Train Wheel Rim Degradation Modeling based on Generalized Linear Mixed Effect Model

Jiayu Zhou, Qi Li*, Mingcheng E, Zengqiang Jiang, Jing Ma
Department of Mechanical Engineering
Beijing Jiaotong University
Beijing, China 100044
liqi@bjtu.edu.cn

***Abstract*—The wear of the train wheel rim is one of the main failure mode of railway train wheel, which could be view as a slow degradation process. Degradation modeling of rim is the foundation of remaining useful life prediction and maintenance decision optimization. In this paper, a generalized linear mixed effect model is proposed to model a non-linear degradation shape with over-dispersion and individual random effects. The results show that the proposed model is effective to deal with non-normal error, non-homogenous variance, and non-linear path degradation data.**

Keywords- degradation; generalized linear mixed effect model; random effect; PHM

I. INTRODUCTION

The government of China is vigorously deducing the high-quality development of the equipment manufacturing industry. Rail transit manufacturing is an important industry of "Made in China 2025" and an important support for the "Belt and Road". For the consideration of system reliability, economy, and safety, Prognostics and Health Management (PHM) has become a research hotspot in the field of reliability. Modeling the degradation process under intermittent monitoring is one of the cores and basic contents of PHM [1].

As one of the important components of the railway train, the wheel directly determines the stability and safety of the train's operating conditions and it is the core of the entire bogie. When the railway train is running at high speed on the rail, especially when the wheel passes the curve, the wheel rim is subjected to a large lateral force from the rail. When the train speed is faster, the rim wear is aggravated. Under the action of centrifugal force, the rim wear exceeding the limit will make the train run unstable, and even cause derailment accidents. Thus, modelling of the degradation process of the rim of the wheel is of vital importance to maintain reliability.

During the use of mechanical parts, their performance indicators are weakened over time. This phenomenon is called degradation. In the existing academic research, the commonly used degradation models mainly include degradation path model, degradation quantity distribution model, and cumulative damage model [2].This paper focuses on the first method. The degenerate path models aim to construct a determined functional relationship to characterize the degradation process of components. Lu et al. [3] based on the linear mixed effects model to model the degradation of the same batch of parts, using random effects to describe the difference among the components. Their research shows that the maximum likelihood method and Bootstrap method are much more accurate than the normal approximation method for interval estimation of parameters in the limited sample. Oliveira [4] based on the linear degradation model of the component, compared the model estimation method, the numerical analysis method, and the pseudo-life approximation method and found that the numerical analysis method is more reliable. Lin [5] et al. analyzed the reciprocal of the slope in the linear model using the Weibull distribution, and used the Bayesian estimation method for parameter estimation. Gebraeel et al. [6] used bearing as the research object, and gave the parameter point estimations and interval estimations based on least squares method and weighted least squares method for the nonlinear degradation model, and analyzed the application conditions of different parameter estimation methods. Wang et al. [7] carried out degradation modeling for the evolution process of metal crack growth, and the model was relatively stable after parameterized time scale transformation, and solved with smooth Gaussian process modeling. Pan et al. [8] studied the parameter estimation problem of failure data under multi-stress degradation experiments based on Wiener and Gamma processes respectively. They used Markov Chain Monte Carlo (MCMC) method to solve complex maximum likelihood functions in parameter estimation. Wang [9] combined the degenerate path method with the degradation quantity distribution method, assuming that the degradation amount follows the Gaussian distribution, and mean and variance of the distribution were modeled as a function of time. Antonio [10] based on the generalized linear mixed effect model, the network customer service data of Kunming is

This work was supported in part by the Beijing Natural Science Foundation under Grant 9184030.

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

modeled and analyzed, and the regional factors are introduced into the model as random variables to effectively determine the main influencing factors of customer status, so as to effectively predict the customer churn problem.

Despite the wider application of the degenerate path models, the nonlinear non-normal degradation modeling problems that consider individual differences are relatively rare in literature, especially in the field of mechanical component degradation. However, due to the large differences in the manufacturing, operation, and maintenance of the rim of the wheels, the degradation data have large dispersion with non-normal residuals, and nonlinear path with heterogeneity of variance. Motivated by this gap between the need of actual engineering and the imperfection of theoretical research, this paper proposed a generalized linear mixed effect model for the degradation process modeling of the rim of the wheels [11].

The rest of the paper is organized as follows: In Section II, the motivating data and the basic method of generalized linear mixed effect models are presented. Results and discussions of the proposed method is presented in Section III, and conclusion is given in Section IV.

II. MATERIALS AND METHOD

A. *Motivating data*

The simulated data used in this case study is the wear data of the rim of the subway wheels under continuous operation conditions. The degradation data of forty wheels are simulated in this study. The inspection interval was 30 days (a month) and continuous measurements were performed for 60 months. Since the subway trains were operated under normal conditions, wear is the main reason for degradation of the rim.

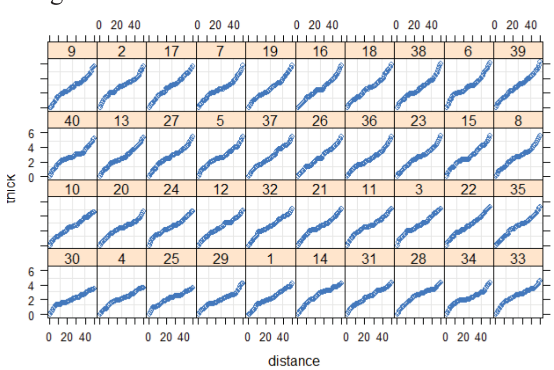

Figure 1. Rim degradation paths of 40 wheels

Fig.1 shows the degradations of the wheel rims according to the inspection time (usage time). It is obvious that the degradation paths are not the same. This phenomenon is common since the manufacturing, the location of the wheel, load, and the force of turning are different among the wheels.

Fig.1 also shows that the degradation process could be roughly divided into two parts, the initial linear degradation and the later non-linear degradation. Sensitivity analysis found that 1mm is a good choice to separate these two parts. Since the remaining useful life prediction of the wheel is much more concerned with the second part, this paper focuses on the modeling and analysis of the degradation process of the wheel rim wear greater than 1 mm. The main reflection factor of the wheel rim wear is the running time. It can be seen from the degraded image that the wheel degradation process is a nonlinear degradation process. Thus, the traditional linear regression models are not suitable to model the wheel rim wear data. As a result, the authors proposed to use generalized linear mixed effect model for degradation modeling.

B. *Structural design of generalized linear mixed effect model*

The general linear regression model mainly models the case where the residual follows the normal distribution, it cannot deal with the case of over-dispersion and heterogeneity of variance. The generalized linear model generalizes the distribution of residuals to exponential distribution families, while could deal with some issues of over-dispersion and nonlinear path of degradation. However, it could not reflect the individual difference and other random effects. Generalized linear mixed effect models (GLMMs) could be viewed as a mixture of generalized linear model and linear mixed model, it could model non-normal, non-linear path, and heterogeneity of variance degradation process with random effect of individuals.

Suppose there are N wheels, and the index i represents the i^{th} wheel $(1 \le i \le N)$. The i^{th} wheel has a number of monitors n_i with equal time interval, and the index j represents the j^{th} $(1 \le j \le n_i)$ monitor (inspection). The random effect of the i^{th} wheel is represented by b_i, and the amount of degradation is represented by y_{ij}. These degradation data $y_{i1}, y_{i2}, \cdots, y_{in_i}$ are assumed to follow a certain distribution from exponential dispersion families, such as normal distribution, gamma distribution, and inverse gaussian distribution.

$$f\left(y_{ij}\middle|b_i,\beta,\phi\right) = \exp\left\{\frac{y_{ij}\theta_{ij} - \psi(\theta_{ij})}{\phi} + c\left(y_{ij},\phi\right)\right\}$$
$$i = 1,2,\cdots,N; j = 1,2,\cdots,n_i \qquad (1)$$

Where $\psi(\cdot)$ and $c(\cdot)$ are known function based on the choice of exponential distribution families. The following equations is defined based on $\psi(\cdot)$ and the dispersion parameter ϕ.

$$\mathrm{E}\left(y_{ij}\middle|b_i\right) = \mu_{ij} = \psi'(\theta_{ij}) \qquad (2)$$
$$\mathrm{Var}\left(y_{ij}\middle|b_i\right) = \phi\psi''\left(\theta_{ij}\right) = \phi V\left(\mu_{ij}\right) \qquad (3)$$

The complete theoretical system of GLMMs can be summarized as follows:

Linear part: a function of the mean of the dependent variable is a linear combination of explanatory variables, represented by $\boldsymbol{\eta} = \boldsymbol{X\beta} + \boldsymbol{Zb}$, where $\boldsymbol{X}$ is the design matrix for fixed effect, $\boldsymbol{Z}$ is the design matrix for random effect, $\boldsymbol{\beta}$ is a vector of fixed effect parameter, and b_i is the random effect of the i^{th} wheel.

Two important distributions: the conditional distribution of the dependent variable $\boldsymbol{y}$ follow a distribution from exponential dispersion families with mean $\boldsymbol{\mu}$ and variance-covariance matrix of $\boldsymbol{R}$, represented by $\boldsymbol{y}|\boldsymbol{b} \sim (\boldsymbol{\mu}, \boldsymbol{R})$. On the other hand, the random effect $\boldsymbol{b}$ follows a multivariate normal distribution with mean $\boldsymbol{0}$ and variance-covariance matrix of $\boldsymbol{G}$, represented by $\boldsymbol{b} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{G})$.

Link function: a monotone differentiable function of the conditional mean of the dependent variable is a linear combination of explanatory variables, represented by $\boldsymbol{g}(\boldsymbol{\mu}|\boldsymbol{b}) = \boldsymbol{\eta}$.

In summary, the structure of the GLMMs could be represented as the follows (in algebraic form):

$$f(y_{ij}|b_i, \beta, \phi) = \exp\left\{\frac{y_{ij}\theta_{ij} - \psi(\theta_{ij})}{\phi} + c(y_{ij}, \phi)\right\}$$

$$\mathrm{E}(y_{ij}|b_i) = \mu_{ij} = \psi'(\theta_{ij})$$

$$\mathrm{g}(\mu_{ij}) = \eta_{ij} = X'_{ij}\beta + Z'_{ij}b_i$$

$$(i = 1,2,\cdots,N; j = 1,2,\cdots,n_i) \tag{4}$$

In the content of modeling degradation process of the rim of the wheels, y_{ij} represents the degradation amount of the rim, X'_{ij}, Z'_{ij} is based on the usage time, θ and ϕ represent the unknown parameters in a certain distribution from exponential dispersion families. θ is the natural parameters, ϕ is constant dispersion parameter. The parameter of p^{th} fixed effect is represented as β_p, and η_{ij} represented the prediction value of the i^{th} wheel based on mixed effect before the linkage transformation.

The main modeling process includes the identification of exponential dispersion families, choice of link functions, design of random effect, parameter estimation, and model selection based on residual analysis and predictive analysis. The following section will demonstrate these procedures with a case study.

III. Results and Discussion

A. *Selection of exponential dispersion families*

It is known from the degradation image that the degradation process of the wheel is nonlinear. The histogram also show that the degradation amount does not follow normal distribution, as shown in Fig.2. In this paper, GLMMs is proposed to model the nonlinear, non-normal degradation process. Thus, the first step is to select the suitable distribution from exponential dispersion families.

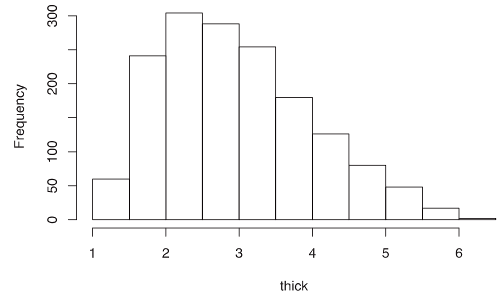

Figure 2. Histogram of the rim data

In exponential dispersion families, the variance of the random variable is a function of the mean of that random variable. Thus, the Box-Cox diagram method could be employed to identify the distribution. Based on the usage time, the degradation data are sorted and grouped into ten subsets. The sample mean and sample standard deviation of each set are calculated respectively. In Box-Cox diagram, the x-coordinate is the group mean while the y-coordinate is the group standard deviation. As shown in Fig.3, the scatter plot is basically on a straight line with a slope of 1.48, which indicates that the inverse Gaussian distribution is a reasonable choice as the distribution of the response v

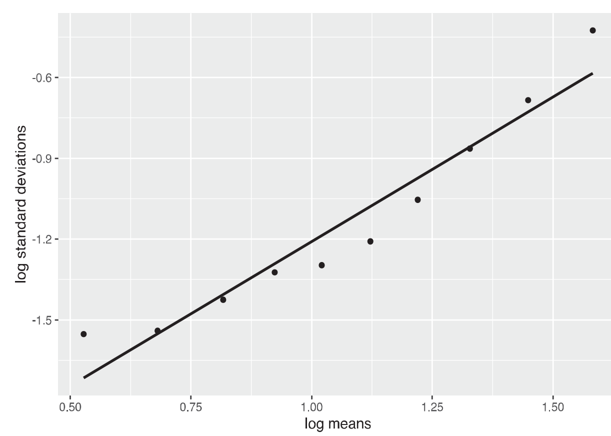

Figure 3. Wheel rim data Box-Cox diagram

B. *Selection of link function*

In the general linear regression models, the mean of the dependent variable is a linear combination of explanatory variables, represented by $E(\boldsymbol{y}) = \boldsymbol{X\beta}$. However, in GLMMs, a monotone differentiable function of the conditional mean of the dependent variable is a linear combination of explanatory variables, represented by $\boldsymbol{g}(E(\boldsymbol{y})|\boldsymbol{b}) = \boldsymbol{\eta}$. In other words, the link function links the expected responses with the linear combination of explanatory variables. The deviance residual and AIC (Akaike Information Criterion) [11] are used for link function selection. The results are summarized in Table I. Log link function is selected based on these results, since smallest AIC and deviance residuals are preferred.

TABLE I. LINK FUNCTION SELECTION

Link function	AIC	Resid. Dev
$g(\mu) = log(\mu)$	938.2	7.6120
$g(\mu) = \mu$	973.97	7.7841
$g(\mu) = 1/\mu$	1308.9	9.5967
$g(\mu) = 1/\mu^2$	1829.7	13.2889

C. *Discussion of random effects*

Since the Inverse Gaussian–GLMMs with log link is a reasonable model for degradation, the random effects could be discussed. In the proposed model, the fixed effects include the intercept and slope of usage time. Thus, the random effects could have three main choices: random intercept (Model 1), random slope (Model 2), random intercept and slope (Model 3). The mathematical expressions of these three models are as follows.

Model 1:

$$\begin{cases} y_{ij}|b_i = Inverse\ Gaussian(\mu_{ij}) \\ \log(\mu_{ij}) = \beta_0 + \beta_1 \times t_{ij} + b_{i0} \\ b_{i0} \sim N(0, \sigma_0^2) \end{cases} \tag{5}$$

Model 2:

$$\begin{cases} y_{ij}|b_i = Inverse\ Gaussian(\mu_{ij}) \\ \log(\mu_{ij}) = \beta_0 + \beta_1 \times t_{ij} + b_{i1} \times t_{ij} \\ b_{i1} \sim N(0, \sigma_1^2) \end{cases} \tag{6}$$

Model 3:

$$\begin{cases} y_{ij}|b_i = Inverse\ Gaussian(\mu_{ij}) \\ \log(\mu_{ij}) = \beta_0 + \beta_1 \times t_{ij} + b_{i1} \times t_{ij} + b_{i0} \\ (b_{i0}, b_{i1}) \sim N(0, \psi) \end{cases} \tag{7}$$

The individual random effect of intercept is represented by b_{i0} and the individual random effect of slope is represented by b_{i1}, and both of them are assumed to follow normal distribution with mean zero and variance σ_0^2 and σ_1^2, respectively. These generalized linear mixed-effects models are solved based on Penalized Quasi-likelihood (PQL) method [11] and the comparison results are summarized in Table II and Table III.

TABLE II. RESULTS COMPARISON OF MODEL 1 AND MODEL 2

Model	Df	AIC	deviance
1	4	838.86	830.86
2	4	831.82	823.82

Table II shows that Model 1 has smaller AIC. It could be interpreted as adding the intercept random effect alone is slightly better than adding the slope random effect alone.

TABLE III. RESULTS COMPARISON OF MODEL 1 AND MODEL 3

Model	Df	AIC	Deviance	P-value
1	4	838.86	830.86	
3	6	821.12	809.12	0.0006426

Table III shows that Model 1 and Model 3 have very significant difference (P-value is almost closed to zero) based on the residual deviance Chi-square test, and Model 3 has much smaller AIC and BIC. It could be interpreted as considering both the intercept random effect and the slope random effect is much better than considering only the slope or the intercept random effect alone. In summary, Model 3 is the best model.

The residual analysis also shows that the proposed GLMMs model (Model 3) is much better than the generalized linear model without random effect. As shown in Fig.4, the residuals of GLMMs are closed to zero with no significant patterns, while the residuals of the generalized linear model have significant patterns.

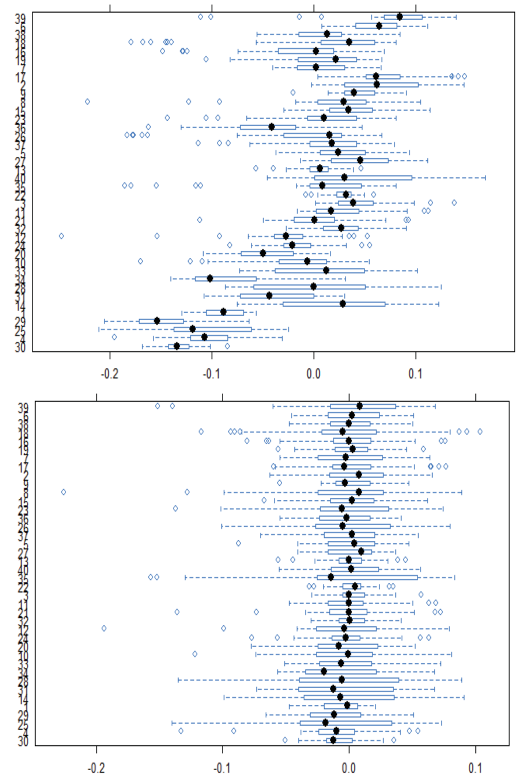

Figure 4. Residual comparison of generalized linear model and generalized linear mixed-effects model

These graphical results are consistent with the numerical calculation. The standardized residual value of GLMMs is 7.612, comparing to 67.892 of the generalized linear models.

D. *Parameter estimation results*

PQL method are employed for parameter estimation. The results are summarized in Table IV. Both methods give good parameter estimation, while the PQL method has better prediction behavior. The prediction is based on $g^{-1}(\hat{\eta}_{in_i+1})$, where $\hat{\eta}_{in_i+1} = x_{in_i+1}\hat{\beta} + z_{in_i+1}\hat{b_i}$.

TABLE IV. MODEL PARAMETER ESTIMATION BASED ON PQL METHOD

Parameter		Estimation	Standard error
Fixed effect	β_0	0.1578475	0.0202435
	β_1	0.0286948	0.0006062
Random effect	σ_0^2	0.0158182	0.1257705
	σ_1^2	1.4027e-05	0.0037452
Residual	σ^2	0.00123112	0.0350874

Based on the analysis, the final model for degradation of the rim of the wheels is as follows:

$$\begin{cases} y_{ij}|b_i = Inverse\ Gaussian(\mu_{ij}) \\ \log(\mu_{ij}) = 0.15784 + 0.02868 \times t_{ij} + b_{i1} \times t_{ij} + b_{i0} \\ (b_{i0}, b_{i1}) \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.01581824 & -0.85 \\ -0.85 & 1.40271e-05 \end{pmatrix}\right) \end{cases} \quad (8)$$

IV. CONCLUSION

In this paper, a generalized linear mixed effect model is proposed to model a non-linear shape of degradation with over-dispersion and individual random effects. The proposed model is illustrated and validated by the degradation data of the rim of the railway train wheels. The analysis shows that an Inverse Gaussian model with log link and considering both the intercept random effect and the slope random effect is the best model. (adaptive) G-H method is employed for parameter estimation, and the high estimation and prediction accuracy indicate the proposed model is appropriate.

The authors are working on the parameter updating and individual remaining useful life prediction based on this model. Bayesian update technology is a natural choice to extend this work since the random effects could provide prior information of the parameters. Since the conjugate distribution of Inverse Gaussian prior is not available, the individual remaining useful life prediction could be done by simulating from the posterior sampling distribution. MCMC method such as Metropolis Hasting algorithm with Gibbs sampling could be employed to avoid multiple integrations. Another further research direction includes interval estimation based on bootstrap method, evaluating other key components of railway train based the proposed framework, and maintenance decision optimization.

REFERENCES

[1] M. G. Pecht. "Current Status and Development of Fault Prediction and Health Management (PHM) Technology," *Acta Aeronautica et Astronautica Sinica,* vol. 26, no. 05, pp. 626-632, May. 2005.

[2] Z. Y. Yang, J. M. Zhao, Z. H. Cheng. "Preventive maintenance decision model for degradation-related multi-component systems," *Systems Engineering and Electronics*, vpl. 10, no. 4, pp. 823-832, Apr. 2018. (In Chinese)

[3] C. J. Lu, Meeker. "A comparison of degradation and failure-time methods for estimating a time-to-failure distribution," *Statistical Sinica*, vol. 6, no. 3, pp. 531-546, Apr. 1993.

[4] V. R. B. D. Oliveira, E. A. Colosimo. "Comparison of Methods to Estimate the Time-to-failure Distribution in Degradation Tests," *Quality & Reliability Engineering International*, vol. 20, no. 4, pp. 363-373, Apr. 2004.

[5] J. Lin , M. Asplund. "Reliability analysis for preventive maintenance based on classical and Bayesian semi-parametric degradation approaches using locomotive wheel-sets as a case study," *Reliability Engineering & System Safety,* vol. 134, pp. 143-156, Apr. 2015.

[6] N. Z. Gebraeel, M. A. Lawley, R. Li. "Residual-life distributions from component degradation signals: A Bayesian approach," *IE Transactions*, vol. 37, no. 06, pp. 543-557, Feb. 2005.

[7] Z. Wang, Q. Wu, X. Zhang. "A generalized degradation model based on Gaussian process," *Microelectronics Reliability*, vol. 85, pp. 207-214, Feb. 2018.

[8] Z. Pan, N. Balakrishnan. "Multiple-Steps Step-Stress Accelerated Degradation Modeling Based on Wiener and Gamma Processes" *Communications in Statistics - Simulation and Computation*, vol. 39, no. 7, pp. 1384-1402, Feb. 2010.

[9] F. K. Wang, T. P. Chu. "Lifetime predictions of LED-based light bars by accelerated degradation test," *Microelectronics Reliability*, vol. 52, no. 7, pp. 1332-1336, Feb. 2012.

[10] K. Antonio, J. Beirlant. "Actuarial statistics with generalized linear mixed models," *Insurance Mathematics & Economics*, vol. 40, no. 1, pp. 58-76, Feb. 2007.

[11] D. E. Myers. "Linear and Generalized Linear Mixed Models and Their Applications," *Technometrics*, vol. 50, no. 1, pp. 93-94, Feb. 2012.

[12] D. Posada, T. R. Buckley. "Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches over Likelihood Ratio Tests," *Systematic Biology*, vol. 53, no. 5, pp. 793-808, Apr. 2004.

[13] A. Benedetti, R. Platt, Juli. Atherton. "Generalized Linear Mixed Models for Binary Data: Are Matching Results from Penalized Quasi-Likelihood and Numerical Integration Less Biased?," PLoS ONE, vol. 9, no. 1, pp. e84601, Feb. 2014.