Data Pre-processing Methods Enhancing Heat Cost Allocator Measurement Usability

Original
Data Pre-processing Methods Enhancing Heat Cost Allocator Measurement Usability / Yang, Qinjiang; Saba, Fabio; Orio, Marina; Santiano, Marco; Audrito, Emanuele; Salenbien, Robbe; Tunzi, Michele. - 1700:(2026), pp. 153-162. (19th International Symposium of District Heating and Cooling, IEA DHC 2025 Genk, Belgium 7-10 September 2025) [10.1007/978-3-032-09844-3_15].

Availability:
This version is available at: 11696/88883 since: 2026-03-02T15:30:25Z

Publisher:
Springer Science and Business Media Deutschland GmbH

Published
DOI:10.1007/978-3-032-09844-3_15

Data Pre-processing Methods Enhancing Heat Cost Allocator Measurement Usability

Qinjiang Yang, Fabio Saba, Marina Orio, Marco Santiano, Emanuele Audrito, Robbe Salenbien, and Michele Tunzi

Abstract Heat cost allocators (HCAs) are devices mounted on radiators to fairly allocate heating consumption among flats within buildings connected to district heating networks or central heating systems. In recent years, HCA data has also been utilized for building heating system analysis, fault detection, diagnosis, and optimization. However, certain inherent limitations of HCAs, such as data truncation and the sparse recording of data points over time, can hinder their direct application in analysis. This underscores the necessity of pre-processing HCA data prior to conducting meaningful analyses. This study aims to develop a methodology for recovering the decimal values of HCA data. By leveraging the continuity of HCA increments and the inverse relationship between external temperature changes and HCA increments, the problem is formulated as an optimization problem. Two case studies were conducted to validate this method. The first case study involved a radiator heating laboratory at the National Metrology Institute of Italy (INRIM), where 40 radiators of various types, geometries, and materials were tested. The lab replicates heating operations typical of real apartment buildings, utilizing specific control strategies and flexible hydraulic connections. The second case study focused on a residential building in Denmark, analyzing HCA data collected from 15 apartments over one month. In both case studies, we used different measures to collect HCA data with decimals as the reference. Results indicate that the proposed method significantly reduces errors and uncertainties associated with data truncation in both laboratory and real-world settings. On average, the root mean square error (RMSE) of the recovered HCA data compared to the reference value decreased by 76.9% and 60.4% when compared to the truncated data in the lab and real buildings, respectively. This demonstrates the

Q. Yang (✉) · M. Tunzi
DTU Construct, Technical University of Denmark, Lyngby, Denmark
e-mail: qyan@dtu.dk

Q. Yang · R. Salenbien
Flemish Institute for Technological Research (VITO NV), Mol, Belgium

EnergyVille, Genk, Belgium

F. Saba · M. Orio · M. Santiano · E. Audrito
National Metrology Institute of Italy (INRIM), Turin, Italy

D. Vanhoudt (ed.), *Proceedings of the 19th International Symposium on District Heating and Cooling*, Lecture Notes in Networks and Systems 1700,
https://doi.org/10.1007/978-3-032-09844-3_15

method's effectiveness in enhancing the usability and reliability of HCA data over short time intervals.

Keywords Heat cost allocators · Data preprocessing · Radiators · Energy meter

1 Introduction

Heat cost allocators (HCAs) are indirect meters installed on radiators and are widely used across European countries [1]. These devices measure the relative heat output of each radiator within space heating (SH) systems, enabling fair distribution of heating costs. According to the Energy Efficiency Directive (EED) [2], all energy meters and submeters in buildings should support remote data access. Given that HCAs share similar properties with energy meters, they hold great potential for heating system analysis and optimization. Previous studies have demonstrated that HCA data can be utilized to identify apartments with high heat demand, determine system supply temperatures [3], and detect non-uniform heat distribution [4], as well as diagnose possible misuse or suboptimal operation within SH systems.

Despite their potential, HCAs were originally designed for billing over entire winter seasons rather than capturing dynamic heating behavior. HCAs accumulate the estimated heat output over time, and the operators collect the information based on a time length equal to the billing period, which is typically one month or even longer. According to the technical specifications, the device displays only integer values on its screen, and the company prefers to align the billing data with what users directly observe [5]. This practice is common among meter manufacturers. This introduces a problem: data is logged only with the integer part, which introduces significant uncertainties when looking at a shorter time interval compared to the billing time range, such as one or a few hours, and in particular for applications of the system optimization. To be more specific, the maximum uncertainty caused by truncation is $+1$ unit. In this study, the consumption of active radiators per day ranged from 4 to 20 units during the sampling period. Therefore, truncation has a substantial impact on the daily analysis, and this effect becomes even more significant for those low-consumption radiators and when shorter time intervals are considered.

To address this problem, this study aims to develop a methodology to reconstruct the missing decimal values in HCA data. By filling in these gaps, the methodology will enhance the precision of dynamic SH system analyses, ultimately improving the reliability of using HCA data for heating system optimization and supporting its future use. Some studies have attempted to infer and reconstruct the decimal parts of smart meter readings [6, 7]; however, to the best of the author's knowledge, there is currently no research focusing on the reconstruction of the decimal parts of HCA data.

2 Methodology

The methodology section will describe the counting basics of the HCA, the assumptions, and the methodology.

HCAs are counting the heat output from the radiators based on the integral of the temperature difference between the radiator surface and the indoor air. According to the European standard of HCAs EN834 [8], the HCA progression can be expressed by Eq. 1:

$$R = K_Q K_C K_T \int \left(\frac{T_{rad} - T_{in}}{60} \right)^n dt = K \int \left(\frac{T_{rad} - T_{in}}{60} \right)^n dt \tag{1}$$

where:

R is the HCA progression, representing the accumulated value that reflects the energy output of the radiator.

K_Q is the rating factor for the thermal output of the radiator, representing the non-dimensional numerical value of the reference output of the radiator.

K_C: The rating factor for the thermal coupling of the sensors.

K_T: The rating factor for rooms with low design indoor temperatures.

K: The total rating factor, the product of the individual rating factors.

T_{rad}: The radiator surface temperature measurement.

T_{in}: The indoor air temperature.

n: The radiator component.

60: The average temperature at the design condition of 90/70/20 °C.

In this study, there is no different thermal coupling of the sensors, nor low temperature setpoints, hence K_C and K_T are treated with default values. As a result, the total rating factor is only determined by the characteristics of the radiators.

The working principle of radiators and the nature of the SH system derive two assumptions that support the development of the methodology.

1. The radiator heat output should be continuously changing from one time interval to another. This is based on the heavy thermal inertia of the hydronic system.
2. The radiator heat output should react to the outdoor temperature in a negative relationship. This is defined based on the fact that the SH system in apartment buildings normally follows a weather compensation control, in which the supply temperature is adjusted against the change of the outdoor temperature.

Hence, the optimization problem is defined as below.

$$\min_{\hat{y}} J(\hat{y}) = \sum_{i=2}^{n-1} (\Delta \hat{y}_{i+1} - \Delta \hat{y}_i)^2 + \lambda \sum_{i=2}^{n} \Delta \hat{y}_i \Delta T_i$$

$$s.t. y_i^{trunc} \leq \hat{y} < y_i^{trunc} + 1 \tag{2}$$

where:

$\hat{y}_i$: Recovered HCA progression at sampling point i

y_i^{trunc}: Observed HCA progression at sampling point i

T_i: Outdoor temperature at sampling point i

$\Delta\hat{y}_i = \hat{y}_i - \hat{y}_{i-1}$: Recovered HCA increment.

$\Delta T_i = T_i - T_{i-1}$: Outdoor temperature increment.

λ: Weighting coefficient between the two terms.

The first term reflects the continuity of the HCA increment change, and it penalizes abrupt changes in energy increments, promoting temporal smoothness. While the second term encourages a negative correlation between energy increments and outdoor temperature increments, ensuring consistency with the weather compensation control in the SH system. Especially, the initial guess of the decimal will start from 0.5 since this is the middle point of the interval.

3 Case Study

In this analysis, two case studies were selected for testing and verifying the methodology. The first one is a laboratory environment, the second one an existing residential building.

3.1 Italian Case—INRIM Lab

The first dataset originates from a laboratory environment at the National Metrology Institute of Italy (INRIM). This laboratory, known as the Heat Accounting Experimental Mock-Up, is specifically designed for the calibration and evaluation of HCAs [9].

The test facility features a central heating system powered by a gas boiler, connected to 40 radiators of various types, sizes, and materials. These radiators are distributed across four floors and linked through hydraulic loops, which can be easily configured for either vertical riser or horizontal connections. Each radiator is equipped with an HCA for data recording and a Direct Heat Meter (DHM) for the reference measurement of the radiator heat output, traceable to the International System of Units (SI). The lab also has a temperature sensor to record the outdoor temperature.

In February 2025, an experiment using all 40 radiators was conducted. For this experiment, only integer values are obtained through the billing system. To acquire

Table 1 Characteristics of radiators in the INRIM lab

Radiator type	Number	Capacity (W)	Original *K*	Modified *K*
Cast aluminum sectional 1	16	1467	1.807	10
Cast aluminum sectional 2	4	815	1.026	10
Four columns of cast iron 1	4	1427	1.923	10
Four columns of cast iron 2	2	713.5	0.997	10
Four columns of tubular steel 1	4	1482	1.859	10
Four columns of tubular steel 2	2	798	1.041	10
Heated towel rail	8	395	0.306	10

decimal values, we manually increased the HCA's rating factor. The specific details of this adjustment are presented in the Table 1.

After increasing the rating factor for all the radiators to the maximum value of 10, the progression of the HCA enlarged to 5.2–32.5 times, which is equivalent to adding the decimals of 0.2–0.03. For large radiators, the effect didn't strictly reach "adding one decimal", but in this study, the HCA data with decimals is only used as references rather than inputs; hence, this will not give a big effect on the final results.

3.2 Danish Case—Viborg Apartment Building

The second dataset originates from a multi-story residential building in Viborg, Denmark, constructed in the 1970s. We focused on a section including 15 apartments equipped with 84 radiators, each mounted with HCA. The data collection period spans January 2021. Typically, HCA data in real-world buildings is only available via the billing system and recorded as integers. However, for this project, we obtained high-resolution backend HCA data with decimal values directly from the HCA provider. Weather data for Viborg was sourced from the Danish Meteorological Institute (DMI).

4 Results and Discussions

The proposed method was applied to raw data from the INRIM laboratory. Results from radiator #4 are shown in Fig. 1, illustrating the HCA increments of the truncated data, the reference values with decimals, and the recovered data. The radiator was randomly selected for illustration purposes. As shown, the original truncated HCA increments fluctuate between integer values, while the recovered data closely follows the reference curve.

Similarly, Figs. 2 and 3 present results from the case study conducted in Viborg. In both cases, the recovered curves show a strong agreement with reference data.

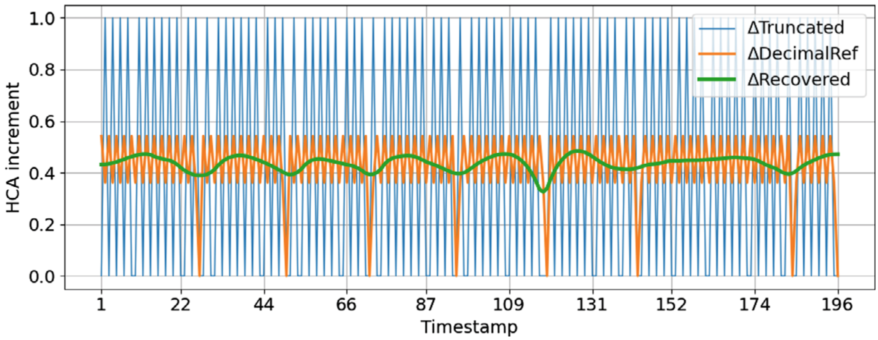

Fig. 1 HCA increment of INRIM radiator #4

Notably, Fig. 3 demonstrates that even when the reference HCA increments exhibit varying jump ranges (e.g., between 0–1 or 0–2), the recovered values still closely track the reference curve, highlighting the robustness of the proposed approach.

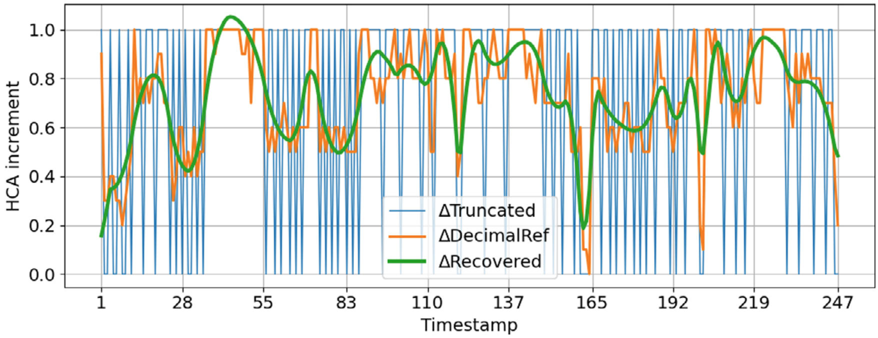

Fig. 2 HCA increment of Viborg radiator #1

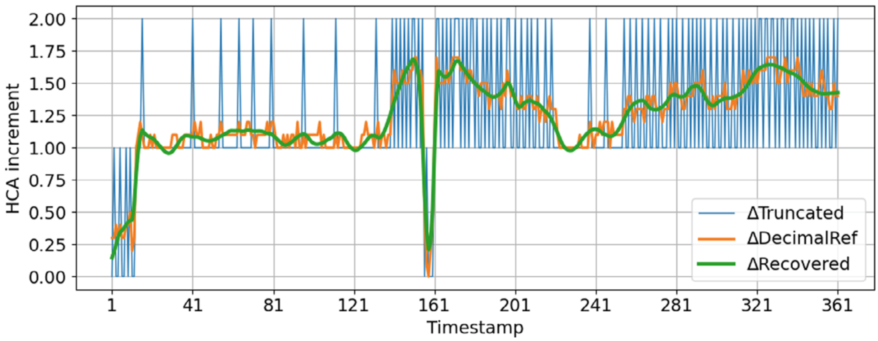

Fig. 3 HCA increment of Viborg radiator #4

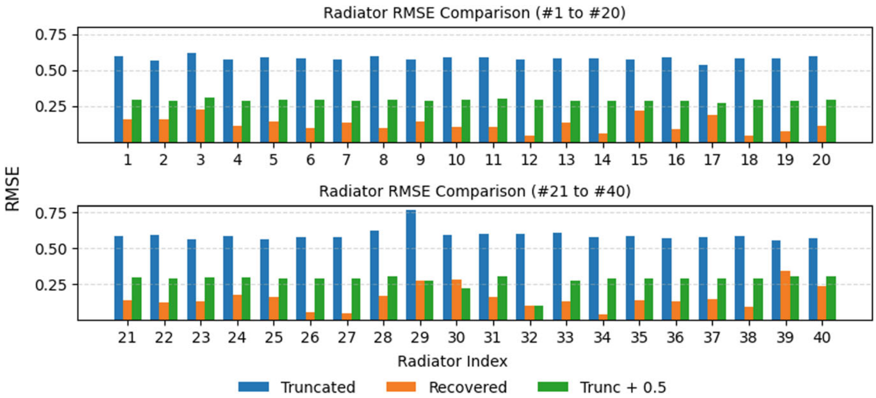

Fig. 4 RMSE between the reference true value and the truncated, recovered, and plus-0.5 guess data for all radiators in the INRIM lab

The data also highlights the difference between lab conditions and real buildings: the heat output from the lab radiator was much more stable compared to that of the radiators in the real building.

For an overall picture of the performance of the methodology across all radiators, Figs. 4 and 5 summarize the root mean square error (RMSE) of the truncated value, recovered value, and the truncated value plus 0.5 compared to the reference value, for both the INRIM and Viborg cases. The value 0.5 comes from the midpoint of the truncation interval and represents a simple guess to address the truncation issue.

In the INRIM lab, recovered data reduced the RMSE by 76.9% compared to the truncated data, while the plus 0.5 guess method reduced it by 51.6%.

The main difference between the lab and real building environments is that in the lab, we can operate all radiators, whereas in real buildings, some users may choose to use only part of the system and keep some radiators turned off. Among the 84 radiators, 18 were completely off, such as #8, #43, and #83. Excluding these radiators, the average RMSE reduction achieved by the recovery method is 60.4%, while the plus 0.5 method achieves a reduction of 44.3%. Specifically, the results for radiators #11 and #32 exhibited an unusual behavior: the truncated data showed a low RMSE, while the recovered data yielded a significantly higher RMSE. Upon examining the original data, it was observed that these radiators alternated between on and off states. Since the current method lacks the ability to detect the operational status of the radiators, it continued to compensate for gaps even during periods when the radiator was off. This inappropriate compensation led to amplified errors and, consequently, a higher RMSE. In future work, we plan to enhance the method by incorporating mechanisms to better identify the radiator's on/off status and thus avoid this issue.

The method demonstrated good RMSE performance in both laboratory and real-building settings, confirming its robustness across different conditions. Building on this, the improved reliability of short-interval HCA data can support more refined

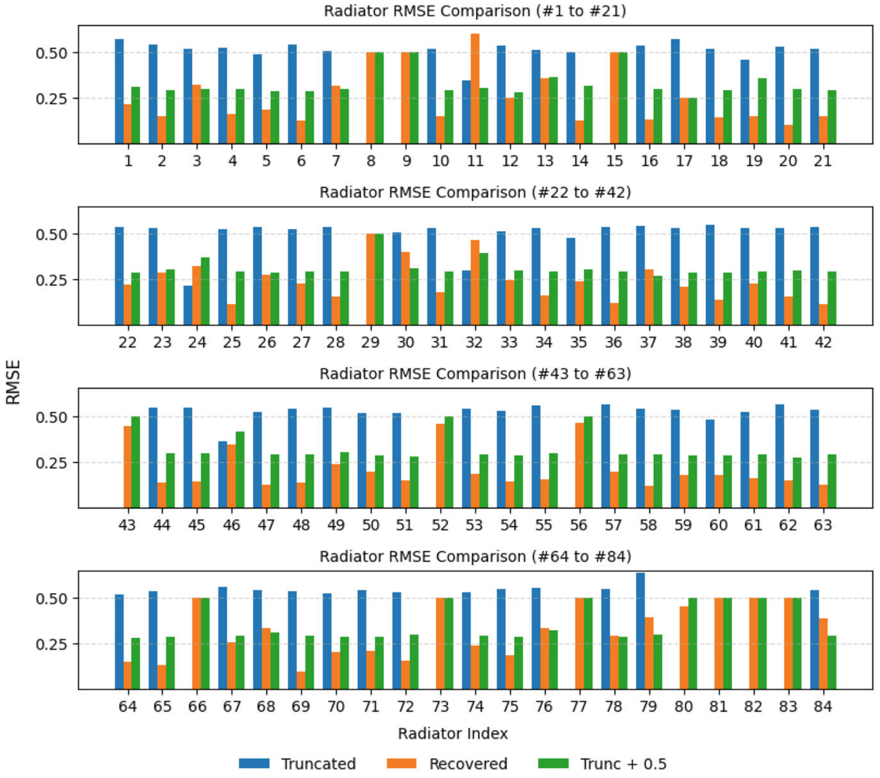

Fig. 5 RMSE between the reference true value and the truncated, recovered, and plus-0.5 guess data for all radiators in the Viborg building

analyses, such as identifying faulty radiators, detecting high-load apartments, and estimating the minimum required supply temperature for a given building. While HCAs are not designed for control purposes, the enhanced data quality can still contribute to radiator-level diagnostics and basic operational observations.

5 Future Work

An interesting direction for future research concerns the role of the weighting coefficient λ in the objective function. In this study, λ was used to balance two terms: temporal continuity and correlation with outdoor temperature. Preliminary analysis, as illustrated in Fig. 6 for one representative radiator, suggests that increasing λ generally leads to a decline in recovery performance. This might indicate that the simple correlation with outdoor temperature does not directly improve accuracy in the current setting.

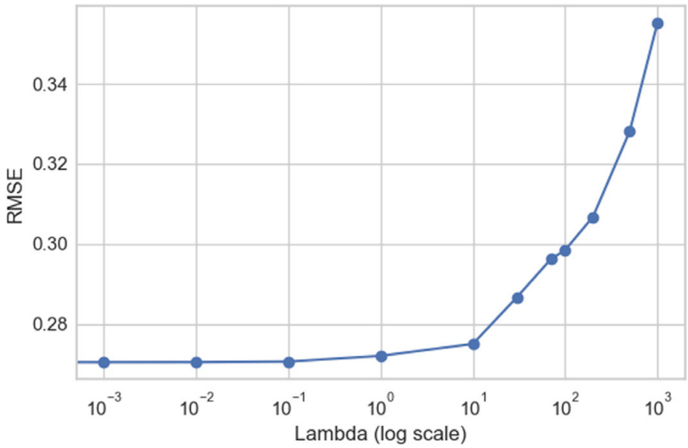

Fig. 6 RMSE between the recovered HCA increment and the true value under different weighting coefficients of Viborg radiator #1

One possible explanation is that the thermal inertia of the heating system introduces a time delay between the radiator's response and the outdoor temperature changes. If this delay is not accounted for, the correlation term may introduce noise rather than guidance. Exploring this potential time shift and more advanced correlation structures is beyond the scope of the current study, but presents a promising avenue for future investigation.

6 Conclusion

Although HCAs are primarily designed for cost allocation rather than energy metering, their data holds potential for future applications such as fault detection and hydronic system optimization, especially in contexts requiring high temporal resolution. This study proposes a method to recover the truncated decimal values in HCA progression. Under laboratory conditions, the method reduces the RMSE of HCA increments by 76.9%, and by 60.4% in real building scenarios. By significantly improving the accuracy of HCA data, the proposed method enhances its usability beyond billing, enabling more reliable high-resolution analyses. As such, it offers an effective solution to the truncation problem and supports the broader application of HCA data in future hydronic system assessments.

Acknowledgements We gratefully acknowledge the support and collaboration of Brunata and ZENNER International, and Viborg Varme.

References

1. Canale L, Dell'Isola M, Ficco G, Cholewa T, Siggelsten S, Balen I (2019) A comprehensive review on heat accounting and cost allocation in residential buildings in EU. Energy Build 202:109398. https://doi.org/10.1016/j.enbuild.2019.109398
2. The European parliament and the council of the European union, directive 2012/27/EU of the European parliament and of the council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC (Text with EEA relevance). http://data.europa.eu/eli/dir/2012/27/oj, 2012
3. Tunzi M, Benakopoulos T, Yang Q, Svendsen S (2023) Demand side digitalisation: a methodology using heat cost allocators and energy meters to secure low-temperature operations in existing buildings connected to district heating networks. Energy 264:126272. https://doi.org/10.1016/j.energy.2022.126272
4. Yang Q, Salenbien R, Smith KM, Tunzi M (2024) Identifying untraced faults associated with high return temperatures from heating systems in buildings connected to district heating networks. Energy 309:133097. https://doi.org/10.1016/j.energy.2024.133097
5. Danfoss, Thermostatic sensors type RA 2000. Accessed Jun. 19, 2025. [Online]. Available: https://assets.danfoss.com/documents/107870/AI015186402271en-010601.pdf
6. Leiria D, Johra H, Marszal-Pomianowska A, Pomianowski MZ (2023) A methodology to estimate space heating and domestic hot water energy demand profile in residential buildings from low-resolution heat meter data. Energy 263:125705. https://doi.org/10.1016/j.energy.2022.125705
7. Schaffer M, Leiria D, Vera-Valdés JE, Marszal-Pomianowska A (2023) Increasing the accuracy of low-resolution commercial smart heat meter data and analysing its error. https://doi.org/10.35490/EC3.2023.208
8. EN 834:2013 Heat cost allocators for the determination of the consumption of room heating radiators—appliances with electrical energy supply (2013)
9. Marinari C, Saba F, Grisotto F, Orio M, Masoero M (2015) The INRiM thermo-hydraulic mock-up for thermal energy measurement devices: design, construction and metrological characterization. Energy Procedia 78:2304–2309. https://doi.org/10.1016/j.egypro.2015.11.378