A DATA MINING APPROACH TO INDOOR ENVIRONMENT QUALITY ASSESSMENT

A study on five detached houses in Finland

Mikko Saarikoski
MSc Thesis
Environmental Science
University of Eastern Finland; Department of Environmental and Biological Sciences
March 2016

UNIVERSITY OF EASTERN FINLAND, Faculty of Science and Forestry
Environmental Science
Mikko Saarikoski: A data mining approach to indoor environment quality assessment
MSc thesis 62 pages, 7 appendixes (10 pages)
Supervisor: Mikko Kolehmainen (D.Sc), University of Eastern Finland, Department of Environmental and Biological Sciences
March, 2016

keywords: indoor environment quality, air quality, data mining, sensor data, residential housing

ABSTRACT

Achieving and maintaining good indoor environment quality (IEQ) while improving energy efficiency of housing is a globally relevant goal. Current developments of sensor networks are increasing the availability of high-resolution data from buildings, which offers opportunities for better understanding of the indoor environment dynamics in different situations. This thesis presents an application of data mining for assessing indoor environment quality in five modern Finnish residential houses, using nine months of observations gathered by static wall-mounted sensors for temperature, relative humidity, carbon dioxide, carbon monoxide and differential pressure. A set of open weather data is integrated into the analysis as background variables.

K-means clustering algorithm is used for partitioning the observations into clusters, based on the multidimensional structure of the data. The clusters are visualized on a two-dimensional plane using Sammon's mapping. Indoor environment quality situations defining the clusters are interpreted by evaluating the distributions of the variables.

Patterns of weather, occupancy and use of household appliances were identified as the main influencing factors responsible for the variations in the IEQ data. Based on the assessment of variable levels against Finnish guidelines, the IEQ in the houses was considered mostly good. Occasional levels beyond the guidelines were recognized and their causes discussed.

The data mining approach in this study can be extended to other built environments. For a better view to the relationships between energy consumption and IEQ, integrating data from the HVAC system to the analysis would be a sensible step for future research.

ITÄ-SUOMEN YLIOPISTO, Luonnontieteiden ja metsätieteiden tiedekunta
Ympäristötiede
Mikko Saarikoski: Tiedonlouhinta sisäilmaston arvioinnissa
Pro Gradu -tutkielma 62 sivua, 7 liitettä (10 sivua)
Tutkielman ohjaaja: professori Mikko Kolehmainen, Itä-Suomen Yliopisto, Ympäristö- ja biotieteiden laitos
Maaliskuu 2016

---

avainsanat: sisäilmasto, tiedonlouhinta, klusterointi, sensoriverkko, asuminen

TIIVISTELMÄ

Sisäilmaston ongelmien ehkäisy sekä energiatalouden parantaminen ovat asuinrakennuksia ajatellen yleisesti tunnistettuja tavoitteita. Viimeaikainen sensoriverkkojen yleistyminen on lisännyt tiheän mittausdatan saatavuutta rakennusten sisältä. Tiheisiin mittauksiin perustuva tutkimus voi parantaa ymmärrystä sisäilmaston vaihtelusta erilaisissa tilanteissa. Tässä pro-gradu -tutkielmassa sovelletaan tiedonlouhintaa viiden suomalaisen omakotitalon sisäilmaston arviointiin. Tutkimuksen havaintoaineisto on yhdeksän kuukauden mittainen ja se on kerätty taloihin asennetuilla lämpötilaa, suhteellista ilmankosteutta, hiilidioksidia, hiilimonoksidia ja paine-eroa mittaavilla sensoreilla. Taustamuuttujina sisäilmaston muutosten arvioinnissa käytetään avointa sääaineistoa.

Havaintoaineisto ryhmiteltiin K-means algoritmin avulla klustereiksi perustuen muuttujayhdistelmien vaihteluun. Klusterit havainnollistettiin kaksiulotteisina Sammonin kuvauksen avulla ja sisäilmaston tilanne kunkin klusterin taustalla tulkitaan muuttujien jakaumien perusteella.

Sisäilmaston muutoksiin eniten vaikuttaneiksi tekijöiksi tunnistettiin vaihtelut säätilassa, rakennusten käyttöasteessa sekä kodin laitteiden käytössä. Aineiston perusteella talojen sisäilmasto on ollut enimmäkseen hyvä. Joissain tilanteissa tunnistettiin ohjearvojen ylityksiä ja ylityksien mahdollisia syitä käytiin läpi. Tutkimuksessa käytettyjä menetelmiä voidaan hyödyntää myös muissa rakennetuissa ympäristöissä. Talotekniikkajärjestelmän muuttujien liittäminen analyysiin vaikuttaa hyvältä aiheelta jatkotutkimusta ajatellen.

FOREWORD

After several years of working outside academia, last autumn I realized time was ripe to finish my Master's studies with a thesis. I contacted the Environmental Informatics group in Kuopio and we figured out that I could attempt to make a contribution to their research of indoor environments using a dataset of sensor observations from residential buildings. The basic dimensions in the dataset (temperature, humidity, CO_2 and CO) were rather familiar to me when starting this work, but looking at the variables one at a time seemed to offer only a narrow representation of the buildings dynamics, considering that there is a need to find new integrated ways to improve both indoor environment quality and energy performance. Therefore, in pursuit of a more holistic picture of the variations, we chose to examine the data from a multidimensional perspective, which was challenging at first. However, after getting used to the tools and methods the selected approach turned out to provide some intriguing insights. It has been exciting to look at common environments from a new scientific perspective, and I learned a lot about buildings and data analysis through this study.

During the whole process of this thesis I've received great support from the Department of Environmental and Biological Sciences, especially the Environmental Informatics group; many thanks for expertise, guidance, data pre-processing help and discussions to Mikko Kolehmainen, Mauno Rönkkö, Marcus Stocker, Robert Ciszek, Markus Johansson, Jukka-Pekka Skön and Mika Raatikainen. Thanks to Mikko Kolehmainen for supervising this work and to Marko Hyttinen for reviewing it. Thanks to Academy of Finland for funding part of this work via the FResCo project. Thanks to Department of Built Environment at Aalto University and Helsinki University Library for providing pleasant working spaces near my home. Finally, thanks to friends and family for support and good times.

CONTENTS

APPENDIXES 1-7

1. INTRODUCTION

Buildings are made to perform a multitude of functions. One of the most basic requirements for residential houses is to provide shelter from the physical environmental stressors such as wind, rain, sunlight and cold or hot temperatures. In modern houses the separation of indoor and outdoor environments is typically created by designs where the building envelope is rather airtight and insulated. The airflow into buildings is usually controlled with mechanical ventilation. Currently people spend significant amount of time indoors, which suggests that the indoor environments have a large potential for health effects (Straube, 2006; Mitchell et al. 2007).

The goal of achieving good, healthy and comfortable indoor conditions is interwoven with the need to reduce the systemic impacts that buildings cause on the environment during their life cycle. In order to achieve improvements regarding both of these goals a holistic understanding of the processes involved is necessary. Current developments in measurement technology, especially sensor networks enable effective collection of data of many parameters which have an effect on the overall indoor environment quality (Mitchell et al. 2007).

Traditionally the main focus in the assessment of indoor measurements has been placed on the univariate behavior of parameters. However, defining standard limit values or even the relative importance of different parameters has turned out to be difficult in many cases, because they can be highly dependent on the context and occupant preferences. The large amount of data by itself possesses challenges for traditional methods of analysis.

This paper presents an example of how data mining and unsupervised machine learning could be used to identify key dynamics which affect the indoor environment in a building. We use K-means algorithm to examine the multidimensional structure of nine months of sensor data from five detached residential houses in Finland. The data is partitioned into sets of clusters which describe different situations in the houses. Results are interpreted using visual and statistical tools, and the implications discussed. The aim of this study is to test the applicability of data mining in the analysis of buildings indoor conditions, which are the main contributor of the indoor environment quality.

2. LITERATURE REVIEW

Buildings interact with their surroundings in many ways. As the scale and complexity of human societies has increased, the total effect of buildings on the environment has become more significant. A 30-40 % share of the global primary energy use and 40-50 % of all greenhouse gas emissions can be attributed to buildings. These insights have led to political actions which aim to reduce the climate impact of buildings throughout their life-cycle. Analysis has indicated that the operating energy accounts for 80-90 % of the buildings life-cycle energy use, so measures have concentrated on achieving good IEQ with minimal energy consumption (Ramesh et al. 2010).

In a review of the health effects of IEQ, Mitchell et al. (2007) concluded that the exposures with potential health effects are significant, and they are resulting from interactions between the building structure, systems, the outdoor environment, the occupants and their activities. The review identified a need for research concerning the circumstances that make the exposures more likely and the effectiveness of interventions, including potential tradeoffs and confounders.

In the following subchapters we will look at the research literature from four perspectives, first the characteristics of dwelling in the Nordic climate, then the various ways of measuring and evaluating IEQ, after that the results of IEQ research (with focus on cold climates) and last the applications of data mining for building-related data.

2.1. DWELLING IN THE NORDIC CLIMATE

Based on the historical nature of buildings as systemic artefacts which change gradually in time, Pirinen (2014) argues that dwelling can be seen as an evolutionary realm. The adaptations of buildings occur in a process which is molded by a various socio-cultural goals and the environment. Assuming energy and IEQ as general goals, Teixeira Chaves (2012) points out the significance of surrounding climate conditions for building designs.

The subarctic climate of Fennoscandia, with temperatures ranging from below -20°C in winter to over +20°C in summer (Pirinen et al. 2012), sets distinctive demands for design and operation of residential buildings. Figure 1, portrays the principles of seasonal thermal dynamics for buildings in northern hemisphere. In order to achieve good energy efficiency, a building should adapt to cold and warm seasons so that the indoor thermal environment naturally inclines towards the comfort zone. In the figure, Building 1 is an example of good thermal performance whereas Building 2 performs poorly in winter and summer.

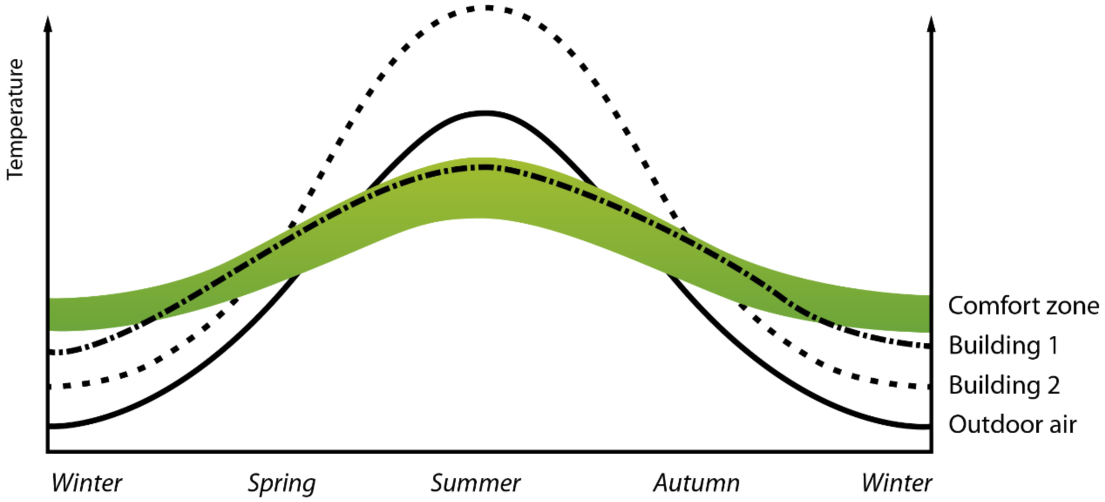


Figure 1. Seasonal thermal dynamics for indoor and outdoor air in the northern hemisphere (modified from Roulet, 2011)

In the agrarian times the Finnish houses have been typically constructed out of wood, with natural ventilation and a wood-burning stove for heating and humidity control. In the 1900's many new materials and designs have been adopted for use in construction. Since the 1970's new buildings have become more airtight and heavily insulated than before. Therefore, in the modern mechanically ventilated buildings, functioning of the HVAC system has become a critical determinant of the indoor air quality. In a study of seven Finnish detached houses, Kaksonen (2012) noted that the operation of the ventilation system has a large influence on the IEQ.

2.2. MEASUREMENT AND EVALUATION OF INDOOR ENVIRONMENT

The quality of the indoor environment is a challenging concept for measurement and analysis because it involves two complex dimensions, the physical/objective and the personal/subjective (Heinzerling et al. 2013). This dichotomy of dimensions can be philosophically invalid in some situations, but in this subchapter it will be used as a working operationalization, in order to shed light on the different approaches used in current research.

Measuring the “objective” dimension of IEQ requires knowledge of several inter-related variables which have uneven temporal and spatial variations in the building. In the Finnish housing health instructions there are three measurement categories: (1) physical conditions, (2) chemical impurities, particles and fibers and (3) microbiological conditions. For each of these categories official guidelines for sampling and measurement are presented. The guidelines are designed to support municipal health officers in investigating potential IEQ problems, with a focus on problem detection and evaluation of compliance through (mostly) manual snapshot sampling (Ministry of Social Affairs and Health, 2003, 2015). Other set of guidelines, which is commonly used in Finland is the 3-step Classification of Indoor Environment (Säteri, 2015).

The subjective dimension of IEQ is characterized by the individual occupants of the building and their personal responses and preferences regarding the desired state of indoor environment. For example Bluyssen (1999) has noted that the occupants themselves are the best source for information concerning the quality of the indoor environment where they live. Looking at the standardization of comfortable indoor conditions from a sociological perspective, Shove (2003) points out that the standard definitions for comfort have had a tendency to develop narrower in time. This has, according to Shove, caused air-conditioned spaces to become the norm for even mild climates, which has in turn lead into expansion of electric air-conditioning systems at the cost of locally adapted building/housing/clothing practices. The adaptation to varying thermal conditions is apparent in an acceptance study of 125 persons living in 32 Hong Kong apartments by Lai et al. (2009), where occupants were noted to accept indoor temperatures up to 29°C by adjusting their clothing.

Relying on subjective measures makes it difficult to identify the causes and effects related to changes in the indoor conditions. For example Sharpe et al. (2014) report that in a UK subpopulation, increased risk of asthma was associated with living in more energy efficient homes. The paper presents hypothetical causative agents which might explain the association, including inadequate heating, ventilation or increased concentrations of biological, chemical or physical contaminants. Yet, working with relatively sparse data researchers recognize that the possible effect is likely to be modified by complex interaction between behavioral and environmental factors. They also note that findings may be confounded by the response rate, demographic and behavioral differences between residents of low and high energy efficiency homes and that the exposures and outcome data is self-reported through the questionnaire.

As a research strategy, a combination of objective and subjective measures seems to be often the most desirable one. Usually the resources for data gathering and analysis limit the scope of study, but useful insights of IEQ can be acquired even with subjective or objective measures only. In many research settings the nature of data presents several common challenges that have to be faced in the investigation.

Heinzerling et al. (2013) point out two potential issues for both objective and subjective measures, with reference to Nicol and Wilson (2011). These are (1) determining the representative period for the gathered data and (2) difficulty of interpreting the results. On the one hand, a snapshot sample is relatively easy to obtain, for example with single survey or a measurement cart, but it offers poor temporal generalizability. On the other hand, continuous measurement requires new tools and methods for processing the data into meaningful summaries.

2.2.1. Some remarks on the models and tools for IEQ research

A big portion of the building systems research is focused on commercial buildings. For example the literature review on IEQ evaluation models by Heinzerling et al. (2013) offers an overview on various IEQ measurement carts and desktop devices used in commercial buildings. The measurement categories for these devices are acoustics, indoor air quality, lightning and thermal comfort. Most common variables for these categories are sound level, CO2, illuminance and air temperature. Measurement with handheld devices or movable carts enables sampling to be adjusted so that it corresponds to occupants movements in the house, therefore potentially giving good picture of the actual exposure. However, for a holistic understanding of the buildings indoor dynamics, continuous measurements are valuable and they are hard to obtain with manual methods. Static sensor networks are currently used in many environmental monitoring tasks and they provide automated cost-effective way to gather data with dense temporal resolution (Stocker et al. 2012).

Focusing on efficient use of IEQ data for the planning of renovations, Bluyssen (2000) describes a software tool (titled EPIQR) which combines occupant complaints inventory (acquired through a questionnaire) with a diagnostic checklist which gives direction for observations regarding the degradation of critical building elements. The eight categories for IEQ in the questionnaire are

1. Winter thermal comfort
2. Summer thermal comfort
3. Air quality
4. Natural lightning
5. Acoustic comfort – outdoor sources
6. Acoustic comfort – indoor sources
7. Water quality
8. Safety

As an example of application of EPIQR, Balaras et al. (2000) used the tool to audit 38 multi-family apartment buildings located in United Kingdom, Switzerland, Netherlands, Greece, Germany, France and Denmark. In addition to the questionnaire and observations, the study measured indoor and outdoor T and RH in eight buildings and concluded that high energy consumption does not always provide satisfactory thermal conditions. Average IEQ was found to be best in Switzerland and worst in Greece, however the n is small so the results cannot be generalized.

Regardless of the tools and variables in use, the data itself often poses distinctive challenges for research. Dissertation of Niska (2012) offers a discussion of the general characteristics of environmental data, outlining various challenges that are often faced in the field of environmental informatics. Diversity of underlying dynamics in environmental systems cause the data to become noisy and chaotic. Green and Klomp (1998) have recognized four categorical sources of complexity in environmental systems; (1) spatial and temporal scales, (2) non-linear interactions and feedback loops, (3) high number of influencing factors and (4) human influence. All of these categories are applicable for IEQ data. In order to cut down some of the layers of complexity, restricting the scope of study to a single building type within a certain climate is one option.

2.3. PREVIOUS RESEARCH ON RESIDENTIAL IEQ IN COLD CLIMATES

Considering buildings from a holistic perspective, the site sets constraints and opportunities for sustainable design decisions, thereby strongly affecting the quest for good IEQ. Many characteristics of the site can be seen as unique manifestations of the interplay between nature, culture and technology. They make a big difference for the intuitive sense-perception which defines the atmosphere in a certain place, but they might be difficult to break down into quantified form (Pallasmaa, 2011). However, as we recognized in previous chapters, climate patterns offer a useful general context for research. This subchapter presents a view to some of the previous residential IEQ research in cold climates.

During years 2002-2004 Vinha et al. (2005) investigated 102 Finnish wooden detached residential buildings, looking at the conditions of humidity, temperature, ventilation and airtightness. The study recognized that in the winter the variability of indoor temperatures was larger than expected and in the summer houses were often significantly too hot. The ventilation rates for 2-person bedrooms were often too low. The variability of total energy consumption in the houses was high, indicating a crucial effect of the residents' lifestyles. Generally the IEQ was found to commonly deviate from the Finnish housing health instructions (Ministry of

Social Affairs and Health 2003) The instructions recommended indoor air temperatures to be between 19°C and 24°C around year, which means that in winter the temperature gradient between outdoor and indoor air can sometimes be over 40°C. In the summer indoor temperatures can easily exceed 24°C when outdoor temperature is high and the sun shading is not sufficient.

Research of built environment is most commonly based on data from a single country, but international projects are not rare either. Du et al. (2015) conducted a research on 16 multi-family buildings in Finland and 20 in Lithuania, all of which were waiting to be renovated. Data of T, RH, CO_2, CO, PM, NO_2, formaldehyde, VOCs, radon and microbial content in settled dust was measured and then analyzed in unison with health and housing quality questionnaire data. The paper concluded that most parameters were within recommended limits, yet differences in the baseline levels between the countries were recognized for thermal conditions, ventilation and the respondents' satisfaction with their residence and IAQ. Thermal conditions and ventilation adequacy were noted as having biggest potential for IEQ improvements in the studied buildings.

In order to investigate the relationship of tenure status and IEQ with both objective and subjective measures, Pekkonen et al. (2015) compared data from a housing and health questionnaire survey to two months of continuous T, R and CO_2 measurements in 28 Finnish apartments using cross tabulations and logistic regressions. The study recognized that housing satisfaction was lower in rental flats than owner-occupied apartments, which supports the notion that social circumstances have a large effect on the experience of IEQ. Complexity and diversity of lifestyles makes the handling of confounding, mediating and suppressing factors a challenging yet important task when looking at the social aspects of housing.

Large parts of Canada fall into the subarctic climate zone and the country's economic development path is relatively close to Finland, therefore offering relevant benchmark for Finnish IEQ research. Recently Sharmin et al. (2014) conducted a research of energy consumption, thermal performance and IEQ in 4-storey residential buildings in the province of Alberta, using sensor data from 12 households. The study detected that CO_2 and RH levels sometimes exceeded ASHRAE limits in the study units, especially during winter season.

2.5. DATA MINING APPROACHES FOR BUILDING-RELATED DATA

Niska (2012) states that the analysis of environmental systems can be approached as an iterative data enrichment process. This suggests that even when the data of interest is not originated from a specific experimental design, it can be processed and analyzed in multiple stages to reveal useful knowledge. Following Fayyad et al. (1996), Niska recognizes five steps that are involved in discovery of knowledge from a database. These are (1) *selection* of target data from the database, (2) *preprocessing* of the data to handle quality problems such as missing values, (3) *transformation* to prepare the data scaling and dimensionality for (4) *data mining*, which produces patterns/models that are examined in the step (5) *interpretation*. Iterative approach means that some or all of these steps can be repeated according to initial and subsequent interpretations.

In 2013, Yu et al. proposed a general framework for data mining building-related data. Their framework includes four components: (1) data analysis techniques/algorithms, (2) potential applications of data mining in building engineering, (3) input (collected data) and (4) output (extracted knowledge). See *Figure 2* for an overview of Yu et al. framework, presented for the context of engineering energy performance improvements.

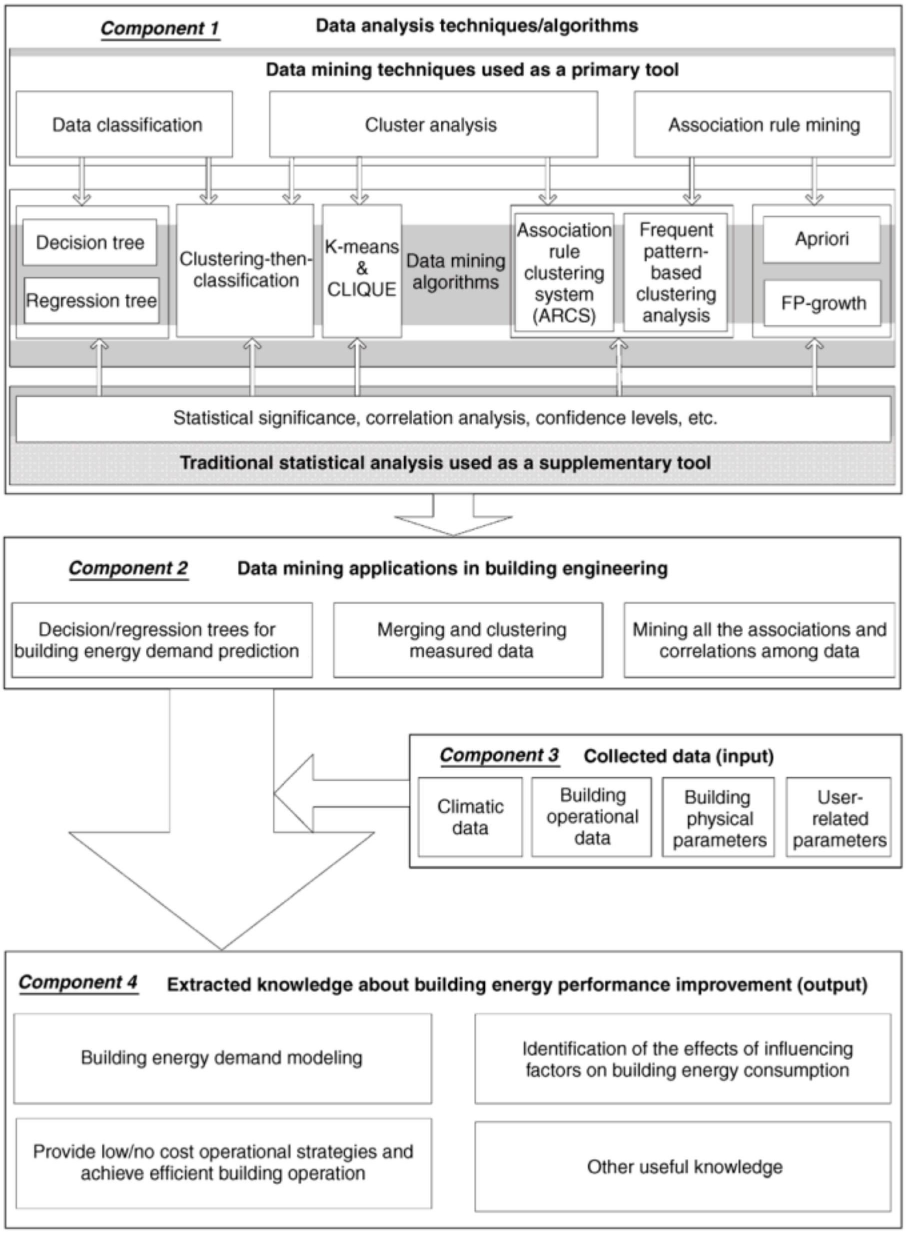


Figure 2. Data mining framework for building-related data, proposed by Yu et al. (2013)

Working with eight months of observations at 15 minute intervals, Xiao and Fan (2014) applied data mining to improve the operational performance of a commercial skyscraper in Hong Kong. Their data consisted of a lot of variables from the building automation system and some outdoor and indoor variables. Method of data mining used in the paper was a combination of K-means clustering and association rule mining, which is an algorithm-based way to identify recurring correlations, ‘association rules’ between the variables. Because the building was used according to the office hours, the association rules were generated separately for weekdays, Saturdays and Sundays. Most of the associations were seen as knowledge that would be easily acquired through domain knowledge, however the study was able to recognize certain abnormal running conditions in some pumps in the HVAC system, which caused waste of energy.

Other studies of data mining building data have focused for example on electricity consumption prediction (Li et al. 2015), energy efficient design (Kim et al. 2011) and the influence of occupant behavior on energy consumption (Yu et al. 2011). Some studies that focus on IEQ will be examined here closer.

Using temperature, humidity and light data from a sensor network installed in a research lab, Wu and Clements-Croome (2007) used data mining to investigate the relationships between the parameters. The original dataset turned out to be very noisy, so a lot of invalid values and some outliers were removed in the initial pre-processing. The target dataset covered business hours of 26 days. In the data mining step of the case study, K-means algorithm was used to sort the observations into four clusters (k=4). Researchers suggested that the results could be used for example to arrange working spaces so that the employees individual thermal comfort preferences would be satisfied in an efficient way. However, the interpretation of the clusters was only shortly discussed in the paper.

Raatikainen et al. (2012) examine the effects of weather conditions and differential air pressure on indoor air quality by analyzing two months of sensor data from a single residential house using data mining methods, namely Self-Organized Map (SOM) and K-means clustering. The study used Davies-Boulding index to assess the number of clusters (k). Data was then clustered at k=9, and interpreted using SOM visualizations and statistical boxplots for each cluster. The clusters corresponded to various situations which were mostly defined by the occupancy status. Self-Organized Maps on *Figure 3* show that T and CO in the living room are correlated, which is probably caused by the use of fireplace. The indoor air quality in the house was good during the study period. The paper concluded that pressure difference has an impact on the indoor air parameters and that the used clustering technique was useful in revealing dependencies between variables.

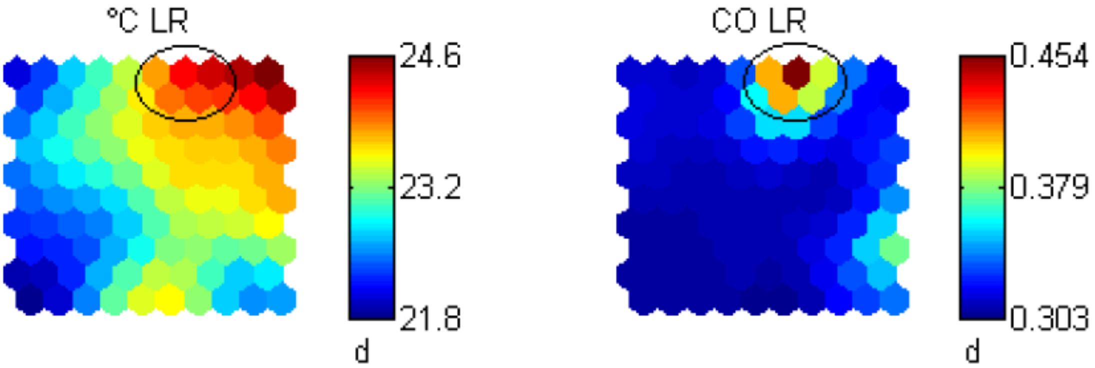


Figure 3. Example of a SOM visualization, used to assess the dependencies of T and CO in Raatikainen et al. (2012)

For a discussion on the possibilities of automation in the context of IEQ data, see Stocker et al. (2012), who present a method for automated representation of knowledge for univariate sensor data from a residential house. Two patterns examined in the paper are (1) the elevations of CO and (2) shower use. The study recognized that computational techniques enable knowledge acquisition from numerical time-series data, and that different problem classes need different approaches. The selection of approach should consider the frequency of available data and formalizability of the phenomena of interest. If necessary, different methodologies can be combined.

3. MATERIALS AND METHODS

The houses investigated in this study are located in Kuopio, Finland. Kuopio is located in the subarctic climate zone, with Köppen-Geiger classification Dfc. See *Figure 4* for the location of Kuopio and the variety of climate zones in Europe.

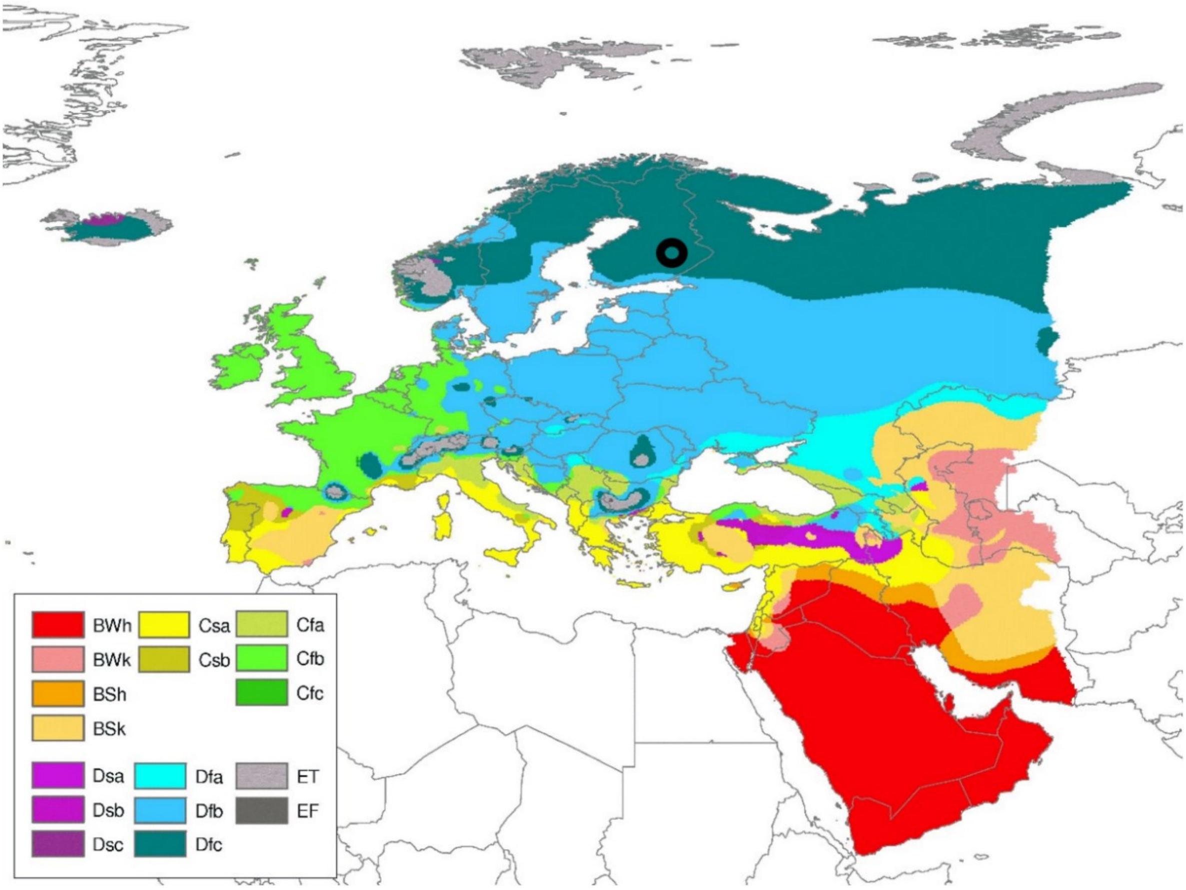


Figure 4. Climate zones in Europe and the location of Kuopio (black circle) (derived from Peel et al. 2007)

3.1. MATERIALS

This study uses data from two main sources. First is the database of the monitoring system originally developed as a part of AsTEKa-project and the other is public weather observation data from the Finnish Meteorological Institute. Additional source for information is the website of the Finnish housing fair, where the houses were displayed in 2010.

3.1.1. The monitoring system

This study uses data collected from 5 detached houses in Kuopio, Finland. The measurements were obtained via a sensor network, which was part of a monitoring system installed in 11 houses at the Kuopio Housing Fair in 2010 (see *Figure 5*). The monitoring system consists of sensors for carbon dioxide, relative humidity, temperature and carbon monoxide. Besides the indoor air parameters, the system gathered data off building pressure, ventilation duct pressure, electricity, water, district heat consumption and occupancy of the houses. Inhabitants could follow the IEQ/environmental performance of their own house through a www interface (Skön et al. 2011).

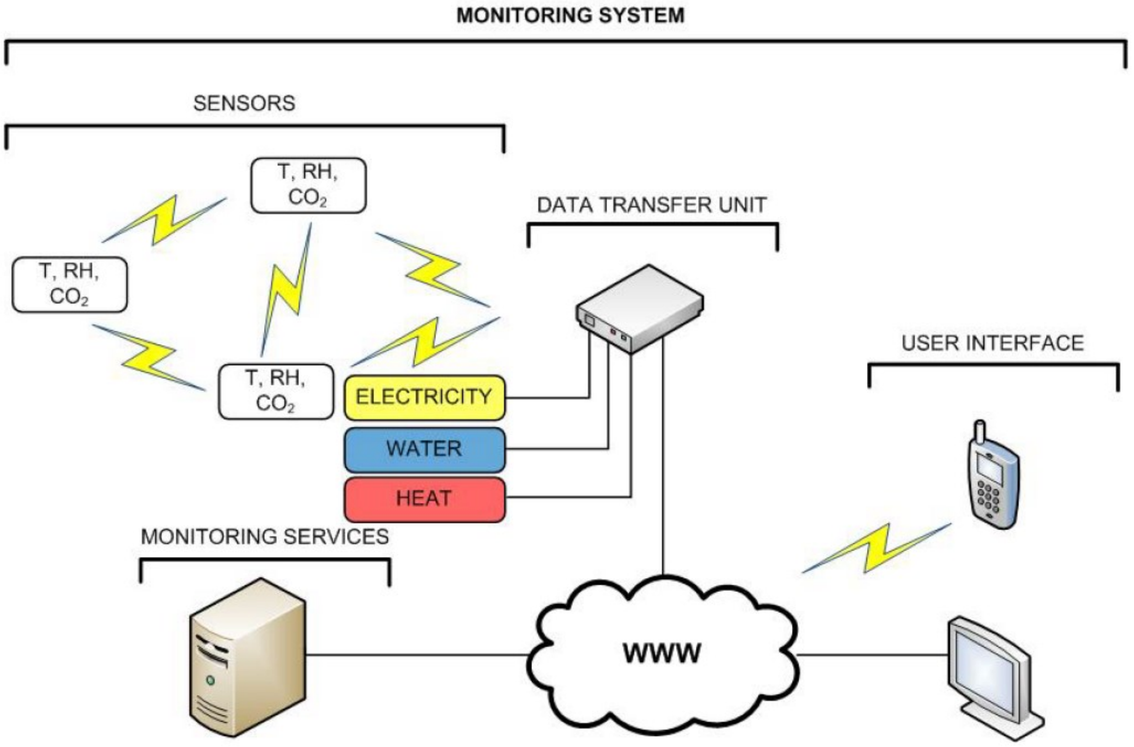


Figure 5. Measurement sensors and the monitoring system (from Skön et al. 2011)

The parameters, units and the sensor devices used for the indoor monitoring are shown on *Table 1*. The sensor for T, RH and CO_2, “EE80” is manufactured by E+E Elektronik and the CO sensor, “F2000TSM-CO-C101” is manufactured by Tongdy Control Technology. The pressure sensor used for both the building pressure (difference) and the in-duct pressure is Dwyer MS-221. For more details of the sensors, see referred datasheets (E+E Elektronik 2014; Tongdy Control Technology 2012; Dwyer Instruments 2009).

Table 1. Monitoring system parameters, units and sensors

Parameter	Unit	Range	Accuracy	Sensor
Temperature	°C	0-50	± 0,3 °C (at 20 °C)	EE80
Relative humidity	%	10-90	± 3 % (for 30…70 % RH at 20 °C) ± 5 % (for 10…90 % RH at 20 °C)	
Carbon dioxide	ppm	0-2000	< ± 50ppm +2% of measuring value (at 25 °C and 1013 mbar)	
Carbon monoxide	ppm	0-99	< ± 1 ppm (at 20±5 °C / 50±20 % RH)	F2000TSM-CO-C100
Pressure	Pa	0-100	±1 % for 0.25˝ (50 Pa)	Dwyer MS-221

3.1.2. Weather data

We were originally expecting to have microclimatic weather data from the site, gathered by two outdoor sensors stations on the different parts of the neighbourhood, but the data had been lost because transfer/server faults. As a backup plan we used weather data from the public database of the Finnish Meteorological Institute (FMI).

The closest weather station to the study location turned out to be six kilometers away from the study area (see *Figure 6*). Weather data was acquired from the FMI open data server using a dedicated open source downloader app (Salmi 2012). Temporal resolution for the raw weather data was 1 observation per 10 minutes. This weather data was considered to present an acceptable approximation to the general conditions at the site at a resolution of +/-1 hour. However the finer resolution effects of microclimatic variations on the site were decided to be left out of this study.

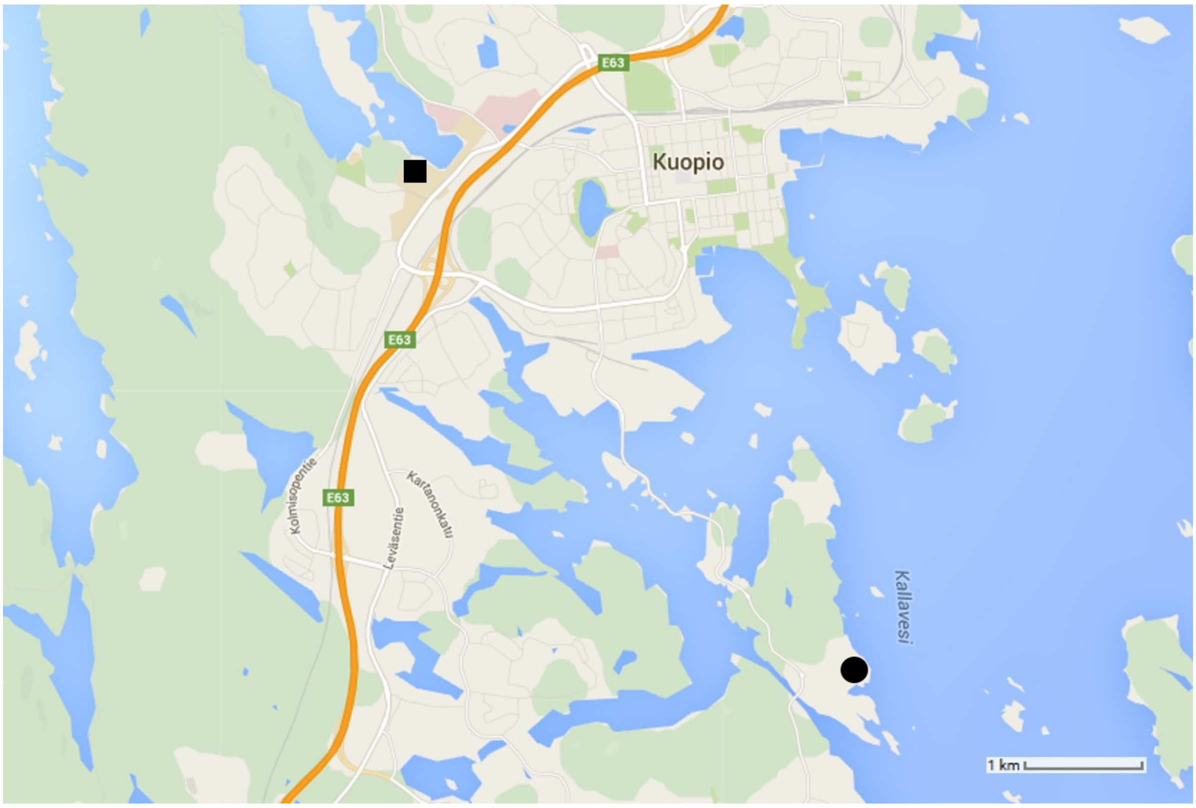


Figure 6. Map of the study site (●) and the weather station (■) (adapted from Google Maps, 2015)

3.2.3. Characteristics of the houses

Information concerning the characteristics of the houses under investigation was retrieved from the Housing Fair website. The basic characteristics concerning the five houses are presented on *Table 2*.

Table 2. Basic characteristics of the houses

House	Floor area (m^2)	Rooms	Occupants	Floors	Wall material	Energy Rating	Heating
A	155	6	5	2	Stone	A	Geothermal + fireplace
B	138,5	4	4	2	Stone	Passive	Fireplace with in-floor air circulation, electric
C	241,5	7	4	2	Stone	A	Geothermal
D	133	5	4	1	Wood	C	District heat
E	128	5	4	1	Wood	A	Hybrid system with fireplace and solar thermal collectors

The houses are built in compliance with The National Building Code of Finland (Ministry of the Environment 2015). The ventilation in all the houses is mechanical with heat recovery. Heating systems include geothermal, district heat, solar and fireplace configurations with various automated controls for hybrid use. The energy ratings of the houses on *Table 2* are determined according to the act 765/2007 by the Ministry of the Environment, which defines energy ratings through estimated total energy use per gross floor area (brm^2). By this standard, to achieve rating A the total energy consumption estimate has to be less than 150 kWh / brm^2 / year. Corresponding consumption for rating C is between 171 and 190 kWh / brm^2 / year. The Passive rating means less than 140 kWh / brm^2 / year, of which less than 25 kWh should go for heating. In 2013 the law concerning the building energy ratings changed, so these ratings are not comparable to the current ratings (Ministry of the Environment 2007 and 2013, Nieminen 2010).

Concerning the energy ratings, it can be noted that Heikkinen (2011) analyzed the energy consumption data from a single house in the housing fair area, using three months of winter data. She found out that the actual energy usage does not necessarily fall into the limits of pre-calculated energy rating.

Neighbourhood of the houses is shown on *Figure 7*. The site is located on gently sloping cape, Lake Kallavesi opening to east. House C is closest to the beachline, around 25 meters away. Distance to the lakeshore from houses B and E is ~100 meters and from houses D and A ~200 meters. The south side of houses D and B is characterized by a rather dense forest with mostly evergreen trees. The northwestern side of house C is also covered with some forest, albeit less dense than the one on the southern side of the district.

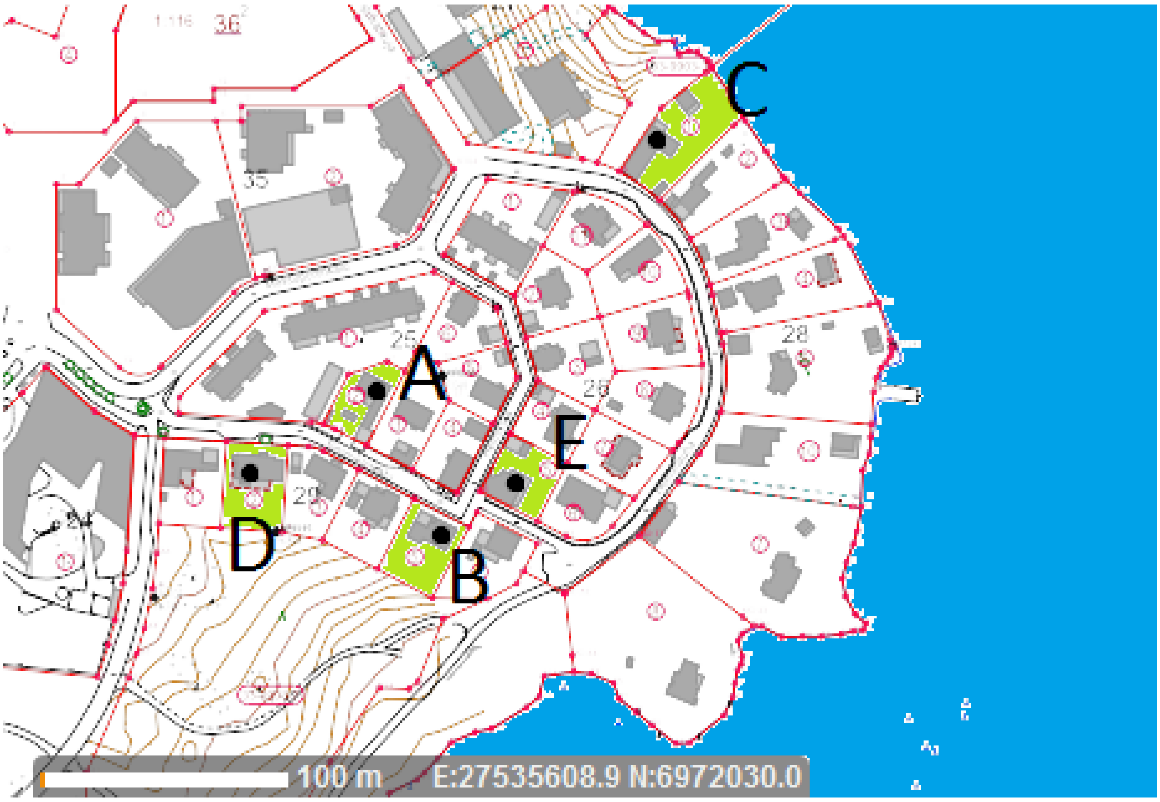


Figure 7. The houses, the neighbourhood and the vicinity of Lake Kallavesi. Black dots next to the letters indicate the houses (Adapted from Kuopio City online Guide Map 2016).

Facades of the houses are pictured on figures 8, 9, 11, 12 and 13. Photograph copyrights by Valokuvia Antero Tenhunen 2010. After each picture the positions of sensor-equipped rooms are described shortly.

Figure 8. House A viewed from west

The living room of house A is located behind the big window facing southwest, and it opens up to the second floor. The living rooms sensors are on the higher part of the open space. The bedroom is in the northeastern corner of the house, first floor, and the bathroom is in the south-facing single-storey extension. The bearing walls of house A are made of lightweight concrete blocks, with insulated concrete blocks on the outer walls. Ceiling structure is cavity slab, with 45 cm of insulation on top (Housing Fair Finland 2010). The insulated concrete blocks are rated for U-value of 0,12-0,15 W/m^{2}K (HB Betoniteollisuus 2012).

Figure 9. House B from southeast

The living room of house B spans both floors on the southern corner of the building. The fireplace is in the first floor, with the hot air circulation inside the floor between the storeys. The living room T/RH/CO_2 sensor is on the second floor eastern wall, and the CO sensor is next to the fireplace on the first floor. Bedroom is located in the first floor under the balcony, and the kitchen above it on the second floor (Housing Fair Finland 2010).

Wall structure of house B consists of passive blocks (see *Figure 10*). The U-value for the wall is 0,10 W/m^{2}K. Leakage rate (n50) of the house is > 0,6 h^{-1}, which is the passive house standard for airtightness.

Figure 10. Passive block before installation (Laakkonen 2015)

Figure 11. House C from northeast

In house C the living room is in the eastern corner of the house and the kitchen in the northern corner, both on the first floor. The bedroom is above the living room. The bathroom is on the second floor, middle of the house, on the northwestern side (Housing Fair Finland 2010).

Figure 12. House D from south

The sensor-equipped rooms in the house D are located as follows. The bedroom is in the southwestern corner of the house and living room next to it on the southern side. Kitchen opens from the living room to the northern side of the building. Next to the kitchen, near the northwestern corner of the house is bathroom. Insulation on the house D wall elements is 31 cm thick, and the ceiling has 50 cm of insulation (Housing Fair Finland 2010).

Figure 13. House E from east

The living room of house E is in the eastern corner of the building and it opens to the kitchen which is in the middle of the southeastern side. The bathroom is in the southern corner and bedroom in the northern corner.

Structure of the wooden wall elements in the house E, in order from inside to outside, includes following nine layers: (1) gypsum board, (2) vapour barrier membrane, (3) supporting vertical framework (48 x 198 mm), (4) insulation (200 mm), (5) horizontal framework (44 x 68 mm), (6) insulation (70 mm), (7) wind barrier (25 mm), (8) vertical frames with air gap (25 mm) and (9) horizontal timber facing (28 x 170 mm). The U-value for this wall structure is rated at 0,14 W/m^{2}K. The ceiling insulation of house E is 50 cm thick (Housing Fair Finland 2010).

3.2. DATA PRE-PROCESSING

As a first step of the pre-processing of the data, we evaluated the time resolution of the raw data. For most of the sensors the raw data from the measurements was saved on the server at intervals between 6 seconds and 1 minute. To determine what sample rate would be suitable for the analyses, a simple evaluation was made concerning the temporal dynamics of some short time span processes which are known to influence IEQ in residential buildings. We mapped out some potential processes which could cause quick changes in the IEQ, and estimated shortest plausible timescales for them, considering the spatial dimensions of our sensor infrastructure and the expected approximate rates of air movement/dilution. The processes and estimates can be seen on *Table 3*. We decided to downsample the higher resolution sensor data to a frequency of 1/min. This sample rate was considered to offer enough resolution without consuming too much data processing resources.

Table 3. Estimates for shortest plausible timescales of some IEQ-influencing processes

Process	Cooking	Sauna	Shower	Fireplace	Washing machines	Smoking	Physical exercise	Candles	Opening windows
Shortest timescale (minutes)	3	20	3	60	30	3	10	30	3

After downsampling the data we proceeded to evaluate the coherency of the time series at buildings level, in order to find good synchronistic data for several houses. This was done by calculating the monthly rates of missing values for each building, all sensors considered, and visually mapping the rates (*r*) at four threshold levels;

(1) $r < 10\ \%$ (dark green)
(2) $10\ \% \geq r < 20\ \%$ (light green)
(3) $r \geq 20 < 100\ \%$ (black)
(4) $r = 100\ \%$ (red)

The mapping is presented on *Table 4.*, and it shows that the amount of missing values turned out to be significant. For filling the gaps we considered a variety of imputation methods described in Junninen et al. (2004). Many gaps in the raw data raw were considered to be too long to fill with reasonable accuracy. The most intact time period turned out to be December 2010 – September 2011, where data for five houses was sufficiently solid, with mostly less than 10 % values missing. The length of the gaps in this target dataset was generally short, so the imputation was performed with straightforward linear interpolation.

Table 4. Evaluation of the rates of missing values in the original dataset. Period of least broken data was found from december 2010 to september 2011, for houses A-E.

Year	Month	House A	House B	House C	House D	House E	House F	House G	House H	House I	House J	House K
2010	5	0.4430	0.8780	0.2879	1.0000	1.0000	0.9189	0.6823	0.7505	0.6909	1.0000	0.7026
2010	6	0.2322	0.6058	0.4743	0.8226	0.0750	0.3442	0.8389	0.1926	0.6314	0.7806	0.9088
2010	7	0.1139	0.9576	0.1008	1.0000	0.0737	0.0970	0.8845	0.9511	0.7341	0.9044	0.9789
2010	8	0.1029	0.9068	0.8665	0.3892	0.0455	0.0901	0.5638	0.5418	1.0000	1.0000	1.0000
2010	9	0.2572	1.0000	0.8355	0.7842	0.4704	0.3396	1.0000	0.9101	1.0000	1.0000	1.0000
2010	10	0.9716	0.5632	0.8239	0.8607	0.0396	0.9982	1.0000	1.0000	0.9957	0.9380	1.0000
2010	11	0.5445	0.5199	0.8664	0.5437	0.0429	1.0000	1.0000	0.5780	0.5958	0.7980	0.8302
2010	12	0.0056	0.0338	0.7262	0.1263	0.0388	0.7285	1.0000	0.0846	0.1692	0.8793	1.0000
2011	1	0.1820	0.0333	0.0092	0.0465	0.0076	0.6497	1.0000	0.1102	0.1993	0.9428	1.0000
2011	2	0.0058	0.0424	0.0077	0.1371	0.0771	0.8354	1.0000	0.1098	0.2063	0.9047	0.9371
2011	3	0.0055	0.2144	0.0060	0.1756	0.1399	0.0271	1.0000	0.3233	0.2770	0.7822	1.0000
2011	4	0.0234	0.1871	0.0121	0.1068	0.0186	0.0233	1.0000	0.9594	1.0000	0.9691	0.6822
2011	5	0.0303	0.1139	0.0767	0.0156	0.0505	0.0275	1.0000	0.9169	1.0000	1.0000	1.0000
2011	6	0.0390	0.0374	0.0112	0.1194	0.0098	1.0000	1.0000	0.0224	1.0000	0.9782	1.0000
2011	7	0.0890	0.1131	0.0175	0.1031	0.0074	1.0000	1.0000	0.1672	1.0000	0.9353	1.0000
2011	8	0.0580	0.0422	0.0368	0.0068	0.0093	1.0000	1.0000	0.1985	1.0000	0.6045	1.0000
2011	9	0.0695	0.2010	0.0057	0.0364	0.0816	1.0000	1.0000	0.3591	1.0000	0.8693	1.0000
2011	10	1.0000	1.0000	0.0667	0.0661	0.1021	1.0000	1.0000	0.2532	1.0000	0.6305	1.0000
2011	11	1.0000	1.0000	0.0478	0.0379	0.0400	1.0000	1.0000	0.1363	1.0000	0.7732	1.0000
2011	12	1.0000	1.0000	0.0109	0.6112	0.0453	1.0000	1.0000	1.0000	1.0000	0.2393	1.0000
2012	1	1.0000	1.0000	0.8001	0.9352	0.5158	1.0000	1.0000	1.0000	1.0000	0.6153	1.0000
2012	2	1.0000	1.0000	0.8278	0.8349	0.0093	1.0000	1.0000	1.0000	1.0000	0.2839	1.0000
2012	3	1.0000	1.0000	0.8279	0.9166	0.0098	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2012	4	1.0000	1.0000	0.8255	1.0000	0.0294	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2012	5	1.0000	1.0000	0.8264	0.8903	0.2654	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2012	6	1.0000	1.0000	0.8267	0.8269	1.0000	1.0000	1.0000	1.0000	0.7835	1.0000	1.0000
2012	7	1.0000	1.0000	0.8252	0.8254	1.0000	1.0000	1.0000	1.0000	0.2072	1.0000	1.0000
2012	8	1.0000	1.0000	0.8253	1.0000	1.0000	1.0000	1.0000	1.0000	0.2564	1.0000	1.0000
2012	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.1310	1.0000	1.0000

For predictive modelling purposes a dataset of several years would have been necessary, but for the descriptive analysis that is the core of this study, this dataset was considered sufficient, because it covers the coldest and hottest part of the year, which usually coincide with the most challenging IEQ situations. Because the target dataset is more than five years old, occupant survey/interview for measuring the subjective experiences of IEQ was not performed in this study. The public weather data was upsampled from 0,1/minute to 1/minute, to even up the resolution with the data from the building sensors. Upsampling was executed by linear interpolation.

Minute-resolution data from both origins, the monitoring system and the weather database was merged into a single MATLAB file including 379632 data points of 152 variables – 8 timestamp variables, 131 sensor observation variables and 13 weather observation variables. APPENDIX 1 describes this target dataset on a table including basic statistics (min, max, mean, standard deviation, kurtosis, skewness) for each variable. Variable ID is identifier for each variable. The statistics were used in the quality screening and general evaluation which guided the selection of variables for the analysis. Besides inspection of the statistics, plots and histograms were drawn on MATLAB to look for obvious anomalies (such as sensor faults) in the data. Column “notes” includes some observations about the statistics and/or plots.

The amount of missing/corrupted data was particularly large for the heating, ventilation and electricity variables. Because of the lack of (clean) data for these variables, the IEQ data became the backbone of the analysis. The hybrid heating system of house E incorporates many special sensors which are used in the automated control of the different heat sources, including water flow, temperature and thermostat parameters in the different parts of the system. This data was left out of the scope of this study, though on a cursory examination the data seemed to be rather clean.

Variable ID on APPENDIX 1 indoor sensors can be read as a combination of two or three strings as follows:

First string is the type of sensor:
Hum = humidity
Tem = temperature
CO2 = carbon dioxide
CO = carbon monoxide.
IVPre = pressure in ventilation duct after the intake fan (Pa)
DifPre = pressure difference over the building envelope (Pa, outdoors-indoors)
PIR = occupancy sensor

The second (complementary) string is the type of room:
OH = living room (finnish ‘olohuone’)
MH = bedroom (finnish ‘makuuhuone’)
PH = shower room (finnish ‘pesuhuone’)
KT = kitchen (finnish ‘keittiö’)
ET = vestibule (finnish ‘eteinen’)

The third string comes after _ and it signifies the house:
13 = house A,

15 = house B
17 = house C
18 = house D
9 or Ast_1_9 = house E

Weekdays are coded 1 = Monday, 2 = Tuesday … 7 = Sunday on the variable 139.

Meteorological variables 140 – 152 are defined as follows:
t2m = temperature
ws_10min = wind speed (10 min avg)
wg_10min = wind gusts (10 min avg)
wd_10min = wind direction (10 min avg)
rh = relative humidity
td = dew point
r_1h = rain (1 hour avg)
r_10min = rain (10 min avg)
snow_aws = snow
p_sea = pressure sea level
vis = visibility
n_man =
wawa = weather description code

Additional 18 variables were created for threshold-based analysis of elevations in CO_2 and CO. In these variables observations from each CO_2 and CO sensor were transformed into discrete values from 1 to 4, where 1 signifies value below the first threshold, and 2, 3 and 4 values exceeding thresholds. For CO_2 the thresholds derived from literature were 1000, 1200 and 1500 ppm, and for CO they were 6,9, 25 and 50 ppm (Ministry of Social Affairs and Health 2003; Satish et al. 2012; WHO 2000). The time of elevations in minutes was calculated simply by counting the data points where discrete value was either 2 OR 3 OR 4 (first threshold exceeded), 3 OR 4 (second threshold) or 4 (third threshold).

3.3. METHODS

The main criteria for the selection of data mining methods in this study, was the need to break down the large target data set into a more concise summary while trying to capture and identify the essential dynamics for each house. In the beginning Microsoft Excel was used to examine the data, but the computing in the Excel environment turned out to consume too much resources considering the system in use (see subchapter 3.3.4.), resulting into software lagging and crashing. However, the more simpler and efficient architecture of MATLAB software could process the big dataset without significant lags, so the main data analysis was decided to be done in that environment.

3.3.1. K-means clustering

The main method of analysis in this study is called clustering. It means using computational methods to assign the observations (in this case made by the sensor network) into groups based on the multivariate combinations in the data (Rencher 2002). In other words we run the data consisting of *n* observations through algorithm in order to partition it into *k* groups (which we call clusters) so that the observations are similar within their own cluster and different compared to other clusters. Therefore our approach can be considered data-driven – our primary driver in the analysis is the data and the dynamics of the houses in it.

The clustering algorithm used is this study originates from MacQueen (1967), and it's called K-means. The algorithm is based on iterative minimization of the sum-of-squares criterion. The procedure of K-means can be described in four steps:

1. Observations are assigned randomly to *k* clusters.
2. The mean values of the clusters define the centroids.
3. Observations are re-assigned to the cluster which centroid is nearest to them, and new centroids computed.
4. Steps 2 and 3 are repeated until the clusters become stationary (or a limit for iteration rounds is met).

The minimization of the sum-of-squares criterion can be expressed by equation:

$$J_{SSE} = \sum_{j=1}^{k} \sum_{i \in S_j} \left\| \boldsymbol{X}_i - c_j \right\|^2$$

Where k is the number of clusters, S_j is a subset of n_j data points, $\boldsymbol{X}_i$ is a vector representing the ith data point and c_j is the centroid of the data points in S_j (Kolehmainen 2014).

3.3.2. Methods for cluster validation

In the evaluation of suitable k, we used Davies-Boulding Index (DB_{nc}) and Silhouette measure. DB_{nc} shows how well the clustering at different k fits the data, in terms of scattering within clusters and the separation between clusters.

DB_{nc} can be formulated with the following functions:

$$DB_{nc} = \frac{1}{n_c} \sum_{i=1}^{n_c} R_i \ , R = \max R_{ij}, i = 1, \ldots, n_c$$

$$R_{ij} = \frac{(s_i + s_j)}{d_{ij}}$$

Where n_c is the number of clusters, R_i is the measure of cluster separation, s_i and s_j are the deviations of C_i and C_j and d_{ij} is the distance between clusters (Davies and Bouldin 1979).

Silhouette measure shows the dissimilarity between all clusters for specific k, and it can be used to look for overlapping clusters where observations can fall into wrong place. Silhouette measure can be defined as:

$$S_k = 1 - \sum_{i}^{N} \frac{h(i) - g(i)}{\max\{g(i), h(i)\}}; i \in C_k$$

Where $g(i)$ is average dissimilarity of vector i compared against all other vectors in the same cluster and $h(i)$ is the minimum of the average dissimilarities of i compared to other clusters. Silhouette measures range from -1 to +1. Negative values indicate that observations might fit better into some other cluster whereas positive values denote a separation from other clusters; the bigger the value, the better the separation (Rousseeuw, 1987).

3.3.3. Sammon's mapping

To visualize the clusters, we apply a dimensionality reduction algorithm called Sammon's mapping. With this technique we can map the p-dimensional data points into 2-dimensional space, while approximately preserving the structure of the data, thereby enabling visual inspection of the basic relative structures between clusters (Sammon Jr, 1969). In our case p is defined by the number of variables used as inputs for the K-means for each house.

As a measure of structure preservation, Sammon's mapping uses following error function:

$$E = \frac{1}{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} d_{ij}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{(d_{ij} - d'_{ij})^2}{d_{ij}}$$

In this function, n is the number of data points, d_{ij} is the (Euclidean) distance between points x_i and x_j in the original space, and d′ is the Euclidean distance between the corresponding points x'_I and x'_j in the 2-dimensional target space (Sammon Jr, 1969).

3.3.4. Hardware and software

The main data processing in this study was done on Matlab R2012a software, which was installed on 64-bit Microsoft Windows 7 Enterprise operating system running on a Lenovo G50 laptop. The computer uses Intel Core i5-4200U processor with 4 GB of RAM. All K-Means and visualization operations were performed using the ctgui 0.71 graphical user interface. Gstool module in the ctgui interface was used to create time-series plots of the cluster distributions. The DB-index runs, where five cluster sets with different random initializations were produced for *k*=1…30, turned out to be most time-consuming of the computing tasks. They took up to two hours of processing time per house.

3.3.5. Number of clusters, input variables and transformation

For each of the houses, all data points were assigned into 19-22 clusters using the K-Means clustering algorithm function in the ctgui 0.71. Input variables for clustering are shown on *Table 5*, where letters A to E signify which variable was used for each house.

Table 5. Input variables for clustering

Parameter	Living room	Bedroom	Kitchen	Bathroom
Temperature	A B C D E	A B D E	C D	A B C D E
Relative humidity	A B C E	A B D E	C D	A B C D E
Carbon dioxide	A B D E	A B D E	C D	
Carbon monoxide	B C D			
Pressure difference	A B D E			

In addition to variables shown on *Table 5*, electricity and water data was included as inputs to clustering for house C. Therefore p values were 9, 10, 10, 12 and 9 for houses A, B, C, D and E respectively.

Several different methods of transformation were used in pre-processing the trial runs of clustering, namely variance scaling (zero mean, standard deviation one), equalization to range [0-1], scaling by vector length and ranking. By evaluating the initial cluster structures and separation through Sammon's mappings and Silhouette's, equalization to range from 0 to 1 was selected as the most suitable methods for pre-processing the data.

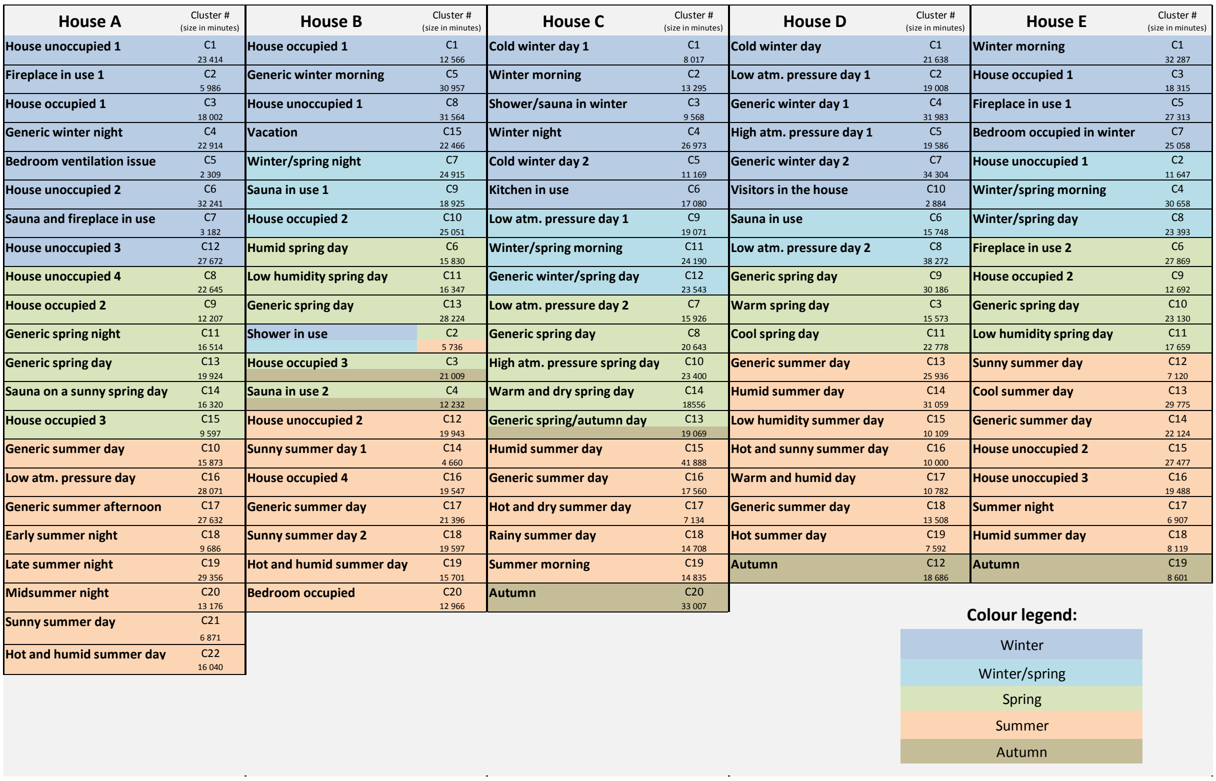
4. RESULTS

The results obtained in this thesis focus on three different aspects of the buildings under study. The first aspect presents a summary of situations defining the clusters for each house. Secondly, the clustered data structures from single building are visualized in more detail, with comparative examples from others. Finally, a univariate analysis of elevations in the levels of carbon oxides is presented. Afterwards, variable levels and probable factors influencing the variations are discussed.

4.1. SUMMARY OF CLUSTERS

All clusters for the houses are presented on *Table 6*.

Table 6. Titles for all clusters, describing interpretations of the data patterns



House A	Cluster # (size in minutes)	House B	Cluster # (size in minutes)	House C	Cluster # (size in minutes)	House D	Cluster # (size in minutes)	House E	Cluster # (size in minutes)
House unoccupied 1	C1 23 414	**House occupied 1**	C1 12 566	**Cold winter day 1**	C1 8 017	**Cold winter day**	C1 21 638	**Winter morning**	C1 32 287
Fireplace in use 1	C2 5 986	**Generic winter morning**	C5 30 957	**Winter morning**	C2 13 295	**Low atm. pressure day 1**	C2 19 008	**House occupied 1**	C3 18 315
House occupied 1	C3 18 002	**House unoccupied 1**	C8 31 564	**Shower/sauna in winter**	C3 9 568	**Generic winter day 1**	C4 31 983	**Fireplace in use 1**	C5 27 313
Generic winter night	C4 22 914	**Vacation**	C15 22 466	**Winter night**	C4 26 973	**High atm. pressure day 1**	C5 19 586	**Bedroom occupied in winter**	C7 25 058
Bedroom ventilation issue	C5 2 309	**Winter/spring night**	C7 24 915	**Cold winter day 2**	C5 11 169	**Generic winter day 2**	C7 34 304	**House unoccupied 1**	C2 11 647
House unoccupied 2	C6 32 241	**Sauna in use 1**	C9 18 925	**Kitchen in use**	C6 17 080	**Visitors in the house**	C10 2 884	**Winter/spring morning**	C4 30 658
Sauna and fireplace in use	C7 3 182	**House occupied 2**	C10 25 051	**Low atm. pressure day 1**	C9 19 071	**Sauna in use**	C6 15 748	**Winter/spring day**	C8 23 393
House unoccupied 3	C12 27 672	**Humid spring day**	C6 15 830	**Winter/spring morning**	C11 24 190	**Low atm. pressure day 2**	C8 38 272	**Fireplace in use 2**	C6 27 869
House unoccupied 4	C8 22 645	**Low humidity spring day**	C11 16 347	**Generic winter/spring day**	C12 23 543	**Generic spring day**	C9 30 186	**House occupied 2**	C9 12 692
House occupied 2	C9 12 207	**Generic spring day**	C13 28 224	**Low atm. pressure day 2**	C7 15 926	**Warm spring day**	C3 15 573	**Generic spring day**	C10 23 130
Generic spring night	C11 16 514	**Shower in use**	C2 5 736	**Generic spring day**	C8 20 643	**Cool spring day**	C11 22 778	**Low humidity spring day**	C11 17 659
Generic spring day	C13 19 924	**House occupied 3**	C3 21 009	**High atm. pressure spring day**	C10 23 400	**Generic summer day**	C13 25 936	**Sunny summer day**	C12 7 120
Sauna on a sunny spring day	C14 16 320	**Sauna in use 2**	C4 12 232	**Warm and dry spring day**	C14 18556	**Humid summer day**	C14 31 059	**Cool summer day**	C13 29 775
House occupied 3	C15 9 597	**House unoccupied 2**	C12 19 943	**Generic spring/autumn day**	C13 19 069	**Low humidity summer day**	C15 10 109	**Generic summer day**	C14 22 124
Generic summer day	C10 15 873	**Sunny summer day 1**	C14 4 660	**Humid summer day**	C15 41 888	**Hot and sunny summer day**	C16 10 000	**House unoccupied 2**	C15 27 477
Low atm. pressure day	C16 28 071	**House occupied 4**	C16 19 547	**Generic summer day**	C16 17 560	**Warm and humid day**	C17 10 782	**House unoccupied 3**	C16 19 488
Generic summer afternoon	C17 27 632	**Generic summer day**	C17 21 396	**Hot and dry summer day**	C17 7 134	**Generic summer day**	C18 13 508	**Summer night**	C17 6 907
Early summer night	C18 9 686	**Sunny summer day 2**	C18 19 597	**Rainy summer day**	C18 14 708	**Hot summer day**	C19 7 592	**Humid summer day**	C18 8 119
Late summer night	C19 29 356	**Hot and humid summer day**	C19 15 701	**Summer morning**	C19 14 835	**Autumn**	C12 18 686	**Autumn**	C19 8 601
Midsummer night	C20 13 176	**Bedroom occupied**	C20 12 966	**Autumn**	C20 33 007				
Sunny summer day	C21 6 871								
Hot and humid summer day	C22 16 040								

Colour legend:

Winter
Winter/spring
Spring
Summer
Autumn

Statistics (minimum, maximum, median, average, standard deviation) for all the variables in each cluster are presented on APPENDIX 2-6, for houses A-E respectively. The averages for some of the key variables are coded with colors/bars to ease the visual recognition of patterns. Each building has its own relative color/bar scale for the House parameters. For T, RH and CO_2, where observations exists for several rooms, the rooms are on a same relative scale, so the colors can be used to visually compare both the patterns between rooms and patterns between clusters in the same room. Legend for color formatting is in APPENDIX 7.

4.2 DETAILED EXAMPLE OF SINGLE HOUSE

For house A, different numbers of clusters (*k*) from 3 to 40 turned out quite steady Davies-Boulding (DB_{nc}) index values around 1,2, which means that the clustering behavior was not distinctly hierarchical for the selected variables. House A DB_{nc} values are shown in *Figure 14*.

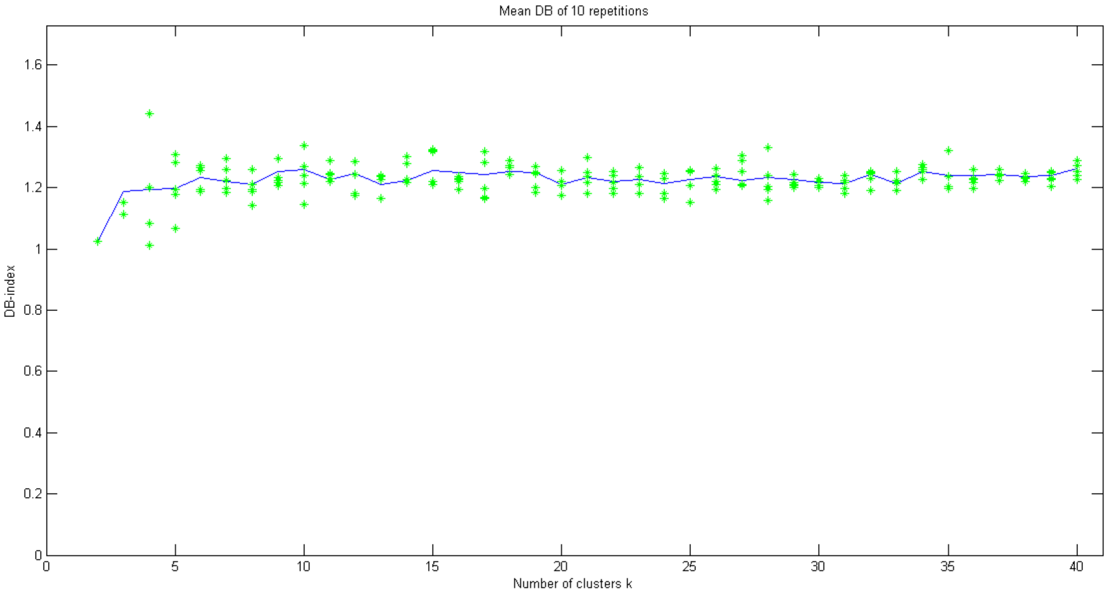


Figure 14. Davies-Boulding index for different *k*-values for house A. Green dots represent the scores for five separate runs. Blue line is the corresponding mean for each *k*.

Because the DB_{nc} didn't suggest any clear optimal value for *k*, different values were tested, and resulting cluster sets evaluated through Silhouettes, Sammon's mappings and cluster statistics. For *k*<15 the basic weather-driven seasonal and circadian dynamics seemed to dominate the clustering process too strongly. On the other hand, analysis for *k*>25 was considered to become too time-consuming in the cluster interpretation step for this study. For house A, *k*=22 was consequently chosen. Sammon's mapping for the clusters of each house are presented in *Figure 15* and *Figure 16*. Silhouettes of the house A clusters are shown in *Figure 17*. In *Figure 18* and *Figure 19* the temporal dynamics of the house A clusters and variables are displayed as time-series plots of the clusters.

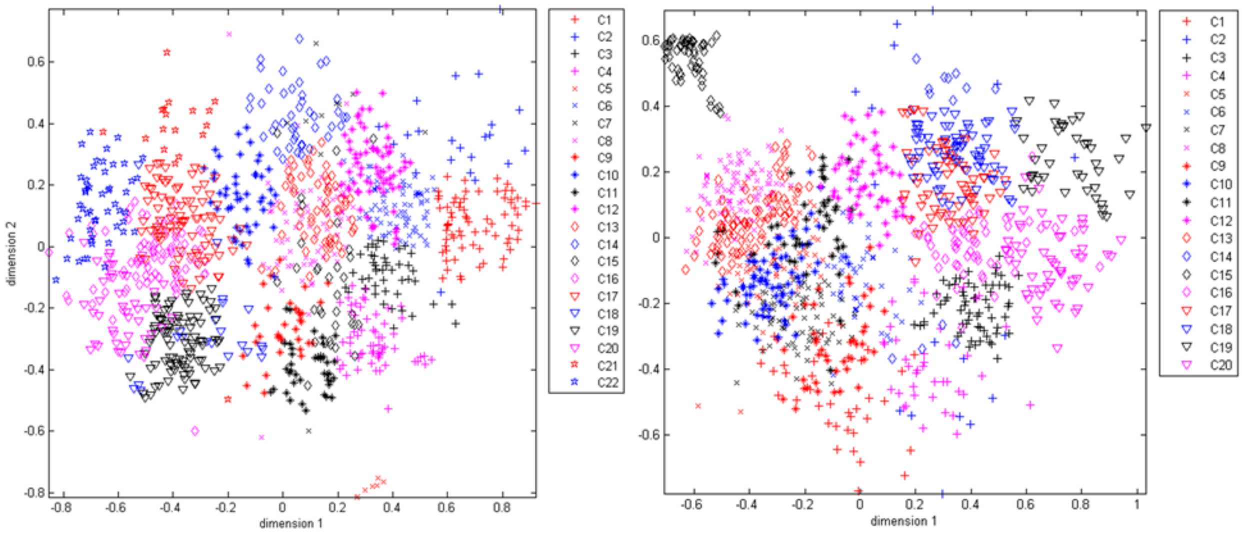


Figure 15. Clusters for houses A and B on Sammon's mapping

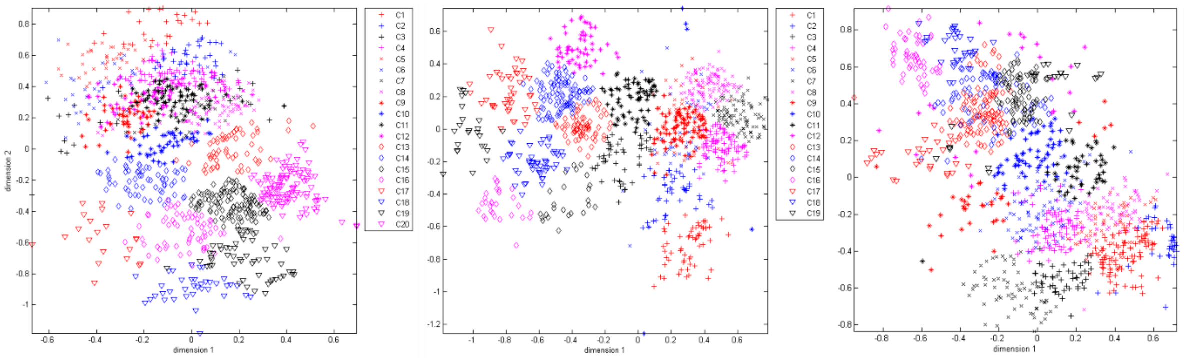


Figure 16. Clusters for house C, D and E on Sammon's mapping

Looking at the Sammon's mappings in *Figure 15* and *Figure 16*, some differences between the multidimensional structures can be noticed.

In the mapping of house A, winter clusters with cool and dry air get positive values on dimension 1 and summer clusters with warm and humid conditions get negative values (note that the axis created by Sammon's mapping are virtual and they do not correspond directly to any of the original variables). On dimension 2 CO_2 dynamics seem to drive the positioning of the clusters, so that nighttime situations with elevated bedroom CO_2 get negative values and empty house situations with low CO_2 get positive values.

In the Sammon's mapping of house B, the 2-dimensional arrangement of clusters is less clear. Dimension 1 corresponds to higher humidity situations on the positive side and lower humidity on the negative side. The cluster positioning on the dimension 2 appers to be driven by CO_2 levels, elevated on the negative side, and lower on the positive side.

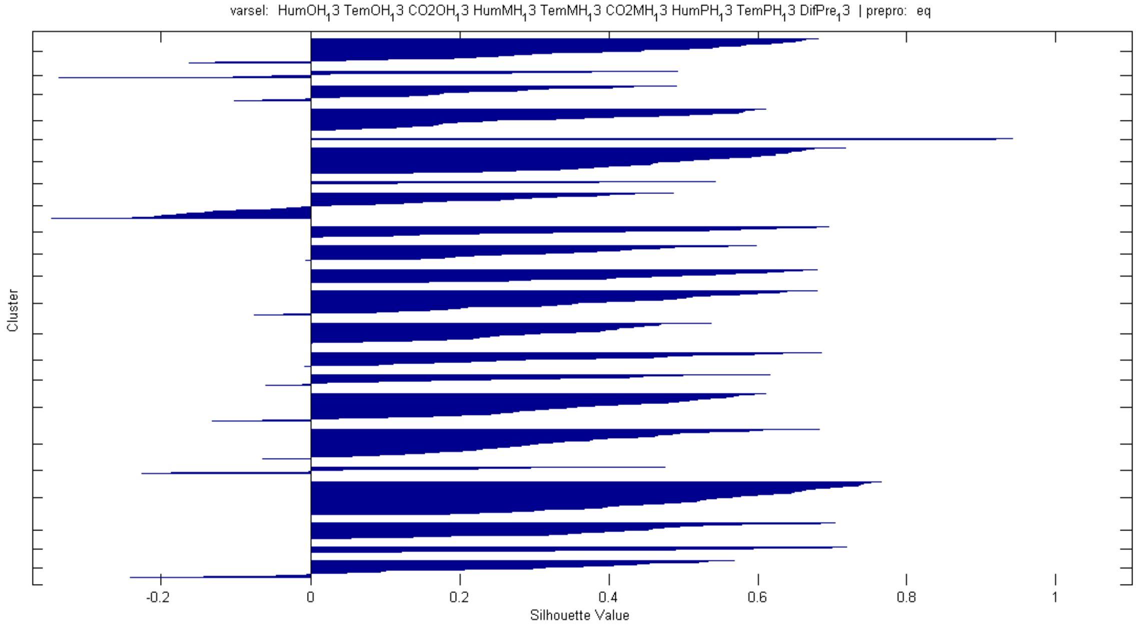


Figure 17. Silhouettes for house A clusters

Silhouettes on *Figure 17* show noticeable but limited overlap between clusters. Only C8 has drawn a significant silhouette on the left hand (negative) side. On the Sammon's mapping it is located in the middle of the 2-dimensional space and temporally the data points are positioned into days of winter and spring.

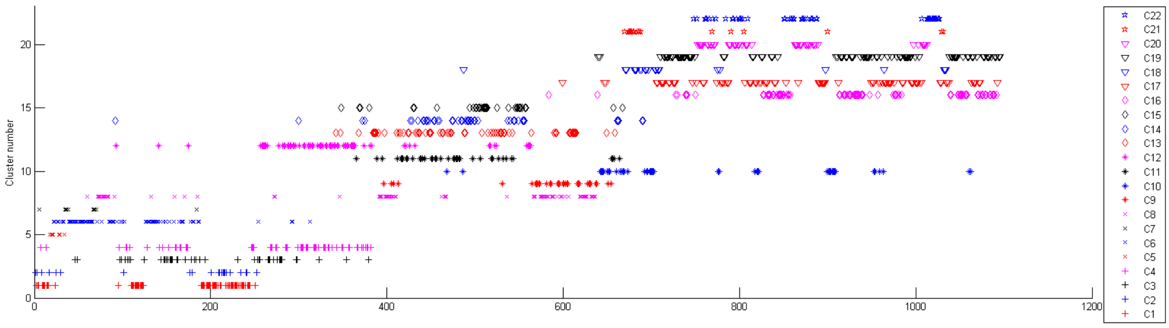


Figure 18. Gstool temporal plot for all house A clusters

The plot in *Figure 18* presents temporal distribution for the clustered data points, showing the occurrence order of each situation on a timeline.

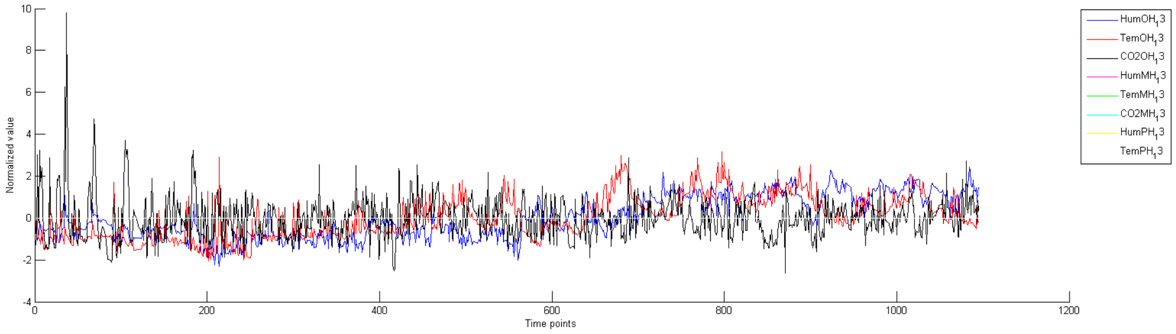


Figure 19. Gstool variable plot for house A

Variable distributions from house A living room are plotted on a normalized scale on *Figure 19*. It shows that the CO_2 spikes in the winter are rather sharp. Both T and RH assume a higher baseline in the summer.

Based on the visualizations and the statistics table on APPENDIX 2, the clusters for house A can be describes as follows:

- Clusters 1, 3, 4, 6 represent generic winter situations with variations in the state of occupation and outdoor temperature.
- Cluster 2 represents the situation where fireplace is in use. Living room T and CO_2 are elevated. Some CO spikes are occurring.
- Cluster 5 describes a situation where ventilation rate for bedroom is too low, and CO_2 gets elevated during the night.
- Cluster 7 represents the use of sauna. Bathroom T and RH are elevated. Living room is occupied during these situations, judging by the slightly elevated CO_2. Some CO spikes fall, into this cluster, so the fireplace is also used.

- Cluster 12 describes a transition from winter to spring. The mean of outdoor temperatures is close to zero and the house is mostly unoccupied.
- Clusters 8 and 9 represent generic spring days. Looking at the CO_2 levels, in the former situation the house is empty and in the latter it's occupied.
- Cluster 10 represents early summer days, where ambient air pressure is somewhat elevated and the indoor air gets warmer.
- Cluster 11 describes a generic spring night.
- Cluster 13 describes a spring day, where living room is occupied.
- Cluster 14 describes a spring day, where ambient air pressure is elevated and humidity is low. Probably sun shines from clear sky and warms up the living room. The bathroom temperature is up, indicating that sauna is used.
- Cluster 15 describes a spring evening, when house is occupied.
- Cluster 16 describes a summer day, where ambient air pressure is low and ambient humidity is up. It's probably cloudy and/or rainy outside. In contrast to clusters 10 and 14, indoor air is less warm and more humid, even though ambient temperature is about the same or higher.
- Clusters 17-20 describe generic situations on a different times of a summer day.
- Clusters 21 and 22 describe hot summer days. Looking at the ambient humidity, the former is a dry and sunny one and the latter is a humid one.

4.3. ELEVATED CONCENTRATIONS OF CARBON OXIDES

The cluster analysis indicated that CO_2 and CO concentrations are sometimes elevated in the studied buildings. All CO sensors were positioned in the living rooms, close to the fireplace, which is the most likely source of CO. For CO_2, the elevated concentrations occurred mostly in bedrooms during night. The elevations were analyzed on three thresholds for both oxides. *Figure 20* and *Figure 22* show the amount of data points where the CO or CO_2 concentrations, respectively, are elevated. Temporal nature of the recurring CO spikes in house B are shown on *Figure 21*.

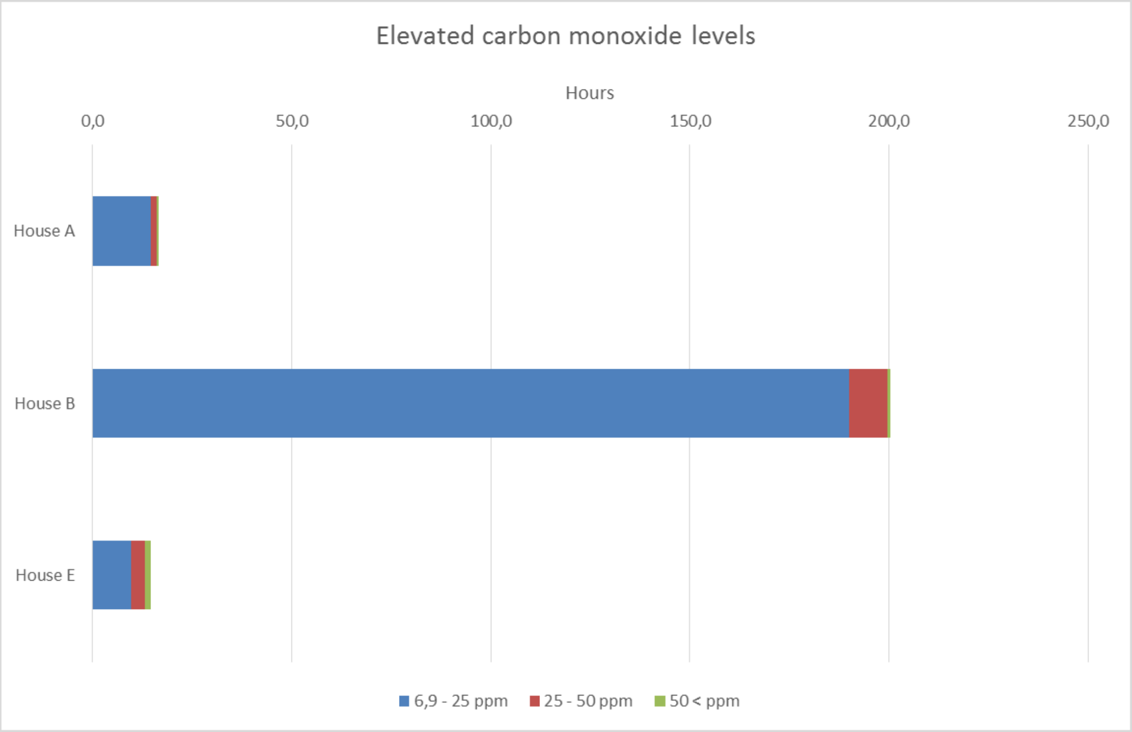


Figure 20. Total times of elevated carbon monoxide concentrations

In house C and D CO values exceeded only the first threshold, for 11 and 37 minutes respectively, so they were left out of the chart in *Figure 20*.

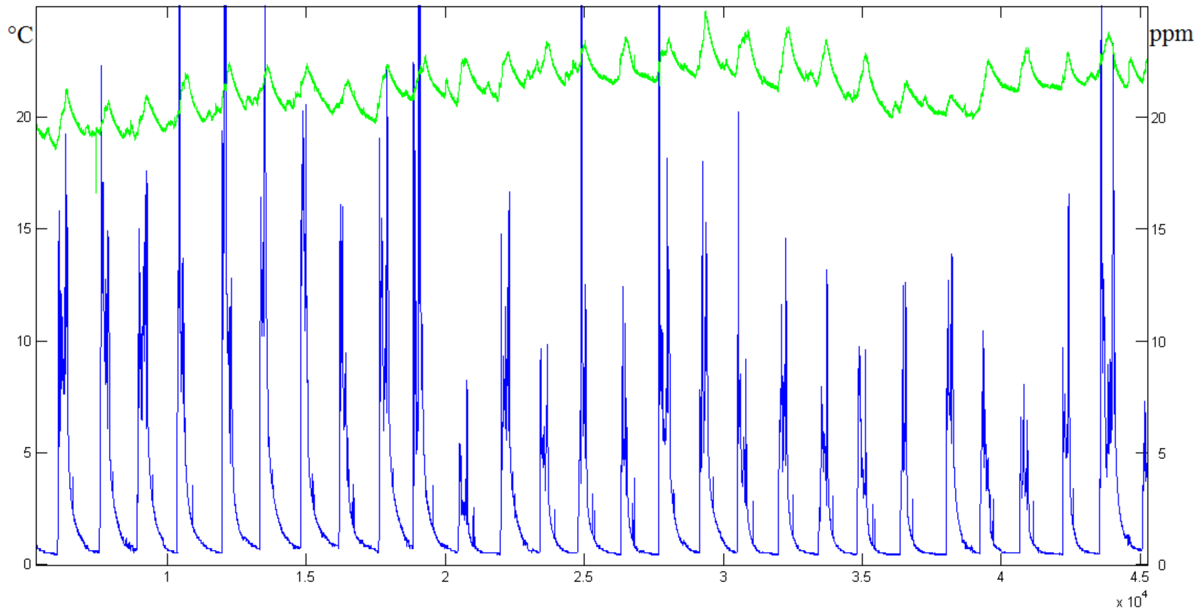


Figure 21. House B CO levels (blue line) and living room temperature (green line), showing the recurring simultaneous occurrence of temperature rise and CO spillage. The scale is numerically same for both variables, signifying °C for temperature and ppm for CO.

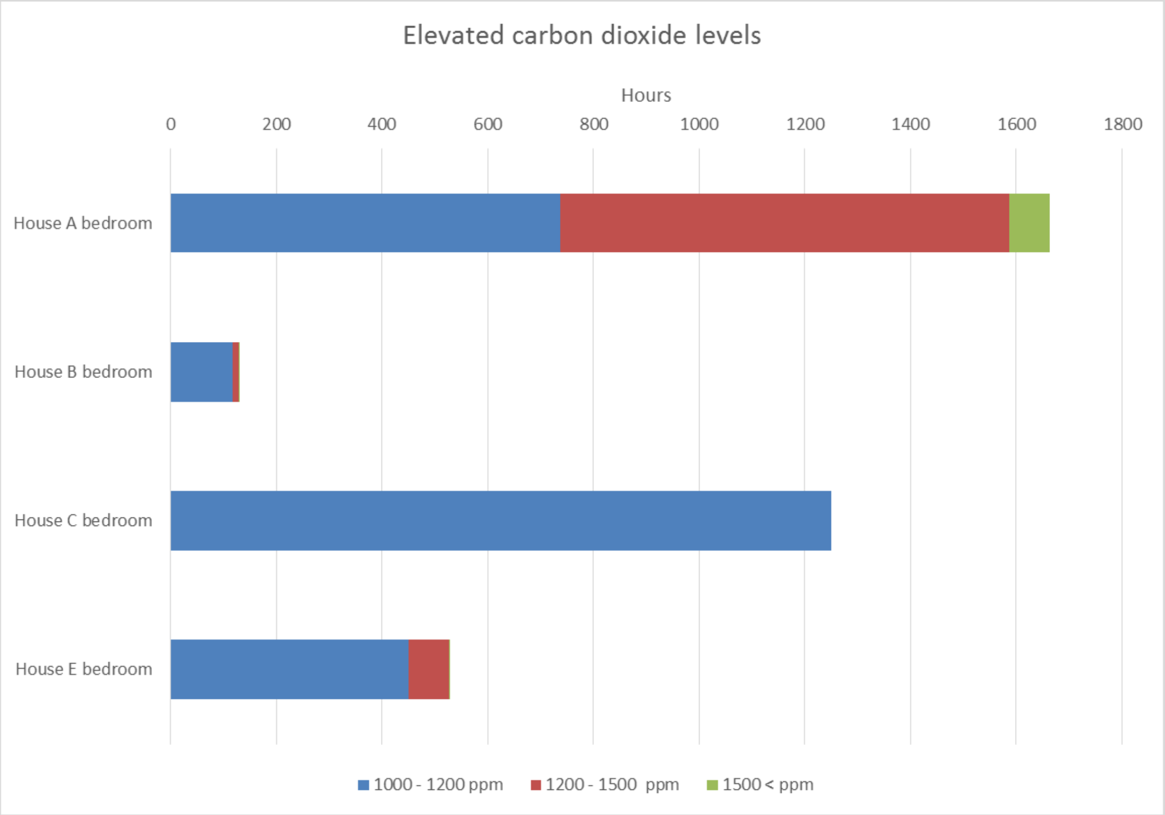


Figure 22. Total times of elevated carbon dioxide concentrations

In house D bedroom CO_2 exceeded 1000 ppm for 30 minutes, and 1200 ppm for 4 minutes, which was considered insignificant and therefore left out of the chart in *Figure 22*.

5. DISCUSSION

The iterative framework of this study yields knowledge about the data in many steps. The steps and the related decisions, insights and problems will be discussed in this chapter.

5.1. DATA SELECTION AND PRE-PROCESSING

The core data for this study originates from the building monitoring system which is described in chapter 3.1.1. The system was developed to provide general information for the occupants concerning the quality of their indoor environment and the energy and water consumption rates. As the data from this particular system has not been previously investigated, and the literature review didn't turn out research reports on these kind of modern detached houses we aimed to produce a holistic picture of the houses, with focus on capturing the key dynamics of the house using the IEQ data through data mining methods.

The selection of our target dataset from the database was done by evaluating the monthly rates of missing values for each house. We found a period of 9 months, where data was clean with mostly less than 10 % missing values. The quality of the database was not investigated on the basis of individual parameters, so (if necessary for future research) it might be possible to acquire wider and/or longer good quality target datasets for some parameters using this raw data.

In this study occupant experiences were not measured because the available sensor data is somewhat dated – since then residents had changed in some houses. Even if same people lived in the house, memorizing the IEQ conditions several years before was considered difficult.

5.2. CLUSTERING

In the clustering process the significance of patterns or parameters is not pre-determined in any explicit way. Therefore the selection of input variables and the pre-processing of data play an important role. To utilize the unsupervised learning properties of the K-means algorithm, offering equal "opportunity" for different phenomena to drive the clustering, we chose to use all available indoor variables and the pressure difference as clustering inputs. The input variables were transformed by equalization into range of 0-1, which was considered to prevent the varying numerical magnitudes of different parameters from influencing the clustering too much.

The Sammon's mappings showed that with the selected input parameters the data was generally scattered quite evenly on 2-dimensional plane. Therefore the distances between neighboring clusters were usually short. The hierarchical structures in the data were not very articulated, which can be seen in the rather flat DB-indexes. Some clusters in the central positions of the vector space overlapped each other significantly, which could confirmed with the Silhouette figures.

Some abnormal situations in the house can be recognized from the Sammon's mapping. These clusters are positioned outside the usual range in the two-dimensional plane, for example cluster five for house A, where bedroom ventilation rate is too low or cluster 15 for house B, where house is unoccupied for a prolonged time.

With K-means clustering input parameters chosen like here, some clusters can become characterized by a distinct change in single room, for example elevation of humidity and temperature in the bathroom caused by shower. These kind of clusters can be dispersed on a rather large area on the Sammon's mapping, for example Cluster 2 of house B.

5.3. GENERAL INFLUENCING FACTORS

Examination of the cluster properties suggests that for all studied houses, three factors seem to account for the main structures in the IEQ data. These factors are (1) weather trends, (2) occupancy patterns and (3) use of energy-intensive household appliances.

5.3.1. Weather trends

The distinctive seasonal difference between the IEQ in the summer and winter was evident in most of the clusters for all houses. For all seasons most of the clusters are organized according to intervals of outdoor air temperature. In winter cold outside temperatures lead to colder and drier indoor air. In spring and summer the effects of outdoor humidity and air pressure seem to increase. One reason for this is probably that these weather variables have a correlation with the amount of clouds blocking the solar heat radiation.

5.3.2. Occupancy patterns

The patterns of CO_2 in the indoor air indicate the relation of ventilation and occupancy rates. Often the occupancy patterns have a circadian nature, such as people leaving the house in the daytime and sleeping in the bedroom during nights. Because the weather changes in circadian patterns, it’s sometimes difficult to estimate whether a cluster is driven by the effects of occupancy or the weather.

5.3.3. Use of energy-intensive household appliances

Clusters indicate that the fireplace, sauna and shower cause significant temporary changes to the indoor environment.

All of these factors are modulated by the characteristics of each house, especially the operation of the HVAC system.

5.4. EVALUATION OF VARIABLE LEVELS

Statistics for clusters provides a way to evaluate the levels of the variables in different situations. In contrast to traditional whole dataset statistics, using the cluster distributions allows simultaneous identification of nonlinear temporal characteristics and relationships to other variables and strip out random outliers. Whether it's reasonable to evaluate conditions using cluster means or other indicators, depends on the variable. In the following subchapters we will look at the IEQ categories present in our dataset and some relevant properties and phenomena. If not otherwise specified, the observations are based on the cluster distributions on APPENDIX 2-6.

The threshold/limit values used in this study are mostly based on the Finnish Housing health instructions from 2003. Some of these guidelines were updated in 2015, which the author unfortunately didn't notice until the research was done. In the update the CO limit was slightly lowered from 6,9 ppm to 7 mg/m3 which is equivalent to 6,11 ppm (at 25°C). For CO2 the action level was reformulated from absolute value of 1500 ppm to 1150 ppm above the outdoor concentration. Acceptable range of thermal conditions was slightly widened with the update.

5.4.1. Thermal conditions

The within-cluster standard deviations of bedroom temperatures are less than 1°C in all the clusters for all houses, so the thermal conditions in the bedrooms have been rather stable. The bathroom and living room temperatures are showing more fluctuations, with standard deviations above 1°C in many cases. Some factors contributing to this are probably:

- Location of intermittent heat sources such as fireplace and sauna
- Orientation of the buildings
- Location of openings such as windows and doors
- Location of thermal mass (for example furniture)
- Control settings of the HVAC system

The average bedroom temperatures for the coldest winter clusters range from 18,1 to 20,8°C. However, in the cooler end of these clusters the bedroom was probably unoccupied because they coincide with low CO_2 levels. For heating season, temperatures from 19 to 23-24°C are considered acceptable in the housing health instructions of 2003. In 2015 the range was widened to 18-26°C. Cluster 1 of house D shows an overheating situation in the winter where whole-house average temperatures are above 25°C and the air is very dry. The houses were almost brand new during the period of examination, which might have two kinds of effects on the thermal performance. On one hand all the HVAC equipment was new and likely to be set up properly during the installation. On the other hand, occupants had only short experience of living in the houses and using the control systems in different situations.

Looking at the summer clusters with peak temperatures, the averages range from 24,8 to 27,4°C in bedrooms and from 26,0 to 28,1°C in the living rooms. The 2003 housing health instructions guideline for maximum indoor temperature is 26°C, though with an added caveat “unless the warming up is caused by hot weather”. The updated instructions define the action level at 32°C. Peak outdoor temperature in the study period was 29,9°C. The summer peak temperatures are mostly affecting the sunny sides of the buildings, indicating a high significance of solar radiation in heating up the indoor air.

5.4.2. Humidity

The driest indoor air in the houses, living rooms and bedrooms considered, occurred in cold winter clusters, with RH averages below 20 % in the wooden houses and below 30 % in the stone houses. The peak humidity in the houses occurred in the summer, with RH from 50,1 to 61,3 %. The 2003 housing health guideline for relative humidity is 20-60%, yet the instructions note that deviations from those levels cannot be considered unhealthy if other indoor conditions are acceptable.

In the summer, weather conditions have a clear effect on the indoor humidity levels. This is apparent for example in house A clusters 21 and 22. Shower/sauna is a potent intermittent source of humidity in the house, with distinctive effects in the bathroom. In other rooms humidity patterns resulting from indoor sources are vague.

5.4.3. Carbon dioxide

As the exhalation of occupants is a regular source for carbon dioxide, it can be used as an occupancy-driven indicator of ventilation adequacy or general air quality. Many clusters reflect the CO_2 variations that are related to different occupancy situations, for example elevated levels in the bedroom during night or low levels in the whole house when unoccupied. Clusters are useful in examining the general CO_2 dynamics, but at least with our input variables, the within-cluster deviations make them somewhat inconvenient for evaluating the potential exposures.

In the housing health instructions CO_2 levels more than 1150 ppm above outdoor concentrations indicate a need to improve ventilation. However, study by Satish et al. (2012) suggests that elevated CO_2 might cause reduced decision-making performance already at 1000 ppm. The temporal extent of elevations was calculated with regards to these three thresholds, 1000, 1200 and 1500 ppm. The ventilation rates in the houses were generally good, with concentrations mostly below 1000 ppm. The higher threshold of 1500 ppm was exceeded sometimes in house A bedroom. Occupants had noticed from stuffiness in the morning that the flow of fresh air was insufficient during these times. The problem was caused by wrong setting of the bedroom air vent. Readjustment of the vent eliminated the problem (Personal communication with the occupant, December 17, 2015).

5.4.4. Carbon monoxide

Carbon monoxide is a toxic gas which is usually produced by incomplete burning of carbonaceous materials (WHO, 2000). The first elevation threshold for CO used in this study is 6,9 ppm. Averages concentrations for all clusters remained below this limit, but looking at the max values in the clusters statistics, it can be seen that there are occasional spikes above 6,9 ppm in all the houses.

In houses B and E the highest measured CO values were 98,2 and 98,0 ppm, respectively. The upper detection limit of the sensor is 99 ppm, therefore it is possible that the actual concentration during these moments was higher than these recorded peak values. Luckily these peak values were only present as single spikes, so the results are not compromised by the sensor range. However, considering that CO has potential for acute poisoning, any indoor monitoring systems with CO sensor should have indication functionality for exceeding some threshold.

In house B the total time of CO sensor readings above 6,9 ppm was 200 hours. *Figure 21* shows that the CO spikes in house B appear simultaneously with a rising temperature pattern and after that fall quite steadily on the baseline level. This suggests that the source for repeated CO spikes in house B was probably too early closing of the fireplace damper. Closing the fireplace damper too early is a common cause for CO elevations in Finnish detached houses (Kokki, 2012). Inhabitants of the house have changed since the study period, so the suspected cause for CO spikes could not be confirmed. The residents are advised to not close the fireplace damper before the last embers have died out. A CO detector indicator is installed in the house (Personal communication with the occupant, December 16, 2015).

Based on the literature of the health effects of CO, the occasional elevations at the levels measured in the investigated houses are not likely to cause any health effects (Raami, 2003; WHO, 2000). Nevertheless, the risk should be recognized and controlled with proper use and maintenance of the fireplace and back-up of functional CO indicators.

5.5. IMPLICATIONS

Based on the analysis of the sensor data, the indoor environment quality has generally been good in all investigated buildings, with only occasional cases of deterioration in certain parameters. Analysis supports the established view that CO_2 is a convenient way to evaluate whether the ventilation rate is adequate with respect to the occupancy status of the house. Some factors which strongly effect on the IEQ, such as noise or lighting were outside the scope of this study. However, considering the location and the architecture of the buildings, it is reasonable to guess that noise and lightning are not very problematic in the investigated houses.

Extending the view of environment quality beyond the indoor space, a paper by Hanski et al. (2012), highlights the significant positive health effects of biodiversity in the homes surroundings. Therefore the vicinity of forests on the site of this study can be claimed to increase environment quality for the whole neighbourhood. The lakeshore gives a distinctive character for the site at least considering aesthetic and microclimatic aspects, but hardly any research has investigated the health or other aspects of living next to lake versus farther away from it.

Identifying opportunities for simultaneous improvement of IEQ and energy efficiency on the basis of our data is difficult, because of the lack of HVAC system data. The design, construction and operation of the houses seems to be of good quality, eliminating obviously wasteful or faulty conditions.

Considering the sensitivity of energy consumption to indoor temperatures in the heating season, lowering room temperatures can be an effective way to save energy. Yet the health effects of room temperatures haven’t been very clearly understood. Experiments by Lee et al. (2014) suggest that cooler sleeping temperatures can cause metabolic benefits by increased formation of brown adipose tissue. Especially if new research offers additional evidence for health benefits, recommendations for lowering room temperatures could become followed more widely. On a related note, avoiding overheating in the winter also helps to prevent dry air. Yet in this context it is necessary to note that sleeping conditions are a highly personal matter, therefore it is desirable that occupants have the possibility to control their own thermal environment.

During summer, excessive peak temperatures could be better prevented by increasing attention on the design of sun shading. Some design methods for improving sun control are presented for example in Prowler and Bourg (2014).

For most parts this thesis has examined the indoor parameters on a rather fine scale, because even the largest variations in the observed data have been relatively close to the comfort zone. From a risk perspective, however, some extreme scenarios should not be ignored because of the severity of potential consequences. Three examples:

- First scenario is the risk of fire getting out of control, which can lead to destruction of entire house. Smoke detector is one good way to reduce the risks of fire, and it’s required by the Finnish law (Ministry of the interior, 2009).
- Second scenario is the risk of carbon monoxide poisoning. Four facts contribute to the risk of CO: (1) toxic properties of the gas, (2) commonness of potential sources, (3) human inability of sensory detection and (4) statistical evidence. At the moment CO

detector indicators are not required by Finnish law, but installing one is a good idea in any household with fireplace or appliances with burning processes, such as a gas stove.

- Third scenario involves a longer term process of adverse microbial growth in the building structures, which can lead to health problems. This problem can be caused by a water damage or moisture control failure, and it builds up in complex interaction with the building materials. Several methods for ensuring moisture control are presented in third chapter of WHO Guidelines for Indoor Air Quality: Dampness and Mould (Seppänen and Kurnitski, 2009).

This thesis has presented an approach for post-hoc assessment of sensor observations in a database. The applicability of these methods to the analysis of real-time data flows is outside the scope of this study, but it should be a good target for future research.

5.6. DIRECTIONS FOR FUTURE RESERCH

The approach of this study could be extended for various applications. Adding a semi-supervised element to the clustering, where acceptable ranges for variables are pre-defined, would be useful for detecting situations where guidelines values are exceeded. In cases where there are several sensor-equipped houses/apartments with identic architectures, the causes for variations could be pointed out with more precision. Inclusion of data from the HVAC system and home appliances would be necessary for gaining more knowledge of the relationships between IEQ and energy. In the ideal setting the indoor sensor data would be accompanied with subjective data describing occupants' experiences and/or health monitoring data. From a practical point of view it's obvious that the focus of research should be directed towards environments where most problems are experienced.

6. SUMMARY AND CONCLUSIONS

This thesis used data mining methods to conduct a post-hoc assessment of the indoor environment quality in five detached houses in Finland. Our database consisted of sensor observations of temperature, relative humidity, carbon dioxide and carbon monoxide from several rooms of the houses. The amount of missing values in the raw data turned out to be relatively big, but a nine month period of coherent synchronistic data was identified and extracted into a target dataset at a frequency of 1/minute.

The target dataset was enhanced with weather variables and partitioned into 19-22 clusters according to the multidimensional structure of the data, using the K-means algorithm. Clusters were considered to represent different IEQ situations in each of the houses. An interpretation for each situation was made by evaluating the clusters variable distributions. Patterns in weather, occupancy and use of household appliances were recognized as main influencing factors, all variables considered. Two-dimensional visual abstraction of the data structures achieved by Sammon's mapping was found to be useful in searching for abnormal situations in the houses. Closer evaluation of variable levels and variations was performed on tables where cluster means for each variable could be easily compared through parameter-specific color-codes. Additionally, temporal distributions of the clusters were inspected on timeline graphs.

The IEQ in the houses was generally good. Occasional cases of variable levels exceeding Finnish guidelines were identified, for example peak temperatures in summer, carbon dioxide levels in a case of ventilation malfunction, carbon monoxide spillage from the fireplace and dry air caused by overheating in winter. Supporting previous research, carbon dioxide concentrations were recognized to provide a good basic indicator of indoor air quality, reflecting the adequacy of ventilation with respect to varying occupancy states.

Sensors have the potential to augment the human sensory system, providing wider awareness of the environment. However, processing the data into situationally appropriate knowledge is a complex challenge. Computational methods, in this case K-means clustering can offer a way to map out multivariate structures in big sensor datasets, which can bring additional holistic insight to the assessment of indoor environment dynamics in buildings.

REFERENCES

Balaras, C.A., Droutsa, K., Argiriou, A.A. and Asimakopoulos, D.N. 2000. EPIQR surveys of apartment buildings in Europe. Energy and Buildings 31: 111-128.

Bluyssen, P.M., 2000. EPIQR and IEQ: indoor environment quality in European apartment buildings. Energy and Buildings 31: 103-110.

Davies, D.L. and Bouldin, D.W. 1979. A cluster separation measure. IEEE Transactions on pattern analysis and machine intelligence, Vol. PAMI-1, No. 2.

Du, L., Prasauskas, T., Leivo, V., Turunen, M., Pekkonen, M., Kiviste, M., Aaltonen, A., Martuzevicius, D. and Haverinen-Shaughnessy, U. 2015. Assessment of indoor environmental quality in existing multi-family buildings in North–East Europe. Environment International 79: 74-84.

Dwyer Instruments. 2009. MS-221 data sheet.

E+E Elektronik. 2014. EE80 data sheet.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. 1996. Knowledge discovery and data mining: towards a unifying framework. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Association for the Advancement of Artificial Intelligence.

Google Maps. 2015. Map view of Kuopio. https://www.google.fi/maps/@62.8786631,27.6482979,12.25z Accessed 29.12.2015.

Hanski, I., von Hertzen, L., Fyhrquist, N., Koskinen, K., Torppa, K., Laatikainen, T., Karisola, P., Auvinen, P., Paulin, L., Mäkelä, M.J., Vartiainen, E., Kosunen, T.U., Alenius, H. and Haahtela, T. 2012. Environmental biodiversity, human microbiota, and allergy are interrelated. PNAS 109 (21): 8334-8339.

HB-Betoniteollisuus. 2012. HB-eristeharkko. suunnittelu- ja työohje. HB-Betoniteollisuus Oy, Jyväskylä.

Hirvonen, J. 2014. Water quality of Lake Taihu: Temporal changes in physical, chemical and biological parameters characterized using computational methods. MSc Thesis. University of Eastern Finland, Faculty of Science and Forestry, Department of Environmental Science

Housing Fair Finland 2010. Web gallery of the housing fair at Kuopio. http://www.asuntomessut.fi/ideagalleria/asunnot Accessed 19.1.2016.

Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen J. and Kolehmainen M. 2004. Methods for imputation of missing values in air quality data sets. Atmospheric Environment 38: 2895-2907.

Kaksonen, P. 2012. Omakotitalojen ilmanvaihdon kunto, huolto ja vaikutus sisäilmaan. Publications of University of Eastern Finland. http://www2.uef.fi/documents/976466/1020718/Kaksonen_Pekka.pdf/9f7c2156-fd63-41ad-81ca-e1f41d594d53 Accessed 16.1.2016. (Reference available only in Finnish.)

Kim, H., Stumpf, A. and Kim, W. 2011. Analysis of an energy efficient building design through data mining approach. Automation in Construction 20: 37-43.

Kokki, E. 2011. Häkä asuintaloissa 2011. Research publications of The Emergency Services College. http://www.nuohoojat.fi/binary/file/-/id/16/fid/89/ Accessed 16.1.2016 (Reference available only in Finnish.)

Kolehmainen, M. 2014. Lecture notes on Environmental Data Mining. Environmental Informatics. University of Eastern Finland.

Kuopio City online Guide Map. 2016. https://karttapalvelu.kuopio.fi/ Accessed 16.1.2016

Lai, A.C.K., Mui, K.W., Wong, L.T. and Law, L.Y. 2009. An evaluation model for indoor environmental quality (IEQ) acceptance in residential buildings. Energy and Buildings 41: 930-936.

Lee, P., Smith, S., Linderman, J., Courville A.B., Brychta, R.J., Dieckmann, W., Werner, C.D., Chen, K.Y. and Celi, F.S. 2014. Temperature-acclimated brown adipose tissue modulates insulin sensitivity in humans. Diabetes 63 (11): 3686-3698.

Li, K., Hu, C., Liu, G. and Xue, W. 2015. Buildings electricity consumption prediction using optimized artificial neural networks and principal component analysis. Energy and Buildings 108: 106-113.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, 281--297, University of California Press, Berkeley.

Ministry of Social Affairs and Health. 2003. Asumisterveysohje. Sosiaali- ja terveysministeriön oppaita 2003. (Reference available only in Finnish and Swedish.)

Ministry of Social Affairs and Health. 2015. Sosiaali- ja terveysministeriön asetus asunnon ja muun oleskelutilan terveydellisistä olosuhteista sekä ulkopuolisten asiantuntijoiden pätevyysvaatimuksista 545/2015. (Reference available only in Finnish and Swedish.)

Ministry of the Environment. 2007. Ympäristöministeriön asetus rakennuksen energiatodistuksesta 765/2007. (Reference available only in Finnish and Swedish.)

Ministry of the Environment 2013. Rakennuksen energiatodistus uudistuu. Presentation. http://www.ym.fi/download/noname/%7BC6C50127-8803-438F-974A-67A5BF3CEEF5%7D/31596 Accessed 16.1.2016.

Ministry of the Interior. 2009. Sisäasiainministeriön asetus palovaroittimien sijoittamisesta ja kunnossapidosta 239/2009. (Reference available only in Finnish and Swedish.)

Mitchell, C.S., Zhang, J., Sigsgaard, T., Jantunen, M., Lioy, P.J., Samson, R. and Karol, M.H. 2007. Current state of the science: health effects and indoor environmental quality. Environmental Health Perspectives 115 (6): 958-964.

Nicol, J.F. and Wilson, M. 2011. A critique of European Standard EN 15251: strengths, weaknesses and lessons for future standards. Building Research & Information 39: 183-93.

Nieminen, J. 2007. Passiivitalo. VTT Technical Research Centre of Finland Ltd. http://passiivitalo.vtt.fi/files/passiivitalon%20maaritelma.pdf Accessed 16.1.2016.

Niska, H. 2012. Predictive Data-Driven Modeling Approaches in Environmental Management Decision-Making. Publications of The University of Eastern Finland. Dissertations in Forestry and Natural Sciences No 60

Pallasmaa, J. 2011. Place, Atmosphere and Architecture -perception, memory and imagination. Inaugural Kenneth Frampton Endowed Lecture. Columbia University GSAPP. https://www.youtube.com/watch?v=MFwrmIljdqo Accessed 26.2.2016.

Peel, M.C., Finlayson, B.L. and McMahon, T.A. 2007. Updated world map of the Köppen-Geiger climate classification. Hydrology and Earth System Sciences 11: 1633-1644

Pekkonen, M., Du, L., Skön, J-P., Raatikainen, M. and Haverinen-Shaughnessy, U. 2015. The influence of tenure status on housing satisfaction and indoor environmental quality in Finnish apartment buildings. Building and Environment 89: 134-140.

Pirinen, A. 2014. Dwelling as a product: perspectives on housing, users and the expansion of design. Aalto University. School of Arts, Design and Architecture. Doctoral dissertations 12/2014. Helsinki.

Pirinen, P., Simola, H., Aalto, J., Kaukoranta, J-P., Karlsson, P. and Ruuhela, R. 2012. Tilastoja Suomen ilmastosta 1981–2010. Reports of the Finnish Meteorological Institute 2012:1.

Prowler, D. and Bourg, J. 2014. Sun Control and Shading Devices. Whole Building Design Guide. National Institute of Building Sciences. https://www.wbdg.org/resources/suncontrol.php Accessed 5.2.2016.

Raami, J. 2003. Hiilimonoksidivaroittimien toiminta ja vaatimukset. Espoo-Vantaa Institute of Technology. http://www.tukes.fi/Tiedostot/julkaisut/Hiilimonoksidivaroittimien%20toiminta%20ja%20vaatimukset.pdf Accessed 5.2.2016 (Reference available only in Finnish.)

Raatikainen, M., Skön, J-P., Johansson, M., Leiviskä, K. and Kolehmainen, M. 2012. Effects of Weather Conditions and Differential Air Pressure on Indoor Air Quality. World Academy of Science, Engineering and Technology 61: 2093-2099.

Ramesh, T., Prakash, R. and Shukla, K.K. 2010. Life cycle energy analysis of buildings: An overview. Energy and Buildings 42: 1592-1600.

Rencher, A.C. 2002. Methods of multivariate analysis. 2nd Edition. A Wiley-Interscience publication, New York.

Roulet, C-A. 2001. Indoor environment quality in buildings and its impact on outdoor environment. Energy and Buildings 33: 183-191.

Rousseeuw, P.J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 20: 53-65.

Salmi, T. 2012. FMI Weather data downloader. http://tumetsu.github.io/FMI-weather-downloader/ Accessed 19.1.2016

Sammon, J.W, Jr. 1969. A nonlinear mapping for data structure analysis. IEEE Transactions on Computers. 48 (5): 401-409.

Satish, U., Mendell, M.J., Shekhar, K., Hotchi, T., Sullivan, D., Streufert, S. and Fisk, W.J. 2012. Is CO_2 an Indoor Pollutant? Direct Effects of Low-to-Moderate CO_2 Concentrations on Human Decision-Making Performance. Environmental Health Perspectives 120 (12)

Seppänen, O. and Kurnitski, J. 2009. Moisture control and ventilation. In: WHO Guidelines for Indoor Air Quality: Dampness and Mould. World Health Organization, Geneva. http://www.ncbi.nlm.nih.gov/books/NBK143947/ Accessed 6.2.2016.

Sharmin, T., Gül, M., Li, X., Ganev, V., Nikolaidis, I. and Al-Hussein, M. 2014. Monitoring building energy consumption, thermal performance, and indoor air quality in a cold climate - region. Sustainable Cities and Society 13: 57-68.

Sharpe, R.A., Thornton, C.R., Nikolaou, V. and Osborne, N.J. 2015. Higher energy efficient homes are associated with increased risk of doctor diagnosed asthma in a UK subpopulation. Environment International 75: 234-244.

Shove, E. 2003. Converging Conventions of Comfort, Cleanliness and Convenience. Journal of Consumer Policy 26: 395-418.

Skön, J-P., Kauhanen, O. and Kolehmainen, M. 2011. Energy Consumption and Air Quality Monitoring System. In: D. Nandagopal, M. Palaniswami, J. Gubbi (eds.). Proceedings of the Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing. 163-167.

Stocker, M., Rönkkö, M. and Kolehmainen, M. 2012. Making Sense of Sensor Data Using Ontology: A Discussion for Residential Building Monitoring. IFIP Advances in Information and Communication Technology 382: 341-350.

Straube, J. 2006. The Building Enclosure. Building Science Digests 018. Building Science Corporation. http://buildingscience.com/documents/digests/bsd-018-the-building-enclosure_revised Accessed 12.2.2016.

Säteri, J. 2015. The Finnish Classification of Indoor Environment and Material Emissions. Presentation of Finnish Society of Indoor Air Quality and Climate. Finnish Green Building Council. http://figbc.fi/wp-content/uploads/2015/05/Sateri-Finnish-classification-on-indoor-environment-2008.pdf Accessed 8.2.2016.

Teixeira Chaves, D.M. 2012. Housing in hostile places – Controlling adverse weather through architectural design in extreme climates: cold, hot/arid and tropical. Extended abstract for Master's degree dissertation in architecture. Universidade Técnica de Lisboa. https://fenix.tecnico.ulisboa.pt/downloadFile/395144322455/Housing%20in%20Hostile%20Places%20-%20Extended%20Abstract.pdf Accessed 16.1.2016.

Tongdy Control Technology. 2012. F2000TSM-CO data sheet.

Vinha, J., Korpi, M., Kalamees, T., Eskola, L., Palonen, J., Kurnitski, J., Valovirta, I., Mikkilä, A. and Jokisalo, J. 2005. Puurunkoisten pientalojen kosteus- ja lämpötilaolosuhteet, ilmanvaihto ja ilmatiiviys. Tampere University of Technology, Department of Civil Engineering, Structural Engineering Laboratory Research report 131. (Reference available only in Finnish.)

WHO. 2000. Air Quality Guidelines - Second Edition, Chapter 5.5 Carbon monoxide. WHO Regional Office for Europe, Copenhagen, Denmark

Wu, S. and Clements-Croome, D. 2007. Understanding the indoor environment through mining sensory data—A case study. Energy and Buildings 39: 1183-1191.

Xiao, F. and Fan, C. 2014. Data mining in building automation system for improving building operational performance. Energy and Buildings 75: 109-118.

Yu, Z., Fung, B.C.M., Haghighat, F., Yoshino, H. and Morofsky, E. 2011. A systematic procedure to study the influence of occupant behavior on building energy consumption. Energy and Buildings 43: 1409-1417.

Yu, Z., Fung, B.C.M. and Haghighat, F. 2013. Extracting knowledge from building-related data — A data mining framework. BUILD SIMUL 6: 207-222.

APPENDIX 1. Statistics for all variables used in the initial quality screening and the selection of variables for the analysis

Var. #	**Variable ID**	**min**	**max**	**mean**	**std**	**kurtosis**	**skewness**	**notes**
1	Years'	2010,0	2011,0	2011,0	0,2	28,6	-5,3	ok
2	'Months'	1,0	12,0	5,0	2,7	2,6	0,4	ok
3	'Days'	1,0	31,0	15,7	9,0	1,8	0,0	ok
4	'Hours'	0,0	23,0	11,5	6,9	1,8	0,0	ok
5	'Minutes'	0,0	59,0	29,5	17,3	1,8	0,0	ok
6	'HumOH_13'	11,5	62,2	34,8	10,0	2,0	0,4	ok
7	'TemOH_13'	18,5	30,5	23,2	2,1	2,6	0,5	ok
8	'CO2OH_13'	135,4	1960,8	587,3	118,7	19,8	2,2	ok
9	'HumMH_13'	15,0	68,7	36,7	10,7	1,9	0,4	ok
10	'TemMH_13'	16,1	29,5	22,2	2,1	2,7	-0,3	ok
11	'CO2MH_13'	284,7	1957,7	762,2	324,7	2,8	0,8	ok
12	'HumPH_13'	9,6	95,4	28,7	13,1	2,7	0,7	ok
13	'TemPH_13'	12,4	40,9	24,5	3,7	3,9	-0,6	ok
14	'COOH_13'	0,7	80,5	0,9	1,0	1662,5	32,8	ok
15	'IVPre_13'	0,1	4,8	0,9	0,5	7,0	1,2	ok
16	'DifPre_13'	-48,8	48,0	4,1	5,7	11,1	2,1	ok
17	'PIROH_13'	0,0	0,0	0,0	0,0	NaN	NaN	all zeroes
18	'Wat_13'	0,0	0,0	0,0	0,0	NaN	NaN	all zeroes
19	'Ele_13'	0,4	2227,8	154,4	129,2	55,7	4,4	not used
20	'Hea_13'	8,1	9000090,0	17679,6	131655,9	1911,6	34,9	not used
21	'Wat_9'	0,0	0,0	0,0	0,0	335,8	16,1	not used
22	'Hea_9'	0,0	0,0	0,0	0,0	NaN	NaN	only zeroes
23	'Ele_9'	4,0	2940,0	192,3	197,9	3,5	0,9	not used
24	'HumKT_15'	0,3	50,8	27,7	7,2	2,2	0,5	single small outlier
25	'TemKT_15'	0,1	0,7	0,1	0,0	19,2	3,8	sensor fault
26	'TemMH_15'	18,6	26,1	22,6	1,3	4,5	-0,7	ok
27	'HumMH_15'	17,9	72,1	36,9	11,5	2,2	0,6	ok
28	'CO2MH_15'	424,5	1555,7	670,0	159,3	2,5	0,5	ok
29	'CO2KT_15'	5,4	1957,4	569,3	130,3	3,1	0,2	single small outlier
30	'HumPH_15'	13,9	97,6	38,6	14,6	3,5	0,9	ok
31	'TemPH_15'	20,4	31,9	24,3	1,5	5,7	0,4	ok
32	'COOH_15'	0,4	98,2	1,2	2,5	102,2	7,4	ok
33	'IVPre_15'	0,2	25,8	1,3	2,3	14,7	2,7	ok
34	'DifPre_15'	32,3	48,3	48,2	0,4	169,9	-10,5	sensor fault
35	'TemOH_15'	16,6	30,0	23,2	2,1	2,6	-0,1	ok
36	'HumOH_15'	19,9	63,4	36,4	9,2	2,1	0,5	ok
37	'CO2OH_15'	408,8	1477,5	666,2	135,5	2,7	0,1	ok
38	'Wat_15'	0,0	0,0	NaN	NaN	291,2	14,7	not used
39	'Ele_15'	0,0	0,0	0,0	0,0	NaN	NaN	only zeroes
40	'Hea_15'	0,0	0,0	0,0	0,0	NaN	NaN	only zeroes

41	'TemKT_18'	18,4	28,7	23,4	1,6	3,1	0,8	ok
42	'HumKT_18'	3,0	66,6	29,4	13,8	2,0	0,4	ok
43	'HumMH_18'	2,5	71,3	29,7	14,5	2,2	0,5	ok
44	'TemMH_18'	18,7	28,5	22,6	1,9	2,8	0,6	ok
45	'CO2KT_18'	387,3	1225,4	495,9	65,8	14,7	2,4	ok
46	'IVPre_18'	0,1	3,5	1,1	0,5	2,5	0,0	ok
47	'HumPH_18'	3,0	90,3	30,1	14,6	1,9	0,3	ok
48	'CO2MH_18'	452,7	1397,0	563,6	59,3	5,2	1,2	ok
49	'TemPH_18'	18,3	32,3	23,3	1,8	3,1	0,7	ok
50	'COOH_18'	0,4	12,8	0,5	0,2	341,8	12,7	ok
51	'DifPre_18'	-35,6	30,9	2,7	4,2	10,4	1,6	ok
52	'TemOH_18'	18,8	29,1	23,6	1,7	2,9	0,7	ok
53	'CO2OH_18'	379,2	1482,6	486,4	70,1	26,8	3,3	ok
54	'HumOH_18'	3,8	64,1	29,7	13,3	2,0	0,4	ok
55	'Wat_18'	0,0	0,0	0,0	0,0	258,8	13,9	not used
56	'Ele_18'	0,1	13,2	1,0	1,2	31,3	4,7	not used
57	'Hea_18'	7,4	60240,0	2298,9	2962,9	105,9	7,1	not used
58	'TemKT_17'	22,0	27,7	24,0	0,9	3,5	0,9	ok
59	'HumKT_17'	17,7	66,7	32,7	10,9	2,3	0,7	ok
60	'HumMH_17'	16,1	61,6	31,8	9,8	2,3	0,6	ok
61	'TemMH_17'	9,9	49,0	17,2	6,3	5,3	1,5	range too wide
62	'IVPre_17'	0,2	24,7	6,3	6,3	1,2	0,1	ok
63	'CO2KT_17'	271,3	1068,5	501,0	106,4	4,2	1,1	min value too low
64	'CO2MH_17'	786,5	1148,4	938,3	67,0	2,3	0,3	range not plausible
65	'HumPH_17'	17,3	95,9	35,0	10,8	2,6	0,5	ok
66	'COOH_17'	0,7	11,4	1,0	0,3	48,6	3,9	ok
67	'TemPH_17'	22,1	28,8	24,6	1,1	3,2	0,6	ok
68	'DifPre_17'	-46,1	48,4	6,8	4,4	14,9	0,2	ok
69	'HumOH_17'	12,0	69,0	32,1	12,9	1,9	0,4	ok
70	'CO2OH_17'	6,7	938,3	186,6	114,8	4,8	1,3	range too low
71	'TemOH_17'	19,1	27,1	23,2	1,2	3,1	0,4	ok
72	'Wat_17'	0,0	1,2	0,0	0,0	231,9	11,9	ok
73	'Ele_17'	9,1	2160,0	343,9	284,3	3,8	1,0	ok
74	'Hea_17'	0,0	0,0	0,0	0,0	NaN	NaN	only zeroes
75	'Fir_3_Ast_1_9'	25,5	60,8	33,7	5,3	5,0	1,6	not used
76	'Fir_4_Ast_1_9'	28,3	58,6	36,4	4,7	5,0	1,5	not used
77	'Fir_5_Ast_1_9'	25,2	49,8	32,8	4,2	4,2	1,4	not used
78	'Fir_6_Ast_1_9'	24,5	44,8	31,5	3,4	3,8	1,2	not used
79	'Fir_7_Ast_1_9'	25,3	60,0	32,8	4,8	6,1	1,8	not used
80	'Fir_8_Ast_1_9'	25,3	59,7	32,7	4,8	6,1	1,8	not used
81	'Fir_9_Ast_1_9'	25,2	58,0	33,0	4,6	5,4	1,6	not used
82	'Fir_10_Ast_1_9'	25,1	60,4	33,0	4,9	5,8	1,7	not used
83	'Fir_11_Ast_1_9'	25,3	54,9	33,0	4,4	4,8	1,5	not used

84	'Fir_12_Ast_1_9'	25,1	52,6	32,5	4,2	4,7	1,4	not used
85	'Fir_13_Ast_1_9'	25,6	53,2	33,8	5,0	3,9	1,3	not used
86	'Fir_14_Ast_1_9'	25,1	54,0	33,4	5,0	4,0	1,3	not used
87	'Fir_15_Ast_1_9'	25,4	50,4	33,3	4,4	3,9	1,3	not used
88	'Fir_16_Ast_1_9'	25,2	52,8	33,3	4,7	4,1	1,3	not used
89	'Col_Wat_Sta_te_Ast_1...'	22,0	38,8	33,5	2,2	2,9	-0,3	not used
90	'Col_Wat_Sta_fl_Ast_1...'	8,0	8,2	8,0	0,0	3,9	-0,9	not used
91	'War_Wat_Sta_te_Ast_1...'	32,9	64,3	37,9	3,4	7,0	1,8	not used
92	'War_Wat_Sta_fl_Ast_1...'	7,9	9,0	8,0	0,0	60,5	0,7	not used
93	'Fir_Wat_te_Ast_1_10...'	28,5	96,8	43,9	11,4	5,1	1,7	not used
94	'Fir_Wat_fl_Ast_1_9'	1,2	9,0	2,9	2,0	4,7	1,5	not used
95	'Sun_Wat_te_Ast_1...'	14,0	71,2	41,8	7,2	3,4	1,1	not used
96	'Sun_Wat_fl_Ast_1_9'	1,2	9,1	2,2	1,7	8,0	2,2	not used
97	'Col_Wat_End_te_Ast_1...'	29,3	58,9	33,9	1,5	30,1	3,5	not used
98	'Col_Wat_End_fl_Ast_1...'	4,0	9,1	8,0	0,5	56,1	-7,2	not used
99	'War_Wat_End_te_Ast_1...'	21,8	38,1	34,1	1,5	7,8	-0,9	not used
100	'War_Wat_End_fl_Ast_1...'	3,9	8,7	7,7	1,1	10,0	-2,9	not used
101	'Und_he_PT10_Ast_1_9'	25,8	41,4	31,9	2,6	2,4	0,5	not used
102	'Sun_Fir_PT10_Ast_1_9'	25,5	78,4	40,9	10,1	4,0	1,3	not used
103	'CO2_OH_Ast_1_9'	346,3	1408,5	576,5	150,5	4,1	1,0	ok
104	'Hum_OH_Ast_1_9'	7,2	60,5	29,6	13,4	1,8	0,4	ok
105	'Tem_OH_Ast_1_9'	20,4	28,3	23,9	1,4	3,3	0,5	ok
106	'CO2_MH_Ast_1_9'	383,7	1568,3	658,7	198,4	3,6	1,0	ok
107	'Hum_MH_Ast_1_9'	7,6	64,1	30,5	13,7	1,9	0,4	ok
108	'Tem_MH_Ast_1_9'	19,7	28,6	23,5	1,6	3,2	0,6	ok
109	'CO_OH_Ast_1_9'	0,1	98,0	0,4	1,5	1405,9	33,5	ok
110	'Dif_Pre_Ast_1_9'	-20,6	48,0	0,3	3,9	36,0	3,7	ok
111	'Wat_fl_Ast_1_1_9'	1,9	11,2	2,1	0,6	19,2	3,8	not used
112	'Wat_fl_Ast_1_2_9'	1,9	2,0	1,9	0,0	18,6	-0,2	not used
113	'Wat_fl_Ast_1_3_9'	1,9	11,2	2,0	0,5	42,0	5,8	not used
114	'Wat_fl_Ast_1_4_9'	1,9	1,9	1,9	0,0	4,7	-1,2	not used
115	'Wat_fl_Ast_1_5_9'	0,0	0,0	0,0	0,0	3833,6	31,8	not used
116	'Wat_fl_Ast_1_6_9'	1,9	1,9	1,9	0,0	4,4	-1,2	not used
117	'Wat_fl_Ast_1_7_9'	1,9	1,9	1,9	0,0	3,9	-1,1	not used
118	'Wat_fl_Ast_1_8_9'	1,9	1,9	1,9	0,0	4,6	-1,2	not used
119	'Wat_fl_Ast_1_9_9'	2,0	11,4	3,7	2,7	4,0	1,5	not used
120	'Wat_fl_Ast_1_10_9'	2,0	2,2	2,0	0,0	4,0	1,3	not used
121	'Wat_te_Ast_1_1_9'	34,0	42,8	37,4	1,9	2,1	0,5	not used
122	'Wat_te_Ast_1_2_9'	34,2	40,4	37,1	1,1	2,3	0,4	not used
123	'Wat_te_Ast_1_3_9'	33,9	44,5	37,7	2,3	2,2	0,6	not used
124	'Wat_te_Ast_1_4_9'	33,6	42,7	37,1	2,0	2,1	0,5	not used
125	'Wat_te_Ast_1_5_9'	34,5	41,3	37,6	1,2	2,4	0,4	not used
126	'Wat_te_Ast_1_6_9'	33,7	43,1	37,3	2,1	2,1	0,5	not used

127	'Wat_te_Ast_1_7_9'	33,7	41,4	36,9	1,6	2,0	0,4	not used
128	'Wat_te_Ast_1_8_9'	34,0	40,9	37,1	1,5	2,0	0,4	not used
129	'Wat_te_Ast_1_9_9'	35,4	41,4	38,2	1,2	2,1	0,3	not used
130	'Wat_te_Ast_1_10_9'	34,1	42,1	38,5	1,0	3,1	0,3	not used
131	'Hum_PH_Ast_1_9'	5,7	94,1	34,0	14,9	1,9	0,3	ok
132	'Tem_PH_Ast_1_9'	20,3	28,4	23,1	1,1	4,4	1,0	ok
133	'PIR_OH_Ast_1_9'	0,0	0,0	0,0	0,0	NaN	NaN	only zeroes
134	'PIR_ET_Ast_1_9'	0,0	1,0	0,9	0,2	18,4	-3,8	not used
135	'Fir_1_Ast_1_9'	25,5	61,7	33,0	5,0	6,1	1,8	not used
136	'Fir_2_Ast_1_9'	25,0	55,5	32,1	4,4	5,8	1,8	not used
137	'Hours_sin'	-1,0	1,0	0,0	0,7	1,6	0,0	ok
138	'Hours_cos'	-1,0	1,0	0,0	0,7	1,5	-0,1	ok
139	'weekday'	1,0	7,0	4,0	2,0	1,7	0,0	ok
140	't2m'	-35,4	29,9	4,7	13,5	2,6	-0,6	ok
141	'ws_10min'	0,0	9,9	2,5	1,3	3,5	0,6	ok
142	'wg_10min'	0,2	23,6	4,6	2,5	3,6	0,8	ok
143	'wd_10min'	0,0	360,0	193,9	94,4	2,0	-0,1	ok
144	'rh'	17,0	100,0	77,4	18,7	3,0	-0,9	ok
145	'td'	-38,9	22,8	0,4	12,1	2,9	-0,6	ok
146	'r_1h'	-1,0	17,4	-0,8	0,6	130,4	7,4	ok
147	'ri_10min'	0,1	40,9	NaN	NaN	158,5	9,0	ok
148	'snow_aws'	-1,0	83,0	25,8	32,6	1,5	0,6	ok
149	'p_sea'	981,5	1042,5	1013,1	10,9	3,0	0,1	ok
150	'vis'	230,0	50000,0	30819,3	16584,3	1,7	-0,4	ok
151	'n_man'	-1,0	9,0	4,2	3,4	1,3	-0,2	ok
152	'wawa'	0,0	87,0	20,0	30,8	2,3	1,1	ok

APPENDIX 2. House A cluster interpretation sheet

SITUATION	Cluster	Months	Hours	TemOH_13	TemMH_13	TemPH_13	HumOH_13	HumMH_13	HumPH_13	CO2OH_13	CO2MH_13	COOH_13	IVPre_13	DifPre_13	weekday	t2m	ws_10min	wd_10min	rh	ri_10min	p_sea
	C1(23414)																				
Winter	Min	1	0	18,52	16,1	12,43	14,57	14,99	9,87	291,24	356,49	0,76	0,11	-11,15	1	-35,4	0	0	70	0,1	994,63
Cold day	Max	12	23	22,74	19,68	22,89	32,75	31,42	62,43	941,03	1070,99	19,27	4,11	48,04	7	-1,4	7,7	356	100	1,8	1039,7
Empty house	Med	2	10	20,18	18,18	16,33	25,5	23,57	13,06	518,86	453,27	0,9	1,64	10,79	3	-19,6	1,81	175,8	84	NaN	1029,01
	Avg	3,4	10,87	20,26	18,06	16,47	23,96	23,54	13,83	534,53	500,09	1,12	1,59	12,28	3,56	-19,6	1,98	207,36	84,3	NaN	1025,83
	Std	3,85	6,49	0,59	0,63	1,9	4,07	2,16	2,85	103,16	122,17	0,9	0,79	8,54	1,95	7,91	1,16	101,18	6,39	NaN	11,89
Winter	C2(5986)																				
Cold evening	Min	1	4	20,96	16,62	12,35	11,5	17,96	9,61	486,12	351,52	0,82	0,11	-5,21	1	-31,7	0	0	71	0,2	994,61
Fireplace in use	Max	12	21	30,34	20,24	23,43	30,08	29,14	77,31	1092,45	972,08	80,51	4,39	48,04	7	-1,4	5,5	351	100	1,35	1039,7
Carbon monoxide spikes	Med	2	16	23,62	18,39	16,06	19,17	22,64	13,59	678,55	415,04	1,8	1,71	15,85	4	-21,02	1,7	165,8	84	NaN	1031,02
Living room CO2 slightly elevated	Avg	5	14,86	23,93	18,39	16,52	19,48	22,86	14,05	703,16	441,49	3,32	1,61	18,73	4,12	-20,6	1,76	197,7	83,79	NaN	1028,32
	Std	4,73	4,04	1,65	0,83	2,39	3,66	2,49	3,58	104,92	82,62	5,85	1,05	10,96	1,97	6,71	0,96	106,09	5,18	NaN	8,68
	C3(18002)																				
Winter	Min	1	0	19,18	17,3	13,35	19,16	20,14	11,17	293,57	603,7	0,72	0,11	-17,33	1	-32	0	0	47	0,13	981,5
Generic day	Max	12	23	24,52	23,02	28,5	37,66	37,47	53,05	1017,77	1193,58	10,93	3,8	41,9	7	8,77	7,8	360	100	7,2	1042,4
House occupied	Med	2	10	21,1	20,08	22,11	28,12	28,8	17,46	618,74	840,79	0,8	1,08	4,96	5	-7,9	2,5	191,79	90	NaN	1012,3
	Avg	2,46	11,55	21,09	20,13	21,83	28,17	28,81	18,02	627,86	844,51	0,91	1,09	6,48	4,36	-9,43	2,71	211,78	88,35	NaN	1011,07
	Std	2,16	7,98	0,59	0,97	2,1	2,92	2,77	3,54	129,74	115,99	0,52	0,51	6,31	2,02	7,45	1,28	92,76	8,69	NaN	13,79
	C4(22914)																				
Winter	Min	1	0	19,27	17,8	13,21	20,51	22,28	11,25	275,53	934,36	0,73	0,11	-6,39	1	-28,74	0	0	51	0,1	981,8
Night	Max	12	23	23,49	23	29,06	35,51	37,75	57,66	1021,34	1583,66	5,66	2,76	43,02	7	9,77	8,1	360	100	8,39	1042,1
	Med	3	4	21,35	20,81	22,97	27,71	31,05	17,03	579,67	1223,03	0,77	1,06	4,99	4	-5,8	2,5	186	92	NaN	1010,88
	Avg	2,67	8,47	21,31	20,72	22,95	27,56	31,08	17,72	593,02	1205,82	0,83	1,03	5,42	4,05	-6,38	2,62	211,62	87,97	NaN	1010,61
	Std	1,64	9,13	0,57	0,96	2,34	2,89	2,83	3	122,21	103,84	0,26	0,31	4,72	2,08	6,04	1,45	91,3	11,43	NaN	13,86
Winter	C5(2309)																				
Cold night	Min	1	0	19,87	18,56	14,65	28,5	24,69	14,18	436	1431,73	0,79	0,11	2,98	1	-26,3	0	0	80	0,2	997,12
Bedroom CO2 high	Max	12	23	22	21,56	25,9	37,2	45,01	47,5	1960,75	1957,7	3,44	0,52	22,02	7	-7,24	5,93	360	97	0,7	1033,4
Ventilation rate too low	Med	12	3	20,79	19,79	17,63	30,53	37,84	16,05	633,25	1954,85	0,91	0,13	6,5	4	-15,2	1,64	282,6	85	NaN	1024,12
	Avg	11,29	6,39	20,78	19,73	18,33	30,59	38,01	16,37	698,71	1890,59	1,08	0,2	7,28	4,07	-18,22	1,91	212,53	85,51	NaN	1022,31
	Std	2,69	8,03	0,4	0,56	2,32	1,37	3,82	1,95	248,37	127,22	0,37	0,1	2,33	1,52	6,39	1,01	111,06	4,39	NaN	9,38
	C6(32241)																				
Winter	Min	1	0	19,55	17,33	16,98	20,6	19,06	11,81	287,09	284,74	0,73	0,11	-24,96	1	-31,96	0	0	59	0,1	983,54
Generic day	Max	12	23	24,22	21,8	26,62	37,48	37,08	72,47	962,66	711,42	7,3	3,88	47,12	7	5,25	7,5	360	100	2,8	1042,5
House empty	Med	1	12	21,07	19,83	21,64	29,78	27,83	17,32	517,06	455,91	0,82	1,12	6,87	4	-9,9	2,26	174	92	NaN	1008,44
	Avg	2,27	11,68	21,1	19,88	21,41	29,62	27,91	18,18	539,96	464,93	0,92	0,92	7,63	4,22	-9,95	2,42	195,64	91,63	NaN	1009,76
	Std	3,03	5,59	0,48	0,7	1,6	2,7	2,26	3,77	123,83	84,26	0,43	0,65	5,1	1,89	6,75	1,35	94,22	5,35	NaN	12,33
Winter	C7(3182)																				
Sauna in use	Min	1	0	20,18	18,16	21,09	25,58	19,2	14,15	659,27	334,53	0,79	0,11	-6,27	1	-27,2	0,68	117	79	0,2	996,2
Living room CO2 elevated (from people and/or fireplace)	Max	12	23	25,42	22,22	39,37	56,18	43,24	87,9	1960,75	1537,79	21,07	4,76	48,04	7	0,8	5,7	347	100	2,9	1026,2
Frequent in weekends	Med	1	11	21,71	20,74	25,32	36,51	30,54	19,73	1029,01	430,56	1,38	0,13	13,45	6	-9,17	3,3	145,6	95,7	NaN	1004,65
CO elevated a little (probably from fireplace)	Avg	2,71	11,77	21,99	20,56	25,9	36,36	30,55	23,07	1107,32	509	2,07	0,45	13,35	5,7	-9,22	3,24	177,32	94,02	NaN	1006,39
	Std	3,83	8,27	0,94	0,88	2,99	3,57	3,43	10,13	303,79	212,86	2,15	0,68	7,07	1,55	7,85	0,98	69,4	5,76	NaN	9,04
	C8(22645)																				
Spring	Min	1	0	19,6	17,21	19,08	20,41	19,47	16,49	168,62	336,8	0,73	0,11	-20,07	1	-26,9	0	0	26	0,12	981,89
Empty house	Max	12	23	24,33	23,1	35,49	51,69	43,3	95,41	973,12	886,79	8,74	3,44	29,32	7	22,2	7,4	359	100	14	1037,8
	Med	4	11	21,86	21,62	23,87	34,73	33,91	28,33	520,26	526,28	0,76	0,95	3,42	3	4,25	2,9	196,8	90	NaN	1005,8
	Avg	3,65	11,72	21,88	21,54	24,15	35,29	34,3	29,38	545,05	558,46	0,9	0,91	3,25	3,62	5,61	2,95	202,44	82,38	NaN	1006,1
	Std	1,71	5,72	0,63	0,74	1,63	3,16	2,98	9,46	102,14	121,71	0,43	0,3	4,14	1,94	7,92	1,21	66,17	17,98	NaN	8,33
	C9(12207)																				
Spring	Min	1	0	20,45	18,83	17,87	27,94	30,58	16,42	314,77	767,03	0,72	0,11	-11,74	1	-25,9	0	0	38	0,1	987,53
Afternoon	Max	12	23	23,9	23,33	32,85	48,01	47,79	82,26	1049,66	1409,22	4,21	1,91	27,46	7	20,35	6,4	359	100	9,08	1028,81
House occupied	Med	5	16	22	21,9	24,28	36,35	38,76	30,75	589,25	1202,38	0,75	0,97	0,45	4	8,32	2,26	212,7	89,5	NaN	1008,49
	Avg	5	12,1	22,04	21,91	24,51	37,02	39,43	30,23	602,79	1143,72	0,77	0,92	1,2	3,75	7,64	2,31	207,93	85,29	NaN	1010,09
	Std	1,09	9,31	0,76	0,61	1,67	3,51	3,52	5,51	75,73	143,41	0,18	0,22	2,73	2,07	4,52	1,05	72,22	13,99	NaN	8,27
	C10(15873)																				
Early summer	Min	1	0	22,64	19,58	22,19	21,51	24,91	18,37	375,82	356,88	0,73	0,11	-33,68	1	-11,04	0	0	26	0,1	998,08
Generic day	Max	12	23	27,64	25,83	35,72	43,78	45,56	93	1005,53	992,28	8,02	1,5	25,53	7	25,7	7,9	360	98	2,2	1037,46
	Med	6	13	25,21	23,76	25,96	34,18	37,25	32,54	572,93	625,51	0,76	0,94	1,19	3	16,07	2,53	150	54	0,24	1017,26
	Avg	6,65	12,73	25,23	23,69	26,05	34,13	37,17	33,27	576,87	629,81	0,79	0,9	1,43	3,48	15,59	2,63	173,87	57,11	0,46	1016,56
	Std	1,22	4,81	0,88	0,83	2,12	2,73	3,16	6,04	93,78	118,23	0,21	0,22	3,94	1,99	4,72	1,34	103,11	16,66	0,4	4,71
	C11(16514)																				
Spring	Min	1	0	21,29	20,24	23,05	20,49	28,29	13,69	267,09	1020,74	0,72	0,11	-18,1	1	-14,61	0	0	36	0,2	990,27
Night	Max	6	23	26,01	24,91	32,01	37,57	41,31	51,93	1804,39	1652,54	3,37	1,45	20,14	7	16,64	9,9	360	100	7,9	1037,4
Bedroom CO2 elevated	Med	4	3	22,66	22,4	26,41	29,89	34,12	21,6	592,46	1395,81	0,76	0,4	1,98	4	2,2	1,92	271	82,9	0,34	1017,48
	Avg	4,29	8,53	22,91	22,62	26,45	29,85	34,34	21,96	595,62	1383,23	0,79	0,5	2,17	3,94	2,81	2,16	235,5	80,84	0,67	1015,64
	Std	0,66	9,34	0,91	0,67	1,45	2,7	2,53	2,71	92,14	107,92	0,16	0,3	2,04	2,23	3,62	1,28	97,79	13,41	1,19	9,92
	C12(27672)																				
Winter	Min	1	0	20,96	18,7	19,6	14,9	17,71	11,95	298,27	328,77	0,73	0,11	-20,62	1	-29	0	0	20	0,1	984,31
Warm day	Max	5	23	25,42	23,34	29,26	30,76	32,38	52,17	928,65	871,37	13,43	3,4	32,07	7	17,84	8,6	359	100	5,53	1042,5
House empty	Med	3	12	22,23	21,6	24,57	23,88	25,11	16,12	529,27	469,82	0,76	0,89	3,64	4	-0,23	3,34	229,2	66,6	0,2	1011,95
	Avg	3,23	12,15	22,47	21,57	24,38	23,85	24,93	16,78	539,68	492,18	0,8	0,85	3,72	3,85	-0,3	3,35	224,9	67,29	0,35	1012,25
	Std	0,71	4,26	0,92	0,57	1,55	2,87	2,43	2,57	84,69	90,64	0,46	0,29	3,75	1,91	5,55	1,37	82,59	18,46	0,31	13,53
	C13(19924)																				
Spring	Min	1	1	21,07	19,21	20,33	21,71	22,2	14,52	234,37	357,21	0,72	0,11	-42,82	1	-27,2	0	0	23	0,1	982,9
Generic day	Max	6	22	25	24	37,78	44,78	38,95	64,34	974,07	939,5	19,87	3,25	22,22	7	22,03	9,4	359	100	11,47	1042,1
Living room occupied	Med	4	13	22,73	22,31	26,28	30,12	30,48	21,96	647,63	591,42	0,76	0,49	2,67	5	4,73	2,49	238	67	0,34	1012,91
	Avg	4,24	12,81	22,84	22,31	26,48	30,15	30,42	23,02	639,83	601,77	0,9	0,59	2,3	4,29	5,75	2,7	226,03	66,5	0,94	1012,95
	Std	0,71	5,18	0,77	0,59	2,01	2,94	2,49	4,73	93,63	127,48	0,9	0,33	3,39	2,08	6,04	1,46	96,64	18,72	1,85	10,02
	C14(16320)																				
Spring	Min	1	3	23,55	20,18	22,92	15,75	18,39	11,38	400,15	336,88	0,74	0,11	-48,85	1	-27,4	0	0	18	0,2	994,4
Day	Max	6	22	28,21	25,56	40,9	33,86	38,23	53,9	889,83	906,62	28,27	4,68	45,23	7	25,9	9,3	359	85,2	6,54	1041,1
Ambient high pressure, low humidity (probably sunshine)	Med	4	13	25,28	23,59	27,97	23,74	28,68	19,51	547,03	469,65	0,78	0,37	1,9	4	11,49	2,87	264	38,8	0,36	1020,91
Sauna in use (maybe after outdoor activities)	Avg	4,5	13,12	25,45	23,48	28,41	23,79	28,49	19,66	554,63	496,65	0,9	0,56	1,83	3,95	11,31	2,97	224,51	40,81	0,66	1020,15
	Std	0,74	3,45	0,94	0,8	2,38	3,02	3,67	4,14	81	92,99	1,55	0,43	4,19	2	5,49	1,45	99,19	12,78	1,1	8,21
	C15(9597)																				
Spring	Min	2	0	21,47	20,21	20,95	17,07	20,56	11,98	282,4	695,33	0,72	0,11	-33,93	1	-15	0	0	26,2	0,2	988,2
Evening	Max	6	23	26,73	24,99	31,72	34,98	35,89	55,17	969,62	1425,14	5,83	1,74	23,44	7	21,34	8,14	360	100	10,85	1038,58
House occupied	Med	5	16	23,83	22,78	26,62	25,63	28,24	20,06	629,04	1000,58	0,76	0,96	2,81	4	3,4	2,38	154,2	62,6	0,3	1019,57
	Avg	4,41	13,65	23,79	22,86	26,65	25,65	28,13	20,12	622,66	1014,84	0,84	0,84	3,24	4,37	3,7	2,5	191,61	62,33	0,49	1018,36
	Std	0,83	8,09	0,93	0,9	1,57	3,57	2,86	3,67	82,53	136,1	0,46	0,38	2,95	2,03	5,56	1,37	109,36	16,77	0,88	7,81
	C16(28071)																				
Summer	Min	2	0	20,45	18,86	21,46	32,66	27,69	29,19	211,58	348,51	0,72	0,11	-17,76	1	-25,6	0	0	36	0,1	991,7
Generic low ambient air pressure day	Max	9	23	26,13	24,76	33,43	61,78	62,38	94,94	1153,6	1006,05	3,63	2,64	29,75	7	28,15	6,5	360	99	40,9	1024,5
Low ambient air pressure	Med	8	12	23,7	23,12	24,57	49,38	50,42	47,04	636,63	652,78	0,75	0,92	0,72	4	15,28	2,22	160,8	86	0,48	1008,51
	Avg	7,7	11,89	23,7	23,07	24,75	49,76	50,79	48,81	624,68	652,44	0,76	0,93	1,18	3,99	15,86	2,31	176,43	84,38	0,98	1007,73
	Std	1,09	5,62	0,77	0,6	1,68	3,11	3,32	7,24	105,73	118,9	0,1	0,12	3,12	1,92	3,6	1,11	90,2	10,6	1,77	6,8
	C17(27632)																				
Summer	Min	4	0	22,08	21,87	22,62	35,5	27,51	26,99	409,63	312,53	0,72	0,11	-19,93	1	1,55	0	0	33	0,1	993,44
Generic afternoon	Max	9	23	27,49	26,36	34,99	57,2	54,73	73,88	903,31	971,05	3,46	1,68	22,38	7	28,6	8	358	99	21,28	1024,49
	Med	7	13	25,14	24,11	26,8	43,98	46,24	39,37	603,24	625,39	0,75	0,87	1,76	3	17,1	2,52	155	69,1	0,28	1013,27
	Avg	7,27	12,56	25,1	24,05	26,78	43,88	46,22	39,99	597,74	633,28	0,76	0,92	2,25	3,39	17,3	2,57	167,48	70,53	0,5	1012,41
	Std	0,92	4,94	0,91	0,7	1,56	2,85	2,49	4,5	83,66	125,71	0,1	0,17	3,38	1,94	3,25	1,31	83,8	14,69	0,63	7,02
	C18(9686)																				
Summer	Min	4	0	23,69	22,68	22,21	30,68	32,54	24,51	414,59	844,16	0,73	0,11	-17,43	1	5,34	0	0	31	0,2	999,4
Early summer night	Max	8	23	28,58	27,75	32,3	47,03	50,82	56,54	927,79	1387,87	3,08	1,36	23,57	7	23,26	6,37	359	99	4,8	1023,3
	Med	6	5	24,96	24,07	26,48	37,65	42,5	33,59	541,17	1194,31	0,76	0,9	2,58	4	12,9	1,9	109	79,05	0,35	1014,3
	Avg	6,45	9,82	25,31	24,56	26,85	37,77	42,16	33,91	552,35	1171,57	0,79	0,92	3,17	3,89	13,42	2,01	140,49	77,39	0,47	1013,61
	Std	0,77	9,2	1,17	1,09	2,31	3,24	3,57	4,03	69,62	109,39	0,12	0,2	3,3	2,09	3,56	1,29	97,36	15,17	0,33	5,38
	C19(29356)																				
Summer	Min	5	0	20,79	20,8	21,55	38,5	44,21	32,57	472,41	849,86	0,72	0,46	-11	1	5,4	0	0	53	0,1	991,4
Late summer night	Max	9	23	25,22	24,88	29,85	56,15	59,49	66,93	1194,65	1453,26	3,62	1,4	26,36	7	20,67	6,3	360	100	6,4	1024,1
	Med	8	5	23,59	23,12	25,28	48,11	51,38	42,79	583,07	1155,14	0,75	0,92	0,75	4	13,58	1,89	149,6	92	0,28	1009,97
	Avg	7,62	9,82	23,6	23,22	25,31	48,02	51,04	43,01	597,88	1140,87	0,78	0,93	1,13	4,15	13,7	1,98	164,98	89,59	0,53	1010
	Std	1,12	9,03	0,64	0,6	1,61	3,2	2,55	4,47	71,34	112,66	0,14	0,11	2,26	1,97	2,48	1,02	86,8	8,86	0,63	8
	C20(13176)																				
Summer	Min	6	0	24,09	23,34	23,96	42,58	46,02	33,91	420,89	771,95	0,73	0,11	-7,48	1	10,93	0	0	44	0,1	1001,41
Midsummer night	Max	8	23	28,06	26,58	32,63	59,42	65,93	62,39	925,77	1501,23	3,78	1,31	13,1	7	27,54	5,3	359	99	4,16	1023,4
	Med	7	5	25,93	24,9	27,54	48,82	53,88	44,36	566,75	1053,88	0,77	0,96	1,83	5	18,45	1,7	154,2	85	0,51	1015,16
	Avg	7,16	9,81	26,02	24,97	27,6	49,57	54,15	45,17	575,24	1059,1	0,79	0,95	2,2	4,52	18,68	1,75	161,44	83,71	0,74	1014,46
	Std	0,54	8,87	0,72	0,56	1,68	3,05	2,92	5,25	64,59	140,3	0,1	0,16	2,02	1,91	2,92	1,08	93,71	11,39	0,65	4,24
Summer	C21(6871)																				
Low humidity day	Min	6	3	25,89	23,54	25,05	27,35	31,66	21,4	402,3	332,73	0,74	0,11	-14,17	1	9,2	0	0	17	0,2	1004,1
Living room and bathroom warm up (probably by sunshine)	Max	8	22	30,54	29,46	36,76	45,53	52,97	86,77	901,65	1195,67	1,35	1,26	21,07	7	29,5	7,3	357	94	4	1022,1
	Med	6	14	28,02	26,27	30,5	36,59	40,71	32,31	547,54	531,28	0,83	0,75	3,03	5	23,9	2,52	131,9	45,2	0,37	1013,5
	Avg	6,47	13,89	28,11	26,02	30,84	36,23	40,95	32,72	549,99	575,45	0,85	0,78	3,43	4,86	22,55	2,57	142,8	47,7	0,59	1013,23
	Std	0,65	3,92	0,94	0,98	2,3	4,05	5,1	6,14	75,75	144,53	0,08	0,19	2,69	1,82	4,79	1,3	88,62	18,44	0,47	3,32
	C22(16040)																				
Summer	Min	6	0	24,18	22,89	23,34	39,15	44,89	33,77	135,37	296,81	0,73	0,11	-23,88	1	13,6	0	0	34	0,1	1002,5
Hot and humid day	Max	8	23	30,14	27,11	35,39	62,16	68,68	95,04	927,07	942,09	3,09	1,46	28,18	7	29,9	7	357	99	38,3	1023,16
	Med	7	13	26,7	25,23	27,46	50,13	54,41	50,03	566,7	597,63	0,8	0,93	2,44	5	23,08	2,39	149	70	0,86	1013,82
	Avg	7,06	13,31	26,89	25,09	27,87	50,13	55,09	50,27	566,88	605,7	0,82	0,93	2,81	4,7	23,17	2,41	166,54	69,88	1,26	1012,63
	Std	0,46	4,72	1,03	0,83	2,5	4,03	3,67	7,5	86,83	115,88	0,11	0,16	3,07	1,69	3,13	1,09	83,17	15,85	2,07	4,64

APPENDIX 3. House B cluster interpretation sheet

Situation	Cluster	Months	Hours	TemMH_1	TemPH_1	TemOH_1	HumMH_1	HumPH_1	HumOH_1	CO2MH_15	CO2OH_15	COOH_15	IVPre_15	weekday	t2m	ws_10min	wd_10mi	rh	ri_10min	p_sea
	C1(12566)																			
Winter	Min	1	0	19,55	22,87	16,64	29,13	25,58	33,95	613,04	540,44	0,42	0,16	1	-29,5	0	0	75	0,2	998,7
Cold day	Max	12	23	23,46	29,39	22,85	49,87	75,19	61,27	1555,68	1223,87	43,38	24,13	7	8,48	5,5	360	98	1,6	1034
House occupied	Med	12	10	21,61	24,62	19,51	36,74	32,77	42,44	825,77	784,5	1,31	0,17	4	-14,7	1,7	142	86	NaN	1024
Christmas & new year	Avg	10,5	10,42	21,58	24,94	19,6	37,34	33,81	42,9	854,11	789,39	2,98	1,73	4,03	-16,73	1,9	179,51	85,8	NaN	1022,1
	Std	3,76	6,93	0,7	1,03	0,9	2,81	5,61	4	140,11	82,74	4,2	3,23	1,99	6,24	0,96	109,92	4,64	NaN	8,33
	C2(5736)																			
Shower	Min	1	0	20,3	22,34	18,44	19,5	56,17	22,28	442,75	430,07	0,42	0,16	1	-31,74	0	0	26	0,18	982,2
Evening	Max	12	21	25,43	31,46	28,93	56,69	97,56	52,14	1164,89	1190,29	20,22	25,05	7	23,4	7,4	360	100	4,86	1042,1
	Med	5	18	22,77	25,05	23,67	35,43	82,17	33,97	638,67	733,41	0,44	0,17	4	6,69	1,86	167	83,1	NaN	1015,2
	Avg	4,62	17,43	22,81	25,6	23,69	36,2	81,04	34,82	650,14	732,2	2,5	2,83	3,76	2,31	1,98	182,67	77,01	NaN	1014,46
	Std	2,64	3,18	0,69	1,51	1,41	7,54	10,89	6,29	105,72	117,35	3,66	4,65	2,01	13,56	1,3	100,22	19,34	NaN	11,84
	C3(21009)																			
Spring/autumn	Min	1	0	21,61	22,72	19,42	36,77	41,22	37,17	702,36	543,68	0,42	0,16	1	-26,7	0	0	50,4	0,1	995,97
House occupied	Max	12	23	24,61	28,19	26,32	58,96	96,87	63,37	1304,85	1012,4	14,92	24,42	7	21,5	5,6	356	100	6,4	1033,8
	Med	8	4	22,91	23,8	24,14	51,31	52,54	48,07	869,4	775,59	0,42	0,17	4	13,3	1,68	153	94	NaN	1010,8
	Avg	7,78	9,84	22,94	23,99	24,12	51,03	53,76	47,63	870,93	778,75	0,49	0,85	3,91	13,13	1,72	163,71	91,58	NaN	1010,92
	Std	1,07	9,46	0,47	0,74	0,75	2,78	5,57	2,54	77,69	77,15	0,65	1,93	1,95	3,39	0,97	89,78	7,43	NaN	7,34
	C4(12232)																			
Spring/autumn	Min	1	0	21,7	24,22	20,78	30,89	31,12	32,08	479,1	461,19	0,42	0,16	1	-21,7	0	0	30	0,2	987,02
Sauna?	Max	12	23	25,91	31,86	27,41	56,7	74,41	51,9	1175,83	1136,84	27,62	23,33	7	22,39	6,4	358	99	5,5	1038,49
	Med	6	20	23,52	26,93	24,7	45,06	42,68	42,27	825,89	736,36	0,42	0,17	4	12,71	2,16	139,4	89,3	NaN	1012,1
	Avg	6,23	14,97	23,58	27,2	24,63	45,75	43,68	43	818,17	736,08	1,24	0,45	3,93	10,27	2,23	149,69	86,12	NaN	1011,34
	Std	2,07	8,79	0,65	1,4	1	4,51	6,19	3,81	105,11	101,03	2,85	1,62	2,22	8,33	1,22	76,8	10,91	NaN	9,05
	C5(30957)																			
Winter	Min	1	0	20,03	21,64	19,73	22,74	13,89	25,21	492,87	547	0,42	0,16	1	-35,4	0	0	28	0,1	984,21
Generic morning	Max	12	23	23,16	27,48	23,29	39,4	61,62	42,16	960,43	1044,77	40,54	13,52	7	18,93	7,6	359	100	4,7	1039,52
	Med	2	8	21,87	24,09	21,49	29,55	28,15	30,88	702,42	723,73	0,58	0,52	4	-7,2	2,2	202,3	88	0,2	1016,24
	Avg	2,63	9,12	21,85	24,21	21,47	29,63	28,57	31,56	704,36	722,17	1,53	2,14	4	-8,69	2,29	216,41	85,63	0,32	1015,18
	Std	2,05	5,96	0,45	0,62	0,67	2,36	3,89	2,87	77,04	77,47	2,74	2,58	2,01	11,98	1,19	91,11	12,52	0,37	12,24
	C6(15830)																			
Spring	Min	1	0	20,48	22,27	21,33	26,89	30,85	26,2	452,79	588,44	0,42	0,16	1	-31,28	0	0	31	0,1	987,3
Living room occupied	Max	12	23	24,66	27,89	26,4	48,84	68,62	46,33	1205,85	1015,99	35,06	23,26	7	20,1	7,2	359	100	6,64	1039,5
	Med	5	12	22,63	24,09	23,71	38,01	40,84	37,17	750,83	813,8	0,42	0,17	4	10,1	2,36	200	81,1	0,41	1010,7
	Avg	4,96	11,32	22,62	24,19	23,71	38,04	42,42	36,99	747,28	811,57	0,99	0,74	4,16	7,66	2,33	204,95	77,2	0,54	1012,72
	Std	1,82	7,88	0,58	0,77	0,88	3,44	7,04	2,93	115,07	73,95	2,03	2,07	2,06	9,18	1,12	84,32	16,52	0,54	8,17
	C7(24915)																			
Winter/spring	Min	1	0	20,73	22,8	20,03	22,52	21,12	20,37	713,04	565,78	0,42	0,16	1	-33,8	0	0	29,6	0,1	982,44
Night	Max	12	23	24,31	27,16	25,38	39,83	72,8	39,75	1339,52	1477,49	64,12	12,15	7	19,27	7,9	359	100	7,48	1038,6
Bedroom occupied	Med	4	6	22,52	24,29	22,07	30,86	30,24	31,08	893,56	781,24	0,51	0,17	4	-0,4	1,97	195,6	88	0,27	1017
	Avg	3,08	10,26	22,55	24,44	22,21	30,82	30,93	31,65	903,6	788,18	1,3	1,5	4,06	-3,31	2,17	211,13	83,06	0,58	1014,66
	Std	1,7	8,96	0,46	0,72	0,82	2,68	4,94	3,02	83,79	87,12	2,73	2,13	2,01	10,03	1,26	94,98	14,92	0,95	11,03
	C8(31564)																			
Winter	Min	1	0	19,9	21,66	19,32	18,99	17,95	19,91	465,51	459,79	0,42	0,16	1	-34,68	0	0	23	0,1	984,7
House unoccupied	Max	12	23	22,82	27,17	23,94	32,51	63,08	35,57	835,75	836,23	98,22	23,97	7	19,19	7,7	360	100	4,6	1039,7
	Med	2	10	21,56	23,83	21,4	25,71	24,12	27,6	567,51	584,07	0,45	2,53	4	-6,5	2,2	193	88	0,2	1012,7
	Avg	2,5	10,43	21,51	23,9	21,37	25,66	24,38	27,66	579,11	589,62	1,09	2,85	3,62	-6,86	2,46	215,67	83,88	0,29	1014,15
	Std	1,33	4,45	0,44	0,49	0,66	2,25	3,12	2,21	62,12	66,92	2,39	2,65	1,96	10,19	1,36	99,58	16,08	0,19	14,8
	C9(18925)																			
Winter/spring	Min	1	0	21,74	24,95	20,9	21,88	20,71	24,38	529,68	503,12	0,42	0,16	1	-32,11	0	0	28	0,1	982
Sauna?	Max	12	23	25,05	31,52	25,8	41,59	60,66	44,29	1316,95	965,67	27,51	23	7	19,08	9,9	359	100	11,5	1041,6
	Med	4	20	23,22	26,89	23,07	33,06	30,25	33,91	841,67	748,66	0,71	0,17	4	-0,1	2,24	160,5	85,8	0,26	1013,81
	Avg	3,28	15,41	23,31	27,15	23,08	33,34	30,9	34,17	841,97	743,01	1,71	1,41	4,16	-2,24	2,32	182,69	79,79	0,97	1013,96
	Std	1,96	8,65	0,55	1,16	1	3,11	5,32	3,24	107,42	83,47	2,37	2,54	2,09	11,69	1,24	90,4	17,3	2,15	12,96
	C10(25051)																			
Winter/spring	Min	1	0	21,99	23,13	20,78	19,67	19,63	20,4	573,47	490,64	0,42	0,16	1	-33,9	0	0	29	0,1	981,5
House occupied	Max	6	23	25,37	28,22	25,36	36,84	62,88	37,77	1006,69	906,57	40,39	25,84	7	19,3	8,5	360	100	6,23	1042,5
	Med	3	15	23,22	24,47	22,43	27,24	27,79	28,55	760,86	705,44	1,19	3,11	4	-4,06	2,67	228,8	88	0,2	1009,01
	Avg	2,71	11,91	23,3	24,64	22,56	27,25	29,09	28,41	759,59	699,64	2,72	3	4,06	-5,26	2,75	214,39	81,31	0,4	1010,98
	Std	1,15	8,33	0,53	0,75	0,7	2,88	5,24	2,56	69,33	70,76	3,88	2,36	1,94	8,44	1,49	97,26	16,82	0,75	14,79
	C11(16347)																			
Spring	Min	1	3	21,35	22,31	22,38	18,31	17,54	23,56	467,61	576,13	0,42	0,16	1	-25,82	0	0	18	0,15	983,5
Living room occupied	Max	6	21	23,97	28,03	26,69	40,23	61,72	36,88	806,16	1165,83	23,31	11,01	7	22	8,5	360	100	7,2	1039,51
	Med	5	12	22,64	24,4	24,21	29,47	30,52	30,19	574,01	783,72	0,42	0,17	4	10,74	3	181	50,8	0,3	1017,9
	Avg	4,55	11,51	22,66	24,43	24,26	29,47	30,79	30,34	584,74	792,89	0,59	0,57	3,66	10,07	3,04	204,16	54,98	0,6	1015,37
	Std	1,09	4,15	0,44	0,66	0,9	3,39	5,09	2,58	59,96	90,07	1,05	1,37	2,07	6,51	1,23	96,87	21,88	0,99	8,8
	C12(19943)																			
Summer	Min	1	0	21,4	22,27	22,04	26	28,71	24,93	428,43	422,23	0,42	0,16	1	-24,58	0	0	21	0,1	987,44
House unoccupied	Max	9	23	24,48	28,04	27,17	46,28	68,34	42,25	885,99	838,99	31,91	25,6	7	27,7	7,9	360	100	14	1035,01
	Med	6	12	23,14	24	24,81	37,69	39,57	35,72	516,91	556,93	0,42	0,17	3	15,72	2,85	169	56,2	0,31	1015
	Avg	6,31	11,62	23,08	24,13	24,77	37,36	40,06	35,46	545,04	567,33	0,5	0,61	3,53	15,73	2,95	186,8	59,4	0,79	1014,28
	Std	1,19	5,27	0,51	0,63	0,83	3,09	5,23	2,71	84,62	82,05	1,01	1,69	1,93	4,59	1,4	99,79	18,98	1,29	6,08
	C13(28224)																			
Spring	Min	1	0	21,51	22,83	21,12	17,87	17,57	19,88	462,24	457,85	0,42	0,16	1	-29,58	0	0	20	0,1	984,8
Generic day	Max	6	23	24,77	27,7	26,2	33,46	62,26	33,41	871	920,24	54,84	11,55	7	21,6	9,4	360	100	10,29	1042,5
	Med	4	12	22,61	24,06	22,73	24,91	25,17	26,4	602,54	614,12	0,5	1,75	4	1,78	2,78	238,8	65	0,26	1011,5
	Avg	3,38	11,62	22,65	24,17	22,82	24,88	25,32	26,37	608,49	620,68	2,04	2,35	4,22	1,13	2,95	226,95	64,63	0,81	1012,42
	Std	1,14	5,12	0,44	0,62	0,89	2,71	3,3	2,4	65,4	66,99	4,17	2,34	2,01	9,32	1,39	97,17	23,8	1,66	12,49
	C14(4660)																			
Summer day	Min	1	0	24,16	23,89	24,67	30,62	31,69	25,91	435,07	426,5	0,42	0,16	1	-3,5	0	0	17	0,2	1003,55
Ambient humidity low	Max	7	23	25,96	29,36	30,03	50,35	92,33	45,88	1189,25	851,51	11,1	23,01	7	29,5	7,3	359	94	2,6	1016,8
Sun heats up the living room?	Med	6	13	25,33	25,24	27,68	41,5	44,63	35,78	510,83	539,93	0,43	0,17	6	22,71	2,3	108	51,1	0,31	1013,4
	Avg	6,09	12,57	25,17	25,38	27,78	41,25	45,84	35,74	592,31	551,9	0,48	0,45	5,46	22,23	2,49	134,14	50,95	0,41	1013,06
	Std	0,39	5,62	0,43	0,8	0,78	4,41	6,51	4,43	163,58	92,66	0,47	1,23	1,27	5,13	1,35	90,95	20,55	0,33	2,42
	C15(22466)																			
House empty	Min	2	0	18,56	20,43	18,75	19,88	20,59	21,8	456,02	432,77	0,41	0,16	1	-14,6	0	0	48	0,2	984,3
Vacation	Max	3	23	20,97	22,14	20,8	26,3	26,01	34,36	689,07	696,47	6,45	6,48	7	4,8	8,6	345	100	2,2	1035,7
	Med	3	11	18,99	20,77	19,14	23,62	22,98	24,74	477,63	453,02	0,42	0,17	4	-2,14	3,47	219,4	82	0,2	1012,74
	Avg	2,82	11,43	19,13	20,85	19,24	23,54	23,07	24,88	478,9	455,91	0,56	0,19	4,25	-2,32	3,54	226,51	80,79	0,24	1012,13
	Std	0,39	6,86	0,47	0,27	0,32	1,57	1,29	1,69	12,82	17,92	0,38	0,29	2,04	3,49	1,4	55,35	13,43	0,13	14,67
	C16(19547)																			
Summer	Min	4	2	21,14	22,61	22,39	36,1	42,46	38,02	441,38	589,19	0,42	0,16	1	1,86	0	0	32,7	0,1	991,7
Living room occupied	Max	9	22	24,52	28,81	27,9	62,92	97,37	58,85	751,94	1199,72	0,62	20,02	7	25,22	7,5	359	99	24	1024,15
	Med	8	10	22,67	23,93	24,73	50,54	51,2	47,52	573,9	800,05	0,42	0,17	3	15,25	2,34	151,8	84	0,3	1010,05
	Avg	7,87	10,88	22,75	24,08	24,75	50,49	53,4	47,46	582,79	806,37	0,43	0,49	3,48	15,61	2,34	172,11	81,77	0,68	1009,3
	Std	1,04	4,99	0,41	0,63	0,7	3,12	7,98	2,78	65,51	87,86	0,02	1,62	1,92	2,74	1,1	85,19	12,23	0,96	8,27
	C17(21396)																			
Summer	Min	6	0	21,35	22,29	21,96	38,45	40,6	38,45	433,2	429,23	0,42	0,16	1	7,66	0	0	39,5	0,1	993,8
Generic day	Max	9	23	23,6	26,94	26,29	60,16	94,64	56,67	741,35	742,83	2,54	18,47	7	27,5	7,2	360	98	5,3	1023,8
	Med	8	12	22,59	23,68	24,29	48,51	49,66	45,74	512,9	544,2	0,42	0,17	4	16,14	2,52	157,4	76,4	0,25	1009,94
	Avg	7,59	11,85	22,6	23,7	24,19	48,64	50,14	45,75	516,81	553,09	0,43	0,71	4,2	16,31	2,61	164,44	75,79	0,55	1009,91
	Std	1,09	4,61	0,45	0,64	0,7	3,06	5,04	2,63	47,19	76,29	0,07	1,72	1,97	2,57	1,21	82,72	12,74	0,64	7,37
	C18(19597)																			
Summer	Min	6	0	22,74	22,53	24,12	37,57	40,96	32,72	424,54	424,23	0,42	0,16	1	9,6	0	0	28,8	0,1	994,26
Sunny day	Max	9	23	25,73	28,9	28,65	56,62	92,58	51,06	796,95	848,51	2,57	19,06	7	28,6	8	358	99	12,5	1023,8
	Med	7	13	23,87	24,19	25,73	47,44	51,5	43,65	514,33	477,54	0,42	0,17	4	18,49	2,49	156,5	64	0,42	1013,8
	Avg	7	12,58	23,91	24,08	25,82	47,26	52,28	43,41	517,54	509,49	0,43	0,24	3,89	19,04	2,5	172,44	66,65	0,7	1013,34
	Std	0,57	5,61	0,47	0,73	0,73	3,32	4,99	2,46	61,27	72,57	0,07	0,88	1,73	4,04	1,2	94,88	17,24	0,66	5,14
	C19(15701)																			
Summer day	Min	6	0	22,37	23,4	23,37	49,3	53,42	42,72	428,66	408,77	0,42	0,17	1	12,4	0	0	34,5	0,2	993,4
Ambient humidity up	Max	9	23	26,05	30,43	28,82	72,06	97,53	61,42	755,6	885,4	2,53	19,73	7	29,9	7	359	99	40,9	1022,96
	Med	7	12	24,89	24,82	26,75	59,42	63,96	53,5	482,78	505,75	0,42	0,17	5	23,11	2,27	149	73,7	0,66	1013,7
	Avg	7,08	11,8	24,79	24,93	26,83	60,21	65,3	53,95	505,28	518,79	0,44	0,3	4,34	22,98	2,29	173,39	73,86	1,36	1012,05
	Std	0,32	5,19	0,64	0,74	0,8	4,28	8,31	3,8	66,55	66,52	0,07	1,5	1,87	3,31	1,17	87,52	16,41	2,86	4,5
	C20(12966)																			
Summer	Min	6	0	22,88	23,18	24,16	45,85	48,94	40,54	565,4	443,05	0,42	0,16	1	11,6	0	0	45,8	0,2	992,86
Bedroom occupied	Max	9	23	26,13	28,72	28,39	67,19	92,63	60,61	1090,32	856,88	2,8	18,35	7	26,81	5,7	359	99	7	1024
	Med	7	12	24,47	24,95	25,86	57,84	59,15	52,56	826,78	630,72	0,43	0,17	5	18,6	1,64	150,3	86	0,96	1011,59
	Avg	7,21	11,64	24,57	25,16	25,99	57,78	60,21	52,31	814,08	619,79	0,53	0,2	4,81	18,92	1,7	157,99	84,88	1,1	1010,89
	Std	0,54	9,12	0,7	0,9	0,96	3,31	5	3,54	70,34	69,29	0,26	0,67	2	2,44	0,93	82,8	9,93	0,83	5,88

APPENDIX 4. House C cluster interpretation sheet

SITUATION	Cluster	Months	Hours	TemKT_17	TemPH_1	TemOH_1	HumKT_1	HumPH_1	HumOH_1	CO2KT_17	COOH_17	DifPre_17	IVPre_17	weekday	t2m	ws_10min	wd_10mi	rh	ri_10min	p_sea
Winter	C1(8017)																			
Cold day 1	Min	1	0	22,08	22,06	19,12	18,46	19,32	12,71	420,8	1,03	-7,72	0,16	1	-30,28	0	0	75	0,2	996,1
	Max	12	23	23,76	23,7	21,93	30,78	75,75	23,44	833,57	6,66	9,76	0,16	7	-4,4	6	360	96	1,7	1038,5
	Med	12	12	22,77	22,6	20,68	21,14	26,8	15,76	557,93	1,27	3,56	0,16	5	-19,94	1,88	261,9	84	NaN	1025,3
	Avg	9,79	12,14	22,78	22,68	20,68	21,13	30,01	15,86	569,84	1,36	2,91	0,16	4,56	-19,36	2,17	208,78	84,28	NaN	1026,59
	Std	4,22	7,05	0,32	0,38	0,55	0,91	8,97	1,06	70,47	0,45	2,75	0	1,96	6,21	1,06	112,88	4,53	NaN	4,9
Winter	C2(13295)																			
Cold morning	Min	1	0	22,19	22,2	19,62	18,15	17,96	12,04	350,9	0,92	-21,03	0,16	1	-32,5	0	0	74	0,2	991,2
	Max	12	23	23,95	24,55	23,04	29,19	49,3	24,31	668,93	4,45	19,1	2,69	7	0,83	7,2	359	99	1,26	1039,7
	Med	2	9	23,06	23,2	21,5	21,38	24,1	16,06	502,32	1,15	3,55	0,16	4	-19,28	2	192,2	85,11	NaN	1028,36
	Avg	2,52	9,04	23,03	23,26	21,45	21,3	24,6	16,31	498,42	1,24	2,72	0,16	4,01	-18,9	2,07	217,24	85,4	NaN	1024,89
	Std	2,91	5,96	0,3	0,34	0,41	2,14	4,42	3,19	51,8	0,29	3,48	0,05	1,93	6,67	1,08	101,93	4,9	NaN	13,13
Winter	C3(9568)																			
Shower / sauna	Min	1	0	22,5	22,88	20,5	18,17	38,65	12,39	397,11	0,82	-34,2	0,16	1	-32,49	0	0	35	0,1	983,14
	Max	12	23	25,23	28,48	24,15	40,18	95,9	41,95	926,23	2,75	23,67	12,97	7	14,6	8,1	360	100	8,36	1039,47
	Med	2	19	23,65	24,44	22,67	25,72	50,73	21,81	645	1,15	5,65	0,16	5	-2,64	2,14	195,2	91	NaN	1007,5
	Avg	2,87	18,21	23,61	24,52	22,66	25,67	53,14	22,39	648,12	1,2	5,84	0,52	4,37	-4,45	2,37	206,43	87,39	NaN	1007,78
	Std	2,43	4,44	0,38	0,88	0,53	2,95	9,43	3,95	88,05	0,24	3,02	1,92	2,02	7,83	1,42	95,91	12,9	NaN	11,72
Winter	C4(26973)																			
Night	Min	1	0	22,57	22,28	20,85	18,45	20,95	12,36	403,45	0,92	-39,31	0,16	1	-32,3	0	0	56,8	0,1	981,5
	Max	12	23	24,69	24,51	23,56	40,19	52,94	30,6	739,85	3,37	30,18	9,83	7	1,3	7,9	360	100	2,9	1038,44
	Med	1	10	23,33	23,27	22,18	23,96	28,57	20,25	596,51	1,15	5,22	0,16	3	-6,9	2,5	193	92,7	NaN	1008,6
	Avg	1,78	10,67	23,39	23,26	22,17	24,23	29,98	19,98	591,86	1,21	4,4	0,22	3,76	-8,3	2,58	206,63	92,33	NaN	1011,85
	Std	2,26	6,95	0,31	0,27	0,41	1,98	5,38	2,53	56,93	0,2	3,71	0,71	1,97	6,42	1,11	86,3	5,23	NaN	11,49
Winter	C5(11169)																			
Cold day 2	Min	1	0	22,78	22,8	20,72	17,92	18,22	12,26	379,28	0,83	-21,74	0,16	1	-35,4	0	0	33	0,2	984,5
	Max	12	23	25	26,32	23,92	31,83	60,49	30,92	922,99	4,23	22,3	15,36	7	14,82	7,6	355	100	4,7	1039,66
	Med	2	11	23,52	23,83	21,84	20,84	24,25	16,13	546,44	1,29	3,85	0,16	4	-20,58	1,5	176,6	85	NaN	1027,4
	Avg	3,83	11,47	23,55	23,91	21,83	21	25,4	15,7	562,27	1,36	3,13	0,25	3,67	-20,86	1,77	205,44	83,66	NaN	1023,25
	Std	4,21	6,98	0,31	0,51	0,4	1,74	5,78	2,34	78,07	0,29	2,76	0,81	1,99	8,39	1,26	98,2	7,36	NaN	12,03
Winter	C6(17080)																			
Kitchen occupied	Min	1	0	22,14	22,27	21,15	19,12	19,33	12,52	639,63	0,82	-27,19	0,16	1	-29,05	0	0	42	0,1	983,2
More frequent on weekends	Max	12	23	25,01	25,43	23,76	43,07	55,59	40,11	1068,52	4,38	29,22	12,58	7	13,9	7,52	360	100	1,82	1033,5
	Med	1	13	23,57	23,4	22,34	25,2	30,86	20,12	754,22	1,27	5,74	0,16	5	-6,5	2,62	158,9	94,6	NaN	1009,2
	Avg	1,98	12,22	23,66	23,48	22,38	25,47	32,32	19,91	770,3	1,27	5,81	0,21	4,57	-6,17	2,62	174,63	93,44	NaN	1010,87
	Std	2,64	6,75	0,38	0,36	0,4	2,47	6,56	3,17	72,27	0,18	2,65	0,59	2,07	5,2	1,27	71,34	6	NaN	11,32
Spring	C7(15926)																			
Ambient air pressure low	Min	1	0	21,98	22,67	20,91	19,64	20,75	15,87	393,09	0,81	-32,2	0,16	1	-17,1	0	0	33	0,1	981,5
	Max	8	23	24,46	25,78	23,65	41,49	51,85	43,85	923,71	3,8	39,71	15,82	7	15,12	8,6	360	100	8,2	1037,99
	Med	4	10	23,13	24,22	22,49	26,12	29,08	25,36	586,74	1,04	6,98	0,16	4	1,09	2,36	214,4	91	0,21	1008,3
	Avg	3,4	11,07	23,16	24,17	22,53	26,29	30,55	24,93	594,95	1,06	7,29	0,56	3,85	0,92	2,59	215,38	86,21	0,36	1007,59
	Std	1,25	7,75	0,34	0,56	0,38	2,88	5,74	3,32	61,08	0,17	3,93	2,15	2,04	4,31	1,46	90,57	14,07	0,42	11,67
Spring	C8(20643)																			
Generic spring day	Min	1	0	22,48	22,96	21,51	18,99	17,51	16,99	337,74	0,82	-46,09	0,16	1	-21	0	0	20	0,1	986,66
	Max	5	23	24,11	26,81	23,59	31,37	45,6	35,76	602,33	11,37	32,82	16,07	7	19	9,9	360	100	7,9	1041,9
	Med	4	12	23,21	24,75	22,64	24,33	23,11	24,6	440,58	1,03	6,48	0,16	4	3,6	2,97	253	62	0,32	1015,3
	Avg	3,93	11,44	23,19	24,8	22,61	24,23	23,79	24,79	442,13	1,05	5,11	2,27	4,14	3,25	3,15	218,06	62,54	0,78	1013,81
	Std	0,74	6,51	0,28	0,55	0,33	1,89	3,3	3,2	56,36	0,25	5,38	4,81	2,03	5,95	1,56	100,94	18,2	1,26	10,07
	C9(19071)																			
Winter/spring	Min	2	0	22,84	23,75	21,58	18,12	17,77	12,23	428,27	0,82	-23,55	0,16	1	-30,22	0	0	21,6	0,1	984,5
Ambient air pressure low	Max	6	23	25,51	26,63	24,71	37,62	47,54	32,57	844,76	3,6	30,65	22,96	7	18,38	7,9	359	100	9,52	1042,1
	Med	3	14	23,77	24,98	22,93	24,75	27,35	21,16	603,38	1,05	5,3	0,16	5	-3,5	2,57	200,5	86,8	0,2	1004,9
	Avg	2,7	13,03	23,8	25,03	22,91	24,6	28,23	21,61	610,47	1,17	5,22	0,43	4,56	-5	2,68	209,35	81,08	0,33	1006,58
	Std	0,79	6,17	0,4	0,44	0,37	2,65	5,4	3,06	60,59	0,26	3,55	1,88	1,99	8,17	1,4	91,46	17,05	0,6	14,15
	C10(23400)																			
Spring	Min	2	0	23,11	23,49	22,44	18,94	17,29	16,34	305,78	0,81	-29,33	0,16	1	-6,54	0	0	18	0,2	987,28
High ambient pressure	Max	9	23	25,27	26,55	24,9	36,45	53,74	39,22	634,24	3,55	30,92	24,7	7	21,6	7,2	360	100	11,5	1037,5
	Med	5	13	23,99	24,91	23,6	26,77	26,18	28,96	450,24	0,93	7,18	11,85	5	10,28	2,23	211,75	49,8	0,36	1021,5
	Avg	4,76	12,4	24,02	24,92	23,58	26,61	26,55	28,57	454,91	0,95	6,96	9,04	4,25	10,21	2,38	213,81	52,53	1,11	1021,81
	Std	0,59	6,53	0,3	0,58	0,33	2,3	4,1	2,74	49,37	0,13	3,08	6,15	2,04	5,01	1,19	98,12	20,79	2,3	7,19
Winter/spring	C11(24190)																			
Morning	Min	2	0	22,01	23,47	21,34	19,56	18,38	17,96	350,25	0,77	-32,79	0,16	1	-24,44	0	0	30	0,2	984,31
	Max	8	23	24,04	25,52	23,88	40,58	50,38	41,57	796,56	4,33	34,54	16,13	7	18,4	8,55	360	100	8,39	1042,5
	Med	4	6	22,96	24,28	22,29	25,35	27,13	25,37	552,04	1,03	6,53	0,16	4	0,9	2,68	235,5	85,6	0,26	1008,88
	Avg	3,68	9,08	22,99	24,33	22,37	25,4	27,84	24,97	550,43	1,01	6,04	0,68	3,87	-0,3	2,85	230,1	82,97	0,48	1008,5
	Std	0,76	7,64	0,25	0,37	0,31	2,2	4,58	2,75	64,71	0,12	4,12	2,55	2,05	5,22	1,53	93,89	14,18	0,93	11,55
Winter/spring	C12(23543)																			
Generic day	Min	2	0	22,55	23,75	21,23	17,68	17,7	12,07	336,79	0,82	-44,13	0,16	1	-32,72	0	0	22,4	0,1	984,75
	Max	6	23	24,64	26,79	23,98	28,8	44,51	31,21	617,94	10,46	28,65	24,31	7	21,4	9,3	360	100	11,32	1042,5
	Med	3	12	23,27	24,88	22,4	22,22	21,76	20,81	467,12	1,04	3,99	0,16	4	-3,9	2,5	257,4	75	0,2	1016,34
	Avg	3	11,28	23,33	24,95	22,41	22,21	22,38	20,48	461,58	1,12	3,49	0,33	3,66	-6,23	2,72	227,87	71,47	0,47	1016,9
	Std	0,75	5,93	0,39	0,49	0,39	2,05	3,05	3,59	58,35	0,3	4,92	1,5	1,76	10,1	1,5	96,11	16,45	0,79	14,31
Spring/autumn	C13(19069)																			
Day	Min	4	0	22	22,91	21,2	26,91	25,73	25,45	313,48	0,74	-28,63	0,16	1	1,4	0	0	27	0,1	996,3
	Max	9	23	24,86	26,36	24,27	46,82	66,35	51,29	726,39	4,25	48,4	16,1	7	22,2	6,4	360	100	9,3	1026
	Med	5	11	23,83	23,97	23,3	35,18	35,43	35,93	450,4	0,83	8,97	11,7	3	10,48	2,68	162,2	80,5	0,3	1011,4
	Avg	6,03	11,4	23,75	24,06	23,23	34,99	36,06	36,11	453,29	0,91	9,9	11,9	3,3	11,86	2,57	170,06	75,77	0,86	1012,12
	Std	1,34	6,96	0,43	0,66	0,42	3,13	4,01	3,92	48,64	0,2	5,27	1,36	1,99	4,24	1,23	87,34	18,92	1,49	6,52
	C14(18556)																			
Early summer	Min	4	0	23,8	24,44	23,12	20,92	21,32	21,52	335,6	0,73	-23,32	0,16	1	4,28	0	0	20	0,1	1002,3
Dry day	Max	9	23	26,34	27,1	25,42	40,56	58,47	45,36	674,28	2,4	44,5	15,9	7	25,9	7,9	360	98	14	1023,1
	Med	6	12	24,87	25,44	24,4	32,79	33,04	33,99	430,73	0,84	7,83	11,67	4	13,38	2,81	216,95	56,4	0,53	1013,12
	Avg	5,59	11,7	24,89	25,5	24,4	32,6	32,97	33,88	441,48	0,87	8	11,49	3,84	13,89	2,96	211,87	60,96	0,69	1012,93
	Std	0,62	6,59	0,35	0,49	0,36	3,29	4,79	3,7	47,96	0,08	3,98	1,26	1,97	3,97	1,38	80,91	19,42	0,75	6,13
Summer	C15(41888)																			
Humid day	Min	5	0	22,66	23,71	22,24	33,74	33,11	33,67	285,51	0,74	-4,46	4,83	1	7,3	0	0	34	0,1	991,4
	Max	9	23	25,64	27,07	25	53,45	87,44	56,35	714,69	2,53	43,2	16,47	7	24,76	7,2	360	100	24	1024,1
	Med	7	11	24,44	24,76	23,98	44,23	43,22	45,37	426,13	0,83	8,49	11,83	3	15,29	2,05	155,26	82	0,29	1011,6
	Avg	7,22	11,45	24,44	24,79	23,98	43,51	43,55	45,26	428,06	0,86	8,68	12,96	3,56	15,79	2,13	170,81	79,32	0,6	1011,04
	Std	1,09	7,08	0,29	0,44	0,32	2,96	3,86	3,03	39,82	0,1	2	1,89	1,9	2,78	1,11	85,01	14,68	0,78	8,21
Summer day	C16(17560)																			
Generic day	Min	6	0	23,65	24,81	23,11	32,94	33,8	33,51	280,65	0,73	-40,51	0,16	1	8,17	0	0	29,4	0,1	1000,3
	Max	9	23	26,23	27,83	26,06	54,45	86,87	55,97	572,24	2,47	48,4	21,33	7	28,6	8	357,8	98	12,5	1023,3
	Med	7	13	25,23	25,87	24,83	43,88	42,57	45,1	411,3	0,82	8,64	11,83	3	17,5	2,66	155,4	67,8	0,34	1014,6
	Avg	6,87	12,55	25,22	25,93	24,83	43,69	43,04	45,32	413,77	0,84	9,56	12,01	3,45	18,01	2,74	166,64	68,24	0,73	1013,96
	Std	0,81	5,98	0,38	0,51	0,37	3,33	3,84	3,56	41,48	0,08	4,47	1,88	1,92	4,41	1,4	89,71	16,18	0,73	5,77
Summer	C17(7134)																			
Hot and dry day	Min	6	0	25,26	26,17	24,75	27,25	26,15	24,15	344,91	0,74	-1,28	10,05	1	11	0	0	17	0,21	1007
	Max	7	23	27,72	28,81	27,13	48,33	53,95	47,21	485,95	2,34	20,05	22,35	7	29,5	7,2	359	95	1,64	1019
	Med	6	13	26,13	27,34	25,75	34,99	35,62	36,52	396,98	0,83	9,06	11,95	4	22,8	2,2	141,8	45,55	0,37	1014,64
	Avg	6,07	12,11	26,13	27,43	25,8	35,85	35,4	37,21	399,44	0,85	9,37	12,78	4,33	22,05	2,15	139,13	48,11	0,45	1014,44
	Std	0,25	6,56	0,47	0,55	0,55	4,48	4,2	4,94	23,57	0,07	1,63	2,84	1,52	4,33	1,22	83,62	19,36	0,31	2,22
Summer	C18(14708)																			
Rainy day	Min	6	0	25,23	25,78	24,72	41,62	39,94	39,33	277,99	0,74	-31,12	0,16	1	12,3	0	0	26	0,2	1002,59
	Max	7	23	27,66	28,65	27,1	65,47	87,05	69,03	560,2	2,36	48,4	22,8	7	29,9	7	357	99	38,3	1018,79
	Med	7	12	26,23	26,97	25,94	54,63	52,85	54,42	415,79	0,82	8,69	11,83	6	22,9	2,2	141	77	1,04	1012,5
	Avg	6,99	11,73	26,24	26,99	25,97	54,88	52,93	54,72	417,89	0,84	9,36	11,93	5,08	22,94	2,21	161,49	73,04	1,43	1011,6
	Std	0,09	6,3	0,37	0,46	0,41	4,24	4,51	3,98	28,87	0,08	3,98	1,25	1,84	3,56	1,14	85,93	16,72	2,05	3,9
Summer	C19(14835)																			
Morning	Min	6	0	23,75	24,58	23,07	45,66	43,26	44,27	271,27	0,73	-11,74	7,55	1	13,1	0	0	42,2	0,2	997,51
	Max	8	23	26,2	27,15	25,76	64,25	91,57	68,82	601,14	2,46	41,64	21,67	7	28,3	6,7	359	99	40,9	1022,55
	Med	7	9	25,06	25,63	24,69	52,73	51,18	53,7	429,94	0,82	8,73	12,46	4	19,48	1,97	150,8	84	0,28	1012,7
	Avg	7,3	10,73	25,09	25,68	24,72	53,42	51,98	54,13	426,72	0,84	9,95	13,17	4,18	19,67	2,08	161,86	81,41	0,88	1012,72
	Std	0,48	7,64	0,37	0,44	0,35	3,8	4,38	3,99	46,7	0,07	4,58	1,88	1,78	2,9	1,17	90,19	12,96	2,23	3,97
Early autumn	C20(33007)																			
	Min	6	0	22,06	22,39	21,17	36,92	34,52	38,1	295,78	0,74	-8,43	9,34	1	5,4	0	0	39,2	0,1	991,7
	Max	9	23	25	26,2	24,24	66,68	92,34	67,06	715,76	4,54	48,4	16,11	7	23,04	5,9	360	99	9,4	1023,4
	Med	8	10	23,82	23,73	23,21	45,75	46,09	48,67	447,22	0,82	8,89	12,47	5	14,06	2,18	142,4	89,8	0,24	1009,38
	Avg	8,2	11,03	23,8	23,75	23,2	46,11	46,95	48,52	453,75	0,91	8,99	13,1	4,43	14,25	2,28	162,32	86,69	0,56	1008,69
	Std	0,6	7,19	0,31	0,47	0,35	3,18	4,16	3,23	55,87	0,32	2,35	1,41	2,04	2,6	1,09	88,87	10,48	0,66	7,41

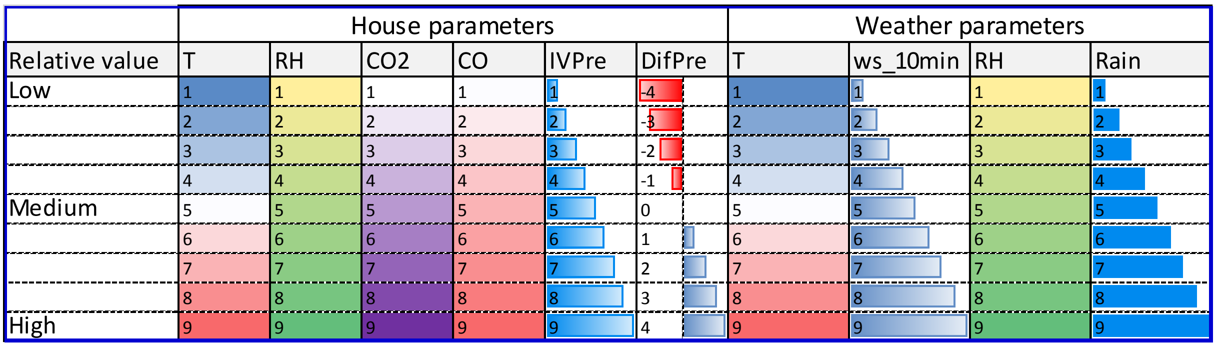
APPENDIX 5. House D cluster interpretation sheet

SITUATION	Cluster	Months	Hours	TemKT_18	TemMH_1	TemPH_18	TemOH_1	HumKT_18	HumMH_1	HumPH_1	HumOH_1	CO2KT_18	CO2MH_1	CO2OH_18	COOH_18	IVPre_18	DifPre_18	weekday	t2m	ws_10min	wd_10mir	rh	ri_10min	p_sea
Winter	C1(21638)																							
Cold day	Min	1	0	23,92	23,53	23,26	24,07	2,99	2,47	2,95	3,84	415,44	457,74	388,38	0,43	0,12	-16,11	1	-29,5	0	0	75	0,1	996,2
Indoor temperature high	Max	12	23	27,92	27,11	28,82	28,24	19,12	19,24	32,1	21,37	1041,31	792,53	967,97	1,96	2,92	26,57	7	0,8	5,82	360	100	1,6	1034
Indoor humidity low	Med	12	12	25,86	25,02	25,64	25,97	9,37	9,29	10,98	10,22	482,02	531,84	480,71	0,46	1,13	3,61	4	-13,02	1,74	144	89,9	NaN	1018,1
	Avg	7,27	11,68	25,85	25,05	25,73	25,94	9,3	8,89	10,28	10,59	494,84	536,67	500,55	0,5	1,15	5,85	4,13	-14,87	2	167,2	88,17	NaN	1016,88
	Std	5,45	6,67	0,66	0,6	1,18	0,59	2,38	2,39	2,82	2,41	61,12	36,18	86,57	0,1	0,42	5,4	1,92	5,83	1,15	93,84	4,92	NaN	10,25
Winter	C2(19008)																							
Low barometric pressure?	Min	1	0	23,23	21,72	21,62	23,41	9,11	8,84	8,3	9,42	400,89	461,81	388,1	0,44	0,13	-22,01	1	-35,4	0	0	18	0,1	994,5
	Max	12	23	26,32	25,35	26,54	26,49	35,85	29,09	46,19	29,43	864,54	793,62	801,53	2,74	2,48	24,12	7	17,06	7,5	360	100	5,97	1036,6
	Med	1	11	24,36	23,48	23,6	24,58	19,24	18,79	17,93	19,55	522,89	592,85	517,41	0,46	1,32	1,99	4	-4,8	2,7	160,2	95	NaN	1006,85
	Avg	1,1	11,42	24,42	23,43	23,65	24,64	19,23	18,48	17,49	19,2	527,15	587,47	524,9	0,48	1,3	2,01	4	-5,16	2,63	176,13	93,5	NaN	1009,56
	Std	0,79	7,09	0,51	0,56	0,96	0,54	3,69	3,56	3,01	3,18	56,59	58,52	65,15	0,09	0,3	2,92	1,97	5,02	1,2	87,87	9,32	NaN	9,65
Spring	C3(15573)																							
Warm day	Min	1	0	23,01	22,42	21,88	23,16	22,21	22,35	19,05	22,91	392,1	452,65	387,62	0,44	0,12	-13,27	1	-11,4	0	0	19	0,1	994,83
	Max	8	23	25,79	25,34	27,32	25,9	36,49	36,96	80,26	37,4	697,37	773,54	659,69	1,75	2,23	26,48	7	28,3	7,9	360	100	4,76	1027,39
	Med	6	13	24,09	23,71	23,59	24,26	31,4	30,93	33,15	31,27	452,95	515,88	440,43	0,44	0,78	2,68	3	14,4	2,64	141,89	56,8	0,47	1015,1
	Avg	5,99	12,13	24,21	23,81	23,85	24,39	31,13	30,64	32,78	30,96	464,51	527,63	453,03	0,45	0,78	3,32	3,15	14,41	2,67	165,88	60,13	0,57	1014,4
	Std	1,13	6,85	0,57	0,58	1,03	0,56	2,75	2,89	3,43	2,69	48,3	44,27	45,5	0,03	0,34	4,14	1,73	6,58	1,44	96,37	20,97	0,53	4,47
Winter	C4(31983)																							
Generic day	Min	1	0	21,87	19,6	19,77	21,97	6,58	6,94	6,39	8,58	393,14	472,95	379,17	0,44	0,12	-27,25	1	-35,4	0	0	23,1	0,2	984,7
	Max	5	23	24,23	23,19	25,19	24,97	25,64	28,06	43,06	30,97	633,88	787,3	829,7	8,75	3,5	25,6	7	16,98	9,3	360	100	6,64	1042,5
	Med	3	11	22,7	21,29	22,25	23,02	15,36	15,33	16,82	15,99	465,73	547,2	463,32	0,45	1,45	2,75	3	-5,41	2,58	241,8	81,11	0,2	1016,6
	Avg	2,7	10,95	22,74	21,38	22,29	23,06	15	14,94	15,74	16,3	466,99	547,36	470,34	0,51	1,43	3,13	3,79	-8,26	2,79	232,05	76,91	0,38	1015,03
	Std	0,9	6,63	0,34	0,51	0,88	0,41	3,04	3,08	4,07	3,02	37,53	40,07	49,23	0,32	0,43	3,77	2,12	9,8	1,51	91,8	16,44	0,82	14,47
Winter	C5(19586)																							
High barometric pressure?	Min	1	0	21,64	19,64	19,58	21,51	9,87	7,8	6,63	13,71	422,61	486,98	462,24	0,43	0,12	-24,23	1	-35,14	0	0	22	0,1	982,7
	Max	6	23	24,26	23,31	25,29	24,77	38,47	37,9	64,88	39,5	784,88	1100,65	1058,72	8,59	2,99	28,1	7	17,76	9,78	360	100	10,07	1042,3
	Med	3	14	22,59	21,26	22,17	22,83	24,11	22,4	19,54	24,28	547,64	595,55	548,77	0,45	1,49	1,87	5	-1,4	2,2	227,2	81	0,2	1019,94
	Avg	2,71	12,96	22,68	21,27	22,22	22,91	24,06	22,08	19,59	24,26	558,13	603,52	563,33	0,56	1,46	1,59	4,28	-6,54	2,38	215,47	78,44	0,34	1018,55
	Std	1,28	6,94	0,45	0,56	0,72	0,63	3,37	4,31	7,33	3,44	55,82	54,73	61,63	0,38	0,36	2,47	2,04	14,01	1,27	92,71	17,65	0,87	13,3
Winter/spring	C6(15748)																							
Sauna	Min	1	0	21,2	19,62	23,34	21,2	9,24	7,95	6,74	16,28	404,31	486,15	390,38	0,44	0,12	-18,33	1	-30,09	0	0	20	0,14	994,01
	Max	5	23	23,69	22,92	32,26	24,13	38,09	38,38	70,3	37,25	716,89	738,1	817,18	2,14	2,44	26,27	7	19,8	7	359	100	8,52	1039,5
	Med	4	7	22,45	21,61	25	22,52	22,56	22,77	22,38	22,59	490,75	569,92	479,1	0,44	1,12	2,09	4	2,46	1,95	174,37	78,78	0,35	1016,02
	Avg	3,92	10,88	22,39	21,52	25,27	22,52	23,19	22,91	22,98	23,22	492,79	566,88	482,63	0,48	1,13	2,67	4,09	0,7	2,13	200,09	75,63	0,99	1016,53
	Std	1,26	8,98	0,46	0,66	1,33	0,59	3,31	3,58	5,53	3,61	46,94	48,52	55,9	0,12	0,42	4,01	2,45	9,68	1,26	93,59	18,61	1,78	9,65
Winter	C7(34304)																							
Generic day	Min	1	0	20,45	19,11	19,14	20,47	6,69	6,88	6,65	10,02	394,88	479,26	385,65	0,43	0,26	-21,87	1	-31,69	0	0	25,6	0,1	981,8
	Max	5	23	23,3	21,74	25,02	23,23	25,3	24,1	45,86	33,65	729,71	962,42	730,12	2,1	3,25	30,87	7	16,4	7,8	360	100	4,7	1042,1
	Med	2	11	21,82	20,28	21,74	21,8	15,77	16,4	15,68	16,74	452,35	537,76	454,98	0,45	1,61	3,27	4	-7,3	2,76	250,05	84,8	0,2	1014,97
	Avg	2,49	10,86	21,77	20,29	21,95	21,76	15,31	15,68	15,36	17,02	457,31	541,88	461,46	0,49	1,61	3,72	3,91	-9,14	3	233,73	79,6	0,33	1015,99
	Std	0,81	6,16	0,39	0,43	0,78	0,49	3,03	3	3,53	2,92	36,96	34,66	44	0,11	0,39	4,61	1,85	8,99	1,48	86,08	15,01	0,31	14,89
Winter/spring	C8(38272)																							
Generic day	Min	1	0	18,38	18,73	18,29	18,85	12,26	12,74	9,02	13,57	396,86	484,75	389,63	0,43	0,28	-27,72	1	-30,7	0	0	22	0,1	981,5
Ambient air pressure low	Max	5	23	22,71	21,48	24,19	22,96	32,5	32,87	73,05	32,78	747,75	789,4	719,18	12,79	3,1	24,56	7	15,88	7,9	360	100	8,77	1039,7
	Med	4	9	21,48	20,21	21,3	21,45	21,94	22,63	24,27	22,34	493,18	580,9	489,4	0,44	1,65	1,69	4	0,9	2,2	226,5	89	0,31	1008,42
	Avg	3,3	10,53	21,49	20,2	21,31	21,46	21,87	22,77	23,79	22,23	495,86	586,7	493,61	0,52	1,63	1,54	4,26	-0,46	2,43	219,51	83,9	0,4	1007,69
	Std	1,2	7,24	0,34	0,5	0,84	0,37	2,27	2,35	4,6	2,32	41,18	46,36	46,44	0,32	0,33	3,17	1,81	6,86	1,38	100,05	16,08	0,36	13,39
Spring	C9(30186)																							
Ambient humidity low	Min	1	0	21,52	20	19	21,49	10,98	10,85	10,2	15,41	393,77	477,11	383,98	0,43	0,12	-35,62	1	-35,4	0	0	18	0,11	983,5
	Max	6	23	24,15	23,36	24,63	24,5	32,67	31,63	73,12	29,91	657,88	721,44	775,8	2,4	3,07	28,74	7	21,6	9,9	359	100	11,5	1042,2
	Med	4	13	22,61	21,79	22,07	22,77	22,66	22,53	24,69	22,84	455,12	530,94	449,16	0,44	1,24	1,85	4	8,9	2,71	243,5	55,6	0,3	1018,4
	Avg	4,26	12,92	22,66	21,84	22,08	22,83	22,72	22,55	24,72	22,74	456,1	538,07	451,47	0,45	1,23	2,33	4,08	7,69	2,78	227,7	58,28	0,93	1015,97
	Std	1,07	5,97	0,41	0,56	0,78	0,47	2,52	2,57	4,64	2,58	32,84	39,13	37,65	0,08	0,4	4,82	2,03	7,71	1,35	91,04	23,59	2	10,94
Winter	C10(2884)																							
Visitors in the house?	Min	1	0	22,08	20,18	19,89	21,83	7,38	7,53	7,83	14,38	462,38	533,8	589,62	0,44	0,12	-14,86	4	-28,6	0	0	32	0,12	1006,3
CO2 up	Max	9	23	25,17	23,92	31,09	26,09	41,49	40,16	50,57	41,88	1225,41	1397,01	1482,57	1,64	2,9	20,48	7	15,3	5,8	358	98,6	1,1	1041,5
weekend	Med	3	14	23,15	21,53	23,02	23,14	25,1	23,64	22,66	25,68	836,9	683,05	881,51	0,56	1,54	1,54	6	-0,1	2,75	143,85	66	0,2	1022,56
	Avg	3,08	13,71	23,2	21,49	23,18	23,04	25,79	24,23	22,89	28,06	834,78	685,53	897,17	0,81	1,42	1,56	5,9	-4,21	2,56	168,78	65,17	0,22	1023,04
	Std	1,18	5,8	0,56	0,54	1,56	0,6	6,52	6,57	7,44	5,46	149,23	96,7	173,86	0,39	0,42	1,96	0,73	11,29	1,21	85,46	19,1	0,08	5,92
Spring	C11(22778)																							
Cool day	Min	1	0	21,24	20,02	20,23	21,24	24,01	21,76	12,29	23,68	398,5	477,42	386,05	0,44	0,12	-26,06	1	-20,44	0	0	27	0,1	996,3
	Max	9	23	23,84	23,51	27,68	24,12	46,38	43,2	77,92	42,89	790,44	794,23	799,78	1,98	2,72	20,4	7	23,4	7,81	360	100	14	1033,3
	Med	5	10	22,64	22,12	22,37	22,78	34,11	34,25	34,84	34,4	456,38	519,32	457,27	0,44	1,17	0,92	3	10,96	2,62	200,95	80,2	0,34	1008,76
	Avg	5,66	10,98	22,65	22,08	22,51	22,74	34,17	34,34	34,85	34,4	471,26	533,08	467,69	0,45	1,13	0,7	3,58	11,51	2,72	199,49	75,5	0,82	1010,82
	Std	1,42	6,9	0,5	0,56	0,84	0,53	3	3,14	3,84	2,9	56,32	46,21	54,32	0,04	0,37	3,65	2,09	4,41	1,2	71,21	17,62	1,31	7,23
Autumn	C12(18686)																							
Generic day	Min	2	0	21,13	20,4	21,04	21,32	39,04	38,7	39,81	40,29	447,62	511,31	411,57	0,44	0,12	-6,92	1	-18,14	0	0	50	0,2	997,4
	Max	9	23	23,45	23,03	24,81	23,61	56,95	59,83	86,17	57,38	821,56	844,85	791,5	1,81	2,72	28,05	7	20,5	7,5	360	99	24	1031
	Med	9	11	22,34	21,76	22,11	22,5	48,82	50,7	50,14	48,48	557,27	655,92	496,76	0,44	1,24	2,88	4	14,1	2,27	142,22	89	0,28	1009,4
	Avg	8,76	11,17	22,29	21,64	22,13	22,43	48,36	50,23	49,72	48,1	561,9	655,78	503,4	0,45	1,26	2,8	3,95	14,14	2,4	147,99	85,94	0,52	1009,65
	Std	0,43	6,96	0,52	0,54	0,69	0,54	2,76	3,36	3,47	2,79	66,43	53,27	52,2	0,04	0,42	2,89	1,76	2,41	1,32	83,49	10,85	0,9	7,79
Summer	C13(25936)																							
Generic day	Min	3	0	22,66	22,04	21,77	23,02	31,18	31,2	26,41	32,63	399,87	476,17	385,83	0,44	0,12	-26,36	1	-4,26	0	0	33	0,1	991,4
	Max	9	23	25,9	24,99	30,51	25,7	48,06	49,49	80,95	48,18	715,88	855,02	707,15	1,74	2,47	27,27	7	28,6	8	358	99	12,5	1024,1
	Med	6	13	24,18	23,73	23,68	24,3	40,9	41,01	41,94	40,68	468,9	528,98	448,91	0,44	0,81	2,72	4	15,32	2,26	180	77,9	0,4	1009,8
	Avg	6,76	12,14	24,12	23,65	23,71	24,27	40,43	40,81	42,16	40,43	476,96	540,23	458,5	0,53	0,76	2,45	4,19	15,56	2,31	175,37	74,92	0,63	1010,21
	Std	0,9	6,87	0,44	0,43	0,78	0,4	2,51	2,67	3,02	2,46	34,91	45,25	34,18	0,17	0,29	3,74	2	3,34	1,23	88,96	17,38	0,64	8,37
Summer	C14(31059)																							
Humid day	Min	3	0	22,13	21,76	21,38	22,37	37,15	36,95	37,78	36,75	410	482,09	397,33	0,43	0,12	-27,57	1	-4,27	0	0	33	0,1	993,5
	Max	9	23	25,12	24,6	27,22	25,39	54,45	56,96	76,74	53,64	843,41	842,13	803,43	1,78	2,43	22,81	7	28,6	7,08	360	100	9,4	1023,4
	Med	8	11	23,61	23,08	23,32	23,8	46,95	48,05	48,5	46,68	515,57	586,52	478,49	0,44	0,83	1,67	3	15,02	2,14	151	86,5	0,25	1012,47
	Avg	7,6	11,06	23,63	23,13	23,54	23,83	46,83	48,08	48,82	46,57	522,87	597,93	487,4	0,45	0,83	1,32	3,63	15,52	2,22	167,56	83,77	0,59	1011,18
	Std	0,78	7,09	0,49	0,55	0,95	0,47	2,17	2,63	3,78	2,16	52,96	67,92	51,63	0,03	0,38	4,09	2,04	2,78	1,03	84,03	12,33	0,7	7,19
Summer	C15(10109)																							
Low ambient humidity	Min	6	0	24,97	24,35	23,76	25,19	25,26	25,53	26,96	26,89	387,28	459,07	385,45	0,44	0,12	-6,23	1	6,1	0	0	17	0,2	1008,6
	Max	7	23	27,79	27,59	31,02	28,01	42,34	42,38	58,93	41,72	790,24	723,9	674,43	1,29	1,5	22,05	7	29,5	7,3	359	96	4,8	1020,1
	Med	6	14	26,32	25,99	25,85	26,48	36,29	35,49	37,12	36,21	484,94	521,89	464,65	0,44	0,21	2,8	4	18,08	2,33	100,4	65,2	0,32	1014,14
	Avg	6,48	12,6	26,32	26,01	25,73	26,49	35,55	34,91	36,9	35,51	482,01	530,36	467,88	0,45	0,29	3,31	4,22	18,11	2,44	123,31	63,23	0,38	1014,73
	Std	0,5	6,71	0,64	0,79	0,77	0,65	3,13	3,28	3,4	3,08	48,77	43,89	45,08	0,03	0,2	3,54	2,15	5,69	1,39	85,74	19,8	0,28	2,68
Summer	C16(10000)																							
Hot and sunny	Min	6	0	26,12	25,67	25,75	26,63	37,84	37,27	38,01	38,28	400,08	471,38	392,06	0,44	0,12	-16,35	1	13,5	0	0	26	0,2	1002,82
	Max	7	23	28,74	28,47	28,26	29,1	54,35	55,48	72,39	54,48	872,78	627,63	887,42	1,4	1,08	22,4	7	29,9	6,7	357	99	38,3	1017,1
	Med	7	12	27,76	27,48	27,24	27,9	47,83	47,48	49,04	46,99	463,58	515,12	465,63	0,45	0,12	3,49	6	22,76	2,2	140,89	71	1	1014,04
	Avg	6,98	11,59	27,75	27,43	27,12	27,9	47,07	46,85	48,07	46,63	479,68	520,82	475,22	0,45	0,16	3,75	5	22,73	2,25	158,44	68,65	1,1	1012,9
	Std	0,13	6,55	0,41	0,43	0,42	0,4	3,44	3,8	3,33	3,04	53,88	27,85	53,43	0,04	0,1	3,66	2,05	4	1,25	94,51	19,51	1,68	2,93
Summer	C17(10782)																							
Warm and humid day	Min	6	0	22,85	22,2	22,14	23,28	46,62	48,13	41,03	46,69	440,92	499,04	420,07	0,44	0,12	-18,36	1	12,9	0	0	44,6	0,1	991,7
Low ambient air pressure	Max	8	23	25,63	25,28	30,73	25,96	66,47	69	89,73	63,07	921,93	855,29	807,06	2,48	2,38	22,59	7	25,9	6,2	359	98	40,9	1022
	Med	8	11	24,32	23,66	23,9	24,55	55,43	58,18	55,31	54,68	513,64	577,58	484,61	0,44	0,78	3,14	5	18,92	1,92	162,4	86,1	0,3	1009,4
	Avg	7,65	11,37	24,35	23,73	24,2	24,56	55,32	58,19	55,38	54,57	523,71	583,05	497,42	0,45	0,87	3,87	4,5	19,24	2,08	181,27	83,2	1,07	1008,8
	Std	0,48	6,77	0,51	0,71	1,31	0,49	2,8	3,39	4	2,63	53,85	52,1	52,98	0,06	0,38	5,03	2,08	2,79	1,07	71,67	12,72	2,51	6,9
Summer	C18(13508)																							
Generic day	Min	6	0	24,31	23,9	23,56	24,59	36,7	36,16	37,81	36,39	432,36	482,31	406,78	0,44	0,12	-27,54	1	11,68	0	0	37	0,1	1002,5
	Max	8	23	27,35	26,78	28,35	27,29	54,33	57,53	90,32	54,9	792,81	801,47	748,98	2,2	1,71	22,71	7	26,7	7	359	100	3,7	1024
	Med	7	12	25,92	25,6	25,7	26,09	44,71	44,56	46,35	44,49	492	537,41	477,87	0,44	0,28	2,6	4	19,64	2,32	156,5	73,7	0,5	1014,2
	Avg	6,77	11,53	25,87	25,52	25,61	26,04	45,5	45,73	46,45	45,17	505,94	543,32	484,07	0,58	0,33	2,39	3,9	19,63	2,32	174,17	72,05	0,93	1014,3
	Std	0,46	7,38	0,48	0,59	0,73	0,48	3,35	3,66	3,09	3,11	51	39,25	45,06	0,4	0,21	3,85	1,82	3,35	1,1	91,17	15,63	0,86	4,78
Summer	C19(7592)																							
Hot day	Min	7	0	24,93	24,56	23,92	25,17	50,41	51,97	48,09	50,21	440,79	494,59	430,14	0,44	0,12	-22,11	2	17,7	0	0	47	0,2	1002,6
	Max	8	23	27,98	27,8	28,83	28,42	66,64	71,3	73,85	64,09	705,71	838,96	725,78	1,35	1,62	22,88	7	28,8	5,3	349	99	34,9	1017,8
	Med	7	12	26,34	26,02	25,86	26,58	58,98	61,71	59,58	58,39	492,86	542,54	479,08	0,44	0,18	3,69	4	22,5	1,85	134,1	81,6	0,5	1008,2
	Avg	7,03	12,04	26,45	26,12	25,9	26,68	58,55	61,32	59,01	57,73	505,51	564,07	489,84	0,45	0,33	5,55	4,13	22,59	1,92	172,16	81,24	1,4	1011,18
	Std	0,16	6,56	0,66	0,7	0,51	0,66	2,8	3,43	3,08	2,57	42,41	70,05	42,54	0,04	0,27	5,66	1,44	2,91	1,07	92,88	13,13	2,24	4,79

APPENDIX 6. House E cluster interpretation sheet

SITUATION	Cluster	Months	Hours	Tem_OH_A	Tem_MH_	Tem_PH_	Hum_OH_	Hum_MH_	Hum_PH_	CO2_OH_	CO2_MH_	CO_OH_	Dif_Pre_Ast_	weekday	t2m	ws_10min	wd_10min	rh	ri_10min	p_sea
	C1(32287)																			
Winter	Min	1	0	20,71	20	20,94	7,19	7,77	5,69	346,31	429,26	0,18	-19,36	1	-29,5	0	0	32	0,2	984,7
Generic morning	Max	12	23	24,11	23,13	24,13	20,07	19,89	38,53	741,74	887,71	98,01	48,03	7	4,2	7,7	360	100	4,7	1042,5
Indoor air dry	Med	2	10	22,53	21,74	22,7	13,78	13,27	15,43	519,34	592,19	0,34	-2,31	4	-11,3	1,95	162	90	NaN	1016,5
	Avg	3,22	10,48	22,56	21,63	22,67	13,42	13,36	15,24	519,43	597,82	0,57	-2,13	4,21	-11,7	2,23	197,44	85,8	NaN	1013,96
	Std	3,72	6,37	0,48	0,49	0,44	2,46	2,29	2,98	58,04	96,87	1,92	3,56	2	7,2	1,36	98,27	12,57	NaN	13,2
	C2(11647)																			
Winter/spring	Min	1	0	20,37	19,72	20,32	7,46	7,63	6,04	359,68	429,1	0,14	-20,59	1	-28,8	0	0	34	0,1	989,4
House empty	Max	12	23	22,82	22,02	22,2	25,2	25,78	39,32	607,95	860,75	14,29	13,16	7	8,2	7,8	360	100	1,87	1032,36
	Med	5	8	21,34	20,67	21,37	18,62	18,82	19,51	372,61	448,77	0,33	-3,12	3	1,85	2,73	177,8	74,3	NaN	1018,3
	Avg	4,28	9,85	21,37	20,79	21,33	18,59	18,56	19,56	436,34	482,57	0,34	-2,78	3,79	0,81	2,94	186,3	70,41	NaN	1016,56
	Std	1,3	6,88	0,51	0,43	0,51	1,73	2,59	2,23	82,34	80,41	0,19	1,77	2,26	5,02	1,33	126,47	17,19	NaN	9,67
	C3(18315)																			
Winter	Min	1	0	21,59	20,7	21,6	9,78	9,73	10,94	524,8	633,05	0,2	-16,16	1	-35,4	0	0	46,4	0,19	984,6
House occupied	Max	12	23	25,43	23,46	24,65	23,13	24,35	56,2	1006,85	1144,28	68,94	48,03	7	2,78	6,7	360	100	4,98	1039,7
Some CO spikes	Med	2	11	23,1	22,32	22,79	15,53	16,94	18,56	749,55	874,58	0,49	-1,29	4	-16,74	2	174	85	NaN	1026,22
	Avg	4,29	11,65	23,13	22,29	22,79	15,77	17,07	19,42	749,63	881,74	0,85	-1,07	4,26	-16,48	2,08	210,74	84,27	NaN	1022,86
	Std	4,14	6,47	0,71	0,47	0,51	1,91	2,17	3,66	73,18	97,47	3,19	3,54	1,88	8,42	1,03	101,07	8,66	NaN	12,4
	C4(30658)																			
Winter/spring	Min	1	0	20,86	21,03	20,54	12,07	10,79	16,06	495,73	543,54	0,16	-15,26	1	-28,1	0	0	23	0,1	984,24
Generic morning	Max	12	23	24,62	23,95	26,31	35,41	33,99	91,2	911,53	1070,56	7	46,34	7	20,2	9,3	360	100	7,79	1040,86
	Med	4	9	22,83	22,54	22,5	23,1	24,29	26,5	638	751,09	0,32	-2,56	4	1,62	2,58	225,7	88	NaN	1011,58
	Avg	3,45	10,37	22,86	22,61	22,54	23,34	24,42	27,49	643,34	759,09	0,35	-2,28	3,86	1,82	2,77	227,09	82,34	NaN	1011,26
	Std	1,25	7,02	0,46	0,43	0,47	2,56	2,67	4,81	66,47	82,24	0,18	2,7	2,1	4,3	1,51	79,4	16,89	NaN	12,93
	C5(27313)																			
Winter	Min	1	0	22,41	20,58	21,84	9,33	8,86	10,23	364,25	415,8	0,17	-17,19	1	-29,3	0	0	30	0,1	981,98
Fireplace in use	Max	12	23	25,9	24,6	26,15	25,31	23,12	58,18	795,15	841,72	74,75	48,03	7	9,68	6,9	360	100	2,95	1041,6
Some CO spikes	Med	1	13	23,8	22,53	23,13	14,37	14,89	17,13	547,14	637,41	0,34	-1,65	4	-9,11	2,4	172	89,5	NaN	1011,51
	Avg	2,39	12,05	23,86	22,57	23,24	14,37	14,88	17,82	547,76	621,01	0,99	-1,21	3,68	-10,18	2,49	205,12	84,71	NaN	1011,68
	Std	2,6	7,09	0,59	0,5	0,54	2,22	2,31	4,1	57,96	85,17	4,03	4,6	1,92	6,51	1,22	100,72	14,99	NaN	11,46
	C6(27869)																			
Spring	Min	1	0	23,08	21,81	21,88	12,95	12,73	13,98	511,66	542,97	0,13	-15,28	1	-21,1	0	0	19	0,1	984,6
Fireplace in use	Max	12	23	26,53	25,9	28,36	33,35	32,45	91,84	942,01	1096,13	16,93	48	7	20,23	9,9	359	100	7,9	1040,8
Slightly more frequent in weekends	Med	4	14	24,45	23,75	23,25	22,19	23,93	27,92	633,4	757,25	0,32	-1,01	4	2,83	2,27	199,2	73,4	NaN	1011,84
	Avg	3,87	12,52	24,5	23,86	23,39	22,51	24,05	29,06	644,33	770,27	0,39	-0,77	4,28	3,48	2,53	211,25	72,54	NaN	1011,76
	Std	1,19	7,71	0,63	0,6	0,69	2,9	3,13	6,52	70,56	94,62	0,24	2,32	1,85	6,37	1,45	90,06	20,78	NaN	11,54
	C7(25058)																			
Winter	Min	1	0	21,12	20,48	21,55	13,56	14,92	14,26	654,1	805,49	0,13	-13,08	1	-35,4	0	0	59,6	0,1	984,5
House occupied	Max	12	23	25,54	24,03	24,83	32,14	34,49	59,53	1401,42	1476,29	3,9	31,33	7	6	8,1	345	100	1,7	1039,7
	Med	2	9	22,74	22,4	22,69	19,8	21,25	22,75	875,84	1095,44	0,44	-1,03	4	-11	2,2	199,2	84,8	NaN	1026,99
	Avg	2,79	11,09	22,82	22,41	22,77	19,63	21,73	22,7	887,25	1099,47	0,58	-1,01	4,1	-13,31	2,47	213,35	84,57	NaN	1021,71
	Std	2,33	8,35	0,72	0,56	0,66	3,48	3,27	4,99	89,97	98,25	0,36	1,73	2,06	10,64	1,38	85,53	8,55	NaN	15,08
	C8(23393)																			
Winter/spring	Min	1	0	21,64	20,3	20,74	13	11,97	14,32	368,54	421,87	0,15	-16,7	1	-16,59	0	0	22,3	0,1	981,5
Generic day/afternoon	Max	12	23	24,48	23,98	24,55	28,96	30,73	57,97	655,32	775,32	76,61	48,03	7	20,8	9,4	360	100	6,78	1041,26
	Med	3	13	22,68	22,14	22,33	19,97	20,25	22,48	514,66	555,49	0,32	-2,16	4	0,83	2,94	249,8	84,4	NaN	1006,55
	Avg	2,93	12,42	22,8	22,19	22,34	20,16	20,44	23,21	505,07	555,44	0,42	-1,37	4,02	2,34	3,01	231,54	73,34	NaN	1009,19
	Std	1,48	5,35	0,49	0,65	0,41	2,01	2,46	3,59	57,61	69,02	1,05	5,36	2,03	5,98	1,44	90,56	23,61	NaN	12,4
	C9(12692)																			
Spring	Min	1	0	22,37	21,72	22,14	16,88	17,78	23,08	593,51	734,67	0,16	-15,43	1	-27,5	0	0	35	0,1	988,62
House occupied	Max	12	23	26,4	25,88	27,41	42,17	41,75	94,13	1408,47	1568,32	3,28	9,38	7	19,56	7	359	100	7,64	1036,8
	Med	5	16	24,18	24,15	22,96	31,57	34,1	38,06	814,18	1005,36	0,32	0,24	4	8,84	2,5	204,1	86,6	0,37	1007,6
	Avg	4,75	12,17	24,29	24,1	23,09	31,08	33,26	38,6	845	1018,74	0,39	0,21	3,75	7,1	2,56	203,86	82,5	0,67	1008,35
	Std	1,29	8,71	0,74	0,63	0,57	4,08	4,09	7,25	131,57	132,45	0,21	1,85	1,87	6,82	1,08	67,97	15,22	0,79	7,63
	C10(23130)																			
Spring	Min	1	0	22,72	22,02	21,58	22,19	16,76	29,52	378,87	408,81	0,16	-13,78	1	-10,15	0	0	27	0,1	988
Generic day	Max	9	23	26,46	25,77	26,56	41,62	43,53	93	786,01	885,82	3,85	39,74	7	25,4	7,4	360	100	14	1040,9
	Med	6	12	24,14	24,09	22,99	33,89	34,2	38,6	553,8	599,91	0,31	1,29	3	13,13	2,8	182,5	68	0,37	1012,5
	Avg	6,03	11,66	24,19	24,14	23,04	33,91	34,3	39,38	550,4	603,5	0,32	1,54	3,24	13,67	2,78	191,29	67,61	0,64	1013,35
	Std	1,14	6,22	0,63	0,58	0,57	2,85	2,96	4,8	73,62	92,5	0,1	2,42	1,8	4,24	1,39	96,5	18,75	0,83	6,63
	C11(17659)																			
Spring	Min	1	0	22,63	21,8	21,77	16,2	13,45	16,25	353,75	408,86	0,16	-9,03	1	-10,63	0	0	18	0,1	988,62
Ambient humidity low	Max	6	23	26,47	25,47	25,77	33,02	33,29	65,57	671,18	769,71	7,06	48,03	7	23,4	7,9	359	100	11,5	1040,9
	Med	5	12	24,38	24	22,87	24,84	25,63	29,39	491,73	532,69	0,32	1,98	4	12,2	3,04	219	43,6	0,55	1016,12
	Avg	4,83	11,94	24,34	23,95	22,96	24,98	25,18	28,97	483,56	529,85	0,34	3,47	4,09	11,95	3,06	215,18	50,58	1,83	1014,52
	Std	0,8	4,55	0,55	0,64	0,51	3,25	3,35	3,69	63,98	65,95	0,24	7,25	2,17	5,68	1,28	86,54	22,15	2,85	8,32
Summer	C12(7120)																			
Warm day	Min	5	0	25,23	25,32	23,6	27,35	26,97	29,87	372,79	393,64	0,15	-2,8	1	6,1	0	0	17	0,2	1003,65
Ambient humidity low	Max	7	23	28,24	28,53	28,03	42,29	43,62	91,3	712,25	827,13	2,01	23,27	7	29,5	7,3	359	95	4,8	1022,49
	Med	6	13	26,81	26,87	25,71	33,13	34,59	37,87	471,46	538,9	0,31	4,61	5	20,56	2,3	107,4	56,6	0,34	1014,15
	Avg	6	12,02	26,74	26,94	25,65	33,65	34,43	38,94	489,11	553,21	0,31	4,61	5,1	19,52	2,4	128,09	56,53	0,38	1014,14
	Std	0,07	7,2	0,72	0,73	0,63	2,2	3,55	5,17	72,23	99,24	0,08	1,88	1,5	6,25	1,5	88,63	22,64	0,27	2,23
	C13(29775)																			
Summer	Min	5	0	22,73	22,63	21,69	39,23	39,6	43,33	392,75	448,34	0,1	-5,4	1	7,4	0	0	42	0,2	991,4
Cool day	Max	9	23	25,52	25,29	24,2	57,81	59,55	85,45	774,32	886,52	2,97	17,54	7	23,1	6,2	360	100	24	1024,1
	Med	8	9	24,15	24,14	22,95	46,78	48,34	52,76	564,75	617,65	0,3	0,74	3	14,61	2,11	155,8	89,1	0,27	1009,8
	Avg	7,6	10,96	24,16	24,09	22,97	47,12	48,37	53,73	567,97	622,72	0,31	0,9	3,66	15,07	2,17	164	85,99	0,6	1009,49
	Std	1,04	7,78	0,47	0,44	0,42	3,37	3,59	4,42	49	68,92	0,11	1,61	2,07	2,63	1,04	78,1	10,56	0,8	8,44
	C14(22124)																			
Summer	Min	6	0	24,44	24,18	22,87	31,27	31,78	34,73	346,52	388,18	0,1	-5,65	1	10,09	0	0	31	0,1	993,9
Generic day	Max	9	23	26,43	26,72	26,85	48,77	51,83	82,46	847,45	980,39	2,05	19,06	7	28,6	8	359	100	12,5	1024
	Med	7	13	25,17	25,19	24,15	43,38	43,77	47,73	476,62	495,49	0,3	3,11	4	16,92	2,23	158	74,37	0,44	1014,1
	Avg	6,81	12,19	25,26	25,27	24,21	42,73	43,03	47,32	481,51	520,98	0,3	3,01	4,2	17,59	2,37	175,33	71,64	0,66	1012,96
	Std	0,58	6,99	0,41	0,51	0,54	2,97	3,83	4,32	100,03	107,25	0,09	1,7	1,76	3,94	1,25	101,69	17,73	0,64	5,61
	C15(27477)																			
Summer	Min	6	0	22,96	22,54	21,53	34,81	35,97	40,77	351,3	396,96	0,1	-8,81	1	7,5	0	0	38	0,1	992,6
House empty	Max	9	23	25,37	25,13	23,85	52,56	55,65	75,76	602,32	681,02	10,4	22,25	7	27,5	7,5	359	99	9,4	1023,3
	Med	8	13	23,92	23,94	22,7	44,63	45,42	50,28	445,4	476,73	0,3	1,73	4	16,28	2,39	144	74,5	0,27	1011,9
	Avg	7,65	12,48	23,98	23,92	22,74	44,62	45,4	50,12	444,55	480,67	0,31	1,98	3,84	16,35	2,46	157,77	74,02	0,6	1010,25
	Std	0,99	5,02	0,47	0,41	0,38	3,22	3,69	3,67	53,58	48,34	0,1	1,98	2	2,74	1,17	79,39	14,62	0,78	8,16
Summer	C16(19488)																			
Hot day	Min	6	0	25,97	25,82	24,66	38,08	39,8	46,63	349,19	383,73	0,14	-1,95	1	12,9	0	0	26	0,2	1002,6
Ambient humidity up	Max	7	23	28,28	28,63	28,19	60,47	64,06	91,42	844,87	836,37	2,15	10,22	7	29,9	7	359	99	38,3	1017,8
House empty	Med	7	13	26,87	26,97	25,72	50,1	52,3	54,93	422,16	448,92	0,31	5	5	22,15	2,15	149	76,4	0,9	1013,44
	Avg	6,99	12,15	26,94	27,04	25,77	51,08	52,24	55,51	441,23	479,32	0,31	5,01	4,41	22,13	2,16	167,99	73,15	0,99	1012,41
	Std	0,07	6,91	0,47	0,57	0,36	3,52	4,97	4,85	69,97	81,89	0,09	1,17	2,11	3,57	1,14	92,51	17,6	1,4	3,85
Summer	C17(6907)																			
Night	Min	6	0	22,64	22,98	21,72	41,93	41,91	44,2	641,51	662,04	0,15	-5,49	1	8,4	0	0	68	0,1	1001,3
House mostly occupied	Max	8	23	25,09	25,36	25,2	56,89	58,68	83	1015,63	1121,02	2,16	5,75	7	21,63	6,3	360	99	3,4	1022,7
	Med	8	5	23,6	23,75	22,41	49,68	51,19	56,65	803,29	853,41	0,31	0,88	5	14,26	1,78	154	90,3	0,26	1017,5
	Avg	7,73	8,88	23,73	23,89	22,58	49,1	50,74	56,53	807,57	860,81	0,32	0,89	4,69	14,43	1,88	154,61	89,62	0,33	1015,6
	Std	0,68	8,15	0,51	0,54	0,54	3,72	4,16	5,06	75,62	80,78	0,11	1,02	1,55	2,39	1,17	79,76	7,29	0,28	6,29
Summer	C18(8119)																			
Humid day	Min	7	0	24,48	24,25	23,21	44,07	42,72	49,92	359,01	400,88	0,07	-4,98	1	15	0	0	46,8	0,2	1002,5
House empty	Max	8	23	26,57	27	25,62	58,39	62,8	83,55	608,78	633,16	2,04	12,83	7	28,8	6,2	359	99	40,9	1017,6
	Med	7	10	25,67	25,66	24,54	53,75	56	58,54	400,07	444,86	0,3	3,57	4	20,9	1,88	150	81	0,83	1009,7
	Avg	7,2	10,95	25,64	25,68	24,52	53,04	55,76	57,87	424,31	462,12	0,3	3,47	4,14	21,36	2,06	171,12	80	1,85	1010,69
	Std	0,4	6,52	0,44	0,49	0,51	2,84	3,4	3,42	58,02	52,68	0,07	1,38	1,59	2,96	1,17	79,74	13,94	3,4	3,69
	C19(8601)																			
Autumn	Min	8	0	21,13	21,43	20,32	36,12	35,06	39,09	363,57	415,84	0,16	-6,59	1	5,4	0	0	47	0,2	997,61
	Max	9	23	24,04	23,82	22,67	54,85	54,29	66,47	622,54	766,37	3,95	26,67	7	19,6	5,9	358	99	5,8	1022,8
	Med	9	10	22,61	22,5	21,63	46,42	47,76	50,13	382,28	443,61	0,31	-0,43	5	13,1	1,9	152	93	0,4	1009,3
	Avg	8,85	10,71	22,44	22,48	21,53	45,84	47,27	49,99	432,84	492,23	0,31	-0,17	4,58	12,27	2,14	169,46	89,89	0,62	1009,8
	Std	0,35	7	0,61	0,51	0,59	3,47	3,69	3,54	80,02	91,55	0,11	2,08	2,25	2,76	1,19	96,65	9,25	0,64	6,55

APPENDIX 7. Legend for colour formatting of the cluster means in APPENDIXES 2-6.



	House parameters						Weather parameters			
Relative value	T	RH	CO2	CO	IVPre	DifPre	T	ws_10min	RH	Rain
Low	1	1	1	1	1	-4	1	1	1	1
	2	2	2	2	2	-3	2	2	2	2
	3	3	3	3	3	-2	3	3	3	3
	4	4	4	4	4	-1	4	4	4	4
Medium	5	5	5	5	5	0	5	5	5	5
	6	6	6	6	6	1	6	6	6	6
	7	7	7	7	7	2	7	7	7	7
	8	8	8	8	8	3	8	8	8	8
High	9	9	9	9	9	4	9	9	9	9