Proceedings of the 1998 Winter Simulation Conference
D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan, eds.

LEAST SQUARES ESTIMATION OF NONHOMOGENEOUS POISSON PROCESSES

Michael E. Kuhl

Department of Industrial and
Manufacturing Systems Engineering
Louisiana State University
Baton Rouge, LA 70803-6409, U.S.A.

Halim Damerdji
James R. Wilson

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

ABSTRACT

We formulate and evaluate weighted and ordinary least squares procedures for estimating the parametric rate function of a nonhomogeneous Poisson process. Special emphasis is given to processes having an exponential rate function, where the exponent may include a polynomial component or some trigonometric components or both. Theoretical and experimental evidence is provided to explain some surprising problems with the weighted least squares procedure. The ordinary least squares procedure is based on a square root transformation of the "detrended" event times; and the results of an extensive Monte Carlo study are summarized to show the advantages and disadvantages of this procedure.

1 INTRODUCTION

In this paper we focus on arrival (counting) processes, and more particularly, arrival processes that can be classified as nonstationary point processes. For such processes we are able to observe each arrival time exactly, and in general the arrival intensity (rate) changes over time. Under certain assumptions a nonstationary arrival process can be represented as a nonhomogeneous Poisson process (NHPP) (Çinlar, 1975). Using NHPPs, we can accurately represent a large class of arrival processes encountered in practice.

An NHPP $\{N(t) : t \geq 0\}$ given by

$$N(t) = \text{ \# of arrivals in } [0, t] \quad \text{for all} \quad t \geq 0$$

is a generalization of the Poisson process in which the instantaneous arrival rate $\lambda(t)$ at time t is a nonnegative integrable function of time. The mean-value function of the NHPP is defined by

$$\mu(t) \equiv \mathrm{E}[N(t)] \quad \text{for all} \quad t \geq 0;$$

and the relationship between the rate function and the mean-value function is

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

The probabilistic behavior of the NHPP is completely defined by the rate or mean-value functions. The literature in this area includes both parametric and nonparametric methods for estimating the NHPP rate function. To model arrival processes having several periodic effects or a long-term trend (or both), Kuhl, Wilson, and Johnson (1997) utilized an NHPP whose rate function is of the type exponential-polynomial-trigonometric with multiple periodicities (EPTMP).

The principle of least squares is a method for estimating the parameters of a statistical model fitted to sample data by minimizing an appropriate sum of squared estimation errors. In this paper we investigate least squares methods for fitting NHPPs to arrival processes having parametric rate functions such as an EPTMP-type rate function of the form

$$\lambda(t) = \exp\{h(t; m, p, \Theta)\}, \qquad t \in [0, S], \tag{1}$$

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 least squares procedure will be used to fit the mean-value function $\mu(t)$ to $N(t)$, the observed cumulative number of arrivals at time $t \in [0, S]$.

Least squares has been widely used to fit distribution functions to observed data. For example, Swain, Venkatraman, and Wilson (1988) successfully used least squares procedures to estimate the parameters of cumulative distribution functions (c.d.f.'s) from the univariate Johnson translation system of distributions based on observed data. Similarly, Wagner and Wilson (1996) found least squares to be an effective and computationally efficient method for fitting a univariate Bézier c.d.f. to sample data. Fitting a mean-value function and a c.d.f. are similar in that both are increasing functions which are fitted to the (possibly rescaled) cumulative frequency of occurrence of relevant sample data. Since certain variants of least squares have proven to be advantageous methods for fitting distribution functions, we are motivated to develop appropriate least squares procedures for estimating the mean-value function of an NHPP.

2 METHODOLOGY

2.1 Setup for Least Squares Estimation of NHPPs

For an NHPP $\{N(t) : t \geq 0\}$ in the interval $[0, S]$, let $\{\tau_i : i = 1, 2, \ldots, N(S)\}$ denote the corresponding arrival times. Throughout this paper, we let $\{\tau_i : i = 1, 2, \ldots\}$ denote a sequence of random arrival times; and a realization of this process (that is, an observed sequence of specific arrival times) we will write as $\{t_i : i = 1, 2, \ldots\}$. If we know the functional form of the mean-value function $\mu(t; \boldsymbol{\Theta})$, then we have the relationship

$$\mu(\tau_i; \boldsymbol{\Theta}) = \mathrm{E}[\mu(\tau_i; \boldsymbol{\Theta})] + \varepsilon_i \quad \text{for } i = 1, 2, \ldots, \qquad (2)$$

where ε_i is the random error, i.e. the statistical variation around the mean, and $\mathrm{E}[\varepsilon_i] = 0$. If the errors $\{\varepsilon_i : i = 1, 2, \ldots\}$ were independent and identically distributed (i.i.d.), then we could calculate the ordinary least squares estimates of the parameters, denoted $\widetilde{\boldsymbol{\Theta}}_{\mathrm{OLS}}$, by minimizing the error sum of squares

$$\mathrm{SS_E}(\widehat{\boldsymbol{\Theta}}) = \sum_{i=1}^{N(S)} \left\{ \mu(\tau_i; \widehat{\boldsymbol{\Theta}}) - \mathrm{E}[\mu(\tau_i; \widehat{\boldsymbol{\Theta}})] \right\}^2$$

over all values of $\widehat{\boldsymbol{\Theta}}$ so that we take $\widetilde{\boldsymbol{\Theta}}_{\mathrm{OLS}} = \arg\min_{\widehat{\boldsymbol{\Theta}}} \mathrm{SS_E}(\widehat{\boldsymbol{\Theta}})$ (Seber and Wild 1989).

In the case of an NHPP, the errors $\{\varepsilon_i : i = 1, 2, \ldots\}$ in (2) are neither independent nor identically distributed — in particular, an NHPP has the following probability structure. Given an NHPP $\{N(t) : t \geq 0\}$ with rate function $\lambda(t)$ and mean-value function $\mu(t)$, the sequence of arrival epochs $\tau_1, \tau_2, \ldots$ are event times of this NHPP if and only if the "detrended" arrival epochs

$$\tau_i^* = \mu(\tau_i; \boldsymbol{\Theta})$$

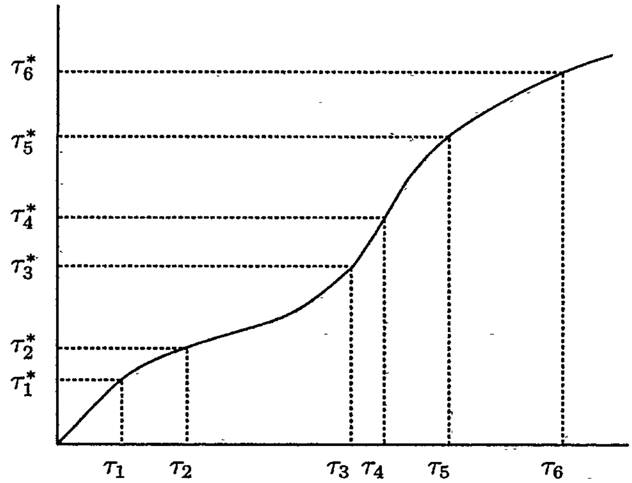

Figure 1: Relationship between Original and Detrended Arrival Times

for $i = 1, 2, \ldots$ constitute a Poisson process with rate 1 (Çinlar 1975). Figure 1 illustrates this relationship.

Since the detrended arrival times $\tau_1^*, \tau_2^*, \ldots$ come from a Poisson process with rate 1, the detrended interarrival times

$$X_i^* = \begin{cases} \tau_i^*, & \text{if } i = 1, \\ \tau_i^* - \tau_{i-1}^*, & \text{if } i = 2, 3, \ldots \end{cases}$$

are i.i.d. exponential random variables with mean 1; and τ_i^* has an i-stage Erlang distribution with scale parameter 1 so that

$$\mathrm{E}[\tau_i^*] = i \quad \text{for} \quad i = 1, 2, \ldots. \qquad (3)$$

Furthermore, the variance of τ_i^* is equal to i, and the covariance between τ_i^* and τ_j^* for $i \leq j$ is

$$\mathrm{Cov}[\tau_i^*, \tau_j^*] = \mathrm{Cov}\left[\sum_{k=1}^{i} X_k^*, \sum_{l=1}^{j} X_l^*\right] = i. \qquad (4)$$

Since the expected value of $\mu(\tau_i; \boldsymbol{\Theta})$ is the constant i, the covariance structure of the errors $\{\varepsilon_i : i = 1, 2, \ldots\}$ in (2) will coincide with the covariance structure of the detrended arrival times given in (4). To exploit the known covariance structure of the "idealized" estimation errors $\{\varepsilon_i : i = 1, 2, \ldots\}$, we developed a weighted least squares (WLS) procedure for estimating the mean-value function of an NHPP as well as an ordinary least squares (OLS) procedure. These methods are examined in the following subsections.

2.2 Weighted Least Squares Estimation of NHPPs

It can be shown that the variance-covariance matrix $\mathbf{V}$ of the idealized residuals $\{\varepsilon_i\}$ in (2) has inverse given by

$$\mathbf{V}^{-1} = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}$$

(Kuhl, Damerdji, and Wilson 1998). In the WLS approach to estimating the mean-value function of a target NHPP, the error sum of squares to be minimized is $\mathrm{SS_E}(\widehat{\Theta}) = \varepsilon^{\mathrm{T}}(\widehat{\Theta})\mathbf{V}^{-1}\varepsilon(\widehat{\Theta})$ over all values of $\widehat{\Theta}$, where the ith element of the vector $\varepsilon(\widehat{\Theta})$ of actual residuals is $\varepsilon_i(\widehat{\Theta}) \equiv \mu(\tau_i;\widehat{\Theta}) - \mathrm{E}[\mu(\tau_i;\widehat{\Theta})]$ for $i = 1, 2, \ldots, N(S)$. In terms of the vector $\mathbf{u}(\widehat{\Theta}) \equiv \mathbf{V}^{-1/2}\varepsilon(\widehat{\Theta})$ of transformed residuals, the WLS estimate of the NHPP parameter vector Θ is given by

$$\widetilde{\Theta}_{\mathrm{WLS}} = \underset{\widehat{\Theta}}{\arg\min} \sum_{i=1}^{N(S)} u_i^2(\widehat{\Theta}),$$

where the ith transformed residual is

$$u_i(\widehat{\Theta}) = \mu(\tau_i;\widehat{\Theta})\sqrt{\tfrac{i+1}{i}} - \mu(\tau_{i+1};\widehat{\Theta})\sqrt{\tfrac{i}{i+1}} \qquad (5)$$

for $i = 1, 2, \ldots, N(S) - 1$; and the last element of $\mathbf{u}$ is

$$u_{N(S)}(\widehat{\Theta}) = \sqrt{\tfrac{1}{N(S)}}\left[\mu(\tau_{N(S)};\widehat{\Theta}) - N(\tau_{N(S)})\right]. \qquad (6)$$

It is clear from (5) and (6) that all information about the discrepancy between the empirical mean-value function $N(\cdot)$ and the fitted mean-value function $\mu(\cdot;\widehat{\Theta})$ has been completely eliminated from the first $N(S) - 1$ elements of $\mathbf{u}$, and only the last element of $\mathbf{u}$ contains any information about the discrepancy between these two functions. It follows that even in the idealized situation in which the weighted least squares estimation procedure starts with perfect (error-free) initial estimates of the unknown parameters so that $\widehat{\Theta} = \Theta$, the value of the objective function $\mathrm{SS_E}(\widehat{\Theta})$ contains relatively little information about how closely the current estimate of the mean-value function approximates the empirical mean-value function. Therefore it should not be surprising if situations arise in which the final WLS estimate of the mean-value function bears almost no reasonable relation to the empirical mean-value function.

Figure 2 shows an example of the anomalous behavior that can result from using the WLS procedure to fit an NHPP. In this example, the empirical mean-value function represents the arrival of patients at a kidney-transplant center in the United States over the time period January 1, 1991 – December 31, 1995. The divergence between the fitted and empirical mean-value functions provides a striking example of the way in which the WLS estimation procedure can fail in practice.

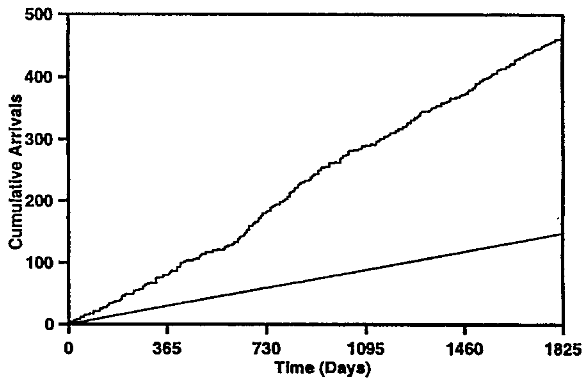

Figure 2: Weighted Least Squares Estimate of the Mean-Value Function (Smooth Curve) versus the Empirical Mean-Value Function of Kidney Transplant Center 103 (Step Function).

2.3 Ordinary Least Squares Estimation of NHPPs

Because of the fundamental problems that we encountered in using the WLS procedure for estimating NHPPs, we developed an alternative approach based on a variance-stabilizing transformation together with an OLS estimation procedure. When building a statistical model for which the variance of the original response variable is proportional to its mean (as in (3) and (4)), a standard variance-stabilizing transformation is to work with the square root of the original response (Box, Hunter, and Hunter 1978). Therefore, we have implemented the following square root transformation to "normalize" and "stabilize the variance" of the dependent variable in our statistical model of the detrended arrival epochs so that the associated idealized residuals have the following form

$$\varepsilon_i = \sqrt{\mu(\tau_i;\Theta)} - \mathrm{E}\left[\sqrt{\mu(\tau_i;\Theta)}\right] \qquad (7)$$

for $i = 1, 2, \ldots$.

In Kuhl, Damerdji, and Wilson (1998) we show that as $i \to \infty$, $\mathrm{E}[\sqrt{\mu(\tau_i;\Theta)}]$ is asymptotic to $\sqrt{i - \frac{1}{4}}$ and $\mathrm{Var}[\sqrt{\mu(\tau_i;\Theta)}] \to \frac{1}{4}$; moreover idealized residuals of the form $\sqrt{\mu(\tau_i;\Theta)} - \sqrt{i - \frac{1}{4}}$ converge in distribution to a normal distribution with mean zero and variance $\frac{1}{4}$,

$$\sqrt{\mu(\tau_i;\Theta)} - \sqrt{i - \tfrac{1}{4}} \xrightarrow[i\to\infty]{\mathcal{D}} N(0, \tfrac{1}{4}).$$

Thus, we see that the square root transformation does in fact stabilize the variance of the idealized residuals $\{\varepsilon_i\}$ in (7); and we obtain the variance stabilized–ordinary least squares estimate for the parameter vector Θ as

$$\widetilde{\Theta}_{\text{OLS}} = \arg\min_{\widehat{\Theta}} \sum_{i=1}^{N(S)} \left(\sqrt{\mu(\tau_i;\widehat{\Theta})} - \sqrt{i - \tfrac{1}{4}} \right)^2. \quad (8)$$

The next step is to identify an appropriate numerical procedure for minimizing the sum of squared errors on the right-hand side of (8).

3 PARAMETER ESTIMATION PROCEDURE

Given an EPTMP-type rate function of the form (1), we must determine the degree m of the polynomial component of the exponent and least squares estimates of the parameters of Θ. To determine m, we will use a sequential model selection procedure. Based on the initial estimates of the parameters, we perform a likelihood ratio test to determine the appropriate degree m. Then we condition the estimation of the parameters on a fixed value of m and compute the final least squares estimate of the parameter vector Θ.

The procedure for obtaining the least squares estimate $\widetilde{\Theta}_m$ conditioned on a fixed value of m involves a numerical search procedure over the relevant parameter space. We have investigated several numerical search procedures including the Levenberg-Marquardt procedure, which is a specialized search gradient-search method for least squares problems (Kennedy and Gentle 1980), and the Nelder-Mead simplex search procedure (Barton 1996, Olsson 1974, Olsson and Nelson 1975), which is a general direct-search method for unconstrained optimization of continuous response functions that may be nondifferentiable. We chose the Nelder-Mead simplex search procedure to perform this numerical optimization because of its ability to handle weighted least squares formulations of our problem. Moreover in the case of least squares estimation of the mean-value function for an NHPP having an EPTMP-type rate function, we found that the performance and computational efficiency of the Nelder-Mead procedure is approximately equivalent to that of the Levenberg-Marquardt procedure.

The initial parameter estimates are based on methods by Kuhl, Wilson, and Johnson (1997) for rapidly approximating the maximum likelihood estimates of the parameters of an EPTMP-type rate function. To determine the degree m of the polynomial, the user can specify the minimum and maximum degree of the polynomial to be fitted. For each degree m of the polynomial, let $\widehat{\Theta}_m$ denote the initial estimate of the parameter vector Θ based on the procedure of Kuhl, Wilson, and Johnson (1997). We use the likelihood ratio test of Lee, Wilson and Crawford (1991) to determine the final estimate of m. Suppose a sequence of n events is 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). For each trial degree m, we let $\mathcal{L}_m\left(\widehat{\Theta}_m \middle| n, \mathbf{t}\right)$ denote the corresponding log-likelihood function evaluated at $\widehat{\Theta}_m$, given $N(S) = n$ and $\mathbf{t} = (t_1, t_2, \ldots, t_n)$. Under the null hypothesis that the current value of m is the true degree of the trend component of the underlying EPTMP-type rate function, the test statistic

$$2\left[\mathcal{L}_{m+1}\left(\widehat{\Theta}_{m+1} \middle| n, \mathbf{t}\right) - \mathcal{L}_m\left(\widehat{\Theta}_m \middle| n, \mathbf{t}\right)\right] \quad (9)$$

has approximately the chi-squared distribution with one degree of freedom provided S and n are sufficiently large. Thus we exploit (9) 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 (9) is not significant at a prespecified level of significance. The corresponding vector $\widehat{\Theta}_m$ provides the initial parameter estimates for the Nelder-Mead simplex search procedure to compute the final least squares estimate $\widetilde{\Theta}_m$ of the parameter vector Θ.

4 EXPERIMENTAL PERFORMANCE EVALUATION

4.1 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 seven NHPPs which represent processes having up to four cyclic components or a general trend over time or both. These cases were chosen based on the set of experimental cases used by Kuhl, Wilson, and Johnson (1997) to evaluate a maximum likelihood estimation procedure for NHPPs with EPTMP-type rate functions. Case 1 is a EPTMP-type rate function with one periodic component. Cases 2 through 4 consist of exponential rate functions with two periodic components. Cases 1 and 2 do not contain a general trend over time. Cases 3, 4, and 5 contain general trends which are represented by polynomials of degree 1, 2, and 3, respectively. Rate functions of type EPTMP with three

Table 1: Parameters of NHPPs Used in the Experimental Evaluation

	Case						
Parameter	1	2	3	4	5	6	7
α_0	3.6269	3.6269	3.6269	3.6269	4.5197	3.6269	3.6269
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	1.0592
ϕ_1	−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
γ_2	—	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
ϕ_2	—	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
ω_2	—	12.5664	12.5664	12.5664	12.5664	12.5664	12.5664
γ_3	—	—	—	—	—	0.2500	0.2500
ϕ_3	—	—	—	—	—	0.2500	0.2500
ω_3	—	—	—	—	—	25.1327	25.1327
γ_4	—	—	—	—	—	—	0.7500
ϕ_4	—	—	—	—	—	—	0.7000
ω_4	—	—	—	—	—	—	3.1416

and four periodic components and no long-term trend are utilized in Cases 6 and 7, respectively.

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 unit of time is taken to be 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` (Kuhl, Wilson, and Johnson 1997). For each case, $K = 100$ independent replications were simulated over the interval $[0, 12]$; and the resulting event-count samples were used first to verify the correct operation of the piecewise inversion scheme implemented in `mp3sim`. Then on each replication of each case, an EPTMP-type rate function was fitted to the observed series of event times. For all of the applications of the estimation procedure, the frequencies of the periodic effects are considered to be known. The user-specified significance level for the approximate likelihood ratio test (9) to determine the appropriate degree of the polynomial was set equal to 0.05. In each case with the exception of Case 7, the initial parameter estimates specified in Section 3 were used in the approximate likelihood ratio test (9). Since the quality of the initial parameter estimates degrades as the number of periodic components increases, the approximate likelihood ratio test based on these initial parameter estimates also begins to fail as the number of periodic components increases. In running these experiments, we found that when the number of periodic components was greater than or equal to four, the performance of the test (9) was unacceptable. Therefore, in Case 7 the true maximum likelihood estimates were used in the test statistic (9). The minimum degree of the fitted polynomial was set to zero and the maximum degree of the fitted polynomial was set to six on every application of the OLS estimation procedure.

4.2 Formulation of Performance Measures

To evaluate the performance of the OLS estimation procedure, we used both visual-subjective and numerical goodness-of-fit criteria. These numerical performance measures were utilized by Kuhl, Wilson, and Johnson (1997) to evaluate the maximum likelihood estimation procedure for fitting an EPTMP-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 ($k = 1, \ldots, K$), the estimated rate function is denoted by $\widetilde{\lambda}_k(t)$ and the estimated mean-value function is denoted by $\widetilde{\mu}_k(t)$.

As defined in Kuhl, Wilson, and Johnson (1997), we let δ_k and δ_k^* respectively denote the average absolute

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

	Case						
	1	2	3	4	5	6	7
$\mu(S)$	586	588	1126	967	968	599	714
$\overline{\delta}$	10.0	11.4	16.0	12.2	14.4	14.4	12.6
V_δ	0.65	0.47	0.36	0.27	0.62	0.40	0.36
Q_δ	0.21	0.23	0.17	0.15	0.18	0.29	0.21
$\overline{\delta^*}$	23.7	29.4	65.5	80.3	56.3	40.7	52.5
V_{δ^*}	0.60	0.47	0.42	0.32	0.57	0.48	0.37
Q_{δ^*}	0.48	0.60	0.70	1.00	0.70	0.82	0.88
$\overline{\Delta}$	12.4	12.8	15.6	12.9	16.1	11.0	11.6
V_Δ	0.84	0.73	0.70	0.59	0.63	0.82	0.59
Q_Δ	0.043	0.043	0.033	0.038	0.036	0.037	0.031
$\overline{\Delta^*}$	25.1	25.4	37.4	33.3	34.7	23.4	24.3
V_{Δ^*}	0.74	0.64	0.65	0.52	0.54	0.73	0.53
Q_{Δ^*}	0.087	0.086	0.081	0.097	0.077	0.078	0.065

error and maximum absolute error that occur in estimating the rate function $\lambda(t)$ on the kth replication of the target NHPP over the time interval $[0, S]$. Similarly, we let Δ_k and Δ_k^* respectively denote the average absolute error and maximum absolute error that occur in estimating the mean-value function $\mu(t)$ on the kth replication of the target NHPP over the time interval $[0, S]$. The sample mean of the observations $\{\delta_k : k = 1, \ldots, K\}$ is denoted by $\overline{\delta}$; and V_δ denotes corresponding sample coefficient of variation. The statistics $\overline{\delta^*}$ and V_{δ^*} are computed similarly from the observations $\{\delta_k^* : k = 1, \ldots, K\}$. The sample statistics $\overline{\Delta}$, V_Δ, $\overline{\Delta^*}$, and V_{Δ^*} are defined in the same fashion. As in Kuhl, Wilson, and Johnson (1997), we also report the "normalized" statistics Q_δ, Q_{δ^*}, Q_Δ, and Q_{Δ^*} to facilitate comparison of results for different rate functions.

In addition to performance measures that indicate the ability of the least squares procedure to fit an EPTMP-type rate and mean-value function to the rate and mean-value function of the underlying NHPP, we have formulated performance measures that indicate the ability of the least squares procedure to fit the observed arrival process. Space limitations preclude elaboration of these performance measures in this paper. For a detailed discussion of these statistics and their application in the present Monte Carlo study, see Kuhl, Damerdji, and Wilson (1998).

Beyond the numerical performance measures of goodness of fit to the underlying arrival process or to a realization of that process, 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. Kuhl, Wilson, and Johnson (1997) provide a detailed description of the method used to construct these tolerance bands.

4.3 Presentation of Results

The statistical results on the estimation of $\lambda(t)$ and $\mu(t)$ for each experimental case are shown in Table 2. These statistics describe the errors in estimating the underlying theoretical rate and mean-value functions. Table 3 shows the frequency distribution of the fitted degree of the polynomial trend taken over 100 replications for each case. Figures 4.3 through 4.3 contain the graphs of 90% tolerance bands for the rate function and mean-value function for cases 1, 5, and 7.

4.4 Analysis of Results

The statistical results in Table 2 seem to be reasonable for the selected measures of performance. Since these experimental cases are based on those of Kuhl, Wilson, and Johnson (1997), we will use their statistical results as a benchmark for evaluating the performance of our least squares estimation procedure.

In general, the performance measures in Table 2 that describe the estimation errors in fitting the underlying rate function (those involving δ) are higher (worse) for the least squares estimation procedure than the corresponding results reported for maximum likelihood estimation. However, the performance measures that describe the errors in fitting the underlying mean-value function (those involving Δ) are approximately the same for the two estimation methods. The larger rate-function estimation errors that were obtained

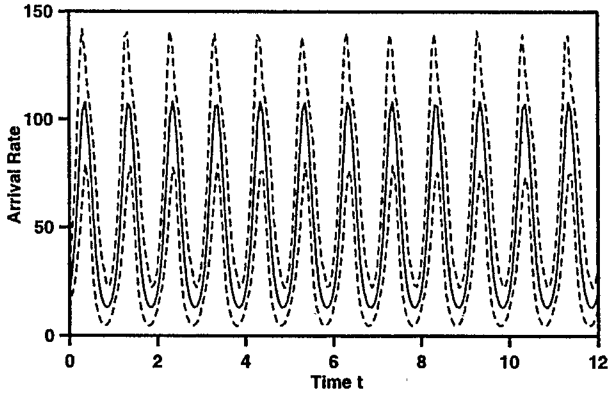

Figure 3: 90% Tolerance Intervals for $\lambda(t)$, $t \in [0, 12]$,

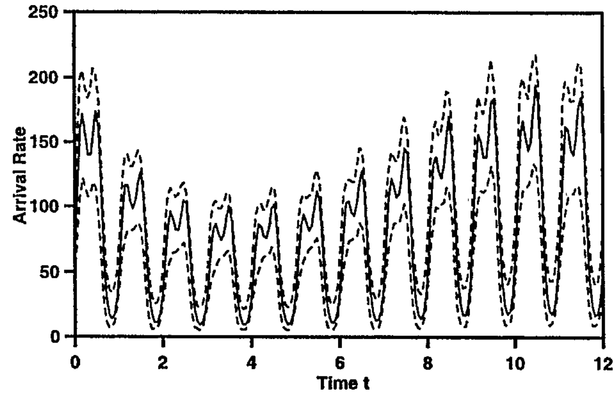

Figure 4: 90% Tolerance Intervals for $\lambda(t)$, $t \in [0, 12]$,

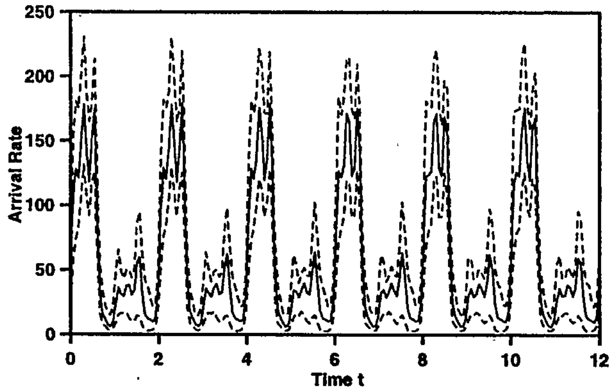

Figure 5: 90% Tolerance Intervals for $\lambda(t)$, $t \in [0, 12]$,

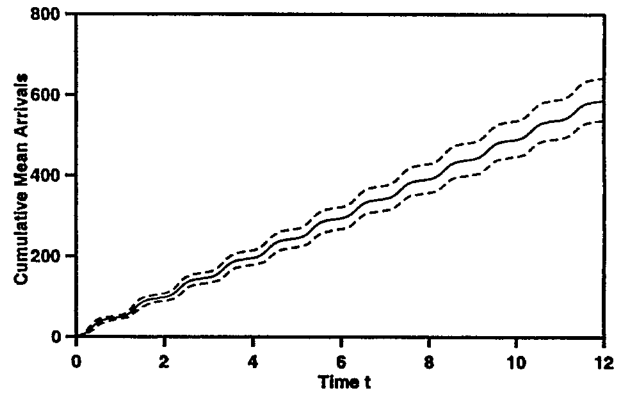

Figure 6: 90% Tolerance Intervals for $\mu(t)$, $t \in [0, 12]$,

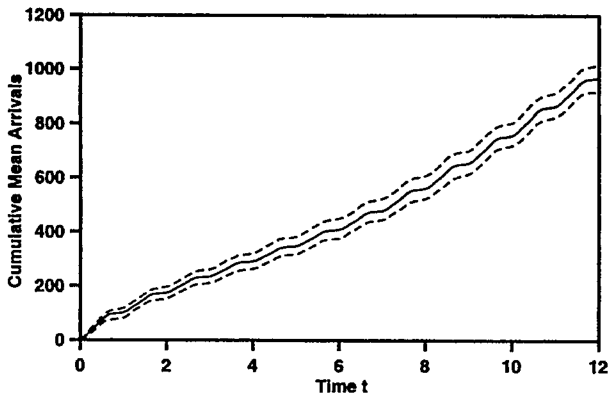

Figure 7: 90% Tolerance Intervals for $\mu(t)$, $t \in [0, 12]$, in Case 5

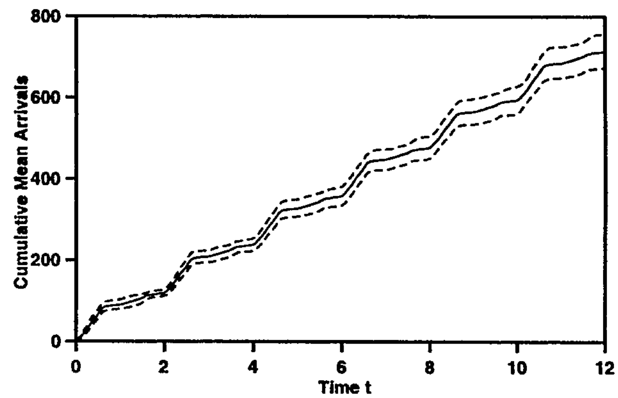

Figure 8: 90% Tolerance Intervals for $\mu(t)$, $t \in [0, 12]$, in Case 7

with the least squares procedure may be due to the fact that the objective function for the least squares procedure is based on the discrepancies between the fitted mean-value function and the empirical mean-value function. Thus a good fit to the mean-value function does not necessarily guarantee that the fit to its derivative, the rate function, will be good. The difference between the quality of the fits for the two methods is also evident in the plots of the rate and mean-value functions. Also, we observe that the errors in estimating the underlying rate and mean-value functions tend to increase as the degree m of the long-term trend and the number p of periodic components increase. One reason for this may be that as the number of periodic components increases, the initial estimates of the parameters begin to degrade, which may cause the numerical search procedure to start too far from the optimum. Poor starting values for the parameter estimates may result in the procedure finding a local minimum least squares estimate and stopping at a suboptimal solution.

Table 3 indicates the ability of the fitting procedure to determine the degree of the exponential-polynomial trend present in the underlying NHPP rate function. These results indicate that the likelihood ratio test based on the initial estimates for the maximum likelihood estimation procedure works well in general for rate functions having up to three periodic components. With more than three periodic components, we were able to achieve similarly good results but at the cost of computing the final maximum likelihood estimates to be used in the likelihood ratio test.

Table 3: Frequency of Fitted Polynomial Degree for $K = 100$ Realizations

	True	Fitted Degree						
Case	Degree	0	1	2	3	4	5	6
1	0	93	7	0	0	0	0	0
2	0	87	13	0	0	0	0	0
3	1	0	94	6	0	0	0	0
4	2	0	0	87	13	0	0	0
5	3	3	0	1	95	1	0	0
6	0	100	0	0	0	0	0	0
7	0	94	6	0	0	0	0	0

The plots of the 90% tolerance bands about the rate functions indicate that the least squares estimation procedure is consistently able to fit a reasonable EPTMP-type rate function to the underlying NHPP. Similar to the results reported by Kuhl, Wilson, and Johnson (1997) for maximum likelihood estimation, the plots of the tolerance bands for least squares estimation are widest at the peaks and valleys of the arrival rate. In addition, the tolerance bands tend to be wider as the number of periodic components increases.

The plots of the 90% tolerance bands about the mean-value function also indicate that the least squares procedure consistently provides reasonable estimates of the underlying NHPP. Also from the plots of the tolerance bands, one can observe that the widths of the tolerance bands increase over time. This behavior is expected. Because the error is cumulative over time, the estimation error increases as the mean-value function increases.

5 CONCLUSION

In this paper we have developed a least squares method for estimating the parameters of an NHPP having an EPTMP-type rate function. This procedure has been implemented in the public domain computer software `mp3ls`. Using this software, we have performed an experimental evaluation of our least squares procedure. The results of this study indicate that the least squares estimation method does a good job of doing what it was designed to do. Namely, the procedure is capable of accurately tracking the empirical mean-value function of an NHPP. In addition, we have developed a weighted least squares formulation of this problem, and have shown theoretically why weighted least squares fails when applied to an estimation problem with a first- and second-order moment structure such as that arising in estimation of the mean-value function for an NHPP.

REFERENCES

Abramowitz, M. and I.A. Stegun. 1972. *Handbook of Mathematical Functions,* Dover Publications Inc., New York.

Barton, R.R. and J.S. Ivey, Jr. 1996. Nelder-Mead Simplex Modifications for Simulation Optimization, *Management Science,* 42:7, 954–973.

Box, G.E.P., W.G. Hunter, and J.S. Hunter. 1978. *Statistics for Experimenters,* John Wiley and Sons, Inc., New York.

Bratley, P., B.L. Fox, and L.E. Schrage. 1987. *A Guide to Simulation, Second Edition,* Springer-Verlag, New York.

Çinlar, E. 1975. *Introduction to Stochastic Processes,* Prentice-Hall Inc., Englewood Cliffs, NJ.

Kennedy, W.J., Jr. and J.E. Gentle. 1980. *Statistical Computing,* Marcel Dekker, Inc., New York.

Kuhl, M. E., J. R. Wilson, and M. A. Johnson. 1997. Estimating and Simulating Poisson Processes Having Trends or Multiple Periodicities. *IIE Transactions* 29:3 201–211.

Kuhl, M. E., H. Damerdji, and J. R. Wilson. 1998. Least-squares estimation of nonhomogeneous Poisson processes. Technical Report, Department of Industrial

Engineering, North Carolina State University, Raleigh, North Carolina.

Lee, S., J.R. Wilson, and M.M. Crawford. 1991. Modeling and Simulating 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, 353–375.

Olsson, D.M. 1974. A Sequential Simplex Program for Solving Minimization Problems, *Journal of Quality Technology,* 6:1, 53–57.

Olsson, D.M., and L.S. Nelson. 1975. The Nelder-Mead Simplex Procedure for Function Minimization, *Technometrics,* 17:1, 45–51.

Seber, G.A., and C.J. Wild. 1989. *Nonlinear Regression,* John Wiley and Sons, Inc., NY.

Swain, J.J., S. Venkatraman, and J.R. Wilson. 1988. Least Squares Estimation of Distribution Functions in Johnson Translation Systems, *Journal of Statistical Computer Simulation,* 29, 271–297.

Wagner, M. A. F., and J. R. Wilson. 1996a. Using univariate Bézier distributions to model simulation input processes. *IIE Transactions* 28 (9): 699–712.

Wagner, M. A. F., and J. R. Wilson. 1996b. Recent Developments in Input Modeling with Bézier Distributions. *Proceedings of the 1996 Winter Simulation Conference,* J.M. Charnes, D.J. Morrice, D.T. Brunner, and J.J. Swain (eds.), Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1448–1456.

AUTHOR BIOGRAPHIES

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

HALIM DAMERDJI is an assistant professor in the Department of Industrial Engineering at North Carolina State University. He received a Ph.D. degree in industrial engineering from the University of Wisconsin-Madison. He has held positions at the Ecole Nationale Polytechnique of Algiers and Purdue University. His research interests are in simulation and applied stochastic processes. He is a member of IIE and INFORMS.

JAMES R. WILSON is a professor 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. Currently he serves as a corepresentative of the INFORMS College on Simulation to the WSC Board of Directors. He is a member of ASA, ACM, IIE, and INFORMS.