety of available devices, and storage options – typically leads to long computing times in optimization problems that often exceed available resources. Possible methods to cope with this task include mathematical approaches like decomposition [1], or model simplifications, for example reducing the number of considered technologies and devices, or the level of detail with which each are modeled [2]. A further option for improving computing times is the aggregation of input data to representative time periods, for example by modeling full year data with twelve monthly averaged typical days [3–5].

1.1. Literature overview

Different clustering methods and combinations of clustering with the other complexity reduction methods mentioned previously are widely used for energy system synthesis.

Yokoyama et al. [1] present a decomposition method for energy system synthesis and apply a season based clustering (SBC) resulting in three representative demand days with hourly sampling intervals for winter, summer and the transitional periods. Their considered energy system consists of absorption chillers (AC), gas boilers (BOI), compression chillers (CC), combined heat and power (CHP) technologies as well as heat pumps (HP). Zhu et al. [6] optimize the energy system of a Chinese airport considering the same technologies and aggregation methods as [1].

Wakui and Yokoyama [7] describe an optimization model for building energy systems consisting of BOI, CHP, electrical resistance heaters (EH) and thermal energy storage (TES) systems. Full year inputs are modeled with three seasonal representatives for winter, summer and transition. Furthermore, they consider a peak day in winter and summer, leading to a total number of five typical demand days with hourly resolution. This modeling of input profiles

* Corresponding author.
E-mail address: tschuetz@eonerc.rwth-aachen.de (T. Schütz).

https://doi.org/10.1016/j.renene.2018.06.028

Nomenclature

Variables and parameters

A	Area, m^2
C	capacity, kWh
$\dot{E}$	Gas consumption, kW
$\dot{Q}$	Heat flow rate, kW
P	Electrical power, kW
V	Volume, m^3
C	Costs, Euro
R	Revenue, Euro
Y	Data point, —

Greek letters

H	Efficiency, %

Subscripts and abbreviations

AB	Apartment building
AC	Absorption chiller
BAT	Battery
BOI	Boiler
CC	Compression chiller
CCHP	Combined cooling, heat and power
CHP	Combined heat and power
DHW	Domestic hot water
EH	Electrical resistance heater
HP	Heat pump
MBC	Monthly based clustering
PV	Photovoltaic
SBC	Season based clustering
SFH	Single-family house
SH	Space heating
SSE	Sum of squared errors
STC	Solar thermal collector
TES	Thermal energy storage
VDI	Association of German engineers
W	Week
WT	Wind turbine
ann	Annualized
D	Typical demand day
dem	Demand
feed	Feed-in
inv	Investment
met	Metering equipment
nom	Nominal
o&m	Operation and maintenance
t	Time period

has also been applied by Wakui and Yokoyama [8] as well as Wakui et al. [9] for building energy systems with extended superstructures.

Moradi et al. [10] optimize multimegawatt CHP systems considering battery (BAT) storages as well as BOI and TES. They represent input profiles with four seasonal representatives that distinguish spring and fall, and differentiate between weekdays and weekends.

Lozano et al. [11] optimize a multimegawatt combined cooling, heat and power (CCHP) system for 5000 apartments in Spain, comprising AC, BOI, CC and CHP. In this study, twelve typical demand days are used that each represent one month of the original input data. A similar methodology is applied by Mavrotas et al. [4] for optimizing the energy system of a hospital in Greece. Schütz et al. [12] also use a month based clustering (MBC) for reducing the inputs for simultaneously optimizing a building energy system and passive building components. This study considers similar technologies as [9] but also accounts for photovoltaic (PV) modules and solar thermal collectors (STC). MBC methods have also been used by Stadler et al. [13] for optimizing energy systems while considering retrofit options. In contrast to the previously mentioned studies [13], uses three typical days per month to account for weekdays, weekends and peak days.

Other studies that employ aggregation methods based on fixed periods of the year to optimize individual energy systems, include Merkel et al. [14] and Buoro et al. [15]. In contrast to the other works mentioned so far, these studies use typical weeks (W) instead of typical days. Merkel et al. [14] use a quarter-hourly time discretization and consider three typical weeks for each season (winter, summer and transition). Buoro et al. [15] use one typical week for each month of the year.

Whereas all previously mentioned studies dealt with single energy systems, Yang et al. [16] optimize an urban area in China considering four energy systems as well as energy transfer between them simultaneously. The implemented superstructure consists of AC, BOI, CC, CHP, PV, TES and wind turbines (WT). In this study, three seasonal representatives are chosen and each day consists of twelve discrete time steps. Other studies that optimize multiple energy systems and employ a SBC with two (summer and winter) [17], three (summer, winter and transition) [5,18–20] as well as four (summer, winter, spring and fall) representative demand days include [21].

Harb et al. [3] also deal with the optimization of energy systems for residential neighborhoods. They reduce the original input data to twelve representative demand days via MBC.

All previously mentioned works focused on the modeling or application of their developed energy system optimization framework. In contrast, the following studies explicitly focus on input data preparation. Fazlollahi et al. [22] proposed a k-means clustering approach for determining typical demand days. Their application case consists of a designing the energy system of a district heating grid with an annual heat demand of 2100 GWh and for computing its operating strategy. They analyzed one to fifteen typical demand days and suggest that the performance only marginally improves with more than four days. A second application is also presented in Fazlollahi et al. [22] regarding the operation scheduling of a district heating system including renewable generation.

In contrast, Domínguez-Muñoz et al. [23] present a clustering method that is based on k-medoids clustering. They used this approach for compressing the heating and cooling demands of a Spanish technology park into a reduced number of typical demand days. The authors suggest to model full year inputs with ten representative demand days.

Schiefelbein et al. [24] also used a k-medoids clustering algorithm to combine the original input data into typical demand days. In their study, they optimized the energy supply of a residential neighborhood of five buildings, considering BOI, CHP and TES. They analyzed the effect of different numbers of typical demand days on the resulting energy systems and conclude that at least seven typical demand days were necessary for their specific application. With less demand days, the optimal energy system layout changed,