Visualizing Time Series Predictability

Cagatay Turkay, *Member, IEEE*

Abstract— Predicting how temporally varying phenomena will evolve over time, or in other terms forecasting, is one of the fundamental tasks in time series analysis. Prediction has gained particular importance with the advent of real time data collection activities. Although there exist several sophisticated methodologies to predict time series, the success of a predictive analysis process remains mostly dependent on whether a particular phenomena is predictable. This paper introduces a methodology where visualizations coupled with a partition-based sampling strategy informs the analyst on the predictability of time series through the communication of prediction results applied on varying parts of data. We then discuss opportunities and research directions in supporting predictive tasks through visualization and interaction.

Index Terms—Time-series prediction, visual analytics, sampling

1 Introduction

Successfully predicting how phenomena varies over time is an intriguing goal that offers valuable insight for both researchers and businesses in several domains from climatology to medicine to economy [2] – which is evident by the vast amount of literature on this topic [4].

There have been several approaches to quantify the predictability of time series [5] and a number of measures have been suggested recently [7]. However, these methods often result in a single value that indicates the complexity of the process – signalling the difficulty in making predictions. Although such measures can be effective in capturing particular aspects of the data and give an overall idea on predictability, they do not provide insight on what makes a time-series hard to predict. A thorough understanding of where and why predictive methods fail or succeed will not only lead to better prediction results but also improve the understanding of the aspects that makes phenomena more predictable.

A mechanism to gain insight into how successful prediction methods operate is to evaluate the results of predictive analysis. One approach that is taken is the use of a *hold-out* strategy [10], where data points from the time series, mostly the most recent ones, are left out from the analysis and the predictions are done on the remaining data points. The results of the prediction are then compared to the initially held-out data points to arrive at a measure of prediction accuracy. Although this method have been applied as a standard approach, there has been little work done on how this process can be decoupled with visualization methods to best understand the predictability of time series.

In this paper, we introduce a methodology where we analyze the predictability of time-series through a three-stage process: *partition*, *predict*, and *visualize*. The methodology uses the hold-out sampling strategy that takes partitions from a time series which are then the input to a prediction algorithm. The predicted output is then compared against the rest of the data to evaluate how successful the prediction is. The comparisons are visualized together with the time-series to enable investigate where and why predictions fail or succeed. Unlike the conventional way of using this methodology, we suggest the use of visualization of multiple prediction results applied on parts of the data that are systematically varied.

2 Visualizing Predictability

Our method starts with a phase where we use a sampling strategy that partitions the data and provides these partitions as inputs to a prediction algorithm. In the conventional use of this methodology, only the last few data points, i.e., most recent, are left out and the rest of the data is used as input to the data. Here, we present three different strategies to generate various partitions of the data as illustrated in Figure 1. In order to simplify the discussion, we refer to the part of the data that is input to the algorithm as the *training partition* and the rest as the *evaluation partition*.

In the first method, the size of the training partition is extended systematically from only covering a small portion of the oldest data points to covering all the points. This method provides an **historic insight** on how predictable the series have always been.

The second method is where the evaluation partition is moved over the data to trace the whole time series while keeping the extend of the training partition constant. This sampling strategy is suitable to detect where predictions fail or succeed and to investigate whether there is any **systematic pattern** that determines these behaviour.

The third option is where we systematically vary the extend of the partition while keeping the scope of mainly bound to the recent data points. This method informs the analyst on **the extent the recent data points influence** prediction results.

To formally define the above strategies, we denote the training par-

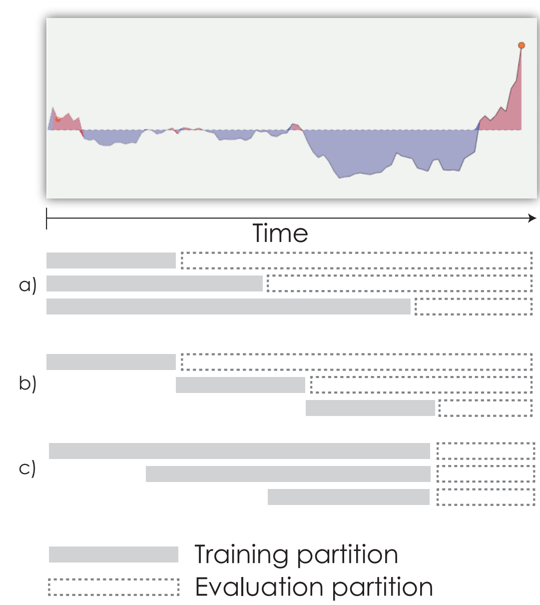

Fig. 1. Three different strategies to partition the data to be used as inputs to a prediction algorithm and to be used for evaluation purposes.

• *Cagatay Turkay, is with the Department of Computer Science at City University London, UK. E-mail: Cagatay.Turkay.1@city.ac.uk.*

Manuscript received 31 Mar. 2014; accepted 1 Aug. 2014; date of publication xx xxx 2014; date of current version xx xxx 2014. For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

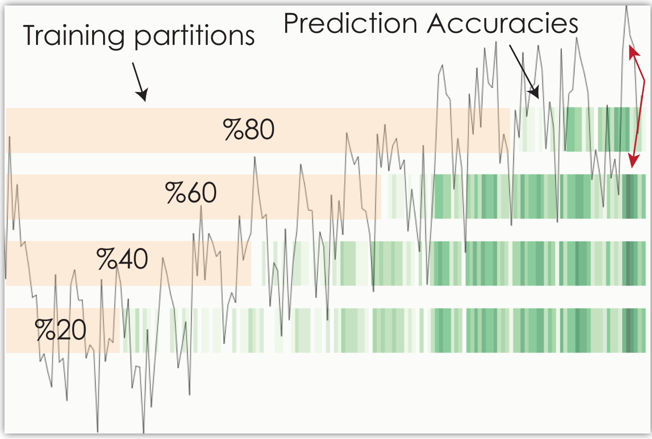

Fig. 2. Different prediction results, computed in accordance with the sampling strategy - a in Figure 1, are displayed in a superimposed style together with the time-series. More saturated green values indicate the areas where the fit is less accurate, i.e., high error. Notice that the three last prediction attempts have failed to predict the recent peak in the time series (red arrows).

tition with T, the evaluation partition with E, and the points in a time series of length n with x_i where $i \in [0,..,n]$. A definition of a single sampling step, where we generate sets T and E, can then be done as follows:

$$T = [x_i, .., x_j], \quad i < j \tag{1}$$
$$E = [x_{j+1}, .., x_n] \tag{2}$$

In a single analysis run, we choose one the three sampling strategies and systematically vary i and j values to generate several T sets.

These different sampling strategies provide the set of alternatives data partitions that can be fed to any prediction algorithm and the results can be compared against the *evaluation partition*. There exists several alternatives that can be used as the prediction algorithm [4] and the partitioning approach is not bound to any particular method in that respect. In order to demonstrate our approach, we utilized an ARIMA estimation model as our prediction algorithm [6]. Once the prediction is done using the *training partition*, we compute a measure to quantify how accurate the prediction fits the *evaluation partition*. There are several measures to evaluate the accuracy of prediction models [8] and for the sake of simplicity in this paper, we calculate the 1D difference between the predicted and the actual value.

We visualize the computed measures to summarize the accuracy of the prediction and we superimpose the time series on top of these visual summaries as seen in Figure 2. Here, we used the sampling strategy where we extend the training partition at each prediction iteration (refer to strategy-a in Figure 1). This is a single example where we can identify where prediction models are likely to fail. In three of these prediction runs (the lower ones with smaller training data), we observe that the prediction algorithm fails to determine the recent peak in the time series, whereas in the first trial, where %80 of the data is used as the sample, the prediction algorithm managed to fit that trend correctly. This is a clear indication that any pattern that is similar to this recently observed pattern is not likely to be predicted by the algorithm while the other patterns in the data are predictable even with very low sample sizes.

3 Discussions & Future Work

The different prediction accuracy patterns observed in our method can benefit analysts to evaluate whether a phenomena is suitable for short-term or long-term forecasting. Through the use of varying lengths of a time-series as the evaluation partition, one can evaluate whether the algorithms are suitable to make long-term or short-term forecasts, e.g., comparing a forecast for the last three data points vs. the last 20 for a series that consists of 100 data points.

A possible further research question is to investigate whether there exist patterns where predictions fail consistently and whether such patterns can be characterized further through the visualizations. Such an insight would require a systematic study of different patterns, possibly artificially generated, and identify how they manifest themselves in the visualizations.

An interesting future direction is to modify these methods to operate with data that update in real-time [3], i.e., streaming data such as stock market fluctuations. The predictions and the visualizations can be configured to update with the newly available data in the streams where the streaming data becomes the new *evaluation partition*.

4 Conclusion

Non-visual automated measures help analysts to make overall judgements on whether a phenomena is predictable, or whether a prediction is accurate. Although such measures can support experts in making decisions about analytical results, they often fall short to provide deep insight in why certain behaviour is observed. Exactly at this point, visualization has lots to offer where several alternative views of the same phenomena needs to be investigated to make insightful observations [9] or where visual summaries are needed to observe imperfections in computational results [11] or where visual guidance is needed to decide which statistical model to employ [1]. The method proposed in this paper exploits the strengths of visualization in making summaries and comparisons over these summaries. Being able to compare the output of how several different predictions perform over data partitions with different characteristics informs analysts to choose suitable prediction algorithms and suitable data that is predictable.

Acknowledgments

We would like to thank Germanos Pappas for the interesting discussions on this topic.

References

[1] M. Bögl, W. Aigner, P. Filzmoser, T. Lammarsch, S. Miksch, and A. Rind. Visual analytics for model selection in time series analysis. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2237–2246, 2013.

[2] P. J. Brockwell and R. A. Davis. *Introduction to time series and forecasting*, volume 1. Taylor & Francis, 2002.

[3] S. Chandrasekaran and M. J. Franklin. Streaming queries over streaming data. In *Proceedings of the 28th international conference on Very Large Data Bases*, pages 203–214. VLDB Endowment, 2002.

[4] J. G. De Gooijer and R. J. Hyndman. 25 years of time series forecasting. *International journal of forecasting*, 22(3):443–473, 2006.

[5] F. X. Diebold and L. Kilian. Measuring predictability: theory and macroeconomic applications. *Journal of Applied Econometrics*, 16(6):657–669, 2001.

[6] J. Durbin and S. J. Koopman. *Time series analysis by state space methods*. Number 38. Oxford University Press, 2012.

[7] J. Garland, R. James, and E. Bradley. Quantifying time-series predictability through structural complexity. *CoRR*, abs/1404.6823, 2014.

[8] R. J. Hyndman and A. B. Koehler. Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4):679–688, 2006.

[9] R. Maciejewski, R. Hafen, S. Rudolph, S. G. Larew, M. A. Mitchell, W. S. Cleveland, and D. S. Ebert. Forecasting hotspotsa predictive analytics approach. *Visualization and Computer Graphics, IEEE Transactions on*, 17(4):440–453, 2011.

[10] F. Schorfheide and K. I. Wolpin. On the use of holdout samples for model selection. *The American Economic Review*, 102(3):477–481, 2012.

[11] C. Turkay, J. Parulek, N. Reuter, and H. Hauser. Interactive visual analysis of temporal cluster structures. In *Computer Graphics Forum*, volume 30, pages 711–720. Wiley Online Library, 2011.