A Bayes-competing risk model for the use of expert judgement in reliability estimation

Reliability Engineering and System Safety 35 (1992) 23–30

A Bayes-competing risk model for the use of expert judgment in reliability estimation*

F. P. A. Coolen,[a] P. R. Mertens[b] & M. J. Newby[a]‡

[a] *Faculty of Industrial Engineering, Technical University Eindhoven, Postbox 530, 5600 MB Eindhoven, The Netherlands*
[b] *Reliability Assurance, Kemira bv, Moezelweg 151, P.O. Box 1015, 3180AA Rozenburg, The Netherlands*

(Received 15 March 1990; accepted 15 August 1990)

In an analysis of the reliability of heat exchangers, a number of causes of failure and a number of potential failure patterns were identified. To allow a simplification of the analysis, five independent groups of failure causes were defined and formed the basis of a competing risks model. Further, lack of data and poor quality data required the experience of 'experts' to be used in a quantitative way through Bayesian methods. Lastly, to simplify the elicitation of the prior densities, a shape, scale and location parameter model for the failure time distributions was adopted.

INTRODUCTION

As part of a general review of reliability in a chemical process plant, a study of the reliability of heat exchangers was undertaken.[1] The heat exchangers fall into two groups: critical and noncritical. The critical exchangers operate at high temperatures and pressures, and any fault requires an immediate shutdown of the plant; the noncritical exchangers operate under less extreme conditions and can tolerate some faults (a few leaking pipes, for example) without requiring a stop in production. Safety regulations also require regular testing of the plant, and a good record of uninterrupted and hazard free operation allows the required intervals between assessments to be extended, in other words, a saving on shutdown and testing costs.

The study of the heat exchangers began with an examination of past failure and maintenance data. It became clear at this stage of the study that a number of factors made the recorded data useless without the mediation of a complex model, or without the incorporation of the experience and knowledge of the users. The problems with the data were manifold: recorded events were not always well defined, there were very few observations associated with any particular failure cause or mode, and there had been a continuous change in the design and the process conditions of the exchangers. However, there was among the maintenance engineers, design engineers and plant operators, a significant amount of knowledge and experience about the performance of the equipment under different conditions, the effects of design changes, and the performance of similar equipment in other plants.

The reliability problem is to specify a conditional reliability for each heat exchanger given its past failure history, its maintenance history and any changes in conditions or design. The function, $\mathcal{R}(t \mid t_0)$ defines the reliability for a period t into the future, given information today, t_0. There are a number of ways of incorporating explanatory variables into a model to reflect the effects of such things as temperature, pressure, choice of material and maintenance policy. In reliability analysis, the proportional hazards model[2,3] or the accelerated failure time model[2,4] are often used to show explicitly the influence of explanatory variables on the hazard rate or reliability. However, the data available in the current study could not support the estimation of parameters in these two particular models. A third approach, commoner in epidemiology, is through the use of competing risks models,[5] some proportional hazards models are also competing risks models, and sometimes stratification is used in proportional hazards models to give a model which lies between the two. Competing risks models are able to work without

* This work was carried out mainly by the first-named author as part of an MSc project while a student at the Technical University, Eindhoven, The Netherlands.
‡ To whom correspondence should be addressed.

Reliability Engineering and System Safety 0951-8320/91/$03·50

requiring independence between the causes or types of failure, but are then hard to interpret and difficult to estimate. In this study, a competing risks framework with independent failure modes was chosen for the underlying failure model.

In the second stage, it was necessary to incorporate the information available within the plant into the model. This was done by using a straightforward Bayesian method after eliciting prior distributions of the parameters of interest from the 'experts'. A second function of the Bayesian approach is to lay the basis for the evaluation of the performance of the experts as judges of failure behaviour and to quantify their performance. Through establishing 'scores', which measure the effectiveness of an expert, a feedback element can be incorporated to help the 'experts' to improve their own performance over time. In the long run the interaction between the experts and the data is the most important feature of this approach. A good and systematic method of data collection and analysis opens the door to better decision making.

COMPETING RISK MODELS

In a competing risks model with independent failures there are n failure types defined, and each failure type is described by the distribution, reliability function, density or hazard rate denoted by F_i, R_i, f_i or h_i, respectively, where

$$F_i(t) = \mathscr{P}[T_i \leq t]$$

$$R_i(t) = 1 - F_i(t)$$

$$f_i(t) = \frac{\mathrm{d}F_i}{\mathrm{d}t}$$

$$h_i(t) = \frac{f_i(t)}{R_i(t)}$$

for $i = 1, \ldots, n$, and T_i is the time of failure due to the i-th cause. The model assumes that the times of failure, as random variables, are T_i and that the observed time of failure is then one of the T_i. The time of failure is thus the minimum of the failure times $T_f = \min(T_1, T_2, T_3, \ldots, T_n)$. Because the failures are assumed to be independent it follows that

$$\mathscr{P}[T_f > t] = \mathscr{P}[T_i > t \forall i] = \prod_{i=1}^{n} R_i(t)$$

so that

$$\mathscr{R}(t) = \prod_{i=1}^{n} R_i(t) \tag{1}$$

and the hazard rate for the process is simply the sum of the individual hazard rates

$$h(t) = \sum_{i=1}^{n} h_i(t) \tag{2}$$

Conversely, the marginal distributions are the F_i. When a satisfactory grouping into independent failure modes is possible, the competing risks model permits the analysis of each group separately so long as the parameters of the failure distributions of each group are functionally independent. Moreover, the Bayesian analysis for each group may also be performed separately. After the failure distributions for each group have been established they may be combined according to eqn (1) or (2) to give an overall reliability function or hazard rate. Expert judgment also enters into this part of the model through the experts' definition of the groups and the assumption of independence.

THE BAYESIAN BACKGROUND

In this section, only a single failure time distribution is mentioned; however, in view of the competing risks assumptions this causes no problem, for each failure group may be treated separately and the results recombined at the end. This approach also requires the experts to give independent assessments of the priors for each failure group.

A standard Bayesian approach was used with the data entering through a likelihood and the judgment entering through the prior density, the combination of the two yielding the posterior density.[6] In this paper, attention is focused on predictive distributions since they are functions of the data alone and, moreover, allow probability statements to be about system performance. For the user of this approach, estimates for parameters themselves are of less interest. The method was extended in a natural way to deal with several priors obtained from a number of experts and to provide a weight or score for each expert that reflected the quality of that expert in describing failure behaviour.[7] Thus an expert whose assessment was close to the observed failure behaviour should receive a higher score than one whose assessment is far from the actual events. In short, the closer the mode of the prior and the likelihood, the higher the score.

One parameter model

To simplify both the elicitation process[8] and calculations, a discrete prior was chosen for the parameter of interest. Further, in the interests of elicitation a shape, scale and location parameter family of distributions was used for the failure model. Such a family is defined by a distribution, $G(t; \beta)$, with a single shape parameter, β, and mean $\mu(\beta)$. A derived distribution, G_1 is obtained by affine transformations of the time variable, namely, $u = \{t - \tau/\vartheta\}$, giving $G_1(t) = G(u; \beta)$. More importantly, in this context, τ represents a 'guarantee period' in which no failure can take place, and the mean of the derived distribution, μ', is given by

$$\mu' = \vartheta\mu(\beta) + \tau,$$

so that

$$\vartheta = (\mu' - \tau)/\mu(\beta)$$

Questions about the shape parameter and location parameter can be put as questions about mean, μ', and the equation inverted to give a prior for ϑ.

In this particular problem the shape parameter β has been taken as a fixed value and interest centres on ϑ and τ. The influence of the shape parameter will be discussed later. The prior density for ϑ is used to define the guarantee period τ as well as the posterior density. There are n experts each of whom specifies the prior for ϑ by choosing a discrete distribution as follows. Each expert specifies in an interview a prior distribution for the mean value of the lifetime of the component. The expert is offered m intervals (not less than 3 and not more than 10) and asked to specify the probability that the mean of the life distribution lies in a particular interval. The intervals are $[0, a_1]$, $[a_1, a_2], \ldots, [a_{m-1}, a_m]$ where, in principle, a_m is infinite, but is here simply taken large enough to encompass all reasonable mean values. Expert j specifies a probability mass function p_{jk} where p_{jk} is the probability that the mean, μ', of the distribution lies in an interval $[a_{k-1}, a_k]$. The location parameter is defined as

$$\tau = \min_j\{\max_k\{a_k \mid p_{jk} = 0\}\}$$

and the scale parameter lies in the interval

$$\left[\frac{a_{k-1} - \tau}{\mu(\beta)}, \frac{a_k - \tau}{\mu(\beta)}\right]$$

with probability p_{jk}. In other words, the location parameter is taken as the lowest plausible time in the opinion of the experts at which a failure can occur.

Let the prior for expert i be p_{ik}, and the likelihood given a parameter value in the k-th interval be $l_k = \textit{likelihood}\ (\text{data} \mid \vartheta_k)$, set $s_i = \sum_k p_{ik} l_k$, where s_i is referred to as the score for expert i, then the Bayes formula for the posterior of expert i is

$$p^*_{ik} = \frac{p_{ik} l_k}{s_i}$$

This one parameter model is particularly easy to formalise as follows:

Intervals for prior	$[0, a_1], [a_1, a_2], \ldots, [a_{m-1}, a_m]$
For calculating means	$\tilde{a} = (\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m)^T$, where $\tilde{a}_j = \xi_j a_{j-1} + (1 + \xi_j) a_j$ for some $\xi_j \in [0, 1]$ and $j = 1 \cdots m$
Prior for expert i	$p_i = (p_{i1}, p_{i2}, \ldots, p_{im})$
Likelihood	$l = (l_1, l_2, \ldots, l_m)^T$
Overall prior	$\mathcal{P} = (p_{ij})_{nm}$
Score vector	$s = (s_1, s_2, \ldots, s_n)^T$
then	$s = \mathcal{P} l$
Setting	$S = \text{diag}(s_i)$, $L = \text{diag}(l_j)$
	$u_n = (1, 1, \ldots, 1)^T$, a vector of n 1's
and	$u_m = (1, 1, \ldots, 1)^T$, a vector of m 1's.

The vectors u_n and u_m are useful for computing sums, with these definitions we have

$$l = L u_m$$

and

$$s = S u_n$$

The posterior densities can be summarised as an unnormalised version, $\mathcal{P}_1$, and a normalised version, $\mathcal{P}_0$, where

$$\mathcal{P}_1 = \mathcal{P} L$$

$$\mathcal{P}_0 = S^{-1}\mathcal{P}_1 = S^{-1}\mathcal{P} L$$

The estimates of the posterior mean are then simply

$$\bar{t} = \mathcal{P}_0 \tilde{a}$$

We evaluate the underlying failure density, f, at an arbitrary value of t and define

$$f_k = f(t \mid \vartheta_k)$$

$$f = (f_1, f_2, \ldots, f_m)^T$$

so that the predictive density for each expert evaluated at t is given in the vector:

$$\bar{f} = \mathcal{P}_0 f$$

In general, if we define a list of q arbitrary times $(t_1, t_2, \ldots, t_q)$ at which the predictive density or reliability is required, and a row vector

$$f_k = (f(t_1 \mid \vartheta_k), f(t_2 \mid \vartheta_k), \ldots, f(t_q \mid \vartheta_k))$$

$$\mathscr{f} = (f_{ki})_{mq}$$

then the predictive densities $\mathscr{f}$ are given by

$$\bar{\mathscr{f}} = \mathcal{P}_0 \mathscr{f}.$$

In the same way, if $\mathfrak{R}$ represents the reliability matrix with the same evaluation points as $\mathscr{f}$,

$$\mathfrak{R} = (R_{ki}) = (R(t_i \mid \vartheta_k))_{mq}$$

then the predictive reliability is

$$\bar{\mathfrak{R}} = \mathcal{P}_0 \mathfrak{R}$$

A single weighted posterior density

The most natural way to construct a single posterior density from $\mathcal{P}_0$ is to construct a weighted sum using a weight vector $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ to yield a posterior

$$\mathbf{P} = \mathbf{w}\mathcal{P}_0$$

Because of the paucity of data, the lack of suitable comparisons, and the different backgrounds of the experts, the scores of the experts were used rather than calibration by the method of 'seed variables'.[7] The weights w_i should reflect the performance of the experts and the easiest way to do this is to choose $w_i \propto s_i$ and to calculate an appropriate norming constant so that the sum of the weights is 1, in our notation $wu_n = 1$. Now

$$s^T\mathcal{P}_0 = (Su_n)^T\mathcal{P}_0 = u_n^T S^T \mathcal{P}_0 = u_n^T S\mathcal{P}_0$$
$$= u_n^T S(S^{-1}\mathcal{P}L) = u_n^T \mathcal{P}L = u_n^T \mathcal{P}_1$$

showing that this version of the posterior is simply the sum of the unnormalised posteriors. To find the appropriate norming constant, calculate the sum

$$s^T\mathcal{P}_0 u_m = u_n^T(\mathcal{P}Lu_m) = u_n^T s$$

Thus, the norming factor is the sum of the scores $|s| = \sum_{i=1}^n s_i$. To conclude, the weighting vector is simply $\mathbf{w} = s^T/|s|$.

In summary:

Weight vector	$\mathbf{w} = s^T/\lvert s\rvert$
Posterior density	$\mathbf{P} = \mathbf{w}\mathcal{P}_0$
Posterior mean	$\bar{t} = \mathbf{w}\bar{t}$
Predictive density	$\bar{\mathscr{f}} = \mathbf{w}\bar{\mathscr{f}}$
Predictive reliability	$\bar{\mathfrak{R}} = \mathbf{w}\bar{\mathfrak{R}}$

Within the Bayesian scheme the elegance of the approach breaks down a little when we consider the hazard rate. It is easiest to show the problem in a general case. Consider a model $f(t \mid \vartheta)$ with likelihood $l(\text{data} \mid \vartheta)$ and prior $g(\vartheta)$ for ϑ. The posterior for ϑ given by

$$g(\vartheta \mid \text{data}) = \frac{l(\text{data} \mid \vartheta)g(\vartheta)}{\int l(\text{data} \mid \vartheta)g(\vartheta)\,\mathrm{d}\vartheta}$$

a predictive density

$$\tilde{f}(t \mid \text{data}) = \int f(t \mid \vartheta)g(\vartheta \mid \text{data})\,\mathrm{d}\vartheta$$

and predictive reliability

$$\tilde{R}(t \mid \text{data}) = \int R(t \mid \vartheta)g(\vartheta \mid \text{data})\,\mathrm{d}\vartheta.$$

The reliability and the density stand in the correct relationship to each other, that is, $\tilde{f} = -\frac{\mathrm{d}}{\mathrm{d}t}[\tilde{R}]$, because we can differentiate under the integral sign with respect to t. The model hazard rate is $h(t \mid \vartheta) = -\frac{\mathrm{d}}{\mathrm{d}t}\ln\{R(t \mid \vartheta)\}$, the hazard rate for the predictive distribution is

$$\tilde{h}(t \mid \text{data}) = -\frac{\mathrm{d}}{\mathrm{d}t}\ln\{\tilde{R}(t \mid \text{data})\},$$

and clearly

$$\tilde{h}(t \mid \text{data}) \neq \int h(t \mid \vartheta)g(\vartheta \mid \text{data})\,\mathrm{d}\vartheta$$

because

$$\frac{\mathrm{d}}{\mathrm{d}t}\ln\left\{\int R(t \mid \vartheta)g(\vartheta \mid \text{data})\,\mathrm{d}\vartheta\right\}$$
$$\neq \frac{\mathrm{d}}{\mathrm{d}t}\left\{\int \ln[R(t \mid \vartheta)]g(\vartheta \mid \text{data})\,\mathrm{d}\vartheta\right\}.$$

In short, the operator $\frac{\mathrm{d}}{\mathrm{d}t}$ which transforms a reliability function into a density is linear and commutes with integration in the present model, and the operator $\frac{\mathrm{d}}{\mathrm{d}t}\ln$ which transforms a reliability function into a hazard rate is nonlinear and does not commute with integration.

A choice has to be made for the hazard rate. However, since Bayesian formalism works completely naturally for the densities and the distributions, the most logical thing to do is to define the hazard rate from the current version of the distribution. In our formulation, the model has a hazard rate $h(t \mid \vartheta) = \frac{f(t, \mid \vartheta)}{R(t \mid \vartheta)}$. Each expert (labelled i) has a predictive hazard rate:

$$\bar{\mathscr{h}}_{ki} = \frac{\bar{\mathscr{f}}_{ki}}{\bar{\mathfrak{R}}_{ki}}.$$

and the overall predictive hazard rate is

$$\bar{\mathscr{h}}_j = \frac{\bar{\mathscr{f}}_j}{\bar{\mathfrak{R}}_j}$$

While this may be a shortcoming in terms of an elegant formalism it is clear that the only sensible thing to do is to work with the densities and reliability functions and then to calculate the appropriate hazard rates.

Lastly, it is useful in assessing the future to have an estimate of the remaining life of the component. In this case, we need to calculate the life expectancy at time t (also called the residual mean life) defined as

$$\mu(t) = \mathscr{E}[T \mid T > t] - t = \frac{1}{R(t)}\int_t^\infty R(u)\,\mathrm{d}u$$

With the predictive reliability available we have an estimate of the life expectancy as

$$\bar{\mu}(t) = (\bar{R}_j)^{-1}\sum_{t_i \geq t_j}\bar{R}_i\Delta t_i \quad \text{for } t_j \leq t \leq t_{j+1}$$

Updating the model
As new data become available the likelihood must be updated, and continuing with the notation above, write the likelihoods in a diagonal matrix, then the likelihood updated from time α to $\alpha+1$ is

$$L_{\alpha+1} = L' L_\alpha$$

the likelihood matrix for data accumulated up to time α multiplied by the likelihood matrix, L', of the new observations. While the data are accumulated the score vector can be updated as

$$s_{\alpha+1} = \mathcal{P} L_{\alpha+1} u_m$$

and the weight vector as

$$\mathbf{w}_{\alpha+1} = s^T_{\alpha+1}/|s_{\alpha+1}|, \quad \text{with } |s_{\alpha+1}| = u_n^T s_{\alpha+1}$$

The score vector provides feedback to the experts who must also periodically update their prior in the light of new data and in the light of known changes in design, maintenance or operating conditions. In this case the prior is updated from $\mathcal{P}$ to $\mathcal{P}'$ and the rest of the changes follow through naturally in imitation of the arguments above. The hope is that over a long period of time the experts will improve their judgment of priors by using the weights as a measure of performance, and so in the ideal steady state situation the weights should be equal, that is $w_\infty = u_n/n$.

Two parameter model

In this situation much of the neat formalism of the one parameter model is lost. However, it is clear that the construction of the predictive density and the reliability function weighted by the scores of the experts should follow the route given above. Suppose the model density is $f(t \mid \beta, \vartheta)$ and that we again work with discrete priors. Write the prior for expert e as

$$p^e_{ij} = probability[\beta \in I_i, \vartheta \in J_j]$$

where I_i and J_j denote the intervals for the parameters used in the elicitation process. Let l_{ij} be the likelihood evaluated for the i-th and j-th intervals of β and ϑ, then the posterior for expert e is

$$\pi^e_{ij} = \frac{l_{ij} p^e_{ij}}{s_e}$$

where $s_e = \sum_{i,j} l_{ij} p^e_{ij}$ is the score for expert e. In line with the one parameter model write:

Prior	$\mathcal{P}_e = (p^e_{ij})$
Likelihood	$L = (l_{ij})$
Scores	$s = (s_e)$
Unnormalised posterior	$\mathcal{P}^e_1 = \mathcal{P}_e \odot L$ where $\odot$ denotes a pointwise multiplication $z_{ij} = x_{ij} \bullet y_{ij}$
Normalised posterior	$\mathcal{P}^e_0 = \frac{\mathcal{P}^e_1}{s_e}$

In line with the above the appropriate posterior to use is the sum of the unnormalised posteriors scaled by the total score, $|s| = \sum_e s_e$, that is,

$$\bar{\mathcal{P}} = \frac{1}{|s|} \sum_e \mathcal{P}^e_1$$

The predictive density at time t is a double sum:

$$\bar{f}(t) = \sum_{i,j} \bar{\mathbf{P}}_{ij} \bullet f(t \mid \bar{\beta}_i, \bar{\vartheta}_j)$$

where $\bar{\mathcal{P}} = (\bar{\mathbf{P}}_{ij})$ and $\bar{\beta}_i$ and $\bar{\vartheta}_j$ are representative values $\bar{\beta}_i \in I_i$ and $\bar{\vartheta}_j \in J_j$. The other results follow easily by analogy with the one parameter model.

CASE STUDY

Here we report on the application of the above approaches to a reliability analysis of the critical heat exchangers. The study was carried out in three stages: the definition of risk groups, the elicitation of priors, and the consolidation into an overall reliability function.

Risk groups

Five risk groups were defined. The division into risk groups is, of course, arbitrary, but was based on extensive interviews with engineering, operational and maintenance management at the plant. The aim was to produce a division into groups, which can plausibly be regarded as statistically independent. The kind of criteria used to define the groups were:

(a) Is the cause internal to the plant or external to the plant?
(b) Is the cause controllable or not controllable?
(c) The cause is a function of intensity of use.

These gave rise to the five defined groups:

1. *Design and manufacturing faults.* For example, problems concerned with materials, and errors in manufacture. They are associated with early failure.
2. *Startup faults.* As the name implies, these faults occur immediately on or shortly after startup. Many human errors, such as a spanner left in the exchanger after maintenance, fall in this category.
3. *Wear and tear other than in pipes.* Failures of baffles and supports, etc.
4. *Wear and tear in the pipes.* Corrosion, leaking seals, erosion in bends. Careful consideration has to be given to the assignment of events as consequences of group 3 or group 4 causes.
5. *Incidental faults.* Consequences of events outside the subject of interest. For example, blockage by contamination coming from another part of the plant.

Groups 1, 2 and 5 refer to a pipe bundle within a heat exchanger, or the heat exchanger as a whole; groups 3 and 4 refer to individual pipes.

Elicitation

The elicitation process has been outlined above with respect to the one parameter model. Failure groups 3 and 4 were assumed to be described by three parameter Weibull distributions with shape parameter 2. A prior for the mean life within failure groups 3 and 4 was obtained in interviews with the experts, and there was no confusion in the minds of the interviewees about what was requested (they knew the difference between mean, median and mode).

For failure groups 1 and 2, a failure chance was given because this is so near to being a discrete part of the overall life distribution; this approach implies a memoryless discrete distribution with the chance of failure depending on the number of startup attempts in a given period. The discrete memoryless distribution is the geometric distribution and it is used to calculate the chance of a successful startup given the number of attempts in a period. After the transient period where the chance of a failure during a startup plays a role, the reliability function for groups 1 and 2 is taken as constant, in other words for the rest of the time it is only necessary to consider groups 3 to 5.[9] The effect of the discrete distributions would be seen as discontinuities in the overall reliability function.

For group 5, the failure time was assumed to be exponential because it seemed reasonable to regard it as a sum of external independent causes. In this case too an estimate of the mean time to failure was used, in keeping with the approach for groups 3 and 4.

The choice of a shape parameter $\beta = 2$ for the Weibull distributions was largely determined for pragmatic reasons. It was clear from what data were available that some aging was taking place in groups 3 and 4; Pitner[10] showed that a Weibull fitted well in an analysis of heat exchangers used in nuclear power stations; the value of 2 is conservative over the period up to the mean life. Lastly, since there was in some instances only one failure observed in these groups, we must take either an arbitrary value for some parameters, or use a prior distribution. The use of a prior for more than one parameter leads us into as yet unresolved difficulties of elicitation. One way to resolve the difficulties of elicitation is to attempt, through a mind experiment, to obtain a description of a whole distribution function, but then the Bayesian approach must also change somewhat, with the experts' distribution functions playing the rôle of the model distribution function.[11]

Example

In this study, there were 11 experts, called A to K. In a series of interviews the experts gave their estimates of the prior distribution of the mean life of a particular heat exchanger in the presence of only one of the defined causes of failure. In this early stage the priors for 11 critical heat exchangers have been obtained, and there are plans to extend the study to less critical heat exchangers and other components. To illustrate the application of the approach, an analysis for one of the critical heat exchangers is carried through. The example is very simplified, but serves to show the outputs analyses and outputs available from the model. The unit of time throughout is 1 month.

For failure groups 1 and 2, the mean of the experts' opinion of the failure chance per startup was taken, and for both groups this chance was assessed as 0·058. This corresponds to using a deterministic prior, as for group 5.

The model reliability function for groups 3 and 4 was assumed to be a Weibull:

$$R(t \mid \beta, \vartheta) = \exp\left\{ -\left(\frac{t-\tau}{\vartheta}\right)^{\beta} \right\}$$

with density

$$f(t \mid \beta, \vartheta) = \frac{\beta}{\vartheta}\left(\frac{t-\tau}{\vartheta}\right)^{\beta-1} \exp\left\{ -\left(\frac{t-\tau}{\vartheta}\right)^{\beta} \right\} \quad \text{and } \beta = 2$$

The priors for groups 3 and 4 are given in Tables 1 and 2. The intervals for the priors in both cases are defined by $a_1 = 24$, $a_2 = 48$, $a_3 = 72$, $a_4 = 96$, $a_5 = 120$ and $a_6 = 180$. The last value is rather arbitrary and is used because, in practice, the last interval cannot extend to infinity. Suppose that the available data consist of an observation period of 90 months, which was terminated by a failure associated with group 4. Thus, the likelihood for group 4 is the density function evaluated at the single time $t = 90$, and for group 3 it is the reliability function evaluated at time $t = 90$. The location parameter, τ, for group 3 is easily seen from Table 1 to be 24, and in group 4 it is zero.

Lastly, for group 5, the experts were asked to give a single estimate of the mean of an exponential

Table 1. Priors for Group 3 Causes

Expert	Interval					
	1	2	3	4	5	6
A	0·0	0·0	0·0	0·3	0·4	0·3
B	0·0	0·0	0·0	0·0	0·2	0·8
C	0·0	0·0	0·0	0·0	0·0	1·0
D	0·0	0·0	0·0	0·0	0·2	0·8
E	0·0	0·4	0·4	0·2	0·0	0·0
F	0·0	0·0	0·0	0·0	0·0	1·0
G	0·0	0·05	0·1	0·15	0·2	0·5
H	0·0	0·0	0·25	0·5	0·25	0·0
I	0·0	0·0	0·0	0·0	0·2	0·8
J	0·0	0·0	0·0	0·2	0·4	0·4
K	0·0	0·0	0·0	0·0	0·5	0·5

Table 2. Priors for Group 4 Causes

Expert	Interval					
	1	2	3	4	5	6
A	0·0	0·0	0·0	0·3	0·4	0·3
B	0·0	0·0	0·0	0·5	0·5	0·0
C	0·0	0·5	0·5	0·0	0·0	0·0
D	0·0	0·0	0·0	0·0	0·2	0·8
E	0·0	0·0	0·0	0·3	0·4	0·3
F	0·0	0·5	0·5	0·0	0·0	0·0
G	0·05	0·1	0·15	0·2	0·2	0·3
H	0·0	0·0	0·0	0·0	0·5	0·5
I	0·0	0·0	0·0	0·0	0·25	0·75
J	0·0	0·0	0·0	0·2	0·4	0·4
K	0·0	0·0	0·1	0·2	0·3	0·4

distribution, that is, a deterministic prior was used for the scale parameter of a Weibull distribution with shape parameter 1. This task was less demanding for the experts, but preserved the general approach and allows the priors to be updated later. Although the model is exponential in form, the predictive density is a mixture of exponentials with a nonconstant hazard rate. The chosen mean times between failures were 60, 96, 120, 60, 240, 60, 600, 120, 240, 60 and 120 months respectively. The lack of data meant that a likelihood function, uniform over the parameter values, was used.

In Fig. 1, the predictive reliability functions, density functions and hazard rates are shown. In Table 3 can be seen the failure groups 3 and 4; the lack of data shows in the even weighting given to the experts, the weights being generally close to 1/11. However, the more pessimistic estimates given by E and H show up in their lower weights. In Fig. 2, the overall predictive functions for groups 3–5 are shown. Figures 1 and 2 ignore the discrete part of the distribution, but this is easily applied as a correction to figures obtained from the graphs or by calculation. For example, the

Table 3. Weights for Experts

Expert	Group 3 weights	Group 4 weights
A	0·09	0·10
B	0·11	0·12
C	0·11	0·10
D	0·11	0·07
E	0·03	0·10
F	0·11	0·10
G	0·09	0·09
H	0·06	0·08
I	0·11	0·07
J	0·09	0·09
K	0·10	0·09

reliability after 72 months follows from the reliability of 30% evaluated for groups 3–5 and the reliability for the one initial startup and three (on average) other startups. The reliability after 72 months is

$$R(72) = 0{\cdot}30(1-p_1)(1-p_2)^3$$
$$= 0{\cdot}30 \times 0{\cdot}942 \times 0{\cdot}942^3 = 0{\cdot}16$$

a reliability of 16%. The hazard rate is affected additively by the terms $\ln(1-p_1)$ and $3\ln(1-p_2)$ and the mean residual life is unaffected as a predictor so long as the startups may be assumed to have happened in the past so that they cancel out in the conditioning required to assess the residual mean life.

It is also interesting to calculate the conditional reliability, given that the plant has been working without problem for a given time. In this example, the conditional reliability, given that the plant has run for 36 months without problem, is required. Because we assume that initial startups have all occurred in the past, the reliability function for group 1 remains constant at its value at time $t = 36$ and cancels out of the problem during the conditioning.

To approximate the behaviour of group 2 with failure probability of 0·058 per demand, an exponential

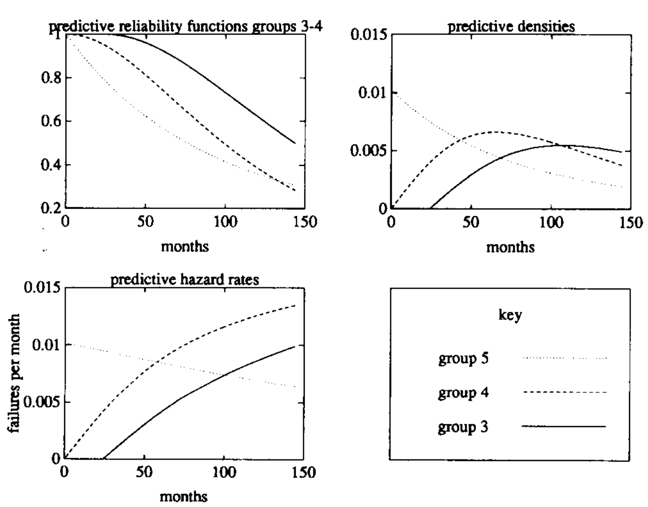

Fig. 1. Plots for groups 3–5.

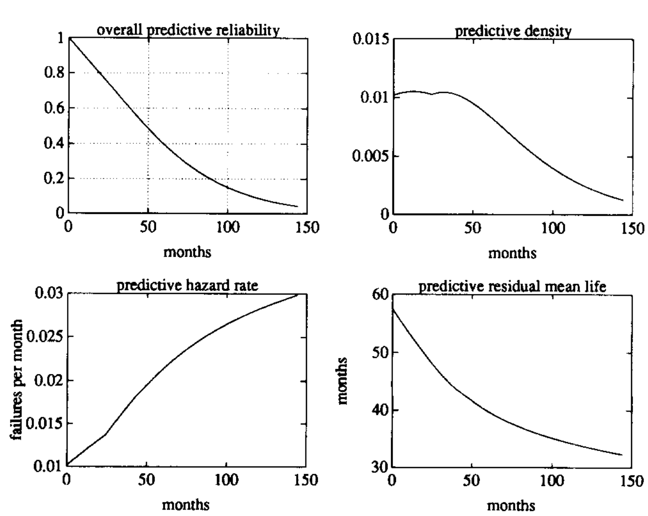

Fig. 2. Overall measures of reliability.

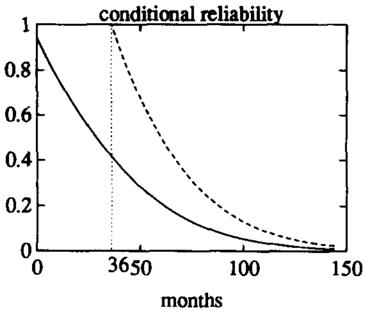

Fig. 3. Conditional reliability after 36 months.

distribution is used; there are approximately two startups per year and so the appropriate exponential has a mean time between failures of 6/0·058. The conditional reliability function is shown in Fig. 3, together with the unconditional reliability. The unconditional reliability jumps immediately from 1 to $1-p_1$ at time $t=0$ because the distribution for group 1 is discrete.[9] The residual mean life is unaffected by conditioning, and the hazard rate is reduced by the overall hazard rate at time $t=36$.

DISCUSSION

This paper has shown a simple and natural use of the Bayesian approach to handling some problems in reliability and maintenance planning in a chemical process plant. In particular, the use of the competing risks idea has allowed the incorporation of some explanatory data in the construction of an overall reliability function. The particular form of model chosen is also easy to implement and does not require extensive numerical algorithms[12] to evaluate the posterior, predictive and marginal densities required in many Bayesian analyses. In practical terms, the model described here is the kernel of a long term data analysis project in which the correspondence between model predictions and observed data will need to be analysed, and also one in which the experts will be asked to update their priors in the light of data and the weights assigned to their priors. Lastly, the presentation of information is through an interactive system on a personal computer and so allows failure patterns to be seen more easily through graphs like those used above than from tables of probabilities.

REFERENCES

1. Coolen, F. P. A., 'Bedrijfszekerheidsanalyse voor Warmtewisselaars' (Reliability Analysis of Heat Exchangers). MSc report, Eindhoven University of Technology, Nov. 1989.
2. Cox, D. R., Regression models and life tables (with discussion). *J. Royal Statistical Soc., Series B,* **34** (1972) 187–220.
3. Cox, D. R. & Oakes, D. A., *The Analysis of Survival Data.* Chapman and Hall, New York, 1984.
4. Lawless, J. F., *Statistical Models and Methods for Lifetime Data.* John Wiley, New York, 1982.
5. David, H. A. & Moeschberger, M. L., *The Theory of Competing Risks,* Griffin's Statistical Monographs, No. 39, London, 1978.
6. Zellner, A., *An Introduction to Bayesian Inference in Econometrics.* John Wiley, New York, 1971.
7. Cooke, R. M. *Expert Opinions in Safety Studies,* Vol. 4, *A Theory of Weights for Combining Expert Opinion,* Delft/Apeldoorn, Delft University of Technology/TNO, 1989.
8. Svenson, O., On expert judgments in safety analyses in the process industries. *Reliability Engineering and System Safety,* **25** (3) (1989) 219–56.
9. Follmann, D., Modelling failures of intermittently used machines. *Appl. Statistics,* **39** (1) (1990) 115–23.
10. Pitner, P., Statistical analysis of steam generator tube lifetime and a probabilistic method for tube bundle inspection. *Reliability Engineering and System Safety,* **21** (1988) 271–92.
11. Van Noortwijk, J. M., Dekker, R., Cooke, R. & Mazzuchi, T. A., Expert judgment in maintenance optimization. *IEEE Trans. Reliability* (submitted).
12. Smith, A. F. M., Skene, A. M., Shaw, J. E. H., Naylor, J. C. & Dransfield, M., The implementation of the Bayesian paradigm. *Communications in Statistical Theory and Methods,* **14** (5) (1985) 1079–102.