# Ensembles and combinations: using multiple models to improve forecasts

## 1 Research beginnings

- Goal
- Application
- Definition

### 1.1 Combinations

Barnard (1963) first proposed an empirical justification of forecast combination by making a simple average of two forecasts in forecasting airline passenger numbers. It was observed that a simple average of two forecasts outperformed each of them. This was the first work to provide an analysis for the problem of point forecast combination.

Bates and Granger (1969) further explored more possibilities for forecast combination by extending the simple average to a weighted combination, which examined various weight determination methods yield from past errors (MSE). It is concluded that combining forecasts with some independent information improves forecast accuracy.

### 1.2 Ensembles

It is difficult to trace the beginning of the history of ensemble forecasting. However, it is clear that ensemble techniques have become a hot topic in various fields, especially weather forecasting, since the 1990s. Lewis (2005) provided a genealogy to depict the scientific roots of ensemble forecasting from several fundamental lines of research.

Ensemble forecasting, which aims at quantifying the forecast uncertainty and generating probability forecasts, achieved its practical implementation in weather forecasting as a response to the limitations of deterministic forecasting without taking account of the initial condition perturbations. Leutbecher and Palmer (2008) discussed the sources of uncertainty in weather forecasting:

uncertainties in the observations, used to initialise the predictions, and in the forecast models themselves.

## 2 Whether a training dataset is required

- Mainly depends on whether the forecasting process includes meta-learner training.

### 2.1 Combinations

Barnard (1963) provided the arithmetic mean of the two individual forecasts in the application of forecasting airline passenger data (not required).

Bates and Granger (1969) extended the simple average to a weighted combination which aims to give higher weight to better forecaster. Several ways of determining combining weights based on past errors are introduced for combining point forecasts (not required).

Newbold and Granger (1974) examined various possibilities of combining three univariate time series forecasting methods (Box-Jenkins, Holt-Winters and stepwise autoregression) over a large collection of economic time series. Their results suggested that a combined forecast generally performs well if the weights calculated based on the relative precision of individual forecasts ignore the effects of correlations among errors (not required). Winkler and Makridakis (1983) further confirmed their conclusions (not required).

Granger and Ramanathan (1984) considered a linear combination of forecasts by performing a regression model having the actual value as the response variable and the individual forecasts as the explanatory variables. They determined combining weights through a constrained ordinary least squares (OLS) and demonstrated its superiority in fitting and out-of-sample forecasting (not required).

Diebold and Pauly (1990) developed a weight estimation technique by using Bayesian shrinkage techniques to incorporate prior information into unrestricted regression-based forecast combination (not required).

Hibon and Evgeniou (2005) proposed a criterion to chose individual methods and combinations using the smallest average sMAPE over validation periods and successfully applied this criterion on the data of the M3-competition (not required).

Collopy and Armstrong (1992) provided one of the pioneering studies in combining forecasts using characteristics. They proposed a rule-based forecasting procedure to develop 99 rules for forecast combination based on 18 features obtained judgementally (not required).

Vokurka et al. (1996) proposed a rule-based expert forecasting system to automatically identify time series features and build a weighted combination of three individual methods (not required).

Prudêncio and Ludermir (2004) used machine learning techniques to define the best linear combination of methods. Specifically, they use the Multi-Layer Perceptron (MLP) network as the learner to associate the time series features and the combining weights for two individual methods (required for building MLP learner).

Lemke and Gabrys (2010) created a feature pool describing both the time series and the pool of individual forecasting methods, and then applied different meta-learning approaches to improve forecasting performance using a ranking-based combination (required for building different meta-learning models).

Andrawis et al. (2011) carefully selected the individual forecasting models and combined them by simple average (not required).

Montero-Manso et al. (2020) employed 42 time series features to estimate the optimal combination weights of nine forecasting methods based on extreme gradient boosting (required for training XGBoost using features and errors).

Kang et al. (2020) proposed GeneRAting TIme Series with diverse and controllable characteristics (GRATIS) with the use of mixture autoregressive (MAR) models. GRATIS can be used as a tool for simulating a diverse set of time series data which serves as a training dataset for forecast combination (required).

Li et al. (2020) transformed time series into images and then extracted features from these images for forecast combination (required for training an XGBoost model to produce weights for individual methods).

### 2.2 Ensembles

Cordeiro and Neves (2009) developed the Boost.EXPOS approach which forecasts time series by combining the use of exponential smoothing methods with the bootstrap resampling technique for time series data (not required).

Bergmeir et al. (2016) presented a bagging exponential smoothing method which ensembles exponential smoothing models estimated on the bootstrapped series (not required).

Petropoulos et al. (2018) implemented six forecasting strategies, four of which focus on isolating each one of the three sources of uncertainty (data, parameter, and model uncertainty) to separately explore the benefits of bagging for time series forecasting for each one of them (not required).

Taieb and Hyndman (2014*a*) proposed a forecasting strategy which adjusts autoregressive forecasts with a direct strategy using a boosting procedure at each horizon (not required).

Barrow and Crone (2016) transformed a time series into a dataset. The dataset consists of several pairs of observations comprised of a lagged autoregressive (AR) realisation vector and a further realisation. A generic AdaBoost algorithm is applied in the dataset to obtain combined forecasts (not required).

Ribeiro and dos Santos Coelho (2020) compared several models to explore the predictive capability of regression ensembles. They considered stacked generalization by training a meta-model which determines combining weights by using individual forecasts as an input set (a training dataset is created by the approach of leave-one-out cross-validation for training meta-model).

## 3 Data resampling for modeling

- Combination forecasting may require a training data set generated using a holdout strategy, simulation techniques or other data generating methods such as GRATIS.
- Ensemble forecasting mainly extracts information from the original series. Data resampling is required for modeling in ensemble forecasting, such as bootstrap in bagging, re-weighting in boosting and data rolling for training dataset generation in stacking.

## 4 Forecast method pool

- Homogeneous methods
- Heterogeneous methods
- Selection of individual forecasting models

## 5 Forecasting structure

- Parallel forecasting
- Sequential forecasting

## 6 Main types of ensembles

- Bagging
- Boosting
- Stacking

## 7 Simple aggregation methods

- Such as mean, trimmed mean, median, trimmed median, ...

## 8 Weight determination

- Data base
- Information used, such as individual forecasts, forecasts diversity, features, errors ...
- How to estimate combining weight, such as information criterion-based, regression-based estimation ...
- Weather to consider correlations in estimating combining weights
- Fixed weights VS time-varying weights

## 9 Uncertainty tackling

- Three sources of uncertainty: model, parameter, and data uncertainty

## 10 Contribution to forecast improvement

- How much do combinations and ensembles contribute to the forecast improvement?

## 11 Point forecasting

## 12 Probabilistic forecasting

- Residual simulation
- Combining interval forecasts
- Combining quantile forecasts
- Combining probability forecasts

## 13 Forecast calibration

- Perfectly calibrated, poorly calibrated (overconfident, underconfident)
- Overfit and overconfident forecasts

## References

Alonzo, B., Tankov, P., Drobinski, P. and Plougonven, R. (2020), ‘Probabilistic wind forecasting up to three months ahead using ensemble predictions for geopotential height’, *International journal of forecasting* **36**(2), 515–530.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207019302018*

Andrawis, R. R., Atiya, A. F. and El-Shishiny, H. (2011), 'Forecast combinations of computational intelligence and linear models for the NN5 time series forecasting competition', *International journal of forecasting* **27**(3), 672–688.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207010001445*

Assaad, M., Boné, R. and Cardot, H. (2008), 'A new boosting algorithm for improved time-series forecasting with recurrent neural networks', *An international journal on information fusion* **9**(1), 41–55.
**URL:** *https://www.sciencedirect.com/science/article/pii/S1566253506000820*

Atiya, A. F. (2019), 'Why does forecast combination work so well?', *International journal of forecasting* .
**URL:** *http://www.sciencedirect.com/science/article/pii/S0169207019300779*

Barnard, G. A. (1963), 'New methods of quality control', *Journal of the Royal Statistical Society. Series A* **126**(2), 255.
**URL:** *https://www.jstor.org/stable/10.2307/2982365?origin=crossref*

Barrow, D. K. and Crone, S. F. (2013), Crogging (cross-validation aggregation) for forecasting — a novel algorithm of neural network ensembles on time series subsamples, *in* 'The 2013 International Joint Conference on Neural Networks (IJCNN)', ieeexplore.ieee.org, pp. 1–8.
**URL:** *http://dx.doi.org/10.1109/IJCNN.2013.6706740*

Barrow, D. K. and Crone, S. F. (2016), 'A comparison of AdaBoost algorithms for time series forecast combination', *International journal of forecasting* **32**(4), 1103–1119.
**URL:** *http://dx.doi.org/10.1016/j.ijforecast.2016.01.006*

Bates, J. M. and Granger, C. W. J. (1969), 'The combination of forecasts', *The Journal of the Operational Research Society* **20**(4), 451–468.
**URL:** *https://www.tandfonline.com/doi/full/10.1057/jors.1969.103*

Ben Taieb, S., Huser, R., Hyndman, R. J. and Genton, M. G. (2016), 'Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression', *IEEE transactions on smart grid* **7**(5), 2448–2455.
**URL:** *http://dx.doi.org/10.1109/TSG.2016.2527820*

Ben Taieb, S. and Hyndman, R. J. (2014), 'A gradient boosting approach to the kaggle load forecasting competition', *International journal of forecasting* .

Bergmeir, C., Hyndman, R. J. and Benítez, J. M. (2016), 'Bagging exponential smoothing methods using STL decomposition and Box–Cox transformation', *International journal of forecasting* **32**(2), 303–312.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207015001120*

Berk, R. A. (2010), *Statistical learning from a regression perspective*, Springer Series in Statistics, Springer, New York, NY.

Brcker, J. and Smith, L. A. (2008), 'From ensemble forecasts to predictive distribution functions', *Tellus A Dynamic Meteorology and Oceanography* **60**(4), 663–678.
**URL:** *http://tellusa.net/index.php/tellusa/article/view/15387*

Bunn, D. W. (1975), 'A bayesian approach to the linear combination of forecasts', *Operational research quarterly* **26**(2), 325.
**URL:** *https://www.jstor.org/stable/3008467?origin=crossref*

Bunn, D. W. (1981), 'Two methodologies for the linear combination of forecasts', *The Journal of the Operational Research Society* **32**(3), 213–222.
**URL:** *https://doi.org/10.1057/jors.1981.44*

Cerqueira, V., Torgo, L., Pinto, F. and Soares, C. (2017), Arbitrated ensemble for time series forecasting, *in* 'Machine Learning and Knowledge Discovery in Databases', Vol. 10535 LNAI, Springer International Publishing, pp. 478–494.

Chan, Y. L., Stock, J. H. and Watson, M. W. (1999), 'A dynamic factor model framework for forecast combination', *Spanish Economic Review* **1**(2), 91–121.
**URL:** *https://doi.org/10.1007/s101080050005*

Che, J. (2015), 'Optimal sub-models selection algorithm for combination forecasting model', *Neurocomputing* **151**(P1), 364–375.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0925231214012119*

Chen, L.-C. G. and van den Dool, H. (2017), 'Combination of multimodel probabilistic forecasts using an optimal weighting system', *Weather and Forecasting* **32**(5), 1967–1987.

Claeskens, G., Magnus, J. R., Vasnev, A. L. and Wang, W. (2016), 'The forecast combination puzzle: A simple theoretical explanation', *International journal of forecasting* **32**(3), 754–762.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207016000327*

Clark, T. E. and McCracken, M. W. (2008), Improving forecast accuracy by combining recursive and rolling forecasts.

Clemen, R. T. (1989), 'Combining forecasts: A review and annotated bibliography', *International journal of forecasting* **5**(4), 559–583.
**URL:** *https://www.sciencedirect.com/science/article/pii/0169207089900125*

Clements, M. P. and Harvey, D. I. (2011), 'Combining probability forecasts', *International journal of forecasting* **27**(2), 208–223.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207010000075*

Collopy, F. and Armstrong, J. S. (1992), 'Rule-based forecasting: Development and validation of an expert systems approach to combining time series extrapolations', *Management science* **38**(10), 1394–1414.
**URL:** *http://pubsonline.informs.org/doi/abs/10.1287/mnsc.38.10.1394*

Cordeiro, C. and Neves, M. (2009), 'Forecasting time series with BOOT.EXPOS procedure', *REVSTAT-Statistical Journal* **7**(2), 135–149.

Dantas, T. M. and Cyrino Oliveira, F. L. (2018), 'Improving time series forecasting: An approach combining bootstrap aggregation, clusters and exponential smoothing', *International journal of forecasting* **34**(4), 748–761.
**URL:** *https://doi.org/10.1016/j.ijforecast.2018.05.006*

Das, M. and Ghosh, S. K. (2017), 'A deep-learning-based forecasting ensemble to predict missing data for remote sensing analysis', *IEEE journal of selected topics in applied earth observations and remote sensing* **10**(12), 5228–5236.
**URL:** *http://ieeexplore.ieee.org/document/8170480/*

De Gooijer, J. G. and Hyndman, R. J. (2006), '25 years of time series forecasting', *International journal of forecasting* **22**(3), 443–473.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207006000021*

de Oliveira, E. M. and Cyrino Oliveira, F. L. (2018), 'Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods', *Energy* **144**, 776–788.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0360544217320820*

de Souza, L. V., Pozo, A. T. R., da Rosa, J. M. C. and Neto, A. C. (2007), The boosting technique using correlation coefficient to improve time series forecasting accuracy, *in* '2007 IEEE Congress on Evolutionary Computation', pp. 1288–1295.
**URL:** *http://dx.doi.org/10.1109/CEC.2007.4424619*

Dickinson, J. P. (1973), 'Some statistical results in the combination of forecasts', *Operational research quarterly* **24**(2), 253.
**URL:** *https://www.jstor.org/stable/3007853?origin=crossref*

Dickinson, J. P. (1975), 'Some comments on the combination of forecasts', *Operational research quarterly* **26**(1), 205.
**URL:** *https://www.jstor.org/stable/3008402?origin=crossref*

Diebold, F. X. and Pauly, P. (1990), 'The use of prior information in forecast combination', *International journal of forecasting* **6**(4), 503–508.
**URL:** *https://www.sciencedirect.com/science/article/pii/016920709090028A*

Diebold, F. X. and Shin, M. (2017), Beating the simple average: Egalitarian LASSO for combining economic forecasts.
**URL:** *https://papers.ssrn.com/abstract=3032492*

Donate, J. P., Cortez, P., Sánchez, G. G. and de Miguel, A. S. (2013), 'Time series forecasting using a weighted cross-validation evolutionary artificial neural network ensemble', *Neurocomputing* **109**, 27–32.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0925231213000209*

Elliott, G. and Timmermann, A. (2004), 'Optimal forecast combinations under general loss functions and forecast error distributions', *Journal of econometrics* **122**(1), 47–79.
**URL:** *https://linkinghub.elsevier.com/retrieve/pii/S0304407603002690*

Franses, P. H. and Legerstee, R. (2011), 'Combining SKU-level sales forecasts from models and

experts', *Expert systems with applications* **38**(3), 2365–2370.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0957417410008079*

Gaba, A., Tsetlin, I. and Winkler, R. L. (2017), 'Combining interval forecasts', *Decision Analysis* **14**(1), 1–20.
**URL:** *https://doi.org/10.1287/deca.2016.0340*

Gaglianone, W. P. and Renato, L. (2012), Constructing optimal density forecasts from point forecast combinations.

Genest, C. and Zidek, J. V. (1986), 'Combining probability distributions: A critique and an annotated bibliography', *Schweizerische Monatsschrift fur Zahnheilkunde = Revue mensuelle suisse d'odonto-stomatologie / SSO* **1**(1), 114–135.
**URL:** *https://projecteuclid.org/euclid.ss/1177013825*

Genre, V., Kenny, G., Meyler, A. and Timmermann, A. (2013), 'Combining expert forecasts: Can anything beat the simple average?', *International journal of forecasting* **29**(1), 108–121.
**URL:** *https://www.sciencedirect.com/science/article/pii/S016920701200088X*

Gneiting, T. and Raftery, A. E. (2005), 'Atmospheric science. weather forecasting with ensemble methods', *Science* **310**(5746), 248–249.
**URL:** *http://dx.doi.org/10.1126/science.1115255*

Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T. (2005), 'Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation', *Monthly Weather Review* **133**(5), 1098–1118.
**URL:** *https://journals.ametsoc.org/mwr/article/133/5/1098/67504*

Gneiting, T. and Ranjan, R. (2013), 'Combining predictive distributions', *European journal of sport science: EJSS: official journal of the European College of Sport Science* **7**(none), 1747–1782.
**URL:** *http://projecteuclid.org/euclid.ejs/1372861687*

Godahewa, R., Bandara, K., Webb, G. I., Smyl, S. and Bergmeir, C. (2020), 'Ensembles of localised models for time series forecasting'.
**URL:** *http://arxiv.org/abs/2012.15059*

Gonçalves, S. and Politis, D. (2011), 'Discussion: Bootstrap methods for dependent data: A review', *Journal of the Korean Statistical Society* **40**(4), 383–386.
**URL:** *https://www.sciencedirect.com/science/article/pii/S1226319211000664*

Granger, C. W. J. (1989*a*), 'Invited review combining forecasts—twenty years later', *Journal of forecasting* **8**(3), 167–173.
**URL:** *http://doi.wiley.com/10.1002/for.3980080303*

Granger, C. W. J. (1989*b*), 'Invited review combining forecasts—twenty years later', *Journal of forecasting* **8**(3), 167–173.
**URL:** *http://doi.wiley.com/10.1002/for.3980080303*

Granger, C. W. J. and Ramanathan, R. (1984), 'Improved methods of combining forecasts', *Journal of forecasting* **3**(2), 197–204.
**URL:** *http://doi.wiley.com/10.1002/for.3980030207*

Grushka-Cockayne, Y. and Jose, V. R. R. (2020), 'Combining prediction intervals in the M4 competition'.
**URL:** *http://dx.doi.org/10.1016/j.ijforecast.2019.04.015*

Grushka-Cockayne, Y., Jose, V. R. R. and Lichtendahl, K. C. (2017), 'Ensembles of overfit and overconfident forecasts', *Management science* **63**(4), 1110–1130.
**URL:** *https://doi.org/10.1287/mnsc.2015.2389*

Hendry, D. F. and Hubrich, K. (2010), Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate.

Hibon, M. and Evgeniou, T. (2005), 'To combine or not to combine: selecting among forecasts and their combinations', *International journal of forecasting* **21**(1), 15–24.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207004000494*

Hoeting, J. A., Madigan, D., Raftery, A. E. and Volinsky, C. T. (1999), 'Bayesian model averaging: A tutorial', *Statistical science: a review journal of the Institute of Mathematical Statistics* **14**(4), 382–401.
**URL:** *http://www.jstor.org/stable/2676803*

Hong, T. and Fan, S. (2016), ‘Probabilistic electric load forecasting: A tutorial review’, *International journal of forecasting* **32**(3), 914–938.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207015001508*

Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A. and Hyndman, R. J. (2016), ‘Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond’.
**URL:** *http://dx.doi.org/10.1016/j.ijforecast.2016.02.001*

Huang, J. and Perry, M. (2016), ‘A semi-empirical approach using gradient boosting and k-nearest neighbors regression for GEFCom2014 probabilistic solar power forecasting’, *International journal of forecasting* **32**(3), 1081–1086.
**URL:** *http://dx.doi.org/10.1016/j.ijforecast.2015.11.002*

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G. and Shang, H. L. (2011), ‘Optimal combination forecasts for hierarchical time series’, *Computational statistics & data analysis* **55**(9), 2579–2589.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0167947311000971*

Ingel, A., Shahroudi, N., Kängsepp, M., Tättar, A., Komisarenko, V. and Kull, M. (2020), ‘Correlated daily time series and forecasting in the M4 competition’, *International journal of forecasting* **36**(1), 121–128.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207019301360*

Inoue, A. and Kilian, L. (2008), ‘How useful is bagging in forecasting economic time series? a case study of U.S. consumer price inflation’, *Journal of the American Statistical Association* **103**(482), 511–522.
**URL:** *https://doi.org/10.1198/016214507000000473*

J., W. (1984), ‘Improved methods of combining forecasts’, *Journal of forecasting* **3**, 197–204.

Jaganathan, S. and Prakash, P. K. S. (2020), ‘A combination-based forecasting method for the m4-competition’, *International journal of forecasting* **36**(1), 98–104.
**URL:** *http://www.sciencedirect.com/science/article/pii/S0169207019301542*

Jiang, Y., Chen, X., Yu, K. and Liao, Y. (2017), ‘Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm’, *Journal of Modern Power Systems and Clean Energy* **5**(1), 126–133.
**URL:** *https://doi.org/10.1007/s40565-015-0171-6*

Kang, Y., Hyndman, R. J. and Li, F. (2020), 'GRATIS: GeneRAting TIme Series with diverse and controllable characteristics', *Statistical analysis and data mining* **13**(4), 354–376.
**URL:** *https://onlinelibrary.wiley.com/doi/abs/10.1002/sam.11461*

Kauppi, H. and Virtanen, T. (2021), 'Boosting nonlinear predictability of macroeconomic time series', *International journal of forecasting* **37**(1), 151–170.
**URL:** *https://doi.org/10.1016/j.ijforecast.2020.03.008*

Kim, D. and Hur, J. (2018), 'Short-term probabilistic forecasting of wind energy resources using the enhanced ensemble method', *Energy* **157**, 211–226.
**URL:** *https://doi.org/10.1016/j.energy.2018.05.157*

Kolassa, S. (2011), 'Combining exponential smoothing forecasts using akaike weights', *International journal of forecasting* **27**(2), 238–251.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207010001032*

Kootale, K. C. and Road, P. (2007), Combining forecasts in hierarchically organized data.

Kourentzes, N., Barrow, D. and Petropoulos, F. (2019), 'Another look at forecast selection and combination: Evidence from forecast pooling', *International Journal of Production Economics* **209**(February 2018), 226–235.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0925527318302196*

Kourentzes, N., Svetunkov, I. and Kolassa, S. (n.d.), 'Beyond summary performance metrics for forecast selection and combination'.

Landgraf, A. J. (2019), 'An ensemble approach to GEFCom2017 probabilistic load forecasting', *International journal of forecasting* **35**(4), 1432–1438.
**URL:** *https://doi.org/10.1016/j.ijforecast.2019.02.003*

Lemke, C. and Gabrys, B. (2010), 'Meta-learning for time series forecasting and forecast combination', *Neurocomputing* **73**(10), 2006–2016.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0925231210001074*

Leutbecher, M. and Palmer, T. N. (2008), 'Ensemble forecasting', *Journal of computational physics* **227**(7), 3515–3539.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0021999107000812*

Lewis, J. M. (2005), 'Roots of ensemble forecasting', *Monthly Weather Review* **133**(7), 1865–1885.
**URL:** *http://journals.ametsoc.org/doi/10.1175/MWR2949.1*

Li, X., Kang, Y. and Li, F. (2020), 'Forecasting with time series imaging', *Expert systems with applications* **160**(113680), 113680.
**URL:** *https://linkinghub.elsevier.com/retrieve/pii/S0957417420305042*

Lichtendahl, K. C., Grushka-Cockayne, Y. and Winkler, R. L. (2013), 'Is it better to average probabilities or quantiles?', *Management science* **59**(7), 1594–1611.
**URL:** *https://doi.org/10.1287/mnsc.1120.1667*

Lichtendahl, K. C. and Winkler, R. L. (2020), 'Why do some combinations perform better than others?', *International journal of forecasting* **36**(1), 142–149.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207019301475*

Lin, Y., Yang, M., Wan, C., Wang, J. and Song, Y. (2019), 'A Multi-Model combination approach for probabilistic wind power forecasting', *IEEE Transactions on Sustainable Energy* **10**(1), 226–237.
**URL:** *http://dx.doi.org/10.1109/TSTE.2018.2831238*

Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2020), 'The M4 competition: 100,000 time series and 61 forecasting methods', *International journal of forecasting* **36**(1), 54–74.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207019301128*

Maqsood, I., Khan, M. R. and Abraham, A. (2004), 'An ensemble of neural networks for weather forecasting', *Neural computing & applications* **13**(2), 112–122.
**URL:** *https://doi.org/10.1007/s00521-004-0413-4*

Massaoudi, M., Refaat, S. S., Chihi, I., Trabelsi, M., Oueslati, F. S. and Abu-Rub, H. (2021), 'A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for Short-Term load forecasting', *Energy* **214**, 118874.
**URL:** *https://doi.org/10.1016/j.energy.2020.118874*

Mayrink, V. and Hippert, H. S. (2016), A hybrid method using exponential smoothing and gradient boosting for electrical short-term load forecasting, *in* '2016 IEEE Latin American Conference on Computational Intelligence (LA-CCI)', IEEE, pp. 1–6.
**URL:** *http://dx.doi.org/10.1109/LA-CCI.2016.7885697*

Montero-Manso, P., Athanasopoulos, G., Hyndman, R. J. and Talagala, T. S. (2020), 'FFORMA: Feature-based forecast model averaging', *International journal of forecasting* **36**(1), 86–92.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207019300895*

Moon, J., Jung, S., Rew, J., Rho, S. and Hwang, E. (2020), 'Combination of short-term load forecasting models based on a stacking ensemble approach', *Energy and Buildings* **216**, 109921.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0378778819327938*

Moretti, F., Pizzuti, S., Panzieri, S. and Annunziato, M. (2015), 'Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling', *Neurocomputing* **167**, 3–7.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0925231215005603*

Nagy, G. I., Barta, G., Kazi, S., Borbély, G. and Simon, G. (2016), 'GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach', *International journal of forecasting* **32**(3), 1087–1093.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207015001521*

Newbold, P. and Granger, C. W. J. (1974), 'Experience with forecasting univariate time series and the combination of forecasts', *Journal of the Royal Statistical Society. Series A* **137**(2), 131.
**URL:** *https://www.jstor.org/stable/2344546?origin=crossref*

Nowotarski, J. and Weron, R. (2016), To combine or not to combine? recent trends in electricity price forecasting, Technical Report HSC/16/01.
**URL:** *https://ideas.repec.org/p/wuu/wpaper/hsc1601.html*

Oliveira, M. and Torgo, L. (2015), Ensembles for time series forecasting, *in* D. Phung and H. Li, eds, 'Proceedings of the Sixth Asian Conference on Machine Learning', Vol. 39 of *Proceedings of Machine Learning Research*, PMLR, Nha Trang City, Vietnam, pp. 360–370.
**URL:** *http://proceedings.mlr.press/v39/oliveira14.html*

Park, S. and Budescu, D. V. (2015), Aggregating multiple probability intervals to improve calibration, Technical report.
**URL:** *https://www.sas.upenn.edu/ baron/journal/14/141223/jdm141223.pdf*

Pawlikowski, M. and Chorowska, A. (2020), 'Weighted ensemble of statistical models', *International*

*journal of forecasting* **36**(1), 93–97.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207019301190*

Petropoulos, A., Chatzis, S. P., Siakoulis, V. and Vlachogiannakis, N. (2017), 'A stacked generalization system for automated FOREX portfolio trading', *Expert systems with applications* **90**, 290–302.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0957417417305493*

Petropoulos, F., Hyndman, R. J. and Bergmeir, C. (2018), 'Exploring the sources of uncertainty: Why does bagging for time series forecasting work?', *European journal of operational research* **268**(2), 545–554.
**URL:** *https://linkinghub.elsevier.com/retrieve/pii/S037722171830081X*

Petropoulos, F. and Svetunkov, I. (2020), 'A simple combination of univariate models', *International journal of forecasting* **36**(1), 110–115.
**URL:** *https://linkinghub.elsevier.com/retrieve/pii/S0169207019300585*

Pinson, P. and Madsen, H. (2009), 'Ensemble-based probabilistic forecasting at horns rev', *Wind Energy* **12**(2), 137–155.
**URL:** *http://doi.wiley.com/10.1002/we.309*

Prudêncio, R. and Ludermir, T. (2004), 'Using machine learning techniques to combine forecasting methods', *Planning perspectives: PP* **1122**, 1127.

Prudêncio, R. and Ludermir, T. (2005), Using machine learning techniques to combine forecasting methods, *in* 'AI 2004: Advances in Artificial Intelligence', Springer Berlin Heidelberg, pp. 1122–1127.

Qiu, X., Zhang, L., Ren, Y., Suganthan, P. and Amaratunga, G. (2014), Ensemble deep learning for regression and time series forecasting, *in* '2014 IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL)', IEEE.
**URL:** *http://ieeexplore.ieee.org/document/7015739/*

Radchenko, P., Vasnev, A. L. and Wang, W. (2020), Too similar to combine? on negative weights in forecast combination.
**URL:** *https://papers.ssrn.com/abstract=3647603*

Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005), 'Using bayesian model averaging to calibrate forecast ensembles', *Monthly Weather Review* **133**(5), 1155–1174.
**URL:** *http://journals.ametsoc.org/doi/10.1175/MWR2906.1*

Ranjan, R. (2009), 'Combining and evaluating probabilistic forecasts'.

Ranjan, R. and Gneiting, T. (2010), 'Combining probability forecasts', *Journal of the Royal Statistical Society. Series B, Statistical methodology* **72**(1), 71–91.
**URL:** *http://doi.wiley.com/10.1111/j.1467-9868.2009.00726.x*

Ray, E. L., Wattanachit, N., Niemi, J., Kanji, A. H., House, K., Cramer, E. Y., Bracher, J., Zheng, A., Yamana, T. K., Xiong, X., Woody, S., Wang, Y., Wang, L., Walraven, R. L., Tomar, V., Sherratt, K., Sheldon, D., Reiner, R. C., Prakash, B. A., Osthus, D., Li, M. L., Lee, E. C., Koyluoglu, U., Keskinocak, P., Gu, Y., Gu, Q., George, G. E., España, G., Corsetti, S., Chhatwal, J., Cavany, S., Biegel, H., Ben-Nun, M., Walker, J., Slayton, R., Lopez, V., Biggerstaff, M., Johansson, M. A., Reich, N. G. and COVID-19 Forecast Hub Consortium (2020), Ensemble forecasts of coronavirus disease 2019 (COVID-19) in the U.S.
**URL:** *https://www.medrxiv.org/content/10.1101/2020.08.19.20177493v1.abstract*

Rendon-Sanchez, J. F. and de Menezes, L. M. (2019), 'Structural combination of seasonal exponential smoothing forecasts applied to load forecasting', *European journal of operational research* **275**(3), 916–924.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0377221718310518*

Ribeiro, M. H. D. M. and dos Santos Coelho, L. (2020), 'Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series', *Applied soft computing* **86**, 105837.
**URL:** *https://www.sciencedirect.com/science/article/pii/S1568494619306180*

Robinzonov, N., Tutz, G. and Hothorn, T. (2012), 'Boosting techniques for nonlinear time series models', *AStA. Advances in Statistical Analysis. A Journal of the German Statistical Society* **96**(1), 99–122.
**URL:** *https://doi.org/10.1007/s10182-011-0163-4*

Saha, M., Santara, A., Mitra, P., Chakraborty, A. and Nanjundiah, R. S. (2021), 'Prediction of the indian summer monsoon using a stacked autoencoder and ensemble regression model',

*International journal of forecasting* **37**(1), 58–71.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207020300479*

Shaub, D. (2019), 'Fast and accurate yearly time series forecasting with forecast combinations', *International journal of forecasting* .
**URL:** *http://www.sciencedirect.com/science/article/pii/S0169207019301566*

Shen, W., Babushkin, V., Aung, Z. and Woon, W. L. (2013), An ensemble model for day-ahead electricity demand time series forecasting, *in* 'Proceedings of the fourth international conference on Future energy systems', e-Energy '13, Association for Computing Machinery, New York, NY, USA, pp. 51–62.
**URL:** *https://doi.org/10.1145/2487166.2487173*

Sivillo, J. K., Ahlquist, J. E. and Toth, Z. (1997), 'An ensemble forecasting primer', *Weather and Forecasting* **12**(4), 809–818.

Smith, J. and Wallis, K. F. (2009), 'A simple explanation of the forecast combination puzzle', *Oxford bulletin of economics and statistics* **71**(3), 331–355.
**URL:** *http://doi.wiley.com/10.1111/j.1468-0084.2008.00541.x*

Smyl, S. (2020), 'A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting', *International journal of forecasting* **36**(1), 75–85.
**URL:** *https://doi.org/10.1016/j.ijforecast.2019.03.017*

Song, H., Gao, B. Z. and Lin, V. S. (2013), 'Combining statistical and judgmental forecasts via a web-based tourism demand forecasting system', *International journal of forecasting* **29**(2), 295–310.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207012000064*

Svensson, M. (n.d.), 'Rank-based selection strategies for forecast combinations: An evaluation study', http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8974226&fileOId=8974925. Accessed: 2021-3-7.

Taieb, S. B. and Hyndman, R. (2014*a*), Boosting multi-step autoregressive forecasts, *in* E. P. Xing and T. Jebara, eds, 'Proceedings of the 31st International Conference on Machine Learning', Vol. 32 of *Proceedings of Machine Learning Research*, PMLR, Bejing, China, pp. 109–117.
**URL:** *http://proceedings.mlr.press/v32/taieb14.html*

Taieb, S. B. and Hyndman, R. (2014*b*), Boosting multi-step autoregressive forecasts, *in* E. P. Xing and T. Jebara, eds, 'Proceedings of the 31st International Conference on Machine Learning', Vol. 32 of *Proceedings of Machine Learning Research*, PMLR, Bejing, China, pp. 109–117.
**URL:** *http://proceedings.mlr.press/v32/taieb14.html*

Talagala, T. S., Li, F. and Kang, Y. (2019), 'FFORMPP: Feature-based forecast model performance prediction'.
**URL:** *http://arxiv.org/abs/1908.11500*

Tashman, L. J. (2000), 'Out-of-sample tests of forecasting accuracy: An analysis and review', *International journal of forecasting* **16**(4), 437–450.
**URL:** *http://dx.doi.org/10.1016/S0169-2070(00)00065-0*

Taylor, J. W. and Buizza, R. (2002), 'Weather ensemble predictions', *IEEE Transactions on Power Systems* **17**(3), 626–632.

Taylor, J. W., McSharry, P. E. and Buizza, R. (2009), 'Wind power density forecasting using ensemble predictions and time series models', *IEEE Transactions on Energy Conversion* **24**(3), 775–782.
**URL:** *http://dx.doi.org/10.1109/TEC.2009.2025431*

Thesis, M. and Pritzsche, U. (2015), 'Benchmarking of classical and Machine-Learning algorithms (with special emphasis on bagging and boosting approaches) for time series forecasting'.

Thorey, J., Mallet, V. and Baudin, P. (2017), 'Online learning with the continuous ranked probability score for ensemble forecasting: Ensemble online learning', *Quarterly Journal of the Royal Meteorological Society* **143**(702), 521–529.
**URL:** *https://onlinelibrary.wiley.com/doi/10.1002/qj.2940*

Timmermann, A. (2006), Chapter 4 forecast combinations, *in* G. Elliott, C. W. J. Granger and A. Timmermann, eds, 'Handbook of Economic Forecasting', Vol. 1, Elsevier, pp. 135–196.
**URL:** *https://www.sciencedirect.com/science/article/pii/S1574070605010049*

Trapero, J. R., Cardós, M. and Kourentzes, N. (2019), 'Quantile forecast optimal combination to enhance safety stock estimation', *International journal of forecasting* **35**(1), 239–250.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207018300918*

Trenkler, G. and Gotu, B. (1998), Combination of forecasts: a bibliography, Technical report, Technical Report.

Vaiciukynas, E., Danenas, P., Kontrimas, V. and Butleris, R. (2020), 'Meta-Learning for time series forecasting ensemble'.
**URL:** *http://arxiv.org/abs/2011.10545*

Vokurka, R. J., Flores, B. E. and Pearce, S. L. (1996), 'Automatic feature identification and graphical support in rule-based forecasting: a comparison', *International journal of forecasting* **12**(4), 495–512.
**URL:** *https://linkinghub.elsevier.com/retrieve/pii/S0169207096006826*

Wallis, K. F. (2011), 'Combining forecasts – forty years later', *Applied Financial Economics* **21**(1-2), 33–41.
**URL:** *https://doi.org/10.1080/09603107.2011.523179*

Wang, X. and Petropoulos, F. (2016), 'To select or to combine? the inventory performance of model and expert forecasts', *International Journal of Production Research* **54**(17), 5271–5282.
**URL:** *https://doi.org/10.1080/00207543.2016.1167983*

Wang, Y., Zhang, N., Tan, Y., Hong, T., Kirschen, D. and Kang, C. (2018), 'Combining probabilistic load forecasts'.
**URL:** *http://arxiv.org/abs/1803.06730*

Weigel, A. P., Liniger, M. A. and Appenzeller, C. (2008), 'Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?', *Quarterly Journal of the Royal Meteorological Society* **134**(630), 241–260.
**URL:** *http://doi.wiley.com/10.1002/qj.210*

Weiss, C. (2018), Essays in Hierarchical Time Series Forecasting and Forecast Combination, PhD thesis.
**URL:** *https://www.repository.cam.ac.uk/handle/1810/274757*

Weiss, Christoph, e., Raviv, E. and Roetzer, G. (2019), 'Forecast combinations in R using the ForecastComb package', *The R journal* **10**(2), 262.
**URL:** *https://journal.r-project.org/archive/2018/RJ-2018-052/index.html*

Winkler, R. L. and Makridakis, S. (1983), 'The combination of forecasts', *Journal of the Royal Statistical Society. Series A* **146**(2), 150.
**URL:** *https://www.jstor.org/stable/10.2307/2982011?origin=crossref*

Wolpert, D. H. (1992), 'Stacked generalization', *Elsevier Oceanography Series* **87545**(505), 1–57.

Xie, J. and Hong, T. (2016*a*), 'GEFCom2014 probabilistic electric load forecasting: An integrated solution with forecast combination and residual simulation', *International journal of forecasting* **32**(3), 1012–1016.
**URL:** *http://dx.doi.org/10.1016/j.ijforecast.2015.11.005*

Xie, J. and Hong, T. (2016*b*), 'GEFCom2014 probabilistic electric load forecasting: An integrated solution with forecast combination and residual simulation', *International journal of forecasting* **32**(3), 1012–1016.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0169207015001405*

Xie, J., Hong, T., Laing, T. and Kang, C. (2017), 'On normality assumption in residual simulation for probabilistic load forecasting', *IEEE transactions on smart grid* **8**(3), 1046–1053.
**URL:** *http://dx.doi.org/10.1109/TSG.2015.2447007*

Yu, L., Wang, S. and Lai, K. K. (2005), 'A novel nonlinear ensemble forecasting model incorporating GLARand ANN for foreign exchange rates', *Computers & operations research* **32**(10), 2523–2541.
**URL:** *https://www.sciencedirect.com/science/article/pii/S030505480400156X*

Zhang, G. P. and Berardi, V. L. (2001), 'Time series forecasting with neural network ensembles: an application for exchange rate prediction', *The Journal of the Operational Research Society* **52**(6), 652–664.
**URL:** *https://doi.org/10.1057/palgrave.jors.2601133*

Zhang, S., Wang, Y., Zhang, Y., Wang, D. and Zhang, N. (2020), 'Load probability density forecasting by transforming and combining quantile forecasts', *Applied energy* **277**, 115600.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0306261920311065*

Zhao, S. and Feng, Y. (2020), 'For2For: Learning to forecast from forecasts'.
**URL:** *http://arxiv.org/abs/2001.04601*

Zhou, Z.-H., Wu, J. and Tang, W. (2002), 'Ensembling neural networks: Many could be better than all', *Artificial intelligence* **137**(1), 239–263.
**URL:** *https://www.sciencedirect.com/science/article/pii/S000437020200190X*

Zhu, Y. (2005), 'Ensemble forecast : A new approach to uncertainty and predictability', *Advances in Atmospheric Sciences* **22**(6), 781–788.