Journal of the Korean Society of Safety
Vol. 36, No. 2, pp. 32-38, April 2021
https://doi.org/10.14346/JKOSOS.2021.36.2.32

ISSN 1738-3803 (Print)
ISSN 2383-9953 (Online)
http://www.kosos.or.kr/jkosos

단순 수명정보를 이용한 IPM의 베이지안 신뢰도 평가 연구

조동철* · 구정서**†

A Study on Bayesian Reliability Evaluation of IPM using Simple Information

Dong Cheol Jo* · Jeong Seo Koo**†

†Corresponding Author
Jeong Seo Koo
Tel : +82-2-970-6878
E-mail : koojs@seoultech.ac.kr

Received : February 4, 2021
Revised : March 2, 2021
Accepted : March 25, 2021

Copyright@2021 by The Korean Society of Safety All right reserved.

Abstract : This paper suggests an approach to evaluate the reliability of an intelligent power module with information deficiency of prior distribution and the characteristics of censored data through Bayesian statistics. This approach used a prior distribution of Bayesian statistics using the lifetime information provided by the manufacturer and compared and evaluated diffuse prior (vague prior) distributions. To overcome the computational complexity of Bayesian posterior distribution, it was computed with Gibbs sampling in the Monte Carlo simulation method. As a result, the standard deviation of the prior distribution developed using simple information was smaller than that of the posterior distribution calculated with the diffuse prior. In addition, it showed excellent error characteristics on RMSE compared with the Kaplan-Meier method.

Key Words : IPM, bayesian reliability, gibbs sampling, RMSE, Kaplan-Meier

1. 서 론

서울교통공사 6호선 전동차의 IPM(Intelligent Power Module)은 추진 장치의 전력반도체이며, 전동차 반입과 함께 수명을 같이하여 20여년에 이르고 있다. 동일 기능을 하는 다른 호선의 부품에 비해, IPM 고장의 빈도는 내부적으로 우수하다고 평가하고 있지만, 고장으로 인한 치명도는 유사하다. IPM의 자체 수명평가는 관측중단 특성을 많이 갖는 철도차량 부품의 특성상 데이터의 부족으로 신뢰성이 평가된 적이 없다. 또한 현재 제조사는 IPM을 생산 중단한 상태이고, 평균수명에 관한 정보도 수회의 질의 끝에, 16년 정도라는 단순 정보만을 얻을 수 있었다. 이러한 이유로 고장시 IPM과 조합된 시스템까지도 교체해야 하는 문제와 16년 정도라는 단순정보로 혼란이 발생되고 있다. 본 논문에서는 이러한 문제에 대한 정량적 정보 제공을 목적으로 데이터가 부족한 상황에서 자주 사용되는 베이지안 통계학을 이용하여 신뢰 수명(MTTF, Mean time to failure)을 추론하고자 하였다. 이를 사용하여 평균수명을 포함한 신뢰도 특성을 추론할 수 있다면, 시스템 개조의 적정 시점에 대한 근거로 사용할 수 있기 때문이다. 이때, 제조사에서 제공한 16년이라는 단순 수명정보를 베이지안 통계의 사전분포(Prior distribution)로 사용하고자 수학적 분석 절차를 거쳤다. 베이지안 통계에서 모호한 사전분포(diffuse prior)보다는 한 가지 정보라도 추가된 사전 분포는 사후분포(Posterior distribution)의 추론 성능을 향상시키기 때문이다.

베이지안 통계는 물리, 경영, 공학을 넘어 현재 통계학 기반의 인공지능분야에서 널리 사용되고 있다. 이는 고전 통계학에 비해 매우 직관적인 특성을 가지고 있으며, 정보가 부족한 상황에서도 추론의 근거를 제공할 수 있다. 원동연 등은 2000년대 이후 발표된 베이지안 논문 38편성에 대한 분석을 통해 표본수와 관련하여 고전 통계학과 베이지안 통계학의 특성을 정리하였고, 평가에 사용된 사전분포는 감마분포가 가장 많았으며, 모델로는 와이블 모델이 많이 사용되는 것을

*서울과학기술대학교 철도차량시스템공학과 박사과정 (Department of Rolling Stock System Engineering, Seoul National University of Science & Technology)
**서울과학기술대학교 철도안전공학과 교수 (Department of railway Safety Engineering, Seoul National University of Science & Technology)

확인하였다[1]. 조현준은 부품의 신뢰도 특성을 추론함에 있어 고전통계와 베이지안 통계학을 이용하여 비교 평가하였다. 이때 추정간의 정확도를 비교함에 있어 와이블 분포를 모델로 하여, 표본의 크기 대비 추정의 정확도를 비교하였다. 표본의 크기가 25개 이하일 때는 베이지안 통계가 고전통계보다 우수한 추론특성을 가짐을 확인하였다[2]. 표본의 크기 대비 추정간의 정확도를 비교함에 있어 모델을 와이블 분포로 하였다. 황정택도 관측중단 특성을 갖는 철도차량용 공기압축기의 신뢰도 특성을 평가함에 있어 와이블 분포를 사용하다. 동일 제품의 공기압축기가 운용환경이 다를 때 발생하는 신뢰도 차이를 베이지안 추론을 사용하여 차이가 있음을 밝혔다[3].

본 논문의 연구절차는 다음과 같다. 2.1절에서 베이지안 및 고전 통계학의 차이점에 대해 이론적으로 간략히 정리하고, 베이지안 통계 과정을 기술한다. 2.2절에서는 신뢰도 분석에서 많이 사용되는 와이블 분포와 베이지안 통계를 적용하여 모수를 산출하는 과정을 수식과 함께 상세히 기술하였다. 2.3절에서는 베이지안 통계학에서 사후분포(Posterior Distribution)의 연산을 위해 사용되는 MCMC(Markov Chain Monte Carlo) 방법 중 깁스 샘플링(Gibbs sampling)에 대하여 정리한다. 2.4절에서는 사전분포를 개발하는 과정을 기술하였다. 특히 제조사에서 제공한 단순 수명 정보를 사용하여 사전분포로 감마분포 개발하는 과정을 기술하였다. 2.5절에서는 실제 IPM 고장 데이터를 사용하여, 사전 정보가 있을 때와 사전정보를 고려하지 않을 때의 신뢰도 특성을 분석해 본다.

2. 본 론

2.1 베이지안 통계 개요

베이지안 통계학은 Fig. 1과 같이 사전분포와 우도함수를 기반으로 베이즈 정리를 사용한다. 베이지안 추론을 사용하면 관찰된 데이터가 증가할수록 정확한 사후분포 얻을 수 있다. 베이지 정리는 식 (1)과 같이 우도와 사전분포의 곱에 비례하며, 분모의 정규화 상수로 사후분포로 분석자에게 정보를 제공하게 된다. 베이지안 통계학은 모수값을 포함한 관찰 정보를 확률 분포로 하며, 만약 데이터가 추가되면 이전 단계에서 얻은 사후분포는 다음 단계의 사전분포로 사용 되고, 이 갱신(Updating) 과정이 고전 통계학과의 큰 차이이다.

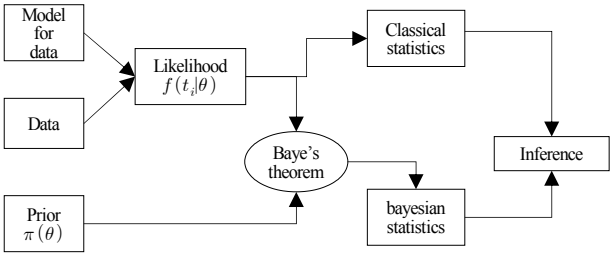

Fig. 1. Classical versus bayesian statistics.

$$\pi(\theta \mid t_1, t_2, \ldots, t_n) = \frac{\prod f(t_i|\theta)\pi(\theta)}{\int \prod f(t_i|\theta)\pi(\theta)d\theta} \quad (1)$$

여기서 θ는 추정하고자 하는 모수이고, t_i은 확률변수 T의 고장관측(측정) 시간을 의미한다. 식 (1) 중 우도함수(Likelihood function)는 미지의 관심모수 θ에 의존하는 확률밀도 함수 $f(t_i|\theta)$를 결합한 함수로 나타낸다. 여기서 $\pi(\theta|t_1, t_2, \ldots, t_n)$는 θ의 사후분포(Posterior distribution)이며, $\pi(\theta)$는 사전분포(Prior distribution)이다.

2.2 관측 중단 자료의 베이지안 적용

2.2.1 우도함수

고전 통계학에서는 우도 함수를 미분하여 직접 모수값을 산출한다. 하지만 베이지안 통계학에서 우도 함수는 Fig. 1과 같은 과정에 포함된다. 관측 중단 특성을 많이 갖는 자료의 모수를 추정하기 위해서는 확률밀도 함수와 신뢰도 함수로 구성된 우도함수를 사용한다[2,3]. 즉, n개의 관측대상에 대해, 확률변수 T 시간까지 관측을 하는 경우 우도함수는 식 (2)와 같다.

$$L(t_i|\alpha, \beta, \delta) = \prod_{t=1}^{n} [f(t_i|\alpha,\beta]^{\delta_i} [S(t_i|\alpha,\beta]^{1-\delta_i} \quad (2)$$

여기서 $\delta \in \{0, 1\}$이고, $\delta = 0$은 고장으로 관측이 종료된 것을, $\delta = 1$는 관측이 중단된 것을 의미한다. 베이지안 통계에서도 고전 통계학과 같이 와이블분포를 우도함수로 주로 사용한다[1,4,5]. 와이블분포와 관련된 신뢰도 함수들은 식 (3), 식 (4)와 같다. 식 (3)은 확률밀도함수(Probability density function), 식 (4)는 신뢰도 함수(Reliability function, Survival function)를 나타낸다.

$$f(t_i|\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{t_i}{\alpha}\right)^{\beta}\right), 0 < \alpha, \beta < \infty \quad (3)$$

$$S(t_i|\alpha,\beta) = \exp\left(-\left(\frac{t_i}{\alpha}\right)^{\beta}\right) \quad (4)$$

여기서 α는 척도모수(Scale parameter)의 추정값, β는 형상모수(Shape parameter)의 추정 값 이고, t_i는 관측시간을 나타낸다.

2.2.2 사후분포

와이블 분포와 같이 2모수 특성 갖는 확률밀도 함수이고, 데이터가 관측 중단 특성을 갖는 경우, 식 (1)은 식 (5)와 같이 변경된다. 또한 사전분포, $\pi(\theta) = \pi(\alpha, \beta)$가 서로 독립이면 $\pi(\alpha)\,\pi(\beta)$로 표현할 수 있다.

$$\pi(\theta|t) = \pi(\alpha,\beta|t) = \frac{\prod_{t=1}^{n}\left[f(t_i|\alpha,\beta)\right]^{\delta_i}\left[S(t_i|\alpha,\beta)\right]^{1-\delta_i}\pi(\alpha)\pi(\beta)}{\int_0^{\infty}\int_0^{\infty}\prod_{t=1}^{n}\left[f(t_i|\alpha,\beta)\right]^{\delta_i}\left[S(t_i|\alpha,\beta)\right]^{1-\delta_i}\pi(\alpha)\pi(\beta)d\alpha d\beta} \quad (5)$$

식 (5)에서 β에 대하여 적분하면 식 (6)과 같이 $\pi(\alpha|t)$의 확률밀도함수를 얻을 수 있고, 식 (6)을 다시 β에 관해 적분하면 α의 기대값인 점추정값 $\hat{\alpha}$을 식 (7)과 같이 계산할 수 있게 된다.

$$\pi(\alpha|t) = \int_0^{\infty}\pi(\alpha,\beta|t)d\beta \quad (6)$$

$$\hat{\alpha} = \int_0^{\infty}\alpha\,\pi(\alpha|t)\,d\alpha = \int_0^{\infty}\alpha\int_0^{\infty}\pi(\alpha,\beta|t)\,d\beta\,d\alpha \quad (7)$$

동일한 절차를 거쳐 식 (8), 식 (9)와 같이 $\hat{\beta}$의 점추정값도 산출된다.

$$\pi(\beta|t) = \int_0^{\infty}\pi(\alpha,\beta|t)d\alpha \quad (8)$$

$$\hat{\beta} = \int_0^{\infty}\beta\,\pi(\beta|t)\,d\beta = \int_0^{\infty}\beta\int_0^{\infty}\pi(\alpha,\beta|t)\,d\alpha\,d\beta \quad (9)$$

2.3 사후분포의 연산(MCMC 시뮬레이션)

사후분포를 나타내는 식 (5)에서 분모를 계산하는 것은 쉽지 않다. 이런 이유로 베이지안 통계는 직접 적분보다는 마르코프 체인 몬테칼로(Markov chain monte carlo, MCMC) 방법을 사용하여, 사후분포에서 샘플링한 결과를 사용한다. MCMC 샘플링 중 깁스 샘플링은 어떠한 튜닝도 필요하지 않다는 장점 때문에 많이 사용된다. 또한 다모수 샘플링이 쉽고, Bugs 및 R(Jags 패키지)코드에서 매우 효율적인 방법 등을 제공하고 있다. 깁스 샘플링의 알고리즘은 Fig. 2와 같다. 여기서 두 개의 매개변수를 갖는 문제에 대하여 고려해보자. 조건부 분포로부터 표본 추출을 할 수 있는 확률변수 θ_1, θ_2가 있다. 이들 사이에 완전조건부분포(Full conditional distribution)는 $p(\theta_1|\theta_2, D)$이고, $p(\theta_2|\theta_1, D)$이다. D는 관측 데이터를 나타낸다. 매개 변수 공간 중 임의의 초기

- $\theta_1^{(k)} \sim p(\theta_1|\theta_2^{(k-1)}, D)$
- $\theta_2^{(k)} \sim p(\theta_2|\theta_1^{(k-1)}, D)$

Loop: $k = 1, \ldots, n$

Fig. 2. Gibbs algorithm for two parameters.

지점에서 시작하여 아래 Fig. 2와 같이 무작위로 $\{(\theta_1^{(0)}, \theta_2^{(0)}), (\theta_1^{(1)}, \theta_2^{(1)}), \ldots, (\theta_1^{(k)}, \theta_2^{(k)})\}$ 나열을 생성한다. 새로운 값이 생성되면 이전 값은 대체된다. $\theta_1^{(0)}$에서 시작하여 $k \to \infty$까지 수행되는데, $\theta_1^{(k)}$의 샘플링값이 특정 값에 수렴하면 표본 값으로 간주할 수 있다.

2.4 Weibull 척도모수의 사전분포의 개발

와이블 확률밀도 함수를 우도함수로 사용하는 경우, 형상모수는 일양분포(Uniform distribution)나 평탄특성을 갖는 감마분포(Gamma distribution) 식 (10), 식 (11)을 사전분포로, 척도모수는 감마분포하여 사전분포로 많이 사용한다. 본 논문에서도 척도모수의 사전분포로 감마 분포를 사용하고자 한다. 감마 사전분포 개발시 MTTF의 값에 대한 정보 중 평균값 식 (12)와 분산 식 (13)을 이용하여, 공액 사전분포인 감마분포의 모수 값을 산출하는 것이 자연스럽게 사용되고 있다[6].

$$f(t|\alpha,\beta) = \frac{\beta^{-a}t^{a-1}\exp(-t/\beta)}{\Gamma(a)} = \gamma(\alpha,\beta) \quad (10)$$

$$F(t|\alpha,\beta) = \frac{1}{\Gamma(\alpha)}\int_0^{t}\beta^{-\alpha}x^{-\alpha-1}\exp(-x/\beta)dx = cumgamma(\alpha,\beta) \quad (11)$$

여기서 $\Gamma(a)$는 감마함수 이고, α는 형상모수, β는 척도모수이다. $0 < x < \infty$이고, $\alpha, \beta > 0$이다[7].

$$Mean = \alpha\beta \quad (12)$$

$$Variance = \alpha\beta^2 \quad (13)$$

감마분포의 100p 분위수(Percentile function)는 식 (14)와 같이 된다.

$$t_p = F^{-1}(p,\alpha,\beta) \quad (14)$$

2.4.1 무정보 사전분포의 적용

Hamada et al., Zaidi et al., Mun et al., Cho et al. 등 많은 연구가들이 와이블분포의 척도모수의 무정보 사전분포로 식 (15)와 같이 감마분포를 따른다고 가정하여 사후분포를 산출하였다[5,8]. 식 (15)는 와이블 분포의

척도모수($\alpha_{weibull}$)인 사전분포 $\pi(\alpha)$의 하이퍼-파라메터(Hyper-parameter) 이다. 6호선 전동차 부품의 수명연구가 부족한 실정이므로 하이퍼-파라메터 값을 작게 하여 모호한 사전분포(diffuse prior) 역할을 하도록 하고자[9-11], α_γ, β_γ의 값을 모두 0.0001을 사용한다. 이는 사전분포보다는 관측데이터가 사후분포의 특성을 지배하도록 하기 위함이다[12].

$$\alpha_{weibull} \backsim \text{dgamma}(\alpha_\gamma, \beta_\gamma) \tag{15}$$

형상모수에 대한 사전분포($\beta_{weibull}$)는 일양분포를 사용하고 식 (16)와 같이 사용한다. 이는 와이블 분포의 형상모수가 대개 0~5의 값을 갖기 때문이다. 식 (15), (16)은 본 논문에서 사용한 R코드의 함수 명으로 형상모수의 사전분포를 표현한 방법이다.

$$\beta_{weibull} \backsim \text{dunif}(0, 5) \tag{16}$$

2.4.2 단순 MTTF를 고려한 감마 사전분포 개발

사전분포의 개발에서 평균값만 알고 있는 경우, 감마분포 식 (10), 식 (11)의 특성을 이용하여 사전분포를 개발하는 방식을 사용할 수 있다[13,14]. 감마분포의 5분위수와 95분위수의 비가 5.8 이상인 특성이 사용된다[14] 식 (10), 식 (11)에서 형상모수(β)가 1인 경우, 감마분포는 지수분포와 동일해 지고, 지수분포의 모수(λ)와 $MTTF$는 역수의 관계를 갖는다. 그러므로 식 (14)를 이용하면 $MTTF_{0.5}$는 $F^{-1}(0.5, \alpha, 1)$이 되고, $MTTF_{0.05}$는 $F^{-1}(0.95, \alpha, 1)$ 같이 된다. MTTF가 16년이라는 단순정보를 이용하여 사전분포를 개발하고자, 확률적 위험과 신뢰도 평가(Bayesian inference for NASA probabilistic risk and reliability analysis)에서는 감마분포의 특성을 이용하여 MTTF 중간값과 하위 0.5% 비(Ratio)가 5.8 이상으로 하여 감마분포의 모수를 산정하였다. 그 예로 MIL-HDBK-217F의 바이폴라 소자 고장률을 지수분포로 예로 들고, 이 비율(Set the reasonable MTTF, the low MTTF, calculate their ratio RT)을 10으로 하여 감마분포의 모수 산정방법을 논하였다. 미국 국립표준기술연구소(NIST, National Institute of Standards and Technology)는 전문가의 검토를 거쳐 이 비율(RT)을 산정하는 방법을 R 코드와 함께 보여주고 있다[13]. 본 논문에서는 이 비율을 10으로 하고, 식 (17)과 같이 Newton-Raphson method으로 수치해석을 하였다.

$$t_{n+1} = t_n - \frac{f(t_n)}{f'(t_n)} \tag{17}$$

Table 1. Proposed prior distribution

	Scale(α)	Shape(β)
Unknown	gamma(0.0001 ,0.0001)	unif(0,5)
Known	gamma(0.439, 996.7)	unif(0,5)

본 논문 중 와이블 분포 척도 값에 대한 사전분포 모수값은 2.4.1절에서 논한 무정보 사전분포와 제조사에서 제공한 단순 수명정보를 사용하여 본 절에서 개발한 단순 사전분포로 Table 1과 같이 정리된다. 또한 형상모수에 대한 일양분포는 식 (16)과 동일하게 [0, 5] 구간으로 한다.

2.5 IPM의 신뢰도 모수 추정

2.5.1 고장데이터 특성

2000년 차량반입 때부터 2018년 12월 31일까지, 6호선 2개 편성(1개 편성 구동차 4개 × 차량당 6개= 24개소)에 대한 자료 48개를 Table 2와 같이 수집한 결과이다. 총 48개 중 고장은 9개이고, 39개의 데이터는 관측 중단 특성을 갖는다. IPM의 경우 데이터의 관리가 잘못되어 다른 편성들의 데이터는 모집하기가 어려움이 있었다. 이 경우 특정한 시점(2018.12.31.)을 기준으로 데이터의 관측을 중단했으므로 1종 관측 중단이 된다.

Table 2. Failure data of the IPM(*Censored, Unit: Years)

18.4*, 13.6, 18.4*, 18.4*, 18.4*, 18.4*, 9.1, 18.4*, 18.4*, 18.4*, 6, 18.4*, 18.4*, 18.4*, 18.4*, 18.4*, 18.4*, 10, 18.4*, 18.4*, 18.4*, 17.9, 16, 18.4*, 18.1*, 18.1*, 18.1*, 18.1*, 18.1*, 18.1*, 18.1*, 8.7, 5.8, 18.1*, 18.1*, 18.1*, 18.1*, 18.1*, 18.1*, 18.1*, 18.1*, 18.1*, 8.7, 18.1*, 18.1*, 18.1*, 18.1*, 18.1*

2.5.2 깁스 샘플링을 사용한 사후분포 해석

식 (5)로 IPM 데이터의 사후분포를 추정하기 위해 2.4절에서 개발한 두 종류의 사전분포와 깁스 샘플링을 사용하였다. 깁스 샘플링에 자주 이용되는 Jags 코드와 R의 CODA 패키지를 사용하여 샘플링 결과를 불러드렸다. 또한 깁스 샘플링의 수렴 여부를 판단하기 위해 Gelman-Rubin 값을 사용하였다. Gelman-Rubin 값은 체인 간 분산과 체인내의 분산의 비(Ratio)로 나타낸다. 이때 통계량이 1에 가까우면 깁스 샘플링, 즉 사후분포의 난수가 수렴한다고 할 수 있으며, 1.2를 기준으로 수렴 여부를 판정하게 된다[15,16]. Fig. 3, Fig. 4에서는 체인 (Chain)을 10개로 설정하였고, 2,000번의 번인과 각 체인에 대에 10,000번의 샘플링을 수행하였다.

Table 3과 같이 최대 Gelman-Rubin 값이 1.07이므로 깁스 샘플링은 잘 수렴한다고 할 수 있을 것이다.

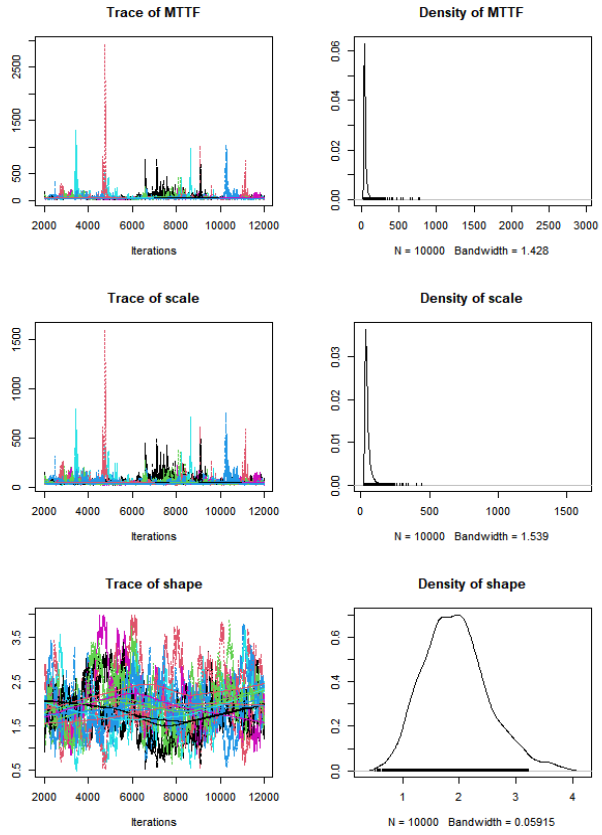

Fig. 3. Gibbs sampling results using jags of unknown prior.

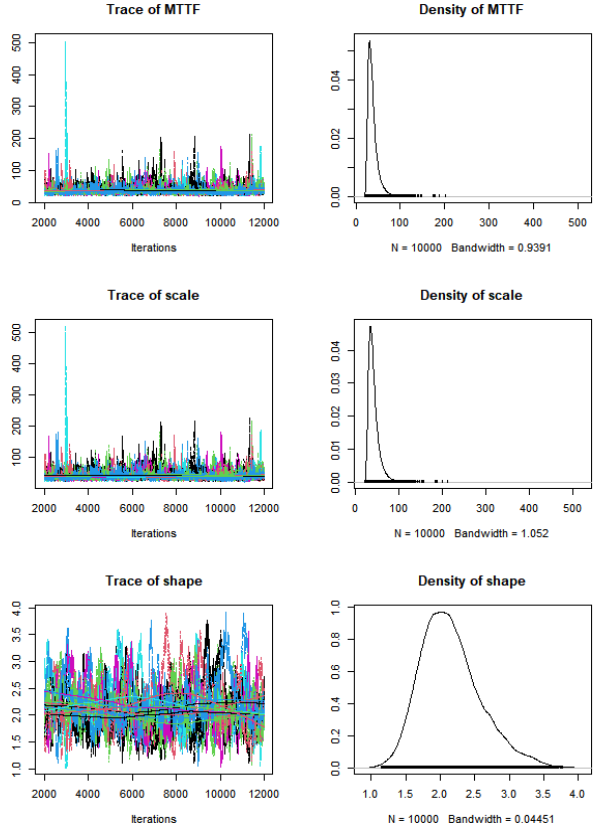

Fig. 4. Gibbs sampling results using jags of known prior.

Table 3. Gelman–Rubin–Brooks Plot

	Unknown	Known
MTTF	1.04	1.01
Shape	1.04	1.01
Scale	1.07	1.01

Fig. 3에서 평가된 확률밀도함수는 와이블분포로 모델링하였기 때문에, 모수값이 2개이다. 여기에 평균수명(MTTF)을 추가하여 샘플링 했는데, 좌측열은 깁스 샘플링 과정을 타점한 Trace plot이고, 우측열은 Trace plot을 확률밀도로 나타낸다. Fig. 3은 무정보 사전분포를 사용하여 얻은 깁스 샘플링 결과이고, Fig. 4는 제조사에서 확인해준 평균수명의 단순정보를 사전분포로 사후분포를 추정한 결과이다.

2.5.3 사후분포 해석

2.5.2절에서 확인한 것처럼 개발된 2가지 사전분포로 평가된 사후분포의 MCMC 모두 Gelman-Rubin 값이 1.2 이하이므로 잘 수렴하는 것을 확인하였다. 하지만 Table 4와 같이 단순 수명정보가 무정보보다는 표준편차가 작아 추론에 우수한 성능을 제공함을 보여준다.

Table 4. Calculation of pont estimate and error

Parameters		Bayesian Inference	
		Unknown	Known
MTTF	Point value	46.952	38.044
	Standard error	42.514	12.412
Shape	Point value	1.936	2.106
	Standard error	0.575	0.385
Scale	Point value	50.363	42.704
	Standard error	32.262	13.508

2.6 단순 정보를 이용한 IPM의 신뢰도 평가

무정보 사전분포와 단순정보를 사용한 베이지안 신뢰도 평가의 성능을 평가하기위해, 비모수적 수명분석에서 많이 사용하는 캐플란-마이어(Kaplan-Meier, KM) 방법을 이용하여 관측 중단이 발생한 시점까지의 신뢰도를 구하고, 깁스 샘플링으로 추정된 모수값과 비교한다. KM 방법은 일정한 간격의 구간 생존율을 구하는 것이 아니라 각 고장이 일어난 시점에서의 생존율을 산출하는 방법이다[17]. KM 방법으로 구해진 고장시점의 신뢰도와 베이지안 방법으로 얻어진 모수값(Table 4)을 입력한 두 가지 신뢰도 함수 값 오차를 식 (17)과 같이 평가한다. 고장발생시 KM 방법으로 산출된 신뢰도 값과 베이지안 통계로 신뢰도 값을 서로 비교할 수 있어, 예측 모델 주변에 얼마나 집중되어 있는지를 판

정하는 근거로 사용할 수 있다. 이는 수학적인 분석이 쉬움과 동시에 직관적 해석이 용이하여 제안된 모형과 발생된 실제값을 비교하는데 많이 사용된다[18]. 식 (18)과 같이 오차(Error)를 제곱(Square)해서 평균(Mean)한 값의 제곱근(Root)으로 값이 작을수록 정밀도가 높다고 판단할 수 있다.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_{KM,i} - x_{Bayes,i}\right)^2}{n}} \qquad (18)$$

여기서 $x_{KM,i}$은 KM 방법으로 평가된 신뢰도 이고, $x_{Bayes,i}$ 베이지안 모델로 평가된 값을 나타낸다. Table 5와 Fig. 5는 R code의 Survival 패키지로 고장 시간대별 평가된 KM 방법의 신뢰도와 각각의 사후모델에서 얻어진 신뢰도 모수값(Table 4)을 식 (18)로 예측한 값이다.

Table 5. Change in usage time(Failure detection)vs reliability

Fail time (Years)	KM	Simple-known	unknown
5.8	0.979	0.985	0.985
6	0.958	0.984	0.984
8.7	0.917	0.966	0.967
8.7	0.896	0.966	0.967
9.1	0.875	0.962	0.964
10	0.854	0.954	0.957
13.6	0.833	0.914	0.924
16	0.813	0.881	0.897
17.9	0.813	0.852	0.874

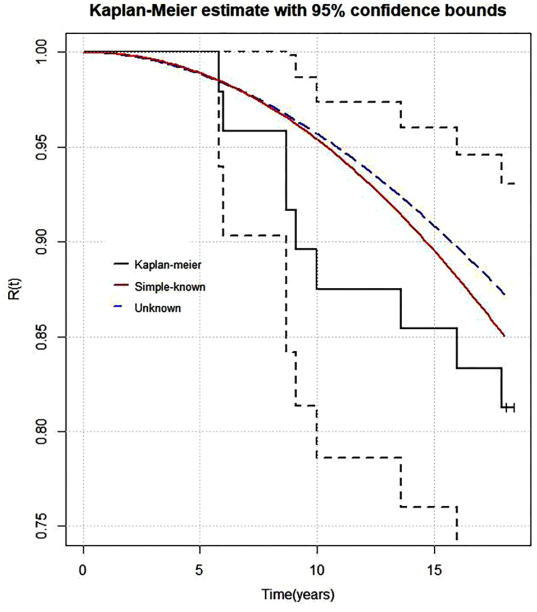

Fig. 5. Comparison of non-parametric reliability and error.

Table 6. Comparison of non-parametric reliability and error

	Unknown	Known
RMSE	0.071	0.065

Fig. 5와 Table 6은 KM과 비교 평가된 오차를 보여준다. 단순 정보를 활용하여 개발한 사전분포가 무정보 사전분포로 연산된 사후분포보다 오차가 약 0.006 만큼 오차가 작음을 보여준다.

3. 결 론

본 논문은 신뢰도 정보부족과 관측중단 특성이 많은 철도차량 부품 중 추진장치의 핵심부품 대한 베이지안 신뢰도 평가 방법에 대하여 논하였다. 제조사에서 단순 제공한 수명 정보를 이용하여 사전분포를 개발하고, 무정보 사전분포와 비교평가를 하였다. 추론 과정 중 사후분포 연산의 복잡함에 대응하고자 MCMC 시뮬레이션 중 깁스 샘플링으로 연산하였다. 깁스 샘플링은 Jags 코드와 R 코드의 CODA 패키지로 평가되었다. 기존 베이지안 신뢰도 평가 중 무정보 사전분포를 사용한 연구와 비교해보면, 제조사 단순 정보를 활용하여 개발한 사후분포가 무정보 사전분포로 연산된 사후분포보다 표준편차가 작고, 케플란 마이어 방법과 비교하여도 오차(RMSE) 특성이 우수함을 보여, 신뢰도 정보부족과 관측중단 특성이 많은 부품의 신뢰도 분석에 활용될 수 있음을 확인하였다.

고장 발생건수가 적고 치명도가 높은 정비 현장의 애로사항을 베이지안 추론 및 MCMC 기법을 이용하고, 무정보 사전분포보다는 제조사 측에서 제공하는 단순정보 혹은 전문가의 의견이 반영된 무정보 사전분포를 하면, 한계수명 예측과 정량적 부품 교체 및 유지보수 주기를 결정하거나 기존 전동차 부품의 검사주기를 변경하는 정책에 활용할 수 있을 것으로 판단된다.

Acknowledgement: This paper was written as a support for internal research fund of Seoul National University of Science & Technology.

References

1) D. Y. Won, J. H. Lim, H. S. Lim, S. I. Sung, H. S. Lim, Y. S. Kim, “A Review on the Analysis of Life Data Based on Bayesian Method”, Journal of Applied Reliability, Vol. 17, No. 3, pp. 213-223, 2017.
2) H. J. Cho, J. H. Lim and Y. S. Kim, “A Comparison of the Reliability Estimation Accuracy between Bayesian

Methods and Classical Methods Based on Weibull Distribution", Journal of the Korean Institute of Industrial Engineers, Vol. 42, No. 4, pp. 256-262, 2016.

3) J. T. Hwang, "Reliability Analysis of Air Compressor using Bayesian Inference and MCMC Simulation when Operating in Different Environments", Journal of the Korean Society for Railway, Vol. 21, No. 3, pp. 233-240, 2018.
4) H. S. Park, J. S. Koo and G. D. Kim, "Determination of Maintenance Period Considering Reliability Function and Mission Reliability of Electromagnetic Valves of EMU Doors Considering Air Leakage Failure", Journal of the Korean Society for Railway, Vol. 20, No. 5, pp. 569-576, 2017.
5) B. M. Jang and J. W. Lee, "Study on the Maintenance Interval Decisions for Life expectancy in Railway Turnout clearance Detector", Journal of the Korean Society for Railway, Vol. 20, No. 4, pp. 491-499, 2017.
6) M. S. Hamada, A. Wilson, C. S. Reese and H. Martz, "Bayesian Reliability", Springer, 2017.
7) https://wiki.analytica.com/index.php?title=Gamma_distribution
8) Antti Penttinen and Robert Piche, "Bayesian Methods", Tampere University of Technology, Finland, pp. 38-40, 2010.
9) M. S. Hamada, A. Wilson, C. S. Reese and H. Martz, "Bayesian Reliability", Springer, pp. 46-48, 97-99, 2008.
10) A. Rodionov, D. Kellyn and J. Uwe-Klügel, "Guidelines for Analysis of Data Related to Ageing of Nuclear Power Plant Components and Systems", JRC Scientific and Technical Reports, EUR, 23954, pp. 88-90, 2009.
11) H. Dezfuli, D. Kelly, C. Smith, K. Vedros and W. Galyean, "Bayesian Inference for NASA Probabilistic Risk and Reliability Analysis", USA, pp. 100-103, 2009.
12) D. L. Kelly and C. L. Smith, "Bayesian Inference in Probabilistic Risk Assessment—the Current State of the Art", Reliability Engineering & System Safety, Vol. 94, No. 2, pp. 628-643, 2009.
13) https://www.itl.nist.gov/div898/handbook/apr/section2/apr25.htm
14) H. Dezfuli, D Kelly, C. Smith, K. Vedros and W. Galyean, "Bayesian Inference for NASA Probabilistic Risk and Reliability Analysis", USA, pp. 49-53, 2009.
15) B. J. Smith, "Boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference", Journal of Statistical Software, Vol. 21, No. 11, pp. 1-37, 2007.
16) A. Gelman, Inference and Monitoring Convergence, Markov Chain Monte Carlo in Practice, USA, Chapman&Hall, pp. 131-143, 1995.
17) J. W. Baik, "Competing Risk Regression Analysis", Journal of Applied Reliability, Vol. 18, No. 2, pp. 130-142, 2018.
18) C. J. Willmott and K. Matsuura, "Advantages of the Mean Absolute Error (MAE) Over the Root Mean Square Error (RMSE) in Assessing Average Model Performance", Climate Research, Vol. 30, pp. 79-82, 2005.