

A Normal Copula Model for the Arrival Process in a Call Center

Nabil Channouf and Pierre L'Ecuyer

Département d'Informatique et de Recherche Opérationnelle
Université de Montréal, C.P. 6128, Succ. Centre-Ville
Montréal, H3C 3J7, CANADA
{channoun,lecuyer}@iro.umontreal.ca

July 29, 2011

Abstract

We propose and examine a probabilistic model for the multivariate distribution of the number of calls in each period of the day (e.g., 15 or 30 minutes) in a call center, where the marginal distribution of the number of calls in any given period is arbitrary, and the dependence between the periods is modeled via a normal copula. Conditional on the number of calls in a period, their arrival times are independent and uniformly distributed over the period. This type of model has the advantage of being simple and reasonably flexible, and can match the correlations between the arrival counts in different periods much better than previously proposed models. For the situation where the number of periods is large, so the number of correlations to estimate can be excessive, we propose simple parametric forms for the correlations, defined as functions of the time lag between the periods. We test our proposed models on three data sets taken from real-life call centers and compare their goodness of fit to the best previously-proposed methods that we know. In the three cases, the new models provide a much better match of the correlations and of the coefficients of variation of the arrival counts in individual periods.

KEYWORDS: call center; arrival process; Poisson process; copula model; correlation; simulation.

1 Introduction

Telephone call centers are a central component in several public and private organizations (Gans et al., 2003; Akşin et al., 2007). They employ approximately 3% of the workforce in North America, so their economic importance is unquestionable. A key aspect of their management is the staffing and scheduling of agents (who answer the calls), which is made difficult in particular by the high uncertainty associated with the volume and time-distribution of external call arrivals. Large call centers are complicated nonstationary stochastic systems and discrete-event simulation is the only truly reliable tool to evaluate their performance and to eventually optimize the staffing, the work schedules of agents, and the routing rules of different call types to the different classes of agents, for example (Gans et al., 2003; Brown et al., 2005; Cezik and L'Ecuyer, 2008; Avramidis et al., 2010). Realistic modeling of call arrival processes is an important ingredient for building valid simulation models of call centers as well as for short-term forecasting methods, which in turn are key tools to support decision making for the management of these centers.

Call arrivals certainly do not follow a stationary Poisson process. The next model that may come to mind is a non-homogeneous Poisson process (NHPP), with a time-varying rate, where the rate may depend on the time of the day and the day of the week, for example. Such a model implies that the call volume over any given time period has a Poisson distribution, which implies in turn that the variance and the mean of that volume should be approximately equal. But data collected from call centers disagree with this property: the variance is typically several times the mean, which means that there is more uncertainty than for a Poisson process (Jongbloed and Koole, 2001; Avramidis et al., 2004; Brown et al., 2005; Steckley et al., 2006, 2009). Another important feature observed in empirical studies is the positive dependence between arrival volumes in disjoint time periods of the same day (Tanir and Booth, 1999; Whitt, 1999; Avramidis et al., 2004; Shen and Huang, 2008). This is illustrated in Figure 1, which shows a scatter plot of the number of calls received after 10 A.M. vs the number received before 10 A.M., using different colors for the different days of the week, in one of our data sets. Neglecting this arrival rate randomness and positive dependence can have a huge impact on the long-term performance measures as well as on its probability distribution on a given day (Deslauriers et al., 2007; Steckley et al., 2009).

Models that account for these two properties have been proposed and studied by Whitt (1999) and Avramidis et al. (2004), for example. Whitt (1999) proposed a doubly stochastic Poisson process (DSPP) model where a time-varying arrival rate function is multiplied by a random *business factor* W of mean 1, for each day. Avramidis et al. (2004) studied this model for the case where W has a gamma distribution, showed that the vector of arrival counts in

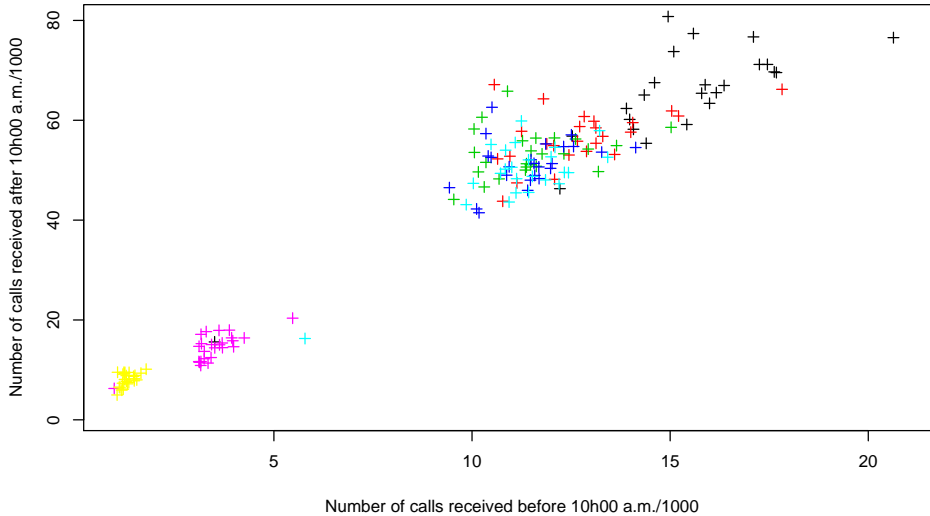


Figure 1: Scatter plot of arrivals before and after 10:00 A.M. in one call center studied.

disjoint time periods then has a negative multinomial distribution, and provided expressions for the maximum likelihood parameter estimators. One weakness of this model is that it tends to give too much correlation across periods of the same day, and that correlation does not depend on the distance between periods, whereas in real life it tends to decrease with that distance (Avramidis et al., 2004). Another drawback is that the variance (and coefficient of variation) of the number of calls in a given time period depends of the distribution of W , which depends (in the model) on what happens in all the periods. As a result, this DSPP model, which we call Model 1 of Avramidis et al. (2004), does not provide a good match for the variance (and marginal distribution) of the number of calls in each period, as we will show in our examples. Jongbloed and Koole (2001) proposed a similar idea, but using independent business factors W across the different time periods of the day, instead of a single factor. This fails to account for the correlation across periods.

Avramidis et al. (2004) introduced two additional models that involve the Dirichlet distribution. The day is partitioned into d time periods of equal length, and the arrival rate is assumed constant in each time period, as is usually done in practice (typically with periods of 15 or 30 minutes). If $\mathbf{X} = (X_1, \dots, X_d)$ is the vector of arrival counts in those periods, the X_j arrivals in period j occur at independent and uniformly distributed time points in that period. In the first model, a dummy variable X_{d+1} is added and the vector $\mathbf{X} = (X_1, \dots, X_d, X_{d+1})$ is assumed to have a negative multivariate distribution whose vector of parameters is random with a Dirichlet distribution. In the second model, the proportion of calls falling in the different time periods obeys a Dirichlet distribution, and their total

number is an independent random variable from another distribution. In their empirical study, these authors take this second distribution as gamma. With these two models, the correlations between arrival counts across periods, for the examples they examine, are closer to the real-life ones than for the model with a single busyness factor, but they are still somewhat too high. Thus, Models 2 and 3 of Avramidis et al. (2004) do not seem to provide enough flexibility to match the correlations well enough.

In this paper, we propose and study an alternative model for the discrete multivariate distribution of \mathbf{X} , where the dependence structure is specified by a normal copula, and the marginal distributions of the X_j 's are specified separately. In some circles, this is known as the NORATA (NORmal To Anything) method. For good general references on copula models, see Joe (1997) and Nelsen (1999). In our implementation, the marginals X_j have a negative binomial distribution, and the arrival times conditional on X_j are independent and uniformly distributed over the time period. This corresponds to a Poisson arrival process whose rate over each time period has a gamma distribution and is constant over the period.

Two important advantages of the normal copula are its simplicity and reasonable flexibility: It is specified simply by specifying a valid correlation matrix. Then the corresponding vector of dependent uniforms is easy to generate, by generating a vector of standard normals (with mean 0 and variance 1) having the given correlation matrix, and applying the probability integral transformation. The correlation matrix is selected so that the rank correlations between the X_j 's matches those observed in the data as closely as possible. However, in d dimensions, the correlation matrix has $d(d - 1)/2$ degrees of freedom (or parameters to estimate). For real-life call centers, d is often as large as 50, in which case we have more than 1000 different correlations (or parameters) to estimate. This may be too many. To reduce the number of parameters in the model (and avoid excessive overfitting), we also restrict the correlation matrices to certain parametric classes, whose choice is a matter of compromise between a better fit (or more flexibility) and fewer parameters. In our experiments, this choice is made based on the Akaike information criterion (Akaike, 1973).

We apply our proposed method to three data sets taken from real-life call centers, two from Bell Canada and one from a major U.S. commercial bank, and compare the goodness-of-fit to that of what we think are the best previously-proposed methods, studied in Avramidis et al. (2004). In all the test cases, our new method provides a much better match of the correlations and of the coefficients of variation (CV) of the arrival counts in individual periods.

The remainder of the paper is organized as follows. In the next section we define the proposed method, explain how to find a correlation matrix for the normal copula to best match a target (linear or rank) correlation matrix for \mathbf{X} , and propose classes of correlation

matrices having a specific structure and a limited number of parameters. In Section 3, we compare the performance of our proposed method to that of other methods on three data sets. A conclusion follows in Section 4. Much of the material of this paper is adapted from Channouf (2008), which also contains other variants and results.

2 A NORTA model for the arrival process

In most call centers, the day is partitioned in d periods of equal length, and the statistics are typically collected only in aggregated form, as averages and counts per period. In particular, statistics on call arrival are usually in the form of number of arrivals in each period.

Let $t_0 < t_1 < \dots < t_d$ be the end points of the time periods, so t_0 and t_d are the opening and closing times of the call center, and period i corresponds to the time interval $[t_{i-1}, t_i)$. For simplicity we take $t_0 = 0$. Let $X(t)$ be the number of calls received in the time interval $[0, t)$, so $X_i = X(t_i) - X(t_{i-1})$ is the number of arrivals in period i , for $i = 1, \dots, d$. The vector of arrival counts during the day is $\mathbf{X} = (X_1, \dots, X_d)$, as defined earlier. For $1 \leq i, j \leq d$, let $\rho_{i,j}^X$ and $r_{i,j}^X$ be the product moment (or Pearson) correlation and the rank correlation (or Spearman coefficient) between X_i and X_j , respectively, and let \mathbf{R}_ρ^X and \mathbf{R}_r^X be the corresponding correlation matrices.

The normal copula (or NORTA) method operates as follows. We first specify a valid correlation matrix \mathbf{R}_ρ^Z , with elements $\rho_{i,j}^Z$, and a marginal distribution function F_i for each X_i . To generate \mathbf{X} , we start by generating a multivariate normal vector $\mathbf{Z} = (Z_1, \dots, Z_d)$ with mean 0 and covariance matrix \mathbf{R}_ρ^Z (so each Z_i has mean 0 and variance 1). We then compute

$$U_i = \Phi(Z_i) \quad \text{and} \quad X_i = F_i^{-1}(U_i), \quad (1)$$

for $i = 1, \dots, d$, where Φ is the standard normal cumulative distribution function. Thus, $\mathbf{U} = (U_1, \dots, U_d)$ and \mathbf{X} are vectors of dependent random variables, and each U_i is uniformly distributed over $(0, 1)$. A standard way of generating \mathbf{Z} is to decompose its correlation matrix as $\mathbf{R}_\rho^Z = \mathbf{A}\mathbf{A}^t$ (where t means ‘‘transposed’’), generate a vector $\mathbf{W} = (W_1, \dots, W_d)$ of independent standard normal variables, and put $\mathbf{Z} = \mathbf{W}\mathbf{A}$ (here we use row vectors).

The marginal distributions F_i can be estimated from the data, and the correlation matrix \mathbf{R}_ρ^Z selected in a way that the resulting correlation matrix \mathbf{R}_ρ^X , or rank correlation matrix \mathbf{R}_r^X , approximates the corresponding matrix from the data. We discuss this in the next section. The rank correlations are generally preferred to the product moment correlations because this yields a simpler and more stable method (Hörmann et al., 2004; Avramidis et al., 2009).

2.1 Fitting the copula

To find a matrix \mathbf{R}_ρ^Z that corresponds to a target matrix \mathbf{R}_r^X for given marginal distributions F_i , we first compute the correlation $\rho_{i,j}^Z$ that yields $r_{i,j}^X$ for each of the $d(d-1)/2$ pairs (i, j) with $i < j$. The other elements can be determined by $\rho_{i,j}^Z = \rho_{j,i}^Z$ and $\rho_{i,i}^Z = 1$. Each $\rho_{i,j}^Z$ is computed as the root of the equation

$$\begin{aligned} r_{i,j}^X &= \text{Corr}(F_i(X_i), F_j(X_j)) = \text{Corr}(F_i(F_i^{-1}(\Phi(Z_i))), F_j(F_j^{-1}(\Phi(Z_j)))) \\ &= \frac{E[F_i(F_i^{-1}(\Phi(Z_i)))F_j(F_j^{-1}(\Phi(Z_j)))] - \mu_{F_i}\mu_{F_j}}{\sigma_{F_i}\sigma_{F_j}} = \frac{g_r(\rho_{i,j}^Z) - \mu_{F_i}\mu_{F_j}}{\sigma_{F_i}\sigma_{F_j}}, \end{aligned} \quad (2)$$

where

$$g_r(\rho_{i,j}^Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i(F_i^{-1}(\Phi(z_i)))F_j(F_j^{-1}(\Phi(z_j)))\phi(z_i, z_j, \rho_{i,j}^Z)dz_idz_j, \quad (3)$$

μ_{F_i} and $\sigma_{F_i}^2$ are the mean and variance of $F_i(X_i)$, $F^{-1}(u) = \inf\{x : F(x) \geq u\}$ for $0 \leq u \leq 1$, and ϕ is the density of the bivariate normal distribution with correlation $\rho_{i,j}^Z$, where the marginals have mean 0 and variance 1.

In the case where both X_i and X_j have a continuous distribution, the mean and variance are $1/2$ and $1/12$, and the above integral equation has the analytic solution (Li and Hammond, 1975) $\rho_{i,j}^Z = 2 \sin(\pi r_{i,j}^X/6)$. But when the marginal distributions are discrete, as is the case here, there is no such analytic solution and we need a root-finding algorithm for a function whose evaluation involves an integral over the (infinite) two-dimensional real space. This is studied by Avramidis et al. (2009), who develop and compare numerical algorithms for this particular application. For our numerical experiments we have adopted one of their algorithms, specifically algorithm NI3 based on the Newton-type method.

Similarly, for the product-moment correlation, we have

$$\begin{aligned} \rho_{i,j}^X &= \text{Corr}(X_i, X_j) = \text{Corr}(F_i^{-1}(\Phi(Z_i)), F_j^{-1}(\Phi(Z_j))) \\ &= \frac{E[F_i^{-1}(\Phi(Z_i))F_j^{-1}(\Phi(Z_j))] - \mu_i\mu_j}{\sigma_i\sigma_j} = \frac{g_l(\rho_{i,j}^Z) - \mu_i\mu_j}{\sigma_i\sigma_j}, \end{aligned} \quad (4)$$

where

$$g_l(\rho_{i,j}^Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i^{-1}(\Phi(z_i))F_j^{-1}(\Phi(z_j))\phi(z_i, z_j, \rho_{i,j}^Z)dz_idz_j, \quad (5)$$

and μ_i and σ_i^2 are the mean and variance of X_i .

Recall that a valid correlation matrix must be symmetric and nonnegative definite, with all its elements in the interval $[-1, 1]$, and the diagonal elements must be 1. After finding all correlations $\rho_{i,j}^Z$ by solving one of two above equations, it often turns out that the resulting

correlation matrix \mathbf{R}_ρ^Z fails to be nonnegative definite. In fact, Ghosh and Henderson (2003) have shown empirically that for a random correlation matrix \mathbf{R}_ρ^X generated uniformly over the set of all valid correlation matrices for uniform marginals F_j , the probability that \mathbf{R}_ρ^Z is nonnegative definite converges to 0 very quickly with the dimension d , and is almost 0 for $d > 15$. This problem can be resolved by modifying \mathbf{R}_ρ^Z slightly to make it nonnegative definite whenever it is not. Various procedures for doing this have been proposed in the literature (Davenport and Iman, 1982; Lurie and Goldberg, 1998; Ghosh and Henderson, 2002). Most are based on the idea of defining a measure of distance between matrices, and computing a nonnegative definite matrix whose distance to the target \mathbf{R}_ρ^Z is as small as possible. This can be formulated and solved as a nonlinear optimization problem with constraints, and the solution is typically quite close to the target. For our numerical experiments, we used the algorithm of Davenport and Iman (1982). This algorithm converges to a nonnegative definite matrix very close to the target in just a few iterations, provided that the number of negative eigenvalues is relatively small compared to the dimension of the matrix and to the number of nonnegative eigenvalues, and that those negative eigenvalues and some of the nonnegative eigenvalues are close to zero. In all our numerical examples, these conditions were verified, and in each case we easily obtained a positive semi-definite matrix that was close to the original (target) matrix. As a typical illustration, in one of our examples, we had a correlation matrix of 48 dimensions, with 3 negative eigenvalues out of 48, all larger than -0.01 , and 4 positive eigenvalues smaller than 0.01 .

2.2 Parameterizing the correlation matrix

There are $d(d-1)/2$ different elements to estimate in the matrix \mathbf{R}_ρ^Z . We refer to the NORTA model with all these $d(d-1)/2$ correlation estimated separately as the *full NORTA model*, or Model N0.

In the context of call centers, the dimension d is typically large; for example, around 50 for 15-minute time periods. With $d = 50$, we have $d(d-1)/2 = 1225$ correlations to estimate in the full NORTA model, quite a large number, which could lead to overfitting unless we have data for a very large number of days.

In what follows, we propose and compare more parsimonious parameterizations of the rank correlation matrix \mathbf{R}_ρ^X . The aim is to reduce the number of quantities (or parameters) that need to be estimated. Brigo (2002) discusses several such parameterizations in the context of interest rate modeling in finance. Some of them reduce the rank of the correlation matrix and other preserve the rank. The parameterizations considered here differ from those in Brigo (2002) and they all preserve the rank.

Let $\boldsymbol{\theta}$ denote the vector of parameters and $r_{i,j}^X(\boldsymbol{\theta})$ the rank correlation as a function of $\boldsymbol{\theta}$.

We will select $\boldsymbol{\theta}$ by least-squares fitting, that is, select $\boldsymbol{\theta}$ that minimizes

$$\min_{\boldsymbol{\theta}} \sum_{i=1}^d \sum_{j=1}^{i-1} (r_{i,j}^X - r_{i,j}^X(\boldsymbol{\theta}))^2. \quad (6)$$

In the full NORTA model, $\boldsymbol{\theta}$ is the vector of all $d(d-1)/2$ correlations and $r_{i,j}^X(\boldsymbol{\theta}) = r_{i,j}^X$.

A natural idea is to model the rank correlations $r_{i,j}^X$ as functions of the difference $|i-j|$ between the time intervals. We expect them to decrease with that difference. A first simple model that captures this property is obtained by taking one distinct parameter for each value of the difference $|j-i|$. We have:

- *Reduced NORTA Model RN1*: $\boldsymbol{\theta} = (p_1, \dots, p_{d-1})$ and

$$r_{i,j}^X(\boldsymbol{\theta}) = p_{|j-i|}, \quad \text{for } j \neq i. \quad (7)$$

This gives $d-1$ parameters to estimate instead of $d(d-1)/2$. Note that there is fewer data to estimate p_k for the values of k close to $d-1$. For $k = d-1$, for example, there is only one observation per day to estimate the correlation of lag k , and this could be problematic if there is only a small number of days.

We also considered a variety of more parsimonious models in which $p_{|j-i|}$ was a parameterized function of $|j-i|$ with less than $d-1$ parameters. We tried polynomial, rational, and exponential (or geometric) functions, and used the curve fitting toolbox *cftool* in MATLAB to find the best fit in each case. These three types of functions gave very comparable fits. We have retained the exponential function, which has only two parameters. This gives:

- *Reduced NORTA Model RN2* (exponential): $\boldsymbol{\theta} = (a, b)$ and

$$r_{i,j}^X(\boldsymbol{\theta}) = a^{|j-i|} + b, \quad \text{for } j \neq i. \quad (8)$$

With this model, the parameterized correlation matrix has the same rank as the original one and is always nonnegative definite.

To compare the different models, we use the *Akaike information criterion* (AIC) (Akaike, 1973), which takes into account both the standard error and the number of parameters estimated in the model. Under the assumption that the model has the form

$$r_{i,j}^X = r_{i,j}^X(\boldsymbol{\theta}) + \epsilon_{i,j}, \quad j \neq i,$$

where the residuals $\epsilon_{i,j}$ are independent and normally distributed with mean 0 and variance σ_ϵ^2 , the AIC is defined by

$$\text{AIC}(s) = n \ln(\hat{\sigma}_\epsilon^2) + 2s, \quad (9)$$

where n is the number of observations (the number of days multiplied by the number of time periods of the day), s is the number of parameters estimated in the model, and $\hat{\sigma}_\epsilon^2$ is the empirical variance of the residuals when the parameters are estimated by the maximum likelihood method. Note that for the full NORTA model, the residues are all zero, so technically the AIC is $-\infty$ in this case and cannot be used for comparison.

3 Numerical examples

In this section, we take data sets from three call centers, fit our proposed models to this data, and compare the goodness-of-fit to that of previously proposed models. In this data, different days of the week have different statistical behavior, so we only considered days having a similar behavior and we removed the data from special days such as Christmas and January 1.

To compare the ability of the different models to match correlations observed in the data, after estimating the model parameters, we computed

$$\rho_m = \text{Corr} \left(\sum_{i=1}^m X_i, \sum_{i=m+1}^d X_i \right), \quad (10)$$

the correlation between the volume of arrivals in the first m periods and that in the remaining $d - m$ periods, for $m = 1, \dots, d - 1$, as in Avramidis et al. (2004). We also computed and compared the coefficient of variation (CV) for each time interval i :

$$\text{CV}(X_i) = \frac{\sqrt{\text{Var}(X_i)}}{\mathbb{E}[X_i]}. \quad (11)$$

With the proposed copula model, these CVs can match the sample values in the data as closely as we want if the marginals are well chosen, but for Model 1 of Avramidis et al. (2004), where an arrival rate function is multiplied by a single random busyness factor for the whole day, all the $\text{CV}(X_i)$'s are determined by the variance of this busyness factor and the mean $\mathbb{E}[X_i]$ over each period, so they cannot be matched as well. We denote by $\hat{\rho}_m$ and $\hat{\text{CV}}(X_i)$ the empirical counterparts of ρ_m and $\text{CV}(X_i)$ observed in the data.

Following Avramidis et al. (2004), the marginal distribution of the number of calls in each time period was assumed to be negative binomial. For each data set, we started by estimating the parameters of these negative binomial marginals, then we computed the rank correlation matrices, fitted the proposed full and reduced NORTA models, and compared the AICs of the various models. We emphasize that with our reduced NORTA Models RN1 and RN2, the correlation matrix \mathbf{R}_ρ^Z is always nonnegative definite. This has the advantage of simplifying the fitting procedure. For the first two call centers, we compare our proposed

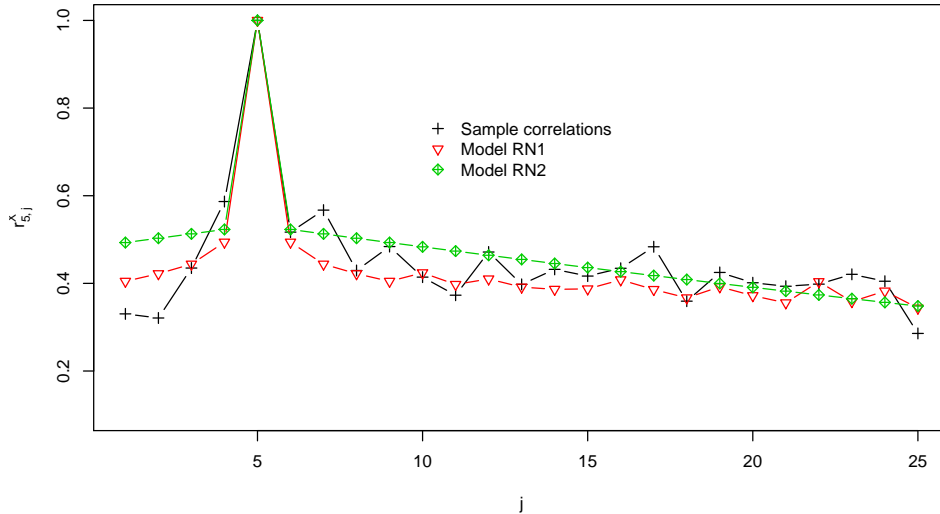


Figure 2: Rank correlations $r_{5,j}^X$ between arrival counts in period 5 and in periods $j = 1, \dots, 25$, for Models RN1 and RN2, for the call center C1. For the full NORTA model, these correlations equal the sample values in the data, shown in the figure.

models to Models 1 to 3 of Avramidis et al. (2004), whereas for the third call center we compare with their Model 1 (we do not have estimations for their other two models).

3.1 C1: A first call center from Bell Canada

Our first data set comes from a call center from Bell Canada that was handling both inbound and outbound calls. This same data set was analyzed and used by Deslauriers (2003) and Avramidis et al. (2004). The data is for 253 days, from May 2001 to March 2002. The arrival counts are aggregated by half hour: we only have the number of incoming calls in each half-hour period (and not the arrival times or call types). The center operates from 8:00 A.M. to 8:30 P.M., for a total of $d = 25$ time periods per day. Based on a preliminary data analysis for this center, it was concluded in Deslauriers (2003) that Tuesday, Wednesday, and Thursday can be considered as similar days, with similar arrival process distributions, so we regrouped those days to obtain a total sample size of 120 days.

As it turns out, the correlations produced by the full and reduced NORTA models do not differ by much. This is illustrated in Figure 2, which gives the rank correlation between X_5 and each other X_i , for the data and Model N0 (those are equal) and for the reduced NORTA Models RN1 and RN2. For Model RN2, the estimated values of (a, b) are $(0.990 \pm 0.002, -0.467 \pm 0.019)$. The trend for Model RN2 in Figure 2 looks linear because a is very close to 1.

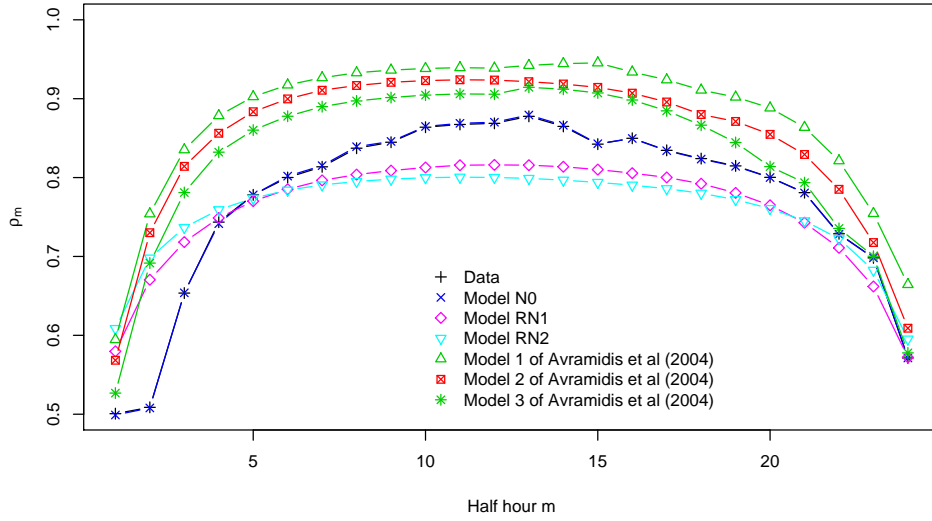


Figure 3: Comparison of models by the correlation ρ_m defined in (10), for the center C1. The values for Model N0 are identical to the empirical values in the data.

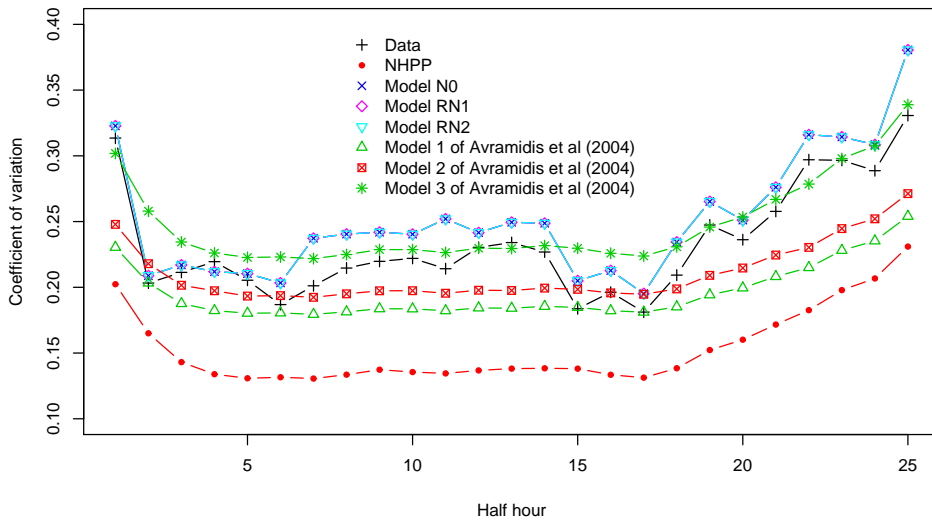


Figure 4: Comparison of models by their values of $CV(X_i)$, in call center C1. For the three NORTA models, the values are identical.

Figures 3 and 4 show the values of ρ_m and $CV(X_i)$, for $m = 1, \dots, 24$, and $i = 1, \dots, 25$, for Models N0, RN1, RN2, and the three models previously proposed in Avramidis et al. (2004) for the same data set. The values of ρ_m and $CV(X_i)$ in the NORTA models were computed exactly, after estimating the parameters. They are compared with the sample values $\hat{\rho}_m$ and $\hat{CV}(X_i)$ in the data. Note that for the full NORTA model N0, we have $\rho_m = \hat{\rho}_m$ for $m = 1, \dots, 24$, and that all NORTA models have identical values of the CVs, because they use the same marginal distributions, independently from the copula. We can observe that the three NORTA models provide a better fit to the empirical correlations and CVs than the other (previous) models, and the full NORTA model fits them more closely than the reduced ones, although it estimates many more parameters. Note that there is more relative variability early and late in the day. Models 1 to 3 of Avramidis et al. (2004) overestimate the correlations and Models 1 and 2 underestimate the variances of the X_i 's. A simple NHPP model also largely underestimates these variances.

For this data set, the AIC took the values -14697.0 and -14355.9 for Models RN1 and RN2, with corresponding standard errors $\hat{\sigma}_\epsilon$ (the standard deviation of residues) of 0.0859 and 0.0913 , respectively. Model RN1 has the smallest AIC between those two.

3.2 C2: A second call center from Bell Canada

Our second set of data concerns a large call center open from 8:00 A.M. to 8:00 P.M. Monday to Friday. We have data on incoming calls for 72 days of type Wednesday, Thursday, and Friday, for the period between January and June, 2005. In this data the observations are aggregated by 15-minute intervals. This gives $d = 48$ time periods.

Figure 5 gives the rank correlation between X_{10} and each other X_i , for the Models N0, RN1, and RN2. These correlations decrease faster with the lag than in the previous example (Figure 2) and they are quite small for large lags. Note that for Models RN1 and RN2, the parameters a and b are not optimized to match only the correlations shown in this figure, but also the rank correlations $r_{i,j}^X$ for other values of i . Figures 6 and 7 show the values of ρ_m and $CV(X_i)$, for $m = 1, \dots, 47$, and $i = 1, \dots, 48$, for various models and for the data. The full NORTA model matches the correlations perfectly, while the reduced NORTA models do not match them as well but nevertheless much better than Models 1 to 3 of Avramidis et al. (2004), which largely overestimate the correlations. Figure 7 indicates that there is more variability early in the day. Again the three NORTA models give identical CVs, whose values are very close to the empirical ones, much closer than those of Models 1 to 3 of Avramidis et al. (2004).

The AIC values for models RN1 and RN2 are -9417.9 and -9416.3 , with standard errors $\hat{\sigma}_\epsilon$ of 0.2526 and 0.2559 . These simplified models are a compromise and not perfect.

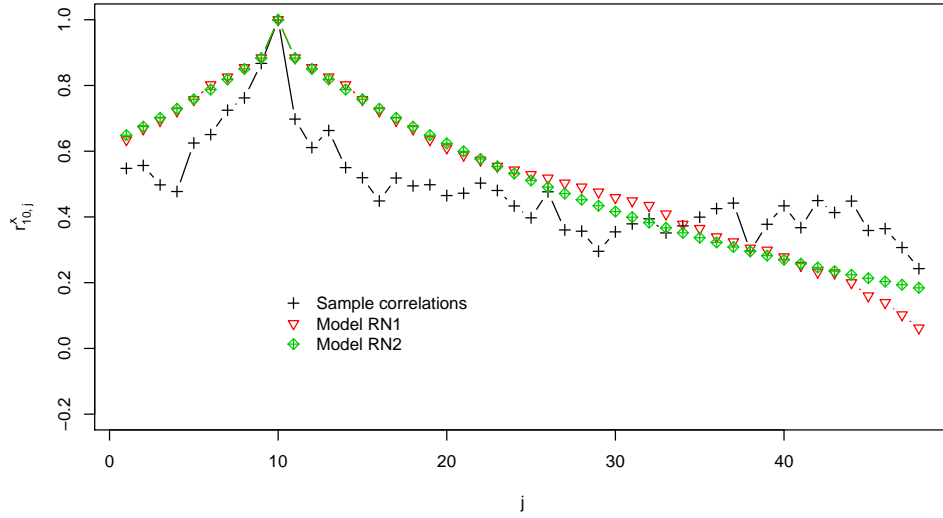


Figure 5: Rank correlations $r_{10,j}^X$ between arrival counts in period 10 and in periods $j = 1, \dots, 48$, for Models RN1 and RN2, for the call center C2. For the full NORTA model, these correlations equal the sample values in the data, shown in the figure.

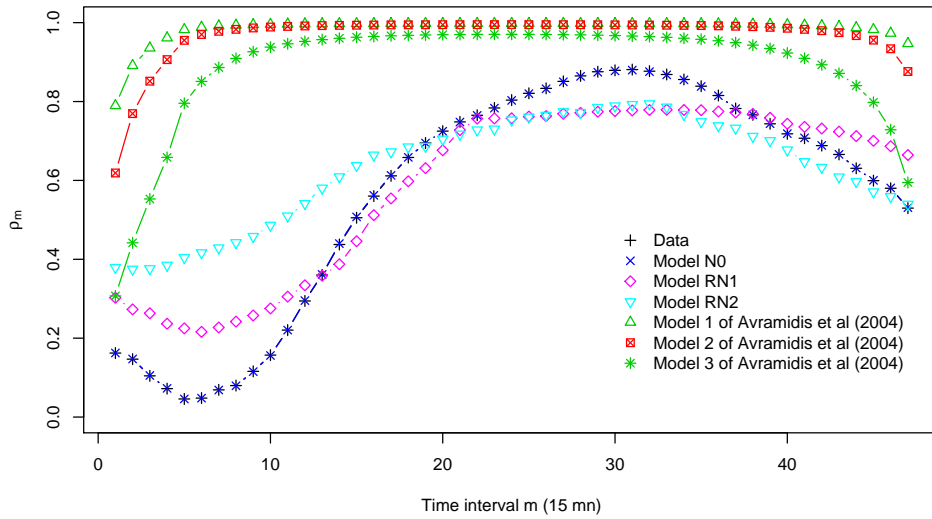


Figure 6: Comparison of models in terms of the correlation ρ_m defined in (10), in center C2. The values for Model N0 are the same as for the data.

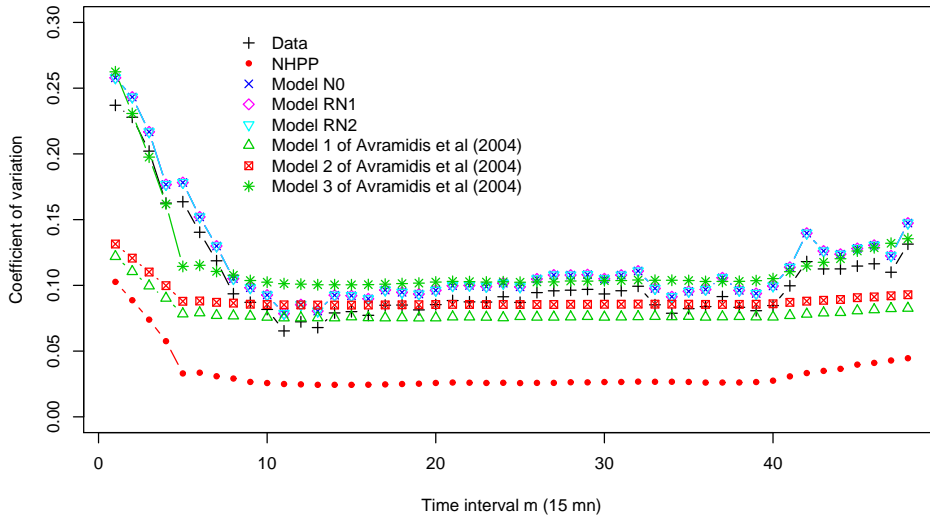


Figure 7: Comparison of models in terms of the $CV(X_i)$'s, in call center C2. The three NORTA models give identical values.

To illustrate this, we show in Figure 8 the rank correlations for pairs having the same differences (2 and 5) $r_{i,i+2}^X$, $i = 1, \dots, 46$, and $r_{i,i+5}^X$, $i = 1, \dots, 43$. We can observe that the correlations of the same lag are about the same, except for correlations in the time periods between 9 and 11 A.M., which corresponds to the peak period of the morning. This is also less true for lag 2 than for lag 5.

3.3 C3: A call center of an American bank

Our third example is a large call center of a commercial bank where calls come from customers in the states of New York, Pennsylvania, Rhode Island and Massachusetts (Trofimov et al., 2004). The center receives up to 300,000 calls per day, soliciting banking services. The data covers a period from March 2001 to April 2003. For our statistical analysis, we consider only the days of type Wednesday, Thursday and Friday, for a total of 143 days. Those days have approximately the same type of arrival patterns. Each day is split into $d = 34$ periods of 30 minutes each, from 7:00 A.M. to midnight.

Figure 9 gives the rank correlation between X_{14} and each other X_i , for the three NORTA models, while Figures 10 and 11 show the values of ρ_m and $CV(X_i)$, for $m = 1, \dots, 33$, and $i = 1, \dots, 34$, for all the models and for the data. The full NORTA model matches the correlations exactly, while the reduced NORTA models do not match them as well but perform much better than Model 1 of Avramidis et al. (2004). Here, we have more relative

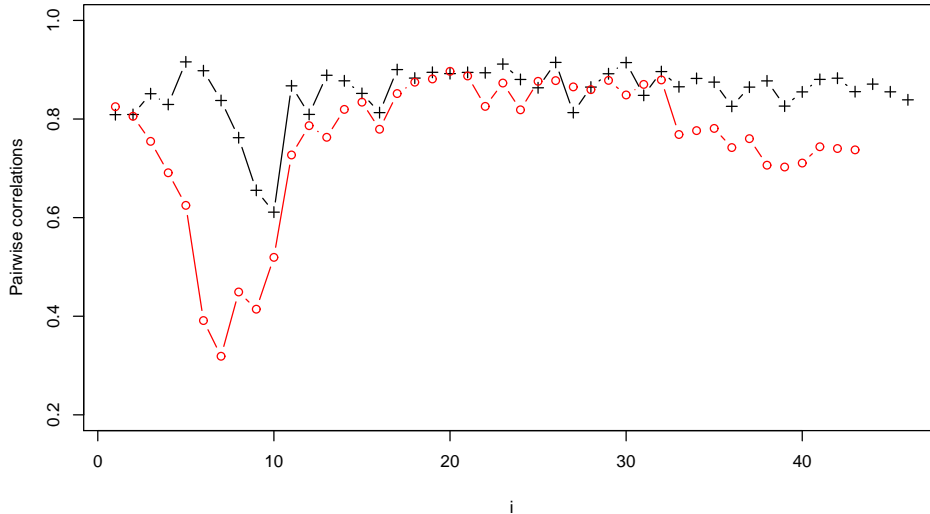


Figure 8: Rank correlations for pairs having the same differences (2 and 5) in call center C2. The symbols “+” are for $r_{i,i+2}^X$ and the symbols “o” are for $r_{i,i+5}^X$.

variability early in the day and also in the last four or five hours in the evening. As in the previous examples, the three NORTA models give identical CV values, and they match the CVs of the data much better than Model 1 Avramidis et al. (2004).

The AIC values for Models RN1 and RN2 are -15626.0 and -15391.1 , with standard errors $\hat{\sigma}_\epsilon$ of 0.1991 and 0.2053, respectively.

4 Conclusion

We developed an arrival process model for a call center, based on a copula model for the vector of arrival counts in the different periods in a given day. We showed empirically that such a model provides a better fit to the correlations between time periods and to the CVs within each period than previously proposed models. An important challenge for future work is the development of extensions for the common situation where several types of calls must be distinguished, and their arrival processes are not independent, in addition to being nonstationary with stochastic rates. Such models are important for simulation (for performance evaluation and optimization) as well as for the construction and real-time update of distributional call-volume forecasts.

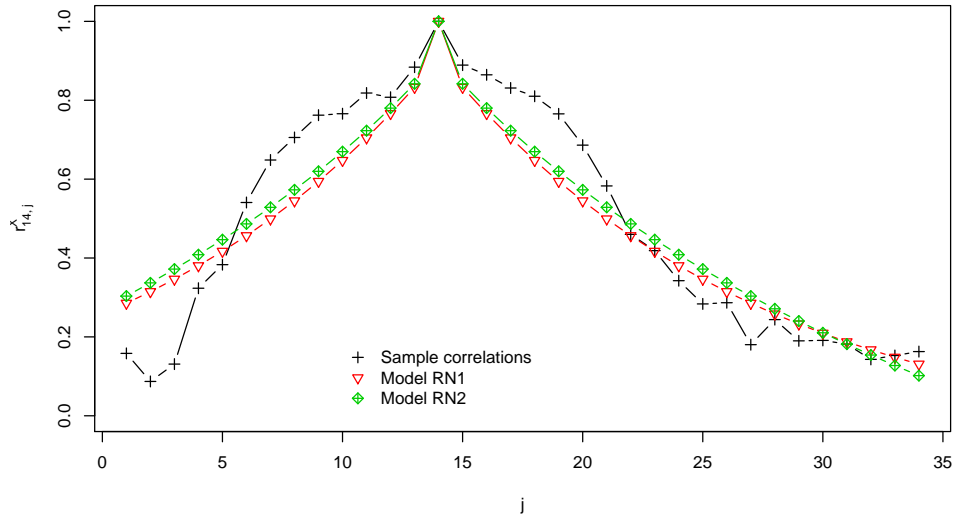


Figure 9: Rank correlations $r_{14,j}^X$ between arrival counts in period 14 and periods $j = 1, \dots, 34$, for Models N0, NR1, NR2, and the data, for call center C3.

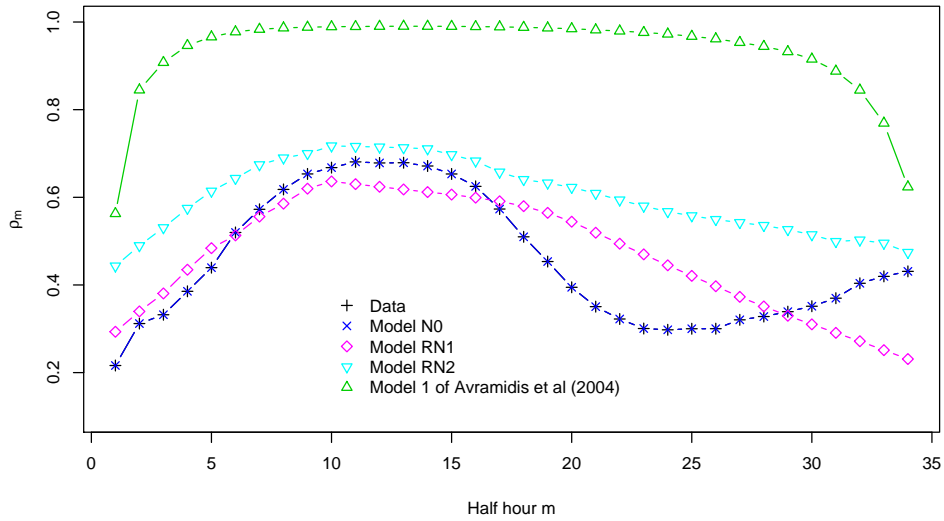


Figure 10: Comparison of models in terms of the correlations ρ_m defined in (10), for call center C3.

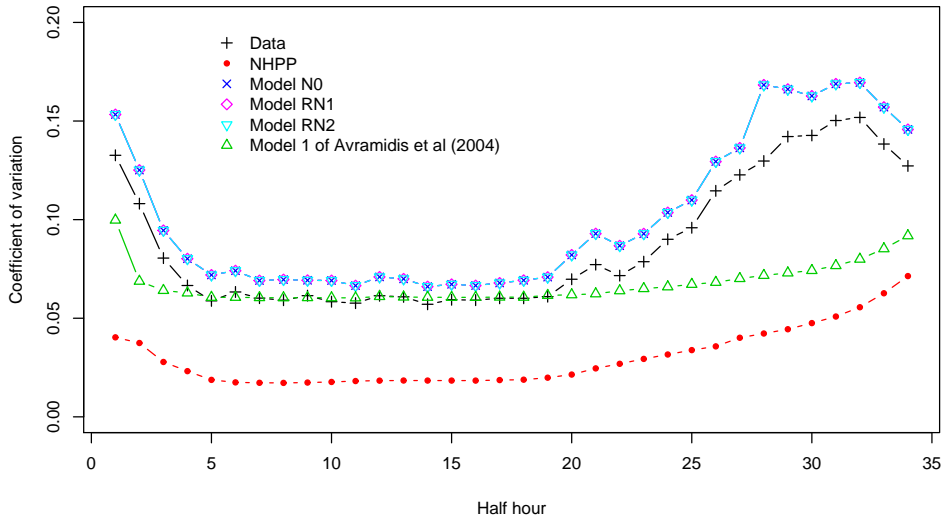


Figure 11: Comparison of models in terms of the $CV(X_i)$'s, for call center C3. The three NORTA models give identical values.

Acknowledgments

This research has been supported by Grants OGP-0110050 and CRDPJ-320308 from NSERC-Canada, a grant from Bell Canada via the Bell University Laboratories, and a Canada Research Chair, to the second author. We are grateful to Avi Mandelbaum for providing us the data for call center C3.

References

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proc. 2nd International Symposium on Information Theory*, (Eds. B. N. Petrov and F. Csaki), *Akademiai Kiado, Budapest*, pages 267–281.
- Akşin, O. Z., Armony, M., and Mehrotra, V. (2007). The modern call center: A multidisciplinary perspective on operations management research. *Production and Operations Management*, 16(6):665–688.
- Avramidis, A. N., Chan, W., Gendreau, M., L’Ecuyer, P., and Pisacane, O. (2010). Optimizing daily agent scheduling in a multiskill call centers. *European Journal of Operational Research*, 200(3):822–832. <http://dx.doi.org/10.1016/j.ejor.2009.01.042>.
- Avramidis, A. N., Channouf, N., and L’Ecuyer, P. (2009). Efficient correlation matching

- for fitting discrete multivariate distributions with arbitrary marginals and normal copula dependence. *INFORMS Journal of Computing*, 21:88–106.
- Avramidis, A. N., Deslauriers, A., and L’Ecuyer, P. (2004). Modeling daily arrivals to a telephone call center. *Management Science*, 50(7):896–908.
- Brigo, D. (2002). A note on correlation and rank reduction. Technical report, Product and Business Development Group, Banca IMI, San Paolo-IMI Group, Milan, Italy.
- Brown, L., Gans, N., Mandelbaum, A., Sakov, A., Shen, H., Zeltyn, S., and Zhao, L. (2005). Statistical analysis of a telephone call center: A queueing-science perspective. *Journal of the American Statistical Association*, 100:36–50.
- Cezik, M. T. and L’Ecuyer, P. (2008). Staffing multiskill call centers via linear programming and simulation. *Management Science*, 54(2):310–323.
- Channouf, N. (2008). *Modélisation et optimisation d’un centre d’appels téléphoniques: étude du processus d’arrivée*. PhD thesis, Département d’Informatique et de Recherche Opérationnelle, Université de Montréal, Canada.
- Davenport, J. M. and Iman, R. L. (1982). An iterative algorithm to produce a positive definite correlation matrix from an approximate correlation matrix. Technical report, Sandia National Laboratories, Albuquerque, New Mexico.
- Deslauriers, A. (2003). Modélisation et simulation d’un centre d’appels téléphoniques dans un environnement mixte. Master’s thesis, Department of Computer Science and Operations Research, University of Montreal, Montreal, Canada.
- Deslauriers, A., Pichitlamken, J., L’Ecuyer, P., Ingolfsson, A., and Avramidis, A. N. (2007). Markov chain models of a telephone call center with call blending. *Computers and Operations Research*, 34(6):1616–1645.
- Gans, N., Koole, G., and Mandelbaum, A. (2003). Telephone call centers: Tutorial, review, and research prospects. *Manufacturing and Service Operations Management*, 5:79–141.
- Ghosh, S. and Henderson, S. (2003). Behaviour of the NORTA method for correlated random vector generation as the dimension increases. *ACM Transactions on Modeling and Computer Simulation*, 13:276–294.
- Ghosh, S. and Henderson, S. G. (2002). Properties of the NORTA method in higher dimensions. In Yücesan, E., Chen, C.-H., Snowdon, J. L., and Charnes, J. M., editors, *Proceedings of the 2002 Winter Simulation Conference*, pages 263–269. IEEE Press.

- Hörmann, W., Leydold, J., and Derflinger, G. (2004). *Automatic Nonuniform Random Variate Generation*. Springer-Verlag, Berlin.
- Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.
- Jongbloed, G. and Koole, G. (2001). Managing uncertainty in call centers using Poisson mixtures. *Applied Stochastic Models in Business and Industry*, 17:307–318.
- Li, S. T. and Hammond, J. L. (1975). Generation of pseudorandom numbers with specified univariate distribution and correlation coefficients. *IEEE Transactions on Systems, Man, and Cybernetics*, 5:557–561.
- Lurie, P. M. and Goldberg, M. S. (1998). An approximate method for sampling correlated random variables from partially-specified distributions. *Management Science*, 44:203–218.
- Nelsen, R. B. (1999). *An Introduction to Copulas*, volume 139 of *Lecture Notes in Statistics*. Springer-Verlag, New York, NY.
- Shen, H. and Huang, J. Z. (2008). Interday forecasting and intraday updating of call center arrivals. *Manufacturing and Service Operations Management*, 10(3):391–410.
- Steckley, S. G., Henderson, S. G., and Mehrotra, V. (2006). Performance measures for service systems with a random arrival rate. In Kuhl, M. E., Steiger, N. M., Armstrong, F. B., and Joines, J. A., editors, *Proceedings of the 2006 Winter Simulation Conference*, pages 2180–2187. IEEE Press.
- Steckley, S. G., Henderson, S. G., and Mehrotra, V. (2009). Forecast errors in service systems. *Probability in the Engineering and Informational Sciences*, 23(2):305–332.
- Tanir, O. and Booth, R. J. (1999). Call center simulation in Bell Canada. In Farrington, P. A., Nemhard, H. B., Sturrock, D. T., and Evans, G. W., editors, *Proceedings of the 1999 Winter Simulation Conference*, pages 1640–1647, Piscataway, New Jersey. IEEE Press. Available on line via www.informs-cs.org.
- Trofimov, V., Feigin, P. D., Mandelbaum, A., Ishay, E., and Nadjharov, E. (2004). Data model for call center analysis. volume 1: Model description and introduction to user interface. Technical report, Technion Israel Institute of Technology, Israel.
- Whitt, W. (1999). Dynamic staffing in a telephone call center aiming to immediately answer all calls. *Operations Research Letters*, 24:205–212.