

Multiple temporal aggregation: the story so far. Part IV: Temporal Hierarchies

By Nikos | July 22, 2017

0 Comment

Temporal Hierarchies

In the previous [post](#) we saw how the Multiple Aggregation Prediction Algorithm (MAPA) implements the ideas of MTA. We also saw that it has some limitations, particularly requiring splitting forecasts into subcomponents (level, trend and seasonality). Although some forecasting methods provide such outputs naturally, for example Exponential Smoothing and Theta, others do not. More crucially, manually adjusted forecasts do not either, and even though it is possible to use MAPAx for that, a simpler approach would be welcome. This is where Temporal Hierarchies become quite useful, which is an alternative way to implement MTA.

Temporal Hierarchies borrow many ideas from cross-section hierarchies and organise the different temporal aggregation levels as a hierarchy. Consider for example four quarterly observations. The first two quarters constitute the first half-year, and the last two quarters constitute the second half-year. The two half-years add up to make a complete year. These connections imply a hierarchy, much like sales of different packet sizes of a product in a supermarket can be organised in a product hierarchy. However, temporal hierarchies have one key advantage over cross-sectional ones, they are uniquely specified by the problem at hand. Suppose I am given monthly data to forecast. There is a single hierarchy across temporal aggregation levels, much like in the quarterly example before, that I need to deal with, irrespective of the item I need to forecast, the way I got the forecast or the properties of the time series. Once this unique hierarchy is defined (and all the data are coming from temporally aggregate views of the original time series), then all that is left is to do is to forecast across the hierarchy, i.e., all temporal aggregation levels and reconcile the forecasts. The act of reconciliation brings together information from all modelling levels, with the MTA benefits discussed in the [previous posts](#).

Some hierarchies are more complex than others. The quarterly hierarchy, from the example above, is a very simple three level hierarchy (quarters, half-years, years). A monthly hierarchy is more complex, because there are more than one ways to reach to yearly data from monthly. For example, one could aggregate by 2 months, then these by 2 (4-monthly level), and then that by 3 (yearly level). Alternatively, one could aggregate to quarterly data, half-yearly and then yearly. The two aggregation paths can happen in parallel. The temporal hierarchy is made up by all possible paths. Note that in contrast to [MAPA](#), levels that do not fully add up to a yearly time series are excluded (intuitively they do not belong in any path from the bottom disaggregate level to the top yearly level). This has the

advantage that any forecasting model/method does not need to deal with series that may have fractional seasonality. Nonetheless, this is an interesting future research avenue.

The following interactive plot provides the temporal hierarchies for common types of time series. Observe that many have multiple pathways to the top yearly level (for example, monthly time series), and some are very simple hierarchies (for example, days in week). Use the highlight option to easily visualise the various pathways. Once visualised, the analogies with cross-sectional hierarchies are apparent.

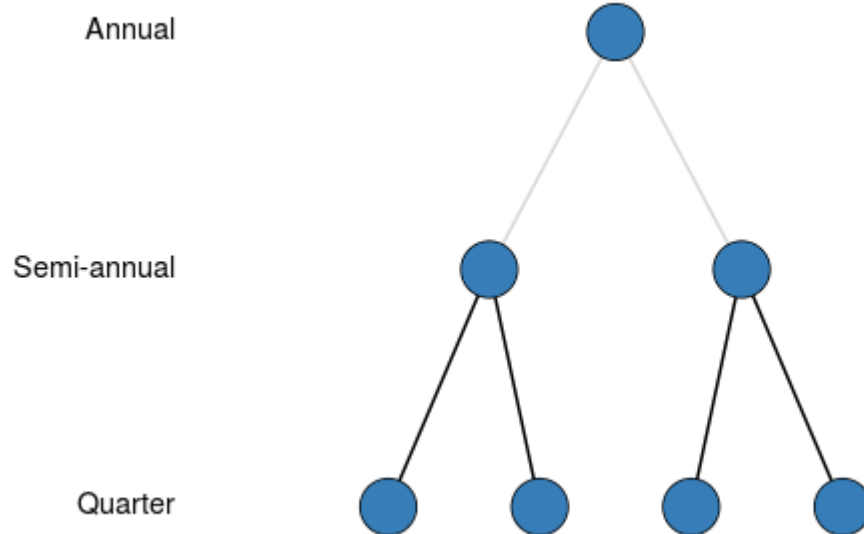
Temporal Hierarchies

Series periodicity

(4) Quarterly ▼

Highlight level

None ▼



Nikolaos Kourentzes (<http://nikolaos.kourentzes.com/>) , 2017

To forecast we need to populate every level of the hierarchy with a forecast. So for example, for the quarterly hierarchy we need to provide 3 sets of forecasts, one for the quarterly time series, one for the semi-yearly and one for the yearly. Imagine that each hierarchy depicts one year's worth of forecasts, but obviously we can produce the same hierarchy for the next year and so on. Mathematically this is just another column of forecasts to be handled by the hierarchy, so in fact it is trivial to do. But an implication is that forecasts are produced in horizons that are multiples of full years (and then any shorter horizons are used accordingly). People are more familiar with two specific cases of temporal hierarchies. One is when we need to produce a total figure over a period, for example for tactical/strategic forecasts. This is simply the bottom-up interpretation of temporal hierarchies: forecasts from the lowest level are summed to a higher level. The other alternative is to produce a forecast and then use a 'profile' to split this further. In supply chain forecasting and call centres this is very common, in breaking weekly forecasts into daily profiles, or daily forecasts into intra-daily profiles. This is merely the top-down interpretation of temporal hierarchies.

Forecasting with Temporal Hierarchies

You may have already noticed that there is nothing to restrict the source of forecasts. They can be based on some statistical model, judgement, mix of both, differ amongst levels, or whatever other exotic source. This is a substantial advantage over MAPA, and temporal hierarchies provide a flexible MTA foundation. In reconciling the forecast there are couple of complications that we deal with in this [paper](#) (the scale and variance of the forecasts are different, which needs to be taken into account during reconciliation). I mentioned earlier that temporal hierarchies are unique. This simplifies substantially the solution, but I will not go into the mathematical details here.

In the following interactive plot you can choose from the usual time series I have been using as examples in this series of posts to produce base (conventional built forecast from a single level, in red) and Temporal Hierarchy Forecasts (THieF, in blue). I provide the forecasts across the various temporal aggregation levels permitted by the hierarchy. Observe how the information across the temporal aggregation levels is shared in the THieF forecasts to achieve better modelling of the series. You can also choose between three different forecasts: exponential smoothing, ARIMA and naive. The naive forecasts are quite illuminating in showing how the multiple views offered by THieF achieve superior results. There other two types of forecasts are quite illustrative as well.

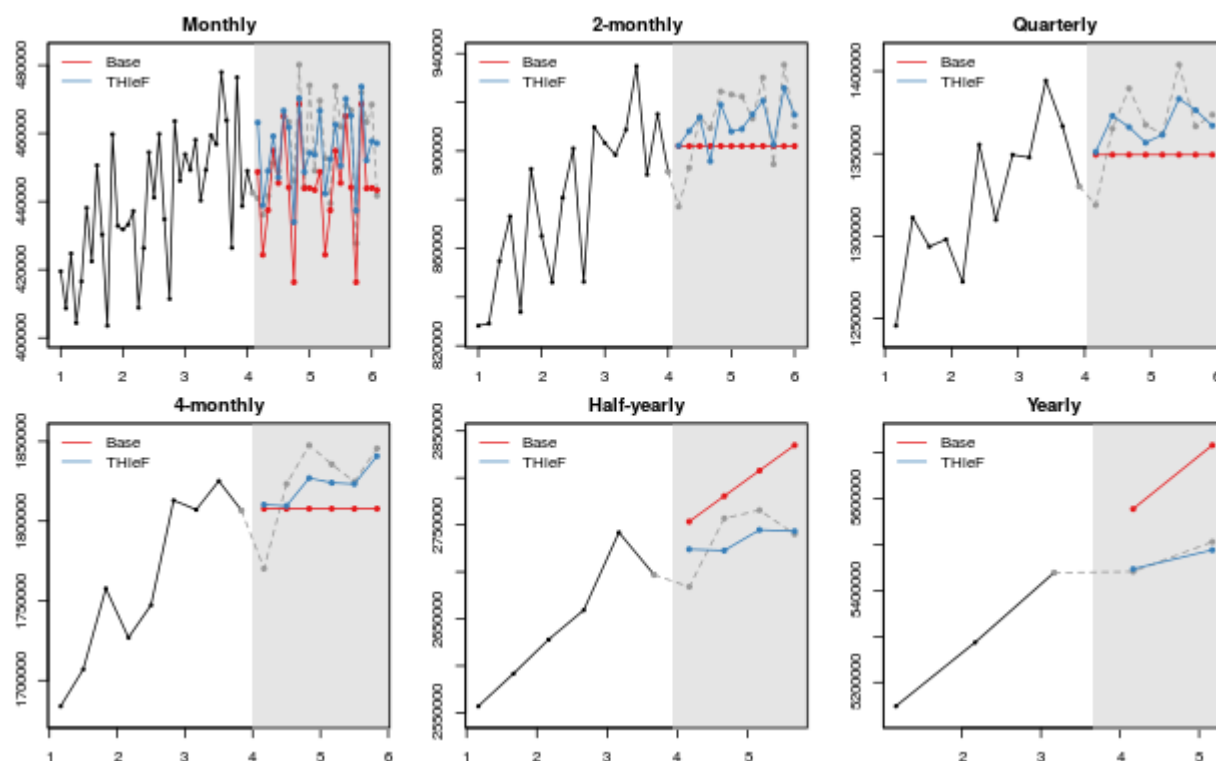
Temporal Hierarchies Forecasting (THieF)

Select Series

A&E Admissions

Forecasting method

Exponential Smoothing



Mean Absolute Error (MAE) on monthly hold-out data

Base ETS: 12508.98

THieF ETS: 8228.9

I also provide Mean Absolute Error (MAE) for the base and THieF forecasts for the disaggregate series. You will observe that on average THieF forecasts are more accurate. The gains improve at more aggregate levels. In the [paper](#) we demonstrate with simulations that in various scenarios of uncertainty (parameter, model) THieF performs better or at least as good as base forecasts.

To sum up, forecasting with temporal hierarchies:

- offers a very flexible framework to implement MTA, with all its advantages;
- is independent of source of forecasts, allowing to provide different additional information at different levels, if available;
- has been shown to offer substantial gains in terms of accuracy over base forecasts, by blending the information available across temporal aggregation levels;
- provides reconciled short term (disaggregate) and long-term (aggregate) forecasts, leading to aligned operational, tactical and strategic planning.

If you want to try it out we have released the [thief](#) package for R.

A final note on THieF. THieF and MAPA both perform very well and neither is a clear winner in terms of forecast accuracy alone. The two MTA alternatives handle information in a different way. MAPA also takes advantage of the 'in-between' levels that THieF excludes. The good performance of both, even though they have some key differences, is exciting: it gives further merit to MTA and offers some clear directions for future work!

Multiple Temporal Aggregation: the story so far: [Part I](#); [Part II](#); [Part III](#); Part IV.

Related Posts

- MAPAx example for R
- New R package nnfor: time series forecasting with neural networks
- Principles of Business Forecasting 2e

Category: Blog Tags: hierarchical forecasting , how to? , MAPA , temporal aggregation , temporal hierarchies