Proceedings of the 2009 Winter Simulation Conference
M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, eds.

A COMPARISON OF MARKOVIAN ARRIVAL AND ARMA/ARTA PROCESSES FOR THE MODELING OF CORRELATED INPUT PROCESSES

Falko Bause
Peter Buchholz
Jan Kriege

Informatik IV
TU Dortmund
D-44221 Dortmund, GERMANY

ABSTRACT

The adequate modeling of input processes often requires that correlation is taken into account and is a key issue in building realistic simulation models. In analytical modeling Markovian Arrival Processes (MAPs) are commonly used to describe correlated arrivals, whereas for simulation often ARMA/ARTA-based models are in use. Determining the parameters for the latter input models is well-known whereas good fitting methods for MAPs have been developed only in recent years. Since MAPs may as well be used in simulation models, it is natural to compare them with ARMA/ARTA models according to their expressiveness and modeling capabilities for dependent sequences. In this paper we experimentally compare MAPs and ARMA/ARTA-based models.

1 INTRODUCTION

One important step in the construction of a valid simulation model is the definition of an accurate input model. Usually input processes are described by independent and identically distributed (i.i.d.) random variables and methods for determining appropriate distributions reflecting specific characteristics, e.g. from measured data, are well known (Kelton and Law 2000). In a variety of applications, like computer networks, input processes can not be described adequately by i.i.d. random variables, since time-dependencies and correlations between events are not captured (Paxson and Floyd 1995, Crovella and Bestavros 1996, Kuhl et al. 2007). For the specification of stationary time-dependent input processes the use of AR (Auto Regressive), ARMA (Auto Regressive Moving Average) and ARIMA (Auto Regressive Integrated Moving Average) models has become common practice, since the work of (Box and Jenkins 1970). AR(I)MA models are linear and result in marginal normal distributions. One way to overcome this restriction is the ARTA (Auto Regressive To Anything) model (Cario and Nelson 1996) for which a flexible fitting algorithm (ARTAFIT) is available (Biller and Nelson 2008).

Considering the modeling of correlated interarrival times only few publications deal with the application of AR/-AR(I)MA/ARTA processes (which we will name AR* processes in the following) and their real use in a simulation model. Usually AR* models are employed as discrete time processes (Leemis 1998) and thus an obvious approach is to use AR* processes as count processes thus modeling the number of arrivals in a specific time interval. According to this approach (Xue et al. 1999) report about the application of F-ARIMA models. In simulation the use of such a counting process is often not sufficient for the modeling of an arrival process, since the accurate times of individual arrivals are not specified. This problem does not occur if the AR* process models the interarrival time process. (Tran and Reed 2004) is one of the very few publications reporting about the application of ARIMA models for the modeling of the interarrival time process.

Another class of processes capable of representing correlated data are Markovian Arrival Processes (MAPs). MAPs are also applied for the modeling of the interarrival time process (Kang et al. 2002, Horvath, Telek, and Buchholz 2005, Heindl, Mitchell, and van de Liefvoort 2006) and exhibit a geometrically decaying correlation structure which makes them at first glance less suitable for the modeling of correlated simulation input processes. A MAP itself is a Markov process where some transitions indicate an arrival event and its application is common in analytical and numerical models, since their integration keeps the Markov property. Similar to the situation of AR* processes only few references can be found on the application of MAPs in simulation models, see e.g. (Tartarelli, Pagano, and Devetsikiotis 2000).

Furthermore, even though there is a multitude of publications dealing with the fitting of on the one hand AR* processes and on the other hand of MAPs, we found no comparison of both kind of process types when being applied for simulation. In this paper we present results of some corresponding experiments. The paper is structured as follows. In the next section we define the type of arrival processes being considered. Section 3 briefly presents the used fitting methods. The main contribution of this paper is given in section 4 where we describe the goodness of the fitting methods being applied to synthetically generated and measured traces and the application of the fitted processes for simulation input modeling. The paper ends with the conclusions in section 5.

2 STOCHASTIC ARRIVAL PROCESSES

Stochastic models of the input for a simulation model have to capture the relevant behavior that is observed in reality. However, until today most simulation models are based on the implicit assumption of independent and identically distributed arrivals, although it is known that this assumption is often violated in practice (Paxson and Floyd 1995, Crovella and Bestavros 1996) since arrivals are correlated and the negligence of correlation can result in a serious loss of validity of the simulation model (Livny, Melamed, and Tsiolis 1993). Although models to describe dependencies in arrival processes are known for a long time (Box and Jenkins 1970) the adequate modeling and the application of the resulting input models in simulation is still a challenge (Biller and Ghosh 2004). The following requirements should be met by a model for arrivals describing correlated input sequences:

- It should adequately capture the marginal distribution of the interarrival time and the correlation structure between subsequent arrivals,
- parameters of the model should be easy to fit according to measured traces, and
- the input model should be easy to integrate into a simulation model.

Two large classes of models for describing input processes exist, namely the AR* processes and Markovian models. Both are briefly introduced in the following paragraphs.

2.1 AR and ARMA Processes

The simplest subclasses of ARMA processes are autoregressive (AR) processes of order p ($AR(p)$) and moving average (MA) processes of order q ($MA(q)$). An $AR(p)$ process is defined as (Box and Jenkins 1970)

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} + a_t$$

with the *innovations* a_t that are normally distributed with mean zero and variance σ_a^2.

While for an $AR(p)$ model z_t is expressed as the weighted sum of the p previous $z_i, i = 1,\ldots,p$, for a $MA(q)$ model z_t is constructed from q previous innovations $a_i, i = 1,\ldots,q$. Thus, the moving average process is defined as

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}$$

A combination of autoregressive and moving average processes results in $ARMA(p,q)$ models defined as

$$z_t = \phi_1 z_{t-1} + \ldots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$$

The autoregressive integrated moving average ($ARIMA(p,d,q)$) model adds to an ARMA(p,q) a third parameter d which describes how observed values are modeled (cf. (Box and Jenkins 1970)). If non-integer values are allowed for d, we finally obtain the fractional autoregressive integrated moving average (F-ARIMA) processes. These models are known to exhibit self-similarity as it can often be observed in network traffic (Sheluhin, Smolskiy, and Osin 2007).

ARIMA models are usually used as discrete-time processes (Leemis 1998) and hence the data from a trace is interpreted as a count process for ARIMA fitting. For example in (Xue et al. 1999) the number of arrivals in given intervals are counted and a F-ARIMA model is used to model the process. Nevertheless, ARIMA processes have also been used to model interarrival time processes in the past, e.g. in (Tran and Reed 2004) ARIMA processes are used for online prediction of the interarrival times of I/O requests.

We also experimented with ARIMA and F-ARIMA models, but found no significant improvements in the fitting of the considered traces. Hence we concentrate in this paper on AR and ARMA processes and on a third class of AR-based processes, namely ARTA processes.

2.2 ARTA Processes

ARTA (Auto Regressive To Anything) Processes (Cario and Nelson 1996) combine a base $AR(p)$ process with an arbitrary marginal distribution and thus can model correlated input processes with a wide variety of shapes for the distribution. They are defined by a marginal distribution F_Y and a base $AR(p)$ process and have the form

$$Y_t = F_Y^{-1}[\Phi(Z_t)], t = 1, 2, \ldots \qquad \text{with} \qquad Z_t = \alpha_1 Z_{t-1} + \alpha_2 Z_{t-2} + \ldots + \alpha_p Z_{t-p} + \varepsilon_t$$

where Φ is the standard normal cumulative distribution function and $\{Z_t; t = 1, 2, \ldots\}$ is a stationary Gaussian $AR(p)$ process. The variance of the series of $\{\varepsilon_t\}$ that are independent $N(0, \sigma_\varepsilon^2)$ random variables is set to $\sigma_\varepsilon^2 = 1 - \alpha_1 Corr(Z_t, Z_{t+1}) - \alpha_2 Corr(Z_t, Z_{t+2}) - \ldots - \alpha_p Corr(Z_t, Z_{t+p})$ such that the marginal distribution of the $\{Z_t\}$ is $N(0,1)$.
Then the probability-integral transformation $U_t = \Phi(Z_t)$ ensures that $U(t)$ is uniformly distributed on $(0,1)$ (cf. (Devroye 1986)) and the application of $Y_t = F_Y^{-1}[U_t]$ yields a time series $\{Y_t, t = 1, 2, \ldots\}$ with the desired marginal distribution F_Y. This approach works for any distribution F_Y, although F_Y^{-1} might have to be approximated by numerical methods in cases where no closed-form expression exists.

2.3 Markovian Arrival Processes

Markovian Arrival Processes (MAPs) have originally been developed as input processes in queuing systems that are solved analytically (Lucantoni, Meier-Hellstern, and Neuts 1990). However, they may as well be used in simulation models. A MAP of order n is described by two $n \times n$ matrices $\mathbf{D}_0$ and $\mathbf{D}_1$ such that $\mathbf{D}_1(i,j) \geq 0$, $\mathbf{D}_0(i,j) \geq 0$ for $i \neq j$ and $\mathbf{D}_0(i,i) = -\sum_{i \neq j} \mathbf{D}_0(i,j) - \sum_{i=1}^{n} \mathbf{D}_1(i,j)$. Furthermore, $\mathbf{Q} = \mathbf{D}_0 + \mathbf{D}_1$ is the generator matrix of an irreducible Markov process (Stewart 1994) and $\mathbf{D}_0$ is non singular. The MAP behaves like a Markov process with generator matrix $\mathbf{Q}$. $\mathbf{D}_0(i,j), i \neq j$ and $\mathbf{D}_1(i,j)$ contain transition rates from state i to j. Transitions from $\mathbf{D}_0$ are silent whereas transitions from $\mathbf{D}_1$ generate an arrival event. Thus, simulation of a MAP implies that the transitions of the process are simulated and arrivals are generated whenever a transition from $\mathbf{D}_1$ occurs. Although MAPs do not allow one to describe long range dependencies, they can be applied to approximate long range dependent processes on every finite time scale arbitrarily good (Horvath and Telek 2002).

3 FITTING METHODS

3.1 Fitting of AR and ARMA

The parameters of AR, MA, ARMA and also ARIMA processes can be computed by a least squares regression approach, using the Yule-Walker equations or by a maximum likelihood approach. Usually, one performs the fitting for different values of p and/or q and chooses the model with the smallest values that gives acceptable fit of the trace. Fitting methods for these models are part of statistical software like R (Chambers 2008) (package `stats`) which is also used for our examples.

3.2 Fitting of ARTA Processes

The first approach to fit ARTA processes is ARTAFACTS (ARTA Fitting Algorithm for Constructing Time Series) (Cario and Nelson 1998). The algorithm uses a given marginal distribution and fits an $AR(p)$ model according to the given autocorrelation structure of the trace. A wide variety of marginal distributions is supported. More recently the ARTAFIT algorithm has been developed (Biller and Nelson 2005, Biller and Nelson 2008). It uses Johnson distributions for the description of the marginal distribution and an advanced method for autocorrelation fitting. The Johnson distribution is completely determined by four parameters which can be chosen such that the distribution can match any finite first four moments (DeBrota et al. 1988). In the mentioned paper it is shown that the restriction to Johnson distributions is sufficient to model a wide variety of processes.

3.3 MAP Fitting

The fitting of MAP parameters according to some traffic trace is a complex optimization problem which is a research topic until today. Several different approaches exist nowadays which reach an appropriate fitting quality. In general one can distinguish between approaches that fit the parameters directly according to the values of the trace which usually means that the likelihood is maximized. Other approaches first derive some quantities like higher order moments, joint moments or lag k autocorrelation coefficients from the trace and fit the parameters of the MAP. Methods of the first type have the advantage of using the complete available information but have to handle the possibly very large trace for fitting. The methods that fit derived quantities like joint moments are usually more efficient but neglect some of the information available in the trace by fitting according to specific measures only. An overview of different fitting approaches can be found in (Horvath and Telek 2002). The question for the best fitting method for MAPs is still open such that several approaches should be tested. For our examples we fit MAPs with two of our own methods. First, we use an algorithm of the expectation maximization type to maximize the likelihood of the MAP with respect to the trace. Thus, if $t_1,\ldots,t_m$ are the interarrival times in the trace, then the following optimization problem is solved

$$\max_{\mathbf{D}_0,\mathbf{D}_1}\left(\pi\left(\prod_{i=1}^{m} e^{\mathbf{D}_0 t_i}\mathbf{D}_1\right)\mathbf{e}^T\right)$$

where $\pi\mathbf{Q}=0$, $\pi\mathbf{e}^T=1$ and $\mathbf{e}=(1,\ldots,1)$. The algorithm is described in (Buchholz 2003) and will be denoted by MAP EM in the following. Alternatively, we use an algorithm that performs the fitting in two steps. In the first step, the distribution of interarrival times is modeled by a so called phase type distribution (Horvath and Telek 2002) which determines matrix $\mathbf{D}_0$ and vector π. Afterwards, the distribution is extended to a MAP by finding an appropriate matrix $\mathbf{D}_1$ which leaves $\mathbf{D}_0$ and vector π unmodified (Horvath, Telek, and Buchholz 2005, Buchholz and Panchenko 2004). The values in $\mathbf{D}_1$ are set to match some lag k autocorrelations found in the trace. Details about the fitting algorithm can be found in (Panchenko and Buchholz 2007) and we name this algorithm MAP MOEA in this paper.

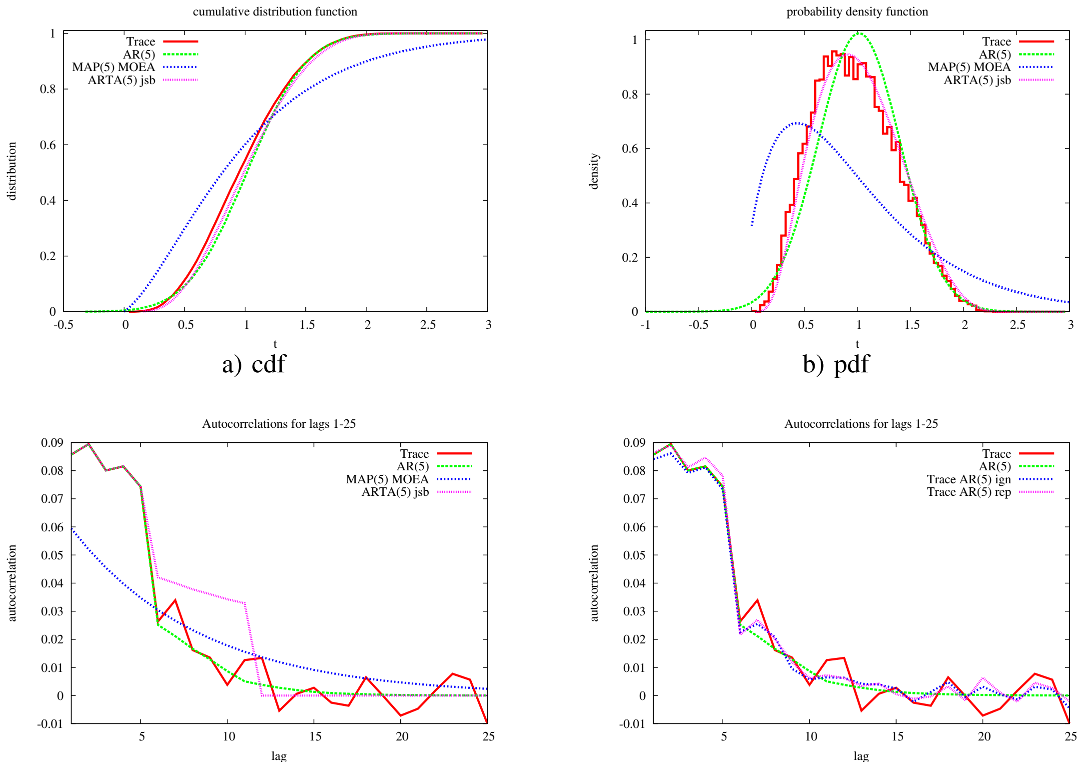
4 EXPERIMENTAL COMPARISON OF FITTING METHODS

In the following we will compare the quality of fitted MAPs, ARMA and ARTA processes for different traces. For our experiments we selected 6 traces. Four of them are synthetically generated traces, two are generated by MAPs and two by ARTA processes. The intention is to check whether processes of one class are able to capture processes of the other class. Two out of the six traces are measured traces which we choose to evaluate the suitability of the process classes and corresponding fitting methods for practical use. As mentioned (Tran and Reed 2004) applied ARIMA models for the modeling of the interarrival time process and (Karki and Hu 2005) reported about the fact that the ARIMA process generated negative values which were ignored in their case. Since MAPs only generate valid interarrival times we also considered this aspect and compared the characteristics of the original traces with those of the traces generated from the fitted processes where we took only feasible values into account by ignoring negative values or by replacing them (by 0). Another possibility to avoid infeasible values is to transform the generated traces of the fitted processes as for example described in (Williamson 1999). We do not consider this approach here since the determination of the parameters for the transformation requires additional expert knowledge and is not straightforward.

4.1 Synthetically Generated Traces

4.1.1 Traces generated from ARTA processes

The first trace considered in our evaluation contains $20,000$ observations from an ARTA process with $AR(5)$ base process and a Johnson bounded marginal distribution. The available tools for the generation of observations from an ARTA model are limited to this trace length and due to the small number of observations the generated trace shows pikes for autocorrelations of lags above 5 which were not specified by the original model. The parameters of the Johnson distribution are chosen such that the distribution only yields positive values, i.e. we set the shape parameters to $\gamma=0.6$ and $\delta=1.4$, the location parameter to $\xi=0.0$ and the scale parameter to $\lambda=2.5$. Since an $AR(5)$ base process was used for trace generation, we selected an autoregressive process of order 5 for AR fitting. For fitting an ARTA model we used the ARTAFIT software



a) cdf

b) pdf

c) autocorrelation

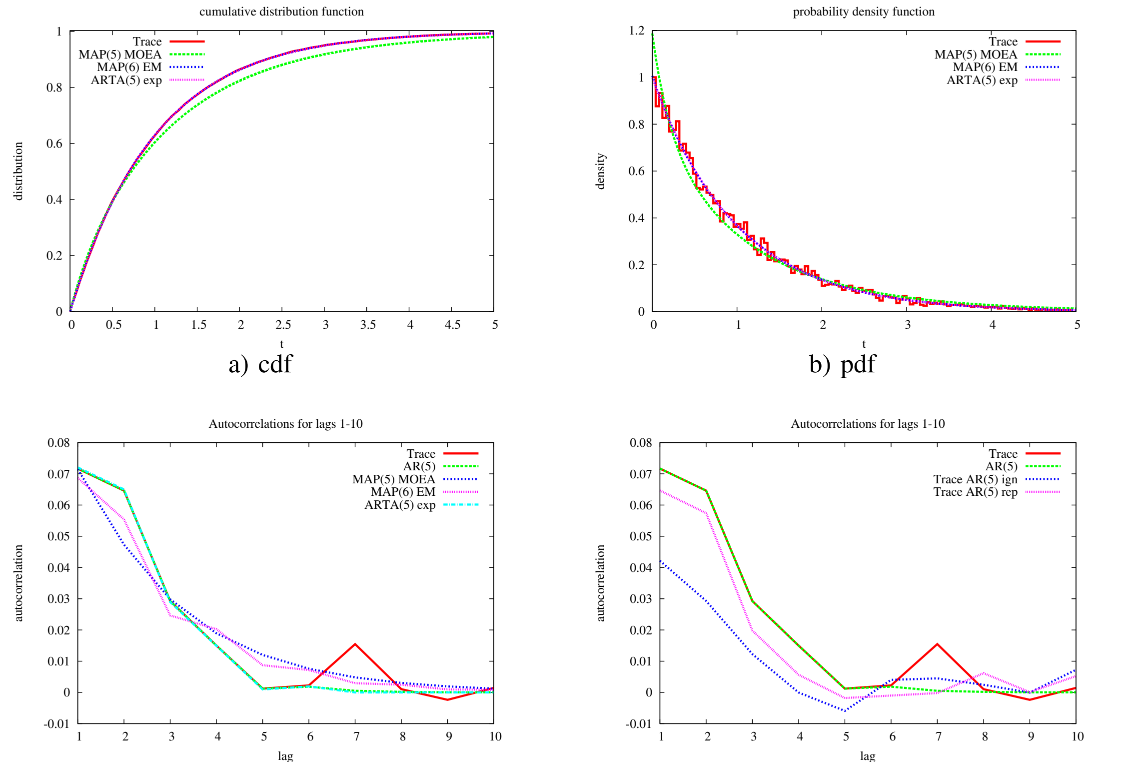
d) autocorrelation of traces generated from AR model

Figure 1: Fitting results for Trace from ARTA model with Johnson distribution

which also tries to fit a Johnson marginal distribution and thus should provide good results for this type of trace. MAP fitting was done with MAP-MOEA for a MAP of order 5.

The results are shown in Figure 1 a)-c). As one can see the ARTA model captures both the autocorrelations and the distribution of the trace well. The AR model provides a good estimation of the autocorrelations too. AR models always assume normal distribution and since the Johnson distribution which was used for trace generation is based on the normal distribution, the AR model captures the distribution as well. But as one can see from Figure 1 b) the distribution of the AR model might yield negative values that we have to deal with when using the AR process for the generation of interarrival times in a simulation model. The obvious choices are to either ignore those values in simulation, i.e. delete them from the generated observations, or to replace them with some non-negative value like 0. Of course this treatment of negative values has impact on the autocorrelations and distributions of the generated observations. For Figure 1 d) we generated $100,000$ observations from the AR model and compared the autocorrelations of the original AR process with the correlations from the generated sequences when ignoring negative values (Trace $AR(5)$ ign) or replacing them with 0 (Trace $AR(5)$ rep). As one can see the effect on the autocorrelations is negligible in this example, but will become more noticeable for the examples presented in the next paragraphs. For MAP fitting this type of trace is actually a difficult task. Although PH distributions can approximate the shape of normal-like distributions reasonably well (cf. e.g. (Thümmler, Buchholz, and Telek 2005)), fitting becomes more sophisticated when autocorrelations have to be considered as well and one has to accept a trade-off between good distribution fitting vs. good fitting of the autocorrelation structure.

The second synthetically generated trace used for the comparison of fitting methods was again generated by an ARTA process. The trace contains $20,000$ observations and was created from an ARTA model with $AR(5)$ base process and exponential marginal distribution with rate parameter $\lambda = 1.0$. For AR and MAP fitting we selected an autoregressive process of order 5 and MAPs of order 5 and 6, respectively. Since ARTAFIT had problems providing a good approximation for the exponential distribution of the trace (this will be discussed in more detail in one of the following examples) we used ARTAFACTS for autocorrelation fitting and set the exponential marginal distribution for the ARTA model manually.



a) cdf b) pdf

c) autocorrelation d) autocorrelation of traces generated from AR model

Figure 2: Fitting results for Trace from ARTA model with exponential distribution

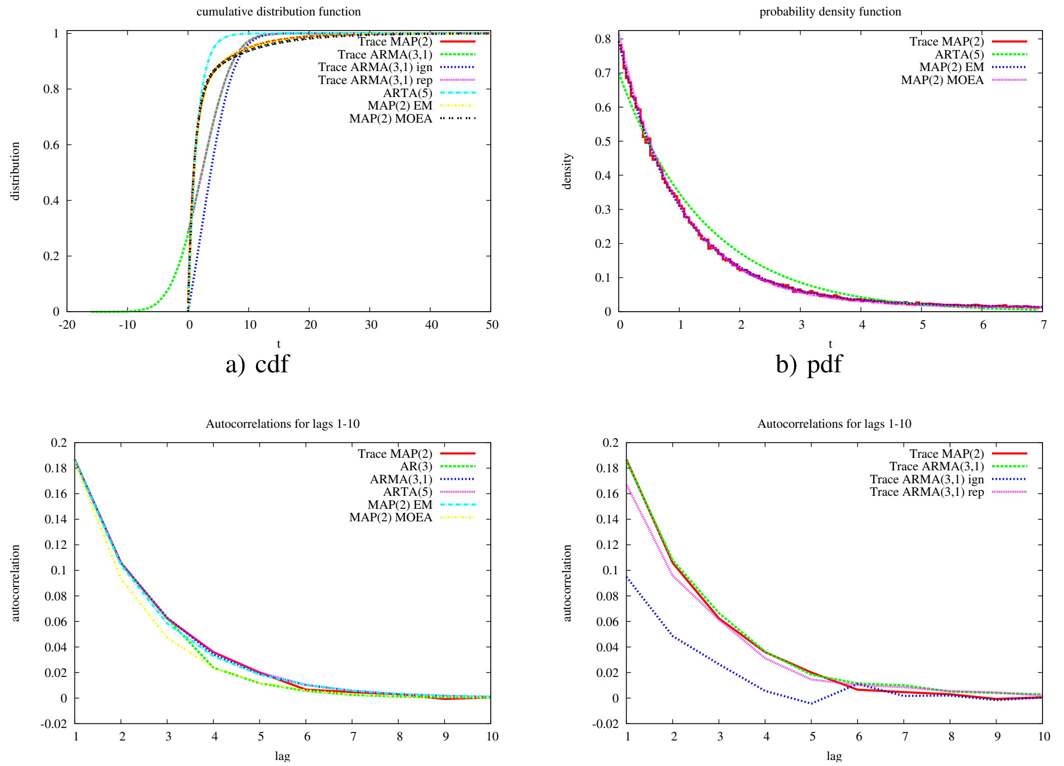
Plots for the distribution and the autocorrelation structure of the fitted model are shown in Figure 2. From Figures 2 a) and b) one can see that the MAP provides a sufficient fitting for the distribution. We omitted the curve for the distribution of the AR process, since it has a Normal shape and therefore is obviously not an adequate approximation of the empirical distribution of the trace. Regarding the autocorrelations of the trace all models provide good results (cf. Figure 2 c). ARTA and AR model capture the first autocorrelations exactly while the MAP underestimates the correlation at lag 2. As already mentioned the AR model does not fit the distribution of the trace and similar to the previous example the simulation of the AR model yields negative values. Figure 2 d) shows the autocorrelation of traces generated from the AR model when negative values are ignored or replaced with 0. It is visible that the treatment of negative values has some serious impact on the autocorrelation for this example.

4.1.2 Traces generated from MAPs

The next two traces considered for our comparison are generated using different MAPs. The first trace with $200,000$ elements was generated from a $MAP(2)$ with matrices

$$D_0 = \begin{bmatrix} -1.00 & 0.03 \\ 0.05 & -0.16 \end{bmatrix}, D_1 = \begin{bmatrix} 0.90 & 0.07 \\ 0.01 & 0.10 \end{bmatrix}$$

and has autocorrelations for smaller lags only. The fitting results are shown in Figure 3. All models provide a good fitting for the autocorrelation, but only the fitted MAPs were able to capture the distribution as well. The distribution of the ARMA models (which is omitted in Figure 3) is normal again and the models yield negative values when used in simulation. The Johnson distribution returned by ARTAFIT has similar problems, though in general the Johnson distribution can take forms which only yield positive values. This would require adding further constraints for the distribution fitting, but since the sources were not available for us, we used ARTAFACTS for the autocorrelations only and fitted an exponential distribution



a) cdf   b) pdf

c) autocorrelation   d) autocorrelation of traces from ARIMA model

Figure 3: Fitting results for Trace from $MAP(2)$

manually. For the ARMA models we cannot fall back to different distributions and thus have to deal with the negative values in simulation. Figure 3 d) shows the impact of ignored and replaced values on the autocorrelation of the $ARMA(3,1)$ model.

The next trace was generated by a $MAP(3)$ with autocorrelation up to lag 20 from (Buchholz 2003) and contains 200,000 elements. The MAP is defined by the two matrices

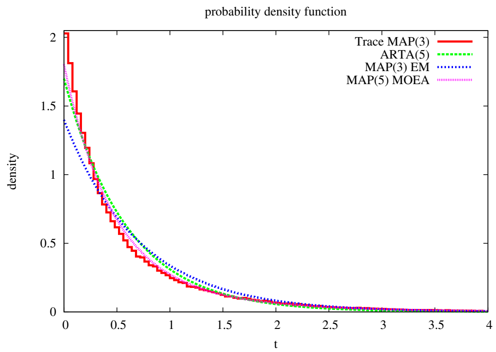
$$D_0 = \begin{bmatrix} -3.721 & 0.500 & 0.020 \\ 0.100 & -1.206 & 0.005 \\ 0.001 & 0.002 & -0.031 \end{bmatrix}, D_1 = \begin{bmatrix} 0.200 & 3.000 & 0.001 \\ 1.000 & 0.100 & 0.001 \\ 0.005 & 0.003 & 0.020 \end{bmatrix}$$

The fitting results (cf. Figure 4) are similar to the results of the trace generated from $MAP(2)$. Again the MAP fitting methods were able to capture both distribution and autocorrelation. For ARTA fitting we had to use ARTAFACTS with a manually fitted distribution. Simulation of the ARMA models resulted in negative values again. The impact of ignoring and replacing those values on the autocorrelation and the distribution function is shown in Figures 4 c) and d), respectively.

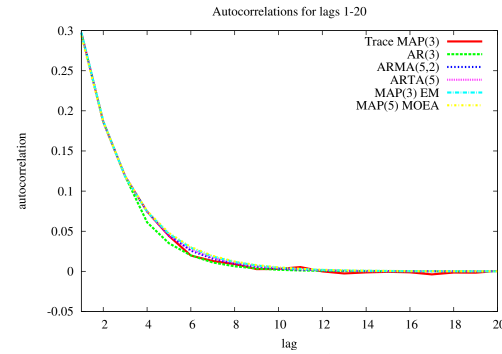
4.2 Measured Traces

In addition to the synthetically generated traces presented in the previous section we compared the fitting methods with real traces. In the following we will present the results for two traces taken from the Internet Traffic Archive (http://ita.ee.lbl.gov/). The trace *BC-pAug89* contains a million packet arrivals observed at the Bellcore Morristown Research and Engineering facility in August 1989. The trace *LBL-TCP-3* (Paxson and Floyd 1995) contains two hours of TCP traffic from the Lawrence Berkeley Laboratory and was recorded in January 1994. For the comparison of fitting tools both traces were normalized to mean 1.0.

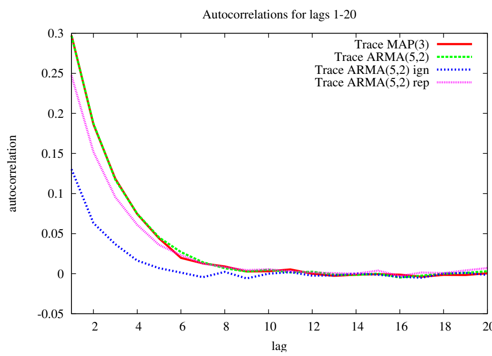
We will start with the result for *BC-pAug89*: Figure 5 a) shows the autocorrelations of different ARMA models that have been used to fit the trace. While the pure AR models ($AR(3)$ and $AR(9)$) only capture the first 3 and 9 lags, respectively,



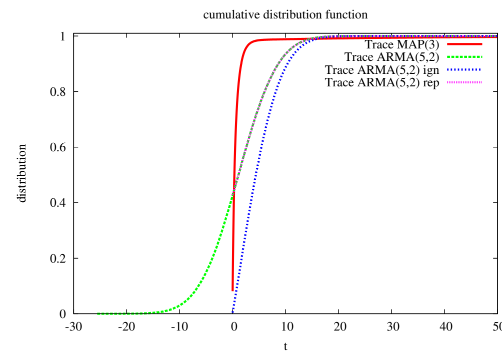
a) pdf



b) autocorrelation



c) autocorrelation of traces from ARIMA model

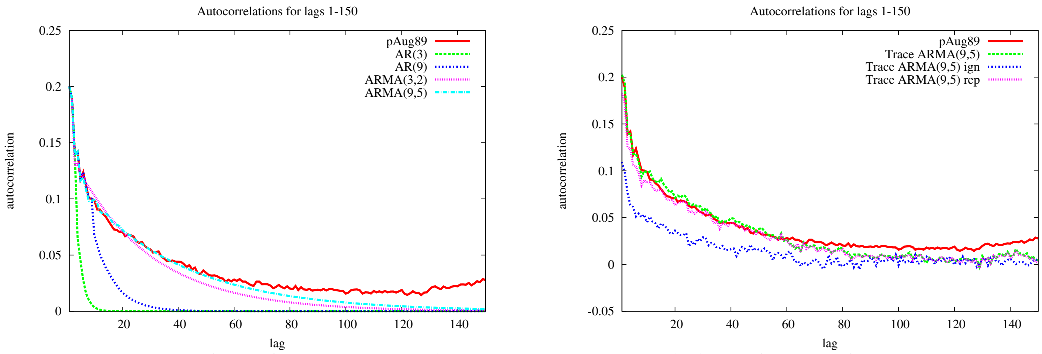


d) cdf of traces from ARIMA model

Figure 4: Fitting results for Trace from $MAP(3)$

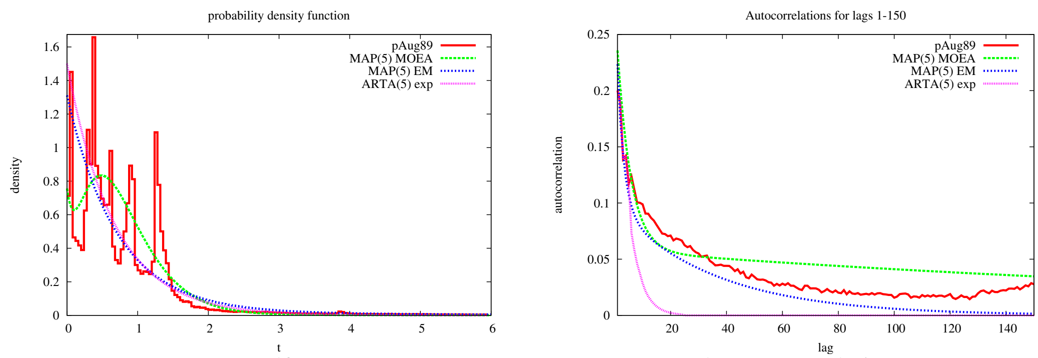
allowing for only few additional moving average terms in the model results in a vast improvement of the fitting quality for higher lags as can be seen for the curves of an $ARMA(3,2)$ and an $ARMA(9,5)$ model, while still keeping the model size small. Again, we have to deal with negative values that the ARMA models might yield when simulated. The impact of ignoring or replacing negative values on the autocorrelation of the simulated $ARMA(9,5)$ model is shown in Figure 5 b). We omitted the plots for the other ARMA models because they show similar results. Since the trace is known to exhibit self-similar behavior, we also used F-ARIMA models for fitting, but the resulting models only showed little improvement for higher lags over ARMA models. The plots for the models resulting from MAP and ARTA fitting are shown in Figure 6. The two MAPs fitted with different techniques either overestimated (MAP MOEA) or underestimated (MAP EM) the higher lag autocorrelations, but provide a good approximation for the lower lags. The ARTA model was fitted again using ARTAFACTS with a manually chosen exponential distribution. Since ARTAFACTS only supports autocorrelation for up to five lags, the resulting model only captures the first five autocorrelations. But since ARTA models use an underlying AR base process it is obvious that capturing autocorrelations up to a reasonable lag would result in a very large base process for the ARTA model.

As already mentioned the second trace from a real system is the *LBL-TCP-3* trace. Figure 7 shows the fitting results for different ARMA models. All the ARMA models provide an adequate fitting of the autocorrelation while, of course, the pure AR models are limited to the smaller lags (cf. Figure 7 a). Figure 7 b) shows the impact on the autocorrelation when ignoring or replacing negative values for the $ARMA(9,5)$ model. Figure 8 shows the fitting results for MAPs and ARTA models. As one can see from Figure 8 a) ARTAFIT was not able to return a Johnson distribution that only has positive values. Using this distribution would lead to similar results as described before for ARMA models. Hence, we used a manually fitted exponential distribution again and used ARTAFACTS only for capturing the autocorrelations (cf. Figure 8 b). The MAPs provided a sufficient fitting of both distribution and autocorrelation for the *LBL-TCP-3* trace and the MAP resulting from MAP MOEA was even able to capture higher lag autocorrelations.



a) autocorrelation of ARMA models b) autocorrelation of traces from ARMA models

Figure 5: ARMA fitting results for BC-pAug89 trace

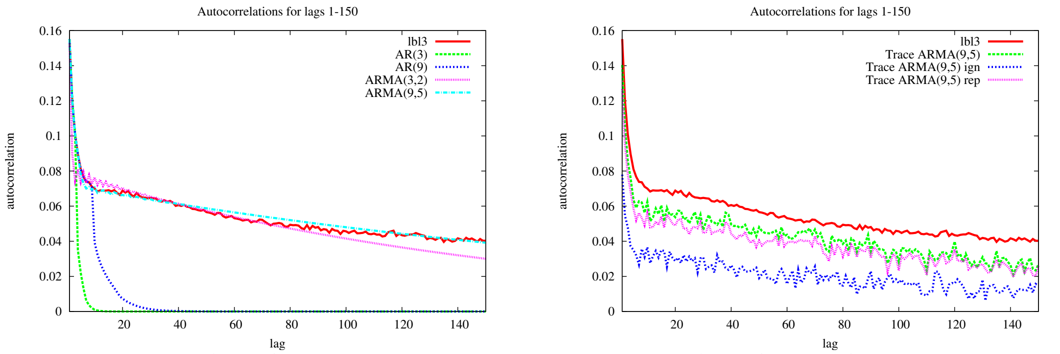


a) pdf b) autocorrelation

Figure 6: MAP and ARTA fitting results for BC-pAug89 trace

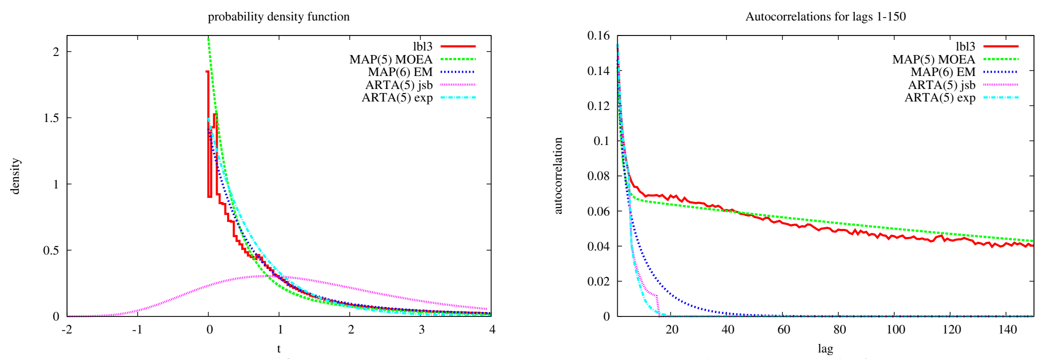
5 CONCLUSIONS

In this paper we analyzed empirically different methods to model correlated input processes in simulation models. Our examples indicate that all the investigated models are able to provide an adequate fitting of the autocorrelation. Of course AR models can only capture the first few lags or would require a large amount of AR coefficients. The same holds for ARTA models which inherit this property from the underlying base AR process. Regarding the distribution ARMA models yield poor results, since they always assume a normal distribution. But even in cases where the empirical distribution of the trace has a normal shape (cf. Figure 1) the fitted ARMA model might yield negative values and requires some additional effort to be usable in a simulation run. Software for ARTA models currently only support an automated fitting of a distribution from the Johnson family, which in our experience was in many cases not sufficient to capture the characteristics of the empirical distribution of the traces. Using distributions other than Johnson provided better results but requires some expert knowledge for selecting an appropriate distribution and its parameters. MAPs provided sufficient results in capturing both distribution and autocorrelation in our experiments but Figure 1 shows that in some cases one has to accept a trade-off between good distribution and good autocorrelation fitting. Table 1 shows a rough rating of our experiment results for comparison of the quality of the used fitting methods. For each trace and fitting method we assessed the quality of the fitted distribution (PDF) and the autocorrelation (AC). The assessment of the quality of autocorrelation fitting consists of four entries, the upper two entries give an indication of the fitting of lower and higher lags, respectively. The lower two entries assess the quality of the lower lag autocorrelations of the traces generated from the fitted processes, taking the procedure of replacing or ignoring negative values into account. As mentioned, MAPs always generate feasible values, so that we omitted those entries for



a) autocorrelation of ARMA models b) autocorrelation of traces from ARMA models

Figure 7: ARMA fitting results for LBL-TCP-3 trace



a) pdf b) autocorrelation

Figure 8: MAP and ARTA fitting results for LBL-TCP-3 trace

them. The same holds for ARTA processes with a given distribution. Other empty entries indicate that we received no useful result from the used tools.

It is noticeable that AR* fitting methods have difficulties to fit the distribution in case it does not belong to the Johnson family of distributions. Unsurprising the resultant processes capture the autocorrelations nearly perfectly up to the user given lag distance p, but higher lag correlations were not captured well. MAPs are in most cases better in fitting the distribution, but are somewhat worse in fitting lower lag correlations. On the other hand our experience was that they can capture higher lag correlations better than the AR* fitting methods which surely depends on the character of the trace's correlation structure.

Concerning the use of the fitted arrival processes in simulation models our evaluation implies a different view. As mentioned, MAPs generate only positive and thus valid interarrival times by their nature, so that the valuation does not change in this respect. AR* fitted processes might generate negative and thus infeasible interarrival time values dependent on the processes distribution. In such situations our experience is that replacing negative values (by 0 in our experiments) is slightly better for autocorrelation fitting than ignoring those values but often results in a bias of the marginal distribution.

In summary, our experiments suggest that the fitting and modeling of arrival processes by MAPs is not only of interest in analytical models, but also offers several advantages for simulation models. A drawback of using MAPs is that the available fitting methods are more time consuming than fitting AR* models. In our experiments fitting of pure AR models was fastest followed by ARTA fitting when only autocorrelations were considered. The time for fitting both distribution and autocorrelation for ARTA models heavily depended on the type of the distribution, i.e. was fast for traces from a Johnson distribution (in the range of CPU minutes) but required much more time for the other empirical distributions (in the range of CPU hours). As already mentioned MAP fitting was most time consuming, but still acceptable: Depending on the size of the trace CPU time was in the range of several minutes up to several hours.

Trace		**Fitting Method**					
		MAP MOEA	MAP EM	AR(p)	ARMA(p,q)	ARTA(p) given dist. (ARTAFACTS)	ARTA(p) (ARTAFIT)
pAug	PDF	–	–	—	—		—
	AC low high	+ +	+ +	++ —	++ +	++ —	++ —
	repl ignore			+ —	+ —		○ —
lbl	PDF	+	+	—	—		—
	AC low high	++ ++	++ –	++ —	++ +	++ —	++ —
	repl ignore			+ —	+ —		○ —
MAP(2)	PDF	++	++	—	—		—
	AC low high	+ +	++ +	+ +	++ +	++ +	++ +
	repl ignore			+ —	+ —		+ —
MAP(3)	PDF	++	+	—	—		
	AC low high	++ ++	++ ++	+ +	++ ++	++ ++	
	repl ignore			+ —	+ —		
ARTA (exp)	PDF	+	++	—	—		—
	AC low high	+ +	+ +	++ ++	++ ++	++ ++	++ +
	repl ignore			+ —	+ —		+ —
ARTA (Johnson)	PDF	○		+	+		++
	AC low high	○ ○		++ ++	++ ++	++ ++	++ ++
	repl ignore			++ ++	++ ++		++ ++

Table 1: Qualitative Comparison of Fitting Methods (Legend: ++ very good; + good; ○ fair; – bad; — very bad)

REFERENCES

Biller, B., and S. Ghosh. 2004. Dependence Modeling for Stochastic Simulation. In *Proceedings of the 2004 Winter Simulation Conference*, ed. R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 153–161. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Biller, B., and B. L. Nelson. 2005. Fitting Time-Series Input Processes for Simulation. *Oper. Res.* 53 (3): 549–559.

Biller, B., and B. L. Nelson. 2008. Evaluation of the ARTAFIT Method for Fitting Time-Series Input Processes for Simulation. *INFORMS Journal on Computing* 20 (3): 485–498.

Box, E., and G. Jenkins. 1970. *Time Series Analysis - forecasting and control*. Holden-Day.

Buchholz, P. 2003. An EM-Algorithm for MAP Fitting from Real Traffic Data. In *Computer Performance Evaluation / TOOLS*, ed. P. Kemper and W. H. Sanders, Volume 2794 of *Lecture Notes in Computer Science*, 218–236: Springer.

Buchholz, P., and A. Panchenko. 2004. A Two-Step EM Algorithm for MAP Fitting. In *ISCIS*, ed. C. Aykanat, T. Dayar, and I. Korpeoglu, Volume 3280 of *Lecture Notes in Computer Science*, 217–227: Springer.

Cario, M. C., and B. L. Nelson. 1996. Autoregressive to anything: Time-series input processes for simulation. *Operations Research Letters* 19 (1): 51–58.

Cario, M. C., and B. L. Nelson. 1998. Numerical Methods for Fitting and Simulating Autoregressive-To-Anything Processes. *INFORMS J. on Computing* 10 (1): 72–81.

Chambers, J. M. 2008. *Software for Data Analysis: Programming with R*. New York: Springer. ISBN 978-0-387-75935-7.

Crovella, M. E., and A. Bestavros. 1996. Self-similarity in World Wide Web traffic: evidence and possible causes. In *SIGMETRICS '96: Proceedings of the 1996 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, 160–169. New York, NY, USA: ACM.

DeBrota, D. J., S. D. Roberts, J. J. Swain, R. S. Dittus, J. R. Wilson, and S. Venkatraman. 1988. Input modeling with the Johnson system of distributions. In *Proceedings of the 1988 Winter Simulation Conference*, ed. M. A. Abrams, P. L. Haigh, and J. C. Comfort, 165–179. New York, NY, USA: ACM.

Devroye, L. 1986. *Non-Uniform Random Variate Generation*. New York: Springer.

Heindl, A., K. Mitchell, and A. van de Liefvoort. 2006. Correlation bounds for second-order MAPs with application to queueing network decomposition. *Perform. Eval.* 63 (6): 553–577.

Horvath, A., and M. Telek. 2002. Markovian modeling of real data traffic: Heuristic phase type and MAP fitting of heavy tailed and fractal like samples. In *Performance 2002*, ed. M. C. Calzarossa and S. Tucci, Volume 2459 of *LNCS*, 405–434: Springer.

Horvath, G., M. Telek, and P. Buchholz. 2005. A MAP fitting approach with independent approximation of the inter-arrival time distribution and the lag correlation. In *QEST '05: Proceedings of the Second International Conference on the Quantitative Evaluation of Systems*, 124. Washington, DC, USA: IEEE: IEEE Computer Society.

Kang, S., Y. H. Kim, D. Sung, and B. Choi. 2002, Apr. An application of Markovian arrival process (MAP) to modeling superposed ATM cell streams. *IEEE Transactions on Communications* 50 (4): 633–642.

Karki, R., and P. Hu. 2005. Wind power simulation model for reliability evaluation. In *Proceedings of the 2005 Canadian Conference on Electrical and Computer Engineering*, 541–544.

Kelton, W. D., and A. Law. 2000. *Simulation Modeling and Analysis*. McGraw Hill.

Kuhl, M. E., E. K. Lada, N. M. Steiger, M. A. Wagner, and J. R. Wilson. 2007. Introduction to modeling and generating probabilistic input processes for simulation. In *Proceedings of the 2007 Winter Simulation Conference*, ed. S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 63–76. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Leemis, L. 1998. Input modeling. In *Proceedings of the 1998 Winter Simulation Conference*, ed. D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, 15–22. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Livny, M., B. Melamed, and A. K. Tsiolis. 1993. The impact of autocorrelation on queueing systems. *Management Science* 39:322–339.

Lucantoni, D. M., K. S. Meier-Hellstern, and M. F. Neuts. 1990. A single server queue with server vacations and a class of non-renewal arrival processes. *Advances in Applied Probability* 22:676–705.

Panchenko, A., and P. Buchholz. 2007. A Hybrid Algorithm for Parameter Fitting of Markovian Arrival Processes. In *Proc. of 14th Int. Conf. on Analytical and Stochastic Modelling Techniques and Applications*, 7–12: SCS Press.

Paxson, V., and S. Floyd. 1995. Wide-Area Traffic: The Failure of Poisson Modeling. *IEEE/ACM Transactions on Networking* 3:226–244.

Sheluhin, O. I., S. M. Smolskiy, and A. V. Osin. 2007. *Self-similar processes in telecommunications*. Wiley.

Stewart, W. J. 1994. *Introduction to the numerical solution of Markov chains*. Princeton University Press.

Tartarelli, S., M. Pagano, and M. Devetsikiotis. 2000. Efficient Estimation of the Cell Loss Probability in a Two-Buffer PGPS Scheduler. In *IEEE International Conference on ICC Communications (3)*, 1320–1324.

Thümmler, A., P. Buchholz, and M. Telek. 2005. A Novel Approach for Fitting Probability Distributions to Real Trace Data with the EM Algorithm. In *DSN*, 712–721: IEEE Computer Society.

Tran, N., and D. A. Reed. 2004. Automatic ARIMA time series modeling for adaptive I/O prefetching. *IEEE Transactions on Parallel and Distributed Systems* 15:362–377.

Williamson, C. 1999. Synthetic Traffic Generation Techniques For ATM Network Simulations. *Simulation* 72 (5): 305–312.

Xue, F., J. Liu, Y. Shu, L. Zhang, and O. W. W. Yang. 1999. Traffic modeling based on FARIMA models. In *Proceedings of the 1999 IEEE Canadian Conference on Electrical and Computer Engineering*, 162–167.

AUTHOR BIOGRAPHIES

FALKO BAUSE holds a Doctoral degree in computer science from the TU Dortmund. His main research interests are in the area of system engineering with emphasis on Stochastic Petri Nets. He defined the Queueing Petri Net formalism, which combines Queueing Networks with Stochastic Petri Nets and coauthored a book with the title "Stochastic Petri Nets – An Introduction to the Theory". Currently he is involved in a project on Markovian Arrival and Service Processes for Performance and Reliability Analysis. His e-mail address is <falko.bause@udo.edu>.

PETER BUCHHOLZ received the Diploma degree (1987), the Doctoral degree (1991) and the Habilitation degree (1996) all from the TU Dortmund, where he is currently a professor for modeling and simulation. His current research interests are efficient techniques for the analysis of stochastic models, formal methods for the analysis of discrete event systems, the development of modeling tools, as well as performance and dependability analysis of computer and communication systems. His e-mail address is <peter.buchholz@udo.edu>.

JAN KRIEGE received the Diploma degree in computer science from the TU Dortmund in 2006. His research interests include the modeling and analysis of logistics networks and computer and communication systems. His e-mail address is <jan.kriege@udo.edu>.