

ORIGINAL RESEARCH

Evaluating the risk of unsatisfied random demand on a time interval

Michael Todinov*

Department of Mechanical Engineering and Mathematical Sciences, Oxford Brookes University, Wheatley, Oxford, UK

Received: August 23, 2015

Accepted: October 20, 2015

Online Published: November 3, 2015

DOI: 10.5430/air.v5n1p67

URL: <http://dx.doi.org/10.5430/air.v5n1p67>

ABSTRACT

This paper focuses on an important and very common problem and presents a theoretical framework for solving it: “determining the risk of unsatisfied request from users placing random demands on a time interval”. For the common case of a single source servicing a number of consumers, a closed-form solution has been derived for the risk of collision of random demands. Based on the closed-form solution, an efficient optimisation method has been developed for determining the optimal number of consumers that can be serviced by a single source, such that the probability of unsatisfied demand remains below a maximal tolerable level. A central part of the proposed theoretical framework is a general equation evaluating the risk of unsatisfied demand by the expected fraction of time of unsatisfied demand. The derived equation covers multiple sources servicing multiple consumers. Finally, the conducted parametric studies revealed an unexpected finding: the risk of collision of random demands on a time interval is practically insensitive to the standard deviations of the durations of demands. This surprising result provides the valuable opportunity to work with random demand times characterised by their means only, without supplying their probability distributions or variances.

Key Words: Probability, Risk, Unsatisfied demand, Sources, Consumers, Random demand, Optimisation

1. INTRODUCTION

Risk is often linked with the collision of demands from consumers on a finite time interval for a particular resource, at random times and for specified durations. Suppose that the source of the supplied resource can only service a single consumer at a time. Unsatisfied demand occurs if one or more demands arrive at a time during which the source is engaged in servicing another demand.

The need for assessing this risk is often present in many manufacturing processes where a number of machine centres demand expensive measuring equipment, control equipment, production equipment or an operator, at a random time during the production process. Because the control equipment

is expensive and unique, it is usually not feasible to equip each machine centre with a separate piece of equipment.

The demands for a resource may occur at random times during a shift. If the start of demand for consumer i is denoted by “ s_i ” and the end of demand for consumer i is denoted by “ e_i ”, Figure 1 illustrates this problem by three machine centres demanding X-ray portable measurement equipment^[1] for measuring the residual stresses at the surface of quenched components, at random times s_1 , s_2 and s_3 for duration intervals (s_1, e_1) , (s_2, e_2) and (s_3, e_3) .

If a single piece of measuring equipment is available, a simultaneous demand from more than a single centre cannot be

*Correspondence: Michael Todinov; Email: mtodinov@brookes.ac.uk; Address: Department of Mechanical Engineering and Mathematical Sciences, Oxford Brookes University, Wheatley, Oxford OX33 1HX, UK.

satisfied. The overlapping region (s_3, e_2) in Figure 1 marks unsatisfied demand on the supplied resource.

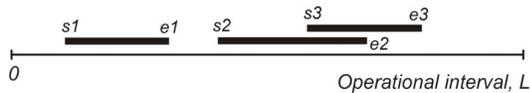


Figure 1. Constraint on the supplied resource for three machines

Here are other examples of the risk associated with collision of random demands.

- Injured or critically ill patients demanding randomly a piece of life-saving medical equipment. The equipment can be used by a single patient at a time.
- Emergency calls for a nurse from critically ill patients in a hospital ward. The nurse can service a single patient at a time (There will be no unsatisfied demands if the times between the emergency calls are greater than the maximum time needed for attending a call).
- A number of machines, each of which places a random demand during a given time interval, for the maintenance services of a single available operator.
- Spare equipment in a warehouse servicing random demands during a specified time interval. After a demand from a customer, the warehouse needs a minimum time to replenish the dispatched equipment before the next demand can be satisfied. In this case, the risk of unsatisfied demand materialises if one or more customers arrive within the critical time interval needed for replenishing the equipment and making it available for the next customer.

In queuing theory, the Poisson process has been traditionally used as a statistical model for random events occurring in a time interval.^[2-6] Neither of these classical comprehensive texts nor more recent comprehensive texts devoted to various problems in probability and queuing theory^[7-12] treats the question related to risk of unsatisfied demand from random requests on a time interval.

From Figure 1, it appears that the problem of unsatisfied demand can be reduced to a problem of geometrical probability where a segment of specified length L is covered by randomly located smaller segments with different lengths. The probability of unsatisfied demand can then be estimated by the probability of an overlap by two or more segments. The expected time of unsatisfied demand is numerically equal to the expected fraction of overlapped area from two or more than two segments.

There are already a number of publications related to covering the circumference of a circle with segments or a linear

segment with segments.^[13-19]

Stevens^[14] derived closed-form expressions for the probability of covering a circle by a specified number of arcs of the same length with random locations and for the probability of existence of a specified number of uncovered gaps. Randomly positioned arcs with different lengths have been considered in Ref.15, where a condition has been derived for covering the circle with probability equal to one.

A segment covered by randomly located smaller segments has been considered in Ref.18, where the probability of existence of a segment which intersects every other segment has been estimated. An asymptotic estimate of the largest number of pairwise non-intersecting intervals among n random intervals has also been provided.

Despite the progress made in problems related to coverage of a large segment with smaller segments with random location, no study seems to exist related to estimating the expected lineal fraction covered by m or more random segments and to the probability of an overlap of segments with different size. The answer to these questions however, is the key to evaluating the risk of unsatisfied demand.

An important aspect of the problem related to risk associated with unsatisfied demand on a time interval is the maximum tolerable level of the probability of unsatisfied demand. The required level of the probability of unsatisfied demand depends on the magnitude of the consequences resulting from unsatisfied demand and must be set individually by the risk experts in the respective application area. Thus, for injured or critically ill patients demanding life-saving medical equipment, the consequences of unsatisfied demand are grave. In this case, unsatisfied demand means human fatalities and the maximum tolerable level of the probability of unsatisfied demand is very low.

In the case where the monitoring services of a single available operator are needed for the successful operation of a number of machines, each of which places random demands during a given time interval, the consequences of unsatisfied demand are significant and the maximum tolerable level of the probability of unsatisfied demand is also low.

For machine centres demanding for example measuring equipment to control the surface roughness of machined work-pieces, the consequences of unsatisfied demand are moderate and the corresponding tolerable level of the probability of unsatisfied demand is moderate.

For machine centres demanding a piece of control equipment measuring the wear-out of the bearings of rotating shafts, the consequences of unsatisfied demand are low because failure

is not imminent. Accordingly, the tolerable probability of unsatisfied demand is high. Questions of significant practical importance, directly related to the problem of unsatisfied random demand, are:

- (1) What is the probability that there will be time during the specified operation interval when the available resource will be simultaneously demanded by more than one consumer?
- (2) What is the maximum number of consumers that can be serviced by a single source so that the probability of unsatisfied demand remains below the maximum acceptable level?
- (3) What is the expected fraction of time during which the $m > 1$ sources will fail to satisfy simultaneous overlapping demands from $m + 1$ or more consumers?

The purpose of this study is to provide answers to these questions.

2. RISK OF COLLISION OF RANDOM DEMANDS ON A TIME INTERVAL

Suppose that n consumers start their demands at random times s_1, s_2, \dots, s_n , for durations d_1, \dots, d_n , during a time interval $0, L(d_1 + d_2 + \dots + d_n < L)$ (see Figure 2a).

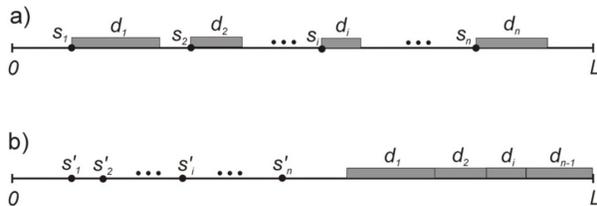


Figure 2. Random demands in a finite time interval $(0, L)$

Suppose that the times of random demands are uniformly distributed along the length of the time interval $(0, L)$. Let A_1, A_2, \dots, A_n denote the events “the last demand has a duration d_1, d_2, \dots, d_n ”, correspondingly. The probability of the event B that there will be no unsatisfied demand can be determined by the following probabilistic argument.

Initially, the conditional probability $P(B|A_n)$ will be determined the probability that there will be no unsatisfied demand, given that the last demand has a duration d_n . Because every consumer has an equal chance to be the last consumer, the probabilities $P(A_i)$ of the events A_i are all equal to $1/n$ ($p(A_1) = 1/n$).

The probability that there will be no overlapping random demands from n users will be derived by using a transform technique. Configurations X where no overlapping demands

are present will be transformed to simpler configurations Y whose probability is easier to determine. Because of the bijection (one-to-one correspondence) between X -configurations and Y -configurations, determining the probability of existence of a configuration Y will measure the probability of a configuration X .

Suppose that an X configuration is present. In this case, there are no overlapping demand times (unsatisfied demand) Figure 2a. The demand intervals can then be “cut out” of the time interval $0, L$ and the remaining parts can be brought together to form a shorter length $L - (d_1 + d_2 + \dots + d_{n-1})$ (see Figure 2b).

As a result of this this operation, the points s_1, s_2, \dots, s_n marking the start of the random demands for configurations X in Figure 2a will transform into a configuration Y for which the points s'_1, s'_2, \dots, s'_n are uniformly distributed along the length $L - (d_1 + d_2 + \dots + d_{n-1})$. In other words, for each X -configuration, there is a corresponding Y -configuration characterised by n uniformly distributed points s'_1, s'_2, \dots, s'_n along the length $L - (d_1 + d_2 + \dots + d_{n-1})$.

Now suppose that a set of n random points s'_1, s'_2, \dots, s'_n are randomly generated along the length L . If all n random points fall within the length $L - (d_1 + d_2 + \dots + d_{n-1})$ (if configuration Y is present), by inserting the corresponding demand intervals $d_i (i = 1, \dots, n - 1)$ after each start of demand s_i , an X -configuration will be obtained with n random non-overlapping demands along the interval $0, L$. In words, for each Y -configuration there is a corresponding X -configuration.

Because of the one-to-one correspondence between a configuration X and configuration Y , the probability of X -configuration can be measured by the probability of the Y -configuration. The probability of Y -configuration is the probability that n uniformly distributed random points s'_1, s'_2, \dots, s'_n on the interval $(0, L)$ will fall in the interval $L - (d_1 + d_2 + \dots + d_{n-1})$ and is given by $\left(\frac{L - (d_1 + d_2 + \dots + d_{n-1})}{L}\right)^n$. This is also the conditional probability that there will be no unsatisfied demand given that the last demand has a duration d_n .

$$P(B|A_n) = \left(1 - \frac{(d_1 + \dots + d_{n-1})}{L}\right)^n \tag{1}$$

The absence of unsatisfied demand (event B) however, can occur in n different ways. The absence of unsatisfied demand can occur given that the demand of length d_n is the last demand, given that the demand of length d_{n-1} is the last demand and so on.

According to the total probability theorem,

$$P(B) = P(B|A_1)P(A_1) + \dots + P(B|A_n)P(A_n) \tag{2}$$

The probabilities $P(B|A_i)$ ($i = 1, \dots, n-1$) are determined in a similar fashion. As a result, the expression

$$P(B) = \frac{1}{n} \left[\left(1 - \frac{d_2 + d_3 + \dots + d_n}{L}\right)^n + \left(1 - \frac{d_1 + d_3 + \dots + d_n}{L}\right)^n + \dots + \left(1 - \frac{d_1 + d_2 + \dots + d_{n-1}}{L}\right)^n \right] \tag{3}$$

is obtained for the probability that there will be no unsatisfied random demand on the interval $(0, L)$. If $D = \sum_1^n d_i$ denotes the sum of durations of all demands, the probability of unsatisfied demand $P(\bar{B}) = 1 - P(B)$ can be obtained as a probability of a complementary event:

$$P(\bar{B}) = 1 - \frac{1}{n} \left[\left(1 - \frac{D - d_1}{L}\right)^n + \left(1 - \frac{D - d_2}{L}\right)^n + \dots + \left(1 - \frac{D - d_n}{L}\right)^n \right] \tag{4}$$

This equation has been confirmed by the results from a simulation algorithm. Thus, for four consumers demanding a particular resource for $d_1 = 5$ min, $d_2 = 10$ min, $d_3 = 20$ min, and $d_4 = 35$ min respectively, during a time interval of 10 hours, the probability of unsatisfied demand calculated from equation (4) is 0.3. This probability has been confirmed by the probability of 0.3 calculated from the simulation.

An important special case is obtained if the durations of all random demands are equal. Substituting $D - 1 = d_2 = \dots = d_n = d$ in equation (4) then yields

$$P(\bar{B}) = 1 - \left(1 - \frac{D - d}{L}\right)^n \tag{5}$$

The analysis of equation (5) reveals a useful result. Suppose that the total length of demand $D = n \times d$ is kept constant and only the number of customers n and the durations d of their demands are varied in such a way that $D = n \times d$ does not change. In other words, $d = D/n$. Equation (5) then becomes

$$P(\bar{B}) = 1 - [1 - (1 - 1/n) \times (D/L)]^n \tag{6}$$

With increasing n , $P(\bar{B})$ tends to unity.

Indeed,

$$\begin{aligned} [1 - (1 - 1/n) \times (D/L)]^n &= \exp\{n \ln((1 - 1/n) \times (D/L))\} \\ \lim_{n \rightarrow \infty} [1 - (1 - 1/n) \times (D/L)]^n &= \\ \lim_{n \rightarrow \infty} [\exp\{n \ln((1 - 1/n) \times (D/L))\}] &= \exp(-\infty) = 0 \end{aligned}$$

Therefore, $\lim_{n \rightarrow \infty} [P(\bar{B})] = 1$.

For $D = 50$ min total demand and time interval $L=600$ min, the dependence presented by equation (6) has been given in Figure 3.

As can be seen from the graph in Figure 3, if the supplied resource is finite and sufficient for a total duration of supply of D hours (e.g. compressed gas in bottles, chemicals, etc.), a strategy involving splitting the total resource among a fewer number of consumers is better than a strategy based on splitting the resource among a larger number of consumers. The first strategy is characterised by a significantly smaller probability of unsatisfied demand.

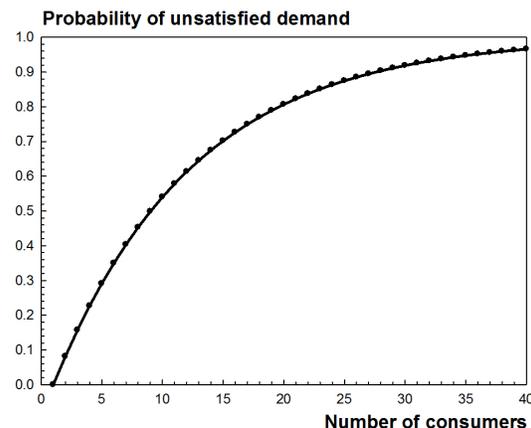


Figure 3. Probability of unsatisfied demand as a function of the number of consumers. The total duration of the demand from all consumers is the same - 50 min.

3. INSENSITIVITY OF THE RISK OF COLLISION OF RANDOM DEMANDS ON THE VARIANCE OF THE DEMAND TIME

Simulation experiments have been conducted involving a constant (fixed) number of random demands on a specified time interval. Each consumer places exactly one demand, randomly located in the specified time interval (see Figure 2a):

Increasing the standard deviation of the duration of random demands reveals a rather unexpected trend (see Figure 4). The simulation results clearly show that *the probability of unsatisfied demand practically does not vary with varying the standard deviation of the demand time*. In the experiments presented in Figure 4, the durations of the random demands follow a normal distribution with a specified mean (80 min,

60 min, 50 min and 40 min) and a standard deviation varying in the interval 0-9 minutes. The results in Figure 4 feature a single source and 4 consumers.

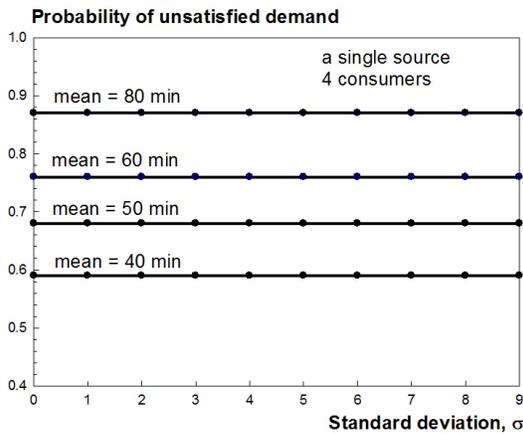


Figure 4. Dependence of the risk of collision of random demands on the standard deviation of the durations of the demand times for a single available source capable of servicing a single consumer at a time. The durations of the demands follow a normal distribution with a specified mean, which is kept constant.

To check whether these results are caused by the symmetry of the Gaussian distribution, the simulation experiments were repeated with an asymmetrical log-normal distribution with mean $\mu = 180$ min and standard deviation varying in the range (0, 36 minutes) and duration of the operation time interval of 300 hours. All 15 consumers were characterised by the same mean demand time of 180 minutes (which was kept constant) and the same standard deviation. The common standard deviation characterising the demand time of each consumer was varied in the interval 0, 36 minutes. The trend was the same; the probability of unsatisfied demand practically did not vary with varying the standard deviation of the consumers' demand time (see Figure 5).

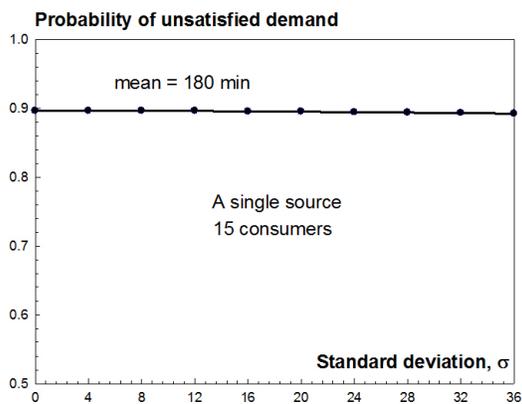


Figure 5. Dependence of the risk of collision of random demands on the standard deviation of the durations of the demands. The durations of random demands follow a log-normal distribution.

These results can be rationalised by using the analytical expression (equation 1) regarding the probability that there will be no unsatisfied demand from n random consumers on a time interval $(0, L)$, given by:

$$p_n^0 = [1 - (x_1 + x_2 + \dots + x_{n-1}) / L]^n$$

where x_i ($i = 1, \dots, n - 1$) are the durations of the random demands from the consumers. The duration of the demand x_n from the last consumer has been discarded, because it cannot possibly contribute to the probability of unsatisfied demand.

If the durations x_1, x_2, \dots, x_{n-1} of the random demands are now realisations of a random variable X following a statistical distribution with mean and standard deviation σ , even for a relatively small number of consumers, the sum of the durations $\sum_{i=1}^{n-1} x_i$ in the above equation can be approximated reasonably well with $\sum_{i=1}^{n-1} x_i \approx (n - 1)\mu$ and, as a result, the probability p_n^0 that there will be no unsatisfied demand will be practically insensitive to the variance of the random variable X standing for the durations of the demand times. The probability p_n^0 will depend only on the expected value $\mu = E(X)$ of the demand times. Because the probability of unsatisfied demand p_n is given by $p_n = 1 - p_n^0$, the probability of unsatisfied demand will be practically insensitive to the variance (standard deviation σ) of the random demand times X .

These results clearly show that reducing the variances of the demand times practically has no impact on the probability of unsatisfied demand. This rather unexpected result provides the valuable opportunity to work with random demand times characterised by their means only without requiring information related to the variance of the demand times.

4. DETERMINING THE OPTIMAL NUMBER OF USERS SERVED BY A SINGLE SOURCE

Consider a finite time interval during which a number of consumers demand a particular service independently and randomly. Suppose that the source needs a minimum time interval to recover and stabilise after a demand from a consumer. Collision among user demands occurs if a random demand arrives while the source is servicing another user.

Decreasing the expected number density of the users arriving randomly during a time interval is an efficient way of decreasing the probability of a collision between the random

demands. Such a decrease however may come at a cost. The number of serviced users may be decreased unnecessarily.

Consider now random demands with density λ (number of requests per unit time) following a homogeneous Poisson process on the time interval $(0, L)$. Assume for the sake of simplicity that a time s is needed for servicing each random demand. The random demands could be requests for using unique equipment; requests for a particular resource (e.g. water vapour, electrical power, compressed air, etc.), requests for a nurse from critically ill patients in a hospital ward, requests for repair because of failed equipment, etc. A basic property of the homogeneous Poisson process, well documented in books on probabilistic modelling^[5,20,21] states: Given that n random requests following a homogeneous Poisson process are present in the finite interval $0, L$, the times of the random requests are distributed uniformly over the time interval $0, L$.

On the time interval $(0, L)$ there can be zero random demands (event A_0), exactly one random demand (event A_1), \dots , exactly k random demands (event A_k), and so on. Events $A_0, A_1, \dots, A_k, \dots$, are mutually exclusive events. The probability of exactly k random demands is given by the Poisson distribution:

$$P(A_k) = \frac{(\lambda L)^k e^{-\lambda L}}{k!} \tag{7}$$

The probability $p(B)$ that there will be no collision of random demands on the time interval $0, L$ is given by

$$p(B) = \sum_{k=0}^{\infty} \frac{(\lambda L)^k e^{-\lambda L}}{k!} p(B | A_k) \tag{8}$$

where $p(B | A_k)$ is the conditional probability that there will be no collision of random demands given that exactly k random demands occur during the time interval $(0, L)$. $p(B | A_k)$ can be determined considering that if the homogeneous Poisson process is conditioned on the number of random demands, the random demands will be uniformly distributed along the time interval $(0, L)$.

The maximum possible number of demands for which time gaps of size at least equal to d between each adjacent pair of demands can still be present is given by $r = (L/d) + 1$, where (L/d) is the greatest integer part of the ratio L/d which does not exceed it. Denote with B the event “there is no unsatisfied demand”, which is equivalent to the event “there are gaps of size at least d between each adjacent pair of random demands”.

According to Equation (1), the conditional probability

$p(B | A_k)$ that there will be no unsatisfied demand given that k random demands are present on the time interval $(0, L)$ is given by

$$P(B | A_k) = \left(1 - \frac{(k-1)d}{L}\right)^k \tag{9}$$

for $k \leq r$, and by $p(B | A_k) = 0$, for $k \geq r + 1$.

The probability of exactly k random demands $k \leq r$ on the time interval $(0, L)$ and no unsatisfied demand is given by

$$P(A_k \cap B) = P(A_k) \times P(B | A_k) = \frac{(\lambda L)^k e^{-\lambda L}}{k!} \times \left(1 - \frac{(k-1)d}{L}\right)^k \tag{10}$$

According to the total probability theorem, the probability that there will be no unsatisfied demand is given by

$$P(B) = P(A_0) \times P(B | A_0) + P(A_1) \times P(B | A_1) + \dots + P(A_r) \times P(B | A_r) \tag{11}$$

which results in

$$P(B) = \sum_{k=0}^r \frac{(\lambda L)^k e^{-\lambda L}}{k!} \times \left(1 - \frac{(k-1)d}{L}\right)^k \tag{12}$$

where $r = [L/d] + 1$.

Expanding the sum in Equation (12) results in

$$p_0 = \exp(-\lambda L) \left(1 + \lambda L + \frac{\lambda^2 (L-d)^2}{2!} + \dots + \frac{\lambda^r [L - (r-1)d]^r}{r!}\right) \tag{13}$$

The probability p_c that two or more random demands will occur within the critical time d is

$$p_c = 1 - \exp(-\lambda L) \left(1 + \lambda L + \frac{\lambda^2 (L-d)^2}{2!} + \dots + \frac{\lambda^r [L - (r-1)d]^r}{r!}\right) \tag{14}$$

This is also the probability that there will be unsatisfied random demand (that two or more random requests will occur within the critical time interval d).

Figure 6 gives the dependence of the probability of unsatisfied demand for a single source and critical time interval $d = 1$ h for servicing a single random demand. The operating time interval is $L = 100$ h. The probability of unsatisfied

demand has been plotted for different values of the number density of the demands.

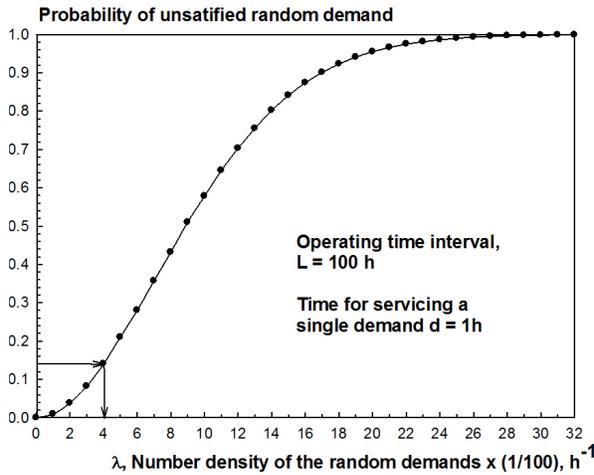


Figure 6. Probability of unsatisfied demand on a finite operational time interval of 100 h. The random demands follow a homogeneous Poisson process and each random demand requires 1 h service time.

For a mean number of 14 demands per 100 h, there is already about 80% probability of unsatisfied demand. Clearly, the probability of unsatisfied demand is substantial and should always be taken into consideration in risk assessments.

The dependence in Figure 6 can be used as a basis of a method for determining the optimal number of users that can be serviced by a single source such that the risk unsatisfied demand remains below a maximal tolerable level.

For a specified maximum acceptable probability of unsatisfied demand p_c^{\max} , for example $p_c^{\max} = 0.14$, the maximum possible number density of demands λ^{\max} is determined by solving:

$$p_c^{\max} = 1 - \exp(-\lambda L) \left(1 + \lambda L + \frac{\lambda^2 (L-d)^2}{2!} + \dots + \frac{\lambda^r [L - (r-1)d]^r}{r!} \right)$$

with respect to λ . This can be done by using the standard repeated bisection method.

Alternatively, for the specified maximum acceptable probability of unsatisfied demand $p_c^{\max} = 0.14$, the maximum possible number density of demands $\lambda^* = 0.04h^{-1}$ can be determined directly from the plotted dependence given by equation (14) (see the arrows in Figure 6). The number density envelope guarantees that whenever the number density of the random demands does not exceed $\lambda^* = 0.04(\lambda \leq \lambda^*)$, the probability of unsatisfied demand will not exceed the

critical level of 14%.

Reducing the probability of collision of random demands below the pre-set level of 14% can be achieved by reducing the number density of the demands, but this will reduce unnecessarily the size of the serviced system and the costs. The optimal level of number density of demands corresponds to the maximum acceptable level (14%) of the risk of unsatisfied demand.

If the average number density of the random demands characterising a single consumer is λ_0 , the total number density of the random demands characterising all n consumers is $\lambda = n \times \lambda_0$. Determining the maximum acceptable demand rate λ^* which guarantees a specified tolerable probability of collision of random demands can be determined from solving equation 14 with respect to λ .

Dividing the maximum acceptable demand rate λ^* to λ_0 yields the maximum acceptable number n^* of consumers that can be serviced by a single source:

$$n^* = \lambda^* / \lambda_0 \tag{15}$$

The purpose of optimising the number density of random demands is not only to minimise the risk of unsatisfied demand below a maximum acceptable level but also to strike an optimal balance between risk and cost. The risk of unsatisfied demand should not be reduced too far below its maximal tolerable level because this incurs extra costs for providing extra sources supplying the particular service in demand.

The described method has wide potential applications in a number of critical situations. For example, if the demands are emergency calls for a nurse arriving from critically ill patients in a hospital, the obtained hazard rate envelope can be used to determine the maximum number of such patients that could be looked after by a single nurse so that the probability of unattended call remains below a maximum acceptable level. The consequences of an unattended patient’s call are grave (an unattended call could result in a human fatality) and to keep the risk low, the specified tolerable probability of unsatisfied demand should be low.

If the average number density of the calls characterising a single critically ill patient is λ_0 , the total number density of the calls characterising all n critically ill patients is $\lambda = n \times \lambda_0$. Determining the maximum acceptable call rate λ^* which guarantees with the specified tolerable probability that there will be no patients’ calls while the nurse is servicing another patient can be determined using the presented method. Dividing the maximum acceptable call rate λ^* to λ_0 yields

the maximum acceptable number n^* of patients that can be looked after by a single nurse (equation 15).

5. ESTIMATING THE RISK OF UNSATISFIED DEMAND BY THE EXPECTED FRACTION OF THE TIME OF UNSATISFIED DEMAND

Suppose that m pieces of a particular resource are available and m simultaneous demands can be satisfied, but not $m + 1$ or more simultaneous demands. If the different demands are represented as overlapping segments with different lengths d_i , the risk of unsatisfied demand can also be estimated with the expected fraction of length covered simultaneously by more than m random segments randomly located along a segment with length L .

If the duration of the demand for the i th consumer is equal to d_i , during the operation period with length L , the ratio of the duration of the demand and the time interval “ L ” will be denoted by $\psi_i = d_i/L$.

Before determining the expected time fraction of unsatisfied demand, the following theorem related to a coverage of space with volume V by n 3-D interpenetrating objects with volumes v_i ($i = 1, \dots, n$), randomly placed in the volume V , will be stated and proved. The volume fractions of the separate objects will be denoted by $\psi_i = v_i/V$. The coverage of a point from the volume V is a “coverage of order k ” if exactly k objects cover the point. The following theorem then holds.

Theorem 1. The expected covered fraction of order k ($k = 0, 1, \dots, n$) from the volume V , by n interpenetrating objects with volume fractions ψ_i , is given by the $k+1$ st term of the expansion $\prod_{i=1}^n [(1 - \psi_i) + \psi_i]$.

Proof. The volume fraction covered by exactly m objects can be determined from the probability that a randomly selected point in the volume V will sample simultaneously m overlapping (interpenetrating) random objects. The probability that a randomly selected point in the volume V will sample simultaneously m overlapping objects is equal to the probability that a fixed point from the volume V will be covered exactly m times by randomly placed objects in the volume V .

Let p_0 denote the probability that the fixed point will not be covered, p_1 denote the probability that the fixed point will be covered by exactly one random object, ..., and p_n denote the probability that the fixed point will be covered by all n random objects.

Because the locations of the random objects are statistically independent events, the probability of the event A_0 that a fixed point in the volume will not be covered by any of the

random objects is given by

$$P(A_0) = \prod_{i=1}^n (1 - \psi_i) \quad (16)$$

which is the probability that the fixed point will not be covered by the first, the second, ..., the n th object.

The probability of the event A_1 that exactly one random object will cover the fixed point, is a sum of the probabilities of the following mutually exclusive events: the first object covers the fixed point and the rest of the random objects do not, the second object covers the fixed point and the rest of the random objects do not and so on. As a result, the probability $P(A_1)$ that the fixed point will be covered by exactly one random object is given by

$$P(A_1) = \sum_{i=1}^n \left(\psi_i \prod_{\substack{k=1 \\ k \neq i}}^n (1 - \psi_k) \right) \quad (17)$$

The probability that exactly two random objects will cover the fixed point is a sum of the probabilities of the following mutually exclusive events: the first and the second random object cover the fixed point and the rest of the objects do not, the first and the third random object cover the fixed point and the rest of the random objects do not and so on, until all possible combination of two objects out of n are exhausted. As a result, the probability that the fixed point will be covered by exactly two random objects is given by

$$P(A_2) = \sum_{i1, i2} \left(\psi_{i1} \psi_{i2} \prod_{\substack{k=1 \\ k \neq i1, k \neq i2}}^n (1 - \psi_k) \right) \quad (18)$$

where $\sum_{i1, i2}$ denotes the sum over all possible combinations of two indices $i1$ and $i2$ out of n . The number of these combinations is $\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2}$.

Continuing this reasoning through the cases 3, 4, ..., n , the probability $P(A_m)$ that the fixed point will be covered by exactly m random objects is given by

$$P(A_m) = \sum_{i1, \dots, im} \left(\psi_{i1} \psi_{i2} \dots \psi_{im} \prod_{\substack{k=1 \\ k \neq i1, \dots, k \neq im}}^n (1 - \psi_k) \right) \quad (19)$$

where $\sum_{i1, \dots, im}$ denotes the sum over all distinct combinations of m indices $i1, i2, \dots, im$ out of n . The number of these combinations is $\binom{n}{m} = \frac{n!}{m!(n-m)!}$.

The fixed point can either remain uncovered, covered by exactly one, two,..., n objects, and there are no other alternatives. Therefore, the events A_0, A_1, \dots, A_n constitute a set of mutually exclusive and exhaustive events. According to the third axiom of the probability theory, their probabilities add up to one:

$$\sum_{i=0}^n P(A_i) = 1 \tag{20}$$

$$\prod_{i=1}^n (1-\psi_i) + \dots + \sum_{i_1, \dots, i_m} \left(\psi_{i_1} \psi_{i_2} \dots \psi_{i_m} \prod_{\substack{k=1 \\ k \neq i_1, \dots, k \neq i_m}}^n (1-\psi_k) \right) + \dots + \prod_{i=1}^n \psi_i = 1 \tag{21}$$

Equation (21) can also be presented as an expansion of the expression $\prod_{i=1}^n [(1 - \psi_i) + \psi_i]$. The theorem has been proved.

Because the proof does not make any reference to the shape of the random objects, the theorem is valid for interpenetrating random objects of any shape. The theorem is also valid in the two-dimensional (2-D) and one-dimensional (1-D) case of area or a segment covered by two-dimensional or one-dimensional objects, correspondingly.

An immediate corollary of the theorem is related to objects with the same volume fraction $\psi = v/V$. In this case, the expected fractions from the volume V covered by the separate random objects with volumes v are given by the separate terms of the binomial expansion of $[(1 - \psi) + \psi]^n$:

$$[(1 - \psi) + \psi]^n = \sum_{i=0}^n \binom{n}{i} \psi^i (1 - \psi)^{n-i} = 1 \tag{22}$$

The expected fraction of the volume covered by exactly m random objects is given by

$$P(A_m) = \binom{n}{m} \psi^m (1 - \psi)^{n-m} \tag{23}$$

Now consider a case where n consumers demand a particular resource, during an operating period with length L . The durations of the demands from consumers are $d_i (i = 1, \dots, n)$. The ratios of the durations of the demands from the separate consumers are given by $\psi_i = d_i/L$. The maximum number of consumers whose demand can be satisfied simultaneously by the sources is m .

Theorem 2. If one source can satisfy only a demand from a

single consumer at a time, the expected fraction of time of unsatisfied demand related to m sources and n consumers is given by the expression

$$1 - \left[\prod_{i=1}^n (1-\psi_i) + \sum_{i=1}^n \left(\psi_i \prod_{\substack{k=1 \\ k \neq i}}^n (1-\psi_k) \right) + \dots + \sum_{i_1, \dots, i_m} \left(\psi_{i_1} \psi_{i_2} \dots \psi_{i_m} \prod_{\substack{k=1 \\ k \neq i_1, \dots, k \neq i_m}}^n (1-\psi_k) \right) \right] \tag{24}$$

Proof. Unsatisfied demand related to m sources and n consumers ($n > m$) is present in the case where more than m consumers require a source simultaneously. Let p_0 denote the probability that a fixed point in the interval $(0, L)$ will not sample any demand, p_1 denote the probability that the fixed point will sample exactly one random demand,..., and p_m denote the probability that the fixed point will sample exactly m random demands.

The probability $p_{\geq m+1}$ that the fixed point will sample more than m random demands, randomly placed in the time interval $(0, L)$, is then given by

$$p_{\geq m+1} = 1 - (p_0 + p_1 + \dots + p_m) \tag{25}$$

According to Theorem 1, the sum of the probabilities $p_0 + p_1 + \dots + p_m$ is given by

$$\prod_{i=1}^n (1-\psi_i) + \dots + \sum_{i_1, \dots, i_m} \left(\psi_{i_1} \psi_{i_2} \dots \psi_{i_m} \prod_{\substack{k=1 \\ k \neq i_1, \dots, k \neq i_m}}^n (1-\psi_k) \right)$$

Hence, the theorem has been proved.

If all random demands are characterised by the same duration d , $\psi = d/L$, the time fraction of unsatisfied demand is given by

$$p_{\geq m+1} = 1 - \sum_{i=0}^m \binom{n}{i} \psi^i (1 - \psi)^{n-i} \tag{26}$$

Note that the sum in the right hand side of equation (26) is part of the binomial expansion of the expression $[(1 - \psi) + \psi]^n$, which is identically equal to unity:

$$[(1 - \psi) + \psi]^n = \sum_{i=0}^n \binom{n}{i} \psi^i (1 - \psi)^{n-i} = 1 \tag{27}$$

In equation (27), $(1 - \psi)^n$ is the expected fraction of time during which no demand is present, $n(1 - \psi)^{n-1}\psi$ is the expected fraction of time during which exactly one random demand is present, $\frac{n(n-1)}{1 \times 2} (1 - \psi)^{n-2}\psi^2$ is the expected frac-

tion of time during which exactly two random demands are present; $\frac{n(n-1)\dots(n-m+1)}{1 \times 2 \times \dots \times m} (1 - \psi)^{n-m} \psi^m$ is the expected fraction of time during which exactly m random demands are present; and finally ψ^n is the expected fraction of time during which exactly n random demands are present.

Here a typical illustrating example related to a particular manufacturing process, during an 8-hour shift, where 12 machines stop once randomly (at any time during the shift) for exactly 50 minutes, to use a particular measurement device. There is only one measurement device available. The time fraction of demand is $\psi = \frac{50/60}{8} = 0.1$; $1 - \psi = 0.9$.

The expected time fraction of no random demand is 0.9^{12} ; the expected time fraction of exactly one random demand is $12 \times 0.9^{11} \times 0.1$; the expected time fraction of exactly two random demands is $\frac{12 \times 11}{1 \times 2} 0.9^{10} \times 0.1^2$.

The expected fraction of unsatisfied demand is $p = 1 - (0.9^{12} + 12 \times 0.9^{11} \times 0.1) = 0.34$. If a second measuring device is introduced, the expected fraction of unsatisfied demand becomes $p = 1 - (0.9^{12} + 12 \times 0.9^{11} \times 0.1 + \frac{12 \times 11}{1 \times 2} 0.9^{10} \times 0.1^2) = 0.11$. By introducing a second measurement device, the expected time of unsatisfied demand has been decreased significantly.

All equations in this section have been verified by a Monte Carlo simulation involving measuring and accumulating directly the multiple intersections. This was achieved by a discrete-event simulator where the variable “cover” increases its value by one, if the current event is a start-consumption event and decreases its value by one if the event is an end-consumption event. The pseudo-code measuring the expected length of demand overlaps of order $m + 1$ is given by Algorithm 1.

The start of resource consumption is a random event simulated by the product Lu_i where L is the length of the operation interval $(0, L)$ and u_i is a uniformly distributed random number in the interval $(0, 1)$.

The duration of the demand is also random and was simulated by sampling a normal distribution with a specified mean μ and standard deviation σ . Algorithms for sampling a normal distribution can for example be found in Refs.21 and 22.

For each user placing a random demand, two types of events are simulated: a “start-consumption” event – marking the start of consumption of the user and “end-consumption” event – marking the end of consumption of the user.

In case of a start-consumption event (see Algorithm 1), if the current content of the variable cover is equal to m , there are already m overlaps and the overlap with the current event

will create $m + 1$ overlaps which means that there will be unsatisfied demand. The s_overlap variable is initialised with the time where the overlap section of order $m + 1$ actually starts (for this section an overlap of more than m random demands is present). The section of more than m overlapping demands will continue until the content of the variable “cover” is reduced back to m (the number of overlapping demands drops back to m). This occurs just after an end-of-consumption event. At this point in time, with the statement $\text{sum} = \text{sum} + (\text{current_time} - \text{s_overlap})$, the size of the section with more than m overlapping demands is calculated and accumulated into the variable “sum”. Thus, for the current simulation history of demands and release of the sources, the lengths of all zones of overlap of order $m + 1$ are accumulated in the variable sum. At the end of the current simulation history, the content of the variable sum is accumulated in the variable global_sum with the statement $\text{global_sum} = \text{global_sum} + \text{sum}$.

Algorithm 1 Determining the expected fraction of time during which unsatisfied demand (more than m demands) exists

```

total_sum=0;
for i=1 to num_simul_trials do
{
  queue_size=0; cover = 0;
  Generate all start-consumption and end-consumption events
  and insert them in the priority queue;

  Insert in the priority queue a notional end-consumption event
  corresponding to the end of the operation interval 0,L;

  sum = 0;
  while (queue_size>0)
  {
    Take event from the priority queue;
    current_time = event.time;
    if (event.type = 'start consumption')
    {
      if (cover = m) {s_overlap = current_time;}
      cover = cover + 1;
      remove the current event from the priority queue;
    }
    else // a case of end-consumption event
    {
      if (cover = m+1) sum = sum + (current_time - s_overlap);
      cover=cover - 1;
      remove the current event from the queue;
    }
  }
  global_sum = global_sum + sum;
}
}
expec_glob_sum = global_sum/num_simul_trials;
expec_frac_uns_demand = expec_glob_sum/L;

```

As a result, at the end of all simulation histories, the variable global_sum contains the sum of the lengths of all sections characterised by more than m overlapping demands. The ratio of the global_sum and the number of simulation tri-

als gives the expected value exp_glob_sum of the total length of the time intervals containing more than m overlapping random demands. The ratio of the expected value exp_glob_sum and the length of the operation interval L gives the expected fraction of unsatisfied demand.

As a result, the risk of unsatisfied demand can be estimated by the expected fraction of unsatisfied demand.

6. CONCLUSIONS

- (1) A theoretical framework based on geometric probability has been created for evaluating the risk of unsatisfied demand arising from random requests on a time interval.
- (2) For the common case of a single source servicing a number of consumers, a closed-form solution has also been derived for the risk of unsatisfied demand.
- (3) Based on the closed-form solution, an efficient optimisation method has been developed for determining the optimal number of consumers that can be serviced by a single source, such that the risk of unsatisfied demand remains below a maximal tolerable level.
- (4) A central part of the proposed theoretical framework is a general equation evaluating the risk of unsatisfied demand by the expected fraction of time of unsatisfied demand. The equation covers multiple sources servicing multiple consumers.
- (5) The conducted parametric studies revealed an unexpected finding: the risk of collision of random demands is insensitive to the standard deviations of the durations of demands. This surprising result provides the valuable opportunity to work with random demand times characterised by their means only without requiring their distributions or variances.
- (6) Given that the available source can satisfy demand for a limited total duration D , a strategy involving splitting the total duration D among fewer number of consumers is preferable to a strategy involving splitting the total duration D among a larger number of consumers. The first strategy is characterised by a significantly smaller risk of unsatisfied demand.

REFERENCES

- [1] Rigaku Corporation. Instruction Manual for Strainflex MSF-2M/PSF-2M; 1994.
- [2] Khintchine AY. Mathematical methods in the theory of Queuing. London: Griffin; 1969.
- [3] Gross D, Harris CM. Fundamentals of queuing theory (2nd ed.). New York: Wiley; 1985.
- [4] Genedenko BV, Kovalenko IN. Introduction to queuing theory, 2nd ed., Birkhauser, Berlin; 1989.
- [5] Thompson WA. Point process models with applications to safety and reliability, Chapman & Hall; 1988.
- [6] Allen AO. Probability, Statistics and Queuing theory with computer science applications, 2nd ed., Academic press Inc; 1990.
- [7] Kalashnikov V. Mathematical methods in queuing theory, Springer-Science+Business Media; 1994.
- [8] Giambene G. Queuing theory and telecommunications, Networks and applications: Springer; 2005.
- [9] Weiss NA. A course in probability, Pearson/Addison Wesley; 2006.
- [10] Venkatesh SS. The theory of probability: explorations and applications, Cambridge University Press; 2013.
- [11] Calin OU. Geometric modelling in probability and statistics: Springer; 2014.
- [12] Stroock DW. Probability theory: An analytical view, 2nd ed., Cambridge University press; 2011.
- [13] Solomon H. Geometric probability, Society for industrial and applied mathematics; 1978.
- [14] Stevens WL. Solution to a geometrical problem in probability. Ann. Eugenics. 1939; 9: 315-20. <http://dx.doi.org/10.1111/j.1469-1809.1939.tb02216.x>
- [15] Shepp LA. Covering the circle with random arcs. Israel J. Math. 1972; 11: 328-45. <http://dx.doi.org/10.1007/BF02789327>
- [16] Coffman EG, Jr L. Flatto P, *et al.* Packing Random Intervals On-Line. Algorithmica. 1998; 22: 448-76.
- [17] Coffman Jr EG, Mallows CL, Poonen B. Parking arcs on the circle with applications to one-dimensional communication networks. The Annals of Applied Probability. 1994; 4(4): 1098-111. <http://dx.doi.org/10.1214/aoap/1177004905>
- [18] Justicz J, Scheinerman ER, Winkler PM. Random intervals. Amer. Math. Monthly. 1990; 97: 881-9.
- [19] Huffer FW, Shepp LA. On the Probability of Covering the Circle by Random Arcs. Journal of Applied Probability. 1987 Jun; 24(2): 422-9. <http://dx.doi.org/10.2307/3214266>
- [20] Ross S. Simulation 2nd edition, Harcourt academic press; 1997.
- [21] David HA. Order statistics (2nd ed.). New York: Wiley; 1981.
- [22] Rubinstein RY. Simulation and the Monte-Carlo method, John Wiley & Sons, New York; 1981.