No. 21

Day-ahead prediction of wind power production with multiple numerical weather prediction data and machine learning algorithms

Gang Huang, Minh-Thang Do
Meteodyn

Objectives

Since the 1950s, Numerical Weather Prediction (NWP) has provided increasingly accurate climate data for both weather forecast and industrial applications. Knowledge of the state and evolution of the atmosphere are essential to a good prediction of wind power production for the benefit of both producers and grid operators. In this study we aim to know to how an energy forecast can be improved by combining NWP data from multiple sources, and the impact of choosing different machine learning (ML) architectures on forecast accuracy and uncertainty.

Methods

Wind power output for 6 months has been collected from a wind farm located in the West Indies. We use weather forecast data from GFS [1], ECMWF [2] and Météo-France (ARPEGE and AROME) [3]. Meso-scale data for the wind farm are extracted using a cubic interpolation. In total, 117 features are created by combining data from 4 NWP sources.

Two machine learning frameworks are used for the forecast task: Keras, with its fully connected neural networks (NN) and LightGBM. NN models deduce the relationship from input features to the target variable through the determination of weights at each node of each layer of the NN. LightGBM, on the other hand, is based on an ensemble of decision trees to decide how each input variable can be used to predict a target value.

The key issue in the training of ML models is to keep learning and extracting patterns from available data without losing its capacity to generalize on unseen data. To this end, several hyper-parameters need to be tuned to find a balance between model complexity and over-fitting (Table 1 and 2).

Hyperparameter	Values
Learning rate	0.01
Batch size	32, 128, 256
Layer configuration	[8], [16], [32], [32, 32], [64, 64], [32, 64, 8]
Dropout	0, 0.05, 0.25

Table 1 Hyperparameters used for NN models. Higher batch size and dropout with simpler layer configurations help with preventing overfitting.

Hyperparameter	Values
Learning rate	0.01
max_depth	2, 3, 4, 5, -1 (no limit)
min_child_samples	200, 500, 1000
colsample_bytree	0.01, 0.2, 0.4, 0.9

Table 2 Hyperparameters used for LightGBM models. Lower "colsample_bytree, "max_depth" and higher "min_child_samples" help with preventing overfitting.

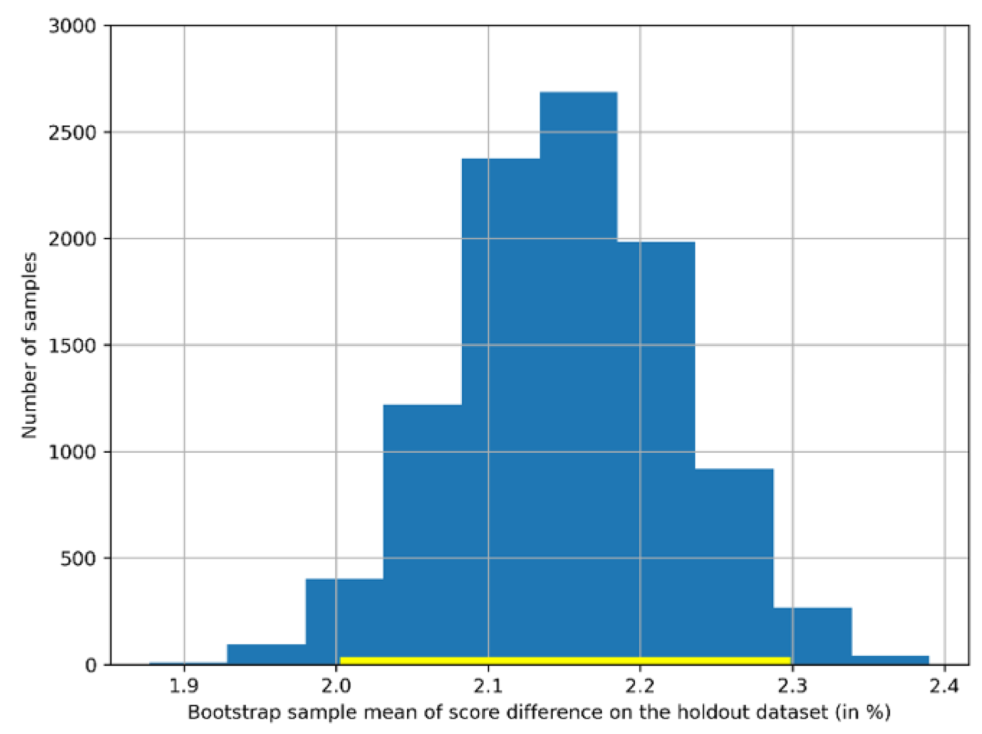

Figure 1 Relative improvement of the forecast score by introducing extra NWP data sources (AROME and ARPEGE). The yellow line marks the 95% confidence interval estimated using the bootstrap technique.

Principal Findings

Models are evaluated by a score, 100% being the perfect score, calculated from the mean daily averaged RMSE (root-mean-square error) over unseen data by five-fold cross-validation. With LightGBM, the improvement with extra NWP data from Météo-France is significant with an average improvement of more than 2 percent (Figure 1). For NN models, complex models do better, since the overall performance of simpler models are negatively impacted by the relatively high number of features and correlations between them (Figure 2). For LightGBM models, the overall performance is better and the spread less than NN models (Figure 3), in accordance with the finding in [4]. An optimal choice of hyper-parameters can be found among the best performing models with the least spread.

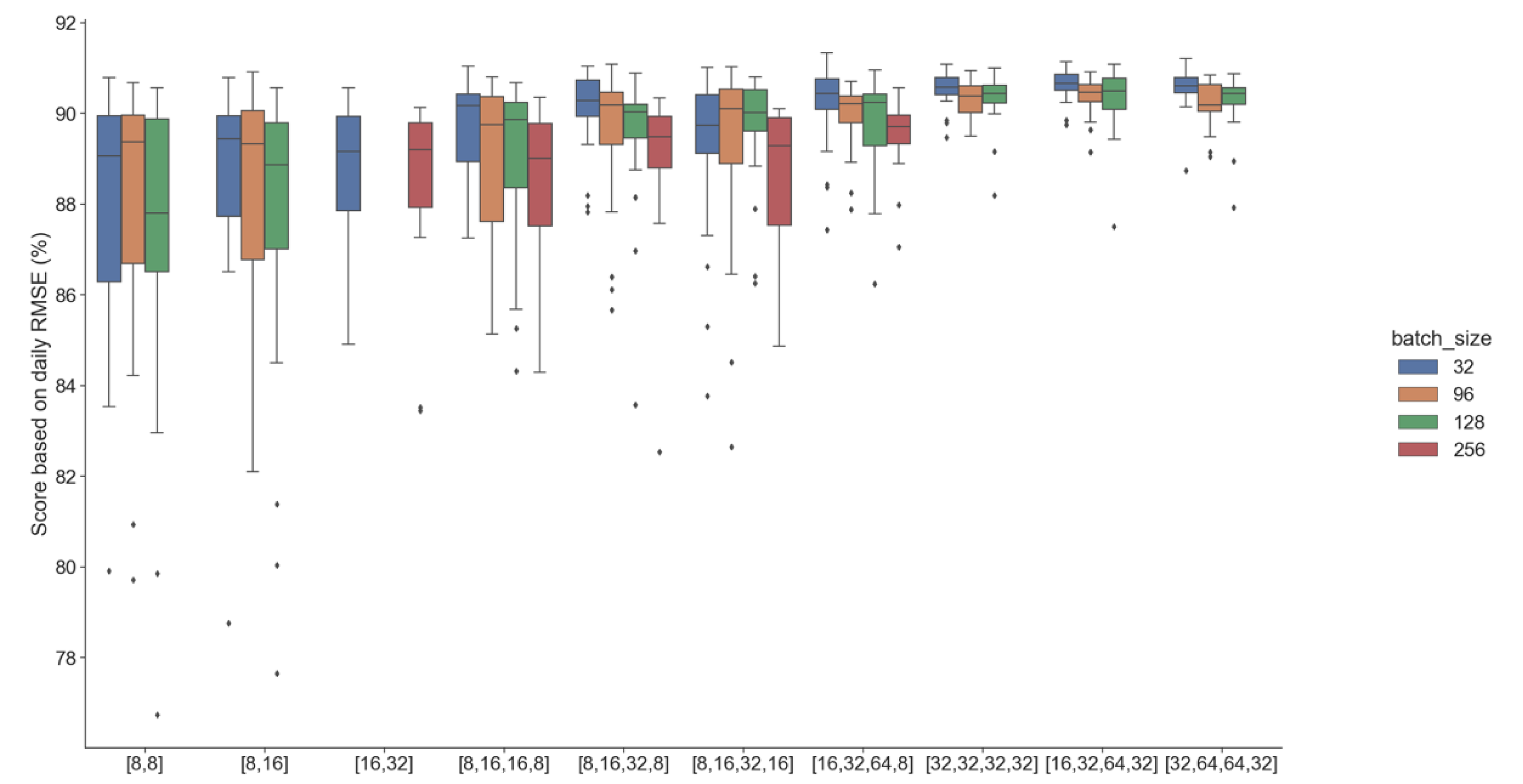

Figure 2 NN forecast score as a function of layer configuration (horizontal axis) and batch size (colors).

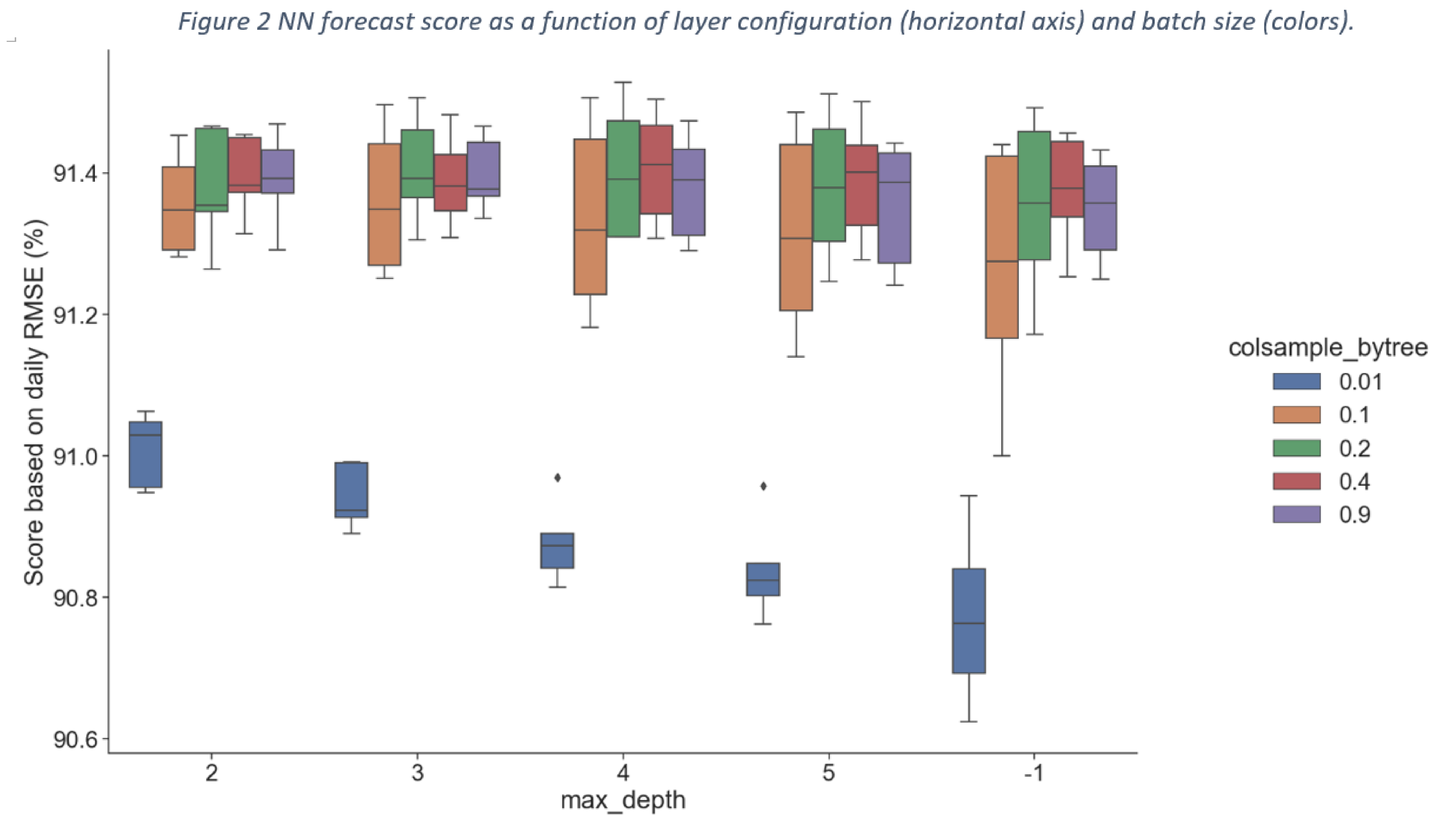

Figure 3 LightGBM forecast score as a function of the hyperparameters max_depth (horizontal axis) and colsample_bytree (colors). For comparison, the best mean score in Figure 2 for NN models is around 90.6%.

Conclusions

Provided with a combination of 4 NWP sources, by varying hyper-parameters and randomization seeds, the LightGBM models, with lower computational cost, are shown to provider more consistent and better predictions on the day-ahead wind power production than the models based on neural networks. It should be noted that the tabular nature and relatively small size of our dataset (less than 50,000 samples) may contribute to the tree-based model's superiority in this benchmark. In the future, a preliminary step of feature selection can be added before the forecast task, and it remains to be seen whether NN models can improve their predictions with more data.

References

1. NOMADS-NOAA Operational Model Archive and Distribution System, accessed 30 January 2023, https://nomads.ncep.noaa.gov.
2. Forecasts ECMWF, accessed 30 January 2023, https://www.ecmwf.int/en/forecasts.
3. Données Publiques de Météo-France – Accueil, accessed 30 January 2023, https://donneespubliques.meteofrance.fr.
4. Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data?. In Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track.

WEMC World Energy & Meteorology Council

ICEM 2023 #icem2023Italy

Reach us at gang.huang@meteodyn.com, twitter @meteodyn and on LinkedIn.

icem 2023