Proceedings of the 2000 Winter Simulation Conference
J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, eds.

NONPARAMETRIC ESTIMATION OF NONHOMOGENEOUS POISSON PROCESSES USING WAVELETS

Michael E. Kuhl
Prashant S. Bhairgond

Department of Industrial and Manufacturing Systems Engineering
3128 CEBA Building
Louisiana State University
Baton Rouge, LA 70803-6409, U.S.A.

ABSTRACT

Nonhomogeneous Poisson processes (NHPPs) are frequently used in stochastic simulations to model nonstationary point processes. These NHPP models are often constructed by estimating the rate function from one or more observed realizations of the process. Both parametric and nonparametric models have been developed for the NHPP rate function. The current parametric models require prior knowledge of the behavior of the NHPP under study for model selection. The current nonparametric estimators, in general, require the storage of all of the observed data. Other hybrid approaches have also been developed. This paper focuses on the nonparametric estimation of the rate function of a nonhomogeneous Poisson process using wavelets. The advantages of wavelets include the flexibility of a nonparametric estimator enabling one to model the nonstationary rate function of an NHPP without prior knowledge or assumptions about the behavior of the process. Furthermore, this method has some advantages of current nonparametric techniques. Thus, using wavelets we can develop an efficient yet highly flexible NHPP rate function. In this paper, we develop the methodology required for constructing a wavelet estimator for the NHPP rate function. In addition, we present an experimental performance evaluation for this method.

1 INTRODUCTION

In this paper, we focus on arrival processes that can be classified, as nonstationary point processes. For these processes we are able to observe the arrival time of each entity where the rate at which arrivals occur changes over time. Under certain assumptions, these nonstationary arrival processes can be represented by a nonhomogeneous Poisson process (see Çinlar 1975).

The main objective of the research is to develop a nonparametric model for the estimation of the rate function of a nonhomogeneous Poisson process (NHPP) using wavelets. There are several advantages to using wavelets as a modeling tool. Wavelets are very flexible in terms of their ability to model complex and irregular behavior that can be found in arrival processes. In addition, wavelets can be used as an estimator without prior knowledge or assumptions about the behavior of the process.

In the following sections, we present the methodology for constructing a wavelet estimator for the NHPP rate function. We then present a method for generating arrivals from an NHPP having a wavelet rate function that utilizes the standard method of thinning. In addition, we present an experimental performance evaluation for the wavelet procedure to evaluate the goodness of fit for the wavelet model. Finally, we present our conclusions and recommendations for future work.

1.1 Nonhomogeneous Poisson Processes

A nonhomogeneous Poisson process $\{N(t): t > 0\}$ is a generalization of the Poisson process in which the instantaneous arrival rate $\lambda(t)$ at time t is a non-negative integrable function of time. The mean-value function (or the integrated rate function) of the NHPP is given as,

$$\mu(t) \equiv E[N(t)] = \int_0^t \lambda(z)dz, \qquad \forall\ t \geq 0 \qquad (1)$$

where $N(t)$ is the number of arrivals in $(0, t]$, for all $t \geq 0$ (Çinlar 1975).

Nonhomogeneous Poisson processes have been applied successfully to model nonstationary point processes for a large class of problems. The distribution of the NHPP is completely defined by the rate (intensity) function. Therefore the objective of this research is to estimate the rate function, $\lambda(t)$. The rate function of these NHPPs may exhibit many types of systematic or irregular behavior. These may include cyclic behavior as well as long-term evolutionary trends. The cyclic behavior may involve multiple periodicities and may be asymmetric in nature. Therefore, a flexible method is needed for modeling the rate function. Both parametric and nonparametric models have been developed for the NHPP rate function. The current parametric models require prior knowledge of the behavior of the NHPP under study for model selection. The current nonparametric estimators, in general, require the storage of all of the observed data. Other hybrid approaches have also been developed. We propose a nonparametric estimate of the rate function of an NHPP using wavelets.

1.2 Wavelets

Wavelet theory is the mathematics associated with building a model for a signal, system, or a process with a set of "special signals" called wavelets. Wavelets are useful in a broad range of applications, such as data compression, signal and image processing, nonparametric statistical estimation, numerical analysis, chemistry, astronomy, oceanography, turbulence, human vision, radar, and earthquake prediction (Bruce, *et al.* 1997).

Wavelets are the functions that satisfy certain mathematical requirements and are used in representing data or other functions. These functions cut the data into different frequencies and then study each component matched to its scale. In wavelet analysis, linear combinations of wavelet functions are used to represent signals or data.

For a function to qualify to be a wavelet, it must be oscillatory and have amplitudes that quickly decay to zero in both the positive and negative directions. The required oscillatory condition leads to using sinusoidal type functions as the wavelet basis functions. These two conditions must be simultaneously satisfied for the function to be a wavelet (Young 1996).

Wavelets consist of two types of functions, scaling function and mother wavelet, that work together to provide wavelet approximations. The scaling function is denoted $\phi(t)$ and has the property that

$$\int \phi(t)dt = 1.$$

The second type of the function used in wavelet analysis is called the mother wavelet or analyzing wavelet and is denoted $\psi(t)$. The mother wavelet has the property that

$$\int \psi(t)dt = 0.$$

The scaling function is comparatively better than the mother wavelet at modeling low frequency and smooth parts of the signal or data. The mother wavelet can be effectively used to approximate the detailed and high frequency parts of the signal or data.

The mother wavelet provides the underlying functional form for the approximation at various levels (resolutions) such that

$$\textit{Mother wavelet} = a^{-1/2}\ \{\textit{Scaled Wavelet}\}$$

or

$$\psi_{ab}(x) = a^{-1/2}\psi\left(\frac{x-b}{a}\right). \tag{2}$$

In equation (2), the factor $a^{-1/2}$ is called an energy normalizing term and is determined such that the mother wavelet and the scaled wavelet have the same amount of energy (Young 1996). Furthermore, wavelets are sets of functions formed by dilations of the mother wavelet, which are controlled by the positive real number $a \in R^+$, and translations of the mother wavelet, which are controlled by the real number b. Visually, the mother wavelet appears to model the local oscillation, or wave, in which most of the energy of the oscillation is located in a narrow region in the physical space. The dilation parameter a controls the width and rate of this local oscillation and can be considered to be controlling the frequency of $\psi(x)$. The translation parameter b moves the wavelet throughout the domain (Erlebacher, *et al.* 1996).

The parameter scale in the wavelet analysis is similar to the scale used in the maps. As in the case of maps, larger scales correspond to a non-detailed global view, and small scales correspond to a detailed view. Similarly, in terms of frequency, low frequency (large scales) corresponds to global information of a signal, and the high frequencies (small scales) correspond to detailed information of short-lived pattern in the signal. The translation and dilation operations applied to the mother wavelet are performed to calculate the wavelet coefficients, which represent the correlation between the wavelet and a localized section of the signal. The wavelet coefficients are calculated for each wavelet segment, giving a time-scale function relating the wavelets correlation to the signal. This process of translation and dilation of the mother wavelet at different scales is referred to as multiresolution analysis where a higher resolution corresponds to a higher level of detail.

2 THE WAVELET ESTIMATION PROCEDURE

In this section, we develop a wavelet estimator to approximate the rate function of an NHPP that represents an arrival process. Since the arrival rate of NHPPs cannot be negative, a positive wavelet estimator has been developed. The devolopment of this methodology is based on Walter and Shen (1998) where the theory for developing the positive wavelet estimator for the approximation of a density function is presented. The properties of the arrival rate of an NHPP are similar in several aspects to the properties of a density function including non-negativity. Therefore, we develop a positive wavelet estimator for approximating the arrival rate of NHPP in a similar manner.

2.1 Selecting Wavelet Basis Functions

The first step in the methodology is the selection of the wavelet system consisting of the mother wavelet and scaling function. From the selected wavelet system, the scaling function will be used for approximation of the rate function of the NHPP.

The mother wavelet must be selected such that it constitutes an orthonormal function. That is, the inner product (dot product) of the mother wavelet must be equal to one. The mother wavelet or the orthonormal basis has the form

$$\psi_{mn}(t) = 2^{-m/2}\,\psi\,(2^{-m}t - n)$$

where m is the level of resolution and n is the shift or translation. The mother wavelet is constructed in conjunction with the scaling function.

The scaling function $\phi(t)$ is a real valued function that is m times differentiable and whose derivatives are continuous and rapidly decreasing. The scaling function to be used for approximation is chosen such that

$$\int \phi(t)\phi(t-n)\,dt = \delta_{0,n}$$

and $\{\phi(t-n)\}$ is an orthonormal basis function where $\delta_{0,n}$ is a delta sequence.

There are different families of wavelets that can be used for approximation. Haar wavelets are simple step wavelets that are non-continuous in nature and exhibit jump discontinuities (Nievergelt 1999). They cannot be effectively used for approximation of the continuous rate function of NHPPs. Another family of wavelets, Daubechies wavelets are continuous in nature and because of this property of continuity, Daubechies wavelets can be used to effectively approximate the rate function. Therefore, we will consider Daubechies wavelets for the approximation of the rate function of NHPPs. Once a wavelet family has been established, the next step in the methodology is to calculate the wavelet coefficients for the basis function at resolution 0 which has support on V_0. The next subsection illustrates these calculations for a relatively simple Daubechies scaling function.

2.2 Calculation of Daubechies Wavelets

To illustrate the methodology of defining the scaling function, we will consider the case of the Daubechies scaling function with support from 0 to 3. A graph of this scaling function is shown in Figure 1.

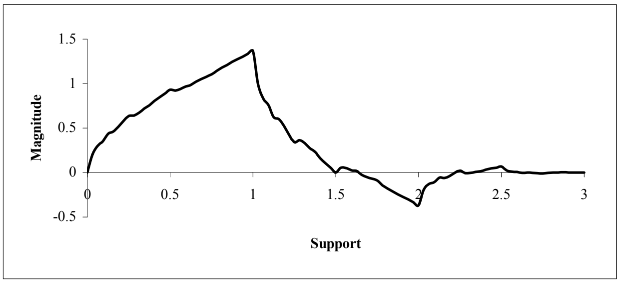

Figure 1: Daubechies Scaling Function with Support [0,3]

The starting values for the Daubechies scaling function having the window width from 0 to 3 are defined as (Nievergelt 1999),

$$\begin{aligned}
&\phi(0) = 0,\\
&\phi(1) = \frac{1+\sqrt{3}}{2},\\
&\phi(2) = \frac{1-\sqrt{3}}{2},\\
&\phi(3) = 0,
\end{aligned}$$

where ϕ satisfies the recurrence relation

$$\begin{aligned}
\phi(r) = {} & \frac{1+\sqrt{3}}{4}\phi(2r) + \frac{3+\sqrt{3}}{4}\phi(2r-1)\\
& + \frac{3-\sqrt{3}}{4}\phi(2r-2) + \frac{1-\sqrt{3}}{4}\phi(2r-3)\ .
\end{aligned}$$

If constants in the recurrence equation are replaced by the following constants

$$\begin{aligned}
h_0 &= \frac{1+\sqrt{3}}{4},\\
h_1 &= \frac{3+\sqrt{3}}{4},\\
h_2 &= \frac{3-\sqrt{3}}{4},\\
h_3 &= \frac{1-\sqrt{3}}{4},
\end{aligned} \qquad (3)$$

then the recurrence relation now becomes

$$\phi(r) = h_0\,\phi(2r) + h_1\,\phi(2r-1) + h_2\,\phi(2r-2) + h_3\,\phi(2r-3)\,. \tag{4}$$

The values of the Daubechies wavelet at other points in [0,3] can be calculated using the initial values in (3) and the recursion (4). Sufficient values of the wavelet are calculated so that a smooth scaling function is obtained.

Using this procedure, we can calculate Daubechies scaling functions that have other supports. The recurrence equation in general terms can be written as

$$\phi(r) = \sum_{i=0}^{n} h_i\,\phi(2r-i)\,. \tag{5}$$

Thus for Daubechies wavelet with support of 0 to 7, the recurrence equation will have 7 terms in the equation (5) and so on for different supports (Nievergelt 1999)

$$\phi(r) = h_0\phi(2r) + h_1\phi(2r-1) + h_2\phi(2r-2) + \ldots + h_6\phi(2r-6)\,.$$

Daubechies (1992) defines the coefficient values for wavelets with different supports. The number r is called a dyadic number if and only if, it is an integral multiple of an integral power of 2 (Nievergelt 1999). Figure 2 shows the Daubechies scaling functions having support 0 to 7.

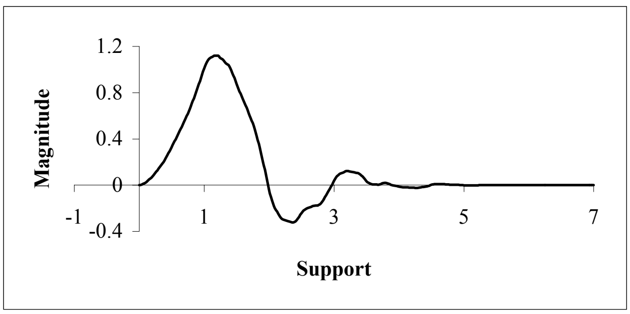

Figure 2: Daubechies Scaling Function with Support [0, 7]

2.3 Estimation Using Positive Basis Functions

Wavelets are in general real valued functions. We can observe from Figures 1 and 2 that the Daubechies wavelets can take on negative values. As stated previously, the rate function of the NHPP will require a non-negative estimator. Walter and Shen (1998) developed a positive wavelet estimator for estimating density functions. The wavelet estimator for the rate function of an NHPP is based on similar approach.

Let $\phi(t)$ be any continuous function having compact support that generates the space V_0. The function $\phi(t)$ generates a partition if

$$\sum_{n\in Z} \phi(t-n) \equiv 1, \qquad t \in R\,. \tag{6}$$

This property is shared by all scaling functions of compact support especially the Daubechies scaling functions and the "Coiflets" (Walter and Shen 1998).

For $0<r<1$, the sumability function or the positive basis function is given by,

$$P_r(t) = \sum_{n\in Z} r^{|n|}\,\phi(t-n)$$

where the constant value r is selected such that this positive basis developed is always greater than or equal to zero. Walter and Shen (1998) have proven the following for the positive basis function.

2.3.1 Lemma 1

Let $\phi(t)$ be continuous and compactly supported. Further assume that $\phi(t)$ satisfies equation (6) and, for positive constants *A,* and *B,* the frame condition

$$0 < A \le \sum_{k\in Z} \left|\hat{\phi}(\omega + 2k\pi)\right|^2 \le B < \infty\ .$$

If V_0 is the closed linear span of $\{\phi(t-n)\}_{n\in Z}$, then there exist $0<r_0<1$ such that $P_r(t)$ satisfies

(i) $P_r(t) \ge 0,\ \forall t \in R$
(ii) $P_r \in V_0$

for every $r_0 \le r < 1$.

The range of $n \in Z$ depends on the scaling function used to find the positive basis function P_r. In this case, we considered the value of n varying from –7 to 7. Figure 3 shows the positive basis function $P_r(t)$ associated with the Daubechies scaling function having support of 0 to 7 for r = 0.3. The truncation for the positive basis is taken from –6 to 10. The value of the positive basis function outside the limit will be zero. The nature of the basis function will depend upon the nature of the scaling function considered for approximation.

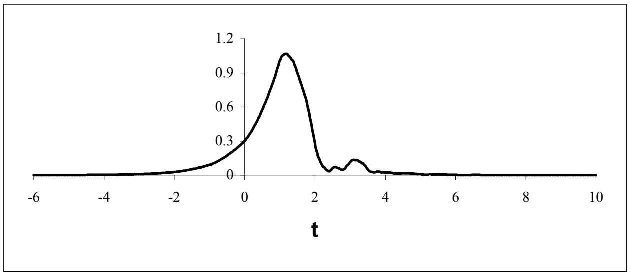

Figure 3: Positive Basis Function Associated with Daubechies Scaling Function

Using $P_r(t)$, a positive reproducing kernel, $k_r(t,t_i)$ in V_0 is constructed from follows,

$$k_{r,0}(t,t_i)=\left(\frac{1-r}{1+r}\right)^2 \sum_{n=-\infty}^{\infty} P_r(2^0 t-n)\, P_r(2^0 t_i-n)$$

which in general form can be written as

$$k_r(t,t_i)=\left(\frac{1-r}{1+r}\right)^2 \sum_{n=-\infty}^{\infty} P_r(t-n)\, P_r(t_i-n)$$

such that,

$$k_r(t,t_i)\ge 0\ ,$$

and

$$\int_R k_r(t,t_i)dt_i=1\ ,$$

and for $f\in L_2(R)$, the approximation in V_0 is given by

$$\lambda_0(t)=\int_{-\infty}^{\infty} k_r(t,t_i)\, f(t_i)dt_i\, .$$

This kernel satisfies the conditions needed to generates a positive delta sequence $\{k_{r,m}\}$ where

$$k_{r,m}(t,t_i)=2^m\, k_r(2^m t, 2^m t_i), \qquad m\in Z\ .$$

The wavelet estimator in V_m can be written in the following form

$$\hat{\lambda}_{m,k}(t)=\sum_{n=-k}^{k}\left\{\sum_{i=1}^{N} P_r(2^m t_i-n)\right\}\left(\frac{1-r}{1+r}\right)^2 2^m P_r(2^m t-n)$$

where t_i are the arrival times of an NHPP whose rate function is to be approximated, N is the number of arrivals in the interval under consideration. The range for n is selected in such a way that the positive basis function $P_r(t)$, can translate through the entire range of arrival times and m, the resolution is selected based on the level of detail of the approximation desired. This wavelet estimator is used to approximate the rate function of an NHPP.

We have written computer code in C which is used to approximate the rate function of an NHPP. Observed data is used for the approximation and is the input to the program. The output is the approximation of this arrival rate.

3 GENERATION OF NHPPS HAVING WAVELET RATE FUNCTIONS USING THINNING

Lewis and Shedlar (1979) proposed a general and simple method known as thinning. This method is commonly used to generate the arrival times of a nonhomogeneous Poisson process. The method of thinning is explained here briefly.

Let $\lambda(t)$ be the approximated arrival rate of the NHPP using the wavelet estimator, for which the arrival times $\{t_i\}$ have to be generated. A stationary Poisson process with constant and finite arrival rate $\lambda^*=\max\{\lambda(t)\}$ is generated with arrival times $\{t_i^*\}$. Figure 4 depicts the two rate functions used in the thinning algorithm which follows.

Step (1): Set $t=t_i^*$

Step (2): Generate U_1 and U_2 as independent and identically distributed U [0,1] of any previous random variates where U_1 is given as

$$U_2=\frac{-1}{\lambda^*}\ln(U_1)$$

Step (3): The next arrival is found out as

$$t_{i+1}^*=t_i^*+U_2$$

Step (4): Replace t by t_{i+1}^*

Step (5): If $U_2\le\frac{\lambda(t)}{\lambda^*}$, we accept t_i^*, otherwise this arrival is thinned out.

Using this algorithm we can easily generate the arrival times of an NHPP of the approximated rate function using the wavelet estimator.

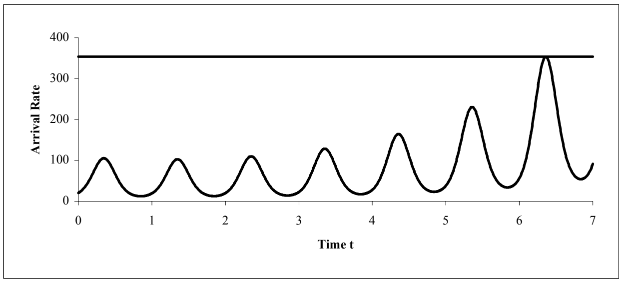

Figure 4: Constant Arrival Rate $\lambda^*(t)$ (thin line) and the Fitted Wavelet Estimator of Arrival Rate $\lambda(t)$ (thick line)

4 EXPERIMENTAL PERFORMANCE EVALUATION

To evaluate the wavelet estimation procedure for fitting the rate function of an NHPP, we have conducted the following experimental performance evaluations and evaluated the goodness-of-fit using both visual and numerical goodness-of-fit measures for the wavelet estimation procedure.

We have considered three different cases that represent the different types of arrival processes. The three cases were selected to evaluate the estimation procedure for fitting rate functions of nonhomogeneous Poisson processes that have periodic components and/or a general trend over time. The cases are chosen based on the set of experimental cases, which were used, by Kuhl, Wilson, and Johnson (1997) to evaluate a maximum likelihood estimation procedure for NHPPs with EPTMP-type rate functions. The underlying EPTMP-type rate functions from which the arrivals are generated have the form

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

with

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

Using this rate function we can generate realizations of an NHPP having long term trend and/or cyclic effects.

Table 1 lists the parameters of the EPTMP-type rate function for each case. Case 1 is an EPTMP-type rate function with one periodic component. Case 1 does not have a long-term trend over time. Case 2 has exponential rate function, which exhibits a long-term general trend and has one periodic component. The general trend in this case is represented by the polynomial of degree 2. Case 3 has an EPTMP-type rate functions with two periodic components.

Table 1: Parameters of NHPP Used in the Experimental Evaluation

Parameter	Case 1	Case 2	Case 3
α_0	3.6269	3.6269	3.6269
α_1	-	-0.1000	-
α_2	-	0.0450	-
γ_1	1.0592	1.0592	1.0592
ϕ_1	-0.6193	-0.6193	-0.6193
ω_1	6.2831	6.2831	6.2831
γ_2	-	-	0.7500
ϕ_2	-	-	0.7000
ω_2	-	-	3.1416

The frequencies used in the experimentation are expressed in radians per unit time such that $\omega_1 = 2\pi$ and $\omega_2 = \pi$ radians per unit time. These frequencies depend upon the unit time considered. If the unit time is considered to be one year, then these frequencies represent annual and biennial effects, respectively.

For these selected NHPPs, the arrival times were generated over the interval [0,7] using the program *mp3sim* (Kuhl, Wilson and Johnson 1997). For each case, K=20 independent replications were simulated. Using the wavelet estimator procedure, the rate function was estimated for each replication. Based on a preliminary empirical study, we have determined that at resolution 4 we achieve an appropriate balance between the level of smoothness and detail desired. Therefore we have carried out the approximation of the rate function and the calculation at resolution 4. Note, however, that in practice the resolution selected will depend on the nature of the data under consideration.

In practice, especially when the data contains a long term trend, we are usually only able to observe one realization of the arrival process. Therefore, this experimental study centers on fitting the wavelet rate function to one realization of data. However, to illustrate the ability of the wavelet rate function estimator to converge to the theoretical rate function, we also investigated the case of multiple realizations of the observed arrival process.

4.1 Formulation of Performance Measures

To evaluate the performance of the wavelet estimation procedure, we used both visual-subjective and numerical goodness-of-fit criteria. Kuhl, Wilson and Johnson (1997) utilized several performance measures to evaluate the maximum likelihood estimation procedure for fitting an EPTMP-type rate function. These performance measures, are restated here for completeness. These performance measures are measures of the procedures ability to fit the underlying theoretical rate and mean-value function and 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 $\hat{\lambda}_k(t)$ and the estimated mean-value function is denoted by $\hat{\mu}_k(t)$. The *average absolute error* in the estimation of the rate function $\lambda(t)$ on the kth replication is

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

and the *maximum absolute error* is

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

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

Aggregate performance measures are computed over all replications of a given experiment. The sample mean of the observations $\{\delta_k : k=1,\ldots,K\}$ is denoted by $\bar{\delta}$ and the corresponding sample coefficient of variation V_δ is given by

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

The statistics $\bar{\delta}^*$ and V_{δ^*} are computed similarly from the observations $\{\delta^*_k : k=1,\ldots,K\}$. We also calculated the normalized performance measures

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

since normalizing by the theoretical average arrival rate over [0,*S*] facilitates comparison of results for different rate functions. Aggregate performance measures analogous to these are also so calculated for the errors in estimating the the mean-value function and denoted by Δ_k and Δ_k^*.

In addition, we have calculated several other performance measures (Kuhl, Damerdji, and Wilson 1998) that are based on the observed arrival process. On the kth replication of the NHPP $(k=1,\ldots,K)$, let $\{t_{i,k} : i=1,\ldots,N_k(S)\}$ denote the arrival times observed in the time interval [0,*S*]. The average absolute error in fitting the mean-value function to the empirical mean-value function on the kth replication is calculated by

$$D_k = \frac{1}{N_k(S)}\sum_{i=1}^{N_k(S)}\left|\hat{\mu}_k(t_{i,k})-i\right|$$

for $(k=1,\ldots\ldots,K)$. Let $\bar{D}$ denote the sample mean of the observed values $\{D_k : k=1,2,\ldots.,K\}$. The maximum absolute error in fitting the mean-value function to the empirical mean-value function on the kth replication is calculated by

$$D_k^* = \max\left\{\left|\hat{\mu}_k(t_{i,k})-i\right| : 1 \le i \le N_k(S)\right\}.$$

The sample mean of the observed values $\{D^*_k : k=1,\ldots,K\}$ is denoted $\bar{D}^*$.

The following measures compute the grand average level of the empirical mean-value functions taken over all K replications to normalize the average performance measures $\bar{D}$ and $\bar{D}^*$,

$$Q_D = \frac{\bar{D}}{\frac{1}{K}\sum_{k=1}^{K}\frac{1}{S}\int_0^S N_k(t)dt} \quad \text{and} \quad Q_{D^*} = \frac{\bar{D}^*}{\frac{1}{K}\sum_{k=1}^{K}\frac{1}{S}\int_0^S N_k(t)dt}.$$

In addition to numerical performance measures, graphical methods are used to provide a visual mean 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. Tolerance intervals are also obtained for the mean-value function $\mu(t)$ at a fixed interval of time $t \in [0,S]$.

4.2 Discussion of Results

The statistics in Table 2 describe the errors in estimating the theoretical rate and mean-value functions. The statistics in Table 3 describe the error in fitting the empirical mean-value functions. Figures 5-10 contain the graphs of 90% tolerance bands for the rate function and the mean-value function for Case 1, Case 2 and Case 3.

The experimental cases are based on those of Kuhl, Wilson, and Johnson (1997), so we consider their statistical results a benchmark for evaluating the performance of our wavelet estimation method. We will also compare our results with least square estimation method (Kuhl, Damerdji, and Wilson 1998). Relative to these benchmarks, the statistical results in Table 2 and 3 seem to be good for the selected measures of performance.

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

Case	1	2	3
$\mu(S)$	343	548	443
$\overline{\delta}$	10.28	14.37	12.5
V_{δ}	0.14	0.15	0.15
Q_{δ}	0.21	0.18	0.20
$\overline{\delta}^*$	33.07	49.01	56.25
V_{δ^*}	0.29	0.33	0.24
Q_{δ^*}	0.67	0.63	0.88
$\overline{\Delta}$	8.49	7.12	10.81
V_{Δ}	0.43	0.40	0.61
Q_{Δ}	0.048	0.034	0.048
$\overline{\Delta}^*$	19.56	23.52	21.85
V_{Δ^*}	0.11	0.10	0.09
Q_{Δ^*}	0.49	0.54	0.49

Table 3: Statistics Describing the Errors in Estimating *N(t)*

Case	1	2	3
$\mu(S)$	343	548	443
$\overline{D}$	1.698	2.321	2.198
$\overline{D}^*$	5.803	8.967	7.632
Q_D	0.0091	0.0081	0.0085
Q_{D^*}	0.0341	0.0592	0.0511

From Table 2 we can see that, in general the estimation errors in fitting the underlying rate function are slightly higher for wavelet estimation method than the corresponding results reported for maximum likelihood estimation and approximately equal to the performance measures obtained using least square estimation.

In Table 2, we observe that $\overline{\delta}$ ranges from 10.28 to 14.37 while V_{δ} ranges from 0.14 to 0.15 over the 3 cases and the normalized absolute error Q_{δ} ranges from 0.18 to 0.21 over the 3 cases. The absolute values of these performance measures are within acceptable limits for most kinds of arrival processes encountered in practice and are not significantly greater than those values found in benchmark studies. In addition to relative consistency of performance measures across the 3 cases, when we compare case 1 which has one periodic component and no long term trend with cases 2 and 3, we observe that the addition of long term trend in case 2 or the addition of second periodic component in case 3 did not adversely affect the performance measures.

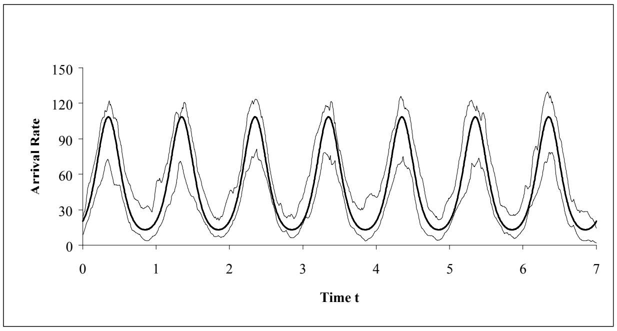

Figure 5: 90% Tolerance Interval for the Arrival Rate, t = (0,7], in Case 1

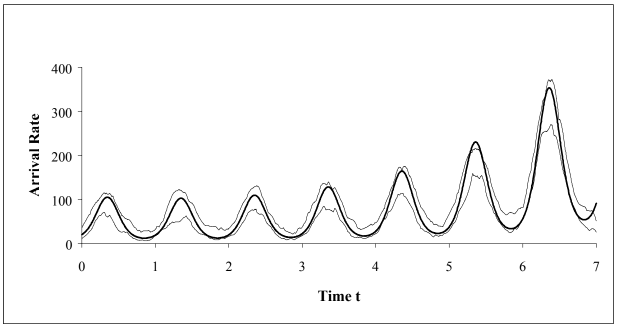

Figure 6: 90% Tolerance Intervals for the Arrival Rate, t = (0,7], in Case 2

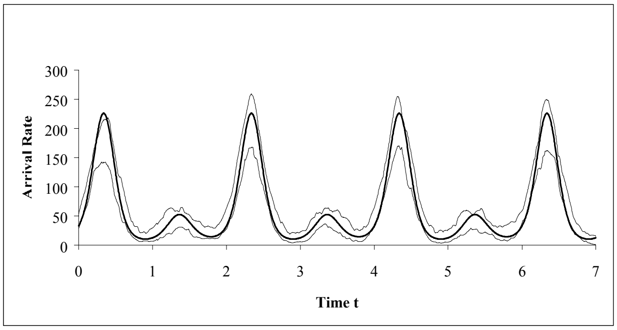

Figure 7: 90% Tolerance Interval for the Arrival Rate, t = (0,7], in Case 3

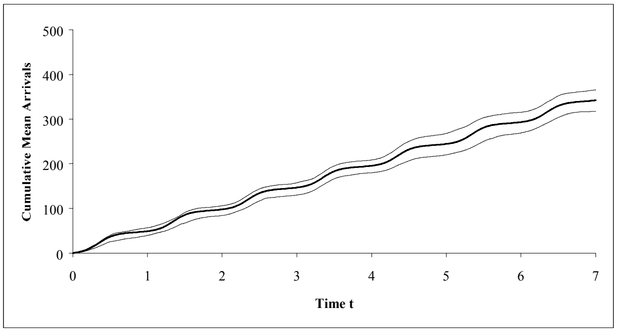

Figure 8: 90% Tolerance Intervals for the Mean-Value Function, t = (0,7], in Case 1

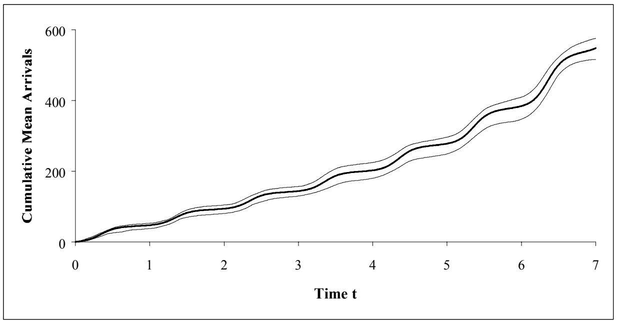

Figure 9: 90% Tolerance Intervals for the Mean-Value Function, t = (0,7], in Case 2

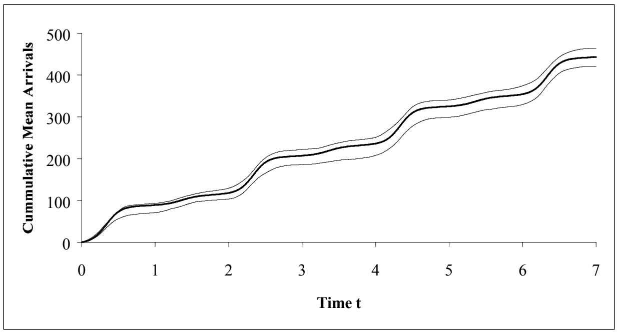

Figure 10: 90% Tolerance Intervals for the Mean-Value Function, t = (0,7], in Case 3

The performance measures in Table 3 describe the errors in fitting the mean value function to the empirical mean value. These performance measures show that the error in fitting the mean-value function to empirical mean-value function is lower than the performance measure reported by Kuhl, Damerdji, and Wilson (1998) that utilized the least square estimation procedures. Kuhl, Wilson, and Johnson (1997) did not report these performance measures for their maximum likelihood estimation procedure.

The plots of the 90% tolerance bands about the mean-value function also indicate that the wavelet estimation procedure provides good estimates of the underlying NHPP. From Figures 6, 8, and 10, we can see that the width of the tolerance band in case of the mean value function increases with time. This is because the error is cumulative over time.

5 CONCLUSION

There are several important objectives we have achieved during this research. We have developed a method for estimating the rate function of a nonhomogeneous Poisson process using wavelets. Utilizing a positive wavelet estimator, we are able to obtain a non-negative estimate of the NHPP rate function.

The wavelet estimator has some advantages over current parametric estimators in that we do not require prior knowledge about the form of the rate function. In addition wavelets provide us with a much more flexible model in that we can represent long term trend as well as asymmetric cyclic behavior and other irregular properties. Although the wavelet estimation procedure requires the storage of the observed data points, the only additional storage is the coefficients of the wavelet basis function.

The empirical experimental performance evaluation has shown that the wavelet estimation procedure has the ability to consistently provide adequate estimates of the rate function of an NHPP. In addition, the experiment has shown that the wavelet estimation procedure was exceptionally good at doing what non-parametric estimators are desired to do which is to fit the observed arrival process.

6 RECOMMENDATIONS FOR FUTURE WORK

This research focused on the development of the methodology for calculating a wavelet estimator for the rate function of an NHPP. Now that this methodology has been established, there are some areas that one may want to investigate further. First, convergence theorems need to be established to show the asymptotic convergence of the wavelet estimator to the theoretical rate function. Second, we need to investigate the choice of the basis function used in conducting the wavelet fitting procedure to determine the advantages and disadvantages in terms of computational effort and convergence properties. Finally, a large-scale experiment should be conducted on a set of real world data to establish the credibility and usefulness of the wavelet estimation procedure in practice.

REFERENCES

Bruce, A. and Gao, H. 1997. *Applied Wavelet Analysis with S-PLUS*, Springer Publishers, New York.

Çinlar, E. 1975. *Introduction to Stochastic Process,* Prentice-Hall Inc., Englewood Cliffs, New Jersey.

Cox, D.R. and Lewis, P.A.W. 1966. *The Statistical Series of Events*. Chapman and Hall, London.

Daubechies, I. 1992. *Ten Lectures on Wavelets*, SIAM CBMS Series No.6, Pennsylvania.

Erlebacher, G., Hussaine, M. and Jameson, L. 1996. *Wavelets, Theory and Applications*. Oxford University Press, New York.

Kuhl, M.E., Wilson, J.R. and Johnson, M.A. 1997. Estimating and simulating Poisson processes having trends or multiple periodicities. . *IIE Transactions*, 29: 201-211.

Kuhl, M.E., Damerdji, H. and Wilson, J.R. 1998. Least Squares Estimation of Nonhomogeneous Poisson Process. In *Proceedings of the 1998 Winter Simulation*

Conference, eds. D.J. Medeiros, E.F. Watson, J.S. Carson, and M.S. Manivannan, 637-645.

Lewis, P.A.W. and Shedler, G.S. 1979. Simulation of nonhomogeneous Poisson processes by thinning. *IBM J.Research Development,* 32: 403-413.

Nievergelt, Y. 1999. *Wavelets Made Easy*. Birkhauser, Boston.

Walter, G.G. and Shen, X. 1998. Positive Estimation with Wavelets, *Contemporary Mathematics*, 216: 63-79.

Young, R.1996. *Wavelet Theory and its Applications*, Kluwer Academic Publishers, New York.

AUTHOR BIOGRAPHIES

MICHAEL E. KUHL is an Assistant Professor in the Industrial and Manufacturing Systems Engineering Department at Louisiana State University. He received his B.S.I.E. from Bradley University and his M.S.I.E./O.R. and Ph.D. from North Carolina State University. He is a member of INFORMS, IIE, and ASEE. His interests include simulation input modeling, design and analysis of simulation experiments, and information systems. His email and web addresses are <mkuhl@lsu.edu> and <www.imse.lsu.edu/kuhl>.

PRASHANT S. BHAIRGOND is a master's student in the Department of Industrial and Manufacturing Systems Engineering at Louisiana State University. He is a member of Alpha Pi Mu and IIE. His email address is <pbhair1@lsu.edu>.