Annals of Nuclear Energy 122 (2018) 137–145

Contents lists available at ScienceDirect

Annals of Nuclear Energy

journal homepage: www.elsevier.com/locate/anucene

Analysis of time-dependent failure rate and probability of nuclear component

Haikuan Guo [a,b], Xinwen Zhao [a], Wenzhen Chen [a,*]

[a] Department of Nuclear Energy Science and Engineering, Naval University of Engineering, Wuhan 430033, China
[b] The 92609th of PLA, Bejing 100077, China

ARTICLE INFO

Article history:
Received 21 September 2017
Received in revised form 8 August 2018
Accepted 10 August 2018

Keywords:
Failure rate
Failure probability
Probabilistic safety assessment
Generalized linear model
Bayes

ABSTRACT

Reliability data of nuclear component are the input parameter and important to the nuclear power plant probabilistic safety assessment. If the failure is weariness in origin, the failure rate and probability will be not constant and have a time trend. Failure rate and probability have been proved to be increasing with time by reliability mathematics. The generalized linear model is built for Poisson and Binomial distribution and applied to study the time-dependent failure rate and probability of nuclear component. The model is proved to be reasonable by the qualitative graph and Bayesian chi-square statistic methods, which can predict the real time trend of nuclear component failure rate and probability correctly. Only analyzing the time dependent failure rate and probability of nuclear component, we will build more accurate general reliability data to analyze the uncertainty of system reliability.

1. Introduction

Reliability data of nuclear component are important to nuclear power plant probabilistic safety assessment (PSA) and have a basic significance to guarantee the quality and reasonable results of PSA model. If the component reliability data of the power plant are used directly, there may exist the problem of little sample space. If the general database is used, it may not reflect the component reliability characteristics of the power plant. In the past years, Bayesian method was employed widely to update the general database through "new evidences" observation failure data of component by the nuclear power plant PSA, and obtain suitable component reliability data for the nuclear power plant PSA model (Dana and Curtis, 2011; Atwood, 1996). Bayesian method not only eliminates the drawbacks that the general database can't represent characteristics of the power plant, but also makes up the defect of a little of certain data and a great uncertainty. So Bayesian method was suggested to determine the reliability parameters by American Nuclear Regulatory Commission (NRC) and American Society of Mechanical Engineering (ASME) for building the reliability general database (Eide et al., 2007; ASME/ANS RA-Sa-2009; Atwood et al., 2003).

Currently, Jeffreys prior has been as the non-informative prior method in processing the nuclear component reliability data. Antonio and Fabrizio (2004) introduced Bayesian method to the reliability analysis of complex repairable systems. Several authors built Jeffreys prior models of Binomial and Poisson distribution on the basis of constant failure rate and probability, but they didn't analyze the time trend of failure rate and probability (He and Zhang, 2013; Shen et al., 2014).

The working nuclear components will exceed the tolerance limit of material and fail due to the cumulative damage in the temperature, mechanical, water, and so on. The failure comes from the weariness, and failure rate and probability aren't constant and relate to time. It is necessary to study the time-dependent problem of nuclear component failure rate and probability, so that we can obtain the time trend of component failure rate and probability correctly. Tan et al. (2005) used the homogeneous Poisson process and Laplace test to analyze the time-dependent failure rates of large-size generating units. Dai et al. (2010) analyzed the variable incidence with negative Binomial distributions. At present, the published literature takes little account of the time-dependent failure rate and probability of nuclear component.

In this study, the inherent characteristics of nuclear component, in which the failure data obey Binomial and Poisson distributions, are analyzed profoundly. The corresponding Bayesian Generalized Linear Model (GLM) is built to analyze the time trend of component failure rate and probability, and two cases are introduced to validate the model correctness. The database is established to analyze the uncertainty of system reliability through time trend of

* Corresponding author.
E-mail address: cwz2@21cn.com (W. Chen).

https://doi.org/10.1016/j.anucene.2018.08.021
0306-4549/

component failure rate and probability, which also can predict the component failure probability correctly.

2. Necessity to study time trend for component failure rate and probability

Nuclear components will fail to work due to the damaged factors such as the impact, vibration, abrasion, corrosion and so on. Some failure data obey the Binomial distribution and their demanding failure probability relates to time, and others obey Poisson distribution and their operating failure rate relates to time. In this section, we analyze the necessity to study time trend for component failure rate and probability.

2.1. Time trend for component failure rate

When environmental stress impacts the nuclear component as Poisson distribution, the component will have three characteristics as follows:

1) Probability of k impacts has nothing to do with time initially and is approximately proportional to the length of the time interval. The constant of proportionality is denoted by lambda (λ).
2) The occurrence of an event in one time interval does not affect the probability of occurrence in another non-overlapping time interval.
3) The probability of simultaneous events in a short time interval is approximately zero.

It is assumed that the components impacted by above environmental stress have x_i failure number in operation time t_i, the probability will be given by the Poisson (λt_i) distribution as follows:

$$P(x_i|\lambda) = \frac{(\lambda t_i)^{x_i} e^{-\lambda t_i}}{x_i!}, \quad x_i = 0, 1, \cdots, \quad n, i = 0, 1, \cdots, m, \quad \lambda > 0 \tag{1}$$

where t_i is the operation time, x_i is the failure number, λ is the failure rate and i is the number of operation time. The reliability and failure function of components are given as follows, respectively

$$\begin{cases} R(k) = \sum\limits_{j=k}^{n} p_j \\ \lambda(k) = \frac{p_k}{\sum\limits_{j=k}^{n} p_j} \end{cases} \quad j = 0, 1, \cdots n \tag{2}$$

$r = \frac{p_{k+j}}{p_k} - \frac{p_{k+j+1}}{p_{k+1}}$, and the kind of λ is measured by r.

$$\frac{p_{k+j}}{p_k} - \frac{p_{k+j+1}}{p_{k+1}} \begin{cases} > 0, \text{ increasing } \lambda \\ = 0, \text{ constant } \lambda \\ < 0, \text{ decreasing } \lambda \end{cases} \tag{3}$$

The formula used to determine the kind of λ for Poisson distribution is given by

$$r = \frac{p_{k+j}}{p_k} - \frac{p_{k+j+1}}{p_{k+1}} = (\lambda t)^j \frac{jk!}{(k+j+1)!} > 0 \tag{4}$$

From Eq. (4) it is seen that the kind of λ for Poisson distribution is increasing, so it is necessary to analyze time trend of failure rate λ which can reflect time-dependent failure rate of component correctly.

2.2. Time trend for component failure probability

The Binomial distribution is often used to describe the demand failure of nuclear component, for example, a valve doesn't change the state in response to a demand. Components which failure data obey Binomial distribution have following characteristics:

1) There are two possible outcomes of each demand: success and failure, components can and can't change the state in response to a demand.
2) There is a constant probability (p) of failure/success on each demand.
3) The result of each demand is independent, the outcomes of earlier demands do not influence that of later demands (i.e., the order of failures/successes is irrelevant).

Assuming the failure probability is p and the success probability is 1-p, there are n demands for components in an experiment and m experiments, random variable X is the failure number, $x = (x_1, \cdots, x_n)$ is the sample of X with Binomial distribution and its distribution function is given as follows:

$$\begin{aligned} p_{x_i} &= P(x_i|p) & x_i &= 0, 1, \cdots, n \\ &= C_n^{x_i} p^{x_i} (1-p)^{n-x_i} & i &= 1, 2, \cdots, m \end{aligned} \tag{5}$$

where i is the experiment number, x_i is the failure number in ith experiment, $C_n^{x_i}$ is the Binomial coefficient and p_{x_i} is the probability with x_i in n demands. The reliability and failure function of components are given by Eq. (2), $r = \frac{p_{k+j}}{p_k} - \frac{p_{k+j+1}}{p_{k+1}}$, the kind of failure probability p is measured by r.

$$\frac{p_{k+j}}{p_k} - \frac{p_{k+j+1}}{p_{k+1}} \begin{cases} > 0, \text{ increasing } p \\ = 0, \text{ constant } p \\ < 0, \text{ decreasing } p \end{cases} \tag{6}$$

The formula used to determine the kind of p for Binomial distribution is given by

$$\begin{aligned} r &= \frac{p_{k+j}}{p_k} - \frac{p_{k+j+1}}{p_{k+1}} \\ &= \frac{(k+1)!(n-k)!p^j}{(k+j+1)!(n-k-j)!(1-p)^j} \left[\frac{k+j+1}{k+1} - \frac{n-k-j}{n-k} \right] > 0 \end{aligned} \tag{7}$$

From Eq. (7) it is seen that the kind of p for Binomial distribution is increasing, so it is necessary to analyze time trend of failure probability p.

3. Time-dependent failure rate of nuclear component

Time-dependent failure rate of nuclear component which failure data obey Poisson distribution is analyzed in this section.

3.1. Jeffreys prior model: Poisson-Gamma model

Jeffreys prior model of Poisson-Gamma is obtained based on Bayes theorem. Second derivative of logarithmic likelihood function for Eq. (1) is given by

$$\frac{d^2 l(\lambda|x_i)}{d\lambda^2} = \frac{d^2 \sum_{i=1}^{n} \ln(P(x_i|\lambda))}{d\lambda^2} = -\frac{1}{\lambda^2} \sum_{i=1}^{n} x_i \tag{8}$$

where $l(\lambda|x_i)$ is logarithmic likelihood function of Poisson distribution. According to Eqs. (1) and (8), Fisher information matrix of parameter λ for Poisson distribution is

$$I(\lambda) = E\left(-\frac{d^2 l(\lambda|x_i)}{d\lambda^2}\right) = E\left(\frac{1}{\lambda^2} \sum_{i=1}^{n} x_i\right) = \frac{nt_i}{\lambda} \tag{9}$$

where $I(\lambda)$ is Fisher information matrix of λ, so the Jeffreys non-informative prior of λ is

$$\pi(\lambda) = (\det I(\lambda))^{1/2} = \left| E\left(-\frac{d^2 l(\lambda)}{d\lambda^2}\right)\right|^{1/2} = \left(\frac{nt_i}{\lambda}\right)1/2 \tag{10}$$

where $\pi(\lambda)$ is the Jeffreys non-informative prior of λ, that is

$$\pi(\lambda) \propto \frac{1}{\lambda^{1/2}} \tag{11}$$

So the Jeffreys non-informative prior of failure rate λ for component failure data obeying Poisson distribution is Gamma(0.5, 0). According to Eq. (11) and Bayes theorem, the posterior distribution of λ is

$$\pi(\lambda|x_i) = \frac{p(x_i|\lambda)\pi(\lambda)}{\int_\Theta p(x_i|\lambda)\pi(\lambda)d\lambda} = \frac{\frac{(\lambda t_i)^{x_i} e^{-\lambda t_i}}{x_i!}\frac{1}{\lambda^{1/2}}}{\int_0^{+\infty}\frac{(\lambda t_i)^{x_i} e^{-\lambda t_i}}{x_i!}\frac{1}{\lambda^{1/2}}d\lambda} = \frac{t_i^{x_i+\frac{1}{2}}}{\Gamma\left(x_i+\frac{1}{2}\right)}\lambda^{x_i-\frac{1}{2}}e^{-\lambda t_i} \tag{12}$$

where $\pi(\lambda|x_i)$ is the posterior distribution of λ for Poisson distribution, $\Gamma(\cdot)$ is Γ function, Θ is the value space of λ. Eq. (12) is the density function of Gamma$(x + 0.5, t)$, so $\pi(\lambda|x_i)$ obeys Gamma$(x + 0.5, t)$, that is

$$\begin{cases} \alpha = x + 0.5 \\ \beta = t \end{cases} \tag{13}$$

where t is the operation time of component, x is the failure number of component in t.

3.2. Failure rate of component based on Poisson-Gamma model

Consider the following virtual data in Table 1 from a particular component in a Chinese reactor. The Jeffreys non-informative prior model is used to analyze the failure rate of component.

As we do for the above example, we begin with a simple graphical check and observe whether λ has an obvious time trend. We plot the 95% posterior credible intervals for λ in each time bin based on uprating Jeffreys prior for λ with Poisson distribution by the failure data from Table 1, the caterpillar curves are seen in Fig. 1. The OpenBUGS script for this model is shown in Table 2.

Fig. 1 does not suggest an obvious time trend for λ. However, it does suggest "random" variability for λ across the bins. So it is necessary to analyze the failure rate from Table 1 by the additional graphical and quantitative check.

The posterior predictive distribution will be adopted in our model check, which is the predictive distribution for the future values of an observed random variable under the conditions of given past empirical data. It is defined as follows:

$$\pi(x_{\text{pred}}|x_{\text{obs}}) = \int_\Theta f(x_{\text{pred}}|\theta)\pi_1(\theta|x_{\text{obs}})d\theta \tag{14}$$

In this equation, x_{obs} is the observed data, x_{pred} is the predicted data, θ is the parameter of the aleatory model that generates the observed data x_{obs}. The schematic diagram of posterior predictive distribution is shown in Fig. 2.

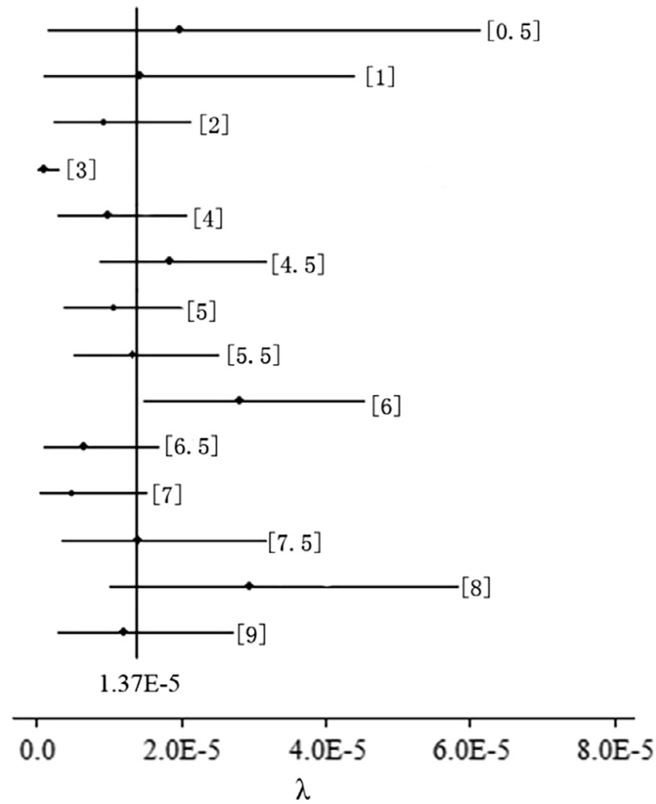

Fig. 1. 95% Posterior credible intervals for λ, based on updating Jeffreys prior by data from Table 1.

In Fig. 2, we observe that x_{obs} updates our prior distribution for θ. Then the updated (i.e., posterior) distribution of θ is used in the aleatory model to generate predicted data x_{pred}. Posterior predictive average probability for Jeffreys prior model is 0.518 and 95% posterior credible interval is (0.286, 0.714). The low posterior predictive probability indicates that this Jeffreys prior model may be inadequate. The check of replicated data from Table 1 for Jeffreys prior model based on posterior predictive distribution is shown in Fig. 3, and the corresponding OpenBUGS script is shown in Table 3.

In Fig. 3, we find that failure numbers 1 and 12 of observed data fall outside the 95% intervals of the replicated data and failure number 5 is on the border, which indicates that Jeffreys prior model with constant λ can't replicate observed data completely.

It isn't enough to check the model ability to replicate observed data only through the graph, a Bayesian chi-square statistic for the count data should be used to check in the quantitative way. In frequentist statistics, a commonly encountered test statistic for the count data is

$$\chi^2 = \sum_i \frac{(y_i - u_i)^2}{\sigma_i^2} \tag{15}$$

Table 1
Operational failure data of component.

Time/y	Failures	Exposure time/h	Time/y	Failures	Exposure time/h
0.5	1	75,936	5.5	6	489,818
1	1	105,924	6	12	446,222
2	3	375,376	6.5	2	378,086
3	1	1,460,112	7	1	306,259
4	4	454,908	7.5	3	249,852
4.5	9	515,112	8	5	186,864
5	5	519,252	9	3	290,800

Table 2
OpenBUGS script for graphical check with Poisson likelihood and Jeffreys prior.

```
model {
  for(i in 1:N) {
  x[i]~dpois(mu[i])                          # Poisson dist. For number of failures in each source
  Mu[i]<-lambda[i]*ExpTime[i]                # Parameter of Poisson distribution
  lambda[i]~dgamma(alpha,beta)               # Jeffreys prior for lambda
  }
}
data
list(x = c(1,1,3,1,4,9,5,6,12,2,1,3,5,3), alpha=0.5, beta=0.0001, N=14, ExpTime = c(75936, 105924, 375376, 1460112,
   454908, 515112, 519252, 489818, 446222, 378086, 306259, 249852, 186864, 290800))
```

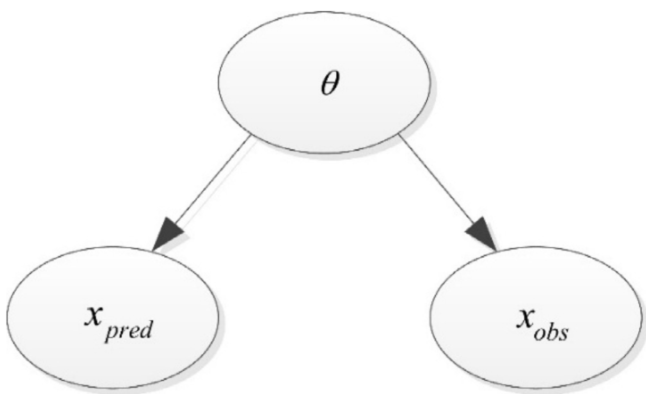

Fig. 2. Schematic diagram of posterior predictive distribution.

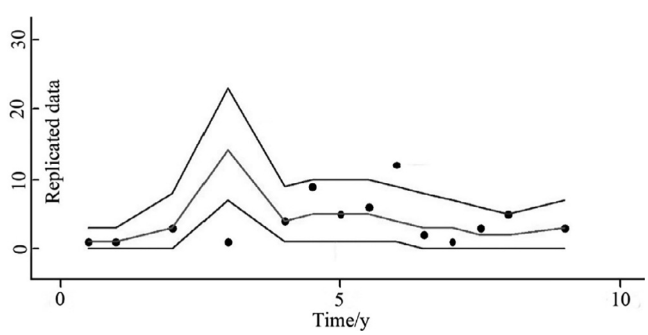

Fig. 3. 95% Posterior predictive intervals for replicated data in each time bin for Jeffreys prior model.

In this equation, y_i is the i-th observed value, μ_i is the i-th expected or mean value, and σ_i^2 is the i-th variance. The distribution of χ^2 is often approximately chi-square with degrees of freedom related to the sample size. We use the observed values of X to form the statistic:

$$\chi^2_{\text{obs}} = \sum_i \frac{(x_{\text{obs},i} - u_i)^2}{\sigma_i^2} \tag{16}$$

where $x_{\text{obs},i}$ is the i-th observed value. Then we generate the predicted values of X from its posterior predictive distributions, and construct an analogous statistic as follows:

$$\chi^2_{\text{rep}} = \sum_i \frac{(x_{\text{rep},i} - u_i)^2}{\sigma_i^2} \tag{17}$$

where $x_{rep,i}$ is the i-th replicated value.

We define the Bayesian p value as $P\left(\chi^2_{rep} \geqslant \chi^2_{obs}\right)$. However, instead of choosing an arbitrary p value (e.g., 0.05) and rejecting a model with a p value below this arbitrary cutoff, we will use the p value to select the model with the best replicated data. This will result in the model with Bayesian p value closest to 0.5, which is the value one will obtain if the distributions of the observed and replicated test statistics overlap perfectly. The Bayesian p value (using the chi-square summary statistic) for Jeffreys prior model is 0.004 far below the ideal value of 0.5 (see Fig. 3).

In summary, Jeffreys prior model with constant λ can't analyze the time trend of failure rate from Table 1.

3.3. Time-dependent failure rate based on GLM

A GLM that can incorporate a time trend in λ should be explored due to drawbacks of Jeffreys prior model with constant λ. A link function is used in the GLM to transform the interest parameter

Table 3
OpenBUGS script for replicated data with Poisson likelihood and Jeffreys prior.

```
model {
  for(i in 1:N) {
  x.obs[i]~dpois(mu[i])                      # Poisson dist. for number of failures in each source
  x.rep[i]~dpois(mu[i])                      # Replicate value from post. predictive distribution
  mu[i]<-lambda*ExpTime[i]                   # Parameter of Poisson distribution
  # Model validation
  diff.obs[i]<-pow(x.obs[i]-mu[i],2)/mu[i]
  diff.rep[i]<-pow(x.rep[i]-mu[i],2)/mu[i]
  }
  # Bayesian p-value calculation
  chisq.obs<-sum(diff.obs[])
  chisq.rep<-sum(diff.rep[])
  p.value<-step(chisq.rep-chisq.obs)         # Monitor mean value, should be near 0.5
  lambda~dgamma(alpha,beta)                  # Jeffreys prior for lambda
}
data
list(x.obs = c(1,1,3,1,4,9,5,6,12,2,1,3,5,3), N = 14, alpha = 0.5, beta = 0.0001, ExpTime = c(75936, 105924, 375376,
   1460112, 454908, 515112, 519252, 489818, 446222, 378086, 306259, 249852, 186864, 290800))
```

Table 4
OpenBUGS script for log-linear model for λ with extra-Poisson variation in each age period.

```
model {
  for(i in 1:N) {
  log(lambda[i])<-a+b*age[i]+eps[i]          # Loglinear link function with extra variation for lambda
  eps[i]~dnorm(0,tau.eps)                    # Normal distribution for extra variation
  mu[i]<-lambda[i]*ExpTime[i]                # Parameter of Poisson distribution
  x.obs[i]~dpois(mu[i])                      # Poisson dist. for number of failures in each source
  # Replicate times from posterior predictive distribution
  x.rep[i]~dpois(mu[i])                      # Replicate value from post. predictive distribution
  diff.obs[i]<-pow(x.obs[i]-mu[i],2)/mu[i]
  diff.rep[i]<-pow(x.rep[i]-mu[i],2)/mu[i]
  }
  # Model validation
  log(lambda [N+1])<-a + b*10 + eps[N]       # Used to predict lambda in 10th year
  chisq.obs<-sum(diff.obs[])
  chisq.rep<-sum(diff.rep[])
  p.value<-step(chisq.rep-chisq.obs)         # Mean of this node should be near 0.5
  # Prior distributions
  a~dflat()                                  # Diffuse priors for a and b
  b~dflat()
  tau.eps<-pow(sigma.eps, −2)
  sigma.eps~dunif(0,10)
}
Inits
list(a = −1, b = 0.1)
list(a = 1, b = 0)
data
list(x.obs = c(1,1,3,1,4,9,5,6,12,2,1,3,5,3), N = 14, age = c(0.5,1,2,3,4,4.5,5,5.5,6,6.5,7,7.5,8,90), ExpTime = c(75936, 105924,
   375376, 1460112, 454908, 515112, 519252, 489818, 446222, 378086, 306259, 249852, 186864, 290800))
```

to a related parameter over the entire real axis. The standard link function is the natural logarithm of λ, which has a monotonic trend and additional random variability over time (Atwood et al., 2003). The GLM of λ becomes

$$\log(\lambda_i) = a + bt_i + \varepsilon_i \tag{18}$$

In this equation, ε_i is a random error term in each age period, which we will take to be normally distributed with mean zero and constant (unknown) variance in each period. In each age bin of this model, the distribution of λ is conditional upon a value of error ε_i that describes the random year-to-year variation in $\log(\lambda)$ about a straight line. But ε_i is uncertain, so we model this uncertainty by introducing a prior distribution for the standard deviation of ε_i, which is assumed to be constant over time. We place the prior on the standard deviation in each age period, which is the square root of the variance. Note that $b = 0$ in this model corresponds to a constant λ. If λ increases (decreases) with time, we will have $b > (<)0$ The OpenBUGS script for this model is shown in Table 4. To focus attention on the aleatory model, we use the independent improper flat priors over the real axis for a and b. OpenBUGS refers to this distribution as dflat(). Examining the posterior density for b will help us judge the significance of any trend for λ that might be present: if the marginal posterior distribution of b is mostly in the right (left) of zero, it indicates an increasing (decreasing) trend for λ. By monitoring the b node, we obtain a posterior mean of b with 0.118. Fig. 4 shows the plot of the marginal posterior distribution for b. The most samples of b are greater than zero, which indicates a statistically significant increasing trend in λ.

The check of the replication of observed data from Table 1 for GLM based on posterior predictive distribution is shown in Fig. 5.

In Fig. 5, we find that all the observed data fall in the 95% intervals of replicated data, which illustrates the improved predictive ability of GLM. The Bayesian p value (using the chi-square summary statistic) for GLM is 0.429 and close to the ideal value of 0.5 (see Fig. 5), which also indicates that the GLM has a well predictive ability.

3.4. Failure rates predicted by two models

We find that the GLM can replicate all the observed data, while the Jeffreys prior model with constant λ can't obtain the replicated results by qualitative and quantitative methods. The Bayesian p value of 0.429 for GLM is closer to the ideal value of 0.5 than the Bayesian p value of 0.004 for Jeffreys prior model. The GLM has a

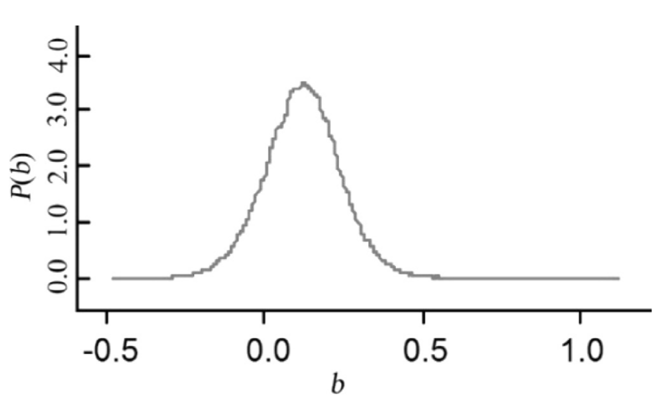

Fig. 4. Marginal posterior density for b.

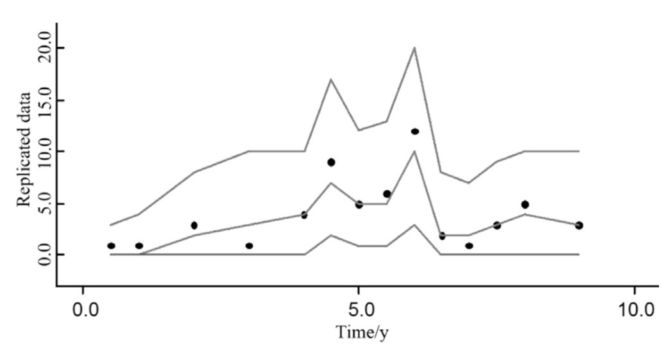

Fig. 5. 95% Credible intervals for predicted failure data under GLM.

Table 5
Failure rate of 10th year predicted by two models.

Model	Mean	Standard Deviation	Error	2.5%th	Median	97.5th
GLM	1.394×10^{-5}	7.257×10^{-6}	5.947×10^{-8}	3.763×10^{-6}	1.265×10^{-5}	3.16×10^{-5}
Jeffreys prior	9.64×10^{-6}	1.283×10^{-6}	3.872×10^{-9}	7.286×10^{-6}	9.584×10^{-6}	1.231×10^{-5}

Table 6
Failure data of MOV.

Time/y	Failures	Demands	Time/y	Failures	Demands
1	4	300	6	3	300
2	2	300	7	4	300
3	3	300	8	9	300
4	1	300	9	6	300
5	4	300			

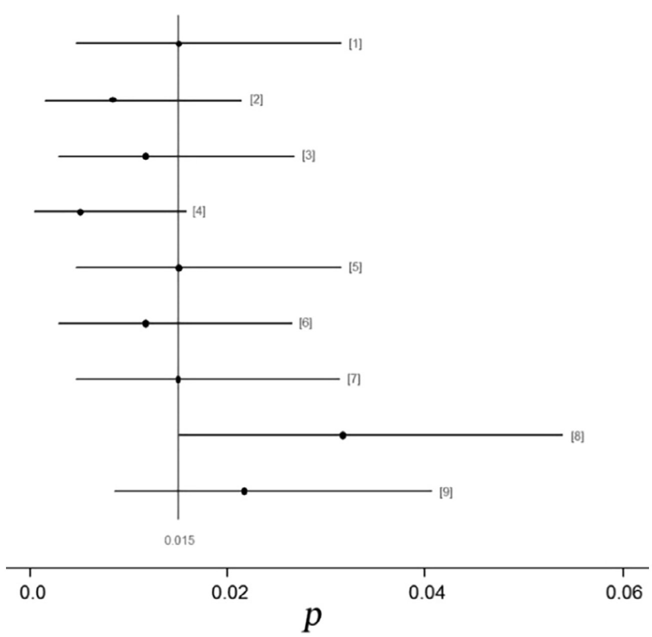

Fig. 6. 95% Posterior credible intervals for p, based on updating Jeffreys prior by data from Table 6.

Table 7
OpenBUGS script for graphical check with Binomial likelihood and Jeffreys prior.

```
model {
  for(i in 1:N) {
  x[i]~dbin(p[i],n[i])        # Binomial distribution for failures in each year
  p[i]~dbeta(alpha,beta       # Jeffreys prior for p
  }
}
Data
list(x = c(4,2,3,1,4,3,4,9,6),n=c(300,300,300,300,300,300,300,300,300),
   alpha = 0.5, beta = 0.5, N= 9 )
```

better replicated ability and reflects the time trend of component failure rate better, so it is suitable to analyze the time-dependent failure rate of nuclear component from Table 1. The failure rate of 10th year is predicted by two models respectively, the results are shown in Table 5.

According to Table 5, we know that the failure rate of 10th year predicted by GLM is close to 1.5 times of Jeffreys prior model, the reason is that the failure rate λ increases with time. So the Jeffreys prior model underestimates the failure rate of component and the failure rate predicted by GLM is closer to the real failure rate of component. The GLM with additional variability in λ should be applied to study the time trend of component failure rate instead of Jeffreys prior model with constant λ.

4. Time-dependent failure probability of nuclear component

Time-dependent failure probability of nuclear component which failure data obey Binomial distribution is analyzed in this section.

4.1. Jeffreys prior model: Beta-Binomial model

Jeffreys prior model of Beta-Binomial is educed based on Bayes theorem. Second derivative of logarithmic likelihood function for Eq. (5) is given by

$$\frac{\mathrm{d}^2 l(p|x)}{\mathrm{d}p^2} = -\frac{\sum_{i=1}^{n} x_i}{p^2} - \frac{n - \sum_{i=1}^{n} x_i}{(1-p)^2} \tag{19}$$

where $l(p|x)$ is logarithmic likelihood function of Binomial distribution. According to Eqs. (5) and (19), Fisher information matrix of parameter p for Binomial distribution is

$$I(p) = E\left(-\frac{\mathrm{d}^2 l(p|x)}{\mathrm{d}p^2}\right) = \frac{n}{p} + \frac{n}{1-p} \tag{20}$$

where $I(p)$ is Fisher information matrix of p, so the Jeffreys non-informative prior of p is

$$\pi(p) = |I(p)|^{1/2} = \sqrt{\frac{n}{p} + \frac{n}{1-p}} = \sqrt{\frac{n}{p(1-p)}} \tag{21}$$

where $\pi(p)$ is the Jeffreys non-informative prior of p, that is

$$\pi(p) \infty p^{-1/2}(1-p)^{-1/2} \tag{22}$$

So the Jeffreys non-informative prior of failure probability p for component failure data obeying Binomial distribution is $\mathrm{Beta}(0.5, 0.5)$. According to Eq. (22) and Bayes theorem, the posterior distribution of p is

$$\begin{aligned}\pi(p|x) &= \frac{p(x|p)\pi(p)}{\int_{\Omega} p(x|p)\pi(p)dp} \\ &= \frac{C_n^x p^x (1-p)^{n-x} p^{-1/2}(1-p)^{-1/2}}{\int_0^1 C_n^x p^x (1-p)^{n-x} p^{-1/2}(1-p)^{-1/2} dp} \\ &= \frac{\Gamma(n+1)}{\Gamma(x+\frac{1}{2})\Gamma(n-x+1/2)} p^{-1/2}(1-p)^{-1/2}\end{aligned} \tag{23}$$

where $\pi(p|x)$ is the posterior distribution of p for Binomial distribution, $\Gamma(\cdot)$ is Γ function, Θ is the value space of p. Eq. (23) is the density function of Beta distribution, so $\pi(p|x)$ obeys $\mathrm{Beta}(x + 0.5, n - x + 0.5)$, that is

$$\begin{cases} \alpha = x + 0.5 \\ \beta = n - x + 0.5 \end{cases} \tag{24}$$

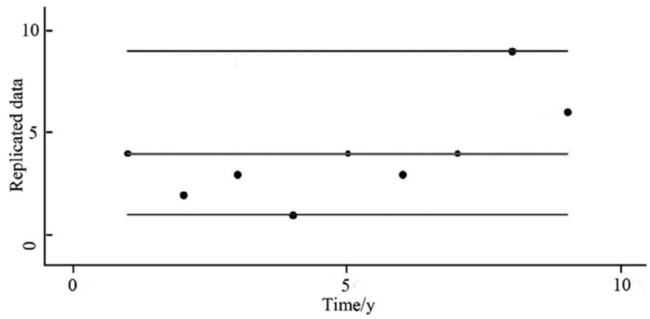

Fig. 7. 95% Credible intervals for posterior predicted valve leakage events, based on Binomial distribution with constant p in each year.

where x is failure number of component, n is the demand number for component. n-x is the success number in n demands for component.

4.2. Failure probability of component based on Beta-Binomial model

The data were originally collected for a study of Motor Operated Valve (MOV) failures, which includes the operating experience of altogether 300 valves in safety systems for a time period of nine years. The data in this example consist of leakage test results of 300 valves at one plant unit, i.e. the annual number of failure in 300 yearly valve tests, as shown in Table 6.

We first carry out a qualitative check to see if there is any systematic time trend in p. We update the Jeffreys prior model for the Binomial distribution with the data from Table 6 for each year, and plot the interval estimates of p, as seen in Fig. 6. The OpenBUGS script for this model is shown in Table 7.

Fig. 6 shows an increasing trend in p with time, but significant uncertainty in the estimates for each individual year clouds this conclusion. Therefore, additional graphical and quantitative checks are needed.

Posterior predictive average probability for Jeffreys prior model is 0.595 and 95% posterior credible interval is (0.333, 0.889). Fig. 7 gives the plot of the replicated event count for a Binomial model with constant p, which indicates that the Jeffreys prior model has some difficulty in replicating the variability of the observed data from Table 6. The Bayesian p value for the Binomial distribution with constant p is 0.219, far from the ideal value of 0.5. The corresponding OpenBUGS script is shown in Table 8.

Table 9
OpenBUGS script for logistic time trend in p.

```
model {
  for(i in 1:N){
  x[i]~dbin(p[i],n[i])                          # Binomial distribution for failures in
                                                each year
  logit(p[i])<-a + b*i                          # Logit link function for p
  # Model validation
  x.rep[i]~dbin(p[i],n[i])                      # Replicate value from post. predictive
                                                distribution
  diff.obs[i]<-pow(x[i]-n[i]*p[i],2)/
   (n[i]*p[i])
  diff.rep[i]<-pow(x.rep[i]-n[i]*p
   [i],2)/(n[i]*p[i])
  }
  logit(p[N+1])<-a+b*(N+1)                      # Used to predict p in 10th year
  # Bayesian p-value calculation
  chisq.obs<-sum(diff.obs[])
  chisq.rep<-sum(diff.rep[])
  p.value<-step(chisq.rep-chisq.                # Mean of this node should be near 0.5
   obs)
  a~dflat()                                     # Diffuse priors for a and b
  b~dflat()
}
Inits
list(a = 5,b = −0.1)
list(a = −5,b = 0.1)
data
list(x = c(4,2,3,1,4,3,4,9,6), n = c(300,300,300,300,300,300,300,300,300), N = 9)
```

4.3. Time-dependent failure probability based on GLM

The GLM is often used to model a monotonic time trend in p. The link functions for p is the logit function (Atwood et al., 2003). In this model, logit (p) is the log of the odds ratio and defined as a linear function of time as follows:

$$\log\left(\frac{p}{1-p}\right) = a + bt \tag{25}$$

Note that $b = 0$ in this model corresponds to a contant p. If p increases (decreases) with time, then we will have $b > (<) 0$. The OpenBUGS script for this model is shown in Table 9. To focus attention on the aleatory model, we use the independent improper flat priors over the real axis for a and b. Examining the posterior density for b will help us judge the significance of any trend for p that might be present: if the marginal posterior distribution of b is mostly in the right (left) of zero, it indicates an increasing (decreasing) trend for p. By monitoring the b node, we obtain that the posterior mean of b is 0.146 and 95% posterior credible interval is (0.015, 0.283). So we

Table 8
OpenBUGS script for replicated data with Binomial likelihood and Jeffreys prior.

```
model {
  for(i in 1:N) {
  x[i]~dbin(p,n[i]                              # Binomial distribution for failures in each year
  # Model validation
  x.rep[i]~dbin(p,n[i])                         # Replicate value from post. predictive distribution
  diff.obs[i]<-pow(x[i]-n[i]*p,2)/(n[i]*p)
  diff.rep[i]<-pow(x.rep[i]-n[i]*p,2)/(n[i]*p)
  }
  # Bayesian p-value calculation
  chisq.obs<-sum(diff.obs[])
  chisq.rep<-sum(diff.rep[])
  p.value<-step(chisq.rep-chisq.obs)            # Monitor mean value, should be near 0.5
  p~dbeta(alpha,beta)                           # Jeffreys prior for p
}
data
list(x = c(4,2,3,1,4,3,4,9,6),n=c(300,300,300,300,300,300,300,300,300),alpha = 0.5,beta = 0.5,N = 9)
```

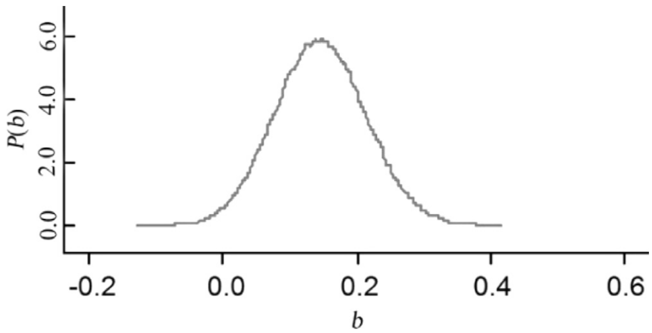

Fig. 8. Marginal posterior density for *b*.

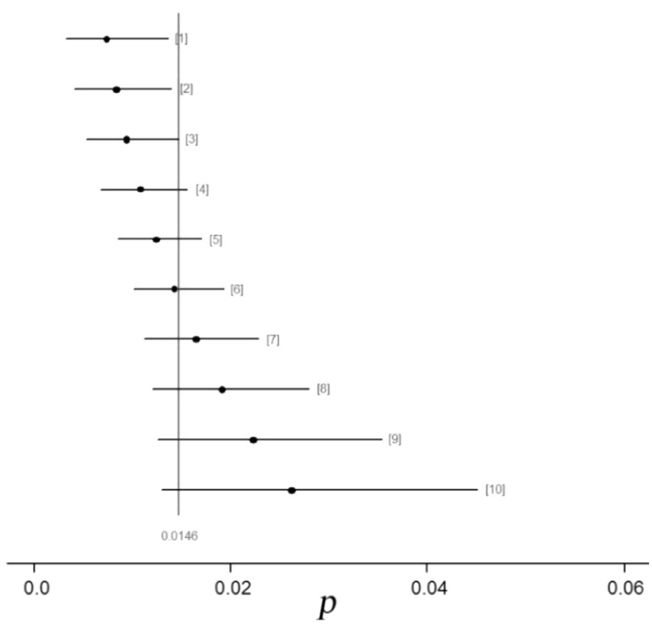

Fig. 9. 95% Posterior credible intervals for p, based on updating GLM by data from Table 6.

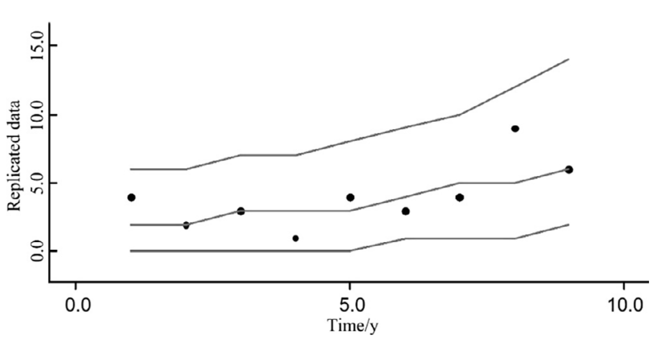

Fig. 10. 95% Credible intervals for predicted failure data, based on GLM with logistic time trend in p.

infer that the posterior probability of at least 0.975 for $b > 0$ indicates a statistically significant increasing trend in p. Fig. 8 shows the plot of the marginal posterior distribution for b.

Fig. 9 gives the plot of 95% posterior credible intervals for p in each time bin based on uprating GLM with failure data from Table 6.

Fig. 9 indicates that there is an increasing trend in p with time. We next check the ability for GLM to replicate the observed data. The posterior predictive plot in Fig. 10 indicates that the model with a logistic time trend in p is better than the Jeffreys prior model without time trend. The Bayesian p value for GLM is 0.505 and much closer to the ideal value of 0.5 than the value of 0.219 for Jeffreys prior model with constant p.

4.4. Failure probabilities predicted by two models

We find that the GLM can replicate all the observed data, while the Jeffreys prior model with constant p can't replicate that by qualitative and quantitative methods. The GLM has a better replicated ability and well reflects the time trend of component failure rate, so it is suitable to analyze the time-dependent failure probability of nuclear component from Table 6. The failure probability of 10th year is predicted by two models respectively, the results are shown in Table 10.

According to Table 10, we know that the failure probability of 10th year predicted by GLM is close to 2 times of Jeffreys prior model, the reason is that the failure probability p increases with time. So the Jeffreys prior model underestimates failure probability of component and the failure probability predicted by GLM is closer to the real value. The GLM with a logistic time trend in p should be applied to study the time trend of component failure probability in comparison with Jeffreys prior model with constant p.

5. Conclusions

The time trend of component failure rate and probability is analyzed. Nuclear components occur failure because of damaged factors such as the impact, vibration, abrasion, corrosion and so on. The failure rate and probability aren't constant and have a time trend. The failure rate and probability are proved to be increasing with time by reliability mathematics in this paper, so it is necessary to study the time-dependent failure rate and probability of nuclear component. It is proved by two cases that Jeffreys prior model is limited in the analysis of the time trend of component failure rate and probability. The GLM with Poisson and Binomial distribution can be used to study the time-dependent failure rate and probability of nuclear component. It is also proved that the GLM can replicate all the observed data by the qualitative graphical and quantitative Bayesian chi-square statistic methods.

The present research results illustrate that the GLM has a well predictive ability and is suitable to analyze the time trend of nuclear component failure rate and probability correctly. These GLM characteristics can be used to build more accurate general reliability data to improve PSA quality and make its results more reasonable.

Table 10
Failure probability of 10th year predicted by two models.

Model	Mean	Standard Deviation	Error	2.5%th	Median	97.5th
GLM	2.606×10^{-2}	8.267×10^{-3}	2.844×10^{-5}	1.292×10^{-2}	2.501×10^{-2}	4.506×10^{-2}
Jeffreys prior	1.351×10^{-2}	2.223×10^{-3}	7.52×10^{-6}	9.513×10^{-3}	1.339×10^{-2}	1.816×10^{-2}

Acknowledgements

This research was financially supported by Foundation of Science and Technology on Reactor System Design Technology Key Laboratory (No. HT-JXYY-02-2014002).

References

Antonio, P., Fabrizio, R., 2004. Bayesian reliability analysis of complex repairable systems. Appl. Stochastic Models Bus Ind. 20, 253–264.

ASME/ANS RA-Sa-2009. Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications. 2009.

Atwood, C.L., 1996. Constrained non-informative priors in risk assessment. Reliab. Eng. Syst. Saf. 53 (1), 37–116.

Atwood, C.L., LaChance, J.L., Martz, H.F., et al., wood et al., NUREG/CR-6823. 2003.. Handbook of parameter estimation for probabilistic risk assessment. NUREG/CR-6823.

Dai, C., Chen, F.Y., Tian, M.Z., 2010. Bayesian inference for infectious incidence obeying negative Binomial distribution. Stat. Decis. 6, 7–9 (in Chinese).

Dana, K., Curtis, S., 2011. Bayesian Inference for Probabilistic Risk Assessment. Springer, London, pp. 6–10.

Eide, S.A., Wierman, T.E., Gentillon, C.D., et al. Industry Average Performance for Components and Initiating Events at U.S. Commercial Nuclear Power Plants NUREG/CR-6928. 2007.

He, J., Zhang, B.B., 2013. Calculation of non-informative prior of reliability parameter and initiating event frequency with Jeffreys method. At. Energy Sci. Technol. 47 (11), 2059–2062 (in Chinese).

Shen, Z.Y., Chen, W., Yuan, J.X., Tang, X.H., Yang, J., 2014. Bayesian method of PSA generic data processing based on Jeffreys prior. Nucl. Power Eng. 35 (6), 84–87 (in Chinese).

Tan, F.R., Jiang, Z.B., Bai, T.S., 2005. A Statistic analysis of early failure rates of large-size generating units. J. Shanghai Jiaotong Univ. 39 (12), 2093–2096 (in Chinese).