How to Bootstrap Time Series without Attracting Attention of Statisticians

Ivan Svetunkov

ISF 2024

1st July 2024

Marketing Analytics and Forecasting | Lancaster University Management School

Introduction

Bootstrapping is a well-known technique

Used when analytical solutions are not available

Classical example – regression analysis to get distribution of parameters

In time series, it can also be used to capture uncertainty

Classical approaches might note be appropriate due to potential dependence of observations.

STL-based Bootstrap

Bergmeir et al. (2016) proposed "Bagged ETS". The idea was:

1. Use STL (Cleveland et al., 1990) to decompose time series;
2. Bootstrap its residuals to recreate new time series;
3. Apply ETS to each of new time series;
4. Produce forecasts for them;
5. Aggregate them using mean.

The method improves accuracy in comparison with conventional ETS.

STL-based Bootstrap

Petropoulos et al. (2018) explained why Bagged ETS works:

- It captures potential parameters uncertainty;
  - e.g. what's the true value of α?
- It addresses model selection uncertainty issue.

Sarris et al. (2020) developed model selection based on STL bagging.

STL-based Bootstrap

STL-based bootstrap:

+ Proven to work well for forecasting on low frequency;
- It assumes a model:
  - If the structure is captured incorrectly, it won't work;
  - Breaks on data with promotions/calendar effects/...;
- Might break on non-seasonal data:
  - Solution: assume another model, i.e. use pure LOWESS;
- Multiple seasonalities?
  - Solution: use a modification, called MSTL;
- Not suitable for intermittent demand.

Maximum Entropy Bootstrap

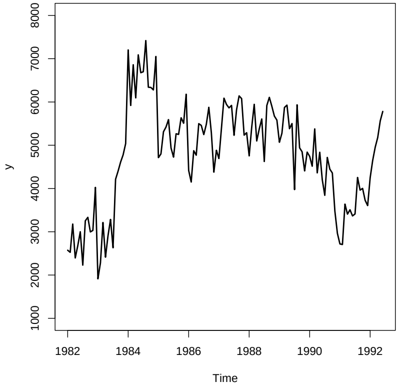
Vinod and de Lacalle (2009) proposed “Maximum Entropy Bootstrap” (MEB)

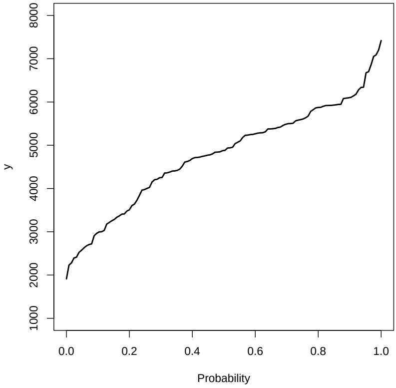
The main idea of the method is:

- Record where the actual values are in the data;
- Sort the values in the ascending order;
- Do some magic, introducing randomness;
- Return back using the original order.

Maximum Entropy Bootstrap

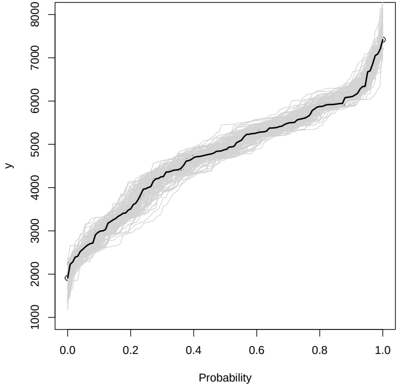
This is the first two steps:

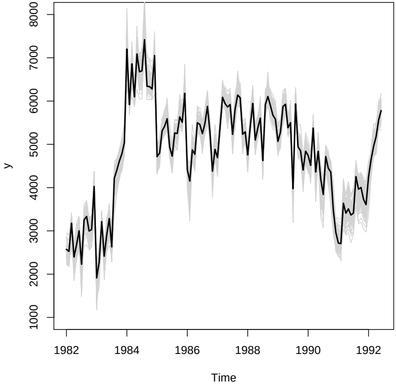




Maximum Entropy Bootstrap

And then we move back:





Maximum Entropy Bootstrap

MEB:

- \+ Doesn't assume a model;
- \+ Works on any data;
- – Might break the structure:
  - If data is positive but low level, will generat negative values;
- – Not suitable for intermittent demand;
- – Not clear whether it works well;
- – Some steps are hard to understand;
- – Wasn't explored in forecasting properly.

The Ideal Time Series Bootstrap

What do we want from a time series bootstrap?

1. Model free;
2. Preserves the structure (trend, seasonality, promotions etc);
3. Does not break autocorrelations;
4. Works on a wide variety of series;
5. Accurate;
6. Fast;
7. Simple.

Why should we care?

1. "Model free" is important;
2. Neither of the existing ones works on intermittent demand;
3. No proper exploration of predictive distributions for ETS in the literature.

New method

New Method: Nonparametric version

1. Take logarithms of the data if it is positive;
2. Use the idea of Vinod and de Lacalle (2009) with ordering, creating series z_j;
3. Take differences of the ordered data and sort them to get Δ_{z_j};
4. Smooth both using Friedman's SuperSmoother (Friedman, 1984)
   - Alternative: LOWESS or whatever else;
5. Sample values from Δ_{z_j} (using uniform distribution) and add them to z_j;
6. Create n new series like that;

New Method: Nonparametric version

7. Scale data to make sure that:
   - Variance is stable over j;
   - Variance is similar to the z_j;
8. Sort values based on the initial indices;
9. Centre the new series around z_j;
10. Take exponent if (1) was done.

In case of intermittent demand, do this separately for demand sizes and intervals.

Regular demand with zeroes? Ignore them and keep them in the new data.

New Method: Parametric version

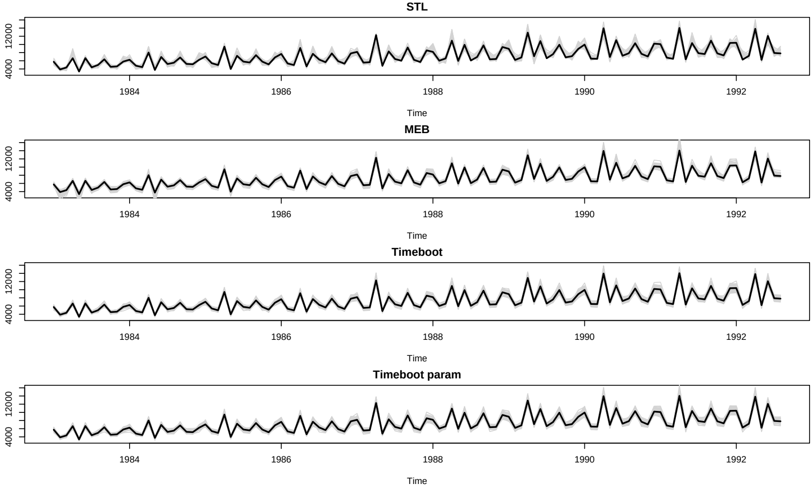
The alternative is to generate differences Δ_{z_j} from a normal distribution.

$\mu = 0$

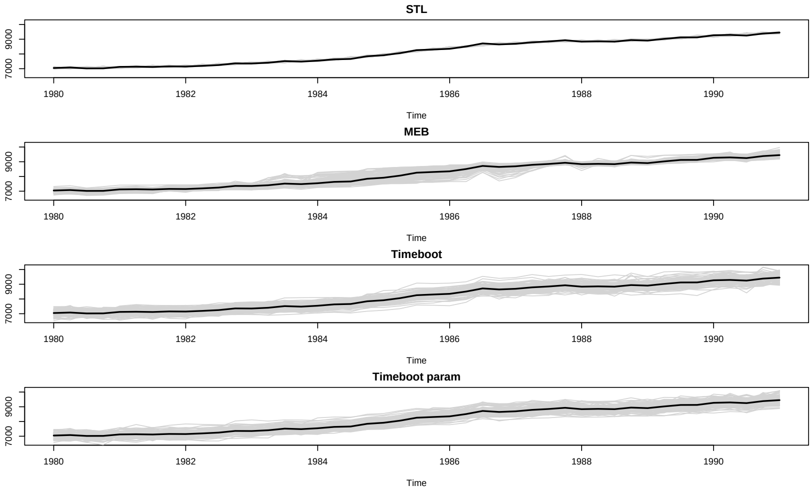
$\hat{\sigma} = \frac{1}{T-1} \sum_{t=1}^{T-1} |y_t - y_{t-1}|$

Mean absolute differences of the original data...

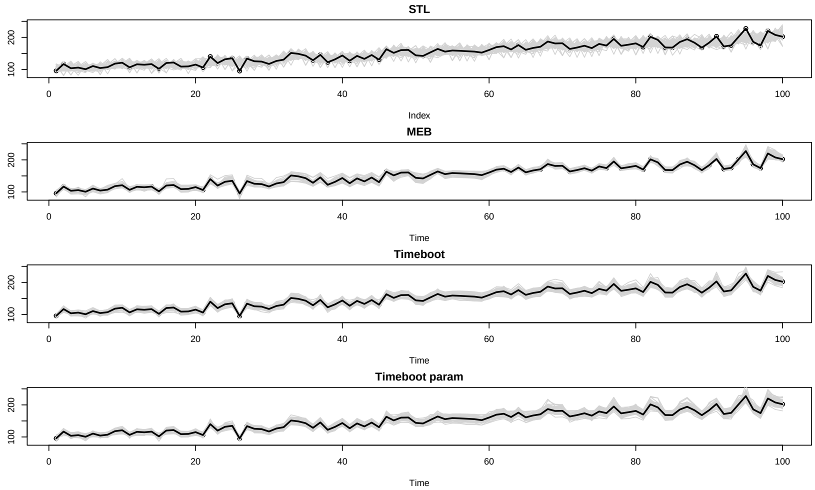
Examples: trend & seasonality



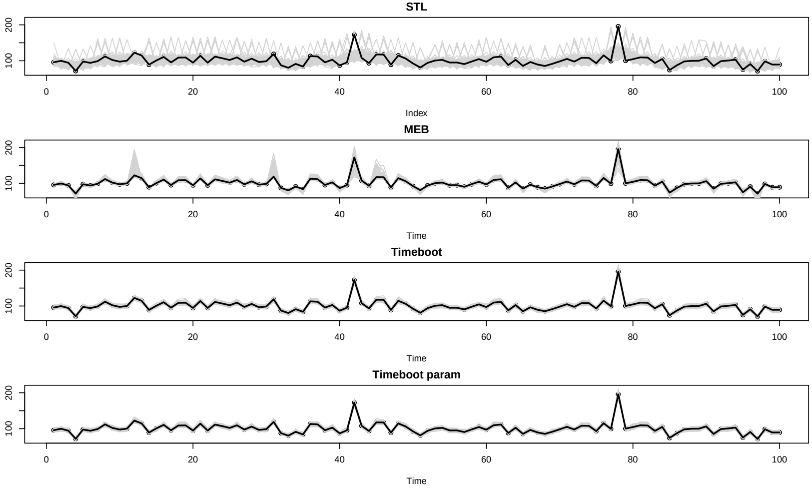
Examples: stochastic trend



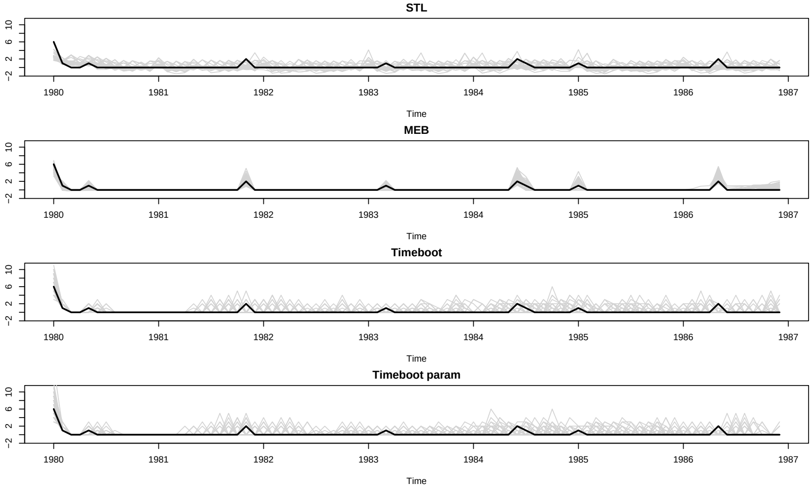
Examples: deterministic trend



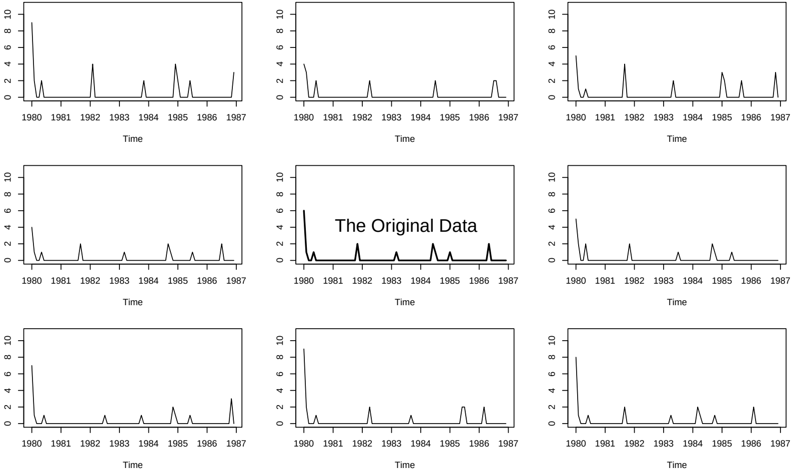
Examples: data with promotions



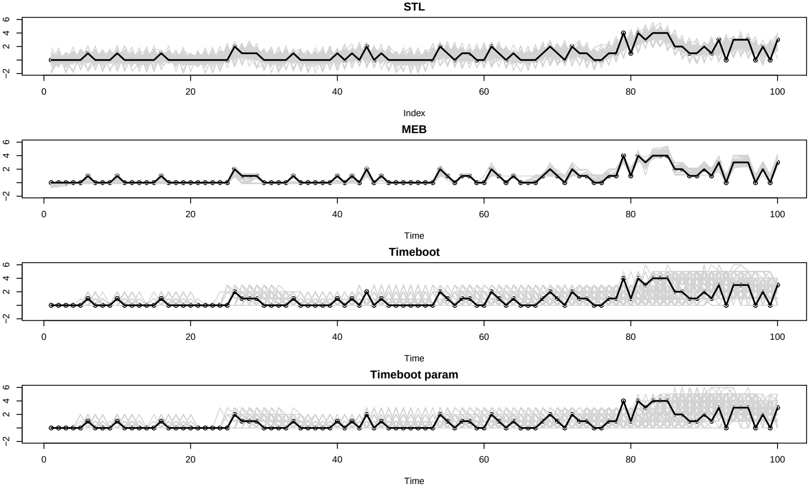
Examples: intermittent demand



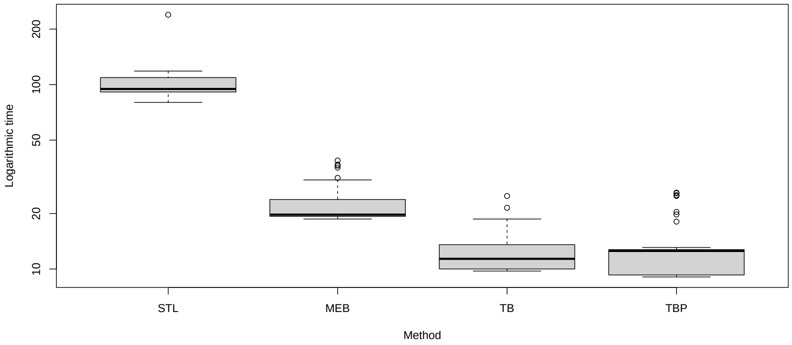
Examples: intermittent demand



Examples: intermittent demand, buildin up



Computational time on time series with 126 obs



An experiment

Experiment: the setting

M1 (Makridakis et al., 1982), M3 (Makridakis and Hibon, 2000) and Tourism (Athanasopoulos et al., 2011) competitions, 5316 time series.

Mixture of yearly, quarterly and monthly data.

`adam()` function from the `smooth` package for R (R Core Team, 2024; Svetunkov, 2024b) to apply ETS.

ETS with AIC weights (Kolassa, 2011).

ETS applied to the data from STL, MEB, TB and TBP.

`forecast` (Hyndman and Khandakar, 2008), `meboot` (Vinod and de Lacalle, 2009) and `greybox` (Svetunkov, 2024a) for STL, MEB and TB respectively.

99 bootstraps per series.

Experiment: what do we produce?

Produce:

- Point forecasts;
- Parametric quantiles from ETS;
- Trimmed means of point forecasts from bootstrapped data (5%);
- Quantiles from bootstrapped models:
  - 1000 scenarios from ETS applied to each bootstrapped data;
  - Take quantiles of scenarios for each horizon.

Generate quantiles from 1% to 99%.

Experiment: what do we measure?

RMSSE and scaled ME for point forecasts.

"Calibration" for quantiles (how many values lie below the quantile).

Scaled pinball for quantiles.

Experiment results: point forecasts

RMSSEs overall:

	mean	min	Q1	median	Q3	max
STL	2.57	1.13	**1.43**	*2.74*	3.35	4.35
MEB	2.58	**1.06**	*1.45*	2.77	3.37	4.42
TB	*2.55*	1.08	1.48	**2.72**	*3.32*	*4.32*
TBP	**2.54**	1.08	1.47	**2.72**	**3.31**	**4.31**
ETS	2.65	*1.07*	1.47	2.79	3.42	4.69
ETS(C,C,C)	2.63	*1.07*	1.47	2.77	3.38	4.65

Experiment results: point forecasts

sME overall:

	mean	min	Q1	median	Q3	max
STL	0.22	-0.04	**0.14**	0.25	**0.31**	0.41
MEB	**0.20**	-0.05	**0.14**	**0.22**	**0.31**	**0.37**
TB	0.23	-0.04	0.15	0.24	0.34	0.39
TBP	0.23	**-0.02**	0.16	0.25	0.34	0.40
ETS	0.23	**-0.02**	0.16	0.26	0.35	0.39
ETS(C,C,C)	0.22	-0.04	**0.14**	0.24	0.33	0.39

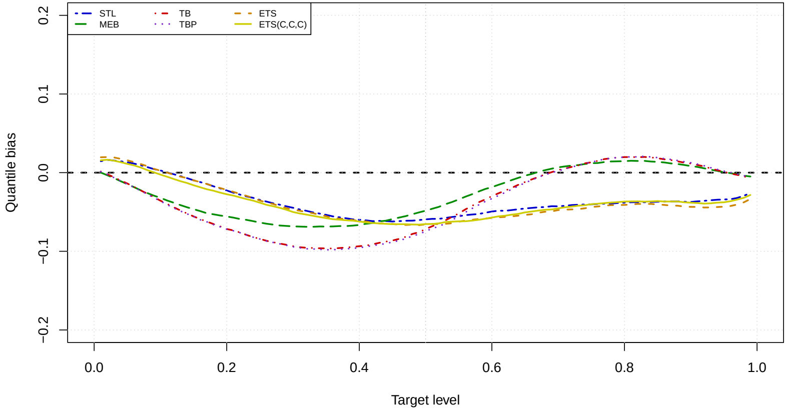
We consistently underforecast!

Experiment results: quantiles, overall

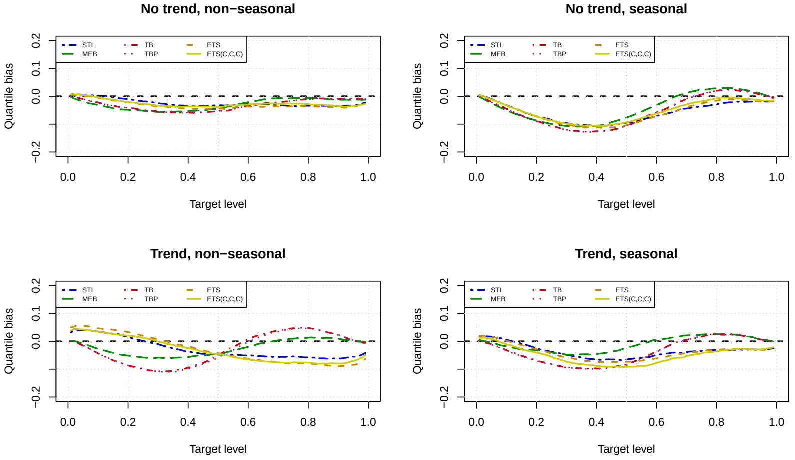
	Quantile bias	Quantile precision	Pinball
STL	-0.0361	0.0383	1549.7885
MEB	**-0.0265**	**0.0322**	**1537.0799**
TB	-0.0369	0.0437	1575.1900
TBP	-0.0377	0.0447	1576.3401
ETS	-0.0396	0.0422	1570.4766
ETS(C,C,C)	-0.0395	0.0414	1539.1906

Quantiles are consistently lower than needed.

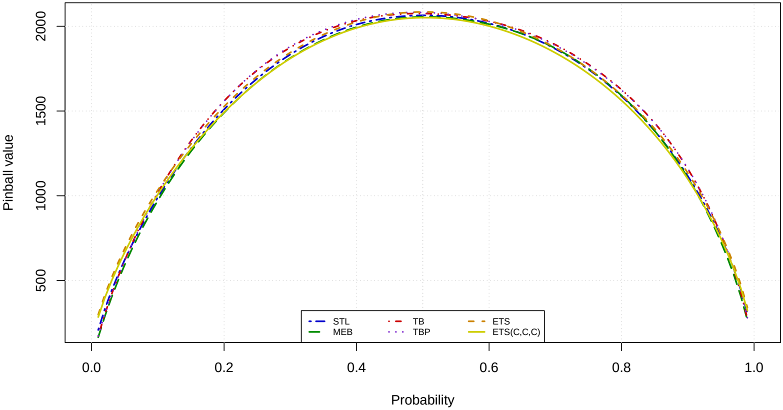
Experiment results: quantile performance, overall



Experiment results: quantile performance, detailed



Experiment results: pinballs, overall



Conclusions

Conclusions

- Time series bootstap approaches are useful;
- They improve accuracy of dynamic models;
- They seem to improve the predictive distribution;
- But the distributions are miscalibrated;
- The new Time series Bootstrap works well;
- More experiments are required...

Thank you for your attention!

Ivan Svetunkov

i.svetunkov@lancaster.ac.uk

https://www.openforecast.org

Marketing Analytics and Forecasting

Lancaster University Management School

References I

Athanasopoulos, G., Hyndman, R. J., Song, H., Wu, D. C., 2011. The tourism forecasting competition. International Journal of Forecasting 27 (3), 822–844.
URL http://dx.doi.org/10.1016/j.ijforecast.2010.04.009

Bergmeir, C., Hyndman, R. J., Benítez, J. M., 2016. Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation. International Journal of Forecasting 32 (2), 303–312.
URL http://dx.doi.org/10.1016/j.ijforecast.2015.07.002

References II

Cleveland, R. B., Cleveland, W. S., McRae, J. E., Terpenning, I., 1990. STL: A seasonal-trend decomposition procedure based on loess. Journal of Official Statistics 6 (1), 3–73.

Friedman, J. H., 1984. A variable span smoother. Laboratory for Computational Statistics, Department of Statistics, Stanford ....

Hyndman, R. J., Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software 27 (3), 1–22.
URL http://www.jstatsoft.org/v27/i03

References III

Kolassa, S., 2011. Combining exponential smoothing forecasts using Akaike weights. International Journal of Forecasting 27 (2), 238–251.
URL http://dx.doi.org/10.1016/j.ijforecast.2010.04.006

Makridakis, S., Andersen, A. P., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., Winkler, R. L., 1982. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. Journal of Forecasting 1 (2), 111–153.
URL http://doi.wiley.com/10.1002/for.3980010202

References IV

Makridakis, S., Hibon, M., oct 2000. The M3-Competition: results, conclusions and implications. International Journal of Forecasting 16 (4), 451–476.
URL https://linkinghub.elsevier.com/retrieve/pii/S0169207000000571

Petropoulos, F., Hyndman, R. J., 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://doi.org/10.1016/j.ejor.2018.01.045

References V

R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL https://www.R-project.org/

Sarris, D., Spiliotis, E., Assimakopoulos, V., 2020. Exploiting resampling techniques for model selection in forecasting: an empirical evaluation using out-of-sample tests. Operational Research 20 (2), 701–721.

Svetunkov, I., 2024a. greybox: Toolbox for Model Building and Forecasting. R package version 2.0.1.
URL https://github.com/config-i1/greybox

References VI

Svetunkov, I., 2024b. smooth: Forecasting Using State Space Models. R package version 4.0.2.
URL https://github.com/config-i1/smooth

Vinod, H. D., de Lacalle, J. L., 2009. Maximum entropy bootstrap for time series: The meboot R package. Journal of Statistical Software 29 (5), 1–19.