Hierarchical Load Forecasting:
2017 Global Energy Competition

Inés Dormoy, Louis Gautier
ICME & MS&E, Stanford University
idormoy@stanford.edu, lgautier@stanford.edu

Dataset. Our project will consist in forecasting energy load on the 2017 Global Energy Forecasting Competition (GEFCom2017) [1] final match dataset. This dataset contains seven years (from 2005 to 2011) of hourly load data measured at three levels of hierarchy (see Appendix A), from individual meters to intermediary stations aggregating the consumption of several meters. The bottom level of the hierarchy includes a total of 183 meters. This dataset also comes with temperature and relative humidity data from 28 weather stations for which the location is not provided.

Project Idea. We will be focusing on load forecasting for all individual meters and aggregation stations in the last year of the dataset (2011), and use the other 6 years for training (originally, the competition evaluated the performance on the year 2012, but we don't have access to this test data).

Background. Load forecasting, which consists of predicting the demand for energy, is crucial to the energy sector. Several load forecasting competitions have been held in 2012 [2], 2014 [3] and 2017 [1]. The 2017 one is the last one organized to date as well as the one for which hierarchical information is the most important. The teams in the competition first implemented data cleaning, weather station selection, and feature engineering. They then used models as varied as Quantile Regression [4], Generalized Additive Models [5], Gradient Boosting [6], Quantile Random Forest [7] and SARIMA [8] models. Few teams took into account the hierarchical aspect of the data, probably because hierarchical forecasting methods weren't widely used in 2017 [1]. Since 2017, hierarchical methods have been further developed, including bottom-up and top-down approaches[9], trace minimization [10], and deep learning approaches [11]. In this project, our goal is to compare baseline methods not leveraging the hierarchical information with hierarchical methods in terms of performance, inference time, and interpretability.

Models. We are first planning to use non-hierarchical models as a baseline for load forecasting. On the one hand, we would like to perform basic feature engineering and train ensemble models like *LightGBM* that don't directly model the time dependency in the data. On the other hand, we would like to fit widely used time series models, like SARIMA models, to forecast the load at each station individually. We would then like to use several models that combine the hierarchical structure of the load measures and the time-related dependency of the load data. First, we are planning to use the optimal combination of hierarchical time series presented in [12] and [10]. These methods allow to reconcile time series forecasts at different levels of hierarchy in an optimal way, better than in a top-down or bottom-up fashion. Then, we are planning to use a deep-learning-based approach, inspired by [11], which will allow us to try deep sequence models on this problem while still taking the hierarchical structure of the data into consideration.

Metrics. The prediction task on this dataset is challenging as the number of times series to forecast is very high. Therefore, we have to carefully define metrics that allow us to benchmark our different models. In Energy load forecasting, balancing demand with supply is crucial, and large forecast errors can lead to serious issues like blackouts. We therefore want to use the RMSE as a metric, as well as MAE and MAPE. To make sure that we are computing each level of hierarchy right, we will compute the sum of the Mean squared errors of all stations at each level of hierarchy.

Software. We will use open-source Python libraries for our project, including *statsmodels* for Time series-based models, and *PyTorch* for deep learning-based sequence models.

Milestone. By the milestone, we expect to have results for baselines that don't exploit the hierarchical structure of the dataset. We would like to compare it with at least one hierarchical model.

References

[1] Tao Hong, Jingrui Xie, and Jonathan Black. Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting. *International Journal of Forecasting*, 35(4):1389–1399, 2019.

[2] Tao Hong, Pierre Pinson, and Shu Fan. Global energy forecasting competition 2012. *International Journal of Forecasting*, 30(2):357–363, 2014.

[3] Tao Hong, Pierre Pinson, Shu Fan, Hamidreza Zareipour, Alberto Troccoli, and Rob Hyndman. Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *International Journal of Forecasting*, 32, 03 2016.

[4] Lingxin Hao and Daniel Q Naiman. *Quantile regression*. Number 149. Sage, 2007.

[5] Trevor J Hastie and Robert J Tibshirani. *Generalized additive models*, volume 43. CRC press, 1990.

[6] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[7] Nicolai Meinshausen and Greg Ridgeway. Quantile regression forests. *Journal of machine learning research*, 7(6), 2006.

[8] M. Braun, T. Bernard, O. Piller, and F. Sedehizade. 24-hours demand forecasting based on sarima and support vector machines. *Procedia Engineering*, 89:926–933, 2014. 16th Water Distribution System Analysis Conference, WDSA2014.

[9] Chul-Yong Lee and Sung-Yoon Huh. Forecasting new and renewable energy supply through a bottom-up approach: The case of south korea. *Renewable and Sustainable Energy Reviews*, 69:207–217, 2017.

[10] George Athanasopoulos Shanika L. Wickramasuriya and Rob J. Hyndman. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, 114(526):804–819, 2019.

[11] Syama Sundar Rangapuram, Lucien D Werner, Konstantinos Benidis, Pedro Mercado, Jan Gasthaus, and Tim Januschowski. End-to-end learning of coherent probabilistic forecasts for hierarchical time series. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 8832–8843. PMLR, 18–24 Jul 2021.

[12] Rob J Hyndman, Roman A Ahmed, George Athanasopoulos, and Han Lin Shang. Optimal combination forecasts for hierarchical time series. *Computational statistics & data analysis*, 55(9):2579–2589, 2011.

A Hierarchical organization of the data

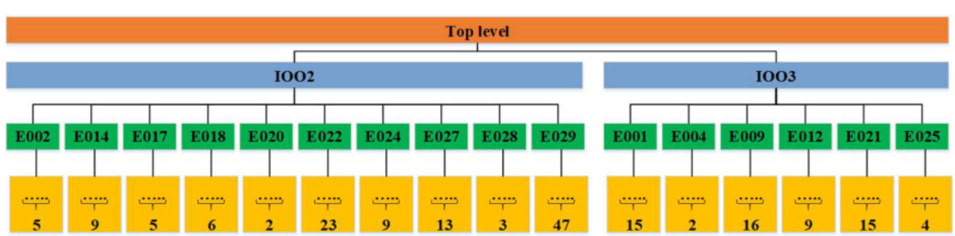

Figure 1: Hierarchical organization of the data, where the numbers at the bottom level indicate the numbers of forecasting locations (total of 183 meters at the base level).