Visual Analytics Methods to Guide Diagnostics for Time Series Model Predictions

Conference Paper · January 2014

7 authors, including:

Some of the authors of this publication are also working on these related projects:

VALID – Visual Analytics im Datenjournalismus View project

[polycube] Towards integrated mental models of cultural heritage data View project

Visual Analytics Methods to Guide Diagnostics for Time Series Model Predictions

M. Bögl, W. Aigner, P. Filzmoser, T. Gschwandtner, T. Lammarsch, S. Miksch, and A. Rind
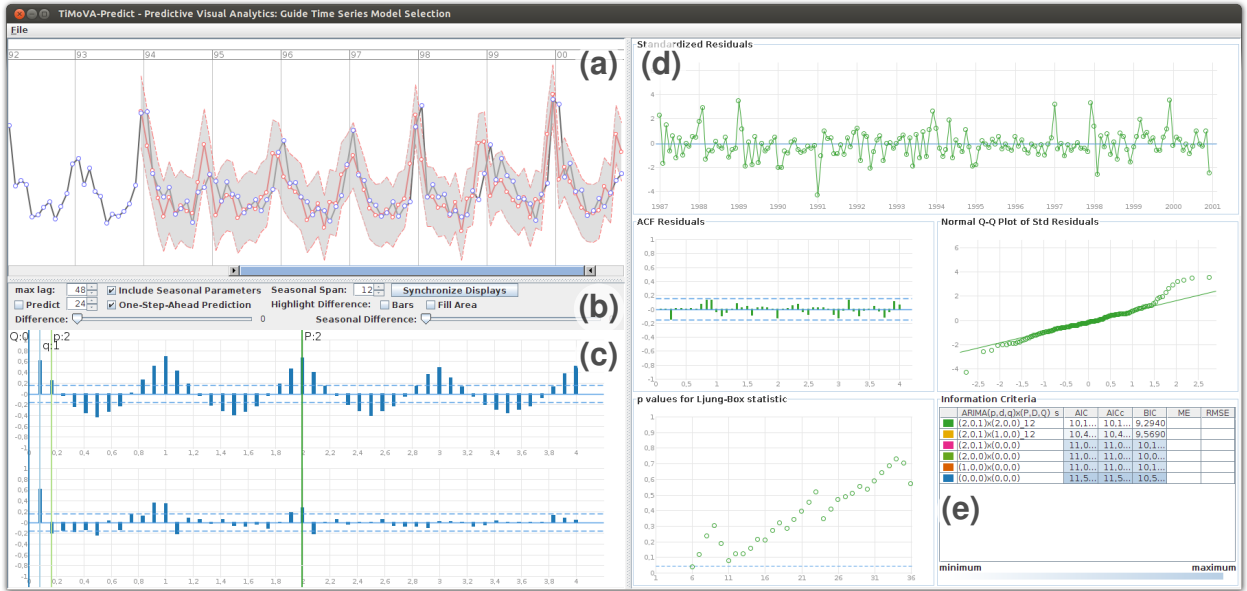
Fig. 1. The graphical user interface of TiMoVA-Predict: (a) time series display showing the one-step-ahead predictions (red) next to the actual values (blue/gray), (b) toolbox for model selection and prediction, (c) autocorrelation and partial autocorrelation plots for selecting the model orders, (d) diagnostic plots for the residual analysis, (e) model selection history including the information criteria.

Abstract—Visual Analytics methods are used to guide domain experts in the task of model selection through an interactive visual exploration environment with short feedback cycles. Evaluation showed the benefits of this approach. However, experts also expressed the demand for prediction capabilities as being already important during the model selection process. Furthermore, good model candidates might show only small variations in the information criteria and structures which are not easily recognizable in the residual plots. To achieve this, we propose TiMoVA-Predict to close the gap and to support different types of predictions with a Visual Analytics approach. Providing prediction capabilities in addition to the information criteria and the residual plots, allows for interactively assessing the predictions during the model selection process via an visual exploration environment.

Index Terms—Visual analytics, time series prediction, model selection, time series analysis.

◆

1 INTRODUCTION

In various domains time series analysis is applied to find a model to a given time series. Predicting future values is one of the main goals in time series analysis [1] and an important area of application for a model [3]. One class of time series models are autoregressive integrated moving average (ARIMA) models [3]. ARIMA models are widely used for prediction tasks, for instance predicting electricity prices [5], system failure analysis [10], and in different financial and medical domains [17]. Finding such models for a given time series is a challenging and cumbersome task that involves iteratively re-

peated computations and visualizations to be judged by domain experts. These users are usually experts in their respective domain with knowledge in statistics and time series analysis, for instance public health officials or economic analysts.

In previous work [2], we introduced a Visual Analytics (VA) approach to support these users in the task of selecting adequate ARIMA models. During evaluation of the prototype it became apparent that including the possibility to perform actual prediction would improve the model selection process considerably. Integrating the prediction capabilities into the exploration environment offers another perspective on the adequacy of the model for a given time series and raises the confidence in the resulting model.

To integrate this prediction functionality in the model selection process, we propose TiMoVA-Predict (Section 2) and demonstrate how domain experts can use TiMoVA-Predict in a usage scenario [16] (Section 3). For the usage scenario we use a dataset about the daily number of deaths from cardiovascular disease of people aged 75 and older in Los Angeles for the years 1987 to 2000 from the NMMAPS study [12, 15].

TiMoVA-Predict guides during the model selection process with the following functionalities:

- *Markus Bögl, Peter Filzmoser, Theresia Gschwandtner, Tim Lammarsch, and Silvia Miksch are with Vienna University of Technology, Austria. E-mail: {boegl, gschwandtner, lammarsch, miksch}@ifs.tuwien.ac.at and peter.filzmoser@tuwien.ac.at.*
- *Wolfgang Aigner and Alexander Rind are with St. Pölten University of Applied Sciences, Austria. E-Mail: {wolfgang.aigner, alexander.rind}@fhstp.ac.at.*

- close coupling of prediction capabilities of the models with the visual model selection interface,
- interactive visual exploration of different types of predictions in direct relation to the input time series,
- visual means to explore the differences between the predicted and the actual values.

2 TIMOVA-PREDICT

In addition to our previous work, TiMoVA-Predict provides means to (1) predict future values with respect to the selected time series model and (2) judge the adequacy of the selected time series model based on the prediction results. The close coupling of the prediction capabilities with the visual model selection interface helps to decide on a less complex model with fewer parameters to prevent overfitting.

Box and Jenkins [3] discuss seasonal ARIMA models together with their Box-Jenkins methodology, which describes how to find an adequate model for a given time series. A seasonal $ARIMA(p,d,q) \times (P,D,Q)_s$ model combines a non-seasonal $ARIMA(p,d,q)$ with a seasonal $ARIMA(P,D,Q)_s$ model multiplicatively. Each of them has an autoregressive component ($AR(p)$ and $AR(P)$), a moving average component ($MA(q)$ and $MA(Q)$), a difference transformation (d and D), and s for the seasonal length. The parameters p, P and q, Q describe the model order of the AR and the MA components and specify the number of parameters that need to be estimated by, for instance, a maximum likelihood estimator. For more details about the ARIMA models cf. [3, 17].

For fitting the time series model and applying the prediction we used the standard libraries integrated in the R project for statistical computing [13] and the R libraries astsa [18] and zoo [20]. Our Java implementation of the TiMoVA-Predict prototype is connected to R using the Java/R Interface [19]. For the visualizations we used the software framework prefuse [9], and as API for time-oriented data we used TimeBench, a software library for time-oriented data [14].

As we use R libraries for the model fitting and the prediction, TiMoVA-Predict is independent from a specific implementation of algorithms and they are basically interchangeable. We employ the astsa library in the prototypical implementation. This fits the model using conditional-sum-of-squares and maximum likelihood and provides finite-history prediction applying Kalman forecast [17].

In general, the application of an ARIMA model for prediction is based on the available observations $x_1, x_2, \ldots, x_n$ at the time points $t_1, t_2, \ldots, t_n$. The predictions of the next m time points $t_{n+1}, \ldots, t_{n+m}$ are then denoted by $\hat{x}_{n+1}, \ldots, \hat{x}_{n+m}$, where m is an integer ≥ 1. Thus, the term predict refers to the predictions of future observations, using the corresponding time series model. If we want to compare predictions with actual values, we mimic this process: The time series is split at time point t_k, with $1 < k < n$, the model parameters are estimated for $x_1, \ldots, x_k$, and the predicted values are $\hat{x}_{k+1}, \ldots, \hat{x}_n$. These values can be compared with the observed values $x_{k+1}, \ldots, x_n$. A variant is the one-step-ahead prediction, where k is set to $n/2$ and step-by-step increased by one until $k = n$. In each step, the model is fit to the data points $x_1, \ldots, x_k$ and the predicted value $\hat{x}_{k+1}$ is derived. In that way, prediction is successively done at only one time point using all previous information. For more details about the estimation of the parameters and the error terms, cf. [17].

From evaluation of earlier work, it became apparent that it is beneficial to not only use diagnostic plots for analyzing the model residuals and the information criteria, but to also apply the time series models for predicting future values (as mentioned in Section 1). This suggests integrating the prediction capabilities of the models in the model selection process. It is important for the model selection task to consider the principle of parsimony [3] to prevent models from getting too complex and avoid overfitting. Therefore, including predictions in the interactive exploration environment during the iterative model refinement loop enables domain experts to judge the prediction capabilities and to select a parsimonious model.

To support the ARIMA model selection with the prediction capabilities of the model, TiMoVA-Predict combines both in an interactive
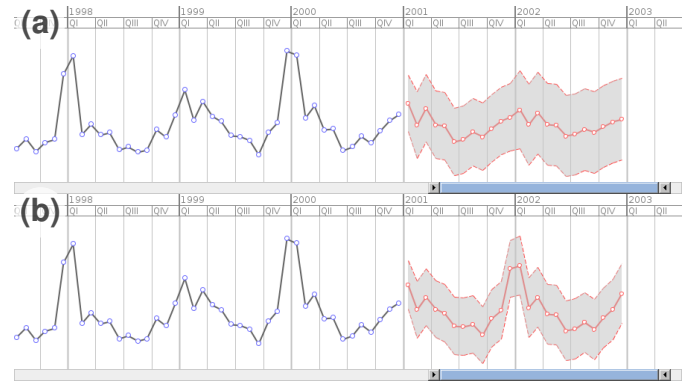

Fig. 2. Prediction of two seasonal cycles applying a model that does not consider the seasonal cycle adequately in (a) and an adequate model in (b). Comparing the peaks at the turn of the years in the prediction and the previous years, the user can identify in (a) that the model does not predict well enough, whereas the model in (b) predicts very well.

exploration environment, shown in Figure 1. The graphical user interface consists of five main areas:

(a) the time series display showing the input time series and, if applied, the model predictions,

(b) the model selection and prediction toolbox,

(c) the autocorrelation and partial autocorrelation (ACF/PACF) plots (the model selection is steered by interactively moving the vertical lines),

(d) the residual plots to perform the model diagnostics and decide for an adequate model, and

(e) the information criteria for the same purpose as (d) and for investigating the model selection history.

A detailed explanation about the residual plots is given in [2] or [17]. In this paper, we focus on the relevant elements for the prediction and how the prediction is integrated into TiMoVA-Predict. The user first explores the input data in the time series display (Figure 1a) and consults the ACF/PACF plot (Figure 1c), which gives a first idea of possible model candidates. The models are selected using the ACF/PACF plot and the toolbox (Figure 1b). The adequateness of the model is diagnosed using the residual plots (Figure 1d) and the information criteria (Figure 1e).

To couple the prediction with the model selection process, we provide prediction controls in a toolbox (Figure 1b). As described before, we include the prediction of future values and one-step-ahead prediction that predict values within the given time series. A checkbox activates the prediction of future values followed by an input field to specify the number of time point to be predicted. In Figure 2 the predicted values are shown in the time series display. In our example data set we have monthly data with yearly cycles. Selecting 24 as the number of time points to be predicted, shows a prediction for 24 months or 2 years. The predicted values are represented by the red points connected with the red line, while the dashed red line shows the upper and the lower confidence interval, in this case $\pm 2 * se$ (standard error). The gray background emphasizes this confidence interval.

Another way to assess the prediction capabilities of a time series model is the one-step-ahead prediction. The time series is split into two parts and the model is fitted to the first part of the time series to predict the second part. Currently, we split the time series, with approximately one half of the time points in each part. However, in the future we want to involve the user in the decision where to split. The only constraint is the minimum length for the earlier/left part of the time series, depending on the model order. The results of the one-step-ahead prediction is shown in Figure 1a. In addition to the actual time series (represented by blue circles connected by gray lines) the predicted values and the confidence interval are plotted by red circles connected by red lines and a gray background band. Using the horizontal range slider below the time series display, the user can zoom
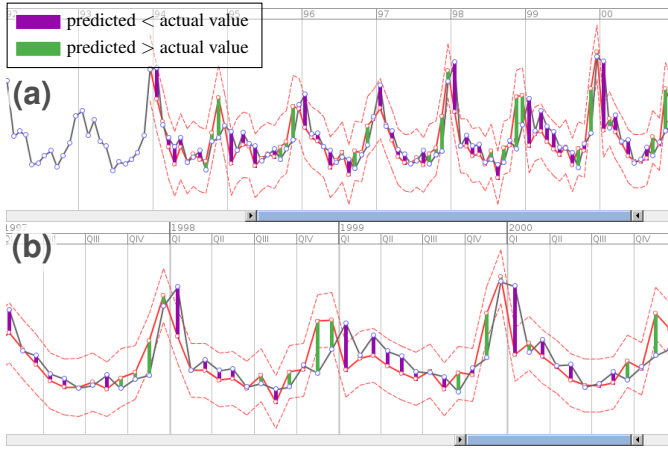
Fig. 3. The difference between the predicted and the actual values are indicated with colored vertical bars. The coloring encodes if the difference is positive or negative i.e., if the prediction is higher or lower as the actual value. Zooming with the time slider into (a) results in (b).

into the time series and compare the actual versus predicted values along with the confidence interval in more detail.

For additional visual support we provide two ways of highlighting the difference between actual and predicted values. Each of them can be added to the one-step-ahead prediction using the checkboxes in the toolbox (Figure 1b). The first option is to highlight the difference with vertical bars. In Figure 3 we show this at two different zoom levels. We used two colors from a diverging color scale to quickly identify if the prediction is higher or lower than the actual value. For demonstration we used purple for differences where the predicted is smaller than the actual value and green for vice versa, i.e. cf. the legend in Figure 3 and 4. The second option to highlight the difference is shown in Figure 4. In this case the area between the predicted value and the actual value is filled. Large differences result in large colored areas, which are recognized by the user even at a higher zoom level. In Section 5 we discuss these modes of highlighting and provide further ideas.

3 USAGE SCENARIO

After introducing the basic concept and the graphical user interface of TiMoVA-Predict, we briefly describe a usage scenario explaining how the model prediction capabilities are used in the model selection process. The scenario is that a public health official needs to predict the number of deaths from cardiovascular disease. We use an example dataset from the MMAPS study [12, 15], cf. Section 1, aggregating the daily number of deaths from cardiovascular disease to monthly sums.

The user first investigates the time series (Figure 1a) and the ACF/PACF plots (Figure 1c). According to the information depicted by the bars in these plots, the model order is selected (Figure 1b,c), which triggers the computation of the residuals (Figure 1d). After increasing the model order of the AR and the MA component to $p = 2$ and $q = 1$, the user recognizes that there is a seasonal behavior that is
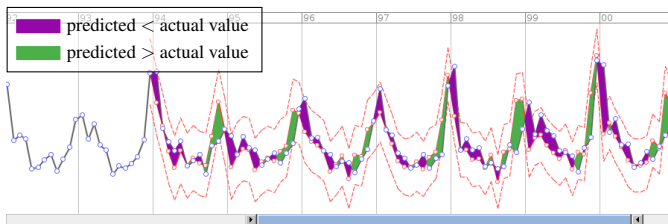

Fig. 4. The residuals are highlighted in a similar way as in Figure 3, but instead of connecting the predicted and the actual values with vertical bars, the area between the two lines is filled with color.

not covered by the model yet. Increasing the order of the seasonal AR component in the model to $P = 1$ results in a better model according to the information criteria and the residual plots. Experimenting with the seasonal AR component the user recognizes that $P = 1$ and $P = 2$ both give very good models according to the information criteria and the residual plots. The information criteria of both models are very close and the residual plots do not differ very much. The user needs to investigate the prediction capability in more detail now. First she/he selects the model with order $p = 2$, $q = 1$, and $P = 1$, and selects the checkbox with two seasonal cycles in TiMoVA-Predict. This triggers the prediction and shows the predicted values (Figure 2a). Comparing the predicted values at the change of the year 2000/2001 and 2001/2002 with the actual values in the change of the year 1998/1999 and 1999/2000 (Figure 2a), the user identifies that the seasonal behavior is not predicted well by the model. Switching the seasonal AR component to $P = 2$ using the vertical lines in (Figure 1c) results in new values (see Figure 2b), where the peaks in the changes of the years are better representing the seasonal pattern.

To investigate the prediction capabilities of the model in more detail, the user clicks on the one-step-ahead prediction checkbox (Figure 1b). TiMoVA-Predict computes additional predicted values as described in Section 2. They are plotted along with the actual values (Figure 1a). The user investigates the differences between the actual and the predicted values by highlighting them (Figure 3 and Figure 4). In both views, the user can immediately see that the predictions are close to the actual values. Therefore, the user is satisfied with the prediciton results and decides that the selected model is adequate for the time series.

4 RELATED WORK

There are several software tools used by mathematicians and statisticians that implement the state-of-the-art methods and models for time series analysis and support prediction, for instance the R project for statistical computing, SAS JMP, MATLAB, EVIEWS, Mathematica, Stata, and Gretl (cf. [2]). Most tools support time series analysis by menu driven user interfaces. The model fitting, model diagnostic, and application for prediction are initiated by the user either by command line or menus and input forms, separating the application of time series model from the model selection process, and provide no direct support for the prediction when selecting the models. Often, it is necessary to decide on a set of models or compute a large bunch of models and arrange the results manually.

One notable solution is TimeSearcher [4], which is a visualization tool to search and explore time series data. With dynamic queries it finds patterns and displays multiple forecasts, provided by similarity-based prediction. TimeSearcher uses a data-driven approach that needs exceptional events to be excluded and requires large datasets compared to model driven methods like ARIMA [4].

Another related work is the x12GUI [11] package for R. It offers an interactive tool for the X-12-ARIMA software for seasonal adjustment. The focus is on the exploration of the time series and the results of the seasonal adjustment as well as the manual editing of outliers [11]. However, for selecting a time series and adjusting the parameters for the X-12-ARIMA call, only form-based input is used. For the computed models there is a history, which allows for loading previous settings, but not to browse and directly compare them. For single models x12GUI provides also the possibility to predict and visualize future values, but this is not integrated in the model selection process.

To the best of our knowledge, there has been no work so far, which closely couples the prediction with interactive visualization and immediate visual feedback during the process of selecting ARIMA models.

5 DISCUSSION AND OPEN CHALLENGES

Predicting future values is an important application area for time series models [3] and one of the main goals in time series modeling [1]. This is a strong argument to integrate the prediction capabilities of a model in the model selection process to find a model that performs well in prediction. Of course, it is possible to just rely on the diagnostic residual plots and the information criteria, but often the values of
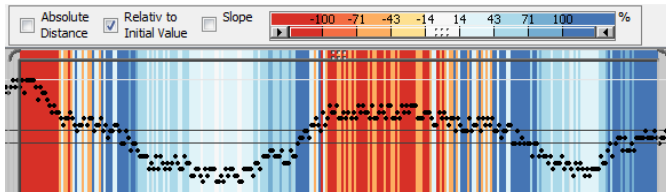
Fig. 5. In CareCruiser [8] a diverging color scale is used to visualize differences in patient parameters in the medical domain. This is a possible technique to highlight the difference between the actual and the predicted values in another way as shown in Figure 3 and 4.

the information criteria differ only little and the structures in the residual plots are not easily recognizable among good model candidates. TiMoVA-Predict provides possibilities to directly show the predicted values and to compare them to the actual values. We plan additional user studies to validate TiMoVA-Predict, including the future ideas discussed below.

A very challenging open topic is the comparison of this rather strict model based predictions to bootstrapping based predictions [17, 6]. This could be important especially for short time series, where the risk of over-parametrization is present. Bootstrapping in time series is not trivial because of the dependency in time. We expect that VA methods may contribute for this challenge by interactive adjustments of different options, like selecting different time points and resizing the bootstrapping samples.

In Figure 3 and 4 we showed different ways of highlighting the difference between actual and predicted values in the one-step-ahead prediction. The idea of this coloring is to grasp how much they differ, even on a higher zoom level. A large difference will generate long bars or large areas, which are visually very prominent in the time series. This indicates a model with inferior predictions. A limitation is that it is not possible to judge details on what the difference actually means, mainly in context of the confidence interval.

CareCruiser [8] uses a diverging color scale to highlight the progress of parameter values from the initial value toward the intended value of applied treatments. They use the whole height of the plot to fill with the color according to the difference. Splitting the confidence interval of the prediction into, for instance, four confidence intervals, like 75%, 90%, and 95% with the difference between the actual and the predicted values in the positive and the negative direction, we could use a similar approach, as shown in Figure 5. That enables to fill the background, using the full height, according to the deviation of the actual versus predicted value.

There is an interesting effect between the bar and the area representation of the difference. The surface of the area depends strongly on the incline of the lines, which depends on the zoom level and the length of the time series. Therefore, the area is rather perceived as a slightly crooked bar than a vertical difference at a specific time point, cf. for instance November 1999 and January 2000 in Figure 4.

Another variant for analyzing the difference is to use a parallel visualization for displaying the difference only. Qualizon graphs [7] can be used as a space efficient visualization, for instance below the time series display (Figure 1a). With Qualizon graphs the qualitative abstractions of the differences according to the confidence intervals using the diverging color scale would show more radical changes.

In case of the one-step-ahead prediction, there is the possibility of quantifying the difference (actual vs. predicted values). Prediction error measures, like the mean error (ME) or the residual mean squared error (RMSE), can be included for the models. Integrating that in the table of the model history, as indicated in the additional columns (Figure 1e), the user can compare the prediction errors of different models.

6 Conclusion

Integrating the prediction capability of time series models for analyzing and selecting ARIMA models, supports the users in the decision for an adequate model. Calculating and visualizing the difference be-

tween the actual and the predicted values guides the user to prevent overfitting and to find a parsimonious and adequate model. We discussed the benefits of including the prediction capabilities and the remaining open challenges and possibilities for future work. In a next step we will evaluate these proposed concepts. TiMoVA-Predict enables the user to judge a model's prediction capability and decide on an adequate model, even if the residual plots do not show recognizable structures and the information criteria differ only slightly.

References

[1] S. Bisgaard and M. Kulahci. *Time Series Analysis and Forecasting by Example*. John Wiley & Sons, Hoboken, 2011.

[2] M. Bögl, W. Aigner, P. Filzmoser, T. Lammarsch, S. Miksch, and A. Rind. Visual analytics for model selection in time series analysis. *IEEE TVCG*, 19:2237–2246, 2013.

[3] G. Box, G. Jenkins, and G. Reinsel. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Hoboken, 4th edition, 2008.

[4] P. Buono, C. Plaisant, A. Simeone, A. Aris, B. Shneiderman, G. Shmueli, and W. Jank. Similarity-based forecasting with simultaneous previews: A river plot interface for time series forecasting. In *Proc. Int. Conf. Information Visualisation*, pages 191–196. IEEE, 2007.

[5] J. Contreras, R. Espinola, F. Nogales, and A. Conejo. ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, 18(3):1014–1020, 2003.

[6] B. Efron and R. J. Tibshirani. *An introduction to the bootstrap*. Chapman and Hall, New York, 1994.

[7] P. Federico, S. Hoffmann, A. Rind, W. Aigner, and S. Miksch. Qualizon graphs: Space-efficient time-series visualization with qualitative abstractions. In *Proc. 12th Int. Working Conf. on Advanced Visual Interfaces (AVI2014)*, pages 273–280. ACM, 2014.

[8] T. Gschwandtner, W. Aigner, K. Kaiser, S. Miksch, and A. Seyfang. Care-Cruiser: Exploring and visualizing plans, events, and effects interactively. In *IEEE PacificVis*, pages 43–50, 2011.

[9] J. Heer, S. K. Card, and J. A. Landay. Prefuse: A toolkit for interactive information visualization. In *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI '05)*, pages 421–430. ACM, 2005.

[10] S. L. Ho, M. Xie, and T. N. Goh. A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. *Computers & Industrial Engineering*, 42(2):371–375, 2002.

[11] A. Kowarik, A. Meraner, D. Schopfhauser, and M. Templ. Interactive adjustment and outlier detection of time dependent data in R. In *Conference of European Statisticians, Work Session on Statistical Data Editing*, Oslo, Norway, 2012. United Nations - Economic Commission for Europe.

[12] R. D. Peng and L. J. Welty. The NMMAPSdata package. *R News*, 4(2):10–14, 2004.

[13] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012.

[14] A. Rind, T. Lammarsch, W. Aigner, B. Alsallakh, and S. Miksch. TimeBench: A data model and software library for visual analytics of time-oriented data. *IEEE TVCG*, 19:2247–2256, 2013.

[15] J. Samet, F. Dominici, S. Zeger, J. Schwartz, and D. Dockery. The national morbidity, mortality, and air pollution study. Part I: Methods and methodologic issues. Research Report 94, Health Effects Institute, Cambridge, 2000.

[16] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE TVCG*, 18(12):2431–2440, 2012.

[17] R. H. Shumway and D. S. Stoffer. *Time Series Analysis and its Applications. With R examples*. Springer, New York, 3rd edition, 2011.

[18] D. Stoffer. *astsa: Applied Statistical Time Series Analysis*, 2012. R package version 1.2, http://CRAN.R-project.org/package=astsa (Sept. 12, 2014).

[19] S. Urbanek. rJava: Low-level R to java interface, 2011. R package version 0.9-6. http://CRAN.R-project.org/package=rJava (Sept. 12, 2014).

[20] A. Zeileis and G. Grothendieck. zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software*, 14(6):1–27, 2005.