Causal Forecasting for Robust Explainable Predictions in Smart Buildings

Simone Rossato, Daniele Meli, Alessandro Farinelli

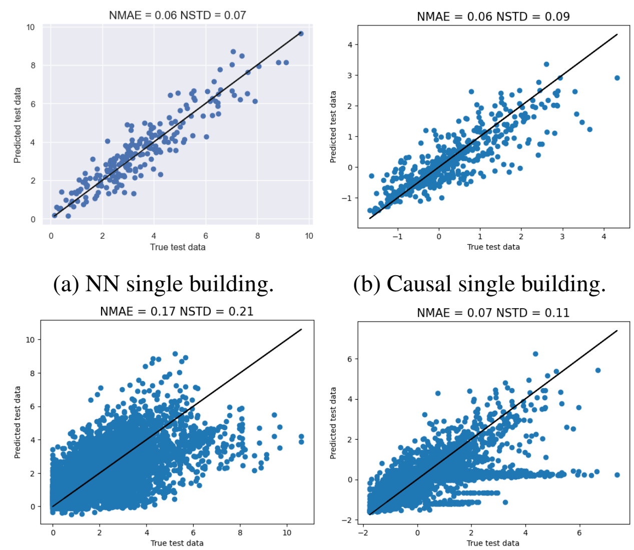
Department of Computer Science, University of Verona, Italy
name.surname@univr.it

High electricity demand raises prices and the overall cost of power distribution networks. The demand mostly comes from civil buildings, e.g., for cooling and heating, with sparse distribution depending on population density (Serrano et al. 2017). It then becomes crucial to optimize power distribution according to the needs of different areas. As evidenced by Jiang et al. 2023, efficient optimization in smart buildings requires the combination of accurate demand forecasting, in order to make the best decisions in the long term. Machine learning, especially deep learning models, have shown promising results in this direction, allowing to accurately forecast the power demand from large sets of distributed sensors (Li and Yao 2021). However, fundamental challenges in the smart buildings domain are the high dimensionality of data, the intricacies introduced by the distributed nature of the system, and the requirement of robust predictions over a medium / long horizon. In addition, explainability is of utmost importance to foster monitoring and predictive maintainance (Kazmi, Fu, and Miller 2023).

In this abstract, we show preliminary results about *cooling demand forecasting in smart buildings*, adopting a *causal discovery approach* (Assaad et al. 2022). Causal discovery allows to capture meaningful spatio-temporal patterns in large datasets, fostering both interpretability and forecasting robustness. We consider the popular benchmark *citylearn* dataset[1], containing data from three buildings. We compare the forecasting capabilities of different models, including random forest, Neural Network (NN) and recurrent NN. Forecasting over *a single building*, NN results in the best Normalized Mean Average Error (NMAE) from Fig. 1a; however, causal discovery[2] attains the same NMAE (0.06), with a slightly superior standard deviation (0.09 vs. 0.07). Interestingly, when training a model on a single building and *forecasting over different ones*, causal discovery achieves the best results, with NMAE 0.07 (Fig. 1d, comparable to Fig. 1b) vs. 0.17 by NN (Fig. 1c). This proves the *superior generalization and robustness of causal forecasting*. Causal discovery also reduces the number of relevant variable links up to 40%, *enhancing explainability*.

[1]https://www.citylearn.net/index.html

[2]https://github.com/jakobrunge/tigramite



(a) NN single building. (b) Causal single building.

(c) NN generalization. (d) Causal generalization.

Figure 1: Forecasting results.

References

Assaad, C.; et al. 2022. Survey and evaluation of causal discovery methods for time series. *Journal of Artificial Intelligence Research*, 73: 767–819.

Jiang, W.; Yi, Z.; Wang, L.; Zhang, H.; Zhang, J.; Lin, F.; and Yang, C. 2023. A Stochastic Online Forecast-and-Optimize Framework for Real-Time Energy Dispatch in Virtual Power Plants under Uncertainty. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 4646–4652.

Kazmi, H.; Fu, C.; and Miller, C. 2023. Ten questions concerning data-driven modelling and forecasting of operational energy demand at building and urban scale. *Building and Environment*, 239: 110407.

Li, X.; and Yao, R. 2021. Modelling heating and cooling energy demand for building stock using a hybrid approach. *Energy and Buildings*, 235: 110740.

Serrano, S.; et al. 2017. Heating and cooling energy trends and drivers in Europe. *Energy*, 119: 425–434.