Validation of a new method to estimate energy use for space heating and hot water production from low-resolution heat meter data

Daniel Leiria[1*], Hicham Johra[1], Evangelos Belias[2,3], Davide Quaggiotto[4], Angelo Zarrella[4], Anna Marszal-Pomianowska[1], Michal Pomianowski[1]
[1]Department of the Built Environment – Aalborg University, Aalborg, Denmark
[2] ESTIA SA, EPFL Innovation Park, Lausanne, Switzerland
[3]Human-Oriented Built Environment Lab, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[4]Department of Industrial Engineering – University of Padova, Padua, Italy
* *corresponding author: dle@build.aau.dk*

Abstract

One of the initiatives to reach the European decarbonization goal is the roll-out of smart heating meters in the building stock. However, these meters often record the total energy usage with only hourly resolution, without distinguishing between space heating (SH) and domestic hot water (DHW) production. To tackle this limitation, this paper presents the validation of a new methodology to estimate the SH and DHW from total measurements in different building types in three countries (Denmark, Switzerland, and Italy). The method employs a combined smoothing algorithm with a support vector regression (SVR) to estimate the different heating uses. The estimation results are compared with the different countries' DHW compliance calculations. The comparison showed that the compliance calculations outperformed this method by considering the validation dataset characteristics.

Introduction

The society is being pressed to become more sustainable. These pressing sustainable challenges are due to the global climate change, pollution issues, and fossil fuel supply curtailment. A "green" transition must occur, especially for the building sector. According to European Commission (2022a), in the European Union (EU), its building sector has an estimated share of 40% of the total energy end-usage, where 79% of it is for space heating (SH) and domestic hot water (DHW) production alone (European Commission (2022b)). It is estimated that 97% of the existing buildings in the EU must be renovated to achieve its 2050 environmental goals (BPIE (2017)). This estimation is based solely on the energy performance certificates (EPC) issued in the different EU member-states. An EPC results from several calculations made by an expert to estimate a building's energy usage and efficiency. These calculations are based on different measurements, assumptions, and standards depending on the country where the building is located. The objective behind these certificates is to raise awareness of energy efficiency among the owners and tenants, promote the refurbishment of the building, and assess the overall country building stock (Iribar et al. (2021); Gonzalez-Caceres et al. (2022)). Even though the EPCs are promising, they usually show a significant difference between the measured and estimated energy usage. This difference is known as the performance gap (Cozza et al. (2021)) and has been studied in several EU countries (Gram-Hanssen and Hansen (2016)). In order to solve this issue, one of the proposed solutions is the usage of actual measurements as additional information for performing the EPC calculations. This manuscript focuses on how the actual building heating measurements can be used to estimate the SH and DHW shares and compare them with the countries' current compliances to estimate the yearly DHW consumption in the EPCs. The countries studied are Denmark, Switzerland, and Italy. Therefore, each country's effort in using energy data to decrease the performance gap is explained below.

As a front-runner country, Denmark is making a great endeavor to install smart heat meters in buildings connected to the district heating (DH) network (Johra et al. (2020)). The resulting data from the meters are the aggregated heating usage (SH and DHW), water consumption, and temperature-weighted volume consumption, resulting in monotonically increasing measurements (Kristensen and Petersen (2021)). Also, these meters typically have hourly measurements, and the data are easily accessible by the utility companies. Although this initiative is a substantial move toward achieving the energetic target set by Denmark (Danish Climate Policies | Energistyrelsen (2022)), it has a downside regarding its data collection. In most buildings, only one device is installed, which collects the total heat usage without differentiating between the energy used for SH or DHW production. Because these two heat uses depend on different factors, it is crucial to disaggregate them to understand better the building and occupancy heat demand (Gram-Hanssen (2014)).

Even though the DHW in Swiss nZEB accounts for 50-70% of the total heat consumption, its monitoring is not required by the local regulations or the EPC (Office fédéral de l'énergie OFEN SuisseEnergie (2016); Flourentzou and Pereira (2021)). On the contrary, in Switzerland, it is common for the buildings to be equipped with one heat meter that measures the total heat consumption, both for SH and DHW, making it challenging to identify the heat required for the DHW or SH production (Flourentzou and Pereira (2021)).

In Italy, in the last years, the Government promoted several energy conservation measures for the building envelope with related incentives due to the prevalence of old buildings. So, as a matter of fact, the energy consumption of the building stock will change in the future (also considering the climate change effect). Consequently, the district heating networks in the main cities of northern Italy, which were built several decades ago and are operating at high temperatures (70-80°C), need to be revised in terms of both production and operating conditions. An example of such an intervention is studied by Vivian, Quaggiotto and Zarrella (2020). The heating and DHW disaggregated profiles will help design and manage these improvements efficiently. In addition, the recent concept of the district heating network integrated with other renewable energy technologies (e.g., heat pumps) in new building districts is a good opportunity (Bordignon et al. (2022)). Also, in this case, the disaggregated profiles can help design and set suitable control strategies to increase energy efficiency.

Another aspect to consider on the importance of knowing these energy shares is regarding refurbishment initiatives. In Pomianowski et al. (2020), the authors argue that global building regulations have stricter SH efficiency rules while overlooking DHW consumption. Therefore, the new buildings, also known as low-energy buildings, have a much higher DHW share due to the continuous decrease of SH usage over the years and the higher levels of comfort concerning heating practices demanded by the residents.

Thus, a better assessment of the thermal appliances can be achieved by disaggregating the energy used in buildings. This contributes to a more detailed understanding and control on the demand side and promotes better decision-making strategies regarding heat production and distribution.

Background

The disaggregation of time-series has been studied since the 1980s regarding electrical appliances metering (Zeifman and Roth (2011)). However, the research has been shifting towards heating meter data. One of the first articles to explore this type of data is Bacher et al. (2016), which presents a statistical methodology to estimate the SH from 10-min resolution total heat measurements. This method is based on the premise that SH demand variates accordingly to the smooth external temperature fluctuations. At the same time, DHW usage fluctuates sporadically with higher peaks due to its short-time hot water draw-off events. The method predicts the SH by applying a kernel smoother to the total measurements, where all values above a defined threshold are due to DHW usage. Although promising, the method needs validation with separated space heating and DHW usage measurements. Also, the need for high-resolution data (10-minutes resolution) to detect the DHW peaks is uncommon to find in the typical installed smart meters.

Unlike the above method, more straightforward methods were developed to disaggregate heating datasets. The articles Lien, Ivanko and Sartori (2020), and Ivanko, Sørensen and Nord (2021) propose different methods to decompose SH and DHW usage based on discovering the DHW profiles when the total heating is assumed to be equal to the DHW usage only (no SH demand). Additionally, considering the relationship between SH demand and the external temperature, the methods were validated with several Norwegian buildings (apartments and hotels) and compared with other existing methods. Also worth mentioning regarding Lien, Ivanko and Sartori (2020) is that besides presenting their developed methodology, they also compared their results with several Norwegian reference data.

In Marszal-Pomianowska et al. (2019), another technique is proposed by assuming that the total heat measurements are equal to the DHW usage during Summer (no SH demand). Their novel approach does not aim to disaggregate the data but to predict the dwelling's daily DHW usage profile. This load-profiling technique seeks to throw light on the customers' DHW practices and how their behavior affects the DH supplier.

In Hedegaard, Kristensen, and Petersen (2021), the weekly SH and DHW usage profiles are predicted using calibrated grey-box models. This method is likely to be the most reliable and accurate of the ones presented in this review, and the authors claim that the model's accuracy can be improved even further. Also worth mentioning is Alzaatreh et al. (2018). A pattern recognition technique was developed and tested in this research work to separate SH measurements from other appliances in two UK single-family dwellings.

This manuscript aims to present the results from the validation of a novel disaggregation methodology described in Leiria et al. (under review). The validation process is constituted by applying the method in three smart heat meter datasets. These datasets are different in terms of measurements resolution (i.e., number of decimal digits), measurements scale (i.e., energy usage in a single apartment or a block of apartments), building type (i.e., residential or commercial buildings), heating systems (i.e., DHW production with or without storage tank) and different countries (i.e., Denmark, Switzerland, and Italy). The current study compares the DHW estimation by the disaggregation methodology with the actual measurements and the DHW compliance calculations of each country.

Following the *Introduction*, the section *Study Case* presents the different validation datasets. In *Methodology*, the applied disaggregation method is explained. The results from the validation are examined in the section *Results and Discussion*. The manuscript closes with *Conclusion* and *Suggestions for Further Work*.

Study Case

For the methodology's validation, three heating datasets are used. All datasets have separated energy measurements of SH, DHW, and the aggregated sum of both (total heat). The differences between datasets are the following.

Danish dataset

This dataset is constituted of 28 single-family apartments. All apartments are from a social housing complex in Aalborg, Denmark. The complex was progressively refurbished to the Nearly Zero-Energy Buildings (NZEB) standard from 2012 to 2020. The interior of the apartments was fully remodeled, and the new SH installation includes radiators in all rooms and kitchens and underfloor heating in the bathrooms and hallways. The heat for SH and DHW is produced at the building block level and distributed to each apartment. Apartments are equipped with single SH and total heat usage meters, and the DHW is calculated through the difference between measurements from the meters. The heated area of the dwellings is between 97 and 112 m^2.

The local weather data (hourly outdoor temperature and the global radiation) is retrieved from the Danish Meteorologic Institute website (Dansk Meterologisk Institut (2022)). The chosen weather station is Tylstrup, the nearest available station to Aalborg.

In this work, the data pre-processing consisted in detecting the number of missing and negative measurements and removing them. In the 28 dwellings dataset (187 123 measurements) with approximately nine months of monitoring for each dwelling, there are 46 661 missing hours (~25% of the dataset). The household with the lowest missing measurements has approximately 3% missing data. Some households have up to 43% of missing data. Regarding negative measurements (incorrect values), few apartments have those. In total, these values only represent 0.013% of the original dataset.

Swiss dataset

This dataset is constituted of an apartment building located in Vevey, Switzerland. The building was built in 20120 and deeply refurbished to reach NZEB standards during 2018-2019. The local district heating network supplies heat for SH and DHW. A heat meter measures the total heat provided by the district heating network. A second, a Flexim ultrasonic portable flowmeter (Fluxus F601), was used to measure the heat consumption for the DHW.

The hourly outdoor temperature and the global radiation data were collected from the Swiss Federal Office of Meteorology and Climatology-MeteoSwiss (Swiss Federal Office of Meteorology and Climatology-MeteoSwiss (2022)). The "Vevey" station was used for the weather data as it was the nearest available. The data were pre-processed in order to identify the missing and negative values and remove them. In total, for 2020, there were five months of available valid data.

Italian dataset

The selected building dataset consists of a theatre and a rehab institution connected to the district heating network of Verona Centro Città, serviced by AGSM. This network supplies heat to residential, tertiary, and industrial customers, operating at constant supply temperature and variable flow rate. Overall there are 247 user substations, but just for 2 (theater and rehab institution) of these 247, the separate monitoring on the use of SH or DHW was provided. The measures all correspond to the primary circuit of the heat exchanger installed at each user substation. The measuring devices installed are all ultrasonic compact energy meters suitable for measuring the energy consumption of district heating systems. The principle of operation of these meters is static and based on the transit time measurement. In particular, ultrasonic meters are characterized by the absence of moving parts, thus preventing mechanical wear of the metering components, low-pressure losses, low start flowrate, and good tolerance to suspended particulates in the water flow. On the whole, the ultrasound principle assures stable and accurate measuring results. The measurement period is from December 1, 2021, to January 31, 2022, for the rehab institution and from January 11 to January 31, 2022, for the theater. The resolution of the measured data is a 15-minute time step.

The local weather data (global solar radiation and air temperature with hourly time step) has been provided by the Arpav Meteorological Institute of Teolo.

Methodology

The methodology starts with the premise that the SH system runs continuously during the heating season while the DHW usage is sporadic throughout the day. Hence, during a day (which has around 24 recorded heating measurements), only a few of those consist of combined SH and DHW usage ($E_{Total} = E_{SH} + E_{DHW}$). The other recorded data points are SH usage alone ($E_{Total} = E_{SH}$). Following this premise, the method has two stages. The first is to segregate the data points with and without DHW production. The second is to estimate the SH share ($E_{SH,estim}$) in the points identified with DHW usage. From the SH estimation, the DHW is calculated through Equation 1:

$$E_{DHW,estim} = E_{Total} - E_{SH,estim} \quad (1)$$

The estimation results are compared with the separated measurements (SH and DHW usage) for each dataset. The DHW values obtained by the disaggregation methodology with the DHW prediction from the different countries' compliance calculations. The disaggregation methodology is disclosed in more detail below. Furthermore, the algorithm presented in this work is coded with the software *Rstudio* (RStudio (2022)).

Energy separation

This first part of the method starts from the same premise as Bacher et al. (2016) that moderate variations of outdoor temperature during the day combined with the inertia of the building environment contribute to smooth SH daily fluctuations. Hence, all peaks recorded by the meters can be accounted for DHW usage. Therefore, the method detects all daily highest points (E_{Total}) and identifies them containing DHW and SH usage ($E_{Total} = E_{SH} + E_{DHW}$). For each day, the method assumes the seven-highest recorded values as DHW usage, while the other measurements are considered SH alone. It is also assumed a sleeping period from 1:00 – 4:00 hours every day. Thus, there is no DHW demand during the sleeping period, and the high values recorded are because of the SH system operation. In Figure 1, one can see a schematic representation of the separation method.

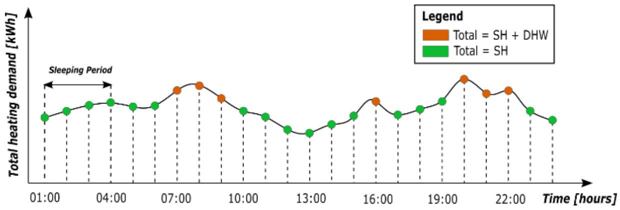

Figure 1: Separation method's representation.

All points identified with DHW production are removed from the dataset in order to have only SH measurements. The remaining SH data points will be used to estimate the SH from the removed recordings. The estimation algorithm is explained in the following subsection.

SH and DHW estimation

At this stage, the smart meters' dataset consists of measurements without DHW production ($E_{Total} = E_{SH}$). The next stage of the methodology is to estimate the SH usage ($E_{SH,estim}$) at the data gaps. After determining the $E_{SH,estim}$, the DHW usage ($E_{DHW,estim}$) is calculated with Equation 1.

From the same starting argument of the energy separation, the SH demand will vary smoothly due to small outdoor temperature oscillations. Therefore, the SH share in the removed data points is predicted from its known neighboring SH measurements that remained in the dataset. To estimate the SH, a smoothed Kalman filter algorithm is applied. This algorithm is based on a structural time series model from the function "StructTS" in the R-package *imputeTS* (Moritz and Bartz-Beielstein (2017)). The package's selected function consists of a linear Gaussian state-space model for univariate time series.

From the results in Leiria et al. (under review), the Kalman smoothing technique is a good method to predict the SH demand in the missing values. However, as mentioned, these values are calculated by their neighboring points. Basing this estimation on the adjacent points raises the risk of inaccuracy when several points are removed sequentially (large gap). To solve this problem, the algorithm is refined to use the smoothed Kalman filter only when the number of hours removed consecutively is equal to or below 2 hours (gap ≤ 2 hours). If the data gap is larger, a support vector regression (SVR) is applied instead. The SVR is a machine learning regressor that is trained with the known SH points that remained in the dataset and takes into account other inputs to calculate the SH usage instead of the neighboring points. The input data to estimate a given SH share is the outdoor temperature and global solar radiation measured two and one hours prior to the missing point and the smart meter measurements before and after the missing point. The SVR model uses a radial kernel function with the parameters C (cost) and γ (gamma) equal to 7 and 0.01, respectively. The SVR algorithm is retrieved from the R-package *e1071* (Meyer et al. (2020)). One can see in Table 1 the details regarding the estimation algorithms.

Table 1: Methods' description.

Method	Parameters	Input	Condition
Kalman filter	Model: StructTS Smoothed: True	E_{Total} *[i]*	Gap ≤ 2 hours
SVR	Kernel: Radial $C = 7$ $\gamma = 0.01$	T_{out} *[i-1, i-2]* *Rad [i-1]* E_{Total} *[i-1, i+1]*	Gap > 2 hours

The final part of the present methodology compares the estimated values from the methodology and the actual measurements. Also, it explicitly compares the DHW method's prediction on the rounded measurements (present case buildings), prediction on the decimal values (from the study in Leiria et al. (under review)), and the DHW estimation from the compliance calculations in the different countries (as described below).

Danish DHW compliance calculations

In Denmark, the DHW consumption in residential buildings is currently estimated using the compliance calculation of 250 liters/m^2 per year (Aggerholm and Skovgaard (2018)). Similarly, the supplied cold water and DHW temperatures are 10ºC and 55ºC, respectively (Dansk Standard (2000)). By using the floor area of the different apartments, the yearly DHW energy production is calculated through Equation 2:

$$E_{DHW}^{DK} = \frac{1}{3600} \cdot 0.25A \cdot \rho_w c_{p,w} \cdot (T_{DHW} - T_c) \quad (2)$$

Swiss DHW compliance calculations

In Switzerland, the DHW consumption in residential apartment buildings is currently predicted using the compliance calculation of 35 liters/day per person, and each person is considered to occupy 30 m^2 of floor area (Société suisse des ingénieurs et des architectes (2015)). Similarly, the supplied cold water and DHW water temperatures are 10ºC and 60ºC, respectively (Société suisse des ingénieurs et des architectes (2015)). Thus, by knowing the building's floor area, the yearly DHW energy production is calculated through Equation 3:

$$E_{DHW}^{CH} = \frac{365}{3600} \cdot \frac{0.035}{30} A \cdot n \cdot \rho_w c_{p,w} \cdot (T_{DHW} - T_c) \quad (3)$$

Italian DHW compliance calculations

In Italy, the DHW consumption in specific (commercial) buildings is estimated using particular compliance calculations and standards. The Italian dataset has the heating measurements of a rehab institution and a theatre,

 https://doi.org/10.1051/e3sconf/202236210001

therefore, the DHW consumption (volume) is calculated accordingly for each building case. The rehab institution accounts for a water volume (V_w) of 80 liters/day per existing bed in the building (Ente Nazionale Italiano di Unificazione (2014)), and the number of days regarding the calculation period (G) is equal to 365. The theatre's DHW consumption (Vw) is given at 3.8 liters/day per person (ISO (2016)). The theatre is divided into zones where the number of people will variate accordingly. The number of days (G) for this case is 251 (ISO (2016)). For both cases, the supplied cold water and DHW temperatures are 13ºC and 40ºC, respectively (Ente Nazionale Italiano di Unificazione (2014)). Thus, by knowing the buildings' bed number or occupants' number (n), the yearly DHW energy production is calculated through Equation 4:

$$E_{DHW}^{IT} = \frac{G}{3600} \cdot 10^{-3} V_w \cdot n \cdot \rho_w c_{p,w} \cdot (T_{DHW} - T_c) \quad (4)$$

Results and Discussion

The first set of results from this research is from applying the energy separation algorithm to identify the measurements with DHW usage. To assess the identification accuracy, the percentage of incorrectly identified measurements is shown per case building in Figure 2.

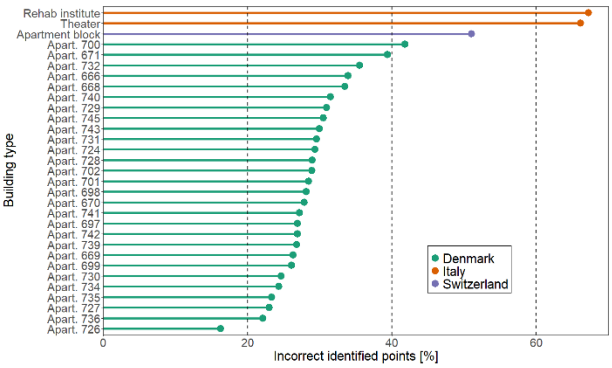

Figure 2: Incorrectly identified points percentage per building type.

The results show that the separation approach is quite inaccurate in identifying the DHW draw-off events. The lowest percentages belong to the Danish cases (single-family apartments), with the lowest value of 16%. The largest inaccurate identified points belong to the Italian cases with the extreme of 67%. The plot corroborates the hypothesis that this separation approach performs better for households than commercial buildings.

The following step in the methodology is the estimation of the SH usage in the detected DHW points. The estimation algorithm combines two methods, smoothed Kalman filter estimator and SVR, as described in the *Methodology*. In Figure 3, it is presented the overall error between the estimated values (SH – upper area; DHW – down area) and the real measurements.

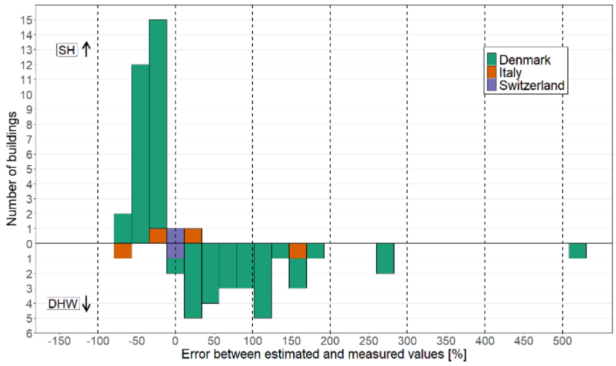

Figure 3: Overall error of the SH and DHW estimation for each building.

As one can see from Figure 3, the overall SH error (upper area) is mainly negative (underestimated) and has a lower error than the estimated DHW, with the buildings having a SH error between -65% and 17%.

Concerning the DHW prediction (down area), the error distribution is much wider than the SH predictions. In this case, 13 of the buildings on the dataset have an overestimation of the DHW demand above +100%. The extreme DHW prediction is one single-family dwelling with an overestimation of +510% and only two apartments being underestimated.

Several reasons can be outlined to explain these energy predictions and their overall error. Foremost, the separation method inaccurately identifies several measurements, decreasing the estimation's accuracy from the start. From Leiria et al. (under review), it is seen that in this research work, the separation approach underperforms more in single-family apartments. This is due to the coarse measurements (rounded values), which hinder the algorithm from finding the maximum values because most data points have the same value (e.g., 1, 2, 3 kWh). Also relevant is that this method has a significant inaccuracy for the commercial buildings. To overcome this challenge, a separation approach can be developed, taking into account the maximum values (as done in this manuscript) and the occupancy schedule. Because these buildings have such strict schedules (e.g., opening and closing hours), a more precise method can be developed to account for these characteristics.

Another factor is the occurrence of missing measurements in the initial dataset. As one can see in the *Study Case*, the different countries' datasets are comprised of large missing measurement gaps (Denmark and Switzerland) or small timespan measurements (Italy). Because the prediction relies on determining the SH usage based on its neighboring points, several missing points negatively impact the overall method's accuracy.

Furthermore, the different heating systems and people's social cultures significantly impact the methodology. As described, some of the DHW systems are of instantaneous heat production (Denmark). However, others have a storage tank (Switzerland), which in itself affects the DHW usage detection. Besides the production system, the unique dwellers' consumption habits or the DHW usage being equal to zero (no occupancy) may influence the method's performance, which might explain the extreme error estimated cases.

The present work also assesses the estimated DHW values from the method with the different countries' compliance calculations used to predict the DHW demand in the buildings. In this comparison, the values presented in the manuscript Leiria et al. (under review) are also displayed ("decimal" column). The results of this comparison are in Table 2.

Table 2: Comparison between the countries' compliance predictions and the method's estimation results. The green-colored cells indicate the best (orange color – the worst) performing method between this research's method (rounded and decimal measurements) and the compliance calculations when comparing with the actual DHW measurements.

Data	Case-building	Error		
		Compliance	Round	Decimal
DK	Apart 666	-47%	97%	0%
DK	Apart 668	-42%	103%	21%
DK	Apart 669	-11%	102%	22%
DK	Apart 670	-72%	21%	-6%
DK	Apart 671	-34%	108%	20%
DK	Apart 697	-76%	12%	-12%
DK	Apart 698	-75%	21%	-7%
DK	Apart 699	-76%	10%	-13%
DK	Apart 700	123%	510%	85%
DK	Apart 701	-1%	93%	18%
DK	Apart 702	87%	182%	32%
DK	Apart 724	-28%	89%	11%
DK	Apart 726	43%	70%	14%
DK	Apart 727	61%	149%	18%
DK	Apart 728	11%	152%	37%
DK	Apart 729	14%	119%	12%
DK	Apart 730	-57%	43%	5%
DK	Apart 731	90%	273%	63%
DK	Apart 732	-60%	24%	-15%
DK	Apart 734	59%	144%	17%
DK	Apart 735	-50%	44%	6%
DK	Apart 736	-51%	40%	1%
DK	Apart 739	-68%	34%	7%
DK	Apart 740	1%	75%	-3%
DK	Apart 741	-30%	59%	7%
DK	Apart 742	0%	121%	15%
DK	Apart 743	-64%	29%	-13%
DK	Apart 745	78%	265%	69%
CH	Apart. block	4%	-9%	-
IT	Rehab inst.	-59%	-79%	-
IT	Theater	-35%	154%	-

As shown in Table 2, there are three calculated errors per DHW usage. The error between the measured DHW usage and the DHW compliance calculations ("Compliance"), the error between the actual measurements and the results from the methodology applied on this manuscript dataset ("Round"), and the error between the DHW measurements and the results from the methodology applied on the Leiria et al. (under review) dataset ("Decimal"). The error calculation is performed using the aggregated DHW usage divided by the number of data points (hours). For the case of the DHW measurements and the disaggregation method, the number of data points is the number of measurement hours in each building. For the compliance case, the number of data points is the number of hours in a year. In most cases, the compliance calculations outperform (green color) the disaggregation methodology when applied to rounded values. However, the total heating values are recorded with decimal number, its accuracy increases and outperforms the DHW compliance prediction. The main reasons behind this performance are stated above. However, it is relevant to highlight that even though the compliance calculations are in some cases more accurate, the methodology should be reviewed or changed because the results are still too high and might be one of the main reasons for the observed building energy performance gap.

Conclusion

This article presents a validation study on a new data-driven methodology to estimate the SH and DHW from hourly resolution heat meters data. The validation novelty is the application of different building cases with different characteristics, e.g., different measurements resolution, building types, heating systems, and countries (consumption habits).

The validation process shows that the method is quite inadequate to detect DHW usage in rounded measurements or commercial buildings. To solve these challenges, it is argued that the measurements cannot be rounded and should be recorded with decimals, and that the separation algorithm must be refined by taking into account the occupancy schedules in large buildings. The overall methodology predicts better the SH demand with an error between -65% and 17%. Concerning DHW prediction, the error is much wider, with most building cases falling between 0% and 200%. Additionally, this study compared the estimated DHW demand predicted by the method with the actual measurements and the DHW compliance calculations used in Denmark, Switzerland, and Italy. This comparison concludes that the compliance estimations outperform this method for most building cases, when the used rounded values. However, it is argued that the compliance calculations must be updated or replaced to estimate more precisely the buildings' DHW demand, hence decreasing the energy performance gap and improving the EPCs' accuracy.

Suggestions for Further Work

A suggestion for further work is the application of this methodology with other datasets for further validation and robustness analysis. Improving the separation methodology for rounded measurements and commercial cases is highly needed.

It is also advised to benchmark this methodology with other existing disaggregation techniques on a common dataset. Additionally, a more extensive endeavor must be made to collect good quality datasets and share them with our research peers.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 893945 (E-DYCE). This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 958345 (PRELUDE). This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement number 754354. This project has received funding from the Interdepartmental Centre "G. Levi Cases" of the University of Padova within the Regional Project GHOTEM (POR-FESR 2014-2020, ID 10064601).

The authors would like to acknowledge the Himmerland housing association and Varmekontrol A/S for their support in data collection in Denmark and to all partners involved in the E-DYCE and PRELUDE projects. Also, the first author would like to express his gratitude to his colleagues in the Ph.D. "Corner Office" for the valuable discussions and support.

The authors declare no conflict of interest

Nomenclature

Acronyms	
CH	Switzerland (country code)
DHW	Domestic hot water
DK	Denmark (country code)
EPC	Energy performance certificate
EU	European Union
IT	Italy (country code)
SH	Space heating
SVR	Support vector regression
Symbols and variables	
A	Floor area [m^2]
C	Cost (SVR parameter) [-]
$C_{p,w}$	Water specific heat capacity – Constant: 4.18 [kJ/kgºC]
E_{DHW}	Measured domestic hot water energy usage [kWh]
$E_{DHW,compl}$	Estimated annual DHW energy usage from any compliance [kWh/year]
$E_{DHW,estim}$	Estimated domestic hot water energy usage [kWh]
E_{DHW}^{DK}	Estimated annual DHW energy usage from Danish compliances [kWh/year]
E_{DHW}^{CH}	Estimated annual DHW energy usage from Swiss compliances [kWh/year]
E_{DHW}^{IT}	Estimated annual DHW energy usage from Italian compliances [kWh/year]
E_{SH}	Measured space heating energy usage [kWh]
$E_{SH,estim}$	Estimated space heating energy usage [kWh]
E_{Total}	Measured total heat usage (smart meter measurements) [kWh]
G	Number of days in calculation period [days]
n	Number of people or beds [-]
Rad	Global solar radiation [W/m^2]
T_c	Temperature of inlet cold water [ºC]
T_{DHW}	Temperature of outlet DHW water [ºC]
T_{out}	Outdoor temperature [ºC]
γ	Gamma (SVR parameter) [-]
ρ_w	Water density – Constant: 1000 [kg/m^3]

References

Aggerholm, S. and M. Skovgaard (2018) *SBi-ANVISNING 213: Bygningers energibehov*. Sbi Forlag. Copenhagen (Denmark).

Alzaatreh, A. and L. Mahdjoubi, B. Gething, and F. Sierra (2018). Disaggregating High-Resolution Gas Metering Data Using Pattern Recognition. *Energy and Buildings 176*, 17–32.

Bacher, P., P. Anton de Saint-Aubain, L. E. Christiansen, and H. Madsen (2016). Non-Parametric Method for Separating Domestic Hot Water Heating Spikes and Space Heating. *Energy and Buildings 130*, 107–112.

BPIE (2017). Factsheet: 97% of Buildings in the EU Need to Be Upgraded. *https://www.bpie.eu/publication/97-of-buildings-in-the-eu-need-to-be-upgraded/*.

Cozza, S., J. Chambers, A. Brambilla, and M. K. Patel (2021). In Search of Optimal Consumption: A Review of Causes and Solutions to the Energy Performance Gap in Residential Buildings. *Energy and Buildings*.

Danish Climate Policies | Energistyrelsen (2022). Accessed March 25. *https://ens.dk/en/our-responsibilities/energy-climate-politics/danish-climate-policies*.

Dansk Meterologisk Institut (2022). DMI Frie. Accessed April 5. *https://www.dmi.dk/frie-data/*.

Dansk Standard (2000). *Norm for Vandinstallationer (DS 439:2009)*.

European Commission (2022a). Energy Performance of Buildings Directive. Accessed January 27. *https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en*.

European Commission (2022b). Heating and Cooling. Accessed January 27. *https://energy.ec.europa.eu/topics/energy-efficiency/heating-and-cooling_en*.

Flourentzou, F. and J. Pereira (2021). Domestic Hot Water Optimizing Potential in Existing or Renovated Multifamily Residential Buildings. *In Journal of Physics: Conference Series, 2042:012144*.

Gonzalez-Caceres, A., J. Karlshøj, T. A. Vik, E. Hempel, and T. R. Nielsen (2022). Evaluation of Cost-Effective Measures for the Renovation of Existing Dwellings in the Framework of the Energy Certification System: A Case Study in Norway. *Energy and Buildings*.

Gram-Hanssen, K. (2014). New Needs for Better Understanding of Household's Energy Consumption - Behaviour, Lifestyle or Practices?. *Architectural Engineering and Design Management 10* (1–2).

Gram-Hanssen, K. and A. R. Hansen (2016). Forskellen Mellem Målt Og Beregnet Energiforbrug Til Opvarmning Af Parcelhuse. Copenhagen.

Hedegaard, R. E., M. H. Kristensen, and S. Petersen (2020). Experimental Validation of a Model-Based Method for Separating the Space Heating and Domestic Hot Water Components from Smart-Meter Consumption Data. In *E3S Web of Conferences, 172:12001*.

Iribar, E., I. Sellens, L. Angulo, J. M. Hidalgo, and J. M. Sala (2021). Nonconformities, Deviation and Improvements in the Quality Control of Energy Performance Certificates in the Basque Country. *Sustainable Cities and Society 75*.

International Organisation for Standardisation (2016). *Energy Performance of Buildings-Schedule and Condition of Building, Zone and Space Usage for Energy Calculation-Part 1: Non-Residential Buildings (ISO 18523-1).*

Ivanko, D., Å. L. Sørensen, and N. Nord (2021). Splitting Measurements of the Total Heat Demand in a Hotel into Domestic Hot Water and Space Heating Heat Use. *Energy 219.*

Johra, H., D. Leiria, P. Heiselberg, A. Marszal-Pomianowska, and T. Tvedebrink (2020). Treatment and Analysis of Smart Energy Meter Data from a Cluster of Buildings Connected to District Heating: A Danish Case. *In E3S Web of Conferences, 172:12004.*

Kristensen, M. H., and S. Petersen (2021). District Heating Energy Efficiency of Danish Building Typologies. *Energy and Buildings 231*.

Leiria, D., H. Johra, A. Marszal-Pomianowska, and M. Z. Pomianowski (under review). A Methodology to Estimate Space Heating and Domestic Hot Water Energy Demand Profile in Residential Buildings from Low-Resolution Heat Meter Data. *Energy*.

Lien, S., D. Ivanko, and I. Sartori (2020). Domestic Hot Water Decomposition from Measured Total Heat Load in Norwegian Buildings. *In BuildSim-Nordic 2020 Selected Papers*.

Marszal-Pomianowska, A., C. Zhang, M. Pomianowski, P. Heiselberg, K. Gram-Hanssen, and A. R. Hansen (2019). Simple Methodology to Estimate the Mean Hourly and the Daily Profiles of Domestic Hot Water Demand from Hourly Total Heating Readings. *Energy and Buildings 184*.

Meyer, D., E. Dimitriadou, K. Hornik, A. Weingessel, F. Leisch, C. Chang, and C. Lin (2020). CRAN - Package E1071. *CRAN - Package. https://cran.r-project.org/web/packages/e1071/index.html.*

Moritz, S. and T. Bartz-Beielstein (2017). ImputeTS: Time Series Missing Value Imputation in R. *R Journal 9 (1),* 207–218.

Bordignon, S., D. Quaggiotto, J. Vivian, G. Emmi, M. De Carli, and A. Zarrella (2022). A Solar-Assisted Low-Temperature District Heating Network Coupled with a Ground-Source Heat Pump. *Energy Conversion and Management 267*.

RStudio (2022). Accessed March 25. *https://www.rstudio.com/products/rstudio/*.

Société suisse des ingénieurs et des architectes (2015). *Données d'utilisation Des Locaux Pour l'énergie et Les Installations Du Bâtiment (SIA 2024).*

The Danish Government (2020). A Green and Sustainable World - Denmark's Global Climate Action Strategy. *https://um.dk/~/media/um/klimastrategi/onepager eng dks globale klimastrategifinal.pdf.*

Ente Nazionale Italiano di Unificazione (2014). *Part 2: Evaluation of Primary Energy Need and of System Efficiencies for Space Heating and Domestic Hot Water Production (UNI/TS 11300).*

Vivian, J., D. Quaggiotto, and A. Zarrella (2020). Increasing the Energy Flexibility of Existing District Heating Networks through Flow Rate Variations. *Applied Energy 275*.

Zeifman, M., and K. Roth (2011). Nonintrusive Appliance Load Monitoring: Review and Outlook. *IEEE Transactions on Consumer Electronics 57 (1)*, 76–84.