

<https://doi.org/10.7250/CONNECT.2025.002>

# APPLYING AI TO HVAC SYSTEMS: A CRITICAL ANALYSIS OF DATA-RELATED CHALLENGES

Dalia Mohammed Talat Ebrahim ALI<sup>1\*</sup>, Violeta MOTUZIENĖ<sup>2</sup>

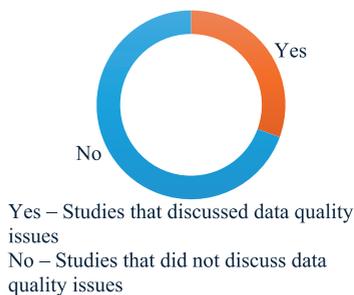
<sup>1,2</sup> Lithuanian Energy Institute, Breslaujos gatvė 3 LT-44403 Kaunas, Lithuania.

Vilnius Gediminas Technical University, Sauletekio Av. 11, LT-10223 Vilnius, Lithuania\*

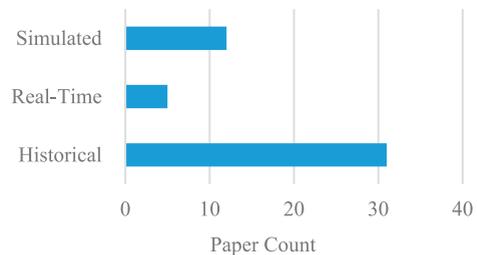
\* **Corresponding author.** Email address: [dalia.ali@lei.lt](mailto:dalia.ali@lei.lt)

**Abstract** – Integrating AI in HVAC systems is a promising approach that helps enhance energy efficiency in buildings, which leads to cost savings and provides environmental benefits. However, the effective performance of these AI models, especially in HVAC systems, depends not only on the model design but also on the data's quality, reliability, size, availability, and management. Data plays an important role in determining the accuracy and reliability of the AI model's performance. This paper analyses recent studies that apply AI models to achieve energy efficiency in HVAC systems from a data perspective, examining various aspects of data management in Deep Learning and Hybrid models applied to HVAC in buildings, such as data availability, the different data sources, type, quality issues, and data splitting methods. Through this analysis, the paper aims to provide insights into data-related challenges and recommend ways to overcome and mitigate them to develop AI models that perform more effectively. The paper highlights the importance of developing better data-handling practices to have more accurate, efficient, and reliable AI models in HVAC systems. The findings reveal that combining multiple data types can enhance model performance and generalizability. Moreover, the analysis concludes that the main data type for residential buildings is simulated data rather than real-world data; this could be due to privacy concerns. Meanwhile, commercial buildings have commonly utilized more structured and reliable dataset sources, enabling more precise modelling. The findings also indicate that data quality is overlooked by researchers in many studies, where only 31 % of the analysed papers discussed quality issues, reflecting that it is not yet a standard practice in this field. Additionally, this analysis addresses the scarcity of reliable and audited data. Therefore, and in response to this issue, this paper recommends accessible and reliable data resources that can be employed in AI applications for HVAC systems in buildings.

**Keywords** – Building management systems; deep learning; energy efficiency; machine learning



(a)



(b)

(a) Proportion of studies discussing data quality issues; (b) number of studies by data type.