Visual Analytics for Model Selection in Time Series Analysis

Markus Bögl, Wolfgang Aigner, Peter Filzmoser, Tim Lammarsch, Silvia Miksch, and Alexander Rind
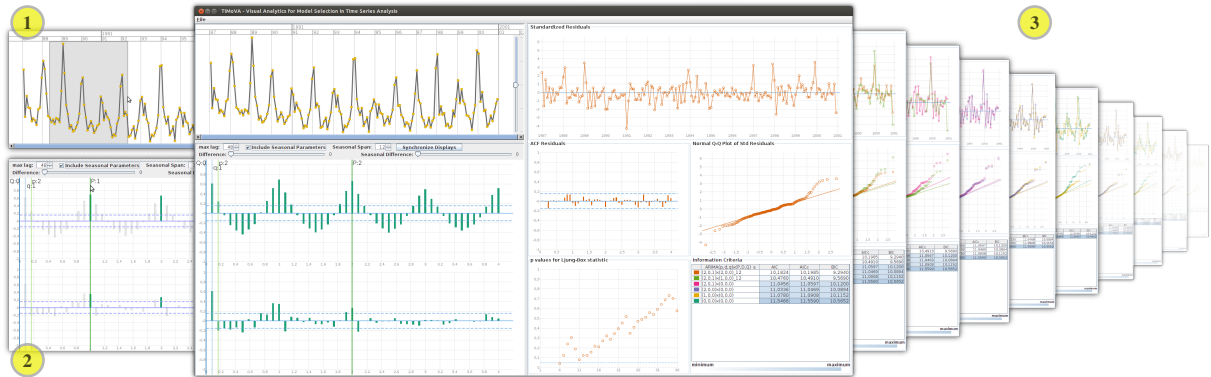

Fig. 1. TiMoVA provides visual guidance for domain experts in the task of model selection, by (1) enabling to choose a certain range of interest in the time series, (2) supporting the model selection interactively via the visual interface, and (3) visualizing the model transitions and giving immediate visual feedback of the model output for the iterative refinements. For more details see Figure 5.

Abstract—Model selection in time series analysis is a challenging task for domain experts in many application areas such as epidemiology, economy, or environmental sciences. The methodology used for this task demands a close combination of human judgement and automated computation. However, statistical software tools do not adequately support this combination through interactive visual interfaces. We propose a Visual Analytics process to guide domain experts in this task. For this purpose, we developed the TiMoVA prototype that implements this process based on user stories and iterative expert feedback on user experience. The prototype was evaluated by usage scenarios with an example dataset from epidemiology and interviews with two external domain experts in statistics. The insights from the experts' feedback and the usage scenarios show that TiMoVA is able to support domain experts in model selection tasks through interactive visual interfaces with short feedback cycles.

Index Terms—Visual analytics, model selection, visual interaction, time series analysis, coordinated & multiple views

---◆---

1 INTRODUCTION

Statistical time series analysis is a challenging task performed by experts in different domains. A practical application scenario is, for example, a public health official predicting the number of people that need to be treated because of cardiovascular reasons in the next year. Another scenario is to be prepared for the number of patients suffering from seasonal flu. The datasets to be analyzed are obtained from observations collected over time, optimally at periodic and equally spaced intervals and ideally without missing values. Such a dataset is called a time series. A range of methods, algorithms, and models to analyze these time series exist in the literature [4, 6, 7, 15, 37]. Moreover, they are implemented in most common software tools for statistical computing. In our work we focus on the most prominent and large class of models, namely ARIMA and seasonal ARIMA models [7]. These models are applied in a variety of practical application domains. The huge amount of work discussing ARIMA models reflects the importance of this model class [4, 6, 7, 15, 16, 37] and there exists an established process for model selection known as Box-Jenkins methodology [6]. We present and discuss this methodology and the theoretical underpinnings briefly in Section 3. However, currently available software tools do not appropriately support the workflow of the Box-Jenkins methodology, as we argue in Section 2. Therefore, support for domain experts would be beneficial for working on model selection in time series analysis.

A potential way to overcome the above-mentioned shortcomings is, in the spirit of Visual Analytics (VA), to "combine automated analysis techniques with interactive visualizations" [21, 40]. This raises the question of how to use VA methods to support the task of model selection for time series analysis. According to the characteristics of design studies [27, 35], we need a comprehensive understanding of the domain problem. We provide its characterization to judge our solution and analyze the tasks in Section 3. In Section 4 we identify the target users, formulate the requirements for the tasks, and specify the data used in time series analysis.

The main objective of this work is to use well-established information visualization techniques [1, 11] and apply them to the particular target problem. For this reason, we defined a VA process based on existing VA process descriptions [21, 25] and implemented a prototype to facilitate it. The design and implementation of the VA process and the prototype were iteratively refined and judged by experts in information visualization and statistics using formative evaluation of user experience [24]. We present the results of this refinement process in Section 5, where we discuss the VA process description and our prototype. We named our prototype TiMoVA, which is an abbreviation of **Ti**me series analysis, **Mo**del selection, and **VA**.

In addition to the formative evaluation and the iterative design, we evaluated the final version of the prototype by defining usage scenarios and applying the prototype to an example dataset. We present the evaluation in Section 6, where we apply the usage scenarios on an example dataset. Furthermore, we also evaluated the user experience

---

• *Markus Bögl, Wolfgang Aigner, Peter Filzmoser, Tim Lammarsch, Silvia Miksch, and Alexander Rind are with Vienna University of Technology, Austria. E-mail: {boegl, aigner, lammarsch, miksch, rind}@ifs.tuwien.ac.at and peter.filzmoser@tuwien.ac.at.*

by performing a feedback session with two external domain experts. We used this informal user feedback to argue how our target users assess TiMoVA [24]. In the discussion of the results (Section 5) and the evaluation (Section 6), we describe the visual encodings and interaction mechanisms used in the prototype and how they fulfill our requirements.

In particular, the main contributions of our paper (Section 7) address issues of interactive visual guidance to ease the model selection in time series analysis by

- selecting the model order inside the autocorrelation and partial autocorrelation function plot, where the domain experts get the information about the model order,
- providing immediate visual feedback of the model results to the domain expert while adjusting the model order, and
- visualizing the model transitions to enable the domain experts to decide whether or not the model improves.

2 RELATED WORK

Following the design study methodology [27, 35], we apply existing and known VA methods and process descriptions to the domain problem of model selection in time series analysis. We based our work on the techniques for visualizing time-oriented data [1, 11] and use state-of-the-art VA process descriptions [21, 25, 40].

All major mathematical and statistical software tools implement the state-of-the-art methods and models for time series analysis, which we describe briefly in Section 3. We considered tools, like the R project for statistical computing [30], SAS JMP [34], MATLAB [39], EVIEWS [20], Mathematica [42], Stata [38], and Gretl [14]. Except for the R base package all these tools support time series analysis and models by menu driven user interfaces. The separate calculations and visualizations need to be initiated by the user either by menus and input forms, or by command line methods. What all these approaches are missing, is the guidance to browse and visually compare models directly when selecting the model. Instead it is necessary to either decide on a set of models or compute a whole bunch of models and arrange them in visualizations by hand to compare them. To summarize, they do not support the repeated execution of the separate steps in the iterative Box-Jenkins methodology [6] well.

One notable solution is the x12GUI [23] package for R. It offers an interactive graphical user interface for the x12 [22] package, which provides a wrapper function to the X-12-ARIMA software. The focus of their approach is to explore the time series and the results of the seasonal time series adjustment as well as to enable the user to do interactive manual editing of outliers [23]. However, the user interface supports the user in selecting a time series and adjusting the parameters for the X-12-ARIMA call using form-based input only. It also provides a history for computed models, which allows for loading previous settings but not to browse and directly compare them.

Because we have the very specific target problem of model selection in statistical time series analysis, to our knowledge, there is only distantly related work in VA, as for example TimeSearcher [9, 10] and the work for visual-interactive time series preprocessing [3]. Both approaches apply VA methods to time-oriented data, but differ in the tasks and their solutions.

Motivated by these findings, we give the necessary background information in the following section to characterize and specify our target problem in more detail and ground our motivation further.

3 PROBLEM CHARACTERIZATION

In this section we provide the characterization of the domain problem [35] and discuss the tasks necessary for the model selection.

Example Dataset and Task. An important domain where time series analysis is used is public health and epidemiology. We have chosen a dataset from this domain for the evaluation in Section 6.1 and illustrate a possible analysis task in the following. The dataset contains 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 [29, 33]. A possible scenario is that a public health official needs to predict the expected number of death from
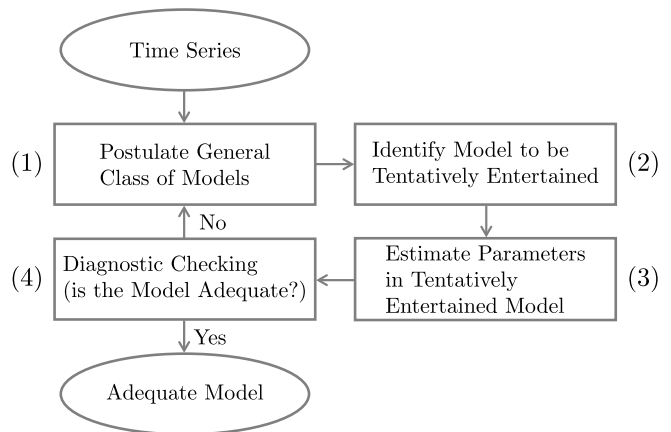

Fig. 2. Box-Jenkins Methodology. An iterative process for model selection of time series [7]. See Section 3.1 for details about the process.

cardiovascular disease to start a prevention program. To do the prediction, the health official has to find a model based on the given data. The Box-Jenkins methodology is a standard method to solve this task. We describe this methodology in Figure 2 and in the following.

3.1 Box-Jenkins Methodology

The Box-Jenkins methodology [6] is an iterative process to select an adequate model for a given time series (see Figure 2). It has been widely used in time series analysis and is an established method for model selection [4, 7, 15, 37]. To find or select a model for a given time series, it is according to Box et al. [7] necessary to:

(1) Use the incomplete theoretical knowledge about the underlying mechanisms and the experience from theory and practice to consider a useful general class of models. By general class of models Box et al. [7] mean any subclass combination of ARIMA and seasonal ARIMA model components. Fitting these models directly to data, would be too extensive and time consuming.

(2) Apply methods to select an appropriate parsimonious (see below) model by deciding on the model order. This determines the number of parameters in the model and gives some rough estimates for them.

(3) Fit the model to the data and estimate its parameters.

(4) Finally, check the model with diagnostics to uncover possible lacks of fit and find their causes.

These steps are repeated until an adequate model is found, which can subsequently be used for forecasting (see Figure 2).

The method is sometimes reduced to a simplified version [15], where the decision for a general class of models (1) and the identification of a model that can be tentatively entertained (2) are combined and entitled as *model specification*. Step (3) is renamed to *model fitting* and step (4) to *model diagnostics*. We present a more detailed description of these separate steps in Section 3.3, 3.4, and 3.5.

The crux of model selection is summarized in the famous quote of George Box that "Essentially, all models are wrong, but some are useful" [5, p. 424]. When introducing the Box-Jenkins methodology, Box and Jenkins [6] use a language that indicates that there is a "useful" and "adequate" model for a time series, but we cannot assume that it is a "true" or "correct" model. The uncertainty of such models is put straight by using the term "tentatively entertaining a model" [4] (see Section 3.4).

Principle of Parsimony. An important principle in the model selection process is the principle of *parsimony* [7]. It is described by the paraphrased quote of Albert Einstein "everything should be made as simple as possible but not simpler" [4, p. 18]. In the process of model selection this means that if there are different candidate models to adequately represent the time series, the model with the least parameters is preferable [15].

In this section, we presented the methodology how to find an adequate model for a given time series. This methodology was introduced for a specific class of models [6], which we describe in the following section.

3.2 ARIMA and Seasonal ARIMA Models

With classical regression it is often not possible to explain a time series sufficiently [37]. Therefore, alternative models exist. We briefly describe the key ideas of the different models and explain them without presenting the full details of the formal definition of these models. These formal definitions and the details are not necessary to understand and argue the design choices and explain the interpretation of the visual representations in Section 5 and 6. For more details and the formal definitions we refer to the literature in time series analysis [4, 6, 7, 15, 37].

Autoregressive (AR) models explain the current value x_t as a function of p past values in the time series. The number of past values p is also called model order, therefore an $AR(p)$ model is an autoregressive model of order p. *Moving average* (MA) models explain the current value x_t as a linear combination of the current white noise term and the q past white noise terms. The number of past white noise terms is again called the model order, therefore, an $MA(q)$ model is a moving average model of order q.

In some cases it is problematic to model a time series with only AR or only MA models, because it would demand a high-order model with many parameters. This is in conflict with the principle of parsimony. For these cases, Box and Jenkins [6] presented *autoregressive moving average* (ARMA) models. To achieve parsimony, ARMA models combine the ideas of AR and MA models. An *autoregressive moving average* $ARMA(p,q)$ model of order p and q, is the combination of an $AR(p)$ model part of order p and a $MA(q)$ model part of order q. It is possible to apply this class of models if the time series is stationary, which means that there is no seasonal effect or trend.

In many practical cases, a time series is non-stationary due to trends. It is possible to transform this time series to a stationary time series by applying a differencing operation, sometimes called detrending. To recover the original time series, the differenced time series needs to be aggregated, or also called integrated. These models are called *autoregressive integrated moving average* (ARIMA) models. An $ARIMA(p,d,q)$ means that the dth difference of the time series is an $ARMA(p,q)$ model.

To include seasonal terms in a model it is necessary to combine an ordinary non-seasonal ARIMA model with an ARIMA model that is extended to the seasonal period s. Therefore, the AR and the MA models are extended to the time shifts, called lag, of the seasonal period s and use capital letters P and Q for the seasonal model order. The seasonal difference D is also applied like the non-seasonal difference d, but with time shifts of the seasonal period s, called seasonal lag. Thus the additive seasonal effects are removed. The resulting models are called *seasonal autoregressive integrated moving average* (SARIMA) models and are denoted as $ARIMA(p,d,q) \times (P,D,Q)_s$.

After introducing the Box-Jenkins methodology in Section 3.1 and the class of models used in that methodology in this section, we discuss the separate steps of this iterative model selection process in the following Sections 3.3, 3.4, and 3.5.

3.3 Model Specification

For the task of model specification, the goal is to decide on a class of models that could be appropriate for the given time series, select the level of differencing and determine the order of the model which specifies the number of parameters used in the model. The first step to achieve this goal is to take a look at the given time series. Usually this is done by viewing the time series in a line plot. After applying all required transformations, such as a difference operation or log transformation, the *autocorrelation function* (ACF) and *partial autocorrelation function* (PACF) plots are checked to support the decision of the model order.
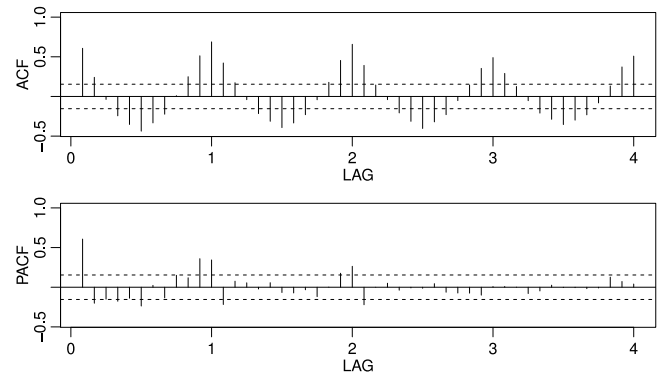
Fig. 3. ACF and PACF over Lags. The behavior of the lags enables domain experts to decide on the order of the model according to Table 1. This plot displays the example dataset (see Section 6) from the NMMAPS study [29, 33].

ACF/PACF Plot. The ACF plot is a spike graph, which is a special type of bar chart, of the ACF as a function over lags. The PACF plot is likewise the PACF as a function over lags. For the formal definition of the ACF and the PACF we refer again to the literature in time series analysis [4, 6, 7, 15, 37] and give a description of the basic ideas in the following: The ACF is the correlation between any two values in a time series with a specific time shift, called lag. The PACF is the correlation between any two points with a specific lag, where the linear effects of the points in between is removed. This PACF plot combined with the ACF plot, where this linear dependence is included, is called ACF/PACF plot and enables us to choose the number of parameters for the model. In addition to the time series plot, the ACF/PACF plot provides a first idea for the level of difference and seasonal difference. In Figure 3 we show an example ACF/PACF plot. The ACF and PACF are plotted on the y-axes and the lags on the x-axes. In this case the labels are seasonal lags, which means that one lag represents one seasonal cycle. The non-seasonal lags are fractions of one, depending on the seasonal length. For example, in a dataset with 12 months in one year and a seasonal length of 12, the seasonal lags are $1, 2, \ldots$ and the non-seasonal lags are $\frac{1}{12}, \frac{2}{12}, \ldots$.

Using the definitions of the autoregressive models, moving average models, ACF, and PACF it is possible to identify the basic behavior of the ACF and the PACF for AR, MA, and ARMA models [37]. Likewise, it is possible to describe the behavior for the seasonal component of the model in a similar way. The behavior is shown in Table 1. If we consider the seasonal lags $1, 2, \ldots$ in Figure 3, we notice that the ACF plot is tailing off and in the PACF plot cuts off at lag 2. According to Table 1 this indicates, that an adequate model for the seasonal component could be an $AR(2)_{12}$ model. Continuing this for the non-seasonal lags, we get a set of possible adequate models.

Table 1. ACF and PACF Behavior for ARMA and Seasonal ARMA Models [37]. The behaviors of the ACF and the PACF indicate which class of model and what number of parameters could be adequate for the non-seasonal and seasonal part of the model.

	AR(p)	**MA**(q)	**ARMA**(p,q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

	AR(P)$_s$	**MA**(Q)$_s$	**ARMA**(P,Q)$_s$
ACF*	Tails off at lags $k \cdot s$	Cuts off after lag Qs	Tails off at lags $k \cdot s$
PACF*	Cuts off after lag Ps	Tails off at lags $k \cdot s$	Tails off at lags $k \cdot s$

* where $k \cdot s$ are multiples of s, for $k = 1, 2, \ldots$

3.4 Model Fitting

In the previous section, we discussed how we select and configure the model order. The result is a model, for example a seasonal ARIMA$(p,d,q) \times (P,D,Q)_s$ model, where the level of difference d, the level of the seasonal difference D, the seasonal length s, the number of parameters p and q, as well as the number of seasonal parameters P and Q are set according to the steps presented above. Note that the parameters p, q, P, and Q determine the order of the model, which is the number of parameters. The differences d and D are transformations to the time series. Box et al. [7] use the term *tentatively entertained model* for this. Once the model is identified, it is fitted to the time series data to estimate the unknown parameters of the model. There are several methods to estimate the parameters. The most important one is the *maximum likelihood-estimation*. Other methods are the *method of moments*, the *least squares estimation*, and the *unconditional least squares*. For details and theoretical discussion we refer to [37, p. 121–140].

3.5 Model Diagnostics

To evaluate how well the model represents the underlying time series, model diagnostic methods are applied. The model is diagnosed by analyzing the residuals, which means the remaining part that is not explained by the model. The exploratory analysis of the residuals is done by plots as shown in (4a-d) of Figure 5. If the model is well fitted to the time series, the remaining part is expected to behave like white noise. This is assessed in four ways: (4a) The standardized residuals are plotted over time. Any non-random episodes can unveil a remaining underlying process. (4b) The ACF of the residuals is calculated and plotted over the lags to check that there is no remaining structure in the residuals. (4c) If the model is well fitted, the standardized residuals are expected to be standard normally distributed. This distribution is checked by the quantile-quantile plot [11]. (4d) The plot of the Ljung-Box statistic [8, 26] is a test that helps to check if the residuals for each lag are independent. If for all lags the p-values are not significant for a pre-specified significance level, indicated by the dashed line, it can be assumed that there is no remaining autocorrelation within the residuals. If all this is fulfilled, the model is well specified, otherwise the model needs to be readjusted. A clear decision is often not possible and is based on human judgement and experience.

Information Criteria. In addition to the diagnostic plots it is possible to examine information criteria. They provide a good basis to decide on the fitness of the model. Criteria that are often used include Akaike's information criterion (AIC), its bias corrected form, the AICc, and the Bayesian information criterion (BIC) [37]. In contrast to the AICc, which does behave very well for smaller samples, the BIC is well suited for larger samples. In order to get an adequate model, as introduced in Section 3.1, the goal is to select a number of parameters for the model, thus minimizing the criteria. Based on these values for different models, it is possible to decide on one of them tentatively. For more details on the information criteria we refer to [37, p. 50–53]. We show such information criteria in the graphical user interface of TiMoVA in area (5) of Figure 5.

According to our findings about the domain problem, related work, and expert feedback, we conclude that these tasks are currently cumbersome to execute by domain experts. It is evident, that combining the computations and visualizations with additional intuitive interactions and visual feedback improves the way to accomplish these tasks and supports the domain experts in their work. To achieve this, we analyzed the requirements and present them in the following section.

4 REQUIREMENTS ANALYSIS

In this section we explicitly describe our target users, distill the tasks and challenges for model selection discussed in Section 3 and formulate them using user stories as well as present the data used.

Target Users. Our target users are domain experts in any field using time series analysis, for example biology, chemistry, epidemiology, economy, or environmental research. The users have to be knowledgeable in statistics and time series analysis. These skills are required

to interpret and understand the visual representations of the time series and the time series models as well as the model fitting and the selection criteria.

Tasks. User stories help to formulate the requirements in a way that is easy to use in discussions within the project team, with customers, or other stakeholders. User stories evolved from the extreme programming (XP) software development methodology [2] and have an important role in other lean and agile software development methodologies [12, 13]. Usually in the beginning only high-level goals and requirements with a broad coverage are known. These goals and requirements are formulated as user stories. Through the process of refinement, the high-level user stories are broken down iteratively to smaller user stories until they are very specific. The high-level user stories with the broad coverage are also called epics [12, 13].

We formulated and refined the user stories in the repetitive meetings of the project team and throughout the formative evaluation. We used these user stories to formulate the VA process and implement the prototype. The stories are written from the perspective of our target users. The high-level goal is formulated in the following epic:

> *As a domain expert (user), I want to find an adequate model for a given time series so that I can use that model for different purposes, e.g., forecasting.*

User stories are defined to get a more detailed understanding of the requirements [13]:

- *As a domain expert (user), I want to select a certain region in the time series so that I can use any subregion of the time series for the model selection step.*
- *..., I want to see all important visualizations of the time series and the model so that I can decide on the model and assess how well the model fits the time series with one glimpse.*
- *..., I want to adjust the model orders at the place where the visualization provides the information about these model orders so that I can intuitively find an adequate model.*
- *..., I want to include and exclude the seasonal components of the model and the seasonal parameter inputs so that I can compare the seasonal influence and if no seasonal components are needed, they do not distract me.*
- *..., I want to see how a new selected model compares to the previous model so that I can decide if one model is better than the other.*
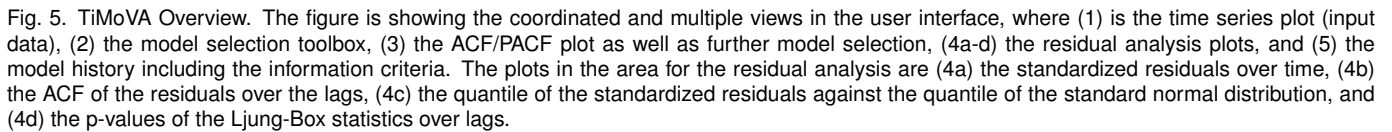
Time Series Data. The considered data for our work are time series as introduced in Section 1. We assume to have univariate data values observed at equidistant discrete time intervals without missing values.

5 VA FOR MODEL SELECTION IN TIME SERIES ANALYSIS

In Section 3 we discussed the characterization and tasks of the problem domain. We identified VA as a basis to define a VA process description to overcome the stated problems. In this section and the following Section 6 we rely on our findings to present the main contributions and results of our work. To do so, we provide the description of a tailored VA process in Section 5.1 that is used for the implementation of the prototype. In Section 5.2 we provide the final design and a discussion of the design choices and interaction techniques in the prototype. These results are designed and implemented to fulfill the requirements specified in Section 4.

5.1 VA Process Description for Model Selection

The basis of our VA process description for model selection is the Box-Jenkins methodology presented in Figure 2 and discussed in Section 3.1. The process description we show in Figure 4 is the application of the VA process for time-oriented data [25] on the domain problem of model selection in time series analysis. The goal of the process is to select an adequate model for a given time series. The details of the theoretical underpinnings of this process in statistical time series analysis are briefly discussed in Section 3. In the following we describe the VA process for model selection in detail to prepare the reader for the discussion in Section 5.2.3 about the connection between the graphical

Fig. 5. TiMoVA Overview. The figure is showing the coordinated and multiple views in the user interface, where (1) is the time series plot (input data), (2) the model selection toolbox, (3) the ACF/PACF plot as well as further model selection, (4a-d) the residual analysis plots, and (5) the model history including the information criteria. The plots in the area for the residual analysis are (4a) the standardized residuals over time, (4b) the ACF of the residuals over the lags, (4c) the quantile of the standardized residuals against the quantile of the standard normal distribution, and (4d) the p-values of the Ljung-Box statistics over lags.

user interface of the prototype (Figure 5), the VA process description for model selection (Figure 4) and the Box-Jenkins methodology (Figure 2).
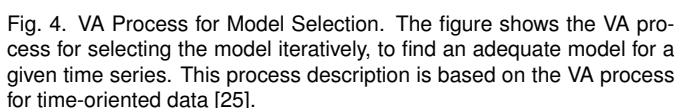
The time series in Figure 2 is *Data* provided as *Input* in Figure 4. The *Domain Knowledge* is based on experience and *Prior Analyses*. The *Interactive Visual Interface* is used to visualize the *Data* (D_i) to decide on the class of models and adjust the number of parameters as well as the level of differencing. To interpret the *Interactive Visual Interfaces*, the *Domain Knowledge* about time series analysis (K_t) and about visualizations of time series and time series models (K_p) is used. Based on this knowledge and the visual representations of the time series and time series model, the *Hypotheses* are formed (V_t). By adjusting the level of differencing (A_d) and the order of the model (A_p), the *Hypotheses* are refined. Based on the *Hypotheses* the model is estimated with the given parameters (B_m) to build a model based on the

input *Data* (A_i). The resulting model is analyzed using the *Domain Knowledge* about time series models and model diagnostics (K_m) as well as the visualizations of the standardized residuals, information criteria, and model parameters (V_d). In this iterative refinement of the process, *Insights* are gained by (1) interpreting the *Interactive Visual Interfaces* (I_v) deciding the fitness of the underlying model that is visualized, (2) the parameter estimations which lead to the adequate model (I_m), and (3) the refinement process of the *Hypotheses* building (I_h). The result is a model with estimated parameters, that is adequate for the given time series, and can be used for forecasting. The *Area of User Interaction* is highlighted in gray and indicates the process steps, where the user is part of the process through user interaction.

5.2 TiMoVA Prototype

Based on the VA process description above and the user stories (Section 4) we implemented our TiMoVA prototype. We refined our design along with the formulation of the user stories to quickly develop a working prototype and acquire user feedback [35]. Here we present the final version.

5.2.1 Implementation Choices

Before we discuss the design choices and implementation ideas in detail, we highlight the most important design decisions for our prototype. We implemented the prototype in Java. For the dynamic visualizations, we used the software framework prefuse [18]. As an API for time-oriented data we used TimeBench, a software library for time-oriented data [31].

One important decision was to use the R project for statistical computing [30] as a comprehensive toolkit for time series analysis and other calculation tasks. R provides a broad variety of methods known from literature and our prototype is designed in a way that allows us to use these methods for the statistical computations. Using Java/R Interface (JRI) enables us to use R in combination with Java. JRI is part of the rJava package [41] in R. We chose Java, because with prefuse, we have more possibilities regarding interactivity than implementing in R. Furthermore the extensibility and interconnectivity of our other projects using TimeBench is given.

Because the calculations in time series analysis can be very time-consuming, especially with large datasets, it is important that the user

Fig. 4. VA Process for Model Selection. The figure shows the VA process for selecting the model iteratively, to find an adequate model for a given time series. This process description is based on the VA process for time-oriented data [25].

interface is still responsive to user input, while calculations are carried out. This is achieved by using Java threads and caching, which allow the computer to pre-compute models and provide them upon request. As a result, the user interface shows good reaction times for user input, even when the calculations are running in the background.

5.2.2 Graphical User Interface

The graphical user interface of the prototype is based on the workflow of the VA process description we defined in Section 5.1. The visualizations are inspired by the visualizations used in R and by well-established visualization techniques [1, 11]. We extended these visualizations so that the user is able to interactively select the models. The result is a prototype that implements the VA process for model selection in time series analysis.

The TiMoVA prototype consists of coordinated and multiple views [32]. An overview of the graphical user interface and the five areas is shown in Figure 5 and in the supplementary video of the usage scenarios in Section 6.1. In Figure 5, (1) displays the time series plotted over time (time series plot). In this view it is possible to explore the time series and select a certain range that is used for the model selection. This range selection is shown in the upper left corner (1) of Figure 1. The details of the interactions for the range selection are discussed in Section 5.2.4. The toolbox in area (2) and the ACF/PACF plot in area (3) are used for the model selection. In the ACF/PACF plot (3) the user can adjust the number of model parameters directly within the plot. The plots in area (4a-d) show the results of the parameter estimation as the plots for the residual analysis. The table in area (5) displays the model history including the information criteria.

In Figure 6 the model selection toolbox (2) and the ACF plot (3) are shown in more detail. These are the areas for the configuration of the model order. In the toolbox the *max lag* input changes the number of lags in the ACF/PACF plot below and in the ACF plot of the residuals in area (4b). The *Include Seasonal Parameters* check box enables or disables the configuration of the seasonal component in the model, which also enables or disables the input for the *Seasonal Span*, as well as the *Seasonal Difference* slider. With the *Difference* slider and the *Seasonal Difference* slider the numbers for the parameter d and seasonal parameter D are selected. The continuous vertical lines in Figure 6 can be dragged along the x-axis to select the order of the model, which is synonymous with the number of parameters. There is one vertical line for p, which is the order of the autoregressive part of the model $AR(p)$. There is another vertical line for q, which is the order of the moving average part of the model $MA(q)$. If the seasonal components are enabled by the check box, two additional continuous vertical lines appear, one for P, which is the order of the autoregressive part of the seasonal component of the model $AR(P)_s$, and another one for Q, which is the order of the moving average part of the seasonal component of the model $MA(Q)_s$. The seasonal span s can be adjusted using the *Seasonal Span* spin box in the toolbox.

TiMoVA shows visual representations for the model diagnostics (4a-d) and (5), in order to evaluate the fitness of the model for the given time series. In this area the results of the parameter estimation are shown. We discussed the model diagnostics and the visual representations used for this task in Section 3.5. The goal is to visually explore the remaining part of the time series that is not described by the model, and check if it is likely to be white noise. The visual representation of the standardized residuals in the user interface of the prototype is inspired by the representation used in R. In Figure 5 the area displaying the plots for the analysis of the residuals are numbered with (4a-d). See Section 3.5 for the details on these plots.

In addition to the residual analysis, we included the information criteria, as introduced in Section 3.5, in the design of TiMoVA. The information criteria table (5) in Figure 5 shows a history of previously selected models. The first column represents the color used in the transitions of the residual plots. The second column describes the model and the other columns show the values of the model information criteria. The cells of the model criteria are colored according to their value indicating whether this criterion for this model is better (minimum) or worse compared to the others in the model history. The legend is
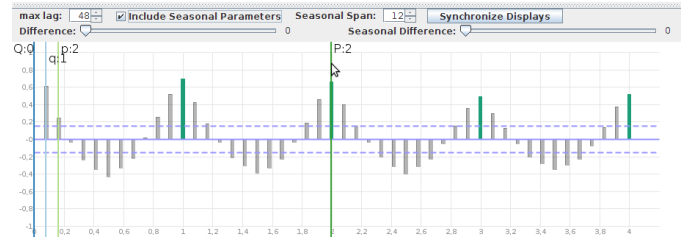

Fig. 6. Model Selection Toolbox and ACF Plot. The toolbox at the top and the four continuous vertical lines are used to select the time series model. In this figure they are set to $p = 2$, $q = 1$, $Q = 0$, and P is currently moved from $P = 1$ to $P = 2$, which is the final model configuration for this time series. The user interface supports the user to focus on the seasonal lags by changing the color and reducing the opacity at the non-seasonal lags.

shown below this table.

Residual analysis and tests for white noise, which are essentially tests for the randomness of a dataset, are manifold in statistics and there are many implementations of these methods in R. In our implementation of the prototype we focused on the standard tests and visualizations used in the literature for time series analysis [4, 6, 7, 15, 37]. It is desirable to enable the user to adjust and customize which tests and visualizations she or he wants to use in the process. This is a possible feature to include in the future work.

5.2.3 Connecting the VA Process Description and the Box-Jenkins methodology

In this paragraph, we describe how the VA process description defined in Section 5.1 is implemented in TiMoVA. We explain in detail how the user interface facilitates the VA process and creates short feedback cycles for the task of model selection. For each transition in the process, we provide the corresponding labels from Figure 4, the number of origin from the original Box-Jenkins methodology in Figure 2, and the number of the affected area in the user interface in Figure 5. By viewing the plots in the user interface, we decide on a general class of models (Figure 2: (1); D_i, V_t). By adjusting the level of difference and the number of model parameters (Figure 5: 2, 3; A_d, A_p, V_t), we identify a so called tentatively entertained model (Figure 2: (2); B_m). The adjustment of the relevant faders triggers the system to estimate the parameters of the model (Figure 2: (3); A_i) and show the resulting diagnostics immediately in the user interface (Figure 5: 4a-d, 5; Figure 2: (4); V_d, D_i, V_t). The insights gained (I_h, I_v, I_m) and the application of the domain knowledge (K_t, K_p, K_m) are part of the user interaction, but not part of the user interface.

5.2.4 Interactive Guided Model Selection Environment

In Section 5.2.2 we introduced the time series plot and the range selection as shown in the upper left corner (1) of Figure 1. The main interaction in this area is the navigation through the time series and the selection of a specific time interval. The horizontal range slider on the bottom allows the user to zoom in and navigate through the time series. When changing the zoom level on the range slider, the time axis is adjusted to show a suitable resolution of time. Details about the time points are provided on demand when moving the mouse cursor to its position. It is possible to specifically select a time interval that is used for the model selection. The user can select, resize, move, and remove the selection using the mouse cursor. The user can select whether or not the selection is connected to the other views using the *Synchronize Displays* button shown in Figure 6. If it is linked to the other views, they are recomputed and the visualizations are updated as soon as the region changes. This feature enables the user to select a certain smaller region of interest from a larger time series for the model selection task and keep a fast reaction time for the model parameter estimation even for larger time series.

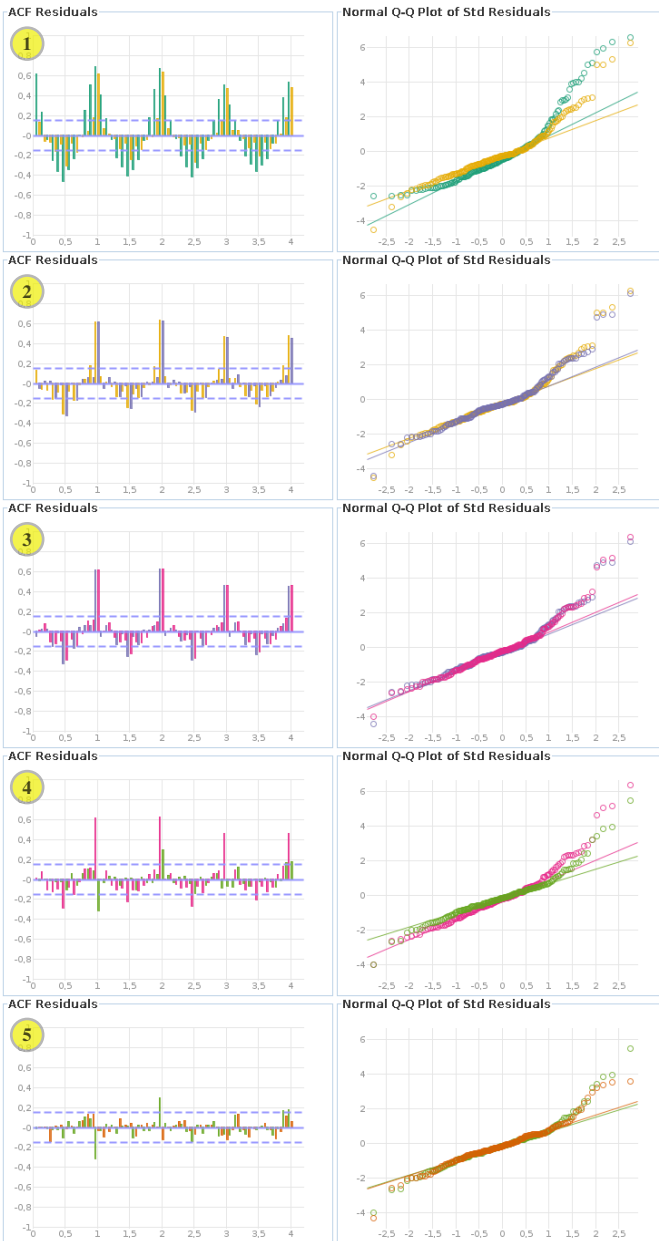Another important design requirement was to visualize the change

Fig. 7. Transitions of Model Selection in Elected Residual Plots. We show the change when selecting a new model in two elected residual visualizations, the ACF and the normal quantile-quantile plot of the standardized residuals. This enables the user to evaluate whether or not the new model improves. Each numbered row represents one transition.

in the plots when adjusting the model order and give the user the control over these transitions. This is supported by direct manipulation [36] using sliders for the level of difference and continuous vertical lines for the model parameters. To focus on the change in the resulting plots, we use animated transitions [19] and different colors that are consistent in the coordinated and multiple views [32]. This process is shown in Figure 7 and the supplementary video. Once the slider or a vertical line is dragged from one value in the direction of the next value, the parameters of the new model are calculated and the plots are seamlessly faded from one display to another by using alpha blending of the bars, points, and lines. The colors for the plots are selected by using ColorBrewer2 [17]. We used this tool to get a qualitative color scheme, which is easy to distinguish on a screen. This set of colors is used as an endless cyclic sequence for the coloration of the plots. This ensures that each parameter combination is a different color, and the

fading process uses always two separate colors.

The toolbox to adjust the parameters and the continuous vertical lines in the ACF/PACF plot are shown in Figure 6. By default the check box to include the seasonal parameters is disabled. In this case, the seasonal span and the seasonal difference input are disabled and the vertical lines for the order of the seasonal autoregressive and the moving average component of the model are not visible. This ensures that the user does not accidentally fit a seasonal model, if a non-seasonal model is needed. By ticking the check box the inputs are enabled and the vertical lines in the ACF/PACF plot for the seasonal order appear. That also ensures to first consider simpler non-seasonal models according to the principle of parsimony and keep the users attention to the relevant class of models.

In Figure 7, we show the fading process when changing the order of the model by dragging one of the continuous vertical lines in the ACF/PACF plot. When sliding the vertical lines for the seasonal order P and Q in the ACF/PACF plot shown in Figure 6, the prototype supports focusing the seasonal lags in the ACF/PACF plot by setting the non-seasonal lags to another color and opacity level. Thus the user can more easily decide the seasonal order of the model. When adjusting the level of differencing we also change the underlying data for the ACF/PACF plot. In this case the ACF/PACF plot and the residual plots are fading from the current to the new configuration.

All four residual analysis plots, area (4a-d) in Figure 5, are included in the interactive fading process presented before. When the user modifies the model configuration, the residual plots are fading from one to the other continuously. This enables the user to see the changes of the model configuration and evaluate if the model fitness improves or worsens. This process is shown in Figure 7.

The information criteria for all previously and currently selected models are shown in area (5) of Figure 5 as described in Section 5.2.2. This history stack is filled during the model selection process. The coloring to immediately find the minimal information criteria is readjusted if a new model is added to the table. So for each transition we can see the values of the criteria which are supported by the color if it is better or worse than the previous one. Additionally, it is possible to see which are the best models according to the information criteria at each point in the model selection task.

5.2.5 Discussion of Design Rationales

The key idea for the layout of the TiMoVA user interface is to map the Box-Jenkins methodology and its workflow. This way of working and thinking is well-established and known by our target users, the domain experts. The main intention to use established standard visualization techniques for the separate steps in the Box-Jenkins methodology, is to avoid confusion and benefit from the experience the domain experts already have. By using familiar visualizations that they know really well, the domain experts can profit from the combination, layout, and especially the interactions of TiMoVA. In this stage of our work this was our goal and is our contribution. For future work it would be exciting to experiment with more advanced visualization techniques, use and include them in TiMoVA in order to further improve the task of model selection in time series analysis.

Another requirement was to use appropriate interactions in TiMoVA, so that it is easy and intuitive for the target users and they can concentrate on their task of model selection. According to Heer and Robertson [19] animated transitions in statistical data graphics are a way to engage the users and improve the perception of changes. They also suggest using alpha blending as a solution for possible occlusion. Because of their findings, we apply animated transitions in TiMoVA and use alpha blending because occlusions may occur in the transitions. These transitions are triggered and steered directly by the user inside the ACF/PACF plot.

One limitation of showing the transitions with animated diagnostic plots is that you actually compare the current model to one of the next possible models. Usually this is good enough because you immediately see if the more specific sub-model you think of is better or worse and by this preview the domain expert can decide if it is worth to go into this direction. To overcome the limitation of comparing only

two models we provide a history table with all previously and currently considered models and the corresponding information criteria as shown in area (5) of Figure 5 and in more detail in the supplementary material. With this overview it is possible to compare more than two models according to their information criteria. In our design we planned for future work to enable the user to select up to three models and load their standardized residuals in the diagnostic plots to directly compare them visually. This should be sufficient, because according to Nazem [28, p. 307] in about 87% of the time series only one or two models remain in the shortlist for adequate models and in about 97.6% three or less models remain.

6 EVALUATION

During the design and implementation phase we evaluated the results by formative evaluations during repetitive meetings of the design team. This iterative refinement process was judged by the team members, experts in information visualization and statistics, and the user experience [24] was discussed by performing demonstrations. In addition to this repetitive internal assessment, the user experience was evaluated by informal user feedback [24] consulting two external domain experts. Besides this first level of evaluation, we evaluated the prototype by defining usage scenarios and applying the prototype on an example dataset. We explain the example dataset below and apply the usage scenarios in Section 6.1. We discuss the insights gained from both levels of evaluation, what we learned, and how we can improve our solution further in Section 6.2. Accordingly we assess the usability and applicability of our solution for its target users.

Example Dataset. In Section 3 we introduced the example dataset of the daily number of deaths from cardiovascular disease from the NMMAPS study [29, 33]. The original dataset contains the data of different cities in the United States of America, but we focused on the number of cardiovascular disease deaths in Los Angeles only. The relevant columns in the dataset are *date* and *cvd*, which is the daily number of deaths from cardiovascular disease. There are no missing values in the dataset. For the evaluation of the prototype, we aggregate the time series to get monthly sums.

6.1 Usage Scenarios

Based on the requirement analysis in Section 4 we used the user stories and defined two usage scenarios for the evaluation. The first one is the high-level task of model selection. The second one is the task of selecting a range in the time series before selecting a model. We describe how the prototype is applied on the example dataset to solve these tasks. We demonstrate the usage scenarios from the perspective of a fictional domain expert. Because it is difficult to show the interactivity and visual feedback in static pictures, we support the textual description with the transitions in two of the residual plots in Figure 7. To clarify the interactions and how all visual representations behave during the transitions, we provide a supplementary video[1] with audio narration. Another supplementary material shows the model diagnostic area from TiMoVA for each of the five transitions that we discuss in the following.

6.1.1 Model Selection

Following the Box-Jenkins methodology presented in Section 3.1, we first consider the time series plot and the autocorrelation function (ACF) and partial autocorrelation function (PACF) plot. According to the time series plot in TiMoVA (Figure 5), we assume that no differencing operation (see Section 3.2) may be needed, because the time series seems to be stationary. Moving the difference slider confirms that the change in the ACF/PACF plot, as well as the residual analysis plots is marginal and, therefore, supports this hypothesis.

We evaluate the ACF/PACF plot, (3) in Figure 5, according to the behavior of the non-seasonal order of the model in Table 1. We show the transitions of this usage scenario in Figure 7 beneath each other with the color code explained in the information criteria table in area (5) of Figure 5. For the non-seasonal component of the model we

---

[1] http://www.cvast.tuwien.ac.at/TiMoVA (July 30, 2013)

decide to have a mixed ARMA model. When sliding the parameter p, the diagnostic plots shows that the adjustment of the non-seasonal AR model to order $p = 1$ results in a more random appearance of the residual time series plot, a more straight line behavior in the normal quantile-quantile plot, and lower lags in the non-seasonal lags of the ACF plot (Transition 1 in Figure 7). In addition to further improvement of those residual plots, the Ljung-Box statistics shows more lags with p-values significantly different from zero, if the order is changed to $p = 2$ (Transition 2 in Figure 7). See the supplementary image and video for further details on this. For both transitions we can see in area (5) of Figure 5 how the information criteria improve. The result is the model configuration $p = 2$, which is an AR(2) model. By adding a MA component of order $q = 1$ to the model (Transition 3 in Figure 7) we get the diagnostic plots for the assumed mixed ARMA model with $p = 2$ and $q = 1$. This configuration advances the model to show more randomness in the diagnostic plots of the standardized residuals, which strengthens the assumption that they are standard normal distributed. The information criteria in area (5) of Figure 5 get minimal worse when adding the MA component $q = 1$, but according to the behavior of the ACF/PACF plot in Figure 5 we assume a mixed ARMA model for the non-seasonal part. To see more details of these transitions, we provide all of them for the diagnostic plots in a supplementary image and in the supplementary video.

The seasonal behavior is not covered by the model yet. Therefore, the next step is to adjust the seasonal parts of the model. We consider an autoregressive model, because sliding the parameter P highlights the seasonal lags and unveils the cut off on seasonal lag 2 in the PACF. This indicates, when consulting Table 1 for the behavior of the seasonal order of the model, that the seasonal component is likely to have order $P = 2$. Sliding from seasonal order $P = 0$ to $P = 1$ (Transition 4 in Figure 7) and from $P = 1$ to $P = 2$ (Transition 5 in Figure 7), shows the improvement of the selected model. In area (5) of Figure 5 we also recognize the abrupt improvement in the information criteria when including the seasonal component with $P = 1$ and the gradual improvement for the transition from $P = 1$ to $P = 2$. With this configuration, we get a seasonal model of the following form:

$$\text{ARIMA}(p = 2, d = 0, q = 1) \times (P = 2, D = 0, Q = 0)_{s=12}$$

Including the estimated parameters of the model, we get the following time series model:

$$(1 - 0.3068B^{12} - 0.5444B^{24})(1 + 0.3143B - 0.3112B^2)x_t = (1 - 0.9072B)w_t$$

In addition, the information criteria in area (5) of Figure 5 indicate that we found an adequate model. We selected the model with the minimum values for the AIC, AICc and BIC using TiMoVA. This goes along with the diagnostics based on the visualization of the standardized residuals.

6.1.2 Range Selection

TiMoVA enables the user to select a range of the time series, as we show in the upper left corner (1) of Figure 1. A possible usage scenario is to select a trend that starts and ends at a defined point in the time series and to consider only this trend for model selection.

In the following usage scenario we want to consider only the range starting with November 1994 and ending with the last data point in the time series. We select the model as described in the previous section. Compared to the complete time series we get a slightly different model, which is simpler and with fewer parameters. We get an intermediate model with $p = 2$ and seasonal $P = 1$. The ACF plot of the residuals indicates a remaining seasonal autocorrelation. In this case we consider a seasonal difference of $D = 1$ as a possible solution to remove this autocorrelation in the residuals. Sliding the seasonal difference fader, improves this model. Finally, we get a model with the following configuration:

$$\text{ARIMA}(p = 2, d = 0, q = 0) \times (P = 1, D = 1, Q = 0)_{s=12}$$

Consulting the information criteria for this model supports that we have found an adequate model.

In addition to the insights gained from the application of TiMoVA in

these two usage scenarios, we discuss the results of the user feedback in the following section.

6.2 Evaluating User Experience

For the evaluation of the user experience [24] we rely on the insights gained from the internal formative evaluation and iterative design during the design and implementation phase, where we had repetitive meetings and discussions on the intermediate stages, and on the demonstration session with two external domain experts who are employed as scientists in an institution for statistics research. They both hold a master in mathematics and one a PhD in statistics. The informal evaluation [24] was performed by demonstrating the prototype with a well-known dataset for time series analysis. The domain experts gave immediate feedback to the features of TiMoVA. This feedback was noted on paper and reflected after the demonstration session, which lasted about one hour. We included the feedback in the design and implementation. We discuss the remaining suggestions in this section and consider them for future work in Section 7. In this reflection our findings from the usage scenarios are included as well.

A very useful feature according to the domain experts is the overview displaying all separate steps of the Box-Jenkins methodology. It provides all information necessary to decide on an adequate model. Because the model selection relies on the behavior of the ACF/PACF plot, it is very helpful to directly select the model order inside this ACF/PACF plot using the continuous vertical lines. The visual support of focusing on the seasonal lags for the selection of the seasonal component in the model was considered very beneficial. Another beneficial feature according to the domain experts is the immediate visualization and preview of the model results when changing the model order and especially the visualization of the transition from one model to another. This enables the domain experts to directly compare the current model to the new model and decide whether the model improves or not. A further benefit is the possibility to select a certain region of interest from a larger time series and consider only this subregion for the model selection task. This is not only very nice for selecting a representative or interesting subregion, but also to have faster reaction times even for originally larger time series.

There are also suggestions to further improve our work. The domain experts would like to see more statistics of the residuals on demand. The numbers are available in memory as a result of the computations and including them in the graphical user interface is going to be future work. Domain experts sometimes prefer to perform different tests and statistics for the residual analysis and would like to customize the selection from a set of test statistics using the graphical user interface. Our solution is prepared for this kind of request and implemented in a way to ensure exchangeable test statistics.

One suggestion during the iterative design and implementation phase was to include a history of the previously selected models and enable the user to get an overview and reload certain models. We included this concept in the design and the result is shown in area (5) of Figure 5. This feature of getting an overview on the previously selected models and compare more than two models on a more abstract level is very beneficial. The concept of loading certain models as an overlay in the diagnostic plots is considered useful and a way to further improve the implementation.

Another feature demanded, is to show how the model performs. This is usually done by taking the first part of the time series and use the model to forecast a certain time range or repetitive single step ahead forecasts. This forecast is then shown along the given time series with the according confidence boundaries. At the moment we focused on the first part in time series analysis, which is finding an adequate model, and not the application of this model. Although the extension of using these forecasts as overlay in the time series plot is considered for further work.

What we learned from the evaluation is that for the target users, with at least a basic knowledge in statistics and time series analysis (ARIMA models), TiMoVA is easy to learn and understandable. For others it is necessary to study time series analysis in order to understand and interpret the visualizations. We achieved this in TiMoVA, as

discussed already in Section 5.2.5, by focusing on the knowledge and experience the domain experts already have and support them with a well-established way of working and thinking, familiar visualizations, and appropriate interactions.

7 CONCLUSION AND FUTURE WORK

The goal of our work is to use VA to support domain experts in the process of model selection. We identified that for the class of ARIMA and seasonal ARIMA models in the Box-Jenkins methodology there is no technique or tool that supports the workflow of the process in an intuitive and user-friendly way. Applying VA methods to this domain problem showed that we can support this task with interactive visual interfaces, short feedback cycles, and the visualization of the model transitions.

By evaluating our work, we discovered that the resulting VA process description and the TiMoVA prototype enables the domain expert to do easy and intuitive visual exploration and selection of time series models. These benefits are achieved by

- enabling the domain expert to select the model order interactively via the visual interface, inside the ACF/PACF plot, which provides a first idea of the model order,
- giving the domain expert immediate visual feedback of the model results while selecting the model order, and
- helping domain experts with the visualization of the model transitions to decide whether or not the model improves.

We also showed that the interactions are appropriate for the task and that the domain experts profit from the usage of a well-established model selection methodology and visualizations from their domain.

In Section 6.2 we discussed the insights gained from the evaluation of our results and developed ideas for future work. One improvement was to include information criteria measures in the graphical user interface in a history. Another idea for future work is to enable using this history to load any previous model and compare two of them. Further improvement is to extend the prototype to directly support different statistical methods for the residual analysis and enable the user to customize the residual analysis for their needs. The diagnostic of the time series model is currently limited to the diagnostic plots. For future work, it would be interesting to include the performance of the model for forecasting the diagnostic step.

Our solution is limited to data with equally spaced time series without missing values. In practical applications, the data often contains missing values and/or are not equally spaced. For future research we consider these limitations as inspiration to apply VA for model selection in time series with missing values and provide visual support for the methods to estimate them. Another interesting challenge for future work is to foresee the model selection support for multivariate time series. Therefore, it is necessary to consider appropriate visualization techniques for this kind of data.

The final model we found in our usage scenarios is rated as a rather complex model by our domain expert. Finding this model using existing statistical software tools would have been very cumbersome and time consuming. Working with TiMoVA reduced the number of models considered, because the immediate visual feedback excluded already certain subclasses of models early in the model selection process.

Based on the insights from the evaluation, we discovered that the well-established visualizations used in the prototype have the benefit that the domain experts are used to work with these visual encodings. Therefore, they can focus on the task of model selection, which is guided by TiMoVA and improves their work.

REFERENCES

[1] W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of Time-Oriented Data*. Springer, London, UK, 2011.

[2] K. Beck. *Extreme Programming Explained: Embrace Change*. Addison-Wesley, Reading, MA, USA, 2000.

[3] J. Bernard, T. Ruppert, O. Goroll, T. May, and J. Kohlhammer. Visual-interactive preprocessing of time series data. In A. Kerren and S. Seipel, editors, *Proc. SIGRAD, Interactive Visual Analysis of Data, Växjö, Sweden*, volume 81, pages 39–48. Linköping University Electronic Press, 2012.

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

[5] G. E. P. Box and N. R. Draper. *Empirical Model Building and Response Surfaces*. John Wiley & Sons, NY, USA, 1987.

[6] G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, CA, USA, 1970.

[7] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Hoboken, NJ, USA, 4th edition, 2008.

[8] G. E. P. Box and D. A. Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332):1509–1526, 1970.

[9] P. Buono, A. Aris, C. Plaisant, A. Khella, and B. Shneiderman. Interactive pattern search in time series. In *Proc. Conf. Visualization and Data Analysis*, pages 175–186, 2005.

[10] 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. Conf. Information Visualisation (IV)*, pages 191–196. IEEE Computer Society, 2007.

[11] W. Cleveland. *Visualizing Data*. Hobart Press, Summit, NJ, USA, 1993.

[12] M. Cohn. *User Stories Applied: For Agile Software Development*. Addison-Wesley, Redwood City, CA, USA, 2004.

[13] M. Cohn. *Succeeding with Agile: Software Development Using Scrum*. Addison-Wesley, Upper Saddle River, NJ, USA, 2010.

[14] A. Cottrell and R. Lucchetti. *Gretl: GNU Regression, Econometric and Time-series Library - User's Guide, Version 1.9.12*, 2013. http://gretl.sourceforge.net (cited Mar 20, 2013).

[15] J. D. Cryer and K.-S. Chan. *Time Series Analysis. With Applications in R*. Springer Texts in Statistics. Springer, New York, NY, USA, 2nd edition, 2008.

[16] J. D. Hamilton. *Time series analysis*. Princeton University Press, Princeton, New Jersey, USA, 1994.

[17] M. Harrower and C. Brewer. Colorbrewer.org: An online tool for selecting colour schemes for maps. *Cartographic Journal*, 40(1):27–37, 2003.

[18] 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)*, pages 421–430. ACM, 2005.

[19] J. Heer and G. G. Robertson. Animated Transitions in Statistical Data Graphics. *IEEE Trans. Visualization and Computer Graphics*, 13(6):1240–1247, 2007.

[20] IHS Global, Inc. *EViews, Version 8*. Irvine, CA, 2013. http://www.eviews.com (cited Mar 20, 2013).

[21] D. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, editors. *Mastering The Information Age - Solving Problems with Visual Analytics*. Eurographics, Goslar, Germany, 2010.

[22] A. Kowarik and A. Meraner. *x12: X12 - wrapper function and structure for batch processing*, 2012. R package version 1.0-3. http://CRAN.R-project.org/package=x12 (cited Mar 12, 2013).

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

[24] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Trans. Visualization and Computer Graphics*, 18(9):1520–1536, 2012.

[25] T. Lammarsch, W. Aigner, A. Bertone, S. Miksch, and A. Rind. Towards a concept how the structure of time can support the visual analytics process. In S. Miksch and G. Santucci, editors, *Proc. Int. Workshop Visual Analytics (EuroVA) in conjunction with EuroVis*, pages 9–12. Eurographics, 2011.

[26] G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.

[27] T. Munzner. Process and pitfalls in writing information visualization research papers. In A. Kerren, J. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization*, volume 4950 of *LNCS*, pages 134–153. Springer, 2008.

[28] S. M. Nazem. *Applied Time Series Analysis for Business and Economic Forecasting*. Statistics: Textbooks and Monographs. Marcel Dekker, Inc., NY, USA, 1988.

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

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

[31] 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 Trans. Visualization and Computer Graphics*, 19(12), 2013. forthcoming.

[32] J. C. Roberts. State of the art: Coordinated & multiple views in exploratory visualization. In *Proc. Conf. Coordinated and Multiple Views in Exploratory Visualization (CMV)*, pages 61–71. IEEE Computer Society, 2007.

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

[34] SAS Institute Inc. *JMP 10 Modeling and Multivariate Methods*. SAS Institute Inc., Cary, NC, 2012. http://www.jmp.com (cited Mar 20, 2013).

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

[36] Shneiderman. Direct Manipulation: A Step Beyond Programming Languages. *IEEE Computer*, 16(8):57–69, 1983.

[37] R. H. Shumway and D. S. Stoffer. *Time Series Analysis and its Applications. With R examples*. Springer Texts in Statistics. Springer, New York, NY, USA, 3rd edition, 2011.

[38] StataCorp. *Stata Statistical Software: Release 12*. StataCorp LP, College Station, TX, 2011. http://www.stata.com (cited Mar 20, 2013).

[39] The MathWorks Inc. *MATLAB (R2009a)*. Natick, MA, 2010. http://www.mathworks.com (cited Mar 20, 2013).

[40] J. J. Thomas and K. A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society, Los Alamitos, CA, USA, 2005.

[41] S. Urbanek. rjava: Low-level r to java interface, 2011. R package version 0.9-3. http://CRAN.R-project.org/package=rJava (cited Mar 14, 2013).

[42] Wolfram Research, Inc. *Mathematica, Version 9.0.1*. Champaign, IL, 2013. http://www.wolfram.com/mathematica (cited Mar 20, 2013).