DECOMPOSITION AND FORECASTING TIME SERIES IN BUSINESS ECONOMY USING PROPHET FORECASTING MODEL

Navratil, M., Kolkova, A.

Miroslav Navratil, Andrea Kolkova / VSB Technical University of Ostrava, Faculty of Economics, Department of Business Administration, Sokolska trida 33, Ostrava, Czech Republic. Email: miroslav.navratil.st1@vsb.cz; andrea.kolkova@vsb.cz.

Abstract

There are many methods of forecasting, often based on the specific conditions of the given time series which are frequently the result of research in scientific centres and universities. Nevertheless, there are also models that were created by scientists in a particular company, examples may be Google or Facebook. The latter one has developed one of the latest Prophet forecasting models published in 2017 by Taylor & Letham. This model is completely new and so it is appropriate to subject it to further research, which is the topic of this article. To accomplish this research objective, the aim of this work is to identify seasonal trends in revenue development in a selected e-commerce segment based on the assessment of the applicability of the Facebook Prophet forecasting tool. To accomplish this goal, the Python Prophet is decomposed with a subsequent two-year forecast. Accuracy of this model is measured by RMSA and coverage. The e-commerce subject selected is active primarily in the field of sales of professional outdoor supplies and organizing outdoor educational courses, seminars and competitions. It is clear from the prediction that the e-commerce entity shows a strong sales period with the beginning of the spring season and then, due to the summer, decline, until the pre-Christmas period. The subject has little growth potential and a new impetus is needed to increase sales and thus restore the growth trend. It has been confirmed that Prophet is a suitable tool for debugging seasonal tendencies.

Keywords: forecasting, prophet, decomposition, RMSE, coverage **JEL Classification**: M21, C53

Introduction

A proper understanding of time series is a key problem for all businesses. This is evident in the fact that enterprise variables forecasting is one of the core areas of enterprise research with many managers attaching great importance to it. There are a number of methods and models that may be used for predicting and often with quite accurate results. Many of these models, however, are driven by the knowledge and skills of an analyst or a business manager and are subsequently less oriented in a number of formulas and conditions. Nevertheless, the use of computer technology is a necessity today. There have been models and methods created in both an academic environment (e.g. all models by Hyndman, Athanasopoulos et al., 2015) and by a particular company (e.g. Google or Facebook) that have attempted to make greater use of computer technology. The latter has developed one of the latest Prophet forecasting models published in 2017 by Taylor & Letham (2018).

The aim of this paper is to reveal seasonal trends in the revenue development of the selected e-commerce segment based on the assessment of the applicability of the Facebook Prophet forecasting tool. This model is an effort to bridge the considerable difficulties of most other models that require high predictive managerial training or experience. In addition, it also tries to bridge the problem present in other models of having to have a single metric for many times (e.g. for each country, because the model is only suitable for specific conditions). It is ultimately one of very few comprehensive models that address the entire forecasting process (from outliers' solution through decomposition, forecast to quantification accuracy).

Large companies worldwide use forecasting to improve their business performance, reduce costs, etc. With the increasing digitalization and new forms of business, such as a shared economy, the role of forecasting is not decreasing, but rather increasing in its importance and demand for it. Forecasting is no longer just a prognosis of basic macroeconomic variables or demand, but is used in a number of other business processes. With the increasing complexity and novelty of business opportunities, the use of forecasting is also growing. An example is a shared economy and Uber, which uses forecasting in three different areas:

- marketplace forecasting predicting both supply and demand in order to direct drivers into high demand areas that will be arising shortly,
- real-time anomaly detection stack tracking various time series both in term of back-end as well as consumer facing time series at scale,
- hardware capacity forecasting predicting the required computing power of highdemand events in order to save costs.

The outcome of such forecasting is then the values from a wide range of variables. Figure 1 gives this outcome for the aforementioned Uber example. Thus, forecasting ceases to be the only tool to predict demand, but also a number of other business factors.

1 Literature review

The breakdown of forecasting methods is, as has been declared, many. As Kolková (2018) states, the most widely used division divides the methods into qualitative and quantitative. Miller & Swinehart (2010) categorized methods into exploratory or normative methods, expert-based methods, evidence-based or assumptions. Moro & Cortez (2015) classified methods for quantitative, semi-quantitative and qualitative methods. Armstrong & Green (2014) divided forecasting methods between simple and complex with a range of other sub-division, as presented in Figure 2.

Figure 1 | Forecasting use cases at Uber



Source: Bell (2018)





Source: Armstrong & Green (2014)

There have been many years of efforts to find the best forecasting model, culminating in such events as time series forecasting competitions by magazines like International Journal of Forecasting and Journal of Forecasting. The oldest forecasting competition was in the 1970s and its conditions are defined by Reid (1969). Due to the communication possibilities of that time, they did not include many participants and not too big forecasts. By 1979,

Makridakis & Hibon (1979) had analysed 111 time series and many forecasting methods, resulting in a rather large discussion in the scientific world. In response to these not always consistent discussions, Makridakis & Hibon (1979) created a competition involving 1001 time series. More researchers joined this competition and it is often referred to as the M-Competition. The main outcome of it is the finding that statistically more sophisticated methods do not produce a more accurate prognosis than simpler ones. Another outcome is that the values of the degree of accuracy of the different methods differ according to what accuracy measure is used, of which has also been investigated by other, more recent studies (Kolková, 2018). The final two outcomes from the M-competition were that (1) the combination of different methods outperforms the use of methods on average and (2) the performance of each method depends on the prognosis horizon.

In 1993, another M-Competition was carried out under the M2-Competition (Makridakis et al., 1993), which consisted of only 29 series, but with much broader information. In 1998 Makridakis & Hibon created another competition, the M3-Competition, according to their final attempt. In this competition, 3003 time series from various areas such as business, demography, finance and economics were applied and 24 researchers were involved, with the Theta model being the winner. The results were published (Makridakis & Hibon, 2000) and confirmed that M3 is consistent with the results of previous competitions, with the exception of the hypothesis that simple methods overcome more complex methods due to the complex Box-Jenkins method and ARIMA models being more accurate than others. Makridakis created another competition, the M4-Competition, in 2018 which included 100.000 time series and new demands being placed on participants who, for example, had to publish their code on GitHub. The results of this competition are already published (Makridakis et al., 2018) and the winning method is a combination of neural networks and exponential smoothing from Uber. These results have further reinforced the discussion of what methods are best for forecasting. For example, data scientist Petr Simeček (Simeček, 2019), who worked with Google to develop the TS models using deep neural networks for multivariate time series forecasting, mentioned at the latest PyCon conference in 2019 that statistical methods are appropriate if we have a few time series with a very long history. In contrast, forecasting methods based on deep learning are useful when we have a lot of short-term time series (datasets). The problem is that we often do not know if the analysed time series are long enough for the right decision on which method to use. For this reason too the latest years of M-competitions (or Kaggle competitions) are regularly focus on hybrid forecasting systems based on both statistical methods and methods using machine learning or neural networks.

Nevertheless, Makridakis was not the only organizer of such competitions. Other examples include the Santa Fe Competitions, Neural Network competition, Kaggle time series competitions and Global energy forecasting competition. However, many questions are still ambiguous, and the future M5 and other researches may answer them.

At present, forecasting has significantly simplified modern statistical programs. Basic forecasting methods allow even simpler statistical programs like SPSS or Excel. Here, however, we come across limitations on possible calculation methods. These focus only on simple statistical models that are already woven to complement and combine with others. If the forecasting model does not work as planned, then our goal is to fine-tune the model's parameters to a particular problem. Tuning these methods requires a thorough

understanding of the underlying time series models. A typical analyst without broad programming knowledge will not know how to modify the commands and change the program. The solution may be to use R programs, Python, etc. with appropriate packages.

There are several packages available to support forecasting in modern statistical programs like R or Python. One of them is the well-known forecast from Hyndman & Khandakar (2008). This includes methods and tools for displaying and analysing univariate time series including exponential smoothing through state space models and automatic ARIMA modelling. Others can be *bfast*, an automatic breakdown for additive seasons and trends, designed for use with remote sensing data (Verbesselt, Zeileis & Hyndman, 2014). The hts package, which contains methods for visualizing, analysing and forecasting hierarchical time series, and its interpretation is contained in Hyndman et al. (2015). Another may be the MEFM, or the Monash Electricity Forecasting Model and the set of tools to implement it, based on Hyndman and Fan (2010). The robets package implements the robust exponential smoothing forecasting series and is based on Hyndman and Khandakar (2008) and is further described in Crevits & Croux (2017). Using stR you can calculate the time series of decomposition, described in Hyndman and Khandakar (2008). The thief is temporal hierarchical forecasting, described in Temporary hierarchies, co-authored with Athanasopoulos et al. (2017). tsfeatures allows time series feature extraction, according to (Hyndman, Wang & Laptev, 2015).

The procedure for using forecasting in different models is summarized by Tassone & Rohani (2017), as shown in Figure 3. The forecast procedure usually begins with the acquisition and collection of data. This phase can be very demanding and time-consuming, but the use of big data can help in alleviating some of these demands. Another downstream phase is cleaning when the data is cleaned up, edited for outliers, defined and resolved for missing values. Another part is the decomposition of acquired data, which can be the focus and main task of forecasting for a given company. Usually, the data breaks down into a major trend, a seasonal component, and a random component. However, this is often insufficient for businesses. Other elements, such as days of the week, holidays, public holidays, major events, and so on, affect business processes, so the classic basic decomposition is often inadequate and needs to be supplemented with holiday adjustment or aggregation to weekly as defined in Figure 4 and what the Prophet can do. The next stage consists in the application of selected forecasting models and their subsequent analysis. The result is daily or weekly forecasts.

The analyst's role is still important in forecasting and decomposition, but many roles can be automated. Only this automation will allow the use of the model on a larger scale, such as for managers and entrepreneurs without broad statistical and programming knowledge. For business or SME models, they need to meet some simplicity and ease of interpretation. Taylor & Letham (2018) defined the best use of human and automated tasks that they applied in business forecasting at scale. This schematic view is in the Figure 4.





Source: Tassone & Rohani (2017)





Source: Taylor & Letham (2018)

Taylor & Letham (2018) began their modeling with an analyst who defines a sufficiently flexible assignment with a subjective interpretation of parameters. Thereafter, the process is automated by which prognosis is simulated and the accuracy of these forecasts is quantified. If there is a situation where accuracy does not reach the required values, or there are problems with other aspects of forecasts, these issues are identified and passed to a human analyst, usually with priority order. The analyst can then visuall inspects the model and adjust it accordingly.

3 Methodology and data description

In 2018, Taylor & Letham (2018) from Facebook developed a Prophet forecasting model. Subsequently, they published this model in the same package in R, Python. This allowed the testing of the model by researchers from other research institutes and universities. Prophet is currently the latest published forecasting model.

3.1 Methodology of Prophet forecasting model

Prophet is an open source library that is based on decomposable models. It provides us with the ability to make time series predictions with good accuracy using simple parameters. Also important is the fact that it has support for including the impact of custom seasonality and holidays. Prophet works with decomposable time series with three main components: trend, seasonality, and holidays.

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t, \text{ where}$$
(1)

g(t) is the trend function, s(t) is periodic change, h(t) is the effects of holidays and ϵ_t is error for any unusual changes not accommodated by the model.

Prophet is applied in this article to the Python scikits-learn (*SKlearn*) concept. This library includes many tools for machine learning and statistical modelling including regression, clustering and dimensionality reduction.

Sklearn is not suitable for reading data as well as a separate Prophet. For this reason, it is necessary to read data through the Python Data Analysis Library (*Pandas*). For example, this library can work with .CSV, TSV, and SQL databases. After loading the file, Pandas creates a table with rows and columns (*dataframe*), which looks very much like a table in Excel.

Another library that is used in this article and is suitable for supporting the Prophet model is *Matplotlib*, which is used to display graphs. Through this library, individual forecasts generated by Prophet can be tailored to individual operator requirements. In this article, the first step was to load CSV data using the *Pandas* library and assign them the appropriate columns (ds, y).

Furthermore, it is possible to import public holidays to Prophet or any other significant days to clarify the prediction. For example, it is possible to implement the period around Christmas, events that are associated with business, such as world sports events (championship, olympics), or regional events such as concerts, exhibitions, etc. This may have an especially intriguing application in the corporate sector. Each business is influenced by other holidays and other events and their selection can significantly simplify forecasting. In the Czech Republic, the Christmas and New Year holidays were used for this application. When using Prophet, each date may not be defined separately, but the *Lower_window* interval may be used. One can also import all holidays using the Python *holiday* library.

Prophet also allows one to generate individual components of the forecast, such as trend and seasonality (yearly and weekly), which can be seen if holidays are included in the code. The weekly seasonal component uses fictitious variables and a seasonal component modelled using the Fourier series.

)

If it is necessary to generate interactive charts, it is advisable to use the external *Plotly* library. This library is provided by a private company of the same name and helps with the visualization of all kinds of charts.

Prophet includes a cross-validation function to measure forecast deviation and allows one to track accuracy with MSE, RMSE, MAE, MAPE, and coverage. RMSE and coverage were used to measure the resulting accuracy in this article.

3.2 Accuracy

Accuracy represents the exactness of a given model. Methods of measuring accuracy can be divided (Kolková, 2019) into scale-dependent measures, measures based on percentage errors, scaled errors (Athanasopoulos et al., 2017), and measures based on relative errors (Hyndman & Koehler, 2006).

The scale-dependent errors include the mean error (hereafter ME), the mean scaled error (hereafter MSE), the root mean squared error (hereafter RMSE), and the mean absolute error (hereafter MAE). The aforementioned methods of measuring accuracy are dependent on the scale of the data used and should not be used to compare data sets with different scales. Can be quantified in the following ways,

$$MSE = mean(e_t^2)$$
, where (2)

 e_t^2 is a prognosis error, i.e. the squared difference between the observed value and its forecast.

$$RMSE = \sqrt{MSE} \tag{3}$$

In the Prophet model, cross-validation including MSE, RMSE, MAE, MAPE and coverage is used to measure accuracy. In this paper the RMSE method will be used as the basic method of measuring accuracy.

Another method of accuracy is coverage. This method expresses coverage of the upper and lower intervals.

3.2 Data description

The input data for Prophet is always a data frame of two columns (ds, y). ds represents the date stamp and the column should be in a format that can be handled by the *Pandas* library. Column y must be numerical and represents the data we want to predict.

This forecast uses data from an e-commerce entity whose name cannot be disclosed due to privacy reasons. Nevertheless, the e-commerce entity is active in the sale of professional outdoor supplies and in the organization of outdoor training courses, seminars and competitions. A large proportion of customers are of the B2B segment. The histogram for this data is shown in Figure 5.





Source: own

The data represent individual daily sales (in the period from 1 October 2014 to 19 May 2019) and are obtained using the e-commerce segment analysis tool, Google Analytics. Basic descriptive statistics were chosen for their description. Their values are given in Table 1.

Descriptive Statistics	Sales
Mean	24171,43
Standard Error	514,66
Median	19527,71
Standard Deviation	21169,99
Sample Variance	448168349,01
Kurtosis	19,88
Skewness	2,93
Range	284181,01
Maximum	284181,01
Minimum	0,00
Sum	40898051,19
Count	1692,00

Table 1| Descriptive Statistic

Source: own elaboration

4 Results

Based on the current prediction, the in-sample period is 730 days and the out-of-sample is set to 365 days with a cut-off of 180 days.

It is clear from the results of the decomposition in Figure 6 that the e-commerce entity shows relatively weak growth potential, which can be seen in the underlying trend. From this, the business could be advised to find a new impetus for increasing sales and thus restoring the growth trend.

Weekly decomposition found Monday to be the strongest day of the week. On this day, the company reaches the highest sales figures. On the following days of the week, daily sales show a gradual decline. The weakest day of the week is Friday, which shows the lowest sales. The company can also use this fact to plan their employees' working hours.



Figure 6 | Decomposition of E-commerce Sales (CZK)

Source: own elaboration



Figure 7 | E-commerce Sales Forecast

Of course, annual sales developments were also analysed and decomposed. It follows from this decomposition that the company has a strong sales period with a start in the spring season (exactly in April) which then declines over the summer, until the pre-Christmas period, when the entity shows the best results. These observations may be attributed to consumer behaviour that favors the purchase of more outdoor goods just before the summer months and the desire to buy presents in the build-up to Christmas. We would recommend that the company utilize these observations to tailor target marketing campaigns for these times of the year in particular.

Source: own elaboration

Figure 7 shows the results of the forecast itself. The black dots represent daily sales in reality and the blue areas represent the forecasted daily sales. The two-sided confidence intervals are set to Prophet's default of 80%. It is clear that the forecasts respect the seasonality of the time series and its underlying growth trend. The forecast created in accordance with the theory defined above is verified by the accuracy of this model. For evaluation of cross-validation, statistic coverage and RMSE was used according to formula 3. The results are given figures 8 and 9.





Accuracy coverage can be seen to be rising to 0.87 in the first few days and declining to 0.66 as the forecasting horizon increases. This is in line with the general assumptions given by other scientific works and M4 competition.



Figure 9 | Cross Validation Metric (RMSE)

Source: own elaboration

The error measured by RMSE is shown in Figure 9. From this figure it is clear that accuracy during days fluctuates and, just like coverage, has the worst accuracy around 200 days.

Source: own elaboration

Consequently, as the prediction horizon increases, the error rate increases slightly (i.e. the accuracy deteriorates).

Conclusion and discussion

The aim of the article was to reveal seasonal trends in the development of revenues of a selected company in the area of the e-commerce segment based on the evaluation of applicability of the forecasting tool Facebook Prophet. To meet this goal, a new Prophet tool and Python libraries were defined to support this model. This tool was decomposed and a clear strong and weak sales period was defined. Subsequently, forecasting also took place in Python. This forecasting was measured by accuracy coverage and RMSE. Accuracy rates showed the anticipated fact that they were rising in the first days of the forecast and then slightly declining with a growing forecast horizon. This is also the biggest limitation of this model. Nevertheless, all forecasting methods can be expected to objectively reduce accuracy with the prediction horizon.

Thus, forecasting using Facebook's Prophet has been shown to be a possible tool for decomposition and prediction in the business economy. The results are relatively easy to interpret, even for managers without expert knowledge of forecasting and statistics. The implementation of decomposition, forecasts, and accuracy measurements in one model greatly simplifies the work of an analyst, or a manager, and allows for a clear and direct interpretation of the outputs obtained. Therefore, the Prophet platform is applicable to both large enterprises and can offer forecast options for medium-sized enterprises. Of course, the use of decomposition and information derived from it is determined by the needs of a particular company. In general, it allows one to forecast the days with the highest sales and then more efficiently plan the workload of one's employees. Last but not least, the basic trend in a forecast was revealed by decomposition. From this, the company under investigation was shown to have a weak growth potential. This should prompt consideration from the top management of the business and motivate it to make changes to the strategy.

From the results of this article, it can be verified that the Prophet meets the requirements of both model accuracy and relative simplicity of interpretation. In our opinion, it would also be beneficial if Prophet were expanded to include some form of artificial intelligence or multivariate analysis, which would extend its use to a wider range of business entities.

For further research, it would be useful to measure the accuracy of this tool compared to classical statistical methods, or with methods based on artificial intelligence. Other directions could be a comparison of forecasting different business variables. Alternatively, external factors can be implemented in the model that can affect forecasts (macroeconomic variables, outdoor temperatures, sentiment tracked by IBM Watson, etc.) to support forecasting by multiple regression analysis.

References

- Armstrong, J., & Green, K. (2014). *Methodology of Tree of Forecasting.* Retrieved from http://forecastingprinciples.com/index.php/methodology-tree.
- Athanasopoulos, G., Hyndman, R., Kourentzes, N., & Petropoulos, F. (2017). Forecasting with temporal hierarchies. *European Journal of Operational Research*, 262(1), 60–74. doi:10.1016/j.ejor.2017.02.046
- Bell, F. (2018). Introduction to forecasting. Retrieved from https://qcon.ai/qconai-2018-presentationvideos
- Crevits, R., & Croux, C. (2017). Forecasting Using Robust Exponential Smoothing with Damped Trend and Seasonal Components. Retrieved from http://ssrn.com/abstract=3068634
- Hyndman, R., & Fan, S. (2010). Density forecasting for long-term peak electricity demand. *IEEE Transactions on Power Systems*, 25(2), 1142-1153. doi:10.1109/TPWRS.2009.2036017
- Hyndman, R., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(3), 1-22.
- Hyndman, R., & Koehler, A. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679-688. doi:10.1016/j.ijforecast.2006.03.001
- Hyndman, R., Athanasopoulos, G., & Shang, H. (2015). hts: AnRPackage for Forecasting Hierarchical orGrouped Time Series. Retrieved from http://cran. unej. ac. id/web/packages/hts/vignettes/hts. pdf
- Hyndman, R., Wang, E., & Laptev, N. (2015). Large-Scale Unusual Time Series Detection. 2015 IEEE International Conference on Data Mining Workshop (ICDMW) 1616-1619. Atlantic City: IEEE. doi:10.1109/ICDMW.2015.104
- Kolková, A. (2018). Indicators of Technical Analysis on the Basis of Moving Averages as Prognostic Methods in the Food Industry. *Journal of competetivess*, 10(4), 102–119. doi:10.7441/joc.2018.04.07
- Kolková, A. (2018). Measuring the accuracy of quantitative prognostic methods and methods based on technical indicators in the field of tourism. *Journal Acta Oeconomica Universitatis Selye*, 7(1), 58-70.
- Makridakis, S., & Hibon, G. (2000). The M3-competition: results, conclusions and implications. International Journal of Forecasting, 16(4), 451-476. doi:10.1016/S0169-2070(00)00057-1
- Makridakis, S., & Hibon, M. (1979). Accuracy of forecasting: an empirical investigation (with discussion). *Journal of the Royala Statistical Society, 142*, 97-145.
- Makridakis, S., Hibon, M., Lawrence, M., Mills, T., Ord, K., & Simmons, L. (1993). The M2-competition: a real-time judgmentally based forecasting study. *International Journal of forecasting*, 9(1), 5-22. doi:10.1016/0169-2070(93)90044-N
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: concerns and ways forward. *PloS ONE*, 13(3), 1-26. doi:10.1371/journal.pone.0194889
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 34, 802–808. doi:10.1371/journal.pone.0194889
- Miller, P., & Swinehart, K. (2010). Technological forecasting: a strategic imperative. JGBM, 6 (2), 1-5.

- Moro, S., & Cortez, P. (2015). Business intelligence in banking: a literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation. *Expert Syst. Appl., 42* (3), 1314-1324. doi:10.1016/j.eswa.2014.09.024
- Reid, D. (1969). A comparative study of time series prediction techniques on economic data. Nottingham: University of Nottingham.
- Rob J Hyndman, S. F. (2010). Density forecasting for long-term peak electricity demand. *IEEE Transactions on Power Systems*, 25(2), 1142-1153. doi:10.1109/TPWRS.2009.2036017
- Šimeček, P. (2019). Statistical vs. Deep Learning Methods for Time Series Forecasting. Retrieved from http://www.mlmu.cz/archiv/
- Tassone, E., & Rohani, F. (2017). *Our quest for robust time series forecasting at scale*. Retrieved from http://www.unofficialgoogledatascience.com/2017/04/our-quest-for-robust-time-series.html
- Taylor, S., & Letham, B. (2018). Forecasting at Scale. *The American Statistician*, *7*2, 37-45. doi:10.1080/00031305.2017.1380080
- Verbesselt, J., Zeileis, A., & Hyndman, R. (2014). *bfast: Breaks For Additive Season and Trend* (*BFAST*). Retrieved from http://CRAN.R-project.org/package=bfast