Analysis of Non-stationary Stochastic Simulations Using Time Series

RITA MARQUES BRANDÃO[1]
Departamento de Matemática,
Universidade dos Açores,
Apartado 1422, 9501-801 Ponta Delgada,
PORTUGAL

ACÁCIO M. O. PORTA NOVA[2]
Departamento de Engenharia e Gestão,
Instituto Superior Técnico,
Av. Rovisco Pais, 1049-001 Lisboa,
PORTUGAL

(October 2, 2003)

[1]E-mail: <rita@notes.uac.pt>. Telephone: (+351) 296650518. Fax: (+351) 296650072.

[2]E-mail: <apnova@ist.utl.pt. Telephone: (+351) 218417174. Fax: (+351) 218417979.

Analysis of Non-stationary Stochastic Simulations Using Time Series

RITA MARQUES BRANDÃO[1] and ACÁCIO M. O. PORTA NOVA[2]

[1]*Departamento de Matemática*
Universidade dos Açores,
Portugal
E-mail: rita@notes.uac.pt

[2]*Departamento de Engenharia e Gestão,*
Instituto Superior Técnico,
Portugal
E-mail: apnova@ist.utl.pt

This paper extends the use of time series models to the output analysis of non-stationary discrete event simulations. A thorough experimental evaluation showed that $ARIMA(p, d, q)$ models are very promising meta-models for simulating queueing systems under critical traffic conditions. In some situations, stationarity-inducing transformations may be required, before this methodology can be used. Our approach for efficient estimation of performance measures of selected responses in the target system is illustrated with a single lane traffic analysis.

1. Introduction

The output analysis of discrete event simulation models is undoubtedly an area of utmost practical importance. It is also a very active research area, and many innovative approaches have been proposed to cope with the highly autocorrelated nature of simulation responses, namely in queueing systems.

Fishman (1971) gave an early example of how reliable measures of response variability (variance, confidence intervals) could be estimated using classical time series models (Box, Jenkins, and Reinsel 1994). In an intermediate step of his procedure, an autoregressive

model of order p, AR(p), was fitted to a selected simulation response. Later, Schriber and Andrews (1984) extended that approach by fitting mixed autoregressive-moving average models, ARMA(p,q), with an automated procedure. However, the authors reported poor performance results, especially in the coverage rates of confidence intervals for simple queueing systems.

In contrast to this apparent incompatibility between queueing systems and time series models, Brandão and Porta Nova (1999) showed that either an insufficient simulation duration, or an excessive initial bias, were responsible for most of those poor results. Eventually, by controlling these two factors, very positive results were obtained for an $M/M/3$ stationary queue under moderate and congested traffic situations.

This paper is organized as follows. In Section 2, we discuss the use of classical time series models to analyze the output of non-stationary discrete event simulations. In Section 3, we investigate the applicability of $ARIMA$ models as potential meta-models for queueing system simulation under critical traffic conditions. In Section 4, we illustrate our approach using a single lane traffic analysis. Finally, in Section 5, we summarize some conclusions and recommendations for future research in this area. Some preliminary results were originally presented in Brandão and Porta Nova (2003a,b).

2. Non-stationary Simulations and Time Series

The specialized literature on the output analysis of discrete event simulations is conspicuously void of non-stationary cases. However, in many situations of undoubted practical interest (rush-hour periods, system breakdowns, etc.), arrival rates do actually exceed processing rates. Occasional references to such systems mention uncontrolled evolution, explosive growth, or unbounded queue lengths... But, would it be possible to (meaningfully) analyze those non-stationary processes? The answer seemed to be no, since most performance measures of simple queueing systems go to infinity, when the utilization factor approaches

one. Probably, these familiar asymptotic results convinced the simulation community at-large that it was worthless to explore non-stationary simulations. However, if they were tractable, much useful information could be extracted and many "what if?" questions could be answered. How many cars can be expected to be waiting in line at the end of a rush-hour period? For a driver arriving halfway through a rush period, how long will he take to pass through a traffic bottleneck? Even better, can we predict the evolution in time of these and other responses, without repeating time consuming simulations and subsequent output analysis? After all, the leitmotif in the pursuit of simulation meta-models is finding simpler but realistic analytical representations of the corresponding computer programs.

2.1. *The Classical Time Series Method*

It is well-known that stochastic simulation models produce strongly autocorrelated time series. So, it should not be surprising that we chose to investigate the applicability of the classical $ARIMA$ time series models, as potential meta-models for non-stationary simulations. However, the independent replications that are readily available in discrete event simulation experiments are in stark contrast to most scientific areas where the Box-Jenkins methodology is widely applied and data is scarcely available. This is the case, for instance, when econometric models are tentatively fit to a single realization of an economical time series. Thus, we can analyze non-stationary simulations by averaging the responses across several runs, instead of relying on a single realization. Consequently: (i) the time series variability is reduced; (ii) the *underlying* evolution in time of the response is easier to identify; and (iii) the meta-model fitting process becomes much more reliable. Additionally, this is a valid approach for fitting meta-models to autocorrelated data, in contrast to regression-based procedures. In some particular situations, it may become necessary to previously apply a variance-stabilizing transformation to the averaged time series.

2.2. *Number of Replications*

To determine an adequate number of independent replications for the averaged series, we analyzed the behavior of the response average length of the $M/M/1$ queue, with $\rho = 2$. We collected data at half minute intervals ($\Delta_t = 0.5$) for different numbers of runs (10, 20, 25, 30, 35 and 40), around 30 replications. As can be observed by looking at Figure 1, the behavior of the average queue length is similar for all cases and indistinguishable from the theoretical curve; see Bailey (1957). The theoretical curve for $\rho = 1$ is also represented in this figure.

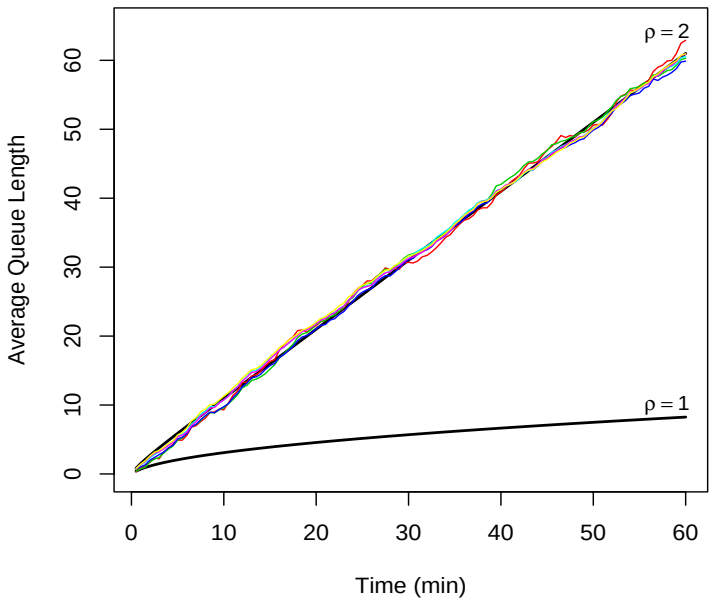

Fig. 1. Average Queue Length ($M/M/1$ Queue, with $\rho = 2$ and $\Delta_t = 0.5$)

In order to complement that analysis, we then estimated the probabilities of coverage of 95% confidence intervals based on the selected number of replications. To do so, we computed 100 such confidence intervals, $\overline{X}_t \pm t_{1-\alpha/2}(n-1)\widehat{\sigma}_{\overline{X}_t}/\sqrt{n}$, at some selected instants t, for the average length of the $M/M/1$ queue, with $\rho = 2$. The results that were obtained are reported in Table 1 and graphically displayed in Figure 2.

Table 1. Coverage of 95% CI for the Average Queue Length ($M/M/1$ Queue, with $\rho = 2$ and $\Delta_t = 0.5$)

	Coverage					
Instant	10 Runs	20 Runs	25 Runs	30 Runs	35 Runs	40 Runs
10	.95	.91	.92	.93	.90	.88
20	.93	.92	.91	.91	.92	.94
30	.90	.94	.91	.95	.93	.93
40	.95	.96	.94	.97	.96	.95
50	.95	.97	.97	.96	.95	.96
60	.94	.96	.96	.96	.96	.96

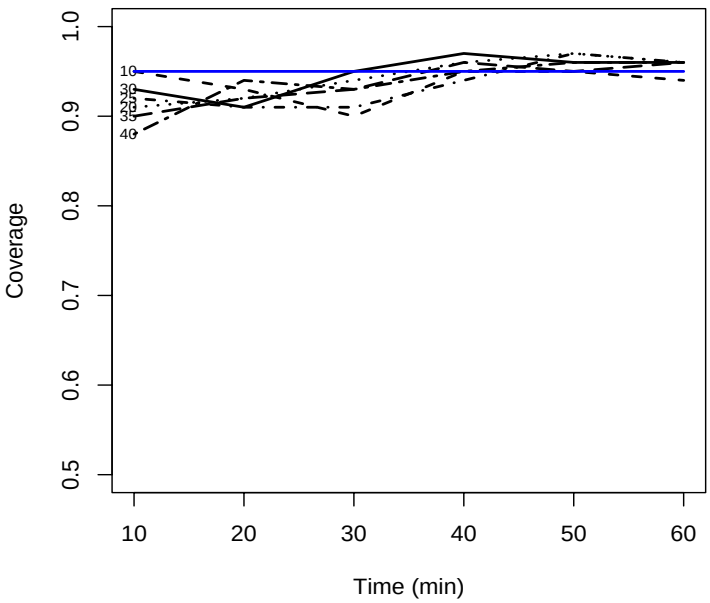

Fig. 2. Coverage of 95% CI for the Average Queue Length ($M/M/1$ Queue, with $\rho = 2$ and $\Delta_t = 0.5$)

We can see that the results are slightly better for 30 runs, and also for statistical stability reasons, we chose 30 as the number of independent replications to use on the remaining experimentations.

3. Experimental Evaluation

In this section, we illustrate and experimentally evaluate our meta-modelling approach to non-stationary simulation using two case studies, two simple queues, with utilization factors greater than or equal to one. We analyze the evolution of two responses, the average queue length and the average time spent in the system. We discuss the applicability of $ARIMA(p, d, q)$ meta-models to the non-stationary simulation of two $M/M/s$ queueing systems—that is, with exponential inter-arrival and service times, and s identical parallel servers. We consider the following two queueing systems, with different values for the utilization factor, $\rho = \lambda/(s\mu)$, where λ and μ represent the arrival and service rates, respectively:

(i) An $M/M/1$ queue with super-critical traffic ($\rho = 2$); and

(ii) An $M/M/2$ queue with critical traffic ($\rho = 1$).

We then performed a Monte Carlo experiment consisting of 100×30 independent replications of the corresponding simulation models, with a *reference* duration of 60 time units. The *actual* duration varied, because each run was only terminated when the last entity that had arrived before 60 time units left the system. Initially, each model was started in an empty and idle state.

Queue lengths were collected at regular time intervals, $\Delta_t = 0.5$ and 1.0, with $t \in (0, 60]$. Sojourn times were sorted according to the time interval in which the arrival had occurred. Then, for each time interval, we collected only the first observation and the corresponding observations across runs were averaged. Additional observations in the same time interval were rejected to maintain independence. This contributed, as well, to a more stable behavior

of the responses, when compared with earlier stages of development of our approach. Finally, we applied the Box-Jenkins methodology to the averaged time series of each response, through identification, estimation and diagnostic checking of the $ARIMA(p, d, q)$ models.

Although the experimentation was performed for both interval widths, we only present the graphs corresponding to $\Delta_t = 0.5$, because the results are very similar.

3.1. *A Super-critical* $M/M/1$ *Queue*

We created a super-critical traffic situation with an inter-arrival rate of $\lambda = 2$ and a service rate of $\mu = 1$, resulting in an utilization factor of $\rho = 2$.

When we analyzed the average queue length for a typical set of 30 runs (see Figure 3, the hardly discernible dash-line), we concluded that the process was non-stationary on the mean. There is a marked linear trend and the sample autocorrelation function (ACF) decreases very slowly to zero. We also represent the sample partial autocorrelation function in the same figure. The results produced by the fitted model are represented as well (in solid line) on the subgraph containing the original series (dash-line).

Differentiating the series (see again Figure 3), we observed that it became stationary, without any statistically significative value on the ACF and PACF. This was also confirmed applying the Box-Ljung test. Thus, we can fit the simplest possible $ARIMA$ model, $ARIMA(0, 1, 0)$, to the average queue length. In this case, the first difference produces white noise. This process is called a random walk with drift, if it has a nonzero expected value, or simply a random walk, otherwise.

The analysis was repeated for the remaining 99 time series (each corresponding to an average of 30 runs) obtaining the results reproduced in Table 2. The same $ARIMA(0, 1, 0)$ model was consistently validated for almost 100% of the cases, when $\Delta_t = 0.5$, and for about 90% of the cases, when $\Delta_t = 1$.

In Figure 4, we represent both the 100 original series and the corresponding fitted series.

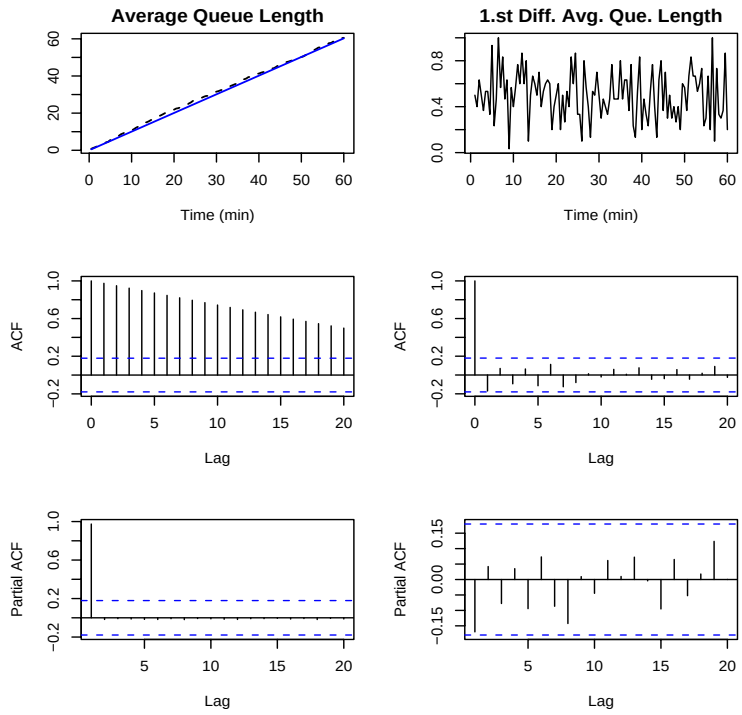

Fig. 3. Average Queue Length ($M/M/1$, $\rho = 2$ and $\Delta_t = 0.5$)

Table 2. Valid Fits for Average Queue Length ($M/M/1$ Queue, with $\rho = 2$)

$ARIMA(0,1,q)$ **FITTED MODEL**	**TIME INTERVAL** $\Delta_t = 0.5$	$\Delta_t = 1.0$	$ARIMA(p,1,0)$ **FITTED MODEL**	**TIME INTERVAL** $\Delta_t = 0.5$	$\Delta_t = 1.0$
$ARIMA(0,1,0)$	97	89	$ARIMA(0,1,0)$	97	89
$ARIMA(0,1,4)$		1	$ARIMA(1,1,0)$		1
$ARIMA(0,1,5)$	1	3	$ARIMA(2,1,0)$		1
$ARIMA(0,1,6)$		1	$ARIMA(3,1,0)$	1	4
$ARIMA(0,1,7)$	1	2	$ARIMA(4,1,0)$		3
$ARIMA(0,1,8)$	1	4	$ARIMA(5,1,0)$		1
			$ARIMA(6,1,0)$	1	1
			$ARIMA(7,1,0)$	1	

Each fitted series was obtained with the model of lowest order (ties were broken in favor of the autoregressive models). Comparing the two subgraphs, it does seem that the $ARIMA$ meta-models managed to capture the underlying trend of the response.

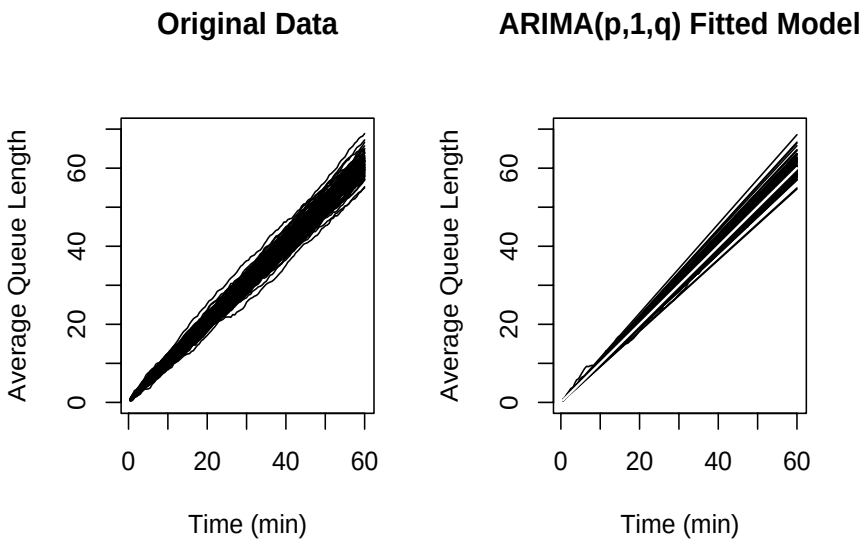

Fig. 4. Original and Fitted Series: Average Queue Length ($M/M/1$ Queue, with $\rho = 2$ and $\Delta_t = 0.5$)

The average values of the differentiated data (for all of the 100 series) are, approximately, 0.5 (for the case $\Delta_t = 0.5$) and 1 (for the case $\Delta_t = 1.0$). This suggests that the average queue length is directly proportional to the elapsed time. This important conclusion is in agreement with a little-known asymptotic result obtained by Bailey (1964) for the $M/M/1$ queue: the expected queue length at instant t can be approximated by $(\lambda - \mu)t$, for $\lambda > \mu$, where λ and μ are the arrival and service rates, respectively. This validates both the approach and the results presented here. In Figure 4, we also represent Bailey's result (the thick white line in the subgraph to the right).

We performed a similar analysis for the other response, the average sojourn time. This time, differentiation failed to produce white noise, except for a few data series and only

for $\Delta_t = 1.0$. In Figure 5, we represent the original series, its first difference and the corresponding ACF and PACF.

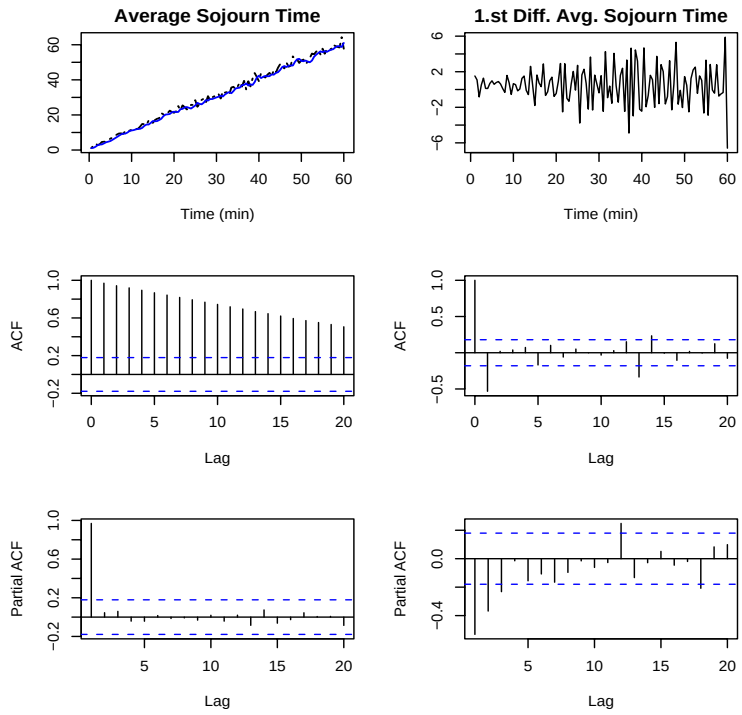

Fig. 5. Average Sojourn Time ($M/M/1$ Queue, with $\rho = 2$ and $\Delta_t = 0.5$)

In Table 3, we report the number of different $ARIMA(p, 1, q)$ models that were actually fitted to the 100 series of average sojourn times. We see that the order of the $ARIMA(p, 1, q)$ model required to cover about 90% of the original series increased substantially. Looking at Figure 6, it becomes clear that the fitted models now keenly capture the global behavior of the data series. However, in this case, no single model seems to dominate the others.

3.2. *A Critical* $M/M/2$ *Queue*

The situation with a critical value for the utilization factor, $\rho = \lambda/(s\mu)$, was created with an arrival rate of $\lambda = 2$ and a service rate of $\mu = 1$, producing an utilization factor of $\rho = 1$.

Table 3. Valid Fits for Average Sojourn Time ($M/M/1$ Queue, with $\rho = 2$)

$ARIMA(0,1,q)$ **FITTED MODEL**	**TIME INTERVAL** $\Delta_t = 0.5$	$\Delta_t = 1.0$	$ARIMA(p,1,0)$ **FITTED MODEL**	**TIME INTERVAL** $\Delta_t = 0.5$	$\Delta_t = 1.0$
$ARIMA(0,1,0)$		24	$ARIMA(0,1,0)$		24
$ARIMA(0,1,1)$	16		$ARIMA(1,1,0)$	13	
$ARIMA(0,1,2)$	16	21	$ARIMA(2,1,0)$	28	13
$ARIMA(0,1,3)$	19	24	$ARIMA(3,1,0)$	8	24
$ARIMA(0,1,4)$	23	26	$ARIMA(4,1,0)$	2	22
$ARIMA(0,1,5)$	13	2	$ARIMA(5,1,0)$	15	10
$ARIMA(0,1,6)$	8	3	$ARIMA(6,1,0)$	17	3
$ARIMA(0,1,7)$	5		$ARIMA(7,1,0)$	10	3
			$ARIMA(8,1,0)$	7	1

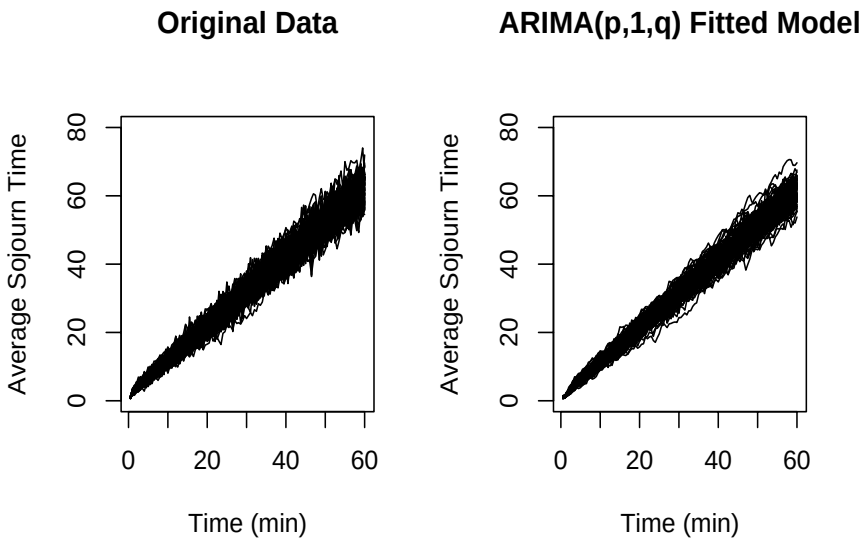

Fig. 6. Original and Fitted Series: Average Sojourn Time ($M/M/1$ Queue, with $\rho = 2$ and $\Delta_t = 0.5$)

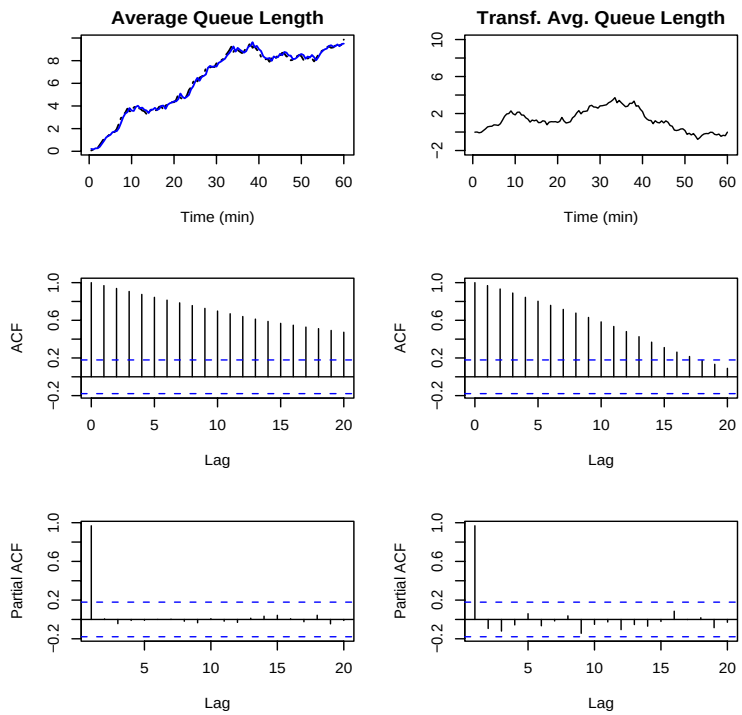

Fig. 7. Average Queue Length ($M/M/2$ Queue, with $\rho = 1$ and $\Delta_t = 0.5$)

Again, a clear linear trend can be observed in the average queue length series (dash-line in Figure 7). As before, the slow decaying of the ACF towards zero indicates a non-stationarity on the mean. When we repeated the application of the Box-Jenkins methodology to the 100 average queue length responses, we were again able to fit a large number of $ARIMA(0,1,0)$ models. This seemed to suggest that simple differentiation would produce white noise. However, in spite of this promising start, this case ended up representing the greatest challenge to our analysis. In fact, the predicted linear evolution failed to detect a marked nonlinear start of the actual responses; see Figure 8. Since differentiation had failed to stabilize the variance of the process, we hypothesized that this response might constitute a trend stationary process (Hamilton 1994):

$$X_t = \alpha + \delta t + \psi(B)\varepsilon_t$$

where the process mean grows linearly with time, $E[X_t] = \alpha + \delta t$, B is the usual backward

Table 4. Valid Fits for Average Queue Length ($M/M/2$ Queue, with $\rho = 1$)

$ARMA(0,q)$ **FITTED MODEL**	**TIME INTERVAL** $\Delta_t = 0.5$	$\Delta_t = 1.0$	$ARMA(p,0)$ **FITTED MODEL**	**TIME INTERVAL** $\Delta_t = 0.5$	$\Delta_t = 1.0$
$ARMA(0,1)$		3	$ARMA(1,0)$	91	89
$ARMA(0,2)$		9	$ARMA(2,0)$	3	1
$ARMA(0,3)$	1	16	$ARMA(3,0)$	1	3
$ARMA(0,4)$	2	23	$ARMA(4,0)$	2	6
$ARMA(0,5)$	5	28			
$ARMA(0,6)$	10	11			
$ARMA(0,7)$	9	7			
$ARMA(0,8)$	13	3			
$ARMA(0,9)$	19				
$ARMA(0,10)$	18				
$ARMA(0,11)$	14				
$ARMA(0,12)$	9				

operator, $\sum_{j=0}^{\infty} |\psi_j| < \infty$, roots of $\psi(z) = 0$ are outside the unit circle and ε_t is a white noise sequence with mean zero and variance σ^2. When this is the case, the appropriate treatment for producing a stationary representation is to subtract δt from X_t. The parameter δ was estimated by the average of the first difference of the original X_t series. We then adjusted $ARMA(p,q)$ models to the transformed series, obtaining the results reproduced in Table 4. We can conclude that the autoregressive model of order $p = 1$, $ARMA(1,0)$, clearly stands out of the remaining models that were fitted to the transformed average queue length series. The moving average models, $ARMA(0,q)$, did not perform nearly as well. In Figure 8, we represent the original series and the corresponding fitted series. A typical data series (dash-line) and the corresponding results produced by the fitted model (solid line) are represented in Figure 7. We can see that the global behavior of the original series is now well represented.

With respect to the average sojourn time in the $M/M/2$ queue, with $\rho = 1$, the results

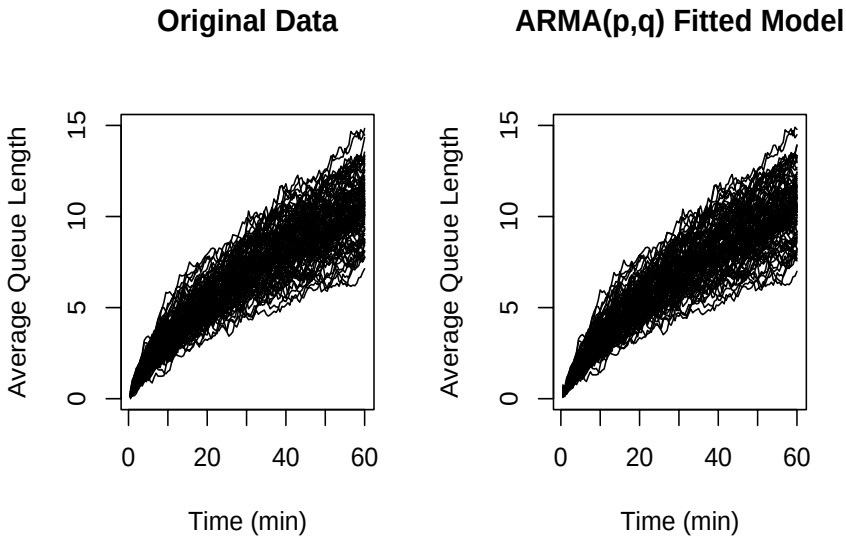

Fig. 8. Original and Fitted Series: Average Queue Length ($M/M/2$ Queue, with $\rho = 1$ and $\Delta_t = 0.5$)

that were obtained were very similar to those of the $M/M/1$ queue (see Figure 9). Thus, we then tried to fit the same type of $ARIMA(p, 1, q)$ models to the 100 series, obtaining the results presented in Table 5.

Again, differentiation did not produce a significant number of cases of white noise. However, since the fitted $ARIMA(p, 1, q)$ models now had higher orders, they aptly captured the marked curvature in the initial behavior of the original series—see Figure 10. In this case, the $ARIMA(0, 1, 1)$ was the most prominent choice.

4. Example: A Single Lane Traffic Analysis

We illustrate our approach with a slight modification of this classical example from the simulation literature. For instance, (Nozari, Arnold and Pegden 1984) used it, as well, in the context of (regression) meta-model estimation. We consider a two-lane road, with the

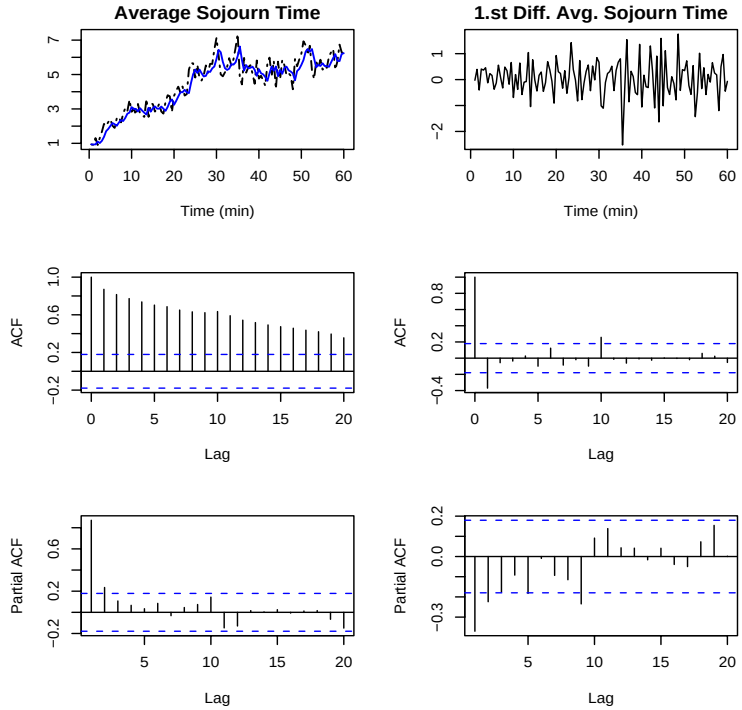

Fig. 9. Average Sojourn Time ($M/M/2$ Queue, with $\rho = 1$ and $\Delta_t = 0.5$)

Table 5. Valid Fits for Average Sojourn Time ($M/M/2$ Queue, with $\rho = 1$)

$ARIMA(0,1,q)$ **FITTED MODEL**	**TIME INTERVAL** $\Delta_t = 0.5$	$\Delta_t = 1.0$	$ARIMA(p,1,0)$ **FITTED MODEL**	**TIME INTERVAL** $\Delta_t = 0.5$	$\Delta_t = 1.0$
$ARIMA(0,1,0)$		30	$ARIMA(0,1,0)$		30
$ARIMA(0,1,1)$	77	57	$ARIMA(1,1,0)$	20	47
$ARIMA(0,1,2)$	9	6	$ARIMA(2,1,0)$	41	16
$ARIMA(0,1,3)$	7	4	$ARIMA(3,1,0)$	20	4
$ARIMA(0,1,4)$	2	2	$ARIMA(4,1,0)$	9	1
$ARIMA(0,1,5)$	2	1	$ARIMA(5,1,0)$	7	2
$ARIMA(0,1,6)$	3		$ARIMA(6,1,0)$	3	

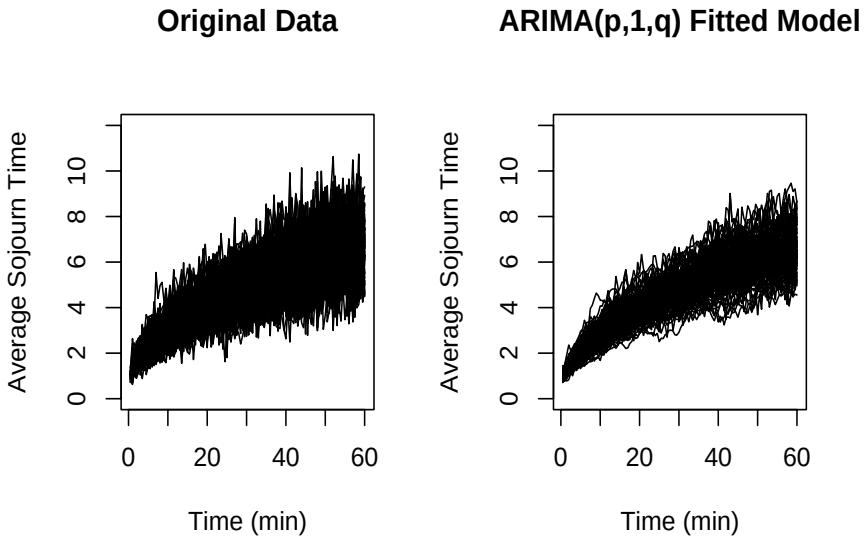

Fig. 10. Original and Fitted Series: Average Sojourn Time ($M/M/2$ Queue, with $\rho = 1$ and $\Delta_t = 0.5$)

traffic flowing from two directions. One of the lanes needed to be repaired, and so it has been closed for 500 m. Traffic lights have been placed at each end of the closed section to control the traffic flow. The lights allow traffic (from only one direction) to flow for a specified time interval. When a light turns green, the waiting cars (from that direction) start and pass the light every 2 sec. When there are no waiting cars and the light is green, any arriving cars will pass through the light without delay. A light cycle consists of green in direction 1, both red, green in direction 2, both red, and then the cycle is repeated. Both lights remain red for 55 sec to allow the cars in transit to leave the section under repair, before traffic from the other direction is initiated. During most of the day, the arrival of cars is exponentially distributed with an expected inter-arrival time of 12 sec from direction 1 and 9 sec from direction 2. However, by the end of the day, there is a two-hour rush period and the expected inter-arrival time from direction 2 is halved to 4.5 sec. The objective is to determine suitable light cycles (for normal and rush-period traffic) so that the average

Table 6. Traffic Analysis: Average Waiting Time *vs.* Cycle Duration

IA Time		**Cycle**	**Waiting Time**			
Dir 1	Dir 2			Dir 1	Dir 2	Global
12	9	55-50-55-60	Average value	78.779	78.685	78.745
			Min. Avg. Value	74.443	70.186	73.431
			Max. Avg. Value	84.682	83.889	82.685
12	4.5	55-50-55-60	Average value	78.786	1441.508	920.602
			Min. Avg. Value	71.167	1240.517	790.169
			Max. Avg. Value	89.912	1589.726	1015.058
12	4.5	55-50-55-135	Average value	200.135	115.173	138.687
			Min. Avg. Value	136.008	77.115	99.889
			Max. Avg. Value	336.050	195.004	197.619

waiting times (for both directions) are minimized.

The first part of the problem is relatively conventional. Simulating green-light times at 5 sec intervals around 60 sec and analyzing the resulting average waiting times, we concluded that the cycle 55-50-55-60 performed better; see Table 6.

Although they do not seem to be very relevant in terms of the problem statement, the average queue lengths (from each of the directions) constitute an invaluable diagnostic measure of the system status. Since the system behavior is cyclic, we collected queue lengths (from one direction) immediately before the light turned green for the other direction. Applying our approach to the data corresponding to the tuned 55-50-55-60 cycle, we were able to conclude that they constituted white noise and that the process was stationary (see Figure 11).

However, the two-hour rush period is a more interesting situation. If we do not change the light cycle, our approach allows us to conclude that the process is no longer stationary: the average queue length from direction 2 will increase linearly (see Figure 12). We fitted the $ARIMA(0,1,0)$ model to the average queue length from direction 2 (from direction 1 is still white noise). Looking at Table 6, we can see that the average waiting time from

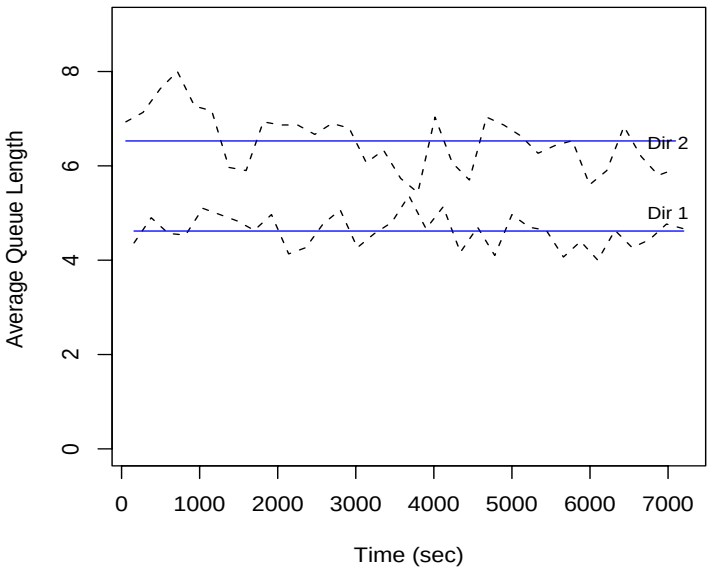

Fig. 11. Original and Fitted Series: Average Queue Length (Single Lane, tuned cycle 55-50-55-60)

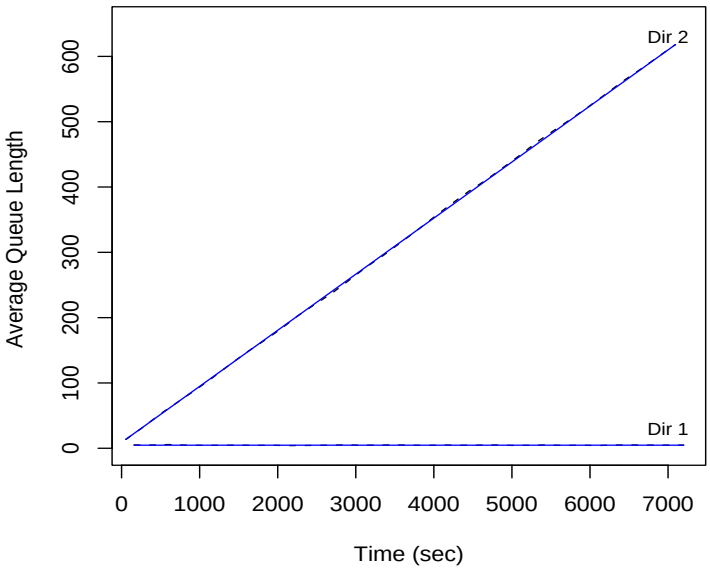

Fig. 12. Original and Fitted Series: Average Queue Length (Single Lane, rush cycle 55-50-55-60)

direction 1 remained unchanged, but from direction 2 increased 18 times. Would it be possible to modify the cycle so that none of the queue lengths or the waiting times become uncontrollable? Considering again 5 sec increments for the green-light time from direction 2 and the ratio of the arrival rates, we found out that the rush cycle 55-50-55-135 performed rather well. The average waiting time from direction 2 is now significantly smaller than the one from direction 1, but, on the other hand, the average queue length is substantially larger (two contradictory goals). In Figure 13, we represent the average queue lengths from both directions, as well as the $ARMA(1,0)$ suitable for representing both series.

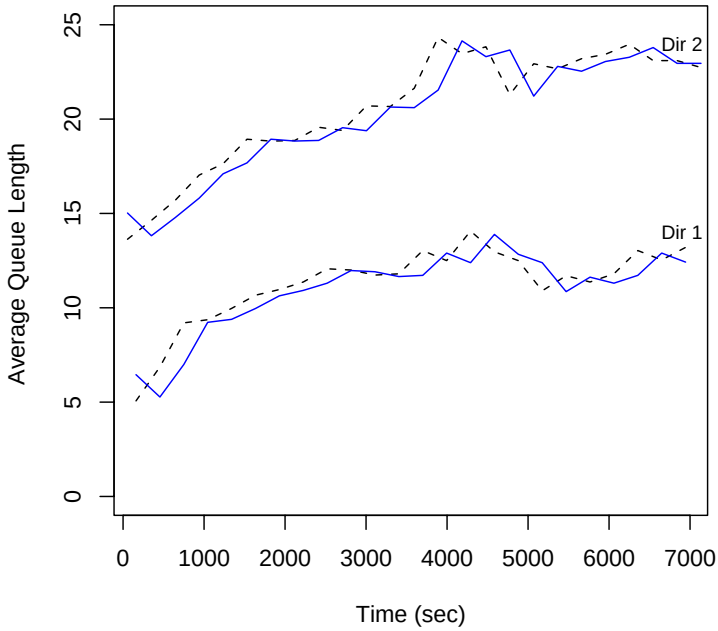

Fig. 13. Original and Fitted Series: Average Queue Length (Single Lane, rush cycle 55-50-55-135)

5. Conclusions and Recommendations

There has not been much research work on the output analysis of non-stationary simulations. But in this paper, we propose an innovative approach, based on classical time series methods, that is valid, and appropriate for stationary simulations as well. An experimental evaluation of two queueing systems showed that our approach is an effective tool for building meaningful meta-models of selected system responses. Its use for decision support was illustrated using a critical traffic situation.

We feel that this may be only the beginning of new significant developments in simulation output analysis. Namely, this approach can be extended to multiple cases of terminating simulations. Additional applications, new analytical meta-models, more comprehensive experimental evaluations... For what it is worth, we intend to continue exploring this promising area of undisputable practical interest.

References

Bailey, N.T.J. (1957) Some further results in the non-equilibrium theory of a simple queue. *J. R. Statist. Soc.*, **B19**, 326–333.

Bailey, N.T.J. (1964) *The Elements of Stochastic Processes with Applications to the Natural Sciences*, John Wiley, New York, NY.

Box, G.E.P., Jenkins, G.M., and Reinsel, G.C. (1994) *Time Series Analysis: Forecasting and Control*, 3rd edn, Prentice Hall, Englewood Cliffs, NJ.

Brandão, R.M., and Porta Nova, A.M.O. (1999) An experimental evaluation of methods for simulation output analysis, in *Proceedings of the 1999 European Simulation Symposium* Horton, G., Möller, D., and Rüde, U. (eds), The Society for Computer Simulation, pp. 601–607.

Brandão, R.M., and Porta Nova, A.M.O. (2003a) Análise de resultados da simulação de filas

de espera não estacionárias usando séries cronológicas, in *Literacia e Estatística – Actas do X Congresso Anual da Sociedade Portuguesa de Estatística*, Brito, P., Figueiredo, A., Sousa, F., Teles, P., and Rosado, F. (eds), Sociedade Portuguesa de Estatística, pp. 133–140.

Brandão, R.M., and Porta Nova, A.M.O. (2003b) Non-stationary queue simulation analysis using time series, to appear in *Proceedings of the 2003 Winter Simulation Conference*, Chick, S., Sánchez, P.J, Ferrin, D., and Morrice, D.J., (eds), Institute of Electrical and Electronics Engineers, Piscataway, NJ.

Fishman, G.S. (1971) Estimating sample size in computing simulation experiments. *Management Science* **18**(1), 21–38.

Hamilton, J.D. (1994) *Time Series Analysis*, Princeton University Press.

Nozari, A., Arnold, S.F., and Pegden, C.D. (1984) Control variates for multipopulation simulation experiments. *IIE Trans.* **16**,159–169.

Schriber, T.J., and Andrews, R.W. (1984) ARMA-based confidence intervals for simulation output analysis. *Am. J. Math. Management Science* **4** (3 & 4), 345–373.