Details about the custom loss function

Zhixian Wang[1], Qingsong Wen[2,*], Chaoli Zhang[2], Liang Sun[2], Leandro Von Krannichfeldt[3], Yi Wang[1,*]
[1] The University of Hong Kong [2] DAMO Academy, Alibaba Group
[3] École Polytechnique Fédérale de Lausanne
{zxwang,yiwang}@eee.hku.hk, {qingsong.wen,chaoli.zcl,liang.sun}@alibaba-inc.com, leandro.vonkrannichfeldt@epfl.ch

1 custom loss function

1.1 Overall process of the custom loss function

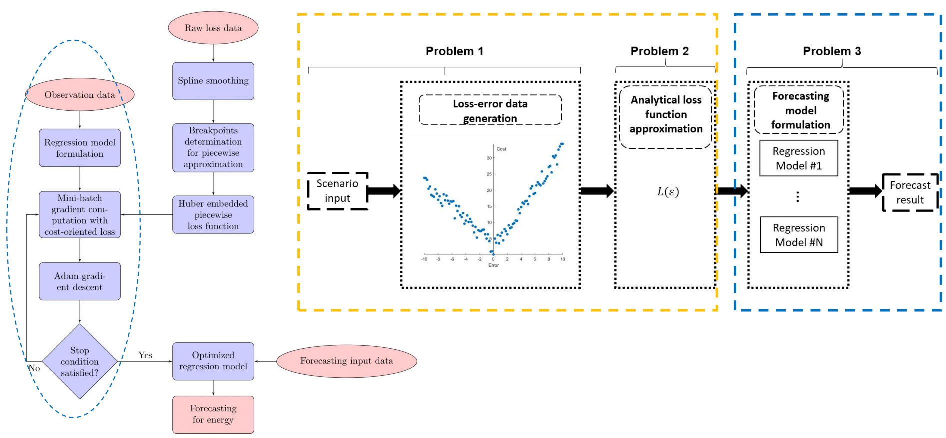

Figure 1: The complete process of constructing a custom loss function, the blue box represents the training process of the regression model (neural network).

An important difference between load forecasting and other forecasting is that load forecasting has a clear downstream task: grid dispatching. Therefore, compared to pursuing higher accuracy, load forecasting focuses more on how to lower the cost of downstream dispatching tasks. In recent years, cost-oriented load forecasting has become a field research focus[1; 2]. According to [3], there exists an asymmetric relationship between forecasting errors and dispatching costs. Specifically, the dispatching cost arising from a predicted value that is higher than the actual value exhibits different characteristics compared to a predicted value that is lower than the actual value, resulting in an asymmetric impact on the overall dispatching process. Based on this situation, a simple and intuitive method is to directly use dispatching cost as a loss function to optimize the model during the training process. To achieve this idea, we can divide the entire process into three parts:

*Corresponding authors.

Submitted to the 37th Conference on Neural Information Processing Systems (NeurIPS 2023) Track on Datasets and Benchmarks. Do not distribute.

generating forecasting error and dispatching cost data, fitting cost-oriented loss functions, and training forecasting models(shown in Figure 1). Among them, problem 2 is to use piecewise linearization to construct a loss function that can be used for neural network training, and its details have been described in section 3.1 in the appendix(our code mainly provides the functionality of constructing loss functions that can be used for training). Problem 3 is the training process of the forecasting model, which is not different from the general training process except for using the function fitted from the second problem as the loss function. Therefore, we will mainly focus on details about the problem 1.

1.2 Details of the problem 1

Although power grid dispatch is crucial for load forecasting, due to legal and privacy limitations, it is difficult to find a perfect match between load data and power grid dispatch data in general. A compromise approach is to convert the mismatched data to obtain simulated results(shown in Figure 2). In [3], a modified IEEE 30-bus test system given in [4] is used to simulate an economic dispatch optimization problem in a regulated municipal-level grid system. Two dispatch optimization problems, namely the DAED problem and the IPB problem(the mathematical formulations of the two problems can be found in [3]), are formulated to compute the loss data. Alike the actual operation process, the DAED problem schedules the power output of all online generators based on the forecasted daily load profile $\hat{y}_i$ in the controlled region. On this basis, the IPB problem with the actual load y_i is implemented to adjust the outputs of generators, charging/discharging behavior of Battery Energy Storage Systems, and load shedding to balance the intra-day load deviation. With established system parameters, the cost obtained by the IPB problem is viewed as the actual total cost $C(\hat{y}_i)$. The actual total costs are then compared against the ideal total costs $C(y_i)$ (i.e., the DAED problem is solved with actual load y_i) to calculate the $C_i\left(\widehat{y}_i, y_i\right) = \frac{C(\widehat{y}_i)-C(y_i)}{C(y_i)} \times 100\%$(the definition here is consistent with our paper, which represents the normalized cost). The load data used here is GEF12[5] and it is assumed to be distributed proportionally to each node based on the magnitude of the load provided in the modified IEEE 30-bus test case. The relevant code can be found in our GitHub.

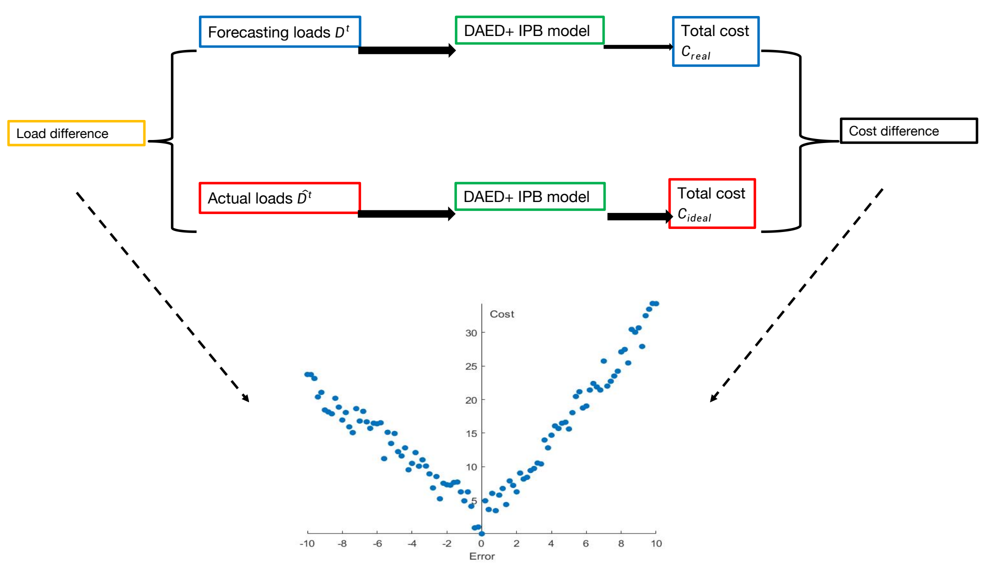

Figure 2: Heuristic data generation.

References

[1] J. Han, L. Yan, and Z. Li, "A task-based day-ahead load forecasting model for stochastic economic dispatch," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5294–5304, 2021.

[2] H. Kebriaei, B. N. Araabi, and A. Rahimi-Kian, "Short-term load forecasting with a new nonsymmetric penalty function," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1817–1825, 2011.

[3] J. Zhang, Y. Wang, and G. Hug, "Cost-oriented load forecasting," *Electric Power Systems Research*, vol. 205, p. 107723, 2022.

[4] P. Hota and A. Naik, "Analytical review of power flow tracing in deregulated power system," *American Journal of Electrical and Electronic Engineering*, vol. 4, no. 3, pp. 92–101, 2016.

[5] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," 2014.