Forecasting energy consumption combining Bagging and Clusters

Tiago Dantas
Fernando Cyrino Oliveira

Pontifical Catholic University of Rio de Janeiro,Brazil

ISF 2018, Boulder-USA, June 17-20

Motivation

Oliveira and Cyrino Oliveira (2018) used Bootstrap Aggregating (Bagging) to generate highly accurate energy consumption forecasts for developed and developing countries

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods

Erick Meira de Oliveira [a, b, *], Fernando Luiz Cyrino Oliveira [a]

Motivation

- The approach generates Bootstrap versions using Sieve Bootstrap (RSB) and Moving Block Bootstrap (MBB)
- Forecasts for each of them are made using ETS and ARIMA models and, subsequently aggregated using the mean and median

Motivation

- Dantas and Cyrino Oliveira, **International Journal of Forecasting - forthcoming**, have demonstrated that for the Bagging + ETS case, the inclusion of a clustering phase in the procedure might potentially reduce the forecast error

- M3 and CIF 2016 data

Motivation

- The aim of this work is to revisit the Oliveira and Cyrino Oliveira (2018) paper, including a clustering phase to reduce the forecast error

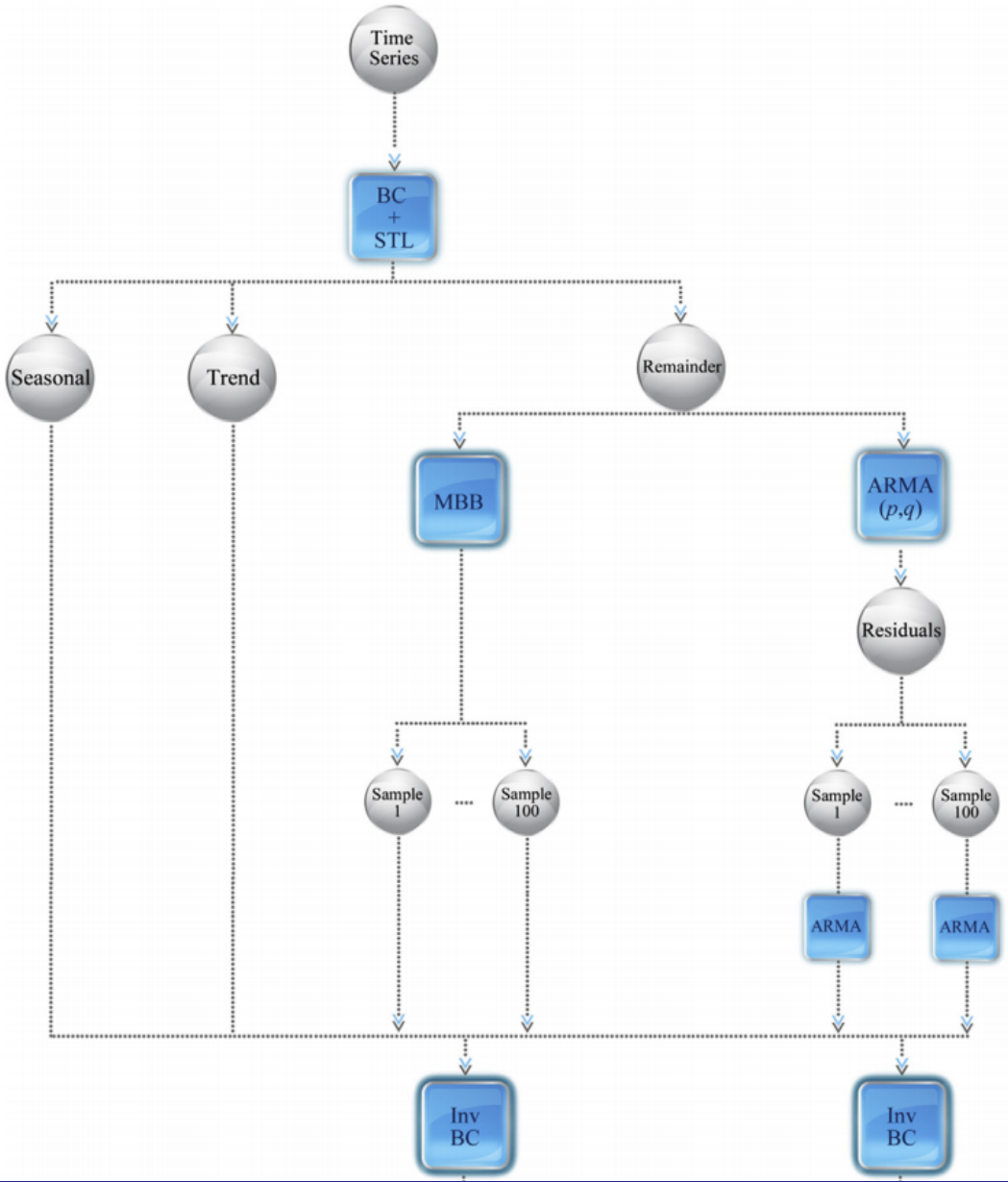
Data

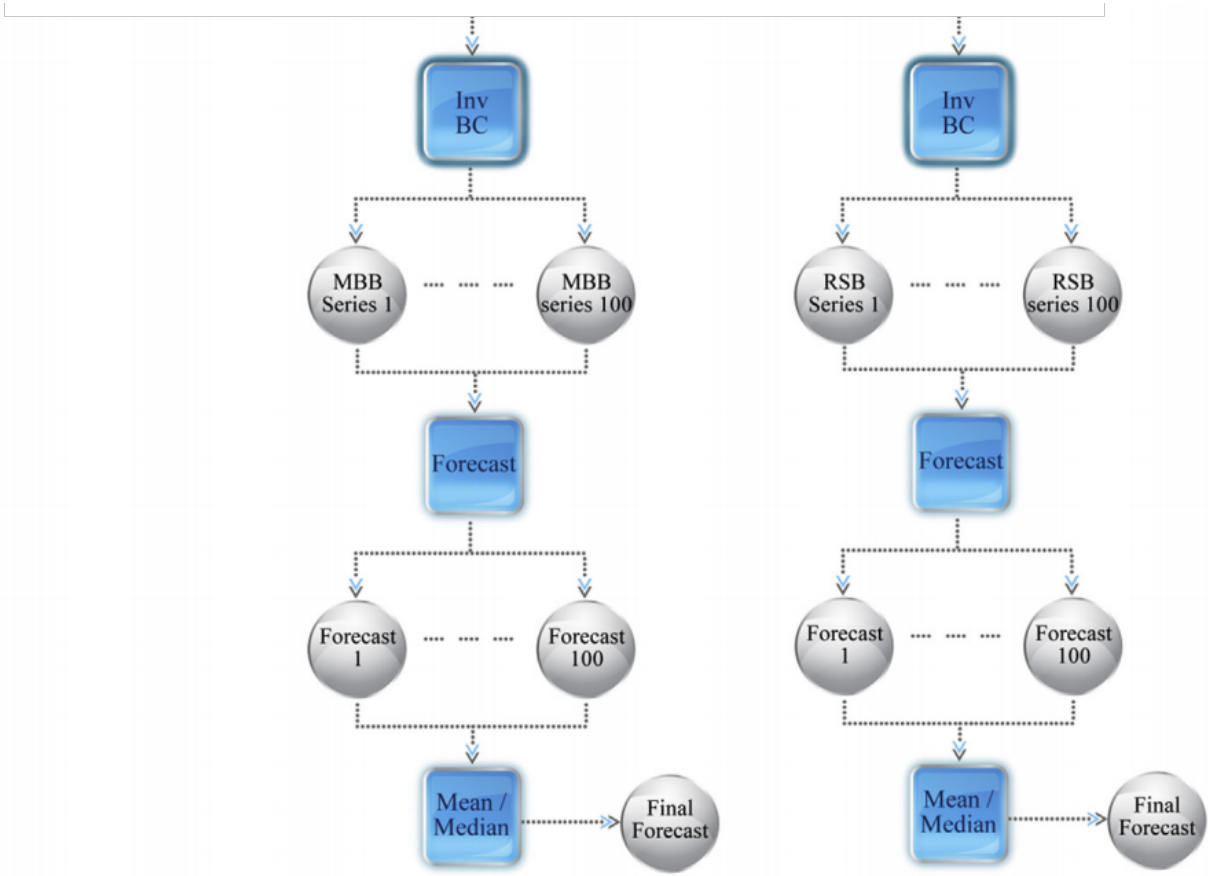
Monthly data of total electricity energy consumption (GWh) in different developed and developing countries

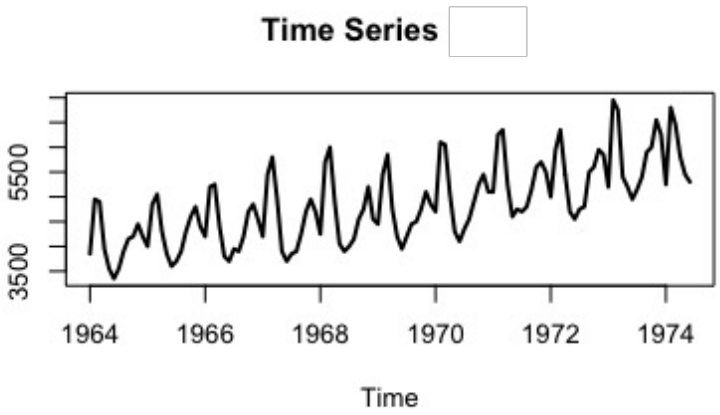
- Countries: Brazil, Canada, France, Italy, Japan, Mexico and Turkey
- Training set: July 2006-December 2014
- Test set: January 2015-December 2016 (24 months)
- Codes developed using R

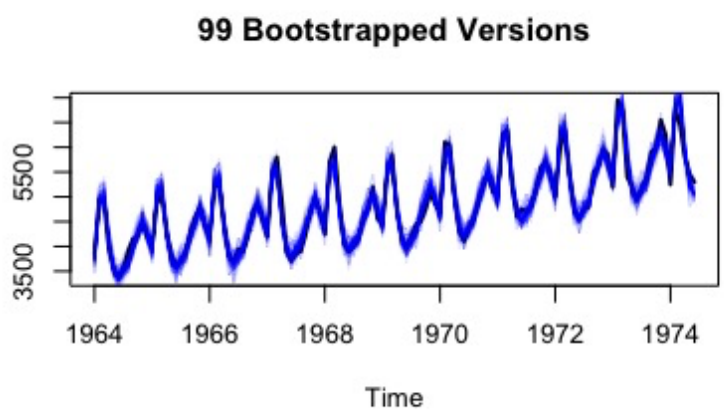
The main approaches in Oliveira and Cyrino Oliveira (2018)

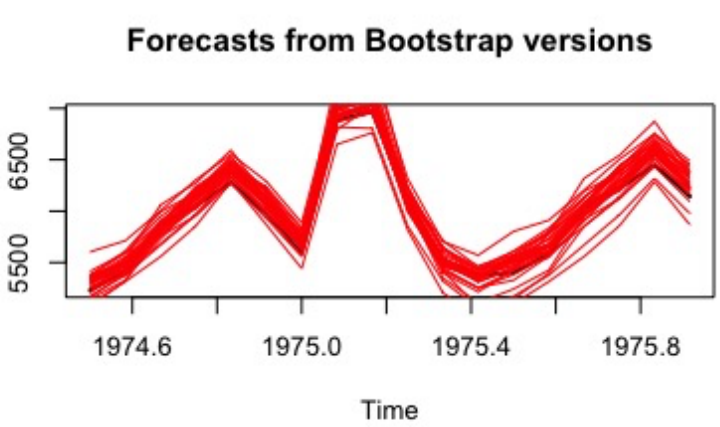
- Bagged.BLD.MBB.ETS (proposed by Bergmeir et al., 2016)
- Bagged.BLD.RSB.ETS
- Bagged.BLD.MBB.ARIMA
- Bagged.BLD.RSB.ARIMA

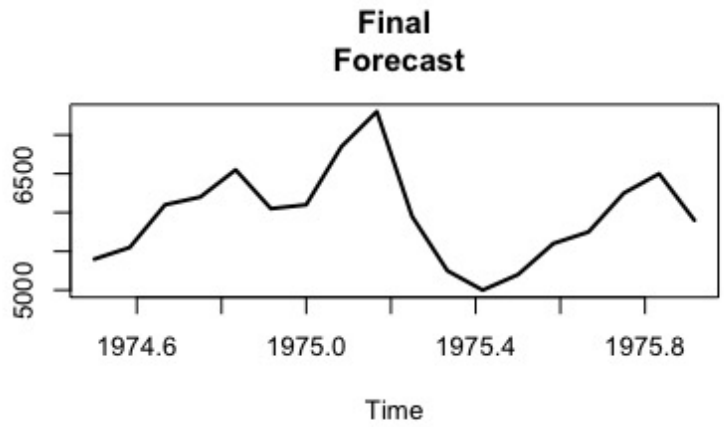

Time Series
BC + STL
Seasonal
Trend
Remainder
MBB
ARMA (p,q)
Residuals
Sample 1
Sample 100
Sample 1
Sample 100
ARMA
ARMA
Inv BC
Inv BC


Inv BC
MBB Series 1
MBB series 100
Forecast
Forecast 1
Forecast 100
Mean / Median
Final Forecast
Inv BC
RSB Series 1
RSB series 100
Forecast
Forecast 1
Forecast 100
Mean / Median
Final Forecast


Time Series
5500
3500
1964
1966
1968
1970
1972
1974
Time

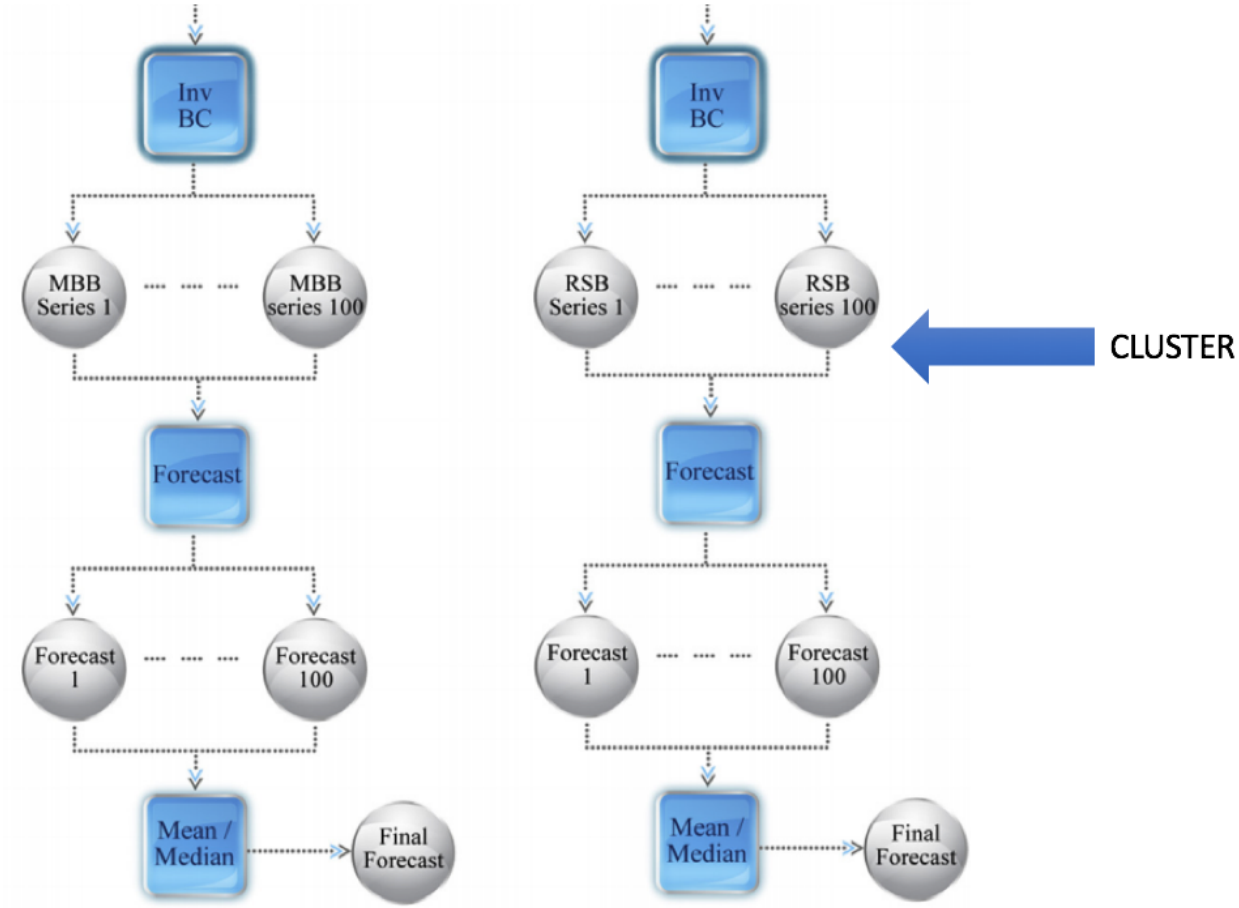

99 Bootstrapped Versions
5500
3500
1964
1966
1968
1970
1972
1974
Time


Forecasts from Bootstrap versions
6500
5500
1974.6
1975.0
1975.4
1975.8
Time


Final
Forecast
6500
5000
1974.6
1975.0
1975.4
1975.8
Time

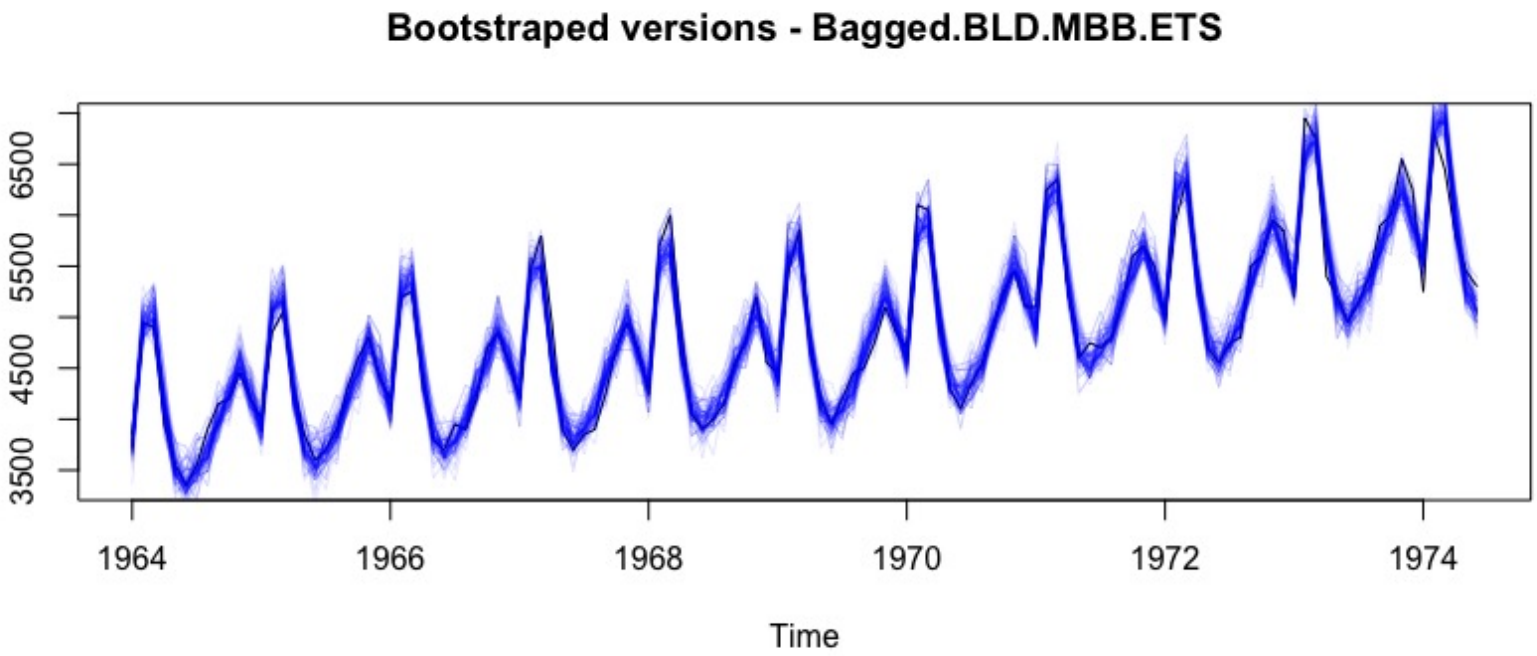
The proposed approach

- Bootstrapped series are generated in the same way as in Oliveira and Cyrino Oliveira (2018)
- The proposed approach considers a less correlated group of time series to make forecasts

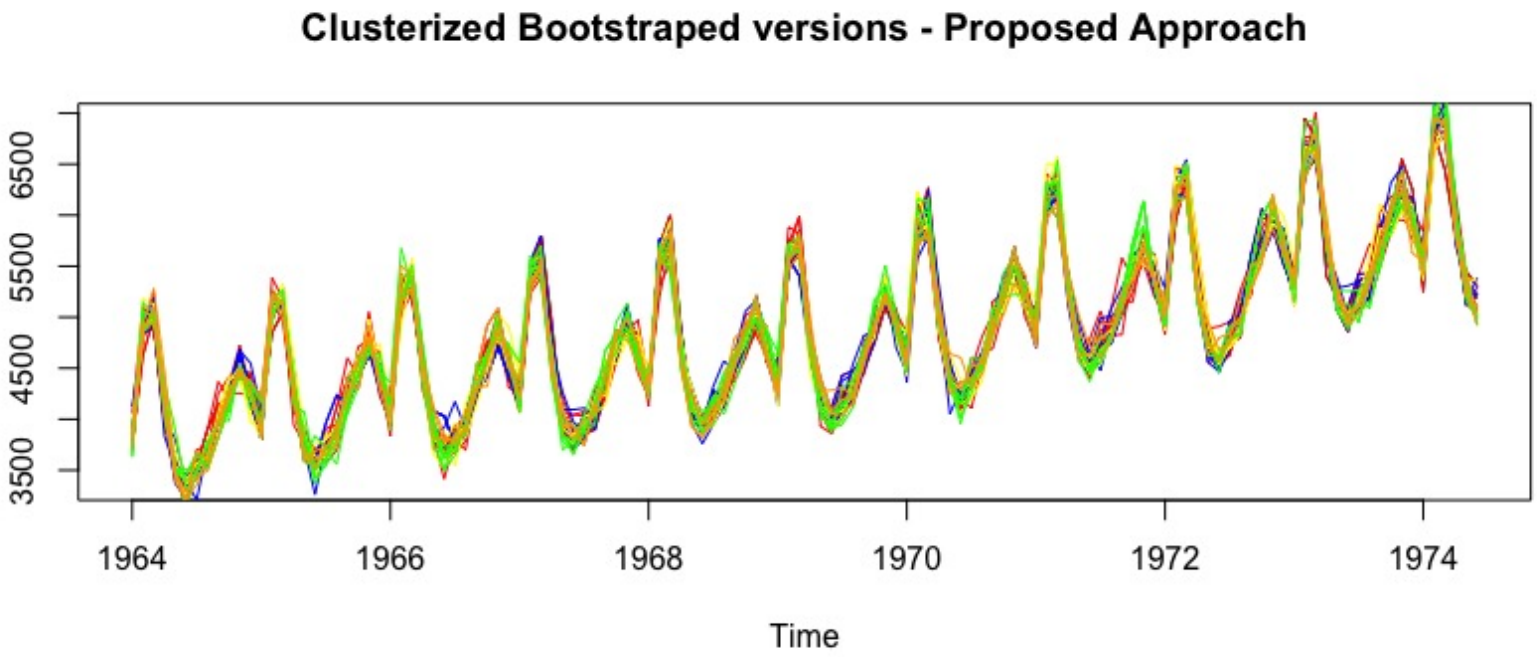

Inv BC
MBB Series 1
MBB series 100
Forecast
Forecast 1
Forecast 100
Mean / Median
Final Forecast
Inv BC
RSB Series 1
RSB series 100
CLUSTER
Forecast
Forecast 1
Forecast 100
Mean / Median
Final Forecast

The modifications - Bagged.Cluster approach

- Larger number of Bootstrap versions (1000)
- Clusters are created using PAM algorithm with euclidean distance
- A validation set is constructed (last 12 observations)
- Forecasts are made for each Bootstrapped versions (series are ranked according sMAPE)
- The best time series (according to validation) in each cluster is selected
- The number of selected time series in each cluster is proportional to the size of each cluster. On total, 100 time series are selected and combined using the mean or median


Bootstraped versions - Bagged.BLD.MBB.ETS
3500
4500
5500
6500
1964
1966
1968
1970
1972
1974
Time

Proposed Approach



Results

Bagging + ETS models

Table: sMAPE - MBB

Methods	Brazil	Canada	France	Italy	Japan	Mexico	Turkey
ETS	6.926	3.942	2.566	1.862	2.323	6.786	3.091
Bagged.BLD.MBB.ETS-Mean	6.144	3.779	2.695	1.819	**3.585**	6.028	**2.549**
Bagged.BLD.MBB.Cluster.ETS-Mean	**5.790**	**3.742**	**2.678**	**1.720**	3.797	**5.930**	2.615
Bagged.BLD.MBB.ETS-Median	6.139	3.833	2.617	1.838	**3.557**	6.122	**2.557**
Bagged.BLD.MBB.Cluster.ETS-Median	**5.953**	**3.767**	**2.565**	**1.745**	3.814	**5.830**	2.638

Table: sMAPE - RSB

Methods	Brazil	Canada	France	Italy	Japan	Mexico	Turkey
ETS	6.926	3.942	2.566	1.862	2.323	6.786	3.091
Bagged.BLD.RSB.ETS-Mean	6.238	4.125	**2.936**	**1.762**	3.903	5.631	2.617
Bagged.BLD.RSB.CLuster.ETS-Mean	**6.015**	**3.946**	3.020	1.902	**3.476**	**5.152**	**2.575**
Bagged.BLD.RSB.ETS-Median	6.310	4.174	**2.995**	**1.840**	3.960	5.864	2.763
Bagged.BLD.RSB.CLuster.ETS-Median	**6.055**	**4.057**	3.139	1.902	**3.412**	**5.515**	**2.654**

Bagging + ARIMA models

Table: sMAPE - MBB

Methods	Brazil	Canada	France	Italy	Japan	Mexico	Turkey
Arima	4.550	4.350	2.725	1.414	3.272	3.120	2.279
Bagged.BLD.MBB.ARIMA-Mean	**5.242**	**3.767**	2.594	**1.799**	**3.594**	3.403	**2.666**
Bagged.BLD.MBB.Cluster.ARIMA-Mean	5.850	3.885	**2.580**	1.971	3.844	**3.291**	2.858
Bagged.BLD.MBB.ARIMA-Median	**5.103**	**3.871**	**2.492**	**1.774**	**3.722**	3.438	**2.680**
Bagged.BLD.MBB.Cluster.ARIMA-Median	6.383	3.915	2.591	1.947	3.931	**3.273**	2.878

Table: sMAPE - RSB

Methods	Brazil	Canada	France	Italy	Japan	Mexico	Turkey
Arima	4.550	4.350	2.725	1.414	3.272	3.120	2.279
Bagged.BLD.RSB.ARIMA-Mean	4.376	4.122	2.659	**2.611**	**3.354**	3.475	**2.747**
Bagged.BLD.RSB.Cluster.ARIMA-Mean	**4.285**	**4.062**	**2.630**	3.414	3.450	**3.308**	2.861
Bagged.BLD.RSB.ARIMA-Median	4.175	4.243	2.730	**2.739**	**3.536**	3.423	**2.756**
Bagged.BLD.RSB.Cluster.ARIMA-Median	**4.017**	**4.126**	**2.641**	3.390	3.706	**3.312**	2.858

Concluding Remarks

- Bagging does improve ETS performance
- The inclusion of a clustering phase gives even better results for the ETS case
- When using ARIMA models, Bagging didn't improve perfomance in the majority of the cases
- Considering ARIMA and RSB, the clustering phase generated better results than Bagging without clusters

Thank You!

t.mendesdantas@gmail.com

References

Bergmeir, C., Hyndman, R.J., Benitez,J.M. (2016). Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation.*International Journal of Forecasting*, 32(2), 303-312.

Dantas, T.M, Cyrino Oliveira, F.L. (Forthcoming). Improving Time Series Forecasting: an Approach Combining Bootstrap Aggregation, Clusters and Exponential Smoothing, *International Journal of Forecasting*

Oliveira, E.M., Cyrino Oliveira, F.L. (2018). Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods, *Energy*, 144, 776-788