Automatic Forecast Model Selection in SAP Business Information Warehouse under Noise Conditions

Introduction

Companies and markets are complex interacting systems. Despite inherent uncertainty it is a necessity for companies to plan ahead. Thus, forecasting based on IT-tools is today an integral part of management. Typical applications include sales and revenue forecasting as inputs for supply chain management and budgeting, respectively. One of the most important vendors for Enterprise Resource Planning (ERP) software in Europe and elsewhere in the world is the SAP AG with its IT-product portfolio centered around the R/3™ system.

Various SAP tools offer forecasting methods as part of their business functionality. Examples include SAP R/3™, SAP Advanced Planner & Optimizer (APO™), and SAP Business Information Warehouse (BW™). The forecasting component examined here is the automatic forecast model selection in SAP BW Business Planning & Simulation (BPS). The automatic model selection fits a forecast model to the available historical data while minimizing some error measure, which is Akaike’s Information Criterion in the case of SAP BW [1, 2].

It is a software feature routinely used by many companies in their forecasting, for instance in sales planning. At first sight, this feature seems to relieve the planner from choosing an adequate statistical forecast model for historical data himself, which can be a time-consuming and difficult task.

Because it is so often applied, one should aim for a good understanding of the strengths and weaknesses of automatic model selection in order to make a well-informed decision when to use it. The experiments outlined briefly in this paper are a contribution to this goal.
**Method**

Three basic functions were used in our forecasting experiments performed with SAP BW BPS release 3.5. The empirical tests are based on time series data showing trend, seasonal and trend-seasonal patterns.¹ The total horizon were 96 month of which 48 were treated as historical data and the other 48 were the planning horizon. Figure 1 gives an illustration of the time series patterns that were used.

---

¹ For a more detailed account of the results see [3].
The data was initially used in an ideal deterministic pattern which should provide a rather trivial task for automatic forecast model selection. Thereafter, varying amounts of normally-distributed noise are added to the historical data in order to see how this affects the quality of model selection and the forecasting result. The additive noise component mimics real-life data quality which can be characterized as a stochastic influence on trend, seasonal or other patterns. Practically, each historical data point \( x \) of the basic functions was modified in the following way:

\[
x' = x + a \cdot N(0,1),
\]

where \( N(0,1) \) represents a standard-normally distributed random number and \( a \) is a noise factor that was varied according to \( a = [0 ; 1 ; 2 ; 5] \). Thus, we conducted four sets of forecasting experiments for all functions. To remove arbitrary results, at each noise level and for each function 30 runs with different random number seeds for the time series modification were performed.

Results

Due to space limitations, results are presented here in the form of tables for each function and noise-level. The tables report on data averaged over all 30 runs for each experiment. The mean squared error (MSE) and the mean absolute percentage error (MAPE) are employed as standard forecast error measures. Furthermore, the percentage of correct models chosen by the automatic forecast model selection is given in the tables.

Set 1: Forecasting without noise (noise factor \( a = 0 \))

<table>
<thead>
<tr>
<th>Function</th>
<th>MSE</th>
<th>MAPE (%)</th>
<th>correct model (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) trend</td>
<td>0,0</td>
<td>0,0</td>
<td>100,0</td>
</tr>
<tr>
<td>(2) seasonal</td>
<td>0,0</td>
<td>0,0</td>
<td>100,0</td>
</tr>
<tr>
<td>(3) trend-seasonal</td>
<td>0,0</td>
<td>0,0</td>
<td>100,0</td>
</tr>
</tbody>
</table>

Table 1: Forecast results with original historical data (noise level = 0), avg. from 30 runs

Without noise added, automatic model selection chose the correct model in every case.

Set 2: Forecasting with noise factor \( a = 1 \)

<table>
<thead>
<tr>
<th>Function</th>
<th>MSE (min, max)</th>
<th>MAPE (%) (min, max)</th>
<th>correct model (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) trend</td>
<td>3,1 (0,0 ; 24,0)</td>
<td>8,5 (0,5 ; 32,5)</td>
<td>80</td>
</tr>
<tr>
<td>(2) seasonal</td>
<td>0,4 (0,1 ; 2,3)</td>
<td>19,2 (8,9 ; 63,6)</td>
<td>90</td>
</tr>
<tr>
<td>(3) trend-seasonal</td>
<td>0,7 (0,2 ; 1,3)</td>
<td>5,7 (2,8 ; 12,6)</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 2: Forecast results with modified data (noise level = 1), avg. from 30 runs
For the trend function, the system chose 24 times the trend model, 5 times a trend-seasonal and once a constant model. For the seasonal function, the system chose a seasonal model 27 times and a trend-seasonal model 3 times. For the trend-seasonal function, the trend-seasonal model was chosen in each case.

Set 3: Forecasting with noise factor $a = 2$

<table>
<thead>
<tr>
<th>Function</th>
<th>MSE (min, max)</th>
<th>MAPE (%) (min, max)</th>
<th>correct model (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) trend</td>
<td>19,3 (0.0 ; 124,3)</td>
<td>25,2 (1.5 ; 73,8)</td>
<td>40</td>
</tr>
<tr>
<td>(2) seasonal</td>
<td>1,1 (0.5 ; 2.0)</td>
<td>33,4 (19,2 ; 49,9)</td>
<td>100</td>
</tr>
<tr>
<td>(3) trend-seasonal</td>
<td>5,5 (0.7 ; 31.9)</td>
<td>14,4 (5.5 ; 42.0)</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 3: Forecast results with modified data (noise level = 2), avg. from 30 runs

For the trend function, the system chose 12 times a trend, 4 times a trend-seasonal and 14 times a constant model. For the seasonal function, the system chose a seasonal model in all cases. For the trend-seasonal function, the trend-seasonal model was chosen 26 times (with 24 times the correct additive model and two times a multiplicative model), while a seasonal and a constant model where both chosen in two cases each.

Set 4: Forecasting with noise factor $a = 5$

<table>
<thead>
<tr>
<th>Function</th>
<th>MSE (min, max)</th>
<th>MAPE (%) (min, max)</th>
<th>correct model (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) trend</td>
<td>44,0 (2,5 ; 168,6)</td>
<td>44,0 (12,6 ; 90,3)</td>
<td>20</td>
</tr>
<tr>
<td>(2) seasonal</td>
<td>5,1 (2.9 ; 15.4)</td>
<td>78,2 (55,5 ; 142.4)</td>
<td>27</td>
</tr>
<tr>
<td>(3) trend-seasonal</td>
<td>26,3 (2,5 ; 86,3)</td>
<td>35,5 (11,9 ; 76,5)</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 4: Forecast results with modified data (noise level = 5), avg. from 30 runs

For the trend function, the system chose 6 times a trend, once a trend-seasonal, once a seasonal and 22 times a constant model. For the seasonal function, the system chose a seasonal model 8 times and a constant model 22 times. For the trend-seasonal function, the trend-seasonal model was chosen in two cases, a seasonal model was chosen 4 times, a trend model 7 times, and a constant model 17 times.

**Discussion**

The performance of the automatic model selection is optimal when no noise is added to the underlying patterns. However, this situation never occurs in practical business forecasting as stochastic influence in historical data is abundant. On a general level, the correctness of the automatic model selection with respect to the underlying function pattern and the forecast quality deteriorates with the amount of noise – a result that can be expected. A closer look at the different noise levels and functions, though, reveals
some interesting details.
When the underlying pattern is a trend, the relative performance appears to be the worst. Even with the low noise factor $a = 1$, in 20% of the test cases the wrong model was chosen by the automatic model selection. Moreover, the error measures MSE and MAPE each display results within a wide interval. The maximum MAPE was as high as 32.5% even at this comparatively low level of noise. This may be interpreted as a significant risk for a bad automatic forecast. The situation is aggravated at higher noise levels.

To a somewhat lesser degree the same applies to the trend-seasonal function, where we have a wide range of MAPE- and MSE-values at the noise level $a = 2$ and higher. Here, two MAPE-results higher than 40% are found, and in three out of 30 cases the MSE is higher than 25, thus again indicating a performance risk for a purely technical forecasting approach.

When analyzing the performance for the seasonal function, one should bare in mind that the MAPE-value depends on the underlying range of correct function values. As these values on average are smaller for the seasonal than for the trend and the trend-seasonal function, a high MAPE should not be overinterpreted in the seasonal case. The data indicates that the MAPE-error is particularly high at lower turning points of the curve. For the seasonal function, the MSE gives a more reliable account of forecast quality. Here, the performance appears reasonable up to a noise factor of $a = 2$.

However, it is worth mentioning that the performance in the seasonal (and the trend-seasonal) case very much depends on the customized seasonal length in the SAP BW forecast profile. We conducted some experiments with slightly wrong seasonal length factors (not documented here for space reasons) that led to very bad results. Even without noise the automatic model selection then generally does not find the correct forecast model. The necessity to enter the correct seasonal length in advance of the forecast limits the practical value of any forecasting tool. This is not an SAP-specific problem, though. One should also consider that in business practice the length of a season may vary over time, once more demonstrating the necessity to closely monitor and supervise automated planning.

For all patterns, the automatic model selection mostly fails at the noise factor $a = 5$, both in terms of the model chosen and the forecast quality achieved. Frequently, the automatic model selection chooses the constant model as no particular pattern can be identified by the system.
Large errors at all noise levels often occur when the automatic model selection chooses a model that does not correspond to the true underlying patterns in the historical data. When classifying this choice as “wrong” prior knowledge of the original function is required that the system does not have. From a purely statistical point of view, based on the available historical data in the individual run, the system’s choice is therefore perfectly justified. This demonstrates the limits of a technical approach to forecasting when really business knowledge is necessary to help deciding the right forecast model, tune it’s parameters or manually adapt the results.

Many companies routinely use automatic forecast model selection, for instance in sales planning, where the number of planning objects is very high. The argument is often that manual planning requires a high effort. A lack of statistical knowledge is another important reason why planners turn to automatic model selection instead of analyzing the data themselves. Our experiment reminds us that a blind trust in the results of a tool can lead to suboptimal business performance.

**Conclusion**

The results presented here may be interpreted as a warning for practical planners not to automate where their individual business knowledge can help to improve the forecast. Automatic model selection is useful where the variance of historical data around clear statistical patterns is relatively small and the importance of the forecast results is not exceptionally high. In all other cases, the planner’s knowledge and skills are important input in forecasting. This suggests that some company planners should reconsider their rather unreflective attitude towards automatic model selection in practical forecasting.

**References:**


**Author:**

Prof. Dr. Volker Nissen
TU Ilmenau, FG Wirtschaftsinformatik für Dienstleistungen,
PF 10 05 65, D-98684 Ilmenau
Phone: +49 (0)3677 69-4047
Fax: +49 (0)3677 69-4219
E-mail: volker.nissen@tu-ilmenau.de