Estimating residential space heating and domestic hot water from truncated smart heat data

D Leiria[1*], M Schaffer[1], H Johra[1], A Marzsal-Pomianowski[1], and M Z Pomianowski[1]

[1] Aalborg University, Department of the Built Environment, Aalborg, Denmark

* Corresponding author: dle@build.aau.dk

Abstract. The EU aims to digitize the building stock across all member states to better understand energy use and achieve energy efficiency goals to address climate change. Smart heat meters are currently used for billing purposes in district heating (DH) grids. Their data is recorded as integer kWh values, which restricts usability for the modeling and analysis of DH networks. Previous research devised a methodology to estimate space heating (SH) and domestic hot water (DHW) energy from total heating data, but the data truncation process reduced accuracy. This study integrates the SPMS (Smooth–Pointwise Move–Scale) algorithm, which estimates decimal values from DH truncated measurements, to improve the accuracy of the DHW and SH disaggregation methods. The study applies these two methodologies to a dataset of 28 Danish apartments and compares the results against full-resolution and truncated data to evaluate performance. Another dataset, named "optimal dataset" is also assessed to determine overall estimation accuracy. Results show that SPMS reduces the disaggregation methodology error of SH and DHW compared to the truncated data. The optimal dataset outperforms the current methodology, indicating a potential for improving and scaling the methodology for larger datasets.

1. Introduction

The building sector plays a major role in our society as it contributes to 40% of the European Union's (EU) total energy end-usage [1]. District heating (DH) systems are cost-effective, flexible, and sustainable energy solutions to fulfill the buildings' heating demand, as already seen, especially in Europe, the USA, Canada, and Asia [2]. This type of system has been through several innovative technological stages throughout the years (named generations). The most current transition is from the 3rd to the 4th generation of district heating (4GDH) systems. The 4GDH is characterized mainly by a lower-temperature heat-carrier fluid supply, bringing the advantage of reducing heat losses in the DH network, increasing the output capacity when integrating with renewable energy sources, minimizing the risk of pipe leakages due to thermal stress, and meeting new building requirements more effectively.

In the 4GDH transition, smart heat meters (SHM) play a vital role in making it possible to manage the energy production and distribution grid, detect faults, and provide information to DH operators and customers (end-users) [3]. Nevertheless, these meters only measure the total heat demand, and to have a better understanding of the DH end-users' heat usage, it is necessary to estimate the space heating (SH) and domestic hot water (DHW) demand separately. This data disaggregation of the energy use in buildings contributes to better decision-making strategies regarding heat production and distribution due to these two heating demands are dependent on different variables.

1.1. State-of-the-art

Several methodologies have been proposed in the last few years to address the problem of disaggregating the energy for SH and DHW from total heating measurements registered by SHM. In Table 1 is listed and summarized these different methodologies based on the state-of-the-art of Leiria et al., 2023 [4].

Table 1. Summary of the existing methods to disaggregate SH and DHW from total heat measurements reproduced from [4].

Ref.	Country	System	Period	Resolution	Dataset	Methodology
[5]	SE	District heating	8 hours for "winter" and "summer" times	5-sec	An apartment building	DH substation model created using Matlab Simulink.
[6]	DK	District heating	1-month	10-min	One single family house (SFH)	SH estimated with a kernel smoother from the total heat measurements.
[7]	UK	Natural gas boilers	2-months	1-min	Two SFHs	Detects the various heat signatures of the appliances.
[8]	DK UK	District heating	Few months to a full year (depending on the building)	1-hour (DK) 10-min (UK)	Apartments and SFHs	Two methods extract the DHW profiles during the no-heating season.
[9]	NO	District heating	1-3 years (depending on the building)	1-hour	58 apartments blocks and 20 hotels	Identifies DHW profiles during the summer period and calculates the SH by subtracting the extrapolated DHW daily profiles from the total values.
[10]	DK	District heating	1-year	1-hour	44 terraced SFHs	Grey-box model-based method to estimate the SH and DHW usage.
[11]	NO	District heating	1-year	1-hour	A hotel with 260 rooms	Two methods use the energy signature (ES) to estimate the SH and DHW. The second employs the singular spectrum analysis to decompose the SH and DHW derived from the ES.
[4]	DK	District heating	3-9 months (depending on the building)	1-hour	28 SFHs	Disaggregation is based on detecting the daily sporadic DHW production and estimating the SH from the neighboring points in the time series.
[12]	DK CH IT	District heating	3-9 months per building (DK) 5-months (CH) 1-2 months (IT)	1-hour	28 SFHs (DK) An apartment block (CH) Other buildings (IT)	Validation of the method [4] with truncated heat measurements and other types of buildings.

Leiria et al. [4] presented a two-step disaggregation algorithm. The first step (separation stage) refers to distinguishing throughout the day, the measurements with simultaneous DHW and SH (SH+DHW data points) usage from those with only SH demand (only SH data points). The second step (estimation stage) is the estimation of the SH in the SH+DHW points based on the neighboring SH values of the data points with only SH usage. Nevertheless, the work was developed based on high-resolution measurements, which is not the reality in most Danish DH cases. The measurements gathered by the DH companies usually have their values rounded down to a 1 kWh resolution (known as truncation) [13]. This means that if a SHM records a specific heating measurement between 1.1 to 1.9 kWh, it is truncated (decimals removed) and recorded as 1.0 kWh. Due to this context, the authors tested their methodology for truncated values in [12]. The conclusions from this second study are that the methodology underperforms significantly and cannot be used for this type of data. Another question raised throughout these two studies [4,12] is the dependency of the SH estimation accuracy (step 2) from the separation stage (step 1). This means that it is envisioned that the SH is better estimated if the separation stage is more precise in segregating the "SH+DHW" measurements from the "only SH" measurements.

1.2. Contributions and novelty of the current study

To tackle the general issue of truncated data in the DH sector, a novel algorithm (named SPMS) is proposed by [14]. It enhances the usability of truncated data by partially restoring the actual underlying trend of the SHM measurements. The work shows that the proposed method can decrease the error from the truncation process, consequently increasing the usefulness of the data. Based on these results, the current study integrates the SPMS method into the disaggregation method from [4] to increase its estimation accuracy when using hourly truncated SHM data.

Regarding the accuracy of the SH estimation (step 2) based on the precision of the separation stage (step 1) question, this manuscript addresses it by assessing the accuracy of the second step of the method [4] (estimation), if the DHW production data points are perfectly detected in the first step (this was done using the validation dataset by categorizing the measurements as "SH+DHW" when the DHW energy demand was larger than 0 kWh). This dataset with perfect detected "SH+DHW" measurements is named "optimal dataset" and is compared with the accuracy assessed when using the method in decimal, truncated, and SPMS retrieved data measurements.

2. Methods

2.1. Dataset description

The study uses a dataset of 28 apartments located in a social housing complex in Aalborg, Denmark. This dataset is the same one used in [4,8,12,14]. The apartments are equipped with individual SH and total heat demand meters. The DHW is thus calculated as the difference between measurements from those two meters. The weather data, the outdoor temperature, and global radiation were taken from the Danish Meteorologic Institute (DMI). The data pre-processing removes missing and negative energy usage measurements.

2.2. Integration of SPMS with the disaggregation method

This section presents the methodology used in this study to estimate the SH and DHW energy use in Danish dwellings when the data is truncated. The method is based on the previous methodology that was developed to disaggregate and estimate the SH and DHW energy use from the total heating measurements recorded by the SHMs [4,12]. However, the current work combines the disaggregation methodology with a data-recovering algorithm named: Smooth – Pointwise Move – Scale (SPMS) method. It retrieves the decimal values from the truncated measurements.

The first stage in the methodology involves the application of the SPMS algorithm. SPMS is a three-step algorithm with the following steps:

1. The data are smoothed using a linear weighted moving average with a centred window of length 5.
2. For each data point, it is ensured that the new value is within ± 0.4 kWh of the original data and that no value is negative.
3. The values obtained are scaled so that they accumulate over one day to the same amount as the original data.

Steps 2 and 3 are repeated in a loop until all conditions are met [14].

After the processed values are retrieved by the SPMS, the newly generated dataset is used in the disaggregation methodology. The method is based on the following assumptions: (1) the SH is in continuous operation, and (2) the DHW is sporadically produced and responsible for the peaks in the heat demand measurements. These assumptions are based on previous studies on the energy use patterns in Danish dwellings and are used to calculate the SH energy use from the disaggregated total heating measurements [6]. Following the initial assumptions, the method starts by categorizing the highest daily heating demand peaks as "SH and DHW combined usage" and the lower heating measurements as "SH usage". Subsequent to the categorization process, the method estimates the SH energy use by applying a combination of a Kalman estimator and support vector regression (SVR) while taking the outdoor conditions (temperature and global solar radiation) as inputs.

The results are then compared with the estimation when the DHW production measurements are perfectly detected (optimal data points detection).

3. Results and discussion

The proposed disaggregation algorithm was tested for four different cases: the original heating measurements with decimal values ("Decimal"), the truncated heating dataset ("Truncated"), the recovered dataset from the truncated data by the SPMS method ("SPMS"), and heating data where the segregation between data points with and without DHW production is performed perfectly ("Optimal"). The results of the application of the estimation method in the four cases are presented and discussed below.

Overall, the results of the experimental evaluation showed that the proposed energy estimation algorithm with the integration of the SPMS method outperforms the performance of the same method when using truncated data. As one can see in Figure 1, the NMBE and CVRMSE indicators display a significant performance improvement in the SH estimation from the truncated measurements to the SPMS recovered dataset. Nevertheless, this improvement is still lower than when using the original decimal dataset, corroborating the argument that, better than applying the SPMS method, the DH companies should address this problem by recording their heating measurements with a higher resolution.

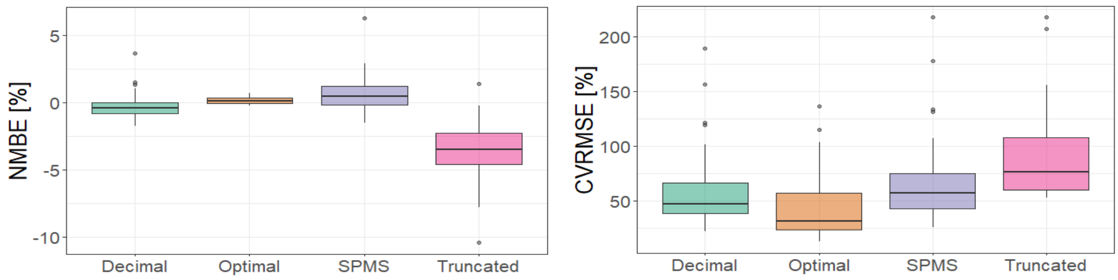

Figure 1. The NMBE and CVRMSE of the SH estimation of the different tested cases.

Regarding the estimation error of overall SH and DHW usage per building, it can be seen in Figure 2 that the best performance is given by the "Optimal" case, while the worst prediction scenario is for the truncated data.

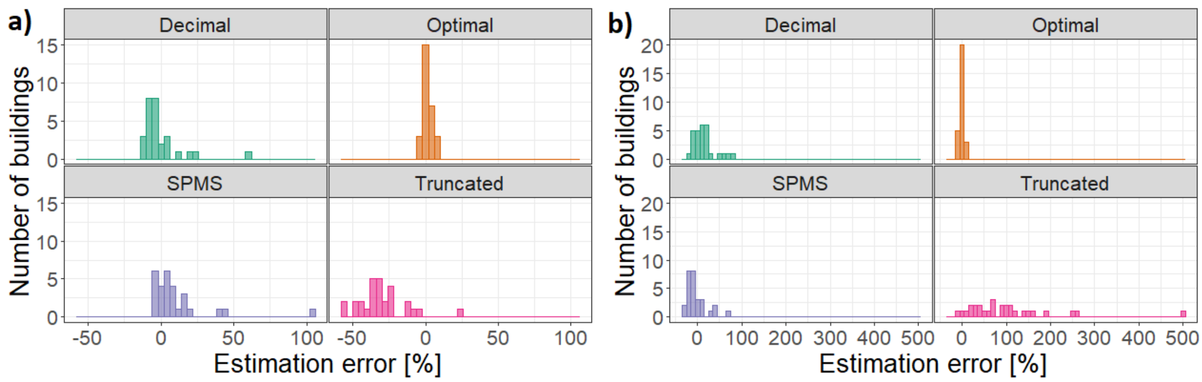

Figure 2. Estimation error for a) SH usage and b) DHW production.

Derived from the results in Figure 2, the Table 2 shows the minimum, median, and maximum error values for the prediction of SH and DHW usage.

Table 2. Minimum, median, and maximum error values per dataset case.

Case	SH			DHW		
	Minimum	Median	Maximum	Minimum	Median	Maximum
Decimal	-13%	-4%	60%	-17%	11%	83%

Truncated	-58%	-32%	23%	-6%	81%	497%
SPMS	-5%	4%	103%	-29%	-10%	73%
Optimal	-6%	1%	9%	-14%	-2%	12%

For SH, the decimal values case has a minimum error of -13% (underprediction) and a maximum error of 60% (overprediction). The truncated scenario has the highest negative error (-58% minimum), while the SPMS scenario has the highest maximum error (103%). The optimal scenario has the smallest range of error (-6% minimum to 9% maximum).

For DHW, the optimal scenario has the smallest absolute median error (-2%), while the truncated data case has the largest median error (81%). The truncated scenario also has the highest maximum error (497%), indicating a large overestimation of DHW energy use. The SPMS scenario has a negative median error (-10%), which means that it more frequently underestimates the DHW production. The optimal scenario has, again, the smallest range of error (-14% minimum to -2% maximum).

Overall, these results suggest that the choice of scenario can significantly impact the accuracy of energy use estimations for both SH and DHW. The results also prove disaggregation methodology can be improved further by developing new DHW usage timing identification algorithms.

4. Conclusion

This article presents the performance evaluation of a heating disaggregation algorithm coupled with a new method to recover (estimate) higher DH data resolution when truncated initially. Four different cases were tested, namely, the original decimal dataset, truncated dataset, SPMS recovered dataset, and the ideal scenario where the data points with and without DHW production are perfectly identified. The results show that the coupled methods outperform the truncated dataset in the estimation of SH and DHW energy use. However, the improvement is still lower than using the original decimal dataset. The optimal scenario performs the best for both SH and DHW estimation, while the truncated dataset case shows the highest median error and maximum error for DHW estimation. These findings highlight the significance of applying the SPMS when having truncated data or using these results to justify the recording of higher-resolution measurements by DH companies for more accurate building energy estimation. This work also emphasizes the potential for further development of the disaggregation algorithm by improving the identification methodology for hourly measurements with DHW usage.

There are several opportunities for further work on the proposed energy disaggregation algorithm. One suggestion is to identify DHW measurement points more accurately, possibly by using additional sensors such as cold-water meters. Another suggestion is to optimize the algorithm for larger datasets by exploring more computationally efficient estimation methods. Another opportunity is the development of a comprehensive evaluation of the present algorithm's performance by comparing it against other existing disaggregation methods (see Table 1) as it could provide insights into its suitability for different scenarios and datasets.

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). The authors would like to acknowledge the Himmerland housing association and Varmekontrol A/S for their support in data collection 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.

Credit-author statement

Daniel Leiria: Conceptualization; Methodology; Software; Validation; Formal analysis; Investigation; Writing - original draft; Visualization. **Hicham Johra:** Conceptualization; Methodology; Resources; Writing - review & editing; Supervision. **Markus Schaffer:** Conceptualization; Methodology; Software; Validation; Formal analysis; Investigation; Data curation; Writing - review & editing. **Anna Marszal-Pomianowska:** Conceptualization; Resources; Writing - review & editing; Supervision.

Michal Zbigniew Pomianowski: Conceptualization; Resources; Writing - review & editing; Supervision; Project administration; Funding acquisition.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] European Commission *Energy Performance of Buildings Directive* [Internet] 2022 [Accessed Jan 27] Available from: https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en

[2] IRENA, IEA and REN21 2018 *Renewable Energy Policies in a Time of Transition* IRENA, OECD/ IEA and REN21

[3] Volkova A, Mašatin V and Siirde A 2018 Methodology for evaluating the transition process dynamics towards 4th generation district heating networks *Energy* **150** 253–261

[4] Leiria D, Johra H, Marszal-Pomianowska A and Pomianowski M Zbigniew 2023 A methodology to estimate space heating and domestic hot water energy demand profile in residential buildings from low-resolution heat meter data *Energy* **263 B**

[5] Yliniemi K, Delsing J and Deventer J 2009 Experimental verification of a method for estimating energy for domestic hot water production in a 2-stage district heating substation *Energy and Buildings* **41** 169–174

[6] Bacher P, Saint-Aubain P Anton de, Christiansen L Engbo and Madsen H 2016 Non-parametric method for separating domestic hot water heating spikes and space heating *Energy and Buildings* **130** 107–112

[7] Alzaatreh A, Mahdjoubi L, Gething B and Sierra F 2018 Disaggregating high-resolution gas metering data using pattern recognition *Energy and Buildings* **176** 17–32

[8] Marszal-Pomianowska A, Zhang C, Pomianowski M, Heiselberg P, Gram-Hanssen K and Hansen A Rhiger 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** 53–64

[9] Lien S, Ivanko D and Sartori I 2020 *Domestic Hot Water Decomposition from Measured Total Heat Load in Norwegian Buildings* BuildSIM-Nordic 2020 Oslo Norway OsloMet

[10] Hedegaard R Elbæk, Kristensen M Heine and Petersen S 2020 Experimental validation of a model-based method for separating the space heating and domestic hot water components from smart-meter consumption data *E3S Web of Conferences* **172**

[11] Ivanko D, Sørensen Å Lekang and Nord N 2021 Splitting measurements of the total heat demand in a hotel into domestic hot water and space heating heat use *Energy* **219**

[12] Leiria D, Johra H, Belias E, Quaggiotto D, Zarrella A, Marszal-Pomianowska A and Pomianowski M Zbigniew 2022 Validation of a new method to estimate energy use for space heating and hot water production from low-resolution heat meter data *E3S Web of Conferences* **362**

[13] Broholt T Hauge, Christensen L Rævdal Lund and Petersen S 2022 *Effect of Measurement Resolution on Data-Based Models of Thermodynamic Behaviour of Buildings* REHVA 14th World Congress CLIMA 2022 Rotherdam The Netherlands TU Delft OPEN Publishing

[14] Schaffer M, Leiria D, Vera-Valdés J Eduardo and Marszal-Pomianowska A 2023 *Increasing the Accuracy of Low-Resolution Commercial Smart Heat Meter Data and Analysing Its Error* 2023 European Conference on Computing in Construction Crete Greece (*Accepted*)