Multi-output Ensembles for Multi-step Forecasting

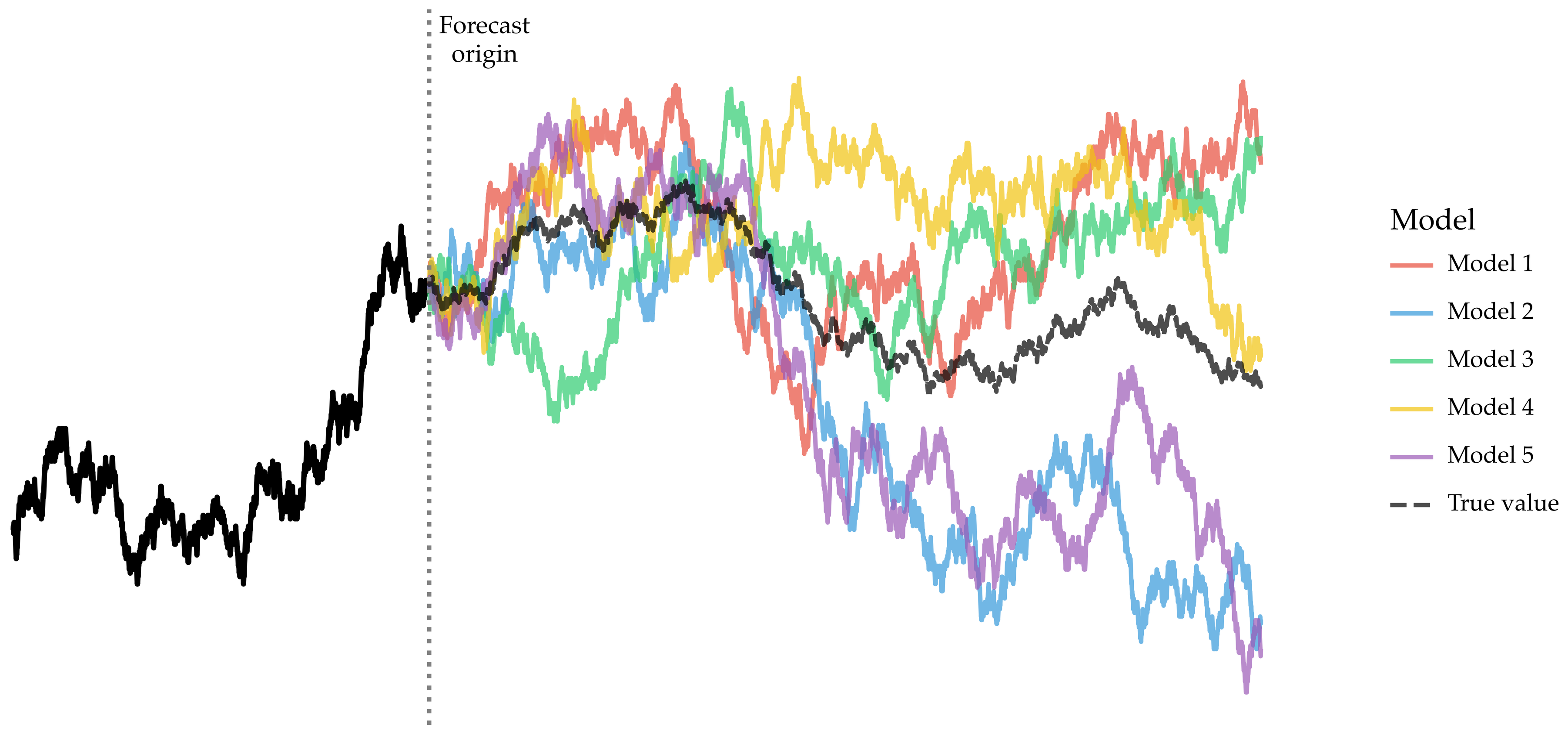
U.PORTO FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Vitor Cerqueira[1,2] Luis Torgo[3]

[1]Faculdade de Engenharia da Universidade do Porto, Portugal [2]Laboratory for Artificial Intelligence and Computer Science (LIACC), Portugal

[3]Dalhousie University, Canada

Center for Responsible AI



CONTEXT

- **Problem:** Multi-step univariate time series forecasting
- **Approach:** Ensemble methods based on multi-output regression algorithms
  - Local auto-regressive methodology (model fit for each individual time series)
  - Individual forecasts are dynamically combined to cope with changes
- **Relevance:**
  - Improved forecasting accuracy via ensembling
  - Reducing the chance of picking the wrong model

GOALS

- **Objective 1:** To determine which dynamic combination rule performs best for multi-step ahead problems, extending previous analyses that focused on one-step ahead tasks.
- **Objective 2:** To identify the optimal approach for computing ensemble weights along the forecasting horizon.

RESEARCH QUESTIONS

- RQ1: How do different dynamic combination methods compare with each other in terms of their rank across all data sets?
- RQ2: What is the best approach for computing the weights along the forecasting horizon?
- RQ3: How does each dynamic combination method compare with a static ensemble that assigns equal weights to all models along the forecasting horizon?

METHODS

- `Simple`: Combination rule which assigns equal weights to all models. In practice, the predictions of the available models are combined using the arithmetic mean;
- `Window`: Dynamic weighted average of the predictions of the available models. The weights are computed according to the forecasting performance in the last λ observations;
- `Blast`: A variant of the `Window` approach. Instead of using past recent performance to weigh the available models, the idea is to select the model with the best performance in the last λ observations;
- `ADE`: A dynamic combination approach based on a meta-learning strategy called arbitrating. The idea is to build a meta model (a Random Forest) for each (base) model in the ensemble. Each meta model is designed to predict the error of the corresponding base model. Then, the models in the ensemble are weighted according to the error forecasts. ;
- `EWA`: A dynamic combination rule based on an exponentially weighted average. This method follows the popular weighted majority algorithm [?];
- `FS`: The fixed share dynamic combination approach. This method is designed to handle non-stationary time series;
- `MLpol`: A dynamic combination method based on a polynomially weighted average;
- `Best`: A baseline which selects the individual model in the ensemble with the best performance in the training data to predict all the test instances;
- `LossTrain`: Another baseline which weights the available models based on the error on the training set. The weights are static and fixed for all testing observations;

COMBINATION ALONG THE HORIZON

Research gaps

Dynamic ensemble methods typically assume immediate feedback. They are mostly designed for one-step ahead forecasting. Thus, it is not clear how the ensemble weights should be computed along the forecasting horizon.

Approaches

We study the following approaches to estimate the weights at each time-step:

- `Complete Horizon (CH)`: The weights of individual models are estimated using their average performance over the complete forecasting horizon;
- `Individual Horizon (IH)`: The ensemble estimates different weights for each horizon;
- `First Horizon Forward (FHF)`: The weights computed for the first horizon are propagated over the rest of the horizon;
- `Last Horizon Backward (LHB)`: For completeness, we include the inverse approach to `FHF`. According to `LHB`, the weights computed for the last horizon are propagated backward to all horizons before it.

PRR Plano de Recuperação e Resiliência | REPÚBLICA PORTUGUESA | Financiado pela União Europeia NextGenerationEU | Center for Responsible AI

Acknowledgements: This work is a result of Agenda "Center for Responsible AI", nr. C645008882-00000055, investment project nr. 62, financed by the Recovery and Resilience Plan (PRR) and by European Union - NextGeneration EU. Funded by the European Union – NextGenerationEU. Funded by the European Union under the Horizon Europe Framework Programme Grant Agreement Nº: 101095387. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the European Health and Digital Executive Agency can be held responsible for them; UID/00027 of the LIACC - Artificial Intelligence and Computer Science Laboratory - funded by Fundação para a Ciência e a Tecnologia, I.P./ MCTES through the national funds. The work of L. Torgo was undertaken, in part, thanks to funding from the Canada Research Chairs program.

RESULTS

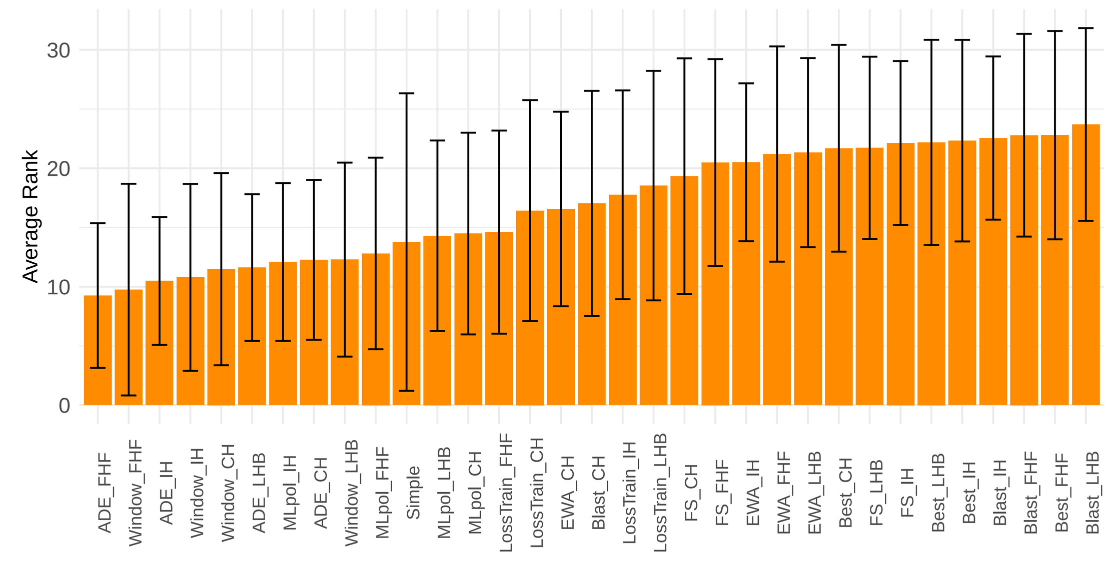


Figure 1. Average rank, and respective standard deviation, of each method across all time series.

RQ1

- The best approaches are variants of `ADE`, `Window`, and `MLpol`.
- Variants of `Blast`, `Best`, `FS`, and `EWA` occupy the bottom positions.
- All methods show a considerable standard deviation of rank.

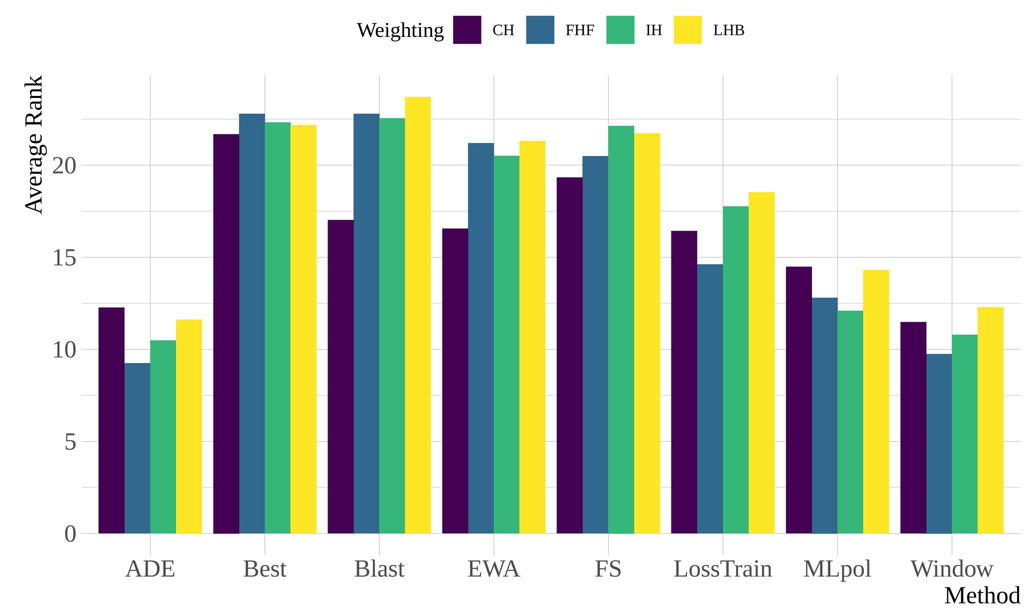


Figure 2. Average rank of each method, aggregated by weighting strategy.

RQ2

- Varying the weighting strategy (`FHF`, `LHB`, `IH`, `CH`) does not have an impact in average rank.
- `ADE` and `Window`, show similar behavior in terms of relative scores for the different weighting strategies. Their best score is achieved with `FHF`, followed by `IH`.

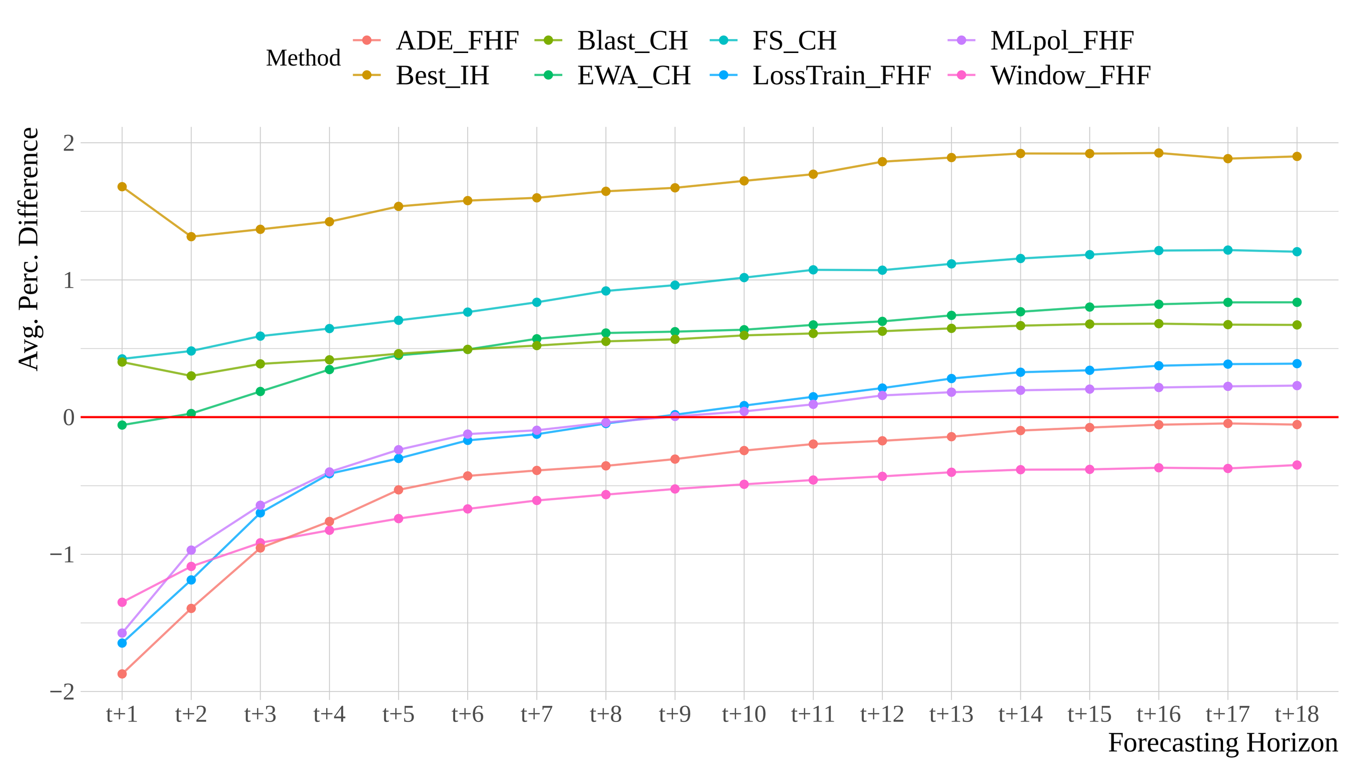


Figure 3. Median percentage difference of each method relative to `Simple` in each forecasting horizon.

RQ3

- Clear trend indicating that dynamic combination methods decrease their performance relative to `Simple` as the forecasting horizon increases.
- For *t+1* (one-step ahead forecasting), 5 out of 8 combination methods outperform `Simple`. But, for long-term forecasting (*t+18*), only two methods achieve better performance.

Faculdade de Engenharia da Universidade do Porto

vitorc.research@gmail.com