FORESIGHT

The International Journal of Applied Forecasting

ARTICLES *by* FORECASTERS *for* FORECASTERS: Q3:2023

Cross-Learning with Short Seasonal Time Series

Join the *Foresight* readership by becoming a member of the International Institute of Forecasters
forecasters.org/foresight/

made available to you with permission from the publisher

Cross-Learning with Short Seasonal Time Series

HUIJING CHEN, JOHN BOYLAN, AND IVAN SVETUNKOV

PREVIEW *Since its introduction by R. G. Brown over 60 years ago, exponential smoothing in its various flavors has been a go-to model for many forecasting professionals. Thanks to its solid performance across 40 years of M competitions, exponential smoothing has earned a secure place in the forecaster's toolbox. The familiar Error-Trend-Seasonality (ETS) taxonomy by Hyndman and colleagues helps define how components of a time series interact with each other, and this new research by Chen, Boylan, and Svetunkov provides an enhanced taxonomy that can aid in cross-learning from similar time series with very short histories.*

INTRODUCTION

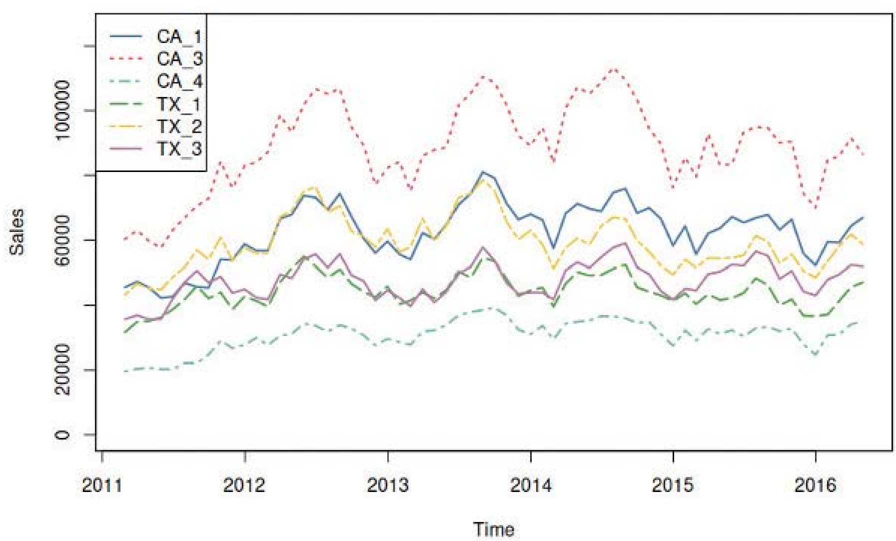
There are many challenges in forecasting demand in the retail sector. Short-term forecasts are required frequently, and the number of products to forecast can be huge. Lengths of data history are typically short, and the product portfolio can vary with some items discontinued while new items are being added. Many products display seasonal variations due to natural seasons (summer and winter), culture and festivities (New Year, Easter, Eid, Christmas, etc.), and other factors. **Figure 1** shows an example of a food item sold in six different Walmart department stores (taken from the M5 competition data). We can observe very clear seasonal variations with similar (but not identical) patterns across the stores.

Intuitively, it seems reasonable to suppose that we can exploit data from similar series to help forecast each individual series. This is sometimes known as "cross-learning" (Smyl, 2020). But what aspects of a series should we try to estimate from similar series, and which aspects should we treat individually? This article provides a framework to help answer this question.

AN ESTABLISHED TAXONOMY: ERROR, TREND, SEASONAL (ETS)

A time series can be decomposed into level, trend, seasonality, and random error components. Based on this

Figure 1. Monthly sales of one item at six Walmart stores.



Key Points

- One of the biggest challenges in forecasting demand in the retail sector is to identify and estimate seasonal variations accurately with very short time-series data. The concept of "cross-learning" is useful in this context, to exploit data from similar time series.
- We develop a new taxonomy, based on vector exponential smoothing (VETS) models, to decide whether *Parameters*, *Initial values*, and *Components* (PIC) should be common across time series.
- This framework performs well on the M5 (Walmart) data in terms of point forecasts and prediction intervals.
- The *legion* package for R has been developed, with an automatic model selection functionality for easy application in practice.

decomposition approach, Hyndman and colleagues (2002, 2008) developed an Error-Trend-Seasonality (ETS) taxonomy, which is a helpful way to define the components and how they interact with each other. According to this taxonomy, the error term can be additive (A) or multiplicative (M), and there are five types of trend: none (N), additive (A), additive damped (Ad), multiplicative (M), and multiplicative damped (Md). Lastly, there are three types of seasonal components: none (N), additive (A) and multiplicative (M). **Table 1** below shows the lettering used in the ETS taxonomy.

Table 1. Lettering for the ETS taxonomy

Element	Lettering
Error	A and M
Trend	N, A, Ad, M and Md
Seasonality	N, A and M

An ETS(A,A,A), for example, denotes a model with an additive error, additive trend, and additive seasonality. There are 30 models in total according to this taxonomy (2 Error x 5 Trend x 3 Seasonality), but not all of them make sense. For example, having an additive error and multiplicative seasonal components is considered an unreasonable model, because it is difficult to expect that with the increase of level, one component will increase by x% while the other will increase by y units. A fuller explanation of "reasonable models" is given by Svetunkov and colleagues (2023). Forecasters can use established packages in R, Matlab or Python to estimate parameters using maximum likelihood (more details later) and to choose appropriate components based on information criteria.

Note that this framework has focused on modeling each series individually. There have been some extensions – for example, de Silva and colleagues (2010) – but to our knowledge, no systematic attempt to extend this framework has been attempted. This article extends the methodology to a multivariate time-series framework allowing cross-sectional information to be shared.

A NEW TAXONOMY: PARAMETERS, INITIAL VALUES, COMPONENTS (PIC)

The ETS modeling framework is now well established and is commonly applied in retail forecasting. In this framework, we need to specify the model *components* – i.e., whether we think the data has trend and/or seasonality (and what form the pattern is in relation to level and to each other). We need to have smoothing *parameters*, one for each component, which regulate how quickly the components evolve over time. We also need to have *initial values* to get the recursive equations started. These three elements – parameters, initial values, and components – can all be modeled separately, for each individual series. These three elements can also be modeled in common, across series. But is this reasonable?

It might make sense to have common smoothing parameters between different time series, so that models applied to them react to noise similarly. There is even empirical evidence that supports common smoothing parameters for level and trend. For example, Fildes and colleagues (1998) reported that common smoothing parameters for level and trend outperformed individually optimised ones (for Holt's method and Damped Holt's on telecommunication series), even though the common smoothing parameters were not optimised.

While the telecommunications data set had no seasonality, Dekker and colleagues (2004) found common seasonal smoothing parameters performing better than the individual Winters method. They explained that sharing common seasonal smoothing parameters meant that the seasonal indices were estimated and updated based on the product family, which was less volatile than individual items, and it made it easier to detect the seasonal patterns. The Dekker study included 56 items from three product families.

When it comes to time-series components, they can potentially be modeled as common, although it is less intuitive to have common level and trend. However, it makes more sense to have common seasonality within a product family. Imagine a family of soft drinks with different flavours and bottle sizes, or shirts with various colours and sizes. The individual levels and trends may differ (certain flavours of the drink and certain colours of the shirt are more popular than others), but it is reasonable to assume that the underlying seasonal indices are common or very similar, so that the products share seasonal patterns. Based on this insight, Ouwehand and colleagues (2007) developed an ETS model that had common (multiplicative) seasonal components but kept levels and trends individual.

As for initial values, when it comes to seasonal component, estimating them might be challenging, because the number of initial seasonal indices to estimate can very quickly become huge within a product family (7 for day of the week, 12 for monthly and 52 for week of the year; then multiplied by the number of series in each group). Wouldn't it be nice to share common initial values across the group? With short data histories we are more likely to get better-estimated values, and the whole task becomes more manageable. And remember, they do not have to remain common; they are allowed to evolve individually. This implies that different series start with similar seasonal pattern, but then they are allowed to evolve based on the individual characteristics of each series.

The thinking behind our PIC taxonomy – Parameters, Initial values, and Components – is to afford flexibility in modeling multiple time series jointly. Sharing cross-sectional information according to our approach is simple but important in contributing towards improving forecasting performance as well as easier implementation for practitioners. **Table 2** briefly outlines the PIC taxonomy:

Table 2. Lettering for the PIC taxonomy

Element	Lettering
Parameters (Smoothing constants for Level, Trend, Seasonality and Damping parameter)	N, L, T, S, D and combinations
Initial values (Level, Trend, Seasonality)	N, L, T, S and combinations
Components (Level, Trend, Seasonality)	N, L, T, S and combinations

The P element can be N if no parameters are common, L for the common level, T for the common trend, S for the common seasonality, and D for a common damping parameter. These letters can be combined; for example, TS indicating common trend and seasonal smoothing parameters, and LTDS meaning all smoothing and damping parameters are common between time series in the group.

For the I and C elements, the options are similar to P, but without D. So PIC(N,N,N) means all parameters, initial values and components are set individually, while PIC(LTDS,LTS,LTS) means everything is common. There are many other options in between to make this a flexible approach to think about how to make effective use of cross-sectional information. The academic literature tends to focus the modeling on the two extremes of either nothing is in common or everything is, but the taxonomy broadens the options and offers practitioners more choice.

The PIC taxonomy is general in the sense that it can be applied to any type of model dealing with parameters, initial values and components. We had the ETS framework in mind when we developed the PIC taxonomy. But it can be applied to a collection of univariate models, which was the case with Ouwehand and colleagues, or to multivariate/vector models. We extended a subset of purely additive and multiplicative ETS models in our academic paper (Svetunkov and colleagues, 2023), making this a VETS-PIC framework. Readers interested in technical aspects, such as model construction and estimation, may refer to that paper.

APPLICATION ON WALMART DATA

To find out how the VETS-PIC framework performs on real data, we conducted an empirical analysis on the M5 data (Makridakis and colleagues, 2022). This dataset comes from the retail giant Walmart and contains groups of products in several categories, so it is suitable for our investigation. We first aggregated the original series to the monthly frequency and to the department store level; this avoided any intermittency in the data and gave us 70 series, each with 63 observations.

Next we needed to find out whether the series are trended and/or seasonal. To do this we used the *adam()* function from the smooth package in R (Svetunkov, 2021). This essentially selected components for the ETS model, for each of the 70 series. **Figure 2** below shows the classification based on the univariate ETS models.

The majority of the models are seasonal. Almost all series have been identified as non-trended (*N*) and only two have been identified as trended (one as AAA and one as MMM). Upon visual inspection, these two series were found to exhibit weak trends, and so we treat them as non-trended. Without trend, it does not make any difference (for an individual series) if seasonality is modeled in an additive or multiplicative way. So, we decided to use multiplicative seasonality because it is more natural to think about it as a certain percentage up or down the base level. Also, very important for common seasonality, it makes sense to model it as a common multiplicative index rather than a fixed quantity added to different products, especially when the base levels are not the same across the group. Therefore, two vector ETS (VETS) models are applied: VETS(M,N,M; LN) for seasonal series and VETS(M,N,N; LN) for non-seasonal ones. LN indicates that the models are applied on log-transformed data. With the natural product group information, **Table 3** shows the detailed breakdown of the number of groups and number of series within each group. Group 12 has only one series, which means we cannot apply a vector model, and it has been removed from the subsequent analysis.

Figure 2. ETS models selected for M5 data by the adam() function

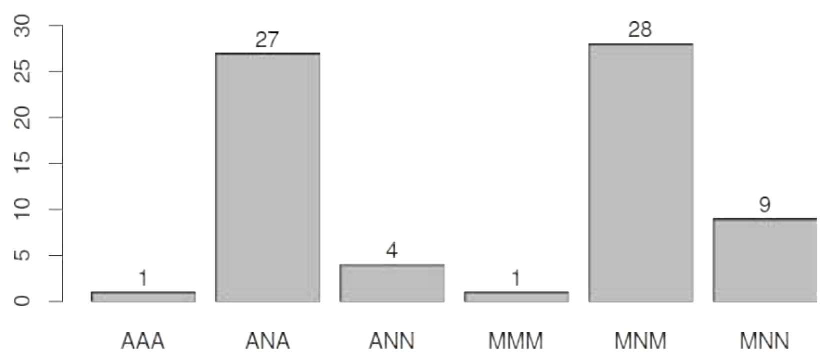


Table 3. Grouping of VETS models

Seasonal: VETS(M,N,M; LN)			Non-Seasonal: VETS(M,N,N; LN)		
Group name	Number of series	Category	Group name	Number of series	Category
G1	6	Hobbies 1	G8	4	Hobbies 1
G2	7	Hobbies 2	G9	3	Hobbies 2
G3	10	Household 1	G10	2	Foods 1
G4	10	Household 2	G11	3	Foods 2
G5	8	Foods 1	G12	1	Foods 3
G6	7	Foods 2			
G7	9	Foods 3			
	Total: 57			**Total: 12**	

With groups defined and VETS models identified, we set out to examine the empirical performance of the VETS models with optimally selected PIC elements. This is done by using the *auto.vets()* function from the *legion* package in R (more on the software package later).

We also used ETS(M,N,N) and ETS(M,N,M) as univariate benchmarks. This is implemented in *ets()* from the *forecast* package (Hyndman and Khandakar, 2008).

For all of the experiments, we have two different data histories.

- Full data: using 51 observations for model estimation and the last 12 for hold-out.
- Short data: using only the last 36 observations (24 for estimation and the same 12 observations as the hold-out sample).

Our expectation was that the models with commonalities would be more beneficial for the shorter series, because methods based on individual series would have less data to rely on.

EMPIRICAL RESULTS

Forecasting accuracy is measured as relative mean absolute error (rMAE), over the Naïve method. We have calculated other error measures (such as MASE and relative RMSE), but the results were very similar to the table shown below, so we skip them for brevity. If it is lower than one, then the method under consideration performs better than Naïve and vice versa. **Table 4** reports the medians of the rMAEs.

Table 4. Overall forecasting accuracy results

Sample size	Model	Overall	Seasonal	Non-seasonal
Long	ETS	0.957	0.924	**1.015**
	VETS	**0.934**	**0.888**	1.029
Short	ETS	1.000	1.003	0.997
	VETS	**0.974**	**0.974**	**0.974**

Overall results show that the VETS models perform strongly and are better than the univariate ETS benchmarks in almost all scenarios. When data history is short, VETS models show consistently superior performance, as expected. Focusing on seasonal data, VETS models outperform ETS probably because the joint estimation of seasonal indices can be done easier in the multivariate setting rather than in the univariate one. As we see, there is clear evidence that using cross-sectional information is beneficial, especially with limited data.

Accuracy measures, important as they are, do not tell the full story. It is also important, perhaps more so, to measure prediction intervals. In this case, we calculated coverage and range (relative value with ETS on long history used as a benchmark) values for 95% prediction intervals, shown in **Table 5**.

Table 5. Overall prediction interval results

Model	Long		Short	
	Coverage	Range	Coverage	Range
ETS	0.937	1.000	0.886	**0.918**
VETS	**0.947**	**0.984**	**0.970**	5.661

The univariate ETS models provide lower coverage than the required 95%, as shown in Table 5. This is worsened when the data history is short, leading to potentially serious and costly consequences when it comes to many forecasting applications. VETS, in comparison, performs better than univariate ETS in terms of coverage. However, this is achieved at the expense of having a wider prediction interval, especially when the data history is short. The VETS results are much better in both coverage and range, when longer data histories are available.

PRACTICAL IMPLEMENTATION OF THE VETS-PIC FRAMEWORK

The VETS-PIC framework has been implemented in the *legion* package for R (***https://github.com/config-i1/legion***). The *auto.vets()* function is particularly useful for practitioners, as the PIC restrictions are automatically selected for the VETS models. For the 11 groups of the Walmart data, **Table 6** shows how often each PIC model was selected.

Table 6. Optimal PIC selection for the Walmart data

PIC	N,N,N	L,N,N	L,L,N	LS,S,N	LS,S,S
Frequency	1	1	2	6	1

Out of the seven groups of seasonal data, PIC(LS,S,N) was selected for six and PIC(LS,S,S) for one. This clearly indicates the benefits of having common smoothing parameters and initial seasonal values. Even for the four non-seasonal groups, we can see the benefits of having common smoothing parameters for level and common initials. Only one out of the 11 groups selected PIC(N,N,N), with the other 10 groups favouring cross-sectional information.

DISCUSSIONS AND REFLECTIONS

The VETS-PIC framework discussed in this paper provides a simple yet important way of understanding how to use cross-sectional information to improve on ETS forecasting. The motivation lies in the complex needs of retail forecasting. Our empirical analysis on Walmart data shows clear benefits in both forecasting accuracy and prediction intervals.

Applying the PIC taxonomy, we observe strong evidence for common smoothing parameters and initial values, for both seasonal and non-seasonal data. The argument for common components is less strong; we found in simulations that common components can become too restrictive and therefore not robust across different data generating processes. The empirical results are consistent with findings from our simulation experiments – please refer to Svetunkov and colleagues (2023) for more detailed discussions.

The *legion* package for R allows easily applying the proposed framework not only in research but also in practice. The different functionality in the package offers various degrees of control, from the fully automatic to manual specifications. The package is maintained by one of the authors and is constantly updated to include more features and to make its functions more robust.

REFERENCES

Dekker, M., Van Donselaar, K. & Ouwehand, P. (2004). How to Use Aggregation and Combined Forecasting to Improve Seasonal Demand Forecasts, *International Journal of Production Economics*, 90 (2), 151–167.

Fildes, R., Hibon, M., Makridakis, S. & Meade, N. (1998). Generalising about Uni-variate Forecasting Methods: Further Empirical Evidence, *International Journal of Forecasting*, 14 (3), 339–358.

Hyndman, R.J., Koehler, A.B., Snyder, R.D. & Grose, S. (2002). A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods, *International Journal of Forecasting*, 18 (3), 439-454.

Hyndman, R.J., Koehler, A.B., Ord, J.K. & Snyder, R.D. (2008). *Forecasting with Exponential Smoothing*, Springer Berlin Heidelberg.

Hyndman R.J. & Khandakar, Y. (2008). Automatic Time Series Forecasting: The Forecast Package for R, *Journal of Statistical Software*, 26 (3), 1-22.

Makridakis, S., Spiliotis, E. & Assimakopoulos, V. (2022). The M5 Competition: Background, Organization, and Implementation, *International Journal of Forecasting*, 38 (4), 1325-1336.

Ouwehand, P., Hyndman, R.J., de Kok, T.G. & van Donselaar, K.H. (2007). A State Space Model for Exponential Smoothing with Group Seasonality, Monash Econometrics and Business Statistics Working Papers 07/07, Monash University, Department of Econometrics and Business Statistics, ***https://www.ideas.repec.org/ p/msh/ebswps/2007-7.html***.

de Silva, A., Hyndman, R.J. & Snyder, R. (2010). The Vector Innovations Structural Time Series Framework, *Statistical Modelling: An International Journal*, 10, 353-374.

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.

Svetunkov, I., Chen, H. & Boylan, J.E. (2023) A New Taxonomy for Vector Exponential Smoothing and its Application to Seasonal Time Series, *European Journal of Operational Research*, 304, 964-980.

Svetunkov, I. & Pritularga, K. F. (2021). Legion: Forecasting Using Multivariate Models, R package version 0.1.1, ***https://www.github.com/config-i1/legion***.

Svetunkov, I. (2021). Smooth: Forecasting Using State Space Models, R package version 3.1.1, ***https://github.com/config-i1/smooth***.

Huijing Chen is a Senior Lecturer in Operations Management at University of Portsmouth (UK). She has a BSc in economics from Shanghai Institute of Foreign Trade, MSc in management science and operational research from Warwick University, and PhD from Brunel University. Her research interests are in short-term seasonal demand forecasting.

huijing.chen@port.ac.uk

John Boylan is Professor of Business Analytics at Lancaster University, Editor-in-Chief of the *Journal of the Operational Research Society*, and President of the International Society for Inventory Research. In 2022 he was named a Fellow of the International Institute of Forecasters.

j.boylan@lancaster.ac.uk

Ivan Svetunkov is a Lecturer of Marketing Analytics at Lancaster University (UK). He has a PhD in management science from Lancaster University and a candidate degree in economics from Saint Petersburg State University of Economics and Finance. His main area of interest is developing statistical methods for forecasting. He is a creator and maintainer of several forecasting- and analytics-related R packages and an author of a monograph: "Forecasting and Analytics with ADAM" (***https://openforecast.org/adam/***).

i.svetunkov@lancaster.ac.uk

In Memoriam: John Boylan

Just as this issue was going to press we learned of the untimely passing of one of our own, John Boylan, Professor of Business Analytics at Lancaster University, and General Chair of the 2022 ISF in Oxford.

John made major contributions to the research and practice of forecasting, particular in regard to intermittent demand and the interface between forecasting and inventory management. He was enthusiastically involved with *Foresight* from the journal's very first issue, serving on the Editorial Board, as Supply Chain Forecasting Editor, and most recently on the *Foresight* Advisory Board. John was also a frequent contributor, with his 2010 paper "Choosing Levels of Aggregation for Supply Chain Forecasts" inducted into the *Foresight* Hall of Fame.

This is a devastating loss to the *Foresight* family. We will honor John with a proper tribute in the Q4 issue.

Commentary: PICS, or, Why Stop at PIC?

STEFAN DE KOK

Chen, Boylan, and Svetunkov (CBS) present a novel idea of a taxonomy for cross-learning. Personally, I found it enlightening. My own experience is deep in both statistical forecasting and data science, and I have experienced the chasm between the two firsthand every single time they coexist. While I've worked hard each time to bridge that chasm, it had never occurred to me that a new taxonomy was in order – until I read CBS.

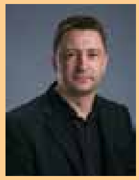
I see the introduction of this taxonomy as an opportunity to bring statistical forecasting and data science closer together: a holy grail of sorts.

CBS introduce the PIC (Parameters, Initial values, Components) taxonomy as an extension to ETS, specifically to be able to classify how, if any, cross-learning is to be applied. Beyond the introduction of the taxonomy itself, I applaud the explicit inclusion of initial values as part of it. In my experience in statistical forecasting settings, parameters and components are always given conscious consideration – even if through brute force, letting some "expert system" determine which methods to include with their intrinsic assumptions on components and with which parameter values. Initial values, however, are rarely given much thought. Making them explicit in the taxonomy will be instrumental in making their consideration part of any curriculum on the topic of forecasting.

This brings me to data science. Most data scientists I encounter are at most a few years out of college and generally lack an understanding of the fundamentals of statistical forecasting. They love to make big, all-inclusive models, and have been taught to consider hyperparameter tuning. Surprisingly, the set of hyperparameters considered rarely includes the forecast parameters themselves, beyond textbook values. Initial value setting is never considered, and the magnitude of the impact this oversight may have on the quality of their results goes unrecognized. The PIC taxonomy, if taught to would-be data scientists, will force the new waves of practitioners to be considerate of these aspects of time series forecasting. I see the introduction of this taxonomy as an opportunity to bring statistical forecasting and data science closer together: a holy grail of sorts.

The taxonomy, however, is missing one dimension that is critical to both data science and cross-learning in general: how to segment the data. Trends may be highly correlated based on the brand of the product, while seasonality may be correlated to both product type and geographical latitude, and promotions to the specific account and country. Each of these is best served with cross-learning from different segmentations. I find this dimension is often given cursory attention by junior data scientists; yet, in my experience, segmentation generally is a dominant factor on accuracy of hierarchical and other cross-learning-type forecasts. To this end, I propose to extend the PIC taxonomy and make it a PICS taxonomy, with the added S for Segment.

The C in PIC already allows extension to other types of components. In data science, the broader equivalent to components would be called drivers as used in driver-based forecasting, with each driver represented by one or more features. Promotions, moving holidays, new-product launches, and any other kind of driver could be seen as additional "components" similar to error, trend, and seasonality. Each of these would have its own parameters and initial values as described by the PIC taxonomy. By further adding segment to the taxonomy, there could finally be a standardized bridge between statistical forecasting and data science, serving both equally well.



Stefan de Kok, CEO and co-founder of Wahupa, has a long career in supply chain planning, with broad experience across many roles and industries. He has been a trailblazer in educating the industry on the need for probabilistic methods to deal with uncertainty. He is passionate about applying probabilistic algorithms to demand planning, inventory optimization, manufacturing, and integrated business planning solutions. Stefan earned an MSc in applied mathematics from Delft University of Technology, is a regular speaker and blogger, and is currently writing his first textbook on probabilistic planning and forecasting.

stefan.dekok@wahupa.com

6 & 7 NOVEMBER 2023 | Harvard Club of New York City

AGENDA: MOFC.UNIC.AC.CY/AGENDA/
REGISTER HERE: MOFC.UNIC.AC.CY/EVENT/THE-FUTURE-OF-FORECASTING-THE-M6-COMPETITION-2/

ORGANIZERS:

UNIVERSITY of NICOSIA   MOFC Learn, Forecast, Compete, Disseminate, Excel.   IIF International Institute of Forecasters

Commentary: Exponential Smoothing in the Spotlight Again

MALVINA MARCHESE

In a world where forecasting seems dominated by big data and machine learning, Chen, Boylan, and Svetunkov (CBS) remind us of a very important truth: many of us work with data sets with a short time series. The authors address this in the context of retail forecasting, where data displays strong seasonality, but similar kinds of data are encountered in various other fields.

The article offers yet another example of the incredible resilience and flexibility of exponential smoothing (ES). When Gardner (1985) published his article on ES's state of the art, many believed that ES should be disregarded because it was either a special case of ARIMA modeling or an ad hoc procedure with no statistical rationale. Today, we know that ES methods are optimal for a very general class of state-space models that is broader than the ARIMA class (Hyndman and colleagues, 2002).

CBS suggest that ES can actually do very nicely in the era of machine learning as well. This is achieved by applying ES with cross-learning – using many series to train a single model, which is unlike standard statistical time series algorithms where a separate model is developed for each series. Cross-learning seems to be a successful attempt to fight the Curse of Dimensionality when applied to Vector ETS models. Multivariate time series models make full use of the dependency structure between different series and as such can greatly overperform univariate models in forecasting when the data set is large enough. But when this is not the case, we need to find smart ways to reduce the number of parameters. CBS address the problem by proposing a taxonomy through which we can choose the level of commonality in a vector ETS model: Parameters (Smoothing constants for Level, Trend, Seasonality, and Damping parameter), Initial values (Level, Trend, Seasonality), and Components (Level, Trend, Seasonality). This is a way to impose restrictions directly on the time series components of the data-generating process.

Today, we know that ES methods are optimal for a very general class of state-space models that is broader than the ARIMA class.

The approach seems quite robust to misspecification and hopefully it will prove to be successful in forecasting practice. As usual, its effectiveness will depend on how commonalities in seasonal demand patterns are identified and imposed. Svetunkov and colleagues (2023) propose an algorithm to this end. As we can expect, their results show that the most restrictive methods perform badly when applied to a flexible multivariate DGP (where for each time series the parameters are different and can cause significant damage in terms of accuracy). However, the good news is that this degradation comes mainly from the common seasonal component element. When this assumption is relaxed to individual seasonal components, even with the restrictions of smoothing parameters and initial seasonal indices, the methods tend to perform better.

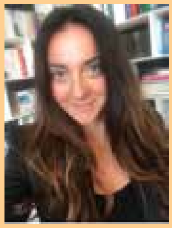
In conclusion, exponential smoothing is in the spotlight again and stronger than ever in its new cross-learning-enhanced version.

REFERENCES

Hyndman, R.J., Koehler, A.B., Snyder, R.D. & Grose, S. (2002). A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods, *International Journal of Forecasting*, 18 (3), 439-454.

Gardner Jr., E.S. (1985). Exponential Smoothing: The State of the Art, *Journal of Forecasting*, 4 (1), 1-28.

Svetunkov, I., Chen, H. & Boylan, J.E. (2023). A New Taxonomy for Vector Exponential Smoothing and its Application to Seasonal Time Series, *European Journal of Operational Research*, 304 (3), 964-980.



Malvina Marchese is a Senior Lecturer in Finance and the Academic Program Director of Finance Degrees at Bayes (formerly Cass) Business School. She is also Invited Associate Professor at the Norwegian University of Science and Technology, Financial Forecasting Consultant with Maersk Brokers Advisory and CBRE Investment Management, and was previously Head of Risk Management at Shell Oil. Malvina holds a PhD in econometrics from the London School of Economics and Political Science. Her research focuses on high-dimensional models for volatility and correlations forecasting with applications to commodities, and on fractional integration in volatility modeling.

Malvina.Marchese@city.ac.uk