Exploring the sources of uncertainty: why does bagging work?

Fotios Petropoulos, Rob Hyndman & Christoph Bergmeir

Cardiff University, UK

Monash University, Australia

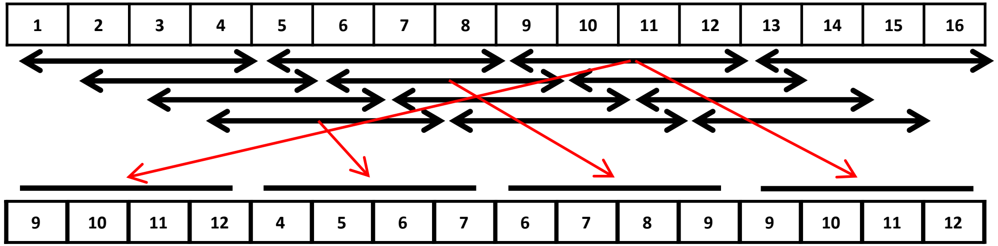
Bagging [Bergmeir et al., 2016]

Bootstrap + **agg**regation = bagging

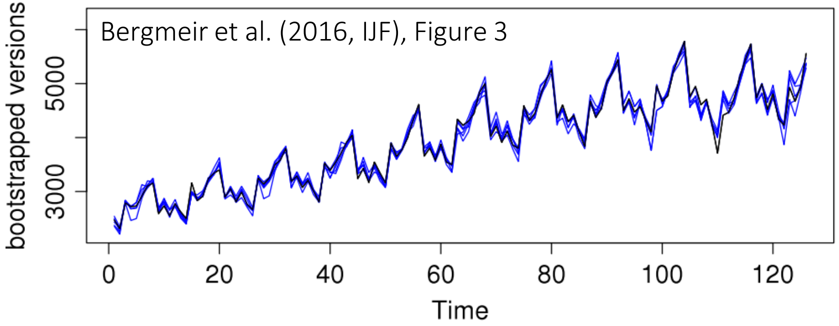
- Bergmeir, Hyndman and Benitez (2016, IJF) successfully employed the bootstrap aggregation technique for improving the performance of exponential smoothing.
- How does bagging work?
- Why does bagging work?

Bootstrapping

- **Box-Cox** transformed data: the variance is stabilised
- **STL** decomposition: seasonal, trend, remainder.
- **Bootstrap** the remainder, using Moving Block boostrap.



- The series is **reconstructed** by its structural components and the new remainder.



Aggregation

- The optimal model and set of parameters is separately identified for the original data and each of the new time series (bootstraps).
- A **different** model and/or set of parameters can be selected for each bootstrap.
- A **set of forecasts** is produced from each series/bootstrap.
- The final forecasts are calculated as the **average (aggregation)** across all series/bootstraps for each horizon.
- How? Arithmetic mean, median, mode, **trimmed mean**.

Why does it work? …sources of uncertainty

- Bootstrapping renders the data **less sensitive** to outliers – the remainder is resampled.
- Model selection and parametrisation **across multiple bootstraps** provides immunity against incorrectly identifying an "optimal" model and set of parameters.
- "All models are **wrong** but some are useful…"
- **Three sources of uncertainty** that bagging mitigates:
  - Data uncertainty, Model uncertainty, Parameters uncertainty

Decomposing the benefits of bagging
Model uncertainty

1. Automatic **model selection** is performed on each series and bootstrap separately.
2. Each **uniquely selected model** form is then applied back to the original data. *Bootstraps are **not used** for producing forecasts.*
3. Forecasts from the selected model forms are combined with **weights** corresponding to the **frequency** that the respective models were identified as optimal.

This is a **weighted model combination approach**, where the selected models and weights are directly derived from automatic model selection on the original data and the bootstraps.

Decomposing the benefits of bagging
Data uncertainty

1. Automatic model selection and parametrisation is performed on the **original data only**.
2. The identified **optimal model form** (with the optimal set of **parameters**) is applied to each one of the bootstraps.
3. Forecasts are produced from **both** original data and bootstraps.
4. The final forecasts are calculated as the **trimmed mean** (5%) across all series/bootstraps for each horizon.

This is a **forecast aggregation approach**, where a single optimal model form and set of parameters is applied to the original series and bootstraps.

Decomposing the benefits of bagging

Parameter uncertainty

1. Automatic model selection is performed on the **original data**.
2. The identified optimal form is applied on the bootstraps, so that **different sets of optimal parameters** are generated.
   Bootstraps are only used for identifying ***sets of parameters****.*
3. The single model form together with the optimised set of parameters is then applied back to the **original series**.
4. Finally, forecasts are calculated as the **trimmed mean** (5%).

This is a **forecast aggregation approach**, where a single optimal model form and various set of parameters are applied to the original series.

Benchmarking

Simple benchmark: Automatic model selection ETS and ARIMA [Hyndman & Khandakar, 2008]

- A **single model** (and a single set of parameters) is selected, based on information criteria.
- A **single set of forecasts** is produced.

Model combination [Kolassa, 2011]

- **All possible models** are fitted on the original time series and parameters are optimised for each model separately.
- The forecasts produced from the different models are combined using **weights** derived from the values of **information criteria**.

Bagging [Bergmeir et al., 2016]

Mapping the sources of uncertainty

	Model Uncertainty	Data Uncertainty	Parameter Uncertainty
Simple benchmark			
Model combination benchmark	✓		
Bagging	✓	✓	✓
Bootstrap model combination	✓		
Bagging: single model & set of parameters		✓	
Bootstrap parametrisation			✓

Data and design

- M- and M3-competition data (**4,000+ time series**).
- Various **frequencies**: yearly, quarterly, monthly, other.
- Multiple forecast **horizons**: up to 18 periods.
- For each series we generate **100 bootstraps**.
- Two **bootstrapping** methods: MBB and LPB. *Similar insights are obtained for both methods*
- Two **forecasting** methods: ETS and AutoARIMA.
- Forecasting **evaluation** using sMAPE and MASE.

Results: 826 **yearly** time series

	ETS		ARIMA	
	sMAPE	MASE	sMAPE	MASE
Simple benchmark	17.36	3.03	17.15	3.08
Model combination benchmark	16.95	3.11		
Bagging	17.17	3.01	17.27	3.04
Bootstrap model combination	17.21	2.98	16.81	2.98
Bagging: single model & set of parameters	17.33	3.04	17.05	3.03
Bootstrap parametrisation	16.78	2.98	18.58	3.49

best worst

Results: 959 **quarterly** time series

	ETS		ARIMA	
	sMAPE	MASE	sMAPE	MASE
Simple benchmark	11.06	1.27	11.96	1.31
Model combination benchmark	11.99	1.34		
Bagging	11.44	1.29	11.50	1.29
Bootstrap model combination	10.99	1.26	11.28	1.27
Bagging: single model & set of parameters	11.20	1.27	11.79	1.30
Bootstrap parametrisation	11.16	1.28	11.86	1.33

best worst

Results: 2045 **monthly** time series

	ETS		ARIMA	
	sMAPE	MASE	sMAPE	MASE
Simple benchmark	14.36	0.93	15.25	0.96
Model combination benchmark	14.19	0.94		
Bagging	14.04	0.90	14.45	0.92
Bootstrap model combination	14.06	0.91	14.54	0.92
Bagging: single model & set of parameters	14.37	0.92	15.00	0.95
Bootstrap parametrisation	14.20	0.92	14.82	0.98

best worst

Overall results: 4,004 time series

	ETS		ARIMA	
	Rel sMAPE	Rel MASE	Rel sMAPE	Rel MASE
Simple benchmark	1.000	1.000	1.000	1.000
Model combination benchmark	0.998	1.015		
Bagging	0.987	0.988	0.957	0.972
Bootstrap model combination	**0.982**	**0.979**	**0.956**	**0.964**
Bagging: single model & set of parameters	1.003	0.999	0.986	0.989
Bootstrap parametrisation	0.989	0.989	0.989	1.057

Final thoughts

- **Bagging** is a new and robust method for forecasting univariate data.
- Bagging's good performance is related with mitigating the **three sources of uncertainty**: model, parameter and data.
- A decomposition and simulation exercise reveals that **model uncertainty** is the basic source of performance improvement.
- We proposed a **new model combination** framework where the weights are based on bootstrapping.

Questions?

PetropoulosF@cardiff.ac.uk – http://fpetropoulos.eu