Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

INPUT UNCERTAINTY QUANTIFICATION FOR SIMULATION MODELS WITH PIECEWISE-CONSTANT NON-STATIONARY POISSON ARRIVAL PROCESSES

Lucy E. Morgan, Andrew C. Titman, David J. Worthington

Statistics and Operational Research Centre for Doctoral Training in Partnership with Industry
Lancaster University
Lancaster, LA1 4YR, UK

Barry L. Nelson

Department of Industrial Engineering & Management Sciences
Northwestern University
Evanston, IL 60208 USA

ABSTRACT

Input uncertainty (IU) is the outcome of driving simulation models using input distributions estimated by finite amounts of real-world data. Methods have been presented for quantifying IU when stationary input distributions are used. In this paper we extend upon this work and provide two methods for quantifying IU in simulation models driven by piecewise-constant non-stationary Poisson arrival processes. Numerical evaluation and illustrations of the methods are provided and indicate that the methods perform well.

1 INTRODUCTION

Within simulation models, more often than not, the true input models driving the system are unknown. These models can sometimes be estimated by observing the real-world system but this causes additional uncertainty to arise within the simulation, known as input uncertainty (IU). Overlooking input uncertainty is still a common error in the simulation community. Once the input distributions are estimated from the data they are typically assumed to be correct; this can be risky if the sample of real-world data is small as the models are unlikely to be correct and could result in misleading outputs. The survey by Barton (2012) showed that in some cases input uncertainty overwhelms stochastic estimation error, the error arising from the generation of random variates during the simulation; it should, therefore, not be ignored.

Recently input uncertainty techniques have been implemented in the commercial software Simio (Simio LLC) making it easier for simulation users to quantify the effect of input uncertainty without having to manually implement a complex statistical procedure. However, this software is limited to i.i.d processes. For a review of input uncertainty quantification techniques see the survey papers by Barton (2012) or Song et al. (2014).

In operations research input models are often arrival processes. Examples include call centers, supply chains or accident and emergency departments where customers or demand can occur according to either a stationary or non-stationary arrival process. Input uncertainty for non-stationary arrival processes is yet to be addressed. This paper aims to fill this gap by quantifying input uncertainty in piecewise-constant, non-stationary Poisson arrival processes. Piecewise-constant arrival rate functions are often used in practice in simulation studies as they provide flexibility and are conveniently fit to count data. They are included in many software packages such as Simio (Simio LLC), SIMUL8 (Simul8 Corporation) and Arena (Rockwell Automation). It is therefore a natural step to want to quantify the uncertainty propagated to the simulation

output due to non-stationary arrival processes. We extend two existing methods for quantifying IU due to i.i.d. input processes to cover non-stationary Poisson processes with piecewise-constant arrival rates estimated from count data. Further, we improve one method by exploiting the knowledge that the process is Poisson allowing it to handle arrival processes with many rate changes. We also demonstrate how change-point analysis can be used to obtain a parsimonious representation of the piecewise-constant arrival rate function.

The paper is organised as follows. In Section 2 we present background on current IU quantification techniques and discuss methods for modelling of non-stationary Poisson arrival processes. Section 3 presents our new methods, building on the work of Cheng and Holland (1997) and Song and Nelson (2015). This is followed by an empirical evaluation and realistic illustration of the methods in Section 4. We finish with conclusions and suggestions for further work.

2 BACKGROUND

An early contribution to the IU literature came from Cheng and Holland (1997) who modelled IU using a Taylor series expansion of the mean response as a function of the input distribution parameters. An adaptation of this method was later given by Lin et al. (2015) making use of internal gradient estimation, derived by Wieland and Schmeiser (2006), to reduce quantification of input uncertainty to a single experiment. An alternative approach was given by Song and Nelson (2015) who present a mean-variance effects model for quantifying IU. This method, although not asymptotically justified, makes intuitive sense as the performance measures are likely to depend greatly on the mean and variance of the input distributions.

There are also Bayesian techniques that can be implemented to assist in quantifying uncertainty. Chick (2001) first employed Bayesian techniques enabling the incorporation of prior knowledge of input distributions into simulation modelling. In this method prior information is used for the selection of the input distributions only and input uncertainty is still calculated using the frequentist approach of finding and subtracting the simulation estimation error from the total uncertainty. Zouaoui and Wilson (2010) extended this technique using the posterior probability of the candidate distributions to weight the simulation response but again use frequentist techniques for IU quantification. Recently Xie et al. (2014) developed a fully Bayesian approach for quantifying uncertainty using Gaussian processes to find the posterior distribution of the simulation performance measure of interest. This is then summarized by a credible interval which can easily be dissected to find an estimate for the input uncertainty.

Modelling non-stationary Poisson arrival processes (NSPPs) is also key to our problem. Using Poisson processes has its advantages: they have good properties that make them easy to simulate using thinning or inversion. Kuhl and Wilson (2009) consider both parametric and non-parametric approximations, with respect to NSPPs.

Our focus in this paper is on count data and we model the rate function, $\lambda(t)$, as a piecewise-constant function over q intervals. The intervals, $(0,t_1],(t_1,t_2],\ldots,(t_{q-1},t_q]$, will represent the intervals over which the rate is unchanged. Chen and Gupta (2011) give a way to identify, from count data, where change points in the rate function occur using hypothesis testing. This technique will be utilised in Section 4 as a pre-processing tool to reduce the number of parameters in our model. Employing piecewise-constant $\lambda(t)$ is justified by Henderson (2003) who showed that asymptotically, increasing the number of observations of a process whilst simultaneously decreasing the interval size leads to the true arrival rate function of interest under mild conditions.

3 METHODS

Before considering IU quantification for piecewise-constant NSPPs we set up our approach by reviewing two existing techniques for quantifying input uncertainty in simulation models with stationary input arrival processes.

For ease of explanation consider simulation of a single queue with two driving processes. Let the true input distributions be denoted by $\mathbf{F}^c$; in reality these distributions are unknown and therefore estimated distributions $\widehat{\mathbf{F}}$ will be used to drive the simulation. We will assume the arrivals follow a Poisson process, with true rate parameter λ^c, denote this by F_λ. The service distribution, depending on the situation, may be estimated by a parametric or non-parametric distribution but for ease of exposition we treat it as a parametric distribution with true parameter $\boldsymbol{\theta}^c$; denote this $\mathrm{F}_{\boldsymbol{\theta}}$. Note that the form of $\mathrm{F}_{\boldsymbol{\theta}}$ will have an effect on the approach we will take. This gives the parameter space $(\lambda^c, \boldsymbol{\theta}^c)$ where $\boldsymbol{\theta}^c$ is a row vector of parameters from the service distribution.

Given real-world data we have independent counts, $N_1, N_2, \ldots, N_{m_\lambda}$ of the arrival process, observed m_λ times over the interval $[0, T)$, and observations $X_1, X_2, \ldots, X_{m_{\boldsymbol{\theta}}}$ of the service process. Therefore $(\lambda^c, \boldsymbol{\theta}^c)$ can be estimated by their maximum likelihood estimators (MLEs) $(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})$. For example assuming the arrivals follow a Poisson process implies that the arrival counts can be represented by a Poisson distribution, $N_1, N_2, \ldots, N_{m_\lambda} \sim \text{Poisson}(\lambda^c \mathrm{T})$, and the MLE of the arrival rate is therefore

$$\widehat{\lambda} = \frac{\Sigma_{i=1}^{m_\lambda} N_i}{m_\lambda \mathrm{T}}.$$

This gives the estimated distributions $\mathrm{F}_{\widehat{\lambda}}$ and $\mathrm{F}_{\widehat{\boldsymbol{\theta}}}$ used to drive the simulation. The simulation goal is to estimate $\eta(\lambda^c, \boldsymbol{\theta}^c)$, the expected value of the output of the simulation given the true input parameters. We describe the output from replication j of the simulation by

$$Y_j(\lambda, \boldsymbol{\theta}) = \eta(\lambda, \boldsymbol{\theta}) + \varepsilon_j(\lambda, \boldsymbol{\theta}) \qquad j = 1, 2, \ldots, n$$

where ε represents stochastic noise and has mean 0 and variance $\sigma^2(\lambda, \boldsymbol{\theta})$. Given the MLEs $(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})$ a nominal performance measure estimate of $\eta(\lambda^c, \boldsymbol{\theta}^c)$ is

$$\bar{Y}(\widehat{\lambda}, \widehat{\boldsymbol{\theta}}) = \frac{\Sigma_{j=1}^{n} Y_j(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})}{n}.$$

This has variance $\mathrm{Var}[\bar{Y}(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})]$ which breaks down into input uncertainty and simulation estimation error. Note that most simulation studies ignore input uncertainty because it is believed to be difficult to quantify. In reality input uncertainty is just the variance of the expected value of the output of the simulation with respect to the estimated parameters $(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})$; this can be denoted by

$$\sigma_I^2 = \mathrm{Var}[\eta(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})] = \mathrm{Var}[\mathrm{E}(Y(\widehat{\lambda}, \widehat{\boldsymbol{\theta}}) | \widehat{\lambda}, \widehat{\boldsymbol{\theta}})].$$

The other contribution to uncertainty in the output of the simulation comes from simulation estimation error caused by the generation of random variates during the simulation. Simulation estimation error is denoted by $\sigma^2(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})/n$ which can be estimated using the sample variance S^2/n.

3.1 Cheng and Holland

Cheng and Holland (1997) consider only parametric distributions as inputs to the simulation model. This simplifies input uncertainty to parameter uncertainty. Using a Taylor Series approximation, if $\eta(\lambda, \boldsymbol{\theta})$ is twice continuously differentiable then to first order it can be expressed as

$$\eta(\widehat{\lambda}, \widehat{\boldsymbol{\theta}}) \approx \eta(\lambda^c, \boldsymbol{\theta}^c) + \nabla\eta(\lambda^c, \boldsymbol{\theta}^c)((\widehat{\lambda}, \widehat{\boldsymbol{\theta}}) - (\lambda^c, \boldsymbol{\theta}^c))^T$$

where $\nabla\eta(\lambda^c, \boldsymbol{\theta}^c)$ is the gradient of the expected value of the performance measure with respect to the input parameters λ and $\boldsymbol{\theta}$. Input uncertainty, $\mathrm{Var}[\eta(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})]$, can then be approximated by

$$\mathrm{Var}[\eta(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})] \approx \nabla\eta(\lambda^c, \boldsymbol{\theta}^c)\mathrm{Var}(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})\nabla\eta(\lambda^c, \boldsymbol{\theta}^c)^T. \tag{1}$$

In reality, none of the terms on the right-hand side of Equation (1) are known and so must be estimated.

If the two input distributions are assumed independent, then $\text{Var}(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})$ can be denoted by

$$\text{Var}(\widehat{\lambda}, \widehat{\boldsymbol{\theta}}) = \begin{pmatrix} \text{Var}(\widehat{\lambda}) & \mathbf{0} \\ \mathbf{0} & \text{Var}(\widehat{\boldsymbol{\theta}}) \end{pmatrix}.$$

This can be estimated by $\widehat{\text{Var}}(\widehat{\lambda}, \widehat{\boldsymbol{\theta}}) = \text{I}^{-1}(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})$, the inverse Fisher information matrix of the MLEs evaluated at $(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})$.

Estimation of the gradient is critical. One method is to use the internal gradient estimator of Wieland and Schmeiser (2006), as seen in Lin et al. (2015). This enables $\nabla\eta(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})$ to be evaluated using no additional simulation effort. We specialise this gradient estimation method to our situation below. Although based on similar ideas, the gradient estimation described here is distinct from the Taylor series expansion of $\eta(\hat{\lambda}, \hat{\boldsymbol{\theta}})$ employed by Cheng and Holland (1997).

Firstly consider a simulation model with a single input distribution, let this describe arrivals to a system and be approximated by real-world data where arrival count observations $N_1, N_2, \ldots, N_{m_\lambda} \sim \text{Poisson}(\lambda^c \text{T})$. We assume arrivals are simulated over the full interval $[0, T)$. From these observations $\widehat{\lambda}$ can be found, this is then, for the purpose of this method, considered to be the true arrival rate λ^c. In replication j, the rate $\widehat{\lambda}$ is used to drive the simulation and the count of the number of simulated arrivals in the interval is recorded. Denote this by d_j for replication j. This count can then be used to re-estimate the arrival rate; we call this estimate $\bar{\lambda}_j$ where $\bar{\lambda}_j = d_j/T$. Note that we assume $\text{E}[\bar{\lambda}_j] = \widehat{\lambda}$ since the parameter $\widehat{\lambda}$ was used to run the simulation over j replications. This results in pairs of observations $(Y_j, \bar{\lambda}_j)$. Assuming the output of the simulation depends on the input models, as is most likely the case, then $(Y_j, \bar{\lambda}_j)$ are expected to be dependent. Moreover, if their joint distribution is assumed to be approximately bivariate normal then

$$\text{E}[Y_j(\widehat{\lambda})|\bar{\lambda}_j] = \eta(\widehat{\lambda}) + \Sigma_{Y\bar{\lambda}}\Sigma^{-1}_{\bar{\lambda}\bar{\lambda}}(\bar{\lambda}_j - \widehat{\lambda}) = \delta_0 + \delta_1\bar{\lambda}_j$$

where $\Sigma_{Y\bar{\lambda}}$ is the covariance between Y_j and $\bar{\lambda}_j$ and $\Sigma_{\bar{\lambda}\bar{\lambda}}$ is the variance of $\bar{\lambda}_j$. Here the derivative of the expected response with respect to λ, the gradient, estimated at $\widehat{\lambda}$ equals $\delta_1 = \Sigma_{Y\bar{\lambda}}\Sigma^{-1}_{\bar{\lambda}\bar{\lambda}}$ which can easily be estimated using least squares regression.

This method can be extended when there are multiple input distributions, which is often the case in simulation models. Recall in our simulation model there is an arrival and service distribution. To find $\bar{\boldsymbol{\theta}}$, for the service distribution, this method is just repeated with respect to $\boldsymbol{\theta}$. This gives n independent and identically distributed (i.i.d) vectors $(Y_j, (\bar{\lambda}_j, \bar{\boldsymbol{\theta}}_j))$, $j = 1, 2, \ldots, n$.

Lin et al. (2015) suggest the joint distribution of $(Y_j, (\bar{\lambda}_j, \bar{\boldsymbol{\theta}}_j))$ should now be considered multivariate normal, a natural extension of the previous approach, which gives

$$\text{E}[Y_j(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})|(\bar{\lambda}_j, \bar{\boldsymbol{\theta}}_j)] = \eta(\widehat{\lambda}, \widehat{\boldsymbol{\theta}}) + \Sigma_{Y(\bar{\lambda}, \bar{\boldsymbol{\theta}})}\Sigma^{-1}_{(\bar{\lambda}, \bar{\boldsymbol{\theta}})(\bar{\lambda}, \bar{\boldsymbol{\theta}})}\left((\bar{\lambda}_j, \bar{\boldsymbol{\theta}}_j) - (\widehat{\lambda}, \widehat{\boldsymbol{\theta}})\right)^T = \delta_0 + \boldsymbol{\delta}_1(\bar{\lambda}_j, \bar{\boldsymbol{\theta}}_j)^T.$$

The gradient of $\nabla\eta(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})$ is $\boldsymbol{\delta}_1$ which, again, can be obtained by least squares regression. We now have estimates of both $\text{Var}(\widehat{\lambda}, \widehat{\boldsymbol{\theta}})$ and $\nabla\eta(\lambda^c, \boldsymbol{\theta}^c)$ and can therefore quantify IU using Equation (1).

3.2 Song and Nelson

Song and Nelson (2015) suggest a different approximation of the mean function. Their approach is applicable to both parametric and non-parametric distributions, unlike the approach of Cheng and Holland (1997). We therefore let the output of the j^{th} replication of the simulation, given the collection of input distributions $\widehat{\mathbf{F}}$, be denoted by $Y_j(\widehat{\mathbf{F}}) = \text{E}[Y(\widehat{\mathbf{F}})|\widehat{\mathbf{F}}] + \varepsilon_j$, where the distribution of ε_j could depend on $\widehat{\mathbf{F}}$.

They assume that the output mean $\mathrm{E}[Y(\widehat{\mathbf{F}})|\widehat{\mathbf{F}}]$ can be represented as a function of the mean, $\mu(\widehat{\mathbf{F}})$, and variance, $\sigma^2(\widehat{\mathbf{F}})$, of the input distributions alone. Since $\mathrm{E}[Y(\widehat{\mathbf{F}})|\widehat{\mathbf{F}}]$ is a random variable dependent on $\widehat{\mathbf{F}}$ it can be thought of as a function, $\eta(\widehat{\mathbf{F}})$, which, in the case of our queueing illustration, Song and Nelson (2015) approximate as

$$\eta(\widehat{\mathbf{F}}) \approx \beta_0 + \beta_\lambda\ \mu(F_{\widehat{\lambda}}) + v_\lambda\ \sigma^2(F_{\widehat{\lambda}}) + \beta_{\boldsymbol{\theta}}\ \mu(F_{\widehat{\boldsymbol{\theta}}}) + v_{\boldsymbol{\theta}}\ \sigma^2(F_{\widehat{\boldsymbol{\theta}}}).$$

This is called a mean-variance effects model, and it can be extended to any number of stationary input distributions.

Song and Nelson (2015) fit this model by generating B bootstrap samples from $\widehat{\mathbf{F}}$, then using the empirical distribution of these bootstrap samples, $\widehat{\mathbf{F}}^\star$, to drive B simulations. Empirical distributions are used to obviate the need to refit a parametric distribution to each bootstrap sample from $\widehat{\mathbf{F}}$, and because it makes certain variance and covariance terms (see below) easier to compute.

Consider our assumption that the observed arrival counts follow a Poisson process with true rate $\lambda\mathrm{T}$. Here the mean and variance of the observed counts, $\mu(\widehat{F}_\lambda)$ and $\sigma^2(\widehat{F}_\lambda)$, both equal $\widehat{\lambda}\mathrm{T}$, simplifying the mean-variance model. Note that Song and Nelson (2015) did not consider the use of counts to estimate the arrival rate, as we do here. Now only one regression coefficient is needed to represent the arrival process

$$\eta(\widehat{\mathbf{F}}) \approx \beta_0 + \beta_\lambda\ \mu(F_{\widehat{\lambda}}) + \beta_{\boldsymbol{\theta}}\ \mu(F_{\widehat{\boldsymbol{\theta}}}) + v_{\boldsymbol{\theta}}\ \sigma^2(F_{\widehat{\boldsymbol{\theta}}}). \tag{2}$$

In addition, in this case the bootstrap samples are easy to fit to a Poisson process using the MLE, $\widehat{\lambda}$. Therefore, the bootstrap simulations can be driven by Poisson processes rather than empirical distributions; this makes the method more accurate. These two insights are key to our approach.

From Equation (2) we derive input uncertainty, σ_I^2,

$$\begin{aligned}\mathrm{Var}[\eta(\widehat{\mathbf{F}})] &= \mathrm{Var}[\beta_0 + \beta_\lambda\ \mu(F_{\widehat{\lambda}}) + \beta_{\boldsymbol{\theta}}\ \mu(\widehat{F}_{\boldsymbol{\theta}}) + v_\theta\ \sigma^2(\widehat{F}_{\boldsymbol{\theta}})],\\ &= \beta_\lambda^2\ \mathrm{Var}[\mu(F_{\widehat{\lambda}})] + \beta_{\boldsymbol{\theta}}^2\mathrm{Var}[\mu(\widehat{F}_{\boldsymbol{\theta}})] + v_{\boldsymbol{\theta}}^2\mathrm{Var}[\sigma^2(\widehat{F}_{\boldsymbol{\theta}})] + 2v_{\boldsymbol{\theta}}\beta_{\boldsymbol{\theta}}\ \mathrm{Cov}[\mu(\widehat{F}_{\boldsymbol{\theta}}), \sigma^2(\widehat{F}_{\boldsymbol{\theta}})],\end{aligned} \tag{3}$$

assuming independence among the input distributions. Expression (3) can be approximated, through the use of bootstrap sampling, by

$$\mathrm{Var}[\eta(\widehat{\mathbf{F}})] = \mathrm{Var}[\eta(\widehat{\mathbf{F}})|\mathbf{F}^c] \approx \mathrm{Var}[\eta(\widehat{\mathbf{F}}^\star)|\widehat{\mathbf{F}}],$$

where

$$\mathrm{Var}[\eta(\widehat{\mathbf{F}}^\star)|\widehat{\mathbf{F}}] = \beta_\lambda^2\ \mathrm{Var}[\mu(F_{\widehat{\lambda}}^\star)|F_{\widehat{\lambda}}] + \beta_{\boldsymbol{\theta}}^2\mathrm{Var}[\mu(\widehat{F}_{\boldsymbol{\theta}}^\star)|\widehat{F}_{\boldsymbol{\theta}}] + v_{\boldsymbol{\theta}}^2\mathrm{Var}[\sigma^2(\widehat{F}_{\boldsymbol{\theta}}^\star)|\widehat{F}_{\boldsymbol{\theta}}] + 2v_{\boldsymbol{\theta}}\beta_{\boldsymbol{\theta}}\ \mathrm{Cov}[\mu(\widehat{F}_{\boldsymbol{\theta}}^\star), \sigma^2(\widehat{F}_{\boldsymbol{\theta}}^\star)|\widehat{F}_{\boldsymbol{\theta}}].$$

Firstly looking at the arrival distribution, if we let $F_{\widehat{\lambda}}^\star$ denote the Possion distribution fitted by the parametric bootstrap sample of arrival counts, then $\mathrm{Var}[\mu(F_{\widehat{\lambda}}^\star)|F_{\widehat{\lambda}}] = \mathrm{Var}[\widehat{\lambda}^\star|F_{\widehat{\lambda}}] = \widehat{\lambda}/m_\lambda\mathrm{T}$. For the service process, which we will assume to be non-parametric, $\widehat{F}_{\boldsymbol{\theta}}^\star$, $\mu(\widehat{F}_{\boldsymbol{\theta}}^\star)$ and $\sigma^2(\widehat{F}_{\boldsymbol{\theta}}^\star)$ are given by the mean and second sample central moment of the bootstrapped sample $X_1^\star, X_2^\star, \ldots X_{m_{\boldsymbol{\theta}}}^\star$. As the number of observations increases this approximation is asymptotically justified and expressions for the variance and covariance can be found by

$$\begin{aligned}\mathrm{Var}[\mu(\widehat{F}_{\boldsymbol{\theta}}^*)|\widehat{F}_{\boldsymbol{\theta}}] &= \frac{M_{\boldsymbol{\theta}}^2}{m_{\boldsymbol{\theta}}}\\ \mathrm{Var}[\sigma^2(\widehat{F}_{\boldsymbol{\theta}}^*)|\widehat{F}_{\boldsymbol{\theta}}] &\approx \frac{M_{\boldsymbol{\theta}}^4 - (M_{\boldsymbol{\theta}}^2)^2}{m_{\boldsymbol{\theta}}}\\ \mathrm{Cov}[\mu(\widehat{F}_{\boldsymbol{\theta}}), \sigma^2(\widehat{F}_{\boldsymbol{\theta}}^*)|\widehat{F}_{\boldsymbol{\theta}}] &\approx \frac{M_{\boldsymbol{\theta}}^3}{m_{\boldsymbol{\theta}}}\end{aligned}$$

where $M_{\boldsymbol{\theta}}^k$ is the k^{th} central moment of $\widehat{F}_{\boldsymbol{\theta}}$ and since $\widehat{F}_{\boldsymbol{\theta}}$ is an empirical distribution $M_{\boldsymbol{\theta}}^k = \Sigma_{i=1}^{m_{\boldsymbol{\theta}}}(X_{\boldsymbol{\theta}i} - \bar{X}_{\boldsymbol{\theta}})^k / m_{\boldsymbol{\theta}}$.

To find the regression coefficients the bootstrap experiments are used to fit a regression model. Least squares regression about $\eta(\widehat{\mathbf{F}})$ can then be used to evaluate $\beta_0, \beta_\lambda, \beta_{\boldsymbol{\theta}}$ and $v_{\boldsymbol{\theta}}$. This gives all components needed to calculate input uncertainty using Equation (3).

When deciding which method to use in practice, the form of the input distributions is key, as is the amount of data available. Cheng and Holland (1997) require all input distributions to be parametric and therefore the method could be said to have less flexibility. Conversely, Song and Nelson (2015) can handle both parametric and non-parametric distributions but difficulty arises in computing the variance and covariance terms needed to quantify IU for some parametric distributions. Note that in our case we exploit the fact that for Poisson distributions this is easy.

The use of bootstrapping by Song and Nelson (2015) means given any number of observations of either process we should be able to obtain the same approximation of IU. Unlike the method by Cheng and Holland (1997) which relies on asymptotic theory, as $m \to \infty$, and therefore may not give a good approximation of input uncertainty when m is small. But being asymptotically justified could be seen as an advantage, Song and Nelson (2015) rely on their intuitive model which may not perform well in situations where the output of the simulation cannot be described well by the first two moments of the input distributions.

It will be of interest to see if the strengths and weaknesses of either method translate to cases where non-stationary arrival processes are included in the simulation model; this will be covered in Section 4.

3.3 Non-stationary Arrival Processes

We now present two methods for quantifying input uncertainty in simulation models driven using at least one piecewise-constant, non-stationary Poisson arrival process. These methods build upon the work of Cheng and Holland (1997) and Song and Nelson (2015) but introduce the idea of modelling the input arrival distributions using arrival count observations instead of inter-arrival time observations. The assumption that these arrival counts follow a Poisson distribution is key to our new methods and leads to a useful simplification in both cases.

Consider a piecewise-constant NSPP with q distinct arrival rates over the intervals $[0,t_1)$, $[t_1,t_2),\ldots,[t_{q-1},T)$. Each interval can be considered as a single input distribution to the simulation with the observation interval matching the simulation interval. Again let us consider a simple queueing model with a stationary service distribution and an arrival process described by a piecewise-constant NSPP. The parameter space is now $(\lambda_1,\lambda_2,\ldots,\lambda_q,\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ is a row vector describing the parameters of the service process.

We start by describing the Taylor series approximation method for quantifying input uncertainty in this situation. Observed arrival counts in each interval are independent implying no dependence between $\widehat{F}_{\lambda_1},\widehat{F}_{\lambda_2},\ldots,\widehat{F}_{\lambda_q}$. Equation (1) therefore becomes

$$\sigma_I^2 = \mathrm{Var}[\eta(\widehat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\theta}})] \approx \nabla\eta(\boldsymbol{\lambda}^c,\boldsymbol{\theta}^c)\mathrm{Var}(\widehat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\theta}})\nabla\eta(\boldsymbol{\lambda}^c,\boldsymbol{\theta}^c)^T.$$

This requires estimation of the gradient, $\nabla\eta(\boldsymbol{\lambda}^c,\boldsymbol{\theta}^c)$, and variance matrix, $\mathrm{Var}(\widehat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\theta}})$. The independence of the q arrival processes gives the following diagonal variance matrix

$$\begin{pmatrix} \mathrm{Var}(\widehat{\lambda}_1) & 0 & \cdots & \mathbf{0} \\ 0 & \mathrm{Var}(\widehat{\lambda}_2) & & \\ & & \vdots & \\ & & \mathrm{Var}(\widehat{\lambda}_q) & \mathbf{0} \\ 0 & & 0 & \mathrm{Var}(\widehat{\boldsymbol{\theta}}) \end{pmatrix}.$$

Since the arrival counts are assumed to be Poisson, closed form-equations exist for each $\mathrm{Var}(\widehat{\lambda}_i)$, $i = 1,2,\ldots,q$. Gradient estimation is also no harder in the non-stationary case using the internal gradient estimation method

of Lin et al. (2015). This requires evaluation of $\bar{\lambda}_i$, for $i = 1,2,\ldots,q$ and least squares regression of Y_j with respect to the parameter space $(\bar{\boldsymbol{\lambda}}_j, \bar{\boldsymbol{\theta}}_j)$. One concern with this approach is the validity of the first-order approximation if q becomes large over many short intervals. A possible way around this would be to merge small intervals with similar arrival rates using change-point analysis within the pre-processing stage of the experiment; this idea is explored further in Section 4.

Our second method, to be referred to as the mean-variance approximation, makes use of a mean-variance effects model in the same way as Song and Nelson (2015) but uses arrival counts to model the input arrival distribution instead of inter-arrival times. Again we consider each interval of the arrival process as a distinct distribution, each with arrival rate λ_i, for $i = 1,2,\ldots,q$. Assuming the arrival counts follow a Poisson process means $\mu(F_{\widehat{\lambda}_i}) = \sigma^2(F_{\widehat{\lambda}_i})$ for $i = 1,2,\ldots,q$, as seen in Section 3.2, allowing a simplification of the mean-variance effects model. The arrival process therefore only contributes q elements, $\mu(F_{\widehat{\lambda}_i})$ for $i = 1,2,\ldots,q$, to the mean-variance effects model, rather than $2q$. This is a significant simplificaiton when there are many intervals. Formulae exist for both $\mu(F_{\widehat{\lambda}_i})$ and $\mathrm{Var}[\mu(F_{\widehat{\lambda}_i})]$ making the method simple to implement.

We have presented two techniques for approximately quantifying input uncertainty in simulation models with piecewise-constant non-stationary Poisson input processes. However, it may also be of interest to determine the overall contribution of the arrival process to IU to evaluate whether it overwhelms the uncertainty contribution from other input distributions or whether there is a specific interval that contributes substantially to the total IU. Similarly in a simulation model with L input distributions it would be useful to establish the relative contribution of the l^{th} input distribution to input uncertainty as this can be used to indicate where more data should be collected if follow-up analysis were to be carried out.

When the input distribution is stationary, Lin et al. (2015) and Song and Nelson (2015) give ways to approximate the contribution of the l^{th} input model. These techniques can also be used alongside our two new methods for finding the contribution to IU of the arrival process. Consider the i^{th} interval of the arrival process, $F_{\widehat{\lambda}_i}$. Using a Taylor series expansion its contribution $c_{\widehat{\lambda}_i}(m_{\widehat{\lambda}_i})$, is given by

$$c_{\widehat{\lambda}_i}(m_{\widehat{\lambda}_i}) = \nabla\eta(\widehat{\lambda}_i)\widehat{\mathrm{Var}}(\widehat{\lambda}_i)\nabla\eta(\widehat{\lambda}_i)^T$$

where the gradient, $\nabla\eta(\widehat{\lambda}_i)$, approximates $\partial\eta(\boldsymbol{\lambda},\boldsymbol{\theta})/\partial\lambda_i$ and the variance is given by $\mathrm{Var}(\widehat{\lambda}_i)$. When the mean-variance approximation method is used this translates to

$$c_{\widehat{\lambda}_i}(m_{\widehat{\lambda}_i}) = \beta^2_{\widehat{\lambda}_i}\mathrm{Var}[\mu(F_{\widehat{\lambda}_i})].$$

Now if we were interested in finding the total contribution of the arrival process this is just the sum of the contributions of the q individual intervals

$$c_{\boldsymbol{\lambda}}(m_{\boldsymbol{\lambda}}) = c_{\lambda_1}(m_{\lambda_1}) + c_{\lambda_2}(m_{\lambda_2}) + \cdots + c_{\lambda_q}(m_{\lambda_q}).$$

Whichever approach is used to quantify input uncertainty, an approximation of the contribution of the l^{th} input model to input uncertainty can be found. From here quantifying the relative contribution of the l^{th} input distribution, $r_l(m_l)$, is simply $r_l(m_l) = c_l(m_l)/\sigma^2_I$. This indicates which input distribution contributes the most to IU and therefore where further input data collection may be required.

4 EMPIRICAL EVALUATION

In this section we empirically evaluate and compare our methods using a tractable $M(t)/M/\infty$ queueing model. An illustration of using the methods to quantify IU in a realistic call center setting is also presented to highlight the need for IU quantification in simulation models with non-stationary input processes.

4.1 $M(t)/M/\infty$ Queueing Model

We firstly evaluate our methods by considering the $M(t)/M/\infty$ queueing model since it has well-known behaviour and calculation of the contribution of the i^{th} input distribution, $\text{Var}[\text{E}(\bar{Y}(\widehat{F}_i))|\widehat{F}_i]$ for $i = 1,2,\ldots,p$, is analytically possible. We can therefore assess the quality of the proposed methods from Section 3 against the true values and compare their respective performance.

We investigated the effect of the size of the observed samples of arrival counts and the speed of convergence to steady state within each interval on the performance of our methods. Notice that fast convergence to steady state is analogous to having q distinct $M/M/\infty$ queues and for stationary input distributions we know the mean-variance (M-V) and Taylor series approximation (TSA) methods both perform well. The system performance measure we selected was the expected number in the system over the whole period, $\text{E}(\bar{N})$, which for an infinite server system is also the expected number of busy servers. This measure is linear in λ_i for $i = 1,2,\ldots,q$ and we therefore expected the approximations to be good.

The experiment is as follows. We considered an $M(t)/M/\infty$ queueing system over a $T = 4$ hour period. The arrival rate was assumed to change hourly according to a piecewise-constant function with rates $\lambda(t) = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$; the service distribution was assumed to be stationary with service rate ψ. To mimic the effect of input uncertainty, the system was "observed" for m days recording the arrival counts and approximately $r = m \times 60 \times (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)$ service times, one service time for each arrival. These provided the data for the fitted input models.

The experiment was split into two sub-experiments with different arrival processes and service rates reflecting "quick" and "slow" speeds of convergence to steady state. Within each sub-experiment we tested different values of m to see if the number of observations of the arrival counts has an effect on the performance of either method. To enable comparability between the two sets of experiments, m and r are chosen such that the total number of arrivals is the same for each level of sample size. The square root of the true analytical contribution from each parameter was recorded, for compatibility with the performance measure estimate, along with the percentage relative error of both methods in each scenario. In the M-V method $B = 40$ bootstrap samples each of $n = 500$ replications of the simulation were run. The entire experiment was repeated for $h = 1000$ macro-replications. The averaged results can be found in Tables 1 and 2.

When calculating the analytical values there is no formula for calculating $\text{Var}[\text{E}(\bar{Y}(\widehat{F}_\psi))|\widehat{F}_\psi]$, although for large enough ψ a very close approximation exists. This approximation was used in Experiment 1 but for Experiment 2, where $\psi = 0.05$, it leads to over-estimation. We therefore simulated 1000 values of $\widehat{\psi}$ and

Table 1: Experiment $1(i)$: The analytical contribution of the i^{th} input distribution and the percentage relative errors of the M-V and TSA methods when the arrival process is $\lambda(t) = \left(\frac{1}{3}, \frac{1}{2}, \frac{5}{12}, \frac{1}{3}\right)$ and service rate $\psi = 0.2$. Here $\text{E}(\bar{N}) = 1.94$.

		$\sqrt{\text{Var}[\text{E}(\bar{Y}(\widehat{F}_i))\|\widehat{F}_i]}$						
Sample Size	Method	λ_1	λ_2	λ_3	λ_4	ψ	Total	Magnitude
$m = 2$	Analytical	6.59	8.07	7.37	6.04	14.4	20.1	$\times 10^{-2}$
$r = 190$	M-V (RE%)	0.41	0.30	-0.12	0.22	-3.22	-1.53	
	TSA (RE%)	0.22	0.79	0.62	0.46	-5.93	-2.69	
$m = 20$	Analytical	2.08	2.55	2.33	1.91	4.54	6.37	$\times 10^{-2}$
$r = 1900$	M-V (RE%)	0.79	0.29	0.28	0.67	-3.31	-1.44	
	TSA (RE%)	-0.09	2.52	-0.18	0.53	-3.73	-1.45	
$m = 100$	Analytical	0.93	1.14	1.04	0.85	2.03	2.85	$\times 10^{-2}$
$r = 9500$	M-V (RE%)	3.65	2.06	1.91	3.17	-1.78	0.38	
	TSA (RE%)	1.67	3.25	0.56	1.32	-3.54	-0.86	

Table 2: Experiment $1(ii)$: The analytical contribution of the i^{th} input distribution and the percentage relative errors of the M-V and TSA methods when the arrival process is $\lambda(t) = \left(\frac{1}{12}, \frac{1}{8}, \frac{5}{48}, \frac{1}{12}\right)$ and service rate is $\psi = 0.05$. Here $\mathrm{E}(\bar{N}) = 1.84$.

		$\sqrt{\mathrm{Var}[\mathrm{E}(\bar{Y}(\widehat{F}_i))\|\widehat{F}_i]}$						
Sample Size	Method	λ_1	λ_2	λ_3	λ_4	ψ	Total	Magnitude
$m = 8$	Analytical	6.59	8.06	7.25	4.50	12.6	18.4	$\times 10^{-2}$
$r = 190$	M-V (RE%)	-0.46	-0.34	-0.03	-0.11	-2.87	-1.20	
	TSA (RE%)	-1.02	-0.57	-0.69	-0.16	-5.76	-2.28	
$m = 80$	Analytical	2.08	2.55	2.29	1.42	4.00	5.84	$\times 10^{-2}$
$r = 1900$	M-V (RE%)	-2.39	-1.27	-2.06	-4.19	-2.27	0.07	
	TSA (RE%)	-1.27	-2.53	-0.93	-1.87	-3.62	-0.77	
$m = 400$	Analytical	0.93	1.14	1.03	0.64	1.80	2.62	$\times 10^{-2}$
$r = 9500$	M-V (RE%)	-5.50	-3.96	-4.67	-10.02	-0.3	2.66	
	TSA (RE%)	-1.49	-3.24	-2.79	-2.64	-1.41	0.74	

calculated $\mathrm{E}(\bar{N})$ using the parameter space $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \widehat{\psi})$. The "analytical" values for $\mathrm{Var}[\mathrm{E}(\bar{Y}(\widehat{F}_\psi))|\widehat{F}_\psi]$ reported in Table 2 are therefore the standard deviation of the 1000 observations of $\mathrm{E}(\bar{N})$.

Notice that the analytically calculated contributions for λ_4 and ψ are smaller in Experiment 2 compared to Experiment 1. When convergence to steady state is slow more work is carried out outside of our window of observation and therefore more service times are truncated by the end of the time period. This causes a reduction in variance which explains the discrepancy between the contributions for λ_4 and ψ across the two experiments. All other values match very closely between the experiments because virtually all the work originating in the first three intervals is completed by the end time, 240 minutes, even in the system that settles to steady state more slowly.

From Tables 1 and 2 it is clear that as the amount of input data m increases the contributions decrease, as they should. However, our interest here is in the relative errors of contribution estimation for the M-V and TSA methods. When the contributions are small, precise estimation of them is harder. However, the TSA method is asymptotically valid as $m \to \infty$ so it tends to hold its relative error level across sample sizes. The M-V approach, on the other hand, has relative errors smaller than TSA when m is small, but as m increases the approximate nature of the mean-variance effects model causes the relative errors to increase somewhat. Overall, the M-V method seems to be better when m is small, and TSA is better as m becomes larger. The speed of convergence of the queue to steady state does not seem to affect the performance of our methods for our chosen performance measure. Overall both methods can be said to perform well with most approximations having relative error less than 5%.

4.2 Healthcare Call Center

We will now illustrate the impact of IU quantification in the simulation of a real-world system with a non-stationary input process. We have data from an NHS 111 healthcare call center. In the UK these call centers are used to advise people who have symptoms of an illness but are unsure where to get treatment. The aim is to reduce congestion in hospital EDs or doctors surgeries caused by minor complaints.

The data was split into 96, 15-minute intervals spanning 24 hours. Of the 6 months of data we decided to consider Wednesdays only as public holidays are unlikely to fall mid-week and therefore we would expect no spikes in the arrival rate. Having 6 months worth of data meant we had $m = 26$ Wednesdays to consider and these were averaged to find the mean arrival rate within each interval which became our initial piecewise-constant arrival rate function. While it is clear from the mean arrivals in each time interval that the process is not stationary, extended periods of time where the rate was approximately constant could

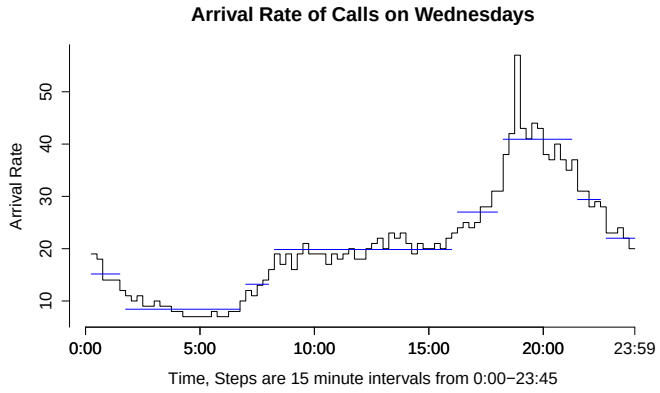

Figure 1: Change point analysis on the arrival rate of calls on Wednesdays.

also be observed. Further, it will be difficult to estimate, and not very meaningful to measure, contributions from 96 tiny intervals. Therefore, rather using a large number of small intervals, or choosing arbitrary large intervals, change-point analysis from Chen and Gupta (2011) was applied to let the data guide when to merge periods where the arrival rates were not significantly different. This resulted in 8 periods of differing length as seen in Figure 1. We would argue that this approach should be used routinely. For the purposes of this analysis, we assume there is no uncertainty in the location of these change points.

In a realistic call center not only does the arrival rate change with time but so too does the number of servers. Therefore, we simulated the 111 call center as an $M(t)/G/s(t)$ queue. From two months of service time-data the mean service time was 8.00 minutes and the standard deviation was 4.33 minutes. A moment matching approach was used to fit a Gamma distribution with shape parameter $\psi_1 = 3.408$ and scale parameter $\psi_2 = 2.347$. Since we wanted to mimic having observed a service time for each arrival, we created a synthetic "observed" service-data set of size $r = 52,711$ observations, corresponding to the expected number of arrivals, and treated this as the real-world data.

The call center's target level of service is $\mathrm{P}(\text{Wait} > 1\,\text{min}) \leq 0.05$ for each caller. Approximately proportional staffing was applied to each time interval and it was found that the waiting time target was met at a level equivalent to 60% utilisation. This is our base case in the experiment as it is likely to be close to the true staffing level the call center used. We also simulated the system with constant staff size tuned to the expected arrival rate over the whole day (Case 1) and to 1.5 times this expected arrival rate (Case 2). These staffing patterns are chosen as they highlight the danger of using stationary approximations of input distributions. In practice someone may use the expected arrival rate over the whole day to set a staffing schedule, ignoring the possibility of fluctuation in the arrival rate.

We investigated performance measures such as the probability of waiting more than 1 minute to be served $\mathrm{P}(\text{Wait} > 1\,\text{min})$, the expected number of people in the queue, $\mathrm{E}(\bar{N})$, and the expected waiting time of customers, $\mathrm{E}(\text{WTime})$ over the whole day. The results for the last of these, $\mathrm{E}(\text{WTime})$, can be seen in Table 3. We used $B = 40$ bootstrap samples, for which $n = 100$ replications of the simulation were carried out for the M-V method. This process was repeated for $h = 1000$ macro-replications of the entire experiment.

Notice first that M-V and TSA give similar, but not identical results. However, they agree on which intervals are the highest and lowest contributors. In Case 1 the contribution of interval 6, $\mathrm{Var}[\mathrm{E}(\bar{Y}(\widehat{F}_{\lambda_6}))|\widehat{F}_{\lambda_6}]$, is much larger than the contribution of any of the other intervals. This coincides with the spike in arrival rate in Figure 1. At this point the queue would be experiencing very high levels of congestion, the number of servers equates to a utilisation of 112.3% which means all servers are always busy. This also seems to have a knock on effect into the next interval, where the contribution of λ_7 is much higher than the contribution of λ_5 even though they have a similar arrival rate. This may be explained by both the backlog

Table 3: The effect of different staffing schemes on the parameter contribution, $\mathrm{Var}[\mathrm{E}(\bar{Y}(\widehat{F}_i))|\widehat{F}_i]$ for input distributions $i = 1,2,\ldots,p$.

		$\mathrm{Var}[\mathrm{E}(\bar{Y}(\widehat{F}_i))\|\widehat{F}_i]$										
Case		λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	$\boldsymbol{\psi}$	Magnitude	E(WTime)
Base	M-V	1.27	9.47	1.31	2.31	0.53	0.51	0.60	0.66	2.41	$\times 10^{-6}$	0.0674
	TSA	1.04	9.37	1.12	2.10	0.34	0.32	0.41	0.45	2.45	$\times 10^{-6}$	0.0674
1	M-V	2.11	2.24	2.22	2.34	2.41	165.53	8.03	2.17	17.55	$\times 10^{-3}$	5.17
	TSA	0.51	0.54	0.57	0.64	0.55	162.16	6.26	0.56	15.57	$\times 10^{-3}$	5.17
2	M-V	3.29	3.27	3.31	3.36	3.28	86.64	3.12	3.15	11.32	$\times 10^{-7}$	0.026
	TSA	2.19	1.99	1.91	2.11	2.09	85.45	1.87	2.09	10.88	$\times 10^{-7}$	0.026

of customers and the arrival rate being above average in the 7^{th} interval. Although the queue is trying to empty, congestion is still high leading to higher uncertainty. By the final interval the system has recovered from the high congestion levels and the contribution of λ_8 is relatively small.

We see a similar but less pronounced effect in Case 2 where again the contribution of λ_6 is larger than the others. This illustrates the importance of understanding the dynamics of IU as the results show how sensitive the overall estimate of performance is to the correct value of arrival rate during the short 6^{th} interval. In the base case we do not see these patterns, the arrival distribution contributions appear to be similar in all but the second interval. When considering E(WTime) the second interval is the most influential; due to the low number of servers this higher contribution was therefore expected.

5 CONCLUSION

This paper presents two methods for quantifying input uncertainty in simulation models with NSPP input processes. The key is the use of count observations to model the arrival processes, meaning each interval of the piecewise-constant rate function can be treated as a distinct, stationary input distribution. From this it is simple to calculate the total contribution to IU of each process and therefore the overall IU. Exploiting the fact that the arrivals are Poisson also allowed us to greatly streamline the method based on a mean-variance effects model.

An evaluation of the performance of the methods was presented using the tractable $M(t)/M/\infty$ model; both methods were seen to perform well. An illustration of a realistic call center scenario was also used to show how input uncertainty quantification in arrival processes may be applied in practice, including the use of change-point analysis to allow the arrival data to guide the choice of time-interval sizes. An open question remains as to the IU that arises from the location of these change points and whether this should be taken into account within the analysis.

REFERENCES

Arena 2015. "Arena". Rockwell Automation. https://www.arenasimulation.com/.

Barton, R. R. 2012. "Tutorial: Input Uncertainty in Output Analysis". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. Uhrmacher, 1–12. Piscataway, New Jersey: IEEE.

Chen, J., and A. K. Gupta. 2011. *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance*. New York: Springer Science & Business Media.

Cheng, R. C., and W. Holland. 1997. "Sensitivity of Computer Simulation Experiments to Errors in Input Data". *Journal of Statistical Computation and Simulation* 57 (1-4): 219–241.

Chick, S. E. 2001. "Input Distribution Selection for Simulation Experiments: Accounting for Input Uncertainty". *Operations Research* 49 (5): 744–758.

Henderson, S. G. 2003. "Estimation for Nonhomogeneous Poisson Processes from Aggregated Data". *Operations Research Letters* 31 (5): 375–382.

Kuhl, M. E., and J. R. Wilson. 2009. "Advances in Modeling and Simulation of Nonstationary Arrival Processes". In *Proceedings of the 2009 INFORMS SSR Workshop*, 1–5.

Lin, Y., E. Song, and B. L. Nelson. 2015. "Single-Experiment Input Uncertainty". *Journal of Simulation* 9:249—-259.

Simio 2015. "SIMIO". Simio LLC. www.simio.com/.

SIMUL8 2015. "SIMUL8". Simul8 Corporation. http://www.simul8.com/.

Song, E., and B. L. Nelson. 2015. "Quickly Assessing Contributions to Input Uncertainty". *IIE Transactions* 47:893–909.

Song, E., B. L. Nelson, and C. D. Pegden. 2014. "Advanced Tutorial: Input Uncertainty Quantification". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 162–176. Piscataway, New Jersey: IEEE Press.

Wieland, J. R., and B. W. Schmeiser. 2006. "Stochastic Gradient Estimation Using a Single Design Point". In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 390–397. Piscataway, New Jersey: IEEE.

Xie, W., B. L. Nelson, and R. R. Barton. 2014. "A Bayesian Framework for Quantifying Uncertainty in Stochastic Simulation". *Operations Research* 62 (6): 1439–1452.

Zouaoui, F., and J. R. Wilson. 2010. "Accounting for Input-Model and Input-Parameter Uncertainties in Simulation". *IIE Transactions* 36 (11): 1135–1151.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the EPSRC funded EP/L015692/1 STOR-i Centre for Doctoral Training, NSF Grant CMMI-1068473 and GOALI sponsor Simio LLC.

AUTHOR BIOGRAPHIES

LUCY E. MORGAN is a Ph.D. student of the Statistics and Operational Research Centre for Doctoral Training in Partnership with Industry at Lancaster University. Her research interests are input uncertainty in simulation models and arrival process modelling. Her email address is l.e.morgan@lancaster.ac.uk.

BARRY L. NELSON is the Walter P. Murphy Professor in the Department of Industrial Engineering and Management Sciences at Northwestern University and a Distinguished Visiting Scholar in the Lancaster University Management School. He is a Fellow of INFORMS and IIE. His research centers on the design and analysis of computer simulation experiments on models of stochastic systems, and he is the author of *Foundations and Methods of Stochastic Simulation: A First Course*, from Springer. His e-mail address is nelsonb@northwestern.edu.

ANDREW C. TITMAN received his Ph.D. from University of Cambridge and currently is Lecturer in Statistics in the Department of Mathematics and Statistics at Lancaster University. His research interests include survival and event history analysis and latent variable modeling, with applications in biostatistics and health economics. His email address is a.titman@lancaster.ac.uk.

DAVID J. WORTHINGTON is a senior lecturer in Operational Research in the Department of Management Science in Lancaster University Management School. He researches the modelling and management of time-dependent queueing systems, and applications of Management Science in healthcare. These research interests often coincide. His email address is d.worthington@lancaster.ac.uk.