

The Effect of Model Uncertainty on Maintenance Optimization

Cornel Bunea and Tim Bedford

Abstract—Much operational reliability data available, e.g., in the nuclear industry, is heavily right-censored by preventive maintenance. The common methods for dealing with right-censored data (Total Time on Test statistic, Kaplan–Meier estimator, adjusted rank methods) assume the s -independent competing-risk model for the underlying failure process and the censoring process, even though there are many s -dependent competing-risk models that can also interpret the data. It is not possible to identify the “correct” competing risk model from censored data. A reasonable question is whether this model uncertainty is of practical importance. This paper considers the impact of this model-uncertainty on maintenance optimization, and shows that it can be substantial. Three competing-risk model classes are presented which can be used to model the data, and determine an optimal maintenance policy. Given these models, then consider the error that is made when optimizing costs using the wrong model. Model uncertainty can be expressed in terms of the “dependence between competing risks,” which can be quantified by expert judgment. This enables reformulating the maintenance optimization problem to account for model uncertainty.

Index Terms—Censored data, competing risks, copula, identifiability, preventive maintenance.

ACRONYMS¹

Cdf	cumulative distribution function
i.i.d.	s -independent and identically distributed
pdf	probability density function
PM	preventive maintenance
RC	replacement cost
RT	replacement time
r.v.	random variable
Sf	survivor function

NOTATION

X	lifetime, $X \geq 0$; a r.v.
Y	PM time, $Y \geq 0$; a r.v.
$f_X(x)$, $F_X(x)$	[pdf, Cdf] of X
$f_Y(y)$, $F_Y(y)$	[pdf, Cdf] of Y
$r_X(t)$	failure rate of X

Manuscript received December 7, 2000; revised July 15, 2001. Responsible Editor: J.-C. Lu.

C. Bunea is with the Faculty of Information Systems and Technology, Delft University of Technology, Delft, The Netherlands (e-mail: C.Bunea@its.tudelft.nl).

T. Bedford is with the Department Management Science, Strathclyde University, Glasgow, UK (e-mail: Tim@mansci.strath.ac.uk).

Digital Object Identifier 10.1109/TR.2002.804486

¹The singular and plural of an acronym are always spelled the same.

$S_X(x)$	$1 - F_X(x)$: Sf of X
$S_Y(y)$	$1 - F_Y(y)$: Sf of Y
$S_X^*(x)$	$\Pr\{X > t, X < Y\}$: sub-Sf of X
$S_Y^*(t)$	$\Pr\{Y > t, Y < X\}$: sub-Sf of Y
$F_X^*(x)$	$\Pr\{X < t, X < Y\}$: sub-Cdf of X
$F_Y^*(t)$	$\Pr\{Y < t, Y < X\}$: sub-Cdf of Y
c_1	cost of critical failure
c_2	cost of planned replacement
$C(x, y)$	copula of X and Y
$\rho(x, y)$	Spearman's ρ
$\tau(x, y)$	Kendall's τ
θ	age replacement time
γ	s -expected cost over θ

DEFINITION

copula: see Section III-C

I. INTRODUCTION

THE COMMON methods (assuming s -independent censoring), used to treat right-censored data, are nonconservative, in the sense that other s -dependent censoring models estimate the underlying failure process more pessimistically [4]. Without making nontestable assumptions (e.g., s -independence of the failure and censoring processes), the true distribution function is not identifiable from the data. Hence, in addition to the usual uncertainty caused by sampling fluctuation, there is the extra problem of model uncertainty.

This paper tests the effect of model uncertainty on the problem of optimizing maintenance.

Assumption 1: Data are available which contain censors from an existing PM program.

The data in assumption #1 are used to estimate an optimal age replacement PM program.

Section III presents 3 model-classes of competing risk. The independent model is used as the most extreme pessimistic model of existing PM. The other extreme model is used for the most optimistic model of existing PM. The dependent competing risk model is used for the general case; the dependence between competing risks is given by a copula. The minimally informative copula with respect to the uniform distribution and Archimedean copula are studied—the later is used to approximate the first one, due to numerical difficulties in working with the minimally informative copula for strong dependence between risks. A method is presented by which expert judgment can be used to quantify model uncertainty.

Section IV recalls the theory of optimal age-replacement policies. Section V presents 3 numerical examples to determine the error that is made when optimizing costs using the wrong model.

Section VI shows that model uncertainty does lead to substantial uncertainty in estimating optimal maintenance intervals and excessive costs.

This paper extends and develops results in [6], in particular by showing how expert judgment can be used to quantify model uncertainty.

II. COMPETING RISK

The competing-risk approach models the data as a renewal process: a sequence of i.i.d. variables Z_1, Z_2, \dots . Each observable Z_i is the minimum of 2 variables and the indicator of which variable was smaller. The lifetime of the component is X : the life that the component would reach if it were not PM'ed. The PM time of the component is Y : the time at which the component would be preventively maintained if it did not fail first. Clearly,

$$Z \equiv [\min[X, Y], I(X < Y)].$$

Usually X is the minimum of several variables giving the time to failure by a particular failure mode: this paper considers the case of 1 failure mode. The observable data allow estimating the sub-Sf,

$$\begin{aligned} S_X^*(t) &\equiv \Pr\{X > t, X < Y\}, \\ S_Y^*(t) &\equiv \Pr\{Y > t, Y < X\}, \end{aligned}$$

but not the true Sf of X and Y . Hence one can not estimate the underlying failure distribution for X without making additional, nontestable, model assumptions. A characterization of these distributions for X that are possible for given sub-Sf is in [5].

By specifying a copula for the underlying joint distribution of X and Y one can identify the marginals (and the full joint distribution) [12]. However the choice of such a copula is difficult to make: [2] suggests doing this by specifying the Spearman's rank correlation between X and Y , and then using the copula with minimum information with respect to the independent copula (i.e., the most-independent copula with the given Spearman rank correlation).

III. THREE MODELS FOR COMPETING RISK

This section presents 3 competing-risk models in which the marginal Cdf are identifiable. Two of them are the extreme cases— independent model and high correlated censoring model; the third one assumes that the dependence between competing risks is given by a copula.

A. Model #1: Independence

Let F_X have pdf, $f_X(t)$, then

$$\begin{aligned} r_X(t) &= \frac{f_X(t)}{S_X(t)} = -\frac{1}{S_X(t)} \cdot \frac{dS_X(t)}{dt} \\ d[\log(S_X)] &= \frac{dS_X}{S_X}, \\ S_X(t) &= \exp\left[-\int_0^t r_X(s) ds\right]. \end{aligned}$$

But from competing risk data a different rate of failure for X is observed. The observed failure rate for X is

$$\begin{aligned} obr_X(t) &\equiv \lim_{\delta \rightarrow 0} \left[\frac{\Pr\{X > t, X < Y, X \in (t, t + \delta) | Z > t\}}{\delta} \right] \\ &= -\frac{1}{S_X^*(t) + S_Y^*(t)} \cdot \frac{dS_X^*(t)}{dt}. \end{aligned}$$

For the most frequently made assumption in the literature, probabilistic independence between X and Y , then

$$\begin{aligned} S_X^*(t) + S_Y^*(t) &= \Pr\{X > t, Y > t\} \\ &= \Pr\{X > t\} \cdot \Pr\{Y > t\} = S_X(t) \cdot S_Y(t). \end{aligned}$$

Using these results, [8] shows that if the competing risks X and Y are s -independent with differentiable Sf, then

$$r_X(s) = obr_X(s).$$

Now, the underlying marginal distributions of X and Y can be identified in terms of the observable sub-Sf,

$$S_X(t) = \exp\left[\int_0^t \frac{dS_X^*(s)}{S_X^*(s) + S_Y^*(s)}\right]. \quad (1)$$

B. Model #2: Highly Correlated Censoring

s -Independent censoring does not capture the notion that PM is done when the equipment gives some sign of future failure.

The most extreme case is: PM aims to prevent component-failure at a time immediately before failure. If that aim is not achieved then the PM action is applied immediately after failure. PM is unsuccessful with probability p and successful with probability $1 - p$, s -independent of the time at which the failure occurs. This is modeled by

- $Y = X + \delta \cdot \varepsilon$,
- $\varepsilon > 0$ is very small but depends on X ,

$\delta = \{1, -1\}$, with probability $[p, 1 - p]$, is s -independent of X . For very small ε Model 2 gives:

$$\begin{aligned} S_X^*(t) &\equiv \Pr\{X > t, X < Y\} = \Pr\{X > t, \delta = 1\} \\ &= \Pr\{\delta = 1\} \cdot \Pr\{X > t\} = p \cdot S_X(t); \\ S_Y^*(t) &\equiv \Pr\{Y > t, Y < X\} = \Pr\{\delta = -1\} \cdot \Pr\{Y > t\} \\ &\approx (1 - p) \cdot \Pr\{X > t\} = (1 - p) \cdot S_X(t). \end{aligned}$$

Hence the *normalized* sub-Sf (normalized so that they equal 1 at $t = 0$) are approximately equal,

$$\frac{S_X^*(t)}{p} \approx \frac{S_Y^*(t)}{1 - p}, \quad (2)$$

and both are equal to $S_X(t)$. This condition can be checked from the data. If it does not hold then model #2 is not correct. Fig. 1 is an example where

$$S_X(t) = \exp(-t^{0.5}). \quad (3)$$

1000 samples were taken for this model, with $p = 1/3$; then the empirical functions $(\hat{S}_X^*(t))/p$, $(\hat{S}_Y^*(t))/(1 - p)$, and the theoretical function, $S_X(t)$, were plotted.

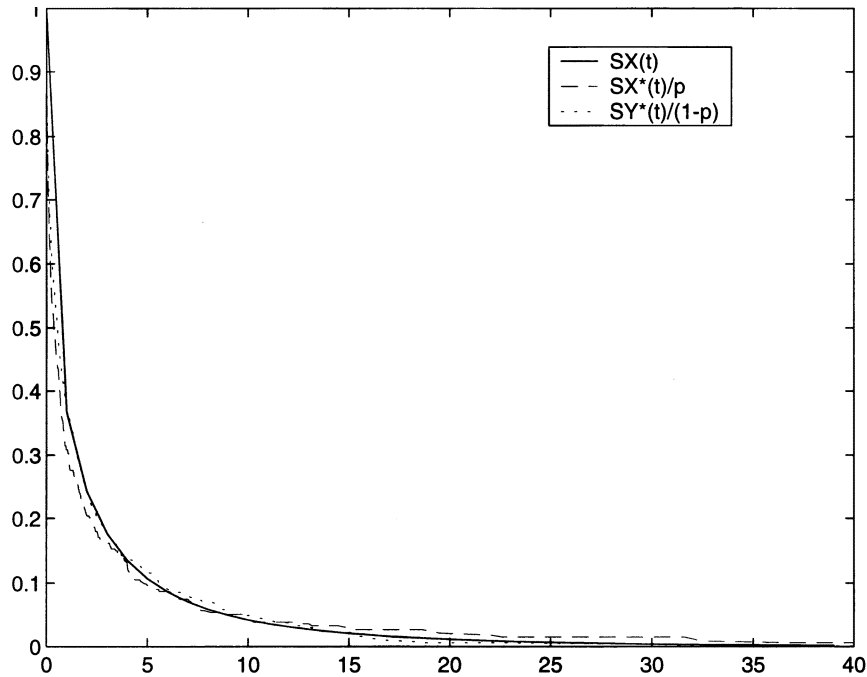


Fig. 1. Highly correlated censoring.

If (2) does hold, then the model might be correct, but the s -independent model might also hold with the same observable data. Assuming model #1 (s -independence) when model #2 holds would lead to an incorrect assessment of the marginals. Proposition 1 is obtained by using (1), [6].

Proposition 1: Let X and Y have a joint distribution described by model #2. Let \tilde{X} and \tilde{Y} be s -independent with

$$S_{\tilde{X}}^*(t) = S_X^*(t), \quad S_{\tilde{Y}}^*(t) = S_Y^*(t).$$

Then

$$S_{\tilde{X}}(t) = [S_X(t)]^p, \quad S_{\tilde{Y}}(t) = [S_Y(t)]^{1-p}.$$

Model #2 is a special case of the random-signs model in [7], and can be used when the sub-Sf satisfy

$$\frac{S_X^*(t)}{p} \geq \frac{S_Y^*(t)}{1-p}, \quad t \geq 0. \quad (4)$$

The random-signs model says that $Y = X - \xi$ (ξ is a r.v.), $\xi \leq X$, $\Pr\{\xi = 0\} = 0$, whose sign is s -independent of X . The failure is observed with probability $\Pr\{X < Y\} = \Pr\{\xi > 0\} = p$.

C. Model #3: Dependent Competing Risks

Definition: Copula: the copula of 2 r.v., X and Y , is the distribution, C , on the unit square $[0, 1]^2$ of the pair $(F_X(X), F_Y(Y))$ (for a continuous r.v., X , with pdf, F_X , the r.v., $F_X(x)$, is always uniformly distributed on $[0, 1]$) [10].

This model assumes that the dependence structure between X and Y is given by a copula. The functional form of $C: [0, 1]^2 \rightarrow \mathbf{R}$ is

$$C(u, v) \equiv H(F_X^{-1}(u), F_Y^{-1}(v)),$$

H is the joint Cdf of (X, Y) ; F_X^{-1} and F_Y^{-1} are the right-continuous inverses of F_X and F_Y .

Under s -independence of X and Y , the copula is

$$C(u, v) = u \cdot v \equiv \Pi,$$

and any copula must fall between

$$M(u, v) \equiv \min(u, v) \quad \text{and} \quad W(u, v) \equiv \max(u + v - 1, 0),$$

the copulas of the upper and lower Fréchet bounds [9]. As in model #1, under the assumption of s -independence of X and Y , the marginal Cdf of X and Y are uniquely determined by the sub-Sf of X and Y . The more general result [12] is: if the copula of (X, Y) is known, then the marginal Cdf of X and Y are uniquely determined by the competing-risk data, as in theorem 1.

Theorem 1: Let the marginal Cdf of (X, Y) be continuous and strictly increasing in $(0, \infty)$. Let the copula C be known, and the corresponding probability measure for any open set of the unit square be positive. Then F_X and F_Y , the marginal Cdf of X and Y , are uniquely determined by the sub-Cdf.

The Appendix shows briefly why the marginals are identifiable when the pdfs and sub-pdfs exist.

The problem of choosing a copula is now considered. There are many measures of association for the pair (X, Y) , which are symmetric in X and Y . The best known measures of association are Kendall's τ and Spearman's ρ ; the more modern term "measure of association" is used instead of "correlation coefficient" for a measure of dependence between r.v.

Kendall's τ for a vector (X, Y) of continuous r.v. with joint Cdf H is defined:

Let (X_1, Y_1) and (X_2, Y_2) be i.i.d. random vectors, each with joint Cdf H ; then the Kendall τ is defined as the proba-

bility of concordance minus the probability of discordance:

$$\tau(X, Y) \equiv \Pr\{(X_1 - X_2) \cdot (Y_1 - Y_2) > 0\} \\ - \Pr\{(X_1 - X_2) \cdot (Y_1 - Y_2) < 0\};$$

or

$$\tau(X, Y) \equiv \Pr\{\text{sgn}(X_1 - X_2) = \text{sgn}(Y_1 - Y_2)\} \\ - \Pr\{\text{sgn}(X_1 - X_2) \neq \text{sgn}(Y_1 - Y_2)\}.$$

The other measure of association, Spearman's ρ , is defined as:

Let X and Y be continuous r.v.; then Spearman's ρ is defined as the product moment correlation of $F_X(X)$ and $F_Y(Y)$:

$$\rho_r(x, y) \equiv \rho(F_X(X), F_Y(Y)) \\ = \frac{\text{Cov}[F_X(X), F_Y(Y)]}{\sqrt{\text{Var}[F_X(X)] \cdot \text{Var}[F_Y(Y)]}}.$$

Simple formulae relating the measures of association to copula density are given in the Appendix.

Because the measure of association is to be treated as a primary parameter, it is necessary to choose a family of copulae which are "smooth" as possible and which model all possible measures of association in a simple way. Reference [5] proposes using the unique copula with the given Spearman's ρ that has minimum information with respect to the s -independent distribution, and gives a method to numerically-calculate this copula. Due to the difficulties of interpreting Spearman's ρ by a nonspecialist, and of quantifying it, Kendall's tau is used as a primary parameter. Kendall's τ has the advantage of a definition which can be explained to a nonspecialist, but the value can not be estimated using only the competing risk data, because of its "identifiability problem."

Thus some prior knowledge or subjective information must be used to obtain information about the value of τ . Expert judgment is used to model the uncertainty over τ , and is discussed later in this section. For now, the way to obtain the copula must be explained.

Reference [12] suggests that the important factor for an estimate of the marginal Sf is a reasonable guess at the strength of the association between competing risks, rather than the functional form of the copula. Thus a class of copula with which it is easy to work from the mathematical view-point is chosen, e.g., Archimedean copula. Some definitions about the Archimedean copula and some properties of Kendall's τ for a certain Archimedean family of copula are explained.

Let X and Y be continuous r.v. with joint Cdf H and marginal Cdf F_X and F_Y . When X and Y are s -independent, then $H(x, y) = F_X(x) \cdot F_Y(y)$; this is the only case when the joint distribution is written as a product of F_X and F_Y . But, there are some families of distributions in which $\lambda(H(x, y)) = \lambda(F_X(x)) \cdot \lambda(F_Y(y))$ [9].

Using $\varphi(t) = -\log(\lambda(t))$, then $[\lambda(\cdot)]$ must be positive on the interval $(0, 1)$, and H is a sum of the marginals F_X and F_Y , $\varphi(H(x, y)) = \varphi(F_X(x)) + \varphi(F_Y(y))$, or in terms of copula $\varphi(C(u, v)) = \varphi(u) + \varphi(v)$.

Copulas of this form are "Archimedean copulas." The function φ is an "additive generator" of the copula. If $\varphi(0) \rightarrow \infty$, then φ is a strict generator and $C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$ is a strict Archimedean copula. For the goal in this paper, choose a

1-parameter family of copulae which has a strict generator. The Gumbel family is defined as:

$$C_\alpha(u, v) \equiv \exp\left(-[(-\log u)^\alpha + (-\log v)^\alpha]^{1/\alpha}\right) \\ \text{for } \alpha \in [1, \infty).$$

The generator is the function: $\varphi_\alpha(t) = (-\log(t))^\alpha$.

The Appendix shows that α can be written as a function of Kendall's τ : $\alpha_\tau = (1 - \tau)^{-1}$.

It remains now to quantify the uncertainty in Kendall's τ using expert opinion. Experts can not be directly asked to quantify their uncertainty over τ , instead they are asked to give uncertainties over physically realizable quantities [3].

- Consider 2 sockets with failure times, X_1 and X_2 , and the PM times Y_1 and Y_2 .
- The expert can be asked for the "probability that an attempt to do PM for socket #1 would occur before the PM for socket #2, given that the failure of socket #1 occurs before the failure of socket #2;" let this probability be q .
- By symmetry the probability is the same for the occurrence of the PM for socket #2 before the PM for socket #1, given that the failure time of socket #1 is greater than the failure time of socket #2.
- The "probability of occurrence of the PM for socket #2 before the PM for socket #1 given that the failure time of socket #1 is smaller than the failure time of socket #2" is $1 - q$.

If the experts can give a distribution for $q \equiv \Pr\{Y_1 > Y_2 | X_1 > X_2\}$, then this can be converted to a distribution over Kendall's τ . Thus,

$$\Pr\{(X_1 - X_2) \cdot (Y_1 - Y_2) > 0\} \\ = \Pr\{(X_1 > X_2) \cap (Y_1 > Y_2)\} \\ + \Pr\{(X_1 < X_2) \cap (Y_1 < Y_2)\} \\ = \Pr\{X_1 > X_2\} \cdot \Pr\{Y_1 > Y_2 | X_1 > X_2\} \\ + \Pr\{X_1 < X_2\} \cdot \Pr\{Y_1 < Y_2 | X_1 < X_2\} = q.$$

Similarly:

$$\Pr\{(X_1 - X_2) \cdot (Y_1 - Y_2) < 0\} = 1 - q;$$

thus $\tau = 2q - 1$.

The q can be considered an observable quantity because q is the approximate average rate for which $\{Y_1^{(n)} > Y_2^{(n)} | X_1^{(n)} > X_2^{(n)}\}$ holds when a large sample of pairs $(X_1^{(n)}, Y_1^{(n)})$, $(X_2^{(n)}, Y_2^{(n)})$ is observed.

For each θ and τ , calculate the long-term specific cost; then optimize this replacement cost by finding the minimal cost. This is discussed in Section IV.

IV. MAINTENANCE OPTIMIZATION

Consider the effect of uncertainty about the underlying life-time distribution on the selection of the maintenance policy. To keep things simple, consider the age-replacement policies. An age-replacement policy is one for which replacement occurs at failure or at age θ , whichever occurs first. Unless otherwise specified, θ is a constant.

In the "finite time-span replacement model" minimize $C(\theta)$ experienced during $[0, \theta]$; the cost is computed in money units,

TABLE I
OPTIMAL MAINTENANCE TIMES AND COSTS

c_1/c_2	10				20			
	0.3		0.1		0.3		0.1	
	RT	RC	RT	RC	RT	RC	RT	RC
D1, M1	0.8280	10.9217	1.2849	9.8756	0.6127	18.8850	0.9508	21.6988
D1, M2	0.3860	7.5921	0.3860	7.5921	0.2574	9.9433	0.2574	9.9433
D2, M1	0.8687	4.9674	1.5047	6.5283	0.5950	7.2682	1.0305	10.1017
D2, M2	0.4758	4.2823	0.4758	4.2823	0.3259	6.1916	0.3259	6.1916
D3, M1	0.3393	3.3011	0.3353	3.3280	0.2322	4.6522	0.2303	4.6821
D3, M2	0.3556	3.2006	0.3556	3.2006	0.2393	4.5459	0.2393	4.5459

time, or an appropriate combination. For an infinite time span, an appropriate objective function is mean-cost per time-unit:

$$\gamma(\theta) \equiv \lim_{\theta \rightarrow \infty} \left[\frac{C(\theta)}{\theta} \right].$$

$N_1(\theta) \equiv$ number of failures during $[0, \theta]$,

$N_2(\theta) \equiv$ number of planned PM during $[0, \theta]$;

$c_1 \equiv$ cost of critical failure,

$c_2 \equiv$ cost for planned replacement.

The mean cost during $[0, \theta]$ is:

$$C(\theta) \equiv c_1 \cdot E[N_1(\theta)] + c_2 \cdot E[N_2(\theta)].$$

Consider only nonrandom age replacement in seeking the policy minimizing the $\gamma(\theta)$ for an infinite time span.

By definition of

$$\gamma(\theta) \equiv \lim_{\theta \rightarrow \infty} \left[c_1 \cdot \frac{E[N_1(\theta)]}{\theta} + c_2 \cdot \frac{E[N_2(\theta)]}{\theta} \right]$$

[1] shows that

$$\gamma(\theta) \equiv \frac{c_1 \cdot F(\theta) + c_2 \cdot S(\theta)}{\int_0^\theta S(t) dt}.$$

Then

$$\gamma(0) \rightarrow \infty, \quad \gamma(\infty) = \frac{c_1}{\int_0^\theta S(t) dt}.$$

Differentiate γ to find the optimum, $d\gamma(\theta)/d\theta = 0$, then

$$r(\theta) \cdot \int_0^\theta S(t) dt - F(\theta) = \frac{c_2}{c_1 - c_2}.$$

When $F_X(x)$ has an increasing failure rate, the optimal θ_0 is the unique solution of this equation. For a r.v. with constant failure rate or decreasing failure rate, the specific cost does not have an optimum:

$$\text{sign} \left(\frac{d\gamma(\theta)}{d\theta} \right) \text{ is constant;}$$

thus this type of maintenance policy is not appropriate for a such r.v.

When the primary parameter is Kendall's τ and the information on τ is $F_\tau(\tau)$ and $f_\tau(\tau)$, the specific cost depends on τ and θ :

$$\gamma(\tau, \theta) \equiv \frac{c_1 \cdot F(\tau, \theta) + c_2 \cdot S(\tau, \theta)}{\int_0^\theta S(\tau, t) dt}.$$

TABLE II
RATIOS OF MAINTENANCE TIMES AND COSTS

c_1/c_2	10				20			
	0.3		0.1		0.3		0.1	
	RT	RC	RT	RC	RT	RC	RT	RC
D1	2.15	1.44	3.33	1.3	2.38	1.9	3.69	2.18
D2	1.83	1.16	3.16	1.52	1.83	1.17	3.16	1.63
D3	0.95	1.03	0.94	1.04	0.97	1.02	0.96	1.03

Thus the long term specific cost, given θ , is

$$\gamma(\theta) = \int_0^1 C(\tau, \theta) \cdot f_\tau(\tau) d\tau$$

and θ_0 is obtaining minimizing $\gamma(\theta)$.

V. NUMERICAL EXAMPLES

Acronyms:

D_i Distribution # i

M_j Model # j

Three numerical experiments show

- the effect of using M1 when M2 actually holds,
- the dependence of replacement cost with the measure of association (Kendall's τ),
- the optimal replacement time of the average specific cost.

A. Part 1

Consider 3 distributions for X :

$$D1: r_X(t) = t^{3/2}$$

$$D2: r_X(t) = t^2$$

$$D3: r_X(t) = t^3.$$

The failure rates are Weibull, and are continuous and increasing.

Because the costs of critical failure can be much higher than those of PM (because of other consequences to the system beyond the simple need to replace the failed unit), let $c_1 \gg c_2$. Because actual plant data show a many PM actions, let p be small:

$$\frac{c_1}{c_2} = 10, 20;$$

$$p = 0.3, 0.1;$$

thus there are 4 cases to compare the models. Both RT and RC are in Table I. The RT are the optimal-RT calculated under the assumption that the model is correct. For M2 the RC are the optimal-RC. For M1 the RC are the RC of M2 (which is actually the correct model), evaluated with the optimal RT calculated for M1. Hence the costs for M1 are always higher than those of M2. Table II gives the ratio of the 2 model-outcomes (M1-outcome

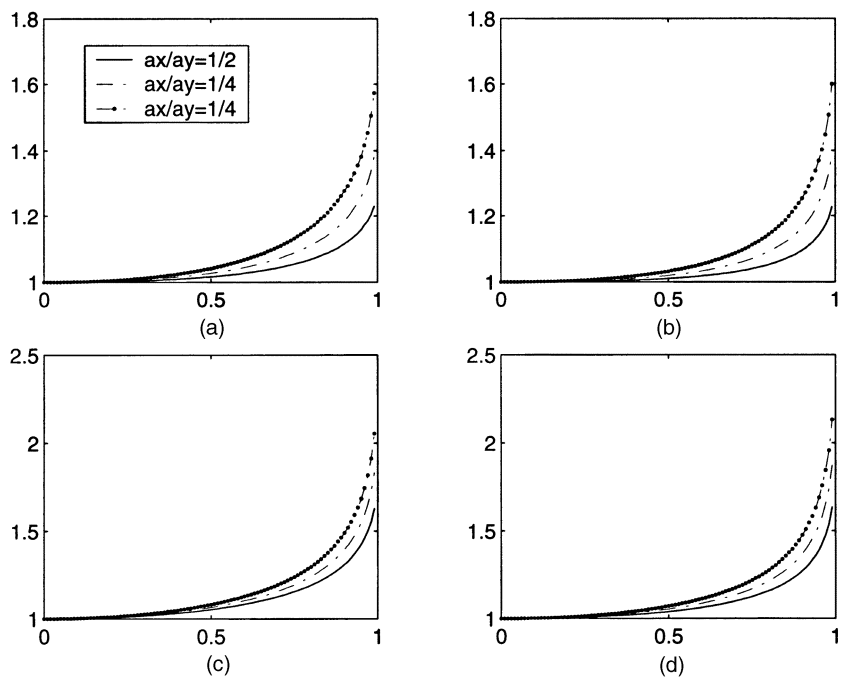


Fig. 2. Dependence between RC and the measure-of-motivation for the pairs of sub-Sf for X and Y given by sub-Sf #1 for X , and the other three for Y . (a) $p=0.3, c_2/c_1=0.1$; (b) $p=0.3, c_2/c_1=0.05$; (c) $p=0.1, c_2/c_1=0.1$; (d) $p=0.1, c_2/c_1=0.05$.

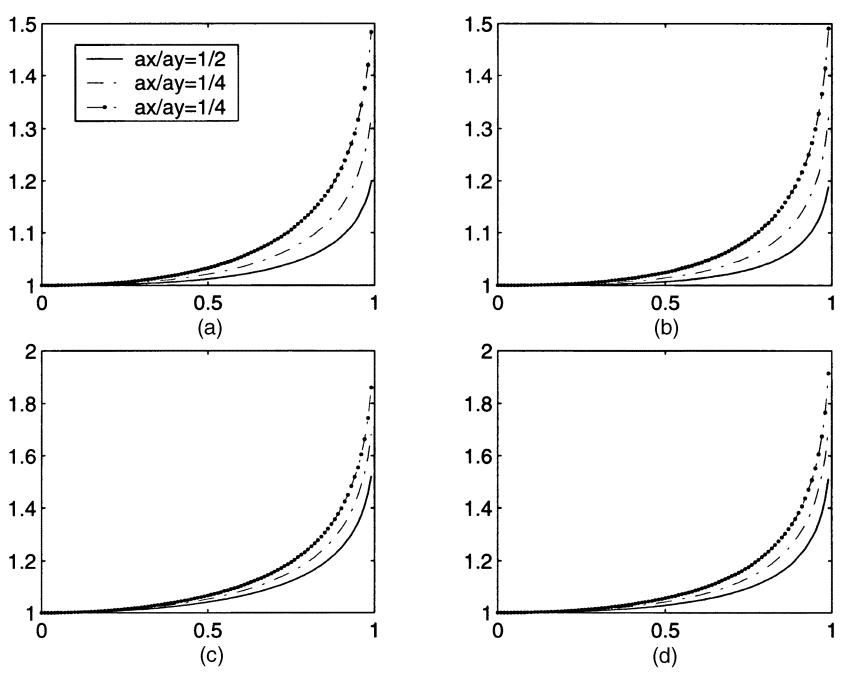


Fig. 3. Dependence between RC and the measure-of-motivation for the pairs of sub-Sf for X and Y given by sub-Sf #2 for X , and the other three for Y . (a) $p=0.3, c_2/c_1=0.1$; (b) $p=0.3, c_2/c_1=0.05$; (c) $p=0.1, c_2/c_1=0.1$; (d) $p=0.1, c_2/c_1=0.05$.

divided by M2-outcome) for the time and costs of each of the distributions.

B. Part 2

Consider 3 sub-Sf for X which for the extreme cases (s -independence and high s -correlation) take the same failure rates for X as in Section V-A and for every sub-Sf of X , take the other 3 sub-Sf Y in such a way that inequality (4) is satisfied;

- use Weibull distributions with the same shape parameter of S_X^* and S_Y^* ,
- the scale parameter of S_Y^* , a_Y , must be greater than the scale parameter of S_X^* , a_X .

Thus, take:

$$\frac{a_X}{a_Y} = \frac{1}{2}, \quad \frac{a_X}{a_Y} = \frac{1}{4}, \quad \frac{a_X}{a_Y} = \frac{1}{8};$$

for p and c_1/c_2 use the same values as in Section V-A.

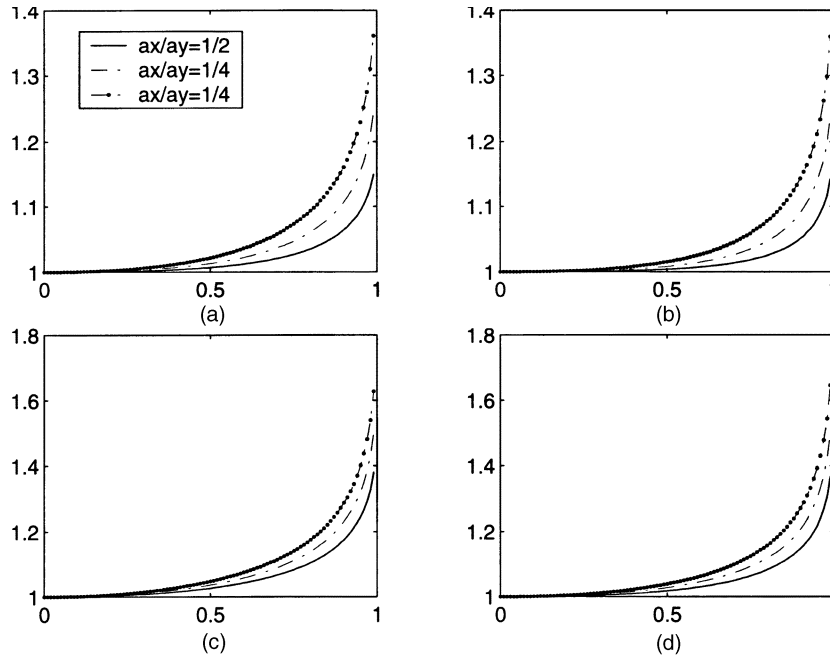


Fig. 4. Dependence between RC and the measure-of-motivation for the pairs of sub-Sf for X and Y given by sub-Sf #3 for X , and the other three for Y . (a) $p=0.3, c_2/c_1=0.1$; (b) $p=0.3, c_2/c_1=0.05$; (c) $p=0.1, c_2/c_1=0.1$; (d) $p=0.1, c_2/c_1=0.05$.

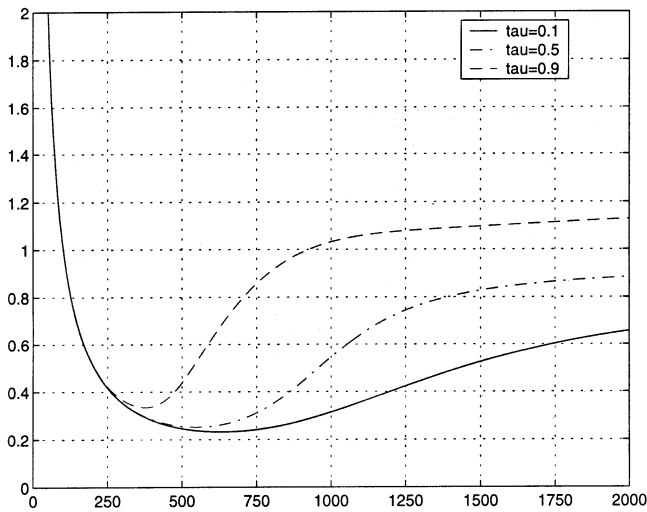


Fig. 5. Specific cost for 3 values of Kendall's τ : 0.1, 0.5, 0.9.

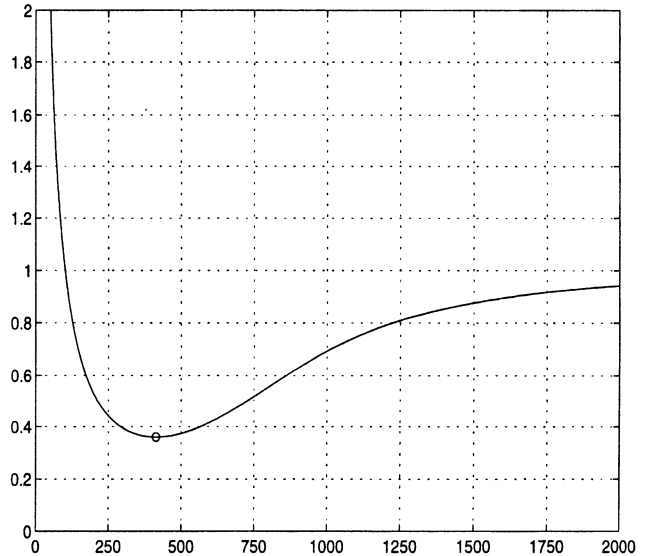


Fig. 6. Average specific cost and optimal replacement time.

Figs. 2–4 show the way in which the RC (normalized by RC for the independent case) depends on Kendall's τ .

To obtain a distribution for Kendall's τ , ask an expert to give quantiles for the q defined in Section III-C. If the expert gives 5% and 95% quantiles then fit a beta distribution; if $\Pr\{q \leq 0.7\} = 0.05$ and $\Pr\{q \leq 0.95\} = 0.95$, then the 5% and 95% quantiles for τ are $\Pr\{\tau \leq 0.4\} = 0.05$ and $\Pr\{\tau \leq 0.9\} = 0.95$. Assume the beta distribution for τ ; then the parameters of this distribution, given the 5% and 95% quantiles, are: $a = 5.6705, b = 2.7322$.

Fig. 5 shows the specific costs for various values of Kendall's τ ; Fig. 6 shows the average specific costs with optimal replacement times.

VI. DISCUSSION

The results in Tables I and II show that the “optimal replacement interval” and “optimal replacement costs” can be dramatically nonoptimal when the wrong model is used to estimate the underlying failure distribution from censored data. The difference is least when the failure rate increases quickly. When the failure rate increase more slowly, the difference is larger. For one case calculated here, the specific costs obtained by using the independent model are more than twice the best possible specific costs using the correct model.

Section V-B considers the effect of model uncertainty due to the impossibility of identifying the “correct” competing risk

model from censored data. Using expert-judgment to quantify the dependence between competing risks, shows that the replacement cost is highly sensitive to Kendall's τ . Figs. 1–3 show that sensitivity is higher for distribution #1 and, for a certain case, RC can be twice that of RC for the independent case. Fig. 4 shows that the difference between optimal replacement costs and optimal replacement time can be more than a factor of 2. Fig. 5 presents the long-term average specific-cost and optimal replacement-time.

This work demonstrates the importance of using good expert judgment from experts with insight into the maintenance process. If the experts can select the correct correlation level then the results in this paper will considerably aid model-selection.

APPENDIX

A. Part 1

This briefly shows why the marginals are identifiable when pdfs and sub-pdfs exist. By definition, the sub-Cdf of X is $F_X^*(t) \equiv \Pr\{X \leq t, X < Y\}$, and

$$\begin{aligned} F_X^*(t) &= \int_0^t \int_x^\infty h(x, y) dx dy \\ &= \int_0^t [H_X(x, \infty) - H_X(x, x)] dx \\ &= F_X(t) - \int_0^t H_X(x, x) dx \\ &= F_X(t) - \int_0^t C_u(F_X(x), F_Y(x)) \cdot f_X(x) dx, \end{aligned}$$

$$\begin{aligned} h(x, y) &\equiv \text{joint pdf of } X, Y, \\ H_X(x, \infty) &\equiv (\partial/\partial x)H(x, y), \text{ calculated in } (x, \infty), \\ H_X(x, x) &\equiv (\partial/\partial x)H(x, y), \text{ calculated in } (x, x). \end{aligned}$$

An analogous formula for F_Y^* is obtained; from this formula it follows that the marginal Cdf F_X and F_Y are solutions of the following system of ordinary differential equations:

$$\begin{aligned} [1 - C_u(F_X(t), F_Y(t))] \cdot F_X'(t) &= F_X^{*\prime}(t) \\ [1 - C_v(F_X(t), F_Y(t))] \cdot F_Y'(t) &= F_Y^{*\prime}(t) \end{aligned}$$

with initial conditions: $F_X(0) = F_Y(0) = 0$,

$$C_u(F_X(t), F_Y(t)) \equiv \frac{\partial C(u, v)}{\partial u},$$

$$C_v(F_X(t), F_Y(t)) \equiv \frac{\partial C(u, v)}{\partial v},$$

both calculated in $(F_X(t), F_Y(t))$.

B. Part 2

To see the relations between the measures-of-association and copula, recall [9, theorems 2–4].

Theorem 2: Let X and Y be continuous r.v. with copula C . Then Kendall's τ for X and Y , denoted by either $\tau(X, Y)$ or τ_C , is

$$\tau(X, Y) = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1.$$

Theorem 3: Let X and Y be continuous r.v. with copula C . Then Spearman's ρ for X and Y [denoted by either $\rho(X, Y)$ or ρ_C] is

$$\rho(X, Y) = 12 \int_0^1 \int_0^1 u \cdot v dC(u, v) - 3,$$

$$\rho(X, Y) = 12 \int_0^1 \int_0^1 C(u, v) du dv - 3.$$

From [9, theorem 5], determines the parameter α (and implicitly the copula) when Kendall's τ is known.

Theorem 4: Let X and Y be r.v. with an Archimedean copula C generated by $\varphi \in \Omega$. Kendall's τ for X and Y is

$$\tau_C = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt.$$

If C_α is a member of the Gumbel family, then for $\alpha \geq 1$,

$$\frac{\varphi(t)}{\varphi'(t)} dt = \frac{t \cdot \log(t)}{\alpha},$$

so that $\tau(\alpha) = 1 - (1/\alpha)$. Thus $\alpha_\tau = 1/(1 - \tau)$.

REFERENCES

- [1] R. E. Barlow and F. Proschan, *Mathematical Theory of Reliability*: Wiley, 1965.
- [2] T. Bedford, "On the use of minimally informative copulae in competing risk problems," in *Statistical Probabilistic Models in Reliability*, C. Ionescu and N. Limnios, Eds: Birkhauser, 1998.
- [3] T. Bedford and R. M. Cooke, *Probabilistic Risk Analysis: Foundations and Methods*: Cambridge, 2001.
- [4] T. Bedford and I. Meilijson, "The marginal distributions of lifetime variables which right censor each other," in *Analysis of Censored Data*, H. L. Koul and J. V. Deshpande, Eds., 1995, vol. 27, IMS Lecture Notes Monograph Series.
- [5] —, "A characterization of marginal distributions of (possibly dependent) lifetime variables which right censor each other," *Ann. Statistics*, vol. 25, pp. 1622–1645, 1997.
- [6] T. Bedford and C. Mesina, "The impact of modeling assumptions on maintenance optimization," in *Mathematical Methods in Reliability*, N. Limnios and M. Nikulin, Eds: Birkhauser, 2000.
- [7] R. M. Cooke, "The total time on test statistic and age-dependent censoring," *Statistics and Probability Lett.*, vol. 18, no. 5, 1993.
- [8] —, "The design of reliability data bases, Part II," *Reliability Engineering and System Safety*, vol. 51, no. 2, pp. 209–225, 1996.
- [9] R. B. Nelsen, *An Introduction to Copulas*: Springer, 1995.
- [10] B. Schweizer and E. F. Wolff, "On nonparametric measures of dependence for random variables," *Ann. Statistics*, vol. 9, pp. 879–885, 1981.
- [11] A. Tsiatis, "A nonidentifiability aspect in the problem of competing risks," *Proc. Nat. Academy of Science USA*, vol. 72, pp. 20–22, 1975.
- [12] M. Zheng and J. P. Klein, "Estimates of marginal survival for dependent competing risks based on an assumed copula," *Biometrika*, vol. 82, pp. 127–138, 1995.

Cornel Bunea (born 1975) received his M.Sc. in 1999 in safety and reliability of power systems from Bucharest University of Technology, and is now a Ph.D. student in Competing Risk Analysis at Delft University of Technology.

Tim Bedford (born 1960) is a Professor at Strathclyde University in Scotland and an Associate Professor at TU Delft in the Netherlands. His current interests are in reliability, risk analysis, and decision making. Research in these areas is carried out often in combination with private industry and/or government institutes. Together with Roger Cooke he has written a book *Probabilistic Risk Analysis: Foundations and Methods*, published by Cambridge University Press.