Proceedings of the 1995 Winter Simulation Conference
ed. C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman

ESTIMATION AND SIMULATION OF NONHOMOGENEOUS POISSON PROCESSES HAVING MULTIPLE PERIODICITIES

Michael E. Kuhl

Department of Industrial Engineering
North Carolina State University
Raleigh, NC 27695-7906, U.S.A.

James R. Wilson

Department of Industrial Engineering
North Carolina State University
Raleigh, NC 27695-7906, U.S.A.

Mary A. Johnson

Department of Mechanical and Industrial Engineering
University of Illinois–Urbana/Champaign
Urbana, IL 61801, U.S.A.

ABSTRACT

We develop and evaluate procedures for estimating and simulating nonhomogeneous Poisson processes (NHPPs) having an exponential rate function, where the exponent may include a polynomial component or some trigonometric components or both. Maximum likelihood estimates of the unknown continuous parameters of the rate function are obtained numerically, and the degree of the polynomial rate component is determined by a likelihood ratio test. The experimental performance evaluation for this estimation procedure involves applying the procedure to 100 independent replications of nine selected point processes that possess up to four trigonometric rate components together with a polynomial rate component whose degree ranges from zero to three. On each replication of each process, the fitting procedure is applied to estimate the parameters of the process; and then the corresponding estimates of the rate and mean-value functions are computed over the observation interval. Evaluation of the fitting procedure is based on plotted tolerance bands for the rate and mean-value functions together with summary statistics for the maximum and average absolute estimation errors in these functions over the observation interval. The experimental results provide substantial evidence of the numerical stability and usefulness of the fitting procedure in simulation applications.

1 INTRODUCTION

In many simulation studies, we encounter time series of events having long-term trends or multiply periodic behavior. For example in analyzing the arrival streams of liver-transplant donors and patients for the UNOS Liver Allocation Model (Pritsker et al. 1995), we found some arrival rates to exhibit significant growth over time as well as daily, semiweekly, weekly, and annual effects—that is, cyclic patterns of behavior with periods of 1, 3.5, 7, and 365 days, respectively. In this paper we propose methods for estimating and simulating such arrival processes.

The scope of this paper encompasses three main objectives. The first objective is to introduce a time-dependent arrival-process model that is sufficiently flexible to represent a wide variety of input processes arising in large-scale simulation experiments. In particular, we seek an model that accommodates both long-term trends and multiply periodic behavior. The second objective is to formulate and implement a methodology for using our input model in simulation experiments. More specifically, we formulate and implement methods for estimating model parameters from time series data and for simulating the arrival process corresponding to a given set of model parameters. The third objective is to carry out and summarize a systematic experimental performance evaluation of our estimation and simulation procedures.

The proposed time-dependent arrival process is a nonhomogeneous Poisson process (NHPP) with an Exponential-Polynomial-Trigonometric (EPT) rate function having Multiple Periodicities—that is, a rate function of the class we shall label EPTMP. Using an exponential rate function is a convenient means of ensuring that the instantaneous arrival rate is always positive. If the arrival rate includes a long-term evolutionary trend, then this trend is naturally represented in the exponent of the rate function by a polynomial component of appropriate degree. Finally if the arrival rate includes some periodic effects, then each periodic effect is naturally represented in the exponent of the rate function by a trigonometric component with the appropriate oscillation frequency, oscillation amplitude, and phase delay.

In Section 2 of this paper we describe the formulation and implementation of our procedures for (a) estimating the parameters of an NHPP with an EPTMP-type rate function, and (b) simulating the arrival process represented by such a model. Maximum likelihood estimates of the continuous model parameters are computed by a Newton-Raphson search. The simulation procedure generates event times according to the specified EPTMP-type rate function by the method of piecewise inversion. This simulation procedure facilitates the application of standard variance reduction techniques (e.g., common random numbers and antithetic variates). Both the estimation and simulation procedures are implemented in portable, public-domain computer programs that are available on request.

In Section 3 of this paper we present the experimental performance evaluation of the proposed procedures for estimating and simulating NHPPs with EPTMP-type rate functions. To study the properties of the estimation procedure for many types of time-dependent arrival behavior encountered in practice, we present the results of applying the estimation procedure to 100 independent replications of nine selected Poisson processes. In terms of a time unit of one year, the performance evaluation includes NHPPs with up to four periodic rate components representing biennial, annual, semiannual, and quarterly effects; moreover the study includes NHPPs with rate components representing trends over time that are constant, linear, quadratic, and cubic, respectively. The performance evaluation is based on visual inspection of tolerance bands for the rate and mean-value functions as well as summary statistics for the maximum and average absolute errors in these functions over the observation interval.

In Section 4 of this paper we recapitulate our main findings and recommend directions for future work. Most of this paper is based on Kuhl (1994).

2 METHODS FOR ESTIMATING AND SIMULATING NHPPs

2.1 Basic Nomenclature

A nonhomogeneous Poisson process $\{N(t) : t \geq 0\}$ is a generalization of a Poisson process in which the instantaneous arrival rate $\lambda(t)$ at time t is a nonnegative integrable function of time. The mean-value function (or the integrated rate function) of the NHPP is defined by

$$\mu(t) \equiv \mathrm{E}[N(t)] = \int_0^t \lambda(z)\,dz \quad \text{for all} \quad t \geq 0.$$

An EPTMP-type rate function has the form

$$\lambda(t) = \exp\{h(t; m, p, \Theta)\} \tag{1}$$

with

$$h(t; m, p, \Theta) = \sum_{i=0}^{m} \alpha_i t^i + \sum_{k=1}^{p} \gamma_k \sin(\omega_k t + \phi_k),$$

where:

$$\Theta = [\alpha_0, \alpha_1, \ldots, \alpha_m, \gamma_1, \ldots, \gamma_p, \phi_1, \ldots, \phi_p, \omega_1, \ldots, \omega_p]$$

is the vector of continuous parameters. The first $m+1$ terms in $h(t; m, p, \Theta)$ define a degree-m polynomial function representing the general trend over time. The next p terms in $h(t; m, p, \Theta)$ are trigonometric functions representing cyclic effects exhibited by the process.

In many simulation applications, the frequencies are known from prior information; however, there is a large class of applications for which such prior information is either unavailable or incomplete. To develop a completely general technique for modeling and simulating an NHPP with an EPT-type rate function, (that is, a rate function of the form (1) in which $p = 1$), Lee, Wilson, and Crawford (1991) assumed that the oscillation frequency is unknown and must be estimated along with all of the other parameters of the rate function. To further generalize this estimation procedure to include multiple cyclic effects, we assume that the frequencies, amplitudes, and phases in the EPTMP-type rate function (1) are unknown. In the case that the frequencies are known, the ω-components (that is, the last p components) of the vector Θ can be dropped before applying the parameter estimation technique.

Suppose a sequence of n events are observed at the epochs $t_1 < t_2 < \cdots < t_n$ in a fixed time interval $(0, S]$ as a realization of an NHPP with a rate function of the form (1). The log-likelihood function of Θ, given $N(S) = n$ and $\mathbf{t} = (t_1, t_2, \ldots, t_n)$, is

$$\mathcal{L}(\Theta|n, \mathbf{t}) = \sum_{i=0}^{m} \alpha_i T_i + \sum_{k=1}^{p} \sum_{j=1}^{n} \gamma_k \sin(\omega_k t_j + \phi_k) - \int_0^S \exp\{h(z; m, p, \Theta)\}\,dz, \tag{2}$$

where $T_i = \sum_{j=1}^{n} t_j^i$ for $i = 0, 1, \ldots, m$; see Cox and Lewis (1966). Following the procedure of Lee, Wilson, and Crawford (1991), we see that the elements of Θ can be determined by conditioning the estimation of Θ on a fixed value of m and solving the system of likelihood equations given in Appendix A of Kuhl (1994) via the Newton-Raphson method. However, general numerical techniques such as the Newton-Raphson method have proven to be unstable when they are applied to the system of likelihood equations outside of a fairly small neighborhood of the optimal solution. Therefore, good initial estimates of the parameters must be obtained to establish reasonable starting values for the Newton-Raphson procedure. As detailed in Subsections 2.2 and 2.3 below, our procedure for computing initial parameter estimates is a natural extension to the case of multiply periodic behavior of the analogous procedure developed by Lee, Wilson, and Crawford (1991) for the case of a single periodic rate component. Finally, the appropriate value of m is determined using the likelihood ratio test discussed in Subsection 2.4.

2.2 Computing Initial Estimates for Trigonometric Parameters

The initial estimates for the frequencies $\{\omega_1, \ldots, \omega_p\}$ can be obtained either from prior information about the process or from a standard spectral analysis of the series of events (Lewis 1970). For some illustrative examples, see Kuhl (1994).

Based on the initial estimates of the frequencies $\{\omega_1, \ldots, \omega_p\}$, the respective initial estimates of the amplitudes $\{\gamma_1, \ldots, \gamma_p\}$ and the phases $\{\phi_1, \ldots, \phi_p\}$ can be determined as follows. Assuming temporarily that there is no long-term evolutionary trend over the observation interval $(0, S]$, we see that the rate function for this multiply periodic process has the form

$$\lambda(t) = \exp\left[\alpha + \sum_{k=1}^{p} \gamma_k \sin(\omega_k t + \phi_k)\right] \tag{3}$$

for $t \in (0, S]$; and the log-likelihood function is

$$\begin{aligned}
&\mathcal{L}(\Theta|n, \mathbf{t}) \\
&= n\alpha - e^{\alpha} \int_0^S \exp\left[\sum_{k=1}^{p} \gamma_k \sin(\omega_k z + \phi_k)\right] dz \\
&\quad + \sum_{k=1}^{p} \sum_{j=1}^{n} \gamma_k \sin(\omega_k t_j + \phi_k),
\end{aligned} \tag{4}$$

where n is the number of arrivals in $(0, S]$.

A key hypothesis of our approach is that we can obtain good initial estimates of the parameters of the EPTMP-type rate function by estimating the parameters of each periodic component independently. The basis for this approach is the following approximation,

$$\begin{aligned}
&\int_0^S \left\{\prod_{k=1}^{p} \exp[\gamma_k \sin(\omega_k z + \phi_k)]\right\} dz \\
&\simeq \prod_{k=1}^{p} \left\{S^{-1/p} \int_0^S \exp[\gamma_k \sin(\omega_k z + \phi_k)]\, dz\right\}.
\end{aligned} \tag{5}$$

Kuhl (1994) provides a heuristic justification for (5). As shown in Kuhl (1994), each factor on the right-hand side of (5) can be written as

$$S^{-1/p} \int_0^S \exp[\gamma_k \sin(\omega_k z + \phi_k)]\, dz = S^{(p-1)/p} I_0(\gamma_k), \tag{6}$$

$k = 1, 2, \ldots, p$, where $I_j(\cdot)$ is a modified Bessel function of the first kind of order j for $j = 0, 1$ (Abramowitz and Stegun 1965). We combine (5) and (6) and insert the result into the formula (4) for the log-likelihood function; moreover in (4) we use the familiar addition formula

$$\sin(\omega_k t_j + \phi_k) = \sin(\phi_k)\cos(\omega_k t_j) + \cos(\phi_k)\sin(\omega_k t_j)$$

to obtain the following approximation to the log-likelihood function,

$$\begin{aligned}
&\mathcal{L}(\Theta|n, \mathbf{t}) \\
&\simeq n\alpha - e^{\alpha} S^{p-1} \prod_{k=1}^{p} I_0(\gamma_k) + \sum_{k=1}^{p} \gamma_k \sin(\phi_k) A(\omega_k) \\
&\quad + \sum_{k=1}^{p} \gamma_k \cos(\phi_k) B(\omega_k),
\end{aligned} \tag{7}$$

where, for $k = 1, 2, \ldots, p$,

$$A(\omega_k) = \sum_{j=1}^{n} \cos(\omega_k t_j), \tag{8}$$

and

$$B(\omega_k) = \sum_{j=1}^{n} \sin(\omega_k t_j). \tag{9}$$

To obtain initial estimates for the parameters, we calculate the following partial derivatives of the approximate log-likelihood function (7) with respect to each parameter and set the partial derivatives equal to zero. Thus we obtain

$$\frac{\partial \mathcal{L}(\Theta|n, \mathbf{t})}{\partial \alpha} \simeq n - e^{\alpha} S^{p-1} \prod_{k=1}^{p} I_0(\gamma_k) = 0; \tag{10}$$

and for $k = 1, 2, \ldots, p$, we have

$$\frac{\partial \mathcal{L}(\Theta|n, \mathbf{t})}{\partial \gamma_k} \simeq -e^{\alpha} S^{p-1} I_1(\gamma_k) \prod_{\substack{j=1 \\ j \neq k}}^{p} I_0(\gamma_j) + \cos(\phi_k) B(\omega_k) + \sin(\phi_k) A(\omega_k) = 0 \quad (11)$$

and

$$\frac{\partial \mathcal{L}(\Theta|n, \mathbf{t})}{\partial \phi_k} \simeq \gamma_k \cos(\phi_k) A(\omega_k) - \gamma_k \sin(\phi_k) B(\omega_k) = 0. \quad (12)$$

Starting from equations (10)–(12) and applying arguments similar to those given in Lee, Wilson, and Crawford (1991) for each periodic rate component separately, we obtain the following initial estimates for the phase and amplitude of each periodic component:

$$\widehat{\phi}_k = \tan^{-1}\left[\frac{A(\omega_k)}{B(\omega_k)}\right], \quad k = 1, 2, \ldots, p, \quad (13)$$

and

$$\widehat{\gamma}_k \text{ is solution of } \frac{I_1(\widehat{\gamma}_k)}{I_0(\widehat{\gamma}_k)} = \frac{\sqrt{A^2(\omega_k) + B^2(\omega_k)}}{n_k} \quad (14)$$

for $k = 1, 2, \ldots, p$, where n_k is the number of events in the time interval

$$\left(0, \left\lfloor \frac{\omega_k S}{2\pi} \right\rfloor \frac{2\pi}{\omega_k}\right],$$

and $\lfloor z \rfloor$ represents the greatest integer $\leq z$ for all real z. Notice that in equation (14), the respective definitions (8) and (9) for $A(\omega_k)$ and $B(\omega_k)$ are modified slightly so that the upper limit n on these summations is replaced by n_k. Moreover, notice in (13) and (14) that the initial estimates of the amplitude and phase for each periodic rate component depend only on the initial estimate of the frequency for that component.

2.3 Computing Initial Values for Polynomial Coefficients

To determine initial values for the coefficients $\{\alpha_i : i = 0, 1, \ldots, m\}$ of the polynomial rate component, we use a generalized version of the moment matching procedure presented by Lee, Wilson, and Crawford (1991). The first $m+1$ moments of the rate function (1) over the interval $(0, S]$ have the form

$$\int_0^S z^i \lambda(z)\, dz = \int_0^S z^i \exp\{h(z; m, p, \Theta)\}\, dz \quad (15)$$

for $i = 0, 1, \ldots, m$. Setting $\partial \mathcal{L}(\Theta|n, \mathbf{t})/\partial \alpha_i = 0$ in the log-likelihood function (2) for $i = 0, 1, \ldots, m$ and solving for T_i yields

$$T_i = \int_0^S z^i \exp\{h(z; m, p, \Theta)\}\, dz \quad (16)$$

for $i = 0, 1, \ldots, m$. Initially, the procedure determines the coefficients $\{c_j : j = 0, 1, \ldots, m\}$ of an ordinary polynomial $\sum_{j=0}^{m} c_j z^j$ of degree m whose first $m+1$ moments match those of $\exp\{h(z; m, p, \Theta)\}$ over the interval $(0, S]$. From (16), it follows that we seek coefficients $\{c_j : j = 0, 1, \ldots, m\}$ yielding

$$T_i = \int_0^S z^i \left(\sum_{j=0}^{m} c_j z^j\right) dz = \sum_{j=0}^{m} \frac{c_j S^{i+j+1}}{i+j+1} \quad (17)$$

for $i = 0, 1, \ldots, m$. The linear system of equations (17) is solved to obtain initial estimates for the polynomial coefficients $\{c_j\}$.

In the next step, the first $m + 1$ moments of the function $\log\left(\sum_{j=0}^{m} c_j z^j\right)$ over the interval $(0, S]$ are matched with the moments of $h(z; m, p, \Theta)$ to determine the initial values for $\{\alpha_i : i = 0, 1, \ldots, m\}$. This yields the equation system

$$\begin{aligned} \int_0^S z^i \log\left(\sum_{j=0}^{m} c_j z^j\right) dz \quad (18) \\ = \int_0^S z^i \left[\sum_{j=0}^{m} \alpha_j z^j + \sum_{k=1}^{p} \gamma_k \sin(\omega_k z + \phi_k)\right] dz \\ = \sum_{j=0}^{m} \frac{\alpha_j S^{i+j+1}}{i+j+1} + \sum_{k=1}^{p} \gamma_k \mathcal{M}_{\sin}(i, S; \omega_k, \phi_k), \end{aligned}$$

for $i = 0, 1, \ldots, m$, where

$$\mathcal{M}_{\sin}(i, S; \omega_k, \phi_k) \equiv \int_0^S z^i \sin(\omega_k z + \phi_k)\, dz,$$

the ith moment of $\sin(\omega_k t + \phi_k)$ over the interval $(0, S]$. Equation system (18) can be rewritten as

$$\int_0^S z^i \log\left(\sum_{j=0}^{m} c_j z^j\right) dz - \sum_{k=1}^{p} \gamma_k \mathcal{M}_{\sin}(i, S; \omega_k, \phi_k) = \sum_{j=0}^{m} \frac{\alpha_j S^{i+j+1}}{i+j+1} \quad (19)$$

for $i = 0, 1, \ldots, m$. Solving the linear equation system (19) yields initial estimates of $\{\alpha_i : i = 0, \ldots, m\}$.

These initial estimates of the polynomial parameters along with the initial estimates of the trigonometric parameters provide a reasonable starting point for the Newton-Raphson scheme. However, this starting point does not guarantee the convergence of the Newton-Raphson method to the global optimum of the log-likelihood function (2).

2.4 Computing Final Parameter Estimates

The final parameter estimates for a specified degree m of the polynomial trend are determined using the Newton-Raphson method. Let $\widetilde{\Theta}_m$ denote the final estimate of Θ yielded by the Newton-Raphson method for a fixed degree m of the polynomial rate component. To determine the appropriate degree of the polynomial to include in the rate function (1), we will use an extension of the likelihood ratio test proposed by Lee, Wilson, and Crawford (1991). This statistical test has the null hypothesis that m is the true degree of the underlying EPTMP-type rate function. Under the null hypothesis, the test statistic

$$2\left[\mathcal{L}_{m+1}\left(\widetilde{\Theta}_{m+1}\middle|n,\mathbf{t}\right)-\mathcal{L}_m\left(\widetilde{\Theta}_m\middle|n,\mathbf{t}\right)\right] \qquad (20)$$

is asymptotically chi-square with one degree of freedom as S (and consequently n) tend to infinity; and we exploit (20) to assess the importance of successive increments of the likelihood function as the degree of the estimated trend component is repeatedly incremented by one. The degree of the fitted EPTMP-type rate function is determined to be the smallest value of m for which the difference (20) is not significant at a prespecified level of significance. The corresponding vector $\widetilde{\Theta}_m$ provides the final parameter estimates for the underlying NHPP.

To estimate the parameters of an NHPP having an EPTMP-type rate function based on an observed series of event times using the method detailed in this section, we developed the public-domain software package `mp3mle`. The structure and operation of `mp3mle` are detailed in Kuhl (1994).

2.5 Simulating NHPPs

Once a rate function is estimated for the NHPP, the public-domain computer program `mp3sim` can be used to simulate independent replications of the process. The structure and operation of `mp3sim` are discussed in Kuhl (1994). This program simulates an NHPP with an EPTMP-type rate function over the interval $(0, S]$ using the method of piecewise inversion.

With the conventional method of inversion (Bratley, Fox, and Schrage 1987) for an NHPP having rate function $\lambda(z)$, $z \in [0, S]$, the cumulative distribution function of the next event time τ_{i+1} conditioned on the observed value $\tau_i = t_i$ of the current event time is given by

$$F_{\tau_{i+1}|\tau_i}(t|t_i) \equiv \Pr\{\tau_{i+1} \le t|\tau_i = t_i\} = \begin{cases} 1-\exp\left[-\int_{t_i}^{t}\lambda(z)\,dz\right], & \text{if } t \ge t_i, \\ 0, & \text{otherwise.}\end{cases}$$

Thus to sample τ_{i+1} by inversion given $\tau_i = t_i$, we generate a random number U_{i+1} from the uniform distribution on the unit interval $(0, 1)$ and we compute

$$\tau_{i+1} = F^{-1}_{\tau_{i+1}|\tau_i}(U_{i+1}|t_i).$$

This amounts to solving for τ_{i+1} in the equation

$$\int_{t_i}^{\tau_{i+1}} \lambda(z)\,dz = -\ln(1-U_{i+1}).$$

The value of τ_{i+1} can be determined using a search procedure over the interval (t_i, S).

The method of piecewise inversion uses a regularly spaced partition of preevaluated points on the mean-value function to identify rapidly the regular subinterval containing the next event. A modified bisection search is used to estimate the time of the next event with sufficient accuracy so that either

(*a*) The absolute difference between the exact and estimated times of the next event is less than a user-specified tolerance; or

(*b*) When compared to the target value of the mean-value function at the exact time of the next event, the mean-value function at the estimated time of the next event has a relative error less than a user-specified tolerance.

Linear interpolation is used on the first six iterations of the modified bisection search to accelerate the estimation of the next event time. Because the piecewise inversion algorithm always starts from the beginning of a regular subinterval to estimate the next event time, errors incurred in estimating the current or previous event times are not propagated to the next event time.

3 EXPERIMENTATION

3.1 Generation of Experimental Data

To evaluate the procedure for fitting an EPTMP-type rate function to a nonhomogeneous Poisson process having multiple cyclic effects, we chose nine NHPPs which represent processes having up to four cyclic components or a general trend over time or both. Cases 1–6 are NHPPs in which the rate function is an EPTMP-type function. Cases 7–9 are NHPPs with polynomial-trigonometric rate functions having multiple periodicities (that is, rate functions of the class we shall label PTMP) and the form

$$\lambda(t) = \sum_{i=0}^{m} \alpha_i t^i + \sum_{k=1}^{p} \gamma_k \sin(\omega_k t + \phi_k) \qquad (21)$$

for all $t \in (0, S]$. For cases 1–6, the parameters for the polynomial function and the first trigonometric component are taken directly from the experimentation done by Johnson, Lee, and Wilson (1994) on the estimation of an EPT-type rate function. Cases 1–4 consist of exponential rate functions with two periodic components. Case 1 does not contain a general trend over time. Cases 2, 3, and 4 contain general trends which are represented by polynomials of degree 1, 2, and 3, respectively. Rate functions of type EPTMP with three and four periodic components are utilized in cases 5 and 6, respectively. The rate functions of type PTMP in cases 7, 8, and 9 consist of 2, 3, and 4 cyclic components, respectively; and the rate functions contain no general trend.

The parameters of the rate function for each case are shown in Table 1. The frequencies used in the experimentation are expressed in radians per unit time such that $\omega_1 = 2\pi$, $\omega_2 = 4\pi$, $\omega_3 = 8\pi$, and $\omega_4 = \pi$ radians per unit time. If the time period is expressed as one year, then these frequencies represent annual, semiannual, quarterly, and biennial effects, respectively.

Realizations of the selected NHPPs were generated over the interval $(0, S]$ using the program `mp3sim`. For each case, $K = 100$ independent replications were simulated over the interval $(0, 12]$. On each replication, an EPTMP-type rate function was fitted to the observed series of event times. For all of the applications of the estimation procedure, the initial estimates of the frequencies were set equal to the actual frequencies with which the events were generated. In contrast to the experimentation presented by Johnson, Lee, and Wilson (1994) which treated the oscillation frequency as a known parameter, we are assuming the oscillation frequencies are unknown parameters that `mp3mle` must estimate. Estimating the frequencies is fundamentally more difficult than assuming the frequencies are known, and the selected cases 1–9 constitute a more stringent test of the capabilities of the estimation procedure. A significance level of 10% was used for the likelihood ratio test (20) in each case; and the maximum degree of the fitted polynomial was set to four on every application of `mp3mle`.

3.2 Formulation of Performance Measures

To evaluate the performance of `mp3mle`, we used both visual-subjective and numerical goodness-of-fit criteria. The numerical performance measures were first formulated and used by Johnson, Lee, and Wilson (1994) to evaluate the procedure for estimating an EPT-type rate function. These include absolute measures of error for each experiment and relative performance measures that can be compared across the different experiments. For replication k of a given case, the estimated rate function is denoted by $\widetilde{\lambda}_k(t)$ and the estimated mean-value function is denoted by $\widetilde{\mu}_k(t)$. The *average absolute error* in the estimation of the rate function $\lambda(t)$ on the kth replication is

$$\delta_k \equiv \frac{1}{S} \int_0^S \left| \widetilde{\lambda}_k(t) - \lambda(t) \right| dt, \qquad (22)$$

and the *maximum absolute error* is

$$\delta_k^* \equiv \max\left\{ \left| \widetilde{\lambda}_k(t) - \lambda(t) \right| : 0 \le t \le S \right\} \qquad (23)$$

for $k = 1, \ldots, K$. A performance measure for the deviation of the estimated mean-value function from $\mu(t)$ is

$$\Delta_k \equiv \frac{1}{S} \int_0^S \left| \widetilde{\mu}_k(t) - \mu(t) \right| dt, \qquad (24)$$

and the corresponding maximum deviation is

$$\Delta_k^* \equiv \max\left\{ \left| \widetilde{\mu}_k(t) - \mu(t) \right| : 0 \le t \le S \right\} \qquad (25)$$

for $k = 1, \ldots, K$.

Johnson, Lee, and Wilson (1994) developed aggregate performance measures based on (22)–(25) that are computed over all replications of a given experiment. The sample mean of the observations $\{\delta_k : k = 1, \ldots, K\}$ is denoted by $\overline{\delta}$; and the corresponding sample coefficient of variation V_δ is given by

$$V_\delta = \left[\frac{1}{K-1} \sum_{k=1}^{K} \left(\delta_k - \overline{\delta} \right)^2 \right]^{1/2} \Big/ \overline{\delta}. \qquad (26)$$

The statistics $\overline{\delta^*}$ and V_{δ^*} are computed similarly from the observations $\{\delta_k^* : k = 1, \ldots, K\}$. Following Johnson, Lee, and Wilson (1994), we also report

$$Q_\delta \equiv \frac{\overline{\delta}}{\mu(S)/S} \quad \text{and} \quad Q_{\delta^*} \equiv \frac{\overline{\delta^*}}{\mu(S)/S} \qquad (27)$$

since normalizing by the theoretical average arrival rate over $(0, S]$ facilitates comparison of results for different rate functions.

Table 1: Parameters of NHPPs Used in the Experimental Evaluation

Case	1	2	3	4	5	6	7	8	9
α_0	3.6269	3.6269	3.6269	3.6269	3.6269	3.6269	35.0000	35.0000	35.0000
α_1	—	0.1000	−0.1000	−0.4743	—	—	—	—	—
α_2	—	—	0.0200	0.0873	—	—	—	—	—
α_3	—	—	—	−0.0041	—	—	—	—	—
γ_1	1.0592	1.0592	1.0592	1.0592	1.0592	1.0592	10.0000	10.0000	10.0000
ϕ_1	−0.6193	−0.6193	−0.6193	−0.6193	−0.6193	−0.6193	−0.6193	−0.6193	−0.6193
ω_1	6.2831	6.2831	6.2831	6.2831	6.2831	6.2831	6.2831	6.2831	6.2831
γ_2	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	10.0000	10.0000	10.0000
ϕ_2	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	−0.6193	−0.6193	−0.6193
ω_2	12.5664	12.5664	12.5664	12.5664	12.5664	12.5664	12.5664	12.5664	12.5664
γ_3	—	—	—	—	0.2500	0.2500	—	10.0000	10.0000
ϕ_3	—	—	—	—	0.2500	0.2500	—	−0.6193	−0.6193
ω_3	—	—	—	—	25.1327	25.1327	—	25.1327	25.1327
γ_4	—	—	—	—	—	0.7500	—	—	4.0000
ϕ_4	—	—	—	—	—	0.7000	—	—	−0.6193
ω_4	—	—	—	—	—	3.1416	—	—	3.1416

Aggregate performance measures analogous to (26)–(27) are also reported for the errors $\{\Delta_k\}$ and $\{\Delta_k^*\}$ incurred in estimating the mean-value function. In particular, when Δ_k (respectively, Δ_k^*) replaces δ_k in the formulas for $\overline{\delta}$ and V_δ, we obtain the definitions of $\overline{\Delta}$ and V_Δ (respectively, $\overline{\Delta^*}$ and V_{Δ^*}). Moreover, the analogues of Q_δ and Q_{δ^*} are

$$Q_\Delta \equiv \frac{\overline{\Delta}}{\frac{1}{S}\int_0^S \mu(t)\,dt} \quad \text{and} \quad Q_{\Delta^*} \equiv \frac{\overline{\Delta^*}}{\frac{1}{S}\int_0^S \mu(t)\,dt},$$

where the overall average error estimates $\overline{\Delta}$ and $\overline{\Delta^*}$ are expressed as percentages of the average level of the mean-value function over the observation interval $(0, S]$. We remark that the definitions for Q_Δ and Q_{Δ^*} given here differ from the definitions for these quantities used in Johnson, Lee, and Wilson (1994), where $\overline{\Delta}$ and $\overline{\Delta^*}$ were expressed as percentages of the average level of the rate function.

In addition to numerical performance measures, graphical methods are used to provide a visual means of determining the quality of the estimates. For each case, the underlying theoretical rate (respectively, mean-value) function is graphed along with a tolerance band for the estimated rate (respectively, mean-value) function. The approximate tolerance band was obtained as follows for the rate function $\lambda(t)$. For a fixed time $t \in (0, S]$, let

$$\widetilde{\lambda}_{(1)}(t) < \widetilde{\lambda}_{(2)}(t) < \cdots < \widetilde{\lambda}_{(K)}(t)$$

denote the ordered estimates of $\lambda(t)$ obtained on all K replications of the estimation procedure. Then, the following approximate $100(1-\beta)\%$ tolerance interval for $\lambda(t)$ is obtained:

$$\left[\widetilde{\lambda}_{(\lceil K\beta/2\rceil)}(t),\ \widetilde{\lambda}_{(\lceil K\{1-\beta/2\}\rceil)}(t)\right],$$

where $\lceil z\rceil$ denotes the smallest integer greater than or equal to z. For example, if $K = 50$ and $\beta = 0.10$, then the estimated 90% tolerance interval for $\lambda(t)$ at a *single fixed* time $t \in (0, S]$ is $\left[\widetilde{\lambda}_{(3)}(t),\ \widetilde{\lambda}_{(48)}(t)\right]$. Similarly, tolerance intervals are obtained for the mean-value function $\mu(t)$ at a fixed time $t \in (0, S]$.

3.3 Presentation of Results

The statistical results on the estimation of $\lambda(t)$ and $\mu(t)$ for each experimental case are shown in Table 2. Figures 1 and 2 are respectively the graphs of 90% tolerance bands for the rate function and mean-value function in Case 1; and Figures 3 and 4 are respectively the graphs of 90% tolerance bands for the rate function and mean-value function in Case 3. A complete set of figures for all cases is given in Kuhl (1994).

Table 2: Statistics Describing the Errors in Estimating $\lambda(t)$ and $\mu(t)$, $t \in (0,\ 12]$

Case	$\mu(S)$	$\overline{\delta}$	V_δ	Q_δ	$\overline{\delta^*}$	V_{δ^*}	Q_{δ^*}	$\overline{\Delta}$	V_Δ	Q_Δ	$\overline{\Delta^*}$	V_{Δ^*}	Q_{Δ^*}
1	588	4.3	0.32	0.09	16.6	0.43	0.34	11.8	0.70	0.039	23.6	0.65	0.079
2	1126	6.3	0.25	0.07	32.1	0.37	0.34	15.0	0.57	0.032	36.1	0.48	0.078
3	967	6.3	0.32	0.08	45.4	0.41	0.56	13.1	0.63	0.038	33.4	0.60	0.097
4	396	5.5	0.33	0.17	26.7	0.40	0.81	13.9	0.58	0.075	27.9	0.52	0.150
5	599	5.3	0.22	0.11	23.9	0.30	0.48	12.3	0.66	0.040	25.3	0.60	0.083
6	714	6.7	0.26	0.11	41.4	0.29	0.70	14.2	0.71	0.037	28.2	0.57	0.074
7	420	3.9	0.28	0.11	13.2	0.31	0.38	10.4	0.81	0.049	20.1	0.73	0.095
8	420	4.9	0.22	0.14	18.9	0.31	0.54	10.4	0.81	0.049	20.6	0.71	0.097
9	420	5.6	0.17	0.16	22.7	0.29	0.65	10.5	0.81	0.049	21.0	0.70	0.098

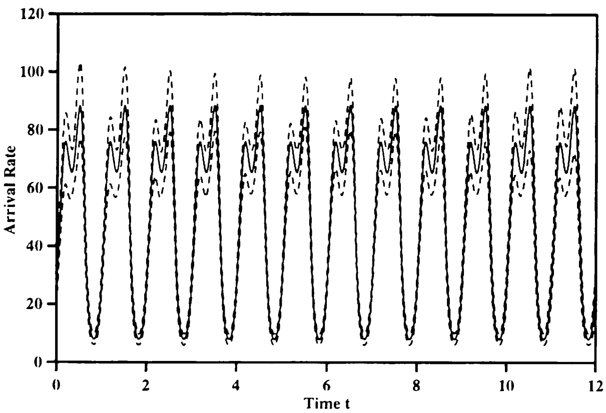

Figure 1: 90% Tolerance Intervals for $\lambda(t)$, $t \in (0, 12]$, in Case 1

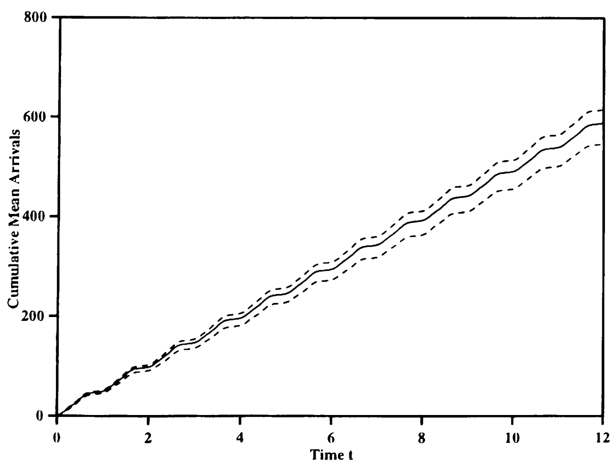

Figure 2: 90% Tolerance Intervals for $\mu(t)$, $t \in (0, 12]$, in Case 1

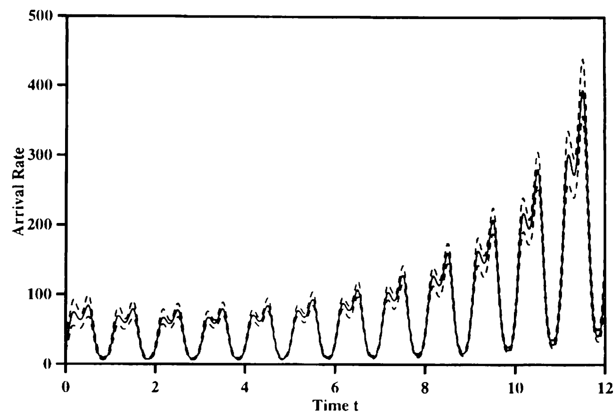

Figure 3: 90% Tolerance Intervals for $\lambda(t)$, $t \in (0, 12]$, in Case 3

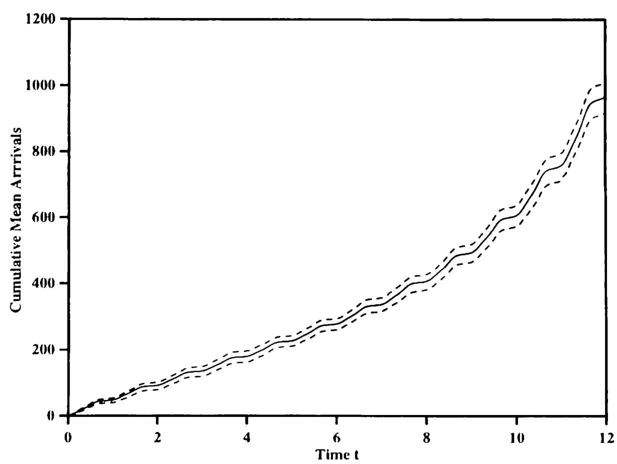

Figure 4: 90% Tolerance Intervals for $\mu(t)$, $t \in (0, 12]$, in Case 3

3.4 Analysis of Results

The numerical results in Table 2 seem to be reasonable values for the selected measures of performance. To evaluate the estimation procedure in general, the normalized statistics can be used. All of the values of these normalized statistics are within the ranges of values presented by Johnson, Lee, and Wilson (1994) for the case of one periodic component. However, some subtle trends do exist in the values of the normalized statistics for these experimental cases. The statistics Q_δ and Q_{δ^*} increase slightly as the number of periodic components p increases. Also, Q_{δ^*} tends to increase as the degree m of the true underlying polynomial trend component increases. The normalized statistics Q_Δ and Q_{Δ^*} for evaluating the estimate of the mean-value function also increase slightly as m increases, and these values are stable as p increases. In addition, the normalized error measures for Cases 7, 8, and 9 (which have PTMP-type rate functions of the form (21)) are generally higher than the corresponding measures for Cases 1, 5, 6 (which have EPTMP-type rate functions with the same values of m and p, respectively). Thus, the PTMP-type rate functions are more difficult to fit than the EPTMP-type rate functions.

Plots of the 90% tolerance bands about the rate function indicate that `mp3mle` is consistently able to fit a reasonable EPTMP-type rate function to the data. However, the plots seem to have the widest tolerance band at the peaks of the rate function. In addition, the plots of the rate function show that the accuracy of the estimation procedure decreases as the number of periodic components increases. This may be due to the key approximation (5), which in general becomes less accurate as p increases. Such a breakdown of (5) may degrade the accuracy of the initial parameter estimates. Thus, the Newton-Raphson procedure may yield a suboptimal solution estimate of the parameter vector Θ.

The plots of the rate functions also indicate that the estimation procedure has more difficulty fitting an EPTMP-type rate function to the underlying PTMP-type rate functions in cases 7, 8, and 9. In particular, the peaks and valleys of the rate function exhibit wide tolerance bands for the fitted rate function. Paralleling the explanation given by Johnson, Lee and Wilson (1994), the major cause of this "ballooning" of the tolerance bands is the symmetrical behavior of the underlying PTMP-type rate function about the polynomial trend, compared to the asymmetric behavior of the fitted EPTMP-type rate function about the corresponding exponential-polynomial function. The error caused by this asymmetric behavior is compounded with the addition of more periodic components. However, the numerical measures and plots do show that an EPTMP-type rate function can be used to approximate reasonably an NHPP with an underlying PTMP-type rate function.

From the plots of tolerance bands for the mean value functions, one can see that each tolerance interval grows steadily over time in each case. This behavior is expected. Since the error is cumulative, the estimation error increases as the mean-value function increases.

4 CONCLUSIONS AND RECOMMENDATIONS

The main contributions of this work are summarized as follows.

- We have introduced an EPTMP-type rate function, which is sufficiently flexible to model a wide variety of input processes that arise in large-scale simulation experiments.

- We have formulated methodologies for estimating parameters of an NHPP with a EPTMP-type rate function and for simulating such a process. Our procedure for obtaining the initial parameter estimates can effectively handle up to four periodic rate components, yielding starting values that are close to the maximum likelihood estimates obtained by the Newton-Raphson method while requiring substantially less computer time than the Newton-Raphson method.

- In building the UNOS Liver Allocation Model, Pritsker et al. (1995) successfully applied `mp3mle` and `mp3sim` to the liver-donor arrival streams at 64 organ procurement organizations and to the liver-patient arrival streams at 111 transplant centers in the United States.

- The estimation procedure and the simulation procedure are implemented in portable, numerically robust software packages which are in the public domain and are available on request.

There are several directions in which this research should be extended.

- The main enhancement to the estimation procedure should be to automate the spectral analysis required to obtain the initial estimates of the frequencies $\{\omega_k : k = 1, 2, \ldots, p\}$.

- The program `mp3mle` should be enhanced to handle more than four periodic rate components and to compute and print the standard errors for each element of $\widetilde{\Theta}_m$.

ACKNOWLEDGMENTS

This work was partially supported by Pritsker Corporation and by the United Network for Organ Sharing (UNOS). The authors thank Dave Martin, Alan Pritsker, and Janet Reust of Pritsker Corporation and Leah Bennett, Pat Daily, Erick Edwards, and Ann Harper of the United Network for Organ Sharing for many helpful suggestions on this work. The authors also thank Halim Damerdji of North Carolina State University for many enlightening discussions on this paper.

REFERENCES

Abramowitz, M., and I. A. Stegun. 1965. *Handbook of mathematical functions.* New York: Dover Publications, Inc.

Bratley, P., B. L. Fox, and L. E. Schrage. 1987. *A guide to simulation.* 2d ed. New York: Springer-Verlag.

Cox, D. R., and P. A. W. Lewis. 1966. *The statistical analysis of series of events.* London: Chapman and Hall.

Johnson, M. A., S. Lee, and J. R. Wilson. 1994. Experimental evaluation of a procedure for estimating nonhomogeneous Poisson processes having cyclic behavior. *ORSA Journal on Computing* 6 (4): 356–368.

Johnson, M. A., S. Lee, and J. R. Wilson. 1994. NPPMLE and NPPSIM: Software for estimating and simulating nonhomogeneous Poisson processes having cyclic behavior. *Operations Research Letters* 15:273–282.

Kuhl, M. E. 1994. Estimation and simulation of nonhomogeneous Poisson processes having multiple periodicities. Master's thesis, Department of Industrial Engineering and Graduate Program in Operations Research, North Carolina State University, Raleigh, North Carolina.

Lee, S., J. R. Wilson, and M. M. Crawford. 1991. Modeling and simulation of a nonhomogeneous Poisson process having cyclic behavior. *Communications in Statistics—Simulation and Computation* 20 (2&3): 777–809.

Lewis, P. A. W. 1970. Remarks on the theory, computation, and application of spectral analysis of series of events. *Journal of Sound and Vibration* 12 (3): 353–375.

Pritsker, A. A. B., D. L. Martin, J. S. Reust, M. A. F. Wagner, J. R. Wilson, M. E. Kuhl, J. P. Roberts, O. P. Daily, A. M. Harper, E. B. Edwards, L. Bennett, J. F. Burdick, and M. D. Allen. 1995. Organ transplantation policy evaluation. In *Proceedings of the 1995 Winter Simulation Conference*, ed. C. Alexopoulos, K. Kang, W. Lilegdon, and D. Goldsman, to appear. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

AUTHOR BIOGRAPHIES

MICHAEL E. KUHL is a Ph.D. student in the Department of Industrial Engineering at North Carolina State University. He received a B.S. degree in industrial engineering from Bradley University in 1992, and he received an M.S. degree in industrial engineering and operations research from North Carolina State University in 1994. He is a licensed Engineer-in-Training in the state of Illinois. He is a member of Alpha Pi Mu and IIE.

JAMES R. WILSON is Professor and Director of Graduate Programs in the Department of Industrial Engineering at North Carolina State University. He received a B.A. degree in mathematics from Rice University, and he received M.S. and Ph.D. degrees in industrial engineering from Purdue University. His current research interests are focused on the design and analysis of simulation experiments. He also has an active interest in applications of operations research techniques to all areas of industrial engineering. From 1988 to 1992, he served as Departmental Editor of *Management Science* for Simulation. He was *Proceedings* Editor for WSC '86, Associate Program Chair for WSC '91, and Program Chair for WSC '92. He has also held various offices in TIMS (now INFORMS) College on Simulation. He is a member of ASA, ACM/SIGSIM, IIE, and INFORMS.

MARY A. JOHNSON is an Assistant Professor in the Department of Mechanical and Industrial Engineering at the University of Illinois–Urbana/Champaign. She received a B.S. degree in mathematics from Adrian College in 1981, and she received M.S. and Ph.D. degrees in industrial engineering from Purdue University in 1986 and 1988 respectively. Her research interests emphasize computational aspects of stochastic models, especially characterizing and modeling stochastic input processes. She is also interested in queueing approximations and performance analysis of computer and communications systems. She is a member of INFORMS and IIE.