Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

MISE-OPTIMAL INTERVALS FOR MNO–PQRS ESTIMATORS OF POISSON RATE FUNCTIONS

Huifen Chen

Department of Industrial and Systems Engineering
Chung-Yuan University
Chung-Li, 320 Taoyuan, Taiwan

Bruce W. Schmeiser

School of Industrial Engineering
Purdue University
West Lafayette, IN 47907, USA

ABSTRACT

A Poisson point process is characterized by its rate function. One family of rate-function approximations is the authors' MNO–PQRS, which is based on a piecewise-quadratic function for each equal-width time interval. Fitting MNO–PQRS is based on the number of observed arrivals in each such interval. Therefore, the first step in fitting is to choose the number of intervals. Previously the authors discussed choosing the number of intervals for piecewise-constant rate functions. Here we extend those ideas to choosing the number of intervals for MNO–PQRS. Typically, the number of MNO–PQRS intervals is smaller than the piecewise-constant number. The results can be applied to non-Poisson arrival times; we do not investigate sensitivity to the Poisson assumptions.

1 INTRODUCTION

Given a set of observed nonhomogeneous Poisson process (NHPP) arrival times $t_1, t_2, \ldots, t_n$ in a known time interval, we are interested in estimating the unknown true rate function $\lambda(t)$ over the time interval. Here n is the observed value of N, the Poisson-distributed number of arrivals. Similarly, if the arrival indices are permuted randomly, then the arrival times are independent, each with density function $\lambda(t)/\overline{\lambda}$, where $\overline{\lambda}$ is the mean number of arrivals, the area under the rate function.

Despite the close relationship between rate and density functions, their estimation differs in that the lower and upper bounds are known for point processes and (often) unknown for density estimation. Typically, estimation methods for both rate and density functions partition the data into contiguous intervals and use the interval counts (i.e. number of observations in each interval) to fit a model family. In density estimation with unknown bounds, the emphasis is on interval (bin) widths; in rate estimation with known bounds, the analogy is the number of intervals.

1.1 Problem Statement

Let k be a dummy variable denoting the number of intervals, let k^* denote the unknown optimal number of intervals, and let $\hat{k}^*$ denote the estimator of k^*. Let $\hat{\lambda}$ denote an approximation to λ. Using observed arrival times from λ to create $\hat{k}^*$ begins by choosing a family of rate functions with k as a parameter and then, for a fixed k value, obtain a $\hat{\lambda}$ approximation by fitting the family to the k interval counts in terms of its other parameters. Our problem is to create an appropriate estimator $\hat{k}^*$ of k^* as a function of the observed arrival times.

1.2 Problem Solution

We follow four steps to estimate k^* based on the observed arrival times. First, choose an expected-value distance metric between λ and $\hat{\lambda}$. Second, derive the metric, say $g(k)$, as a function of the process assumptions (for us, Poisson) and the implied statistical behavior of $\overline{\lambda}$. Third, create an estimate of $g(k)$,

978-1-7281-3283-9/19/$31.00 ©2019 IEEE

say $\hat{g}(k)$, as a function of the observed arrival times. Fourth, choose a method that computes a value $\hat{k}^*$ using $\hat{g}(k)$ values to approximate k^* that minimizes g over $k = 1, 2, \ldots$.

In particular, in this paper our family of rate-function models, and method of fitting to count data from k intervals, is MNO–PQRS, developed in Chen and Schmeiser (2014; 2017) and reviewed in Section 3.1. Our decisions for the four steps to estimate k^* are as follows. First, our distance metric is the mean integrated squared error (MISE) criterion, as used in Chen and Schmeiser (2018) for the family of piecewise-constant rate functions. Second, we derive the MISE value, $g(k)$, for the PQRS family in Section 3.3. Third, we obtain $\hat{g}(k)$, an unbiased estimator of $g(k)$, in Section 3.4. Fourth, rather than defining $\hat{k}^*$ to be the global argmin of $\hat{g}(k)$, we introduce intuitively and empirically better heuristics in Section 3.5. Having chosen the number of intervals, the MNO–PQRS rate-function estimate is fitted using $\hat{k}^*$ intervals.

Two other choices have been made: piecewise-constant family and I-SMOOTH family (Chen and Schmeiser 2011 and 2013), an iterative smoothing family and fitting method. Chen and Schmeiser (2018) consider the piecewise-constant family, defined by the k counts, use the MISE criterion, and define $\hat{k}^*$ using the global minimum. For this paper, we attempted I-SMOOTH. After I-SMOOTH iteration i, a nonnegative piecewise-constant rate function with $2^i k$ intervals is obtained; call this function I-SMOOTH (i). I-SMOOTH is reviewed in Appendix A. In Appendix A, we state g for I-SMOOTH(1). Ideally, we would present these values for I-SMOOTH(i) for $i = 2, 3, \ldots, \infty$, but the derivations become complicated. Based on comparisons in Chen and Schmeiser (2015), for fixed k, the I-SMOOTH(∞) and MNO-PQRS results are similar, so likely their values of k^* are similar. Until analysis for general I-SMOOTH(i) is available, using MNO-PQRS with its $\hat{k}^*$ is more appealing. Or, if I-SMOOTH(i) for larger values of i is used, probably the MNO–PQRS value of $\hat{k}^*$ is more appropriate than the I-SMOOTH(1) value.

2 LITERATURE REVIEW

Rather than the usual literature review, we refer the reader to the review in Chen and Schmeiser (2018). We summarize two updates.

First is Chen and Schmeiser (2018) itself. They derive the gMISE criterion, g_{C}, for piecewise-constant rate functions, create an unbiased estimator $\hat{g}_{\mathrm{C}}$, and estimate the MISE-optimal number of intervals with $\hat{k}^*_{\mathrm{C}}$(GM), the global minimum (GM).

Second, Morgan (2019) fits a spline family with equal-width knot intervals using maximum likelihood. The number of intervals is a given algorithm parameter. She compares her spline rate function with the piecewise-linear rate function of Zheng and Glynn (2017) and with MNO–PQRS using $\hat{k}^*_{\mathrm{C}}$(GM) intervals as suggested in Chen and Schmeiser (2018). Unlike last year, where we recommended using $\hat{k}^*_{\mathrm{C}}$(GM) intervals to fit MNO–PQRS, we now recommend using $\hat{k}^*_{\mathrm{Q}}$(LLM2-PC), as discussed in Section 3.5. We look forward to seeing Morgan et al. (2019).

3 MNO–PQRS: ESTIMATING THE NUMBER OF INTERVALS AND FITTING

MNO–PQRS, which fits a piecewise-quadratic function to k arrival counts, is detailed in Section 3.1. Then Sections 3.2 through 3.5 explain our approach to the four steps to determine an appropriate value of k from the arrival times $t_1, t_2, \ldots, t_n$.

A difference between determining the number of intervals k and MNO–PQRS fitting is that fitting treats arrival counts as deterministic whereas determining the number of intervals assumes distributional properties on the arrivals. Hence, we make a notational change from Chen and Schmeiser (2017), where the number of arrivals in the ith interval was denoted by (the constant) λ_i. In this paper, where the number of intervals is a decision variable, $C_i(k)$ will denote the ith (random variable) count when there are k intervals.

The arrival counts $C_i(k)$ do not depend on time scaling. Nevertheless, we use two time scales: Unit-width intervals are central to both. For determining an appropriate number of intervals, k, we assume that arrival times lie in $[0,1]$; for fitting MNO–PQRS to k counts, we assume that the arrival times lie in $[0,k]$. That is, in Sections 3.2 through 3.5 the ith interval is $((i-1)/k, i/k]$, but in Section 3.1 the ith interval is

$(i-1,i]$. Therefore, the rate function for determining the number of intervals is k times larger than the fitted MNO–PQRS rate function.

3.1 MNO–PQRS Review

Chen and Schmeiser (2014; 2017) propose the MNO–PQRS algorithm for smoothing a piecewise-constant rate function with rates $\lambda_i = C_i(k)$, for $i = 1,2,\ldots,k$ for the k equal-width intervals. MNO–PQRS, with no user-specified parameters, returns a smoother rate function that maintains the piecewise-constant rates.

MNO–PQRS proceeds in two steps: PQRS (Piecewise-Quadratic Rate Smoothing) returns a continuous and differentiable piecewise-quadratic function without regard to negativity. If negative rates occur, then MNO (Max Nonnegativity Ordering) returns the maximum of zero and another piecewise-quadratic function. Chen and Schmeiser (2015) propose an efficient inverse-transformation method to generate arrival times of a NHPP having the MNO–PQRS rate function.

MNO–PQRS considers five time-horizon contexts: Context 0 (cyclic) and finite-horizon Contexts 1–4 (two end-time rates, only left end-time rate, only right end-time rate, and no end-time rate specified, respectively). For Context 0, the rate cycles with period length k. For the four finite-horizon contexts, the rate function is defined only in time interval $[0,k]$. Although we have results for all five contexts, we discuss only Contexts 0 and 4 in this paper, so we define here no notation for the end-time rates.

We now state the MNO–PQRS rate function explicitly. Let $q(x : a,b,c) = ax^2 + bx + c$, $x \in [0,1]$, denote a quadratic function in the unit interval. The PQRS function is

$$\tau_0(t : \underline{a}, \underline{b}, \underline{c}) = q(x : a_i, b_i, c_i),$$

where $j = \max\{0, \lceil t/k \rceil - 1\}$ is the number of previous cycles, $i = \max\{1, \lceil t - jk \rceil\}$ is the interval number, and $x = t - jk - i + 1$ is the fractional time within interval i (except that the time intervals are closed on the right, so $x = 1$ when t is integer).

The PQRS function has $3k$ parameters: $\underline{a} = (a_1, a_2, \ldots, a_k)$, $\underline{b} = (b_1, b_2, \ldots, b_k)$, and $\underline{c} = (c_1, c_2, \ldots, c_k)$, which are chosen to maintain the piecewise-constant rates, continuity of the PQRS function and continuity of the derivatives. In particular, for Contexts 0 and 4 the $3k$ parameters are the linear functions

$$a_i = \sum_{j=1}^{k} \alpha_{ij} C_j(k), \quad b_i = \sum_{j=1}^{k} \beta_{ij} C_j(k), \quad c_i = \sum_{j=1}^{k} \theta_{ij} C_j(k). \tag{1}$$

The constant weights α_{ij}'s, β_{ij}'s, and θ_{ij}'s depend only on the context, as shown in Chen and Schmeiser (2017).

If the PQRS function τ_0 is non-negative everywhere, the MNO–PQRS function $\tau = \tau_0$. Otherwise, the MNO logic modifies τ_0 by

$$\tau^+(t : \underline{a}^+, \underline{b}^+, \underline{c}^+) = \max\{0, q(x : a_i^+, b_i^+, c_i^+)\},$$

where i and x are defined as for τ_0. The parameters $(\underline{a}^+, \underline{b}^+, \underline{c}^+)$ are computed using the MNO logic discussed in Chen and Schmeiser (2017). The MNO-PQRS rate function is then $\tau = \tau^+$.

3.2 The MISE Criterion

The MISE criterion is the mean integrated squared distance between the true rate function λ and the estimated rate function $\hat{\lambda}$. Under the assumption that arrival times are observed in the unit time interval, the MISE criterion is

$$\mathrm{E}\Big\{ \int_0^1 [\lambda(t) - \hat{\lambda}(t)]^2 dt \Big\}.$$

Rather than using MISE directly, we use the gMISE criterion

$$g(k) = \mathrm{E}\Big\{ \int_0^1 [\hat{\lambda}^2(t) - 2\lambda(t)\hat{\lambda}(t)] dt \Big\}.$$

Because $\int_0^1 [\lambda(t)]^2 dt$ is a constant, minimizing gMISE yields the same result as minimizing MISE. We define the optimal number of intervals to be

$$k^* \equiv \text{argmin}_{k=1,2,\ldots} g(k).$$

3.3 The gMISE Criterion for PQRS

Here we specialize $g(k)$ to PQRS. We ignore the complication that arises from MNO's non-negatively logic. In practice, the fitted PQRS rate function is seldom negative. Ignoring negativity, however, harms the fitted MNO–PQRS rate function when the true rate is close to zero.

Fix the number of equal-width intervals, k, in $\{1,2,\ldots\}$. The arrival count in the ith interval $((i-1)/k, i/k]$ is $C_i(k)$ for $i = 1,2,\ldots,k$. The PQRS rate-function estimator at time $t \in [0,1]$ is then

$$\hat{\lambda}_Q(t;k) = k\,q(x : a_i, b_i, c_i), \tag{2}$$

where $i = \max\{1, \lceil kt \rceil\}$ is the interval number, and $x = kt - i + 1$ is the magnified fractional time within the ith magnified interval $((i-1), i]$. The $3k$ parameters $(\underline{a}, \underline{b}, \underline{c})$ are as shown in Equation (1).

Result 1 states the gMISE function for the PQRS rate-function estimator, as a function of k.

Result 1 Consider a NHPP on the unit interval with rate function λ. The gMISE criterion for the PQRS estimator $\hat{\lambda}_Q(t;k)$ is

$$g_{\mathrm{Q}}(k) = k\left\{\overline{\lambda} - \Big(\sum_{i=1}^{k} \mathrm{E}^2[C_i(k)]\Big) + h_Q(k)\right\}, \tag{3}$$

where

$$\begin{aligned} h_Q(k) \quad = \quad & \sum_{i=1}^{k} \frac{4\mathrm{E}(a_i^2)}{45} + \frac{\mathrm{E}(b_i^2)}{12} + \frac{\mathrm{E}(a_i b_i)}{6} + \frac{2\mathrm{E}(a_i)\mathrm{E}[C_i(k)]}{3} + \mathrm{E}(b_i)\mathrm{E}[C_i(k)] \\ & -2\mathrm{E}(b_i)\int_{(i-1)/k}^{i/k} (kt-i+1)\lambda(t)dt - 2\mathrm{E}(a_i)\int_{(i-1)/k}^{i/k} (kt-i+1)^2 \lambda(t)dt, \end{aligned}$$

$\mathrm{E}(a_i) = \sum_{j=1}^k \alpha_{ij}\mathrm{E}[C_j(k)]$, $\mathrm{E}(b_i) = \sum_{j=1}^k \beta_{ij}\mathrm{E}[C_j(k)]$, $\mathrm{E}(a_i b_i) = \sum_{j=1}^k \alpha_{ij}\beta_{ij}\mathrm{E}[C_j(k)] + \mathrm{E}(a_i)\mathrm{E}(b_i)$, $\mathrm{E}(a_i^2) = \sum_{j=1}^k \alpha_{ij}^2 \mathrm{E}[C_j(k)] + \mathrm{E}^2(a_i)$, $\mathrm{E}(b_i^2) = \sum_{j=1}^k \beta_{ij}^2 \mathrm{E}[C_j(k)] + \mathrm{E}^2(b_i)$, $\mathrm{E}[a_i C_i(k)] = \mathrm{E}[C_i(k)][\alpha_{ii} + \mathrm{E}(a_i)]$, and $\mathrm{E}[b_i C_i(k)] = \mathrm{E}[C_i(k)][\beta_{ii} + \mathrm{E}(b_i)]$. (For simplicity, the PQRS $\underline{c}$ is expressed as a function of $\underline{a}$ and $\underline{b}$, and hence, does not appear in $h_Q(k)$).

Although we do not provide the proof here, we note that this analytical form of gMISE for PQRS arises for three reasons. First, the integration of the quadratic function q is tractable. Second, the parameters $\underline{a}$ and $\underline{b}$ are linear combinations of the counts $C_1(k), C_2(k), \ldots C_k(k)$. Third, we assume the Poisson properties, in particular $\mathrm{E}[C_i(k)] = \mathrm{var}[C_i(k)]$ and $\mathrm{cov}[C_i(k), C_j(k)] = 0$ for $i \neq j$.

3.4 Estimating gMISE for PQRS

In practice, λ is unknown; therefore gMISE is unknown. Here we create an unbiased estimator of $g_{\mathrm{Q}}(k)$ based on NHPP arrival times $T_1, T_2, \ldots, T_N$ in the unit interval. For any number of intervals k, $N = \sum_{i=1}^k C_i(k)$. In Section 3.5, we minimize the unbiased estimator, $\hat{g}_{\mathrm{Q}}$, to estimate k^*.

For fixed k, Result 2 provides an unbiased estimator $\hat{g}_{\mathrm{Q}}(k)$ of the gMISE criterion $g_{\mathrm{Q}}(k)$ for the PQRS estimator $\hat{\lambda}_Q$ in Equation (2). In this subsection only, we simplify the notation $C_i(k)$ to C_i.

Result 2 Consider arrival times $T_1, T_2, \ldots, T_N$ from the NHPP with unknown rate function λ on the unit interval. Let T_{ij} denote the jth arrival time in the ith time interval $((i-1)/k, i/k]$, $i = 1,\ldots,k$, $j = 1,\ldots,C_i$.

An unbiased estimator of gMISE, shown in Equation (3), for the PQRS rate-function estimator $\hat{\lambda}_Q(t;k)$ is

$$\hat{g}_{\mathrm{Q}}(k) \quad = \quad k\left\{2N - \Big(\sum_{i=1}^{k} C_i^2\Big) + \hat{h}_Q(k)\right\},$$

where

$$\begin{aligned}
\hat{h}_Q(k) \quad = \quad & \sum_{i=1}^{k} \frac{4a_i^2}{45} + \frac{b_i^2}{12} + \frac{a_i b_i}{6} \\
& -2(a_i - \alpha_{ii})\left[k^2\Big(\sum_{j=1}^{C_i} T_{ij}^2\Big) - 2k(i-1)\Big(\sum_{j=1}^{C_i} T_{ij}\Big) + \Big(i^2 - 2i + \frac{2}{3}\Big)C_i\right] \\
& -2(b_i - \beta_{ii})\left[k\Big(\sum_{j=1}^{C_i} T_{ij}\Big) - \Big(i - \frac{1}{2}\Big)C_i\right].
\end{aligned}$$

We provide no proof here, but Result 2 depends upon two logic threads. First, any (unordered) arrival time in the ith interval has density function $\lambda(t)/\mathrm{E}(C_i)$ for $(i-1)/k \le t \le i/k$. Therefore, $\sum_{j=1}^{C_i} T_{ij}$ and $\sum_{j=1}^{C_i} T_{ij}^2$ are unbiased estimators for $\int_{(i-1)/k}^{i/k} t\lambda(t)dt$ and $\int_{(i-1)/k}^{i/k} t^2\lambda(t)dt$, respectively. Second, Section 3.2 of Chen and Schmeiser (2018) says that $\hat{g}_{\mathrm{C}} = k(2N - \sum_{i=1}^{k} C_i^2)$ is an unbiased estimator for $g_{\mathrm{C}} = k[\overline{\lambda} - \sum_{i=1}^{k} \mathrm{E}^2(C_i)]$. In addition, the proof uses C_i as an unbiased estimator for $\mathrm{E}(C_i)$, $C_i^2 - C_i$ unbiased for $\mathrm{E}^2(C_i)$, $(a_i C_i - \alpha_{ii} C_i)$ unbiased for $\mathrm{E}(a_i)\mathrm{E}(C_i)$, and $(b_i C_i - \beta_{ii} C_i)$ unbiased for $\mathrm{E}(b_i)\mathrm{E}(C_i)$.

The estimator $\hat{g}_{\mathrm{Q}}$ is similar to the unbiased piecewise-constant result in Chen and Schmeiser (2018), where the unbiased estimator is $\hat{g}_{\mathrm{C}}$. The additional term is an unbiased estimator of $kh_Q(k)$ in Equation (3).

3.5 Estimating k^*

We now turn to estimating $k^* \equiv \mathrm{argmin}_{k=1,2,\ldots}\, g_{\mathrm{Q}}(k)$, the MISE-optimal number of intervals, using the unbiased estimated gMISE values $\hat{g}_{\mathrm{Q}}(k)$. For piecewise-constant (PC) rate functions, Chen and Schmeiser (2018) use the GM $\hat{k}_{\mathrm{C}}^*(\mathrm{GM}) \equiv \mathrm{argmin}_{k=1,2,\ldots}\, \hat{g}_{\mathrm{C}}(k)$. The analogy for PQRS is $\hat{k}_{\mathrm{Q}}^*(\mathrm{GM}) \equiv \mathrm{argmin}_{k=1,2,\ldots}\, \hat{g}_{\mathrm{Q}}(k)$.

For eight Monte Carlo realizations from the shown cyclic rate function (solid curve), Figure 1 shows $\hat{k}_{\mathrm{Q}}^*(\mathrm{GM})$ and the associated MNO-PQRS fit (dash curve). The optimal number of intervals is $k^* = 4$. Four of the eight realizations yield $\hat{k}_{\mathrm{Q}}^*(\mathrm{GM})$ values of three, four, and five, with corresponding good fits. The other four realizations lie in the distribution's long right tail, with corresponding fitted rate functions with undesirable fluctuations. The implication is that avoiding large numbers of intervals is good. Therefore, we now discuss alternatives to the global minimum. That is, we consider alternative definitions of $\hat{k}^*$.

As a guiding principle, $\hat{k}^*$ needs to be defined so that $\hat{g}(\hat{k}^*)$ is a local minimum. Then alternatives to the global minimum exist whenever the global minimum is not the only local minimum. Pasupathy and Schmeiser (2010), in using MSER (Marginal Standard Error Rule) estimators to choose the amount of initial data to delete in steady-state simulation, compare two alternatives to the global minimum. The simpler is the leftmost local minimum (LLM). More complicated, and therefore stated explicitly in Appendix B, is the leftmost local minimum of the local minima (LLM2). We add a third alternative, $\hat{k}_{\mathrm{Q}}^*(\mathrm{LLM2\text{-}PC})$, which constrains the LLM2 search to the range $\{1, 2, \ldots, \hat{k}_{\mathrm{C}}^*(\mathrm{LLM2})\}$, where $\hat{k}_{\mathrm{C}}^*(\mathrm{LLM2})$ is the LLM2 of $\hat{g}_{\mathrm{C}}$ for PC rate functions. From their definitions, for every realization $\hat{k}_{\mathrm{Q}}^*(\mathrm{LLM}) \le \hat{k}_{\mathrm{Q}}^*(\mathrm{LLM2\text{-}PC}) \le \hat{k}_{\mathrm{Q}}^*(\mathrm{LLM2}) \le \hat{k}_{\mathrm{Q}}^*(\mathrm{GM})$. Therefore, LLM, LLM2-PC, and LLM2, are heuristics having less right tail than the statistical distribution of $\hat{k}_{\mathrm{Q}}^*(\mathrm{GM})$. As $\overline{\lambda}$ goes to infinity, $\hat{g}_{\mathrm{Q}}(k)$ converges to $g_{\mathrm{Q}}(k)$, and all four alternatives converge to their analogies of g_{Q}. Therefore, asymptotically the global minimum is the safest alternative, despite not performing well when $\overline{\lambda}$ is small.

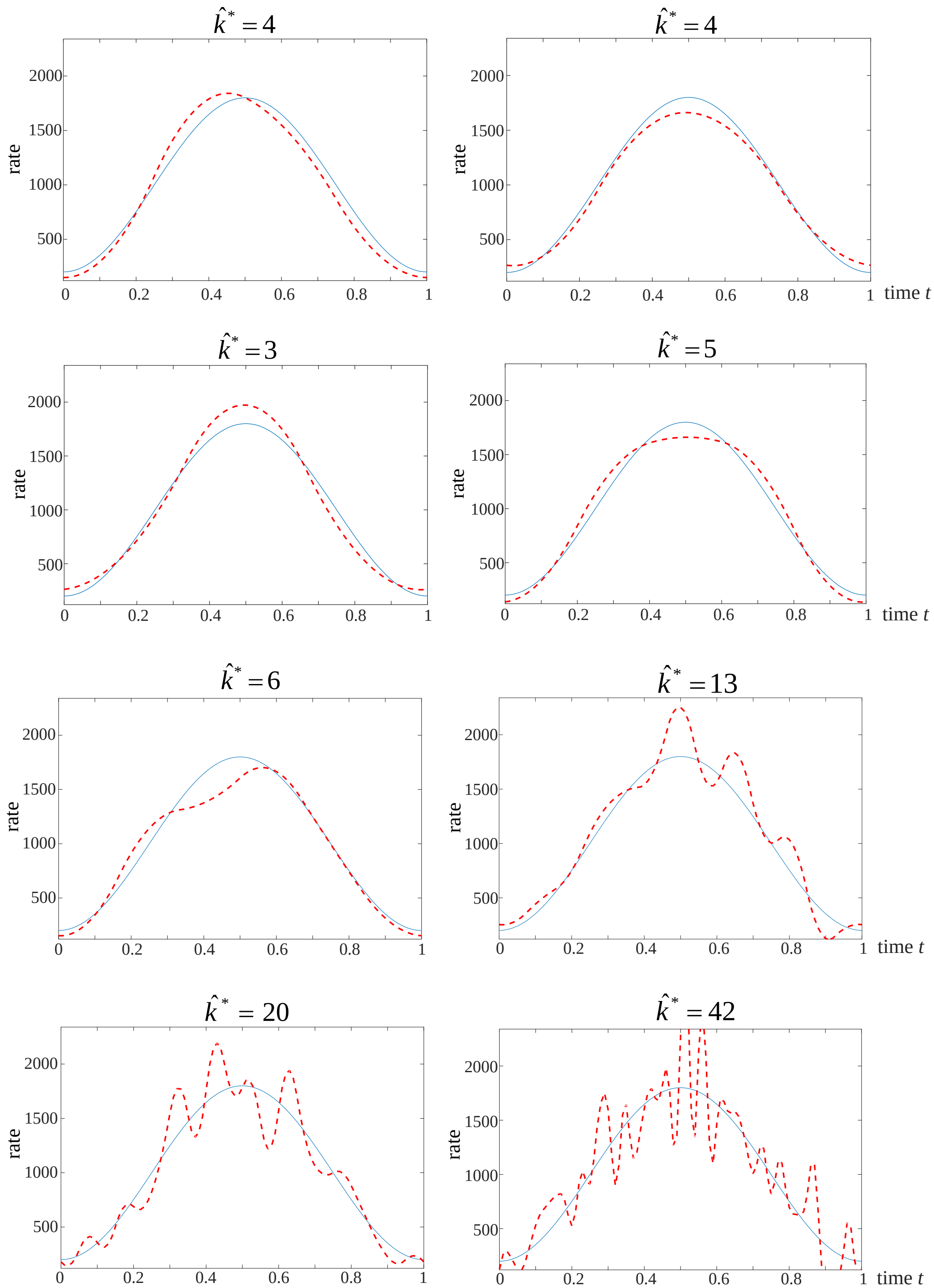

Figure 1: Eight realizations of the MNO-PQRS fit using the global-minimum estimated number of intervals.

4 EMPIRICAL RESULTS

We now provide empirical results, based on the two rate-function examples in Section 4.1. In Section 4.2 we compare piecewise-constant, I-SMOOTH(1), and PQRS models using the gMISE function. In Section 4.3 we compare the k^* estimation alternatives.

4.1 An Example Rate Function

To make various comparisons, both numerical and Monte Carlo, we use the cosine rate function

$$\lambda(t) = \mu + \delta \cos(2\pi(\eta t + \xi))$$

over the unit interval. The parameters μ, δ, ξ, and η are the overall mean, amplitude, phase, and frequency, respectively. Since $\lambda(t)$ must be nonnegative, $|\delta| \le \mu$. Chen and Schmeiser (1992) propose an inverse-transformation method to generate arrival times from a NHPP with the cosine rate function.

We consider two examples. Both use parameter values $\mu = \overline{\lambda} - \delta[\sin(2\pi(\eta+\xi)) - \sin(2\pi\xi)]/(2\pi\eta)$ (so that the mean number of arrivals is $\overline{\lambda}$), $\delta = .8\overline{\lambda}$, and $\xi = 0.5$. Example 1 is Context 0 (cyclic) with $\eta = 1$. Example 2 is Context 4 (finite horizon with no specified end-time rate) with $\eta = 0.5$.

4.2 Comparison Over Rate-Function Models

We compare k^* and gMISE for the piecewise-constant, I-SMOOTH(1), and PQRS rate-function estimators in Table 1 using Example 1. Column 1 is the $\overline{\lambda}$ value; Columns 2 to 4 are the value of k^* for the three rate-function estimators; and Columns 5 to 7 are the corresponding scaled gMISE criterion, $g(k^*)/(\overline{\lambda}^2)$. The entries are printed up to the digit that shows the differences.

For all $\overline{\lambda}$ values, the k^* values decrease from piecewise constant, to I-SMOOTH(1), to PQRS. The gMISE values decrease similarly; Chen and Schmeiser (2018) derived the gMISE $g_{\mathrm{C}}(k) = k\{\overline{\lambda} - \sum_{i=1}^{k} \mathrm{E}^2[C_i(k)]\}$ for piecewise-constant rate functions; the I-SMOOTH(1) gMISE $g_{\mathrm{S}}(k)$ is in Appendix A. The comparison between piecewise constant and PQRS is explicit, being the $kh_Q(k)$ term in Result 1. (Similarly, the difference between the piecewise-constant and I-SMOOTH(1) gMISE is $kh_S(k)$ in Appendix A.) Both decreases makes sense in that these rate-function models use k, $2k$, and $3k$ parameters, respectively.

For all three rate-function estimators, the optimal MISE criterion goes to zero as the mean number of arrivals increases. From Table 1, $g(k^*)/\overline{\lambda}^2$ decreases to the constant -1.32. Here, for Example 1, $\int_0^1 [\lambda(t)]^2 dt = 1.32\overline{\lambda}^2$, and hence, the limiting value of the optimal MISE criterion is $g(k^*) + \int_0^1 [\lambda(t)]^2 dt = 0$. That is, all three models (using their own optimal number of intervals) converge to the true rate function λ at the same speed. For finite $\overline{\lambda}$, however, the relative $g(k^*)$ differences among the three models decrease as $\overline{\lambda}$ increases, as shown in the last three columns of Table 1.

Table 1: Using Example 1, comparisons of the MISE-optimal grouping for piecewise-constant, I-SMOOTH(1), and PQRS rate-function models.

	k^*			$g(k^*)/(\overline{\lambda}^2)$		
$\overline{\lambda}$	piecewise constant	I-SMOOTH(1)	PQRS	piecewise constant	I-SMOOTH(1)	PQRS
1	1	1	1	0	0	0
10	3	3	3	−0.92	−0.94	−0.95
10^2	6	4	3	−1.23	−1.26	−1.28
10^3	13	8	4	−1.301	−1.307	−1.315
10^4	28	17	5	−1.316	−1.317	−1.319
10^5	59	36	7	−1.3191	−1.3194	−1.3199
10^6	128	78	9	−1.31981	−1.31987	−1.31999

In addition to having better gMISE values, PQRS has three other advantages. First, the PQRS rate function is continuous. Second, I-SMOOTH requires a stopping rule. Third, the PQRS k^* grows very slowly with $\overline{\lambda}$; its $3k^*$ parameters are fewer than the optimal number of parameters for piecewise constant and I-SMOOTH(1) when $\overline{\lambda}$ is one thousand or bigger. In addition to parsimony, using fewer intervals reduces the probability of empty cells.

4.3 Comparison of $\hat{k}^*$ Alternatives

Here we compare four PQRS $\hat{k}^*$ alternatives: GM, LLM, LLM2, and LLM2-PC. Simulation experiments are run with 1000 replications using Examples 1 and 2. The Monte Carlo results are correct within one unit of the right-most displayed digit.

The results are shown in Tables 2 and 3. In both tables, Column one is the mean number of arrivals, $\overline{\lambda}$; Column two is the true MISE-optimal number of intervals, k^*; Columns three and four are the estimated mean and standard deviation of $\hat{k}^*$; Column five is the estimated mean squared error of $\hat{k}^*$ divided by $(k^*)^2$; Column six is the estimated scaled excess gMISE (SE-gMISE) multiplied by $\overline{\lambda}$; and Column seven is the maximum observed value of $\hat{k}^*$ over 1000 replications. (Chen and Schmeiser (2018) define SE-gMISE $= \{\mathrm{E}[g(\hat{k}^*)] - g(k^*)\}^2/\overline{\lambda}^2$.)

Table 2 shows that for this cyclic example, LLM performs badly by highly under estimating k^*; worse, its SE-gMISE does not go to zero as $\overline{\lambda}$ goes to infinity. This is because $k = 1$ is a local minimum of the PQRS gMISE in this example. Compared to the global minimum $\hat{k}^*$, the LLM2 estimator has lower bias and reduces 44% to 58% of $\hat{k}^*$'s standard deviation; the mean squared error mse$(\hat{k}^*)$ is reduced for 67% to 83%. This substantial improvement is also shown in SE-gMISE. Although SE-gMISE decreases with $\overline{\lambda}$ for both the global minimum and LLM2, the limiting value (SE-gMISE)$\overline{\lambda}$ of LLM2 is lower. The maximum observed value max$(\hat{k}^*)$ of LLM2 is about an half of that of the global minimum, showing that the distribution of LLM2 has a shorter right tail. Hence, LLM2 is an obvious improvement of the global minimum. LLM2-PC performs better than LLM2 when $\overline{\lambda}$ is small. For $\overline{\lambda}$ being 10 thousand or higher, the LLM2-PC and LLM2 have the same results. Hence, we conclude that the LLM2-PC performs the best.

Table 2: Using Example 1, comparisons of the GM, LLM, LLM2, and LLM2-PC estimators.

	$\overline{\lambda}$	k^*	E$(\hat{k}^*)$	std$(\hat{k}^*)$	mse$(\hat{k}^*)/(k^*)^2$	(SE-gMISE)$\overline{\lambda}$	max$(\hat{k}^*)$
GM	10^2	3	4.3	2.9	1.11	1.20	38
	10^3	4	4.9	3.2	0.71	1.88	42
	10^4	5	5.9	3.1	0.41	1.71	43
	10^5	7	7.6	3.3	0.23	2.00	50
	10^6	9	9.4	2.7	0.09	1.69	35
LLM	10^2	3	1.4	1.0	0.4	25	7
	10^3	4	1.4	1.1	0.5	271	7
	10^4	5	1.5	1.4	0.6	2,757	7
	10^5	7	1.8	1.9	0.6	26,746	9
	10^6	9	1.9	2.3	0.7	277,111	12
LLM2	10^2	3	3.8	1.3	0.25	0.60	11
	10^3	4	4.4	1.4	0.14	1.35	15
	10^4	5	5.4	1.4	0.08	1.17	15
	10^5	7	6.9	1.4	0.04	1.30	14
	10^6	9	8.9	1.5	0.03	1.39	16
LLM2-PC	10^2	3	3.6	1.0	0.15	0.43	11
	10^3	4	4.4	1.3	0.11	1.29	13
	10^4	5	5.4	1.4	0.08	1.17	15
	10^5	7	6.9	1.4	0.04	1.30	14
	10^6	9	8.9	1.5	0.03	1.39	16

Table 3 shows that for the finite-horizon example, LLM does not perform as badly as in Table 2. However, LLM still under estimates k^*. Compared to the global minimum, LLM2 reduces 0% to 78%, 28% to 42% and 49% to 74% of $\hat{k}^*$'s bias, standard deviation and mean square error, respectively. The (SE-gMISE)$\overline{\lambda}$ of LLM2 is lower than that of the global minimum for all $\overline{\lambda}$ values shown. The LLM2-PC performs slightly better than LLM2 when $\overline{\lambda}$ is less than 100 thousand.

Both Tables 2 and 3 show that LLM2 and LLM2-PC are substantially better than the GM $\hat{k}^*$ in terms of the bias, variance, and extreme value; LLM2-PC performs slightly better than LLM2 when $\overline{\lambda}$ is small.

Table 3: Using Example 2, comparisons of the GM, LLM, LLM2, and LLM2-PC estimators.

	$\overline{\lambda}$	k^*	E($\hat{k}^*$)	std($\hat{k}^*$)	mse($\hat{k}^*$)/(k^*)^2	(SE-gMISE)$\overline{\lambda}$	max($\hat{k}^*$)
GM	10^2	2	3.2	3.1	2.9	1.47	37
	10^3	4	4.4	3.1	0.6	1.87	27
	10^4	6	6.3	3.3	0.3	1.77	36
	10^5	8	9.1	3.6	0.2	2.10	36
	10^6	12	12.5	3.3	0.1	1.96	36
LLM	10^2	2	2.2	0.6	0.09	0.29	5
	10^3	4	2.5	0.9	0.20	1.68	6
	10^4	6	4.6	1.3	0.10	5.52	9
	10^5	8	7.3	1.2	0.03	1.70	12
	10^6	12	10.2	1.4	0.04	2.94	16
LLM2	10^2	2	2.8	1.5	0.73	0.92	15
	10^3	4	3.9	1.9	0.23	1.41	18
	10^4	6	5.7	1.8	0.10	1.19	18
	10^5	8	8.2	2.1	0.07	1.50	19
	10^6	12	11.8	2.4	0.04	1.60	24
LLM2-PC	10^2	2	2.4	1.0	0.30	0.44	15
	10^3	4	3.6	1.6	0.16	1.23	12
	10^4	6	5.6	1.7	0.08	1.12	14
	10^5	8	8.2	2.1	0.07	1.50	19
	10^6	12	11.8	2.4	0.04	1.60	24

5 DISCUSSION

We consider using the MNO–PQRS model to estimate the rate function of a NHPP from a set of arrival times. Our topic is how to estimate an appropriate number of time intervals using the given arrival times. For k intervals we derive the gMISE criterion $g_{\mathrm{Q}}(k)$, present an unbiased estimator $\hat{g}_{\mathrm{Q}}(k)$, and recommend the LLM2-PC heuristic for estimating k^*, the global minimum of $g_{\mathrm{Q}}(k)$, using the $\hat{g}_{\mathrm{Q}}(k)$ values.

Our Poisson assumption works in practice in two ways. First, the Poisson conditions often are a good approximation to reality; Poisson-distributed and independent interval counts arise simply from entities not coordinating with each other. Further, merged (superposed) non-Poisson point processes are (with few conditions) asymptotically Poisson (Cinlar 1972). Second—conditional upon a set of arrival times—our intuition is that the value of k^* is not sensitive to small violations of the Poisson assumptions. Sensitivity to the Poisson assumption and derivations for non-Poisson processes are topics of interest.

ACKNOWLEDGMENT

This research was partially supported by Taiwan's Ministry of Science and Technology through Project MOST105-2221-E-033-050-MY3.

A I-SMOOTH AND *g*MISE FOR I-SMOOTH(1)

After reviewing I-SMOOTH, we state the *g*MISE criterion for I-SMOOTH(1). As in Section 3, assume the arrival times lie in interval $[0,k]$ for fitting and in $[0,1]$ for determining k.

Chen and Schmeiser (2011, 2013) introduce I-SMOOTH, which smooths a piecewise-constant function with rates $\lambda_i = C_i(k)$, $i = 1,\ldots,k$, for k consecutive unit intervals. I-SMOOTH iteratively (the "I") smooths (the "SMOOTH") a piecewise-constant function by bisecting time intervals to obtain an updated piecewise-constant function with twice as many intervals. At each iteration, each interval's integral is maintained by decreasing the left half's rate by γ_i while increasing the right half's rate by γ_i. The value of γ_i is chosen to minimize the sum of squared second differences of the new rates. The γ_i values depend upon which of ten contexts applies: cyclic-horizon and nine finite-horizon contexts (caused by allowing zero, one, or two fixed rates on each of the left and right ends). At each iteration, negative rates are avoided by limiting $|\gamma_i|$ to being no more than the current rate.

We now discuss the *g*MISE criterion for the I-SMOOTH(1) rate-function estimator $\hat{\lambda}_S$. For a fixed value of k, $\hat{\lambda}_S$ in the unit time interval is

$$\hat{\lambda}_S(t;k) = \begin{cases} k[C_i(k)-\gamma_i] & \text{for } (i-1)/k < t \le (2i-1)/(2k) \\ k[C_i(k)+\gamma_i] & \text{for } (2i-1)/(2k) < t \le i/k \end{cases},$$

for $i = 1,2,\ldots,k$. The increment $\gamma_i = \sum_{j=1}^{k} w_{ij} C_j(k)$, where the constant weights w_{ij}'s depend only on the context, as shown in Chen and Schmeiser (2013).

Result 3 shows the *g*MISE criterion for the I-SMOOTH(1) rate-function estimator.

Result 3 Consider a NHPP on the unit interval with rate function λ. The *g*MISE criterion for the I-SMOOTH(1) estimator $\hat{\lambda}_S(t;k)$ is $g_S(k) = k\left\{\overline{\lambda} - \sum_{i=1}^{k} \mathrm{E}^2[C_i(k)] + h_S(k)\right\}$, where

$$\begin{aligned} h_S(k) \;\; = \;\; & \sum_{i=1}^{k}\left\{ \sum_{j=1}^{k} w_{ij}^2 \mathrm{E}[C_j(k)] + \left(\sum_{j=1}^{k} w_{ij}\mathrm{E}[C_j(k)]\right)^2 \right. \\ & \left. +2\left(\sum_{j=1}^{k} w_{ij}\mathrm{E}[C_j(k)]\right)\left(\mathrm{E}[C_{2i-1}(2k)] - \mathrm{E}[C_{2i}(2k)]\right)\right\}. \end{aligned}$$

The proof of Result 3 is similar to that for the piecewise-constant estimator in Chen and Schmeiser (2018). The term $h_S(k)$ could be estimated analogously to $h_Q(k)$ for PQRS using individual arrival times.

B LLM2 COMPUTER CODE

Subroutine `llm2` is logic for the LLM2 (leftmost minimum of the local minima) of `ghat`, an external function. The argmin is `kllm2`; its function value is `yllm2`. The maximum allowed value of `kllm2` is `kmax`, a user-specified parameter. As used in Section 3.5, the external function is $\hat{g}(k)$ and `kllm2` is $\hat{k}^*$.

To read this Fortran as pseudocode, know the following. An exclamation point begins a comment. Variables beginning with i, j, k, l, m, and n are integers; all other variables are doubles. `do` loops are analogous to `for` loops. All variables are local, with no reliance upon earlier calls to `llm2`.

```
subroutine llm2( kmax, kllm2, yllm2 )
! huifen chen and bruce schmeiser, february 15, 2019.
! find the location kllm2 of the Leftmost Local Minimum of
!   the all Local Minima (llm2) of values ghat(k), k=1,2,... .
! llm2 originated in pasupathy and schmeiser, wsc 2010.
! input parameter
!   kmax: kllm2 belongs to {1,2,...,kmax}
! output parameters
!   kllm2: llm2 location (decision variable)
!   yllm2: kllm2 function value
! external function
!   ghat(k): external application-specific function
!     returning the objective function for argument k
implicit double precision (a-h,o-z)
implicit integer (i-n)
kllm2 = 0       ! kllm2 = 0 is an error code
if (kmax .lt. 1) return
        ! ====== now k > 0 ======
y     = ghat(1) ! get the first function value
kllm2 = 1       ! initialize llm2 location
yllm2 = y       ! initialize llm2 value
if (kmax .eq. 1) return
endif
        ! ====== now k > 1 ======
yold1 = y       ! yold1 = ghat(1)
y     = ghat(2) ! get the second function value
if (y .lt. yold1) then
  kllm2 = 2       ! if k=2, then ghat(2) is a local minimum
  yllm2 = y       ! ghat(2) less than ghat(1)
endif
if (kmax .eq. 2) return
endif
        ! ====== now k > 2 ======
if (y .lt. yold1) kllm2 = - 1 ! ghat(2) is not yet a local minimum
do k = 3, kmax  ! loop over k
  yold2 = yold1   ! yold2 = ghat(k-2)
  yold1 = y       ! yold1 = ghat(k-1)
  y     = ghat(k) ! get this function value
  ! check whether yold1 is a local minimum
  if (yold1 .le. yold2 .and. yold1 .le. y) then
    if (kllm2 .ge. 1 .and. yold1 .ge. yllm2) return
    kllm2 = k - 1   ! update llm2 location
    yllm2 = yold1   ! update llm2 value
  endif
enddo
        ! ====== now k = kmax ======
if (y .lt. yllm2) then
  kllm2 = kmax    ! llm2 location is kmax
  yllm2 = y       ! llm2 value is the global minimum
endif
return
end
```

REFERENCES

Chen, H. and B. W. Schmeiser. 1992. "Simulation of Poisson Processes with Trigonometric Rates". In *Proceedings of the 1992 Winter Simulation Conference*, edited by J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson, 609–617. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chen, H. and B. W. Schmeiser. 2011. "I-SMOOTH: Iteratively Smoothing Piecewise-Constant Poisson-Process Rate Functions". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu, 469–480. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chen, H. and B. W. Schmeiser. 2013. "I-SMOOTH: Iteratively Smoothing Mean-Constrained and Nonnegative Piecewise-Constant Functions". *INFORMS Journal on Computing* 25(3):432–445.

Chen, H. and B. W. Schmeiser. 2014. "Piecewise-Quadratic Rate Functions: The Cyclic Context". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. D. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 486-497. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chen, H. and B. W. Schmeiser. 2015. "The MNO–PQRS Poisson Point Process: Generating the Next Event Time". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 575–585. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chen, H. and B. W. Schmeiser. 2017. "MNO–PQRS: Max Nonnegativity Ordering—Piecewise-Quadratic Rate Smoothing". *ACM Transactions on Modeling and Computer Simulation*, 27(3), Article 18:1–19.

Chen, H. and B. W. Schmeiser. 2018. "MISE-Optimal Grouping of Point Process Data with A Constant Dispersion Ratio". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 1563-1574. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Cinlar, E. 1972. "Superposition of Point Processes". In *Stochastic Point Processes: Statistical Analysis, Theory, and Applications*, edited by P. A. W. Lewis, 549–606. New York: John Wiley & Sons, Inc.

Morgan, L. E. 2019. *Quantifying and Reducing Input Modelling Error in Simulation*. PhD Thesis, STOR-i Centre for Doctoral Training, Lancaster University, Lancaster, United Kingdom.

Morgan, L. E., A. C. Titman, D. J. Worthington, and B. L. Nelson. 2019. A Spline-based Method for Modelling and Generating a Nonhomogeneous Poisson Process. In *Proceedings of the 2019 Winter Simulation Conference,* edited by N. Mustafee, K.-H. G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, submitted.

Pasupathy, R. and B. W. Schmeiser. 2010. "The Initial Transient in Steady-State Point Estimation: Contexts, a Bibliography, the MSE Criterion, and the MSER Statistic". In *Proceedings of the 2010 Winter Simulation Conference*, edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yucesan, 184–197. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Zheng, Z. and P. W. Glynn. 2017. "Fitting Continuous Piecewise Linear Poisson Intensities via Maximum Likelihood and Least Squares". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, 1740–1749. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

HUIFEN CHEN is a Professor in the Department of Industrial and Systems Engineering at Chung-Yuan University, Taiwan. She completed her Ph.D. in Industrial Engineering at Purdue University in 1994 and master in statistics at Purdue University in 1990. Her research interests include stochastic root finding, point-process rate smoothing and estimation, statistical process control, and public health. Her email address is huifen@cycu.edu.tw. Her website is https://huifenchen.github.io/.

BRUCE SCHMEISER is Professor Emeritus in the School of Industrial Engineering at Purdue University. His research interests center on developing methods for better simulation experiments. He is a fellow of INFORMS and IIE, as well as recipient of the Informs Simulation Society's 2014 Lifetime Professional Achievement Award. A long-time participant in the WSC, he served as the 1983 Program Chair and the 1988–1990 President of the Board of Directors. His e-mail address is bruceschmeiser@gmail.com.