Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P9781-cd

RUL prediction using Bayesian polynomial regression

Kirill Ivanov
Informatics, Alpen-Adria Universität Klagenfurt, Austria. E-mail: kirill.ivanov@infineon.com

Horst Lewitschnig
Infineon Technologies Austria AG, Austria. E-mail: horst.lewitschnig@infineon.com

Maintaining the reliability of complex systems is crucial in today's technological landscape. Maintenance strategies have evolved from corrective and time-based maintenance to condition-based maintenance and prognostics and health management. Typical remaining useful lifetime (RUL) prediction methods require substantial historical data, posing challenges in data-limited scenarios. To address this, we propose an efficient Bayesian polynomial regression approach with informative priors that predicts RUL even with sparse data. Regression parameters are continuously updated as new data are collected, ensuring accuracy and responsiveness. We validate our algorithm on simulated power module run-to-failure degradation data.

Keywords: Bayesian Regression, Prognostics and Health Management, Remaining Useful Lifetime.

1. Introduction

In today's fast-evolving technological landscape, the reliability of complex systems and equipment is of paramount importance. Reliability can be addressed from the design, the production, and from the maintenance point of view. Maintenance can be defined as any activity including tests, measurements, replacements, adjustments, and repairs. It was intended to retain or restore a functional unit to a specified state in which the unit can perform its required functions Kordestani et al. (2021).

Historically, maintenance started as corrective maintenance, where the system was serviced and repaired after a failure occurred. This approach, while straightforward, leads to unexpected downtimes and increased costs Wang et al. (2014). Subsequently, time-based maintenance emerged, where maintenance is performed at scheduled intervals regardless of the system condition. Although this method reduced unexpected failures, it was still not optimal because it did not account for the actual wear and tear of the components Ahmad and Kamaruddin (2012). With the advancement of sensor technologies, condition-based maintenance (CBM) became the new paradigm. CBM involves collecting sensor data from the system, deriving health indicators, and performing maintenance actions based on these indicators Navarra et al. (2007). Despite its effectiveness, CBM does not predict future failures, leading to the development of Prognostics and Health Management (PHM) or CBM+ Jaw and Merrill (2008). One feature of PHM is the prediction of Remaining Useful Lifetime (RUL) of components by analyzing the health indicators and forecasting future degradation.

The RUL of a system or a component can be defined as a random variable that depends on the age, operation environment, and the observed health indicators Si et al. (2011). RUL prediction approaches can be broadly classified into three categories: data-driven, model-based, and hybrid algorithms Heng et al. (2009). Data-driven algorithms rely on historical and real-time data to predict the RUL by identifying patterns and relationships in the data using statistical and machine learning techniques Medjaher et al. (2012). Model-based algorithms, on the other hand, involve the development of physical or mathematical models that describe the degradation process of the system or component Jaw and Wang (2006). Hybrid algorithms combine elements of both data-driven and model-based approaches to leverage the strengths of each method Liao and Kottig (2014).

Before building a model for RUL, features that indicate the state of health and serve as independent variables to predict the RUL need to be identified. This can be based on domain knowledge, or specific methods like Principal Component Analysis (PCA) for dimensionality reduction can be applied Rehab et al. (2021). PCA can help in extracting the most significant features from high-dimensional data, which can then be fed into predictive models for RUL estimation.

One of the most popular data-driven methods are neural networks, which are capable of capturing complex nonlinear relationships in the data. For instance, deep learning techniques, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, have been effectively used for RUL prediction due to their ability to learn from sequential data and extract meaningful features automatically Cheng et al. (2020). Support Vector Machines (SVMs) are another data-driven method for RUL prediction. SVMs are robust to overfitting, especially in high-dimensional spaces. Ensemble learning techniques, such as Random Forests and Gradient Boosting Machines, have also been employed for RUL prediction Dong et al. (2014). These methods combine the predictions from multiple models to improve accuracy and robustness.

RUL prediction methods typically require substantial amounts of historical data to train accurate models Liao and Kottig (2014). The dependency on large data sets can be a significant limitation, especially in scenarios where historical failure data are sparse or incomplete. To address this data-dependency challenge, Bayesian approaches can incorporate prior knowledge into the predictive model and therefore reducing the reliance on extensive historical data Wang et al. (2023). Typically, these methods require comparable high computational resources which makes them difficult or even impossible to use them in so-called "edge" devices.

Our proposed approach addresses situations of sparse data and limited calculation resources. It leverages Bayesian polynomial regression, which incorporates prior knowledge into the predictive model, and reduces the reliance on extensive historical data. Furthermore, we combine Bayesian regression with linear polynomial regression for more efficient RUL predictions. This approach improves their robustness. Moreover, it provides a probabilistic framework that accounts for uncertainties, which is crucial for effective maintenance planning and decision-making.

2. Regression Update with Data

Consider a generalized polynomial regression model where the relationship between the dependent variable y and the independent variable x is given by a polynomial of degree γ,

$$g(y) = \beta_0 + \beta_1 x + \beta_2 x^2 + ... + \beta_p x^\gamma + \epsilon, \quad (1)$$

where ϵ is the error term and g is the link function. For the simple linear case, g is the identity function and ϵ is assumed to be normally distributed with mean zero and variance σ^2:

$$\epsilon \sim \mathcal{N}(0, \sigma^2). \quad (2)$$

Without loss of generality, in this work we use a two-parametric model with β_0 and $\beta_p = \beta_1$. As new data points (x_i, y_i), $i \leq n$, are observed, we update the coefficients $\widehat{\beta}_0$ and $\widehat{\beta}_1$ iteratively. For each new data point, we determine $\widehat{\beta}_0$ and $\widehat{\beta}_1$ based on the least squares criterion

$$\min_{\beta_0, \beta_p} \sum_{i=1}^{k} \left(y_i - (\beta_0 + \beta_1 x_i^\gamma)\right)^2, \quad (3)$$

where $k \in \{l, ..., n\}$, $l \leq n$, and $l, n \in \mathbb{N}$.

For low values of l, i.e., when only few degradation data are available, we use a Bayesian regression. For this, we apply normally distributed priors to the regression coefficients,

$$\begin{aligned} \beta_0 &\sim \mathcal{N}(\mu_{\beta_0}, \sigma^2_{\beta_0}), \\ \beta_1 &\sim \mathcal{N}(\mu_{\beta_1}, \sigma^2_{\beta_1}). \end{aligned} \quad (4)$$

In fact, we set $\sigma^2_{\beta_0}$ and $\sigma^2_{\beta_1}$ as functions of x,

$$\begin{aligned} \beta_0 &\sim \mathcal{N}(\mu_{\beta_0}, \sigma^2_{\beta_0}(x)), \\ \beta_1 &\sim \mathcal{N}(\mu_{\beta_1}, \sigma^2_{\beta_1}(x)). \end{aligned} \quad (5)$$

By varying $\sigma^2_{\beta_0}(x)$ and $\sigma^2_{\beta_1}(x)$, the impact of the prior can be controlled.

Using the Bayes' theorem, we get:

$$p(\theta|y) = \frac{L(\theta|y)p(\theta)}{p(y)}, \quad (6)$$

where $p(y) = \int L(\theta|y)p(\theta)d\theta$. Because $p(y)$ does not depend on θ, the Eq. 6 can be rewritten as

$$p(\theta|y) \propto L(\theta|y)p(\theta). \tag{7}$$

This can further be expanded to

$$\begin{aligned} &L(\theta|X_{i+1}) \times L(\theta|X_i) \times p(\theta) \\ &= L(\theta|X_{i+1}) \times p(\theta|X_i), \end{aligned} \tag{8}$$

where $p(\theta)$ is a prior for θ before knowing X_i and $L(\theta|X_{i+1}) \times p(\theta|X_i)$ is the posterior of θ given prior $p(\theta|X_i)$.

So, we use the posterior after the first update as the prior for the second iteration, and so on.

3. Dynamics of Regression Updates

The general framework of our approach for RUL prediction integrates Bayesian regression with informative priors for initial predictions and transitions to linear polynomial regression as more data become available.

At the start of the operation, we employ a Bayesian regression with priors derived from comparable systems. This stage is crucial for providing robust predictions when historical data are limited, and the system has just started to operate. As the operation continues, sensor data are continuously collected. The Bayesian regression model is periodically updated with the new data, refining the posterior distributions of the model parameters and improving the accuracy.

We are observing degradation over time or cycles, resp. Thus, we set $x = t$. With increasing t, more data are accumulated and the need for Bayesian regression diminishes. At a predefined point in time, we transition from Bayesian regression to linear polynomial regression. The linear polynomial regression model is trained using the accumulated data, providing a simple and fast calculation.

In this work, we identified the parameter V_{DS} as the dependent variable for the degradation model. We first apply a log transformation to the output parameter. This linearizes the relationship between the independent and dependent variables, making the data more suitable for polynomial regression. Specifically, we transform V_{DS} as follows:

$$y = log(V_{DS} - min(V_{DS}) + \varphi), \tag{9}$$

where a constant $\varphi = 10^{-7}$ is added to ensure positive values. After the transformation, we apply a backward-forward model selection algorithm. This results in a cubic polynomial model of the form

$$y = \beta_0 + \beta_1 t^3 + \epsilon. \tag{10}$$

Bayes' theorem states that the posterior can be defined as

$$posterior \propto \frac{likelihood \times prior}{evidence}. \tag{11}$$

Because the $evidence$ is independent of θ (Eq. 6) and we want to maximize the posterior, we can drop the evidence term and Eq. 11 becomes

$$posterior \propto likelihood \times prior. \tag{12}$$

Considering the cubic polynomial model, Eq. 12 can be rewritten as

$$\begin{aligned} f(\beta_0, \beta_1|Y, X) \propto \prod_{i=1}^{n} L(y_i|\beta_0 + \beta_1 t^3) \cdot \\ \cdot f(\beta_0) \cdot f(\beta_1). \end{aligned} \tag{13}$$

The importance of the priors can be controlled either by altering the likelihood L or by making the prior more informative. Historic data are limited, especially for the case when the system has just started operating. Therefore, down-scaling the likelihood L by a factor $\gamma \in (0, 1)$ has shown to lead to unreliable predictions. Therefore, the prior importance is controlled by making the priors more informative. This can be achieved by changing the standard deviation of the prior distributions, σ_{β_i}, $i = 0, 1$, see also Eq. 5. Smaller σ_{β_i} makes the probability mass more concentrated around the means μ_{β_i}, $i = 0, 1$. This results in more informative prior distributions. On the other hand, uninformative prior leads to simple linear regression.

Due to the high number of data points in the dimension of 10^4, we start with σ_{β_i}, $i = 0, 1$, in the range of 10^{-10}. Subsequently, $\sigma_{\beta_i}(t)$ is function of time. This is, $\sigma_{\beta_i}(t)$ gradually increases with t to reduce systematically the impact of the

prior and gives more weight to the data in Eq. 13. The algorithm shifts from Bayesian regression to linear regression when the degradation signature is visible in the collected data. This point can be determined using, e.g., change point detection methods. The rate at which the $\sigma_{\beta_i}(t)$ increases is controlled by domain knowledge.

The prior distributions for β_0, β_1 start from

$$\begin{aligned} \beta_{0_1} &\sim \mathcal{N}(\mu_{\beta_0}, 10^{-10}), \\ \beta_{1_1} &\sim \mathcal{N}(\mu_{\beta_1}, 10^{-10}) \end{aligned} \tag{14}$$

and change over time to

$$\begin{aligned} \beta_{0_n} &\sim \mathcal{N}(\mu_{\beta_0}, \sigma^2_{\beta_0}), \\ \beta_{1_n} &\sim \mathcal{N}(\mu_{\beta_1}, \sigma^2_{\beta_1}), \end{aligned} \tag{15}$$

where n is the last step before end of life.

4. Case Study

To validate the proposed method, we simulated run-to-failure data with typical degradation behavior based on Di Nuzzo et al. (2023).

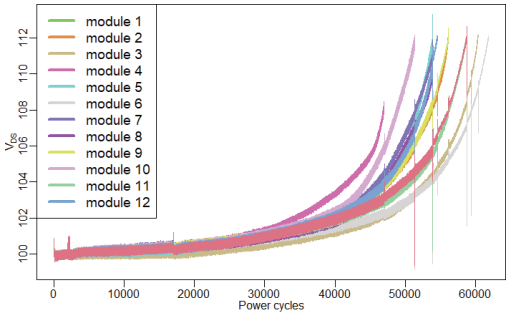

Fig. 1. Power module degradation data.

Figure 1 shows simulated representative power module data, where the drain source voltage (V_{DS}) is the feature that reflects the degradation over time. Once V_{DS} is above a certain defined threshold, the module is considered to have reached its end of life. In this work, this threshold is chosen arbitrarily. Typically, power modules show a constant linear drift in the V_{DS} at early stages, followed by an exponential degradation pattern.

Figure 2 gives the transformed power module data using Eq. 9.

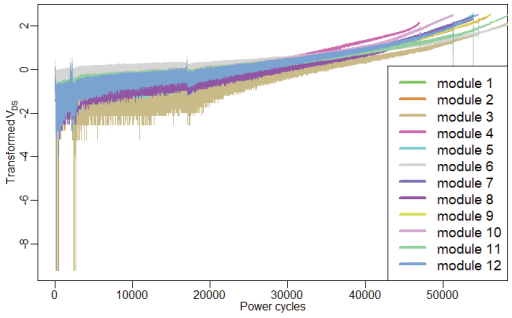

Fig. 2. Transformed power module degradation data.

In this work, the main parameter for the model assessment is the accuracy of the RUL prediction.

Figure 3 shows comparative results of Bayesian and linear cubic regression models defined in Eq. 10 for one module. The prior distribution is based on the regression coefficients computed for the first 6 modules. The colored dots are the percentage of the full data set compared to the difference between the true and predicted failure time. It can be seen that even at 10% of the full data set, Bayesian regression (red dots) provides accurate RUL predictions. As more data become available, linear regression (blue dots) outperforms Bayesian regression because it is not affected by priors.

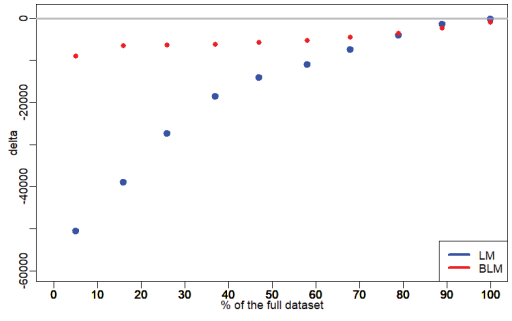

Fig. 3. Bayesian regression compared to linear regression for a single power module.

Figure 4 gives the results for all power modules. It can be seen that the for regression models adding informative priors essentially improves the performance and reduce data-dependency on sparse historic data.

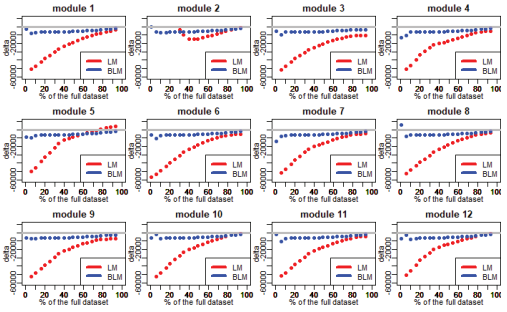

Fig. 4. Bayesian regression compared to linear regression for all power modules.

Figure 5 depicts how the actual and predicted failure values are computed. We apply various methods to the data when 20%, 50%, and 90% of the lifetime is achieved. Then, we make predictions until end of life. The difference between actual and predicted failure time is computed.

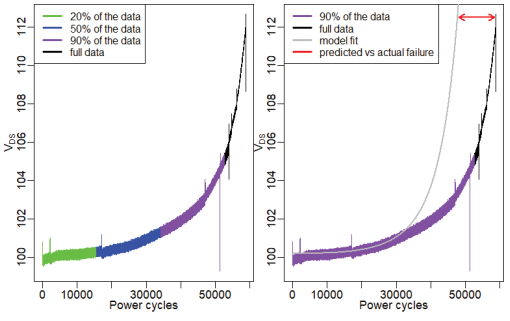

Fig. 5. The process of computing actual vs predicted failure metric.

Table 1 shows the comparison of different models for the power module data set. Values are averaged across all power modules. Negative values mean that the model underestimates the RUL, whereas positive values mean that the model overestimates it. NA means that the prediction was unstable.

The linear regression showed that the error terms are heteroscedastic and have a skewed distribution. We compared our approach with two Generalized Linear Models (GLMs). Both GLMs are from the Gamma family, because the Gamma distribution is tailored for continuous, skewed,

Table 1. Model comparison.

Model	20%	50%	90%
BLM	-6409	-5306	-1775
log-linked Gamma GLM	-5693	12906	13996
inverse-linked Gamma GLM	-6685	11658	14474
FB Prophet	NA	109882	22523
ARIMA	NA	173159	19378
GPR	NA	29390	6906

and positive-valued data. These GLMs can effectively transform and linearize data. The first GLM utilized the log-link function which linearizes the exponential growth pattern of the data. The second GLM uses the inverse link which accommodates the increasing variance frequently seen in degradation processes. The next model is the Facebook Prophet (FB Prophet) model Taylor and Letham (2021) which works well for data with non-linear shapes because it incorporates change points where the growth rate shifts. The ARIMA Hyndman and Khandakar (2008) model is designed to capture the underlying patterns and structures in time series data through its autoregressive (AR), differencing (I), and moving average (MA) components. By focusing on the intrinsic temporal dependencies and trends within the data, an ARIMA model can effectively model the continuous and progressive nature of degradation. The final model is the Gaussian Process Regression (GPR) which accounts for both temporal correlations in the data and the mean trend Erickson (2025).

Table 1 shows that the BLM model performs best in comparison to the other investigated methods.

5. Discussion and Outlook

The new method shows advantages of using Bayesian regression with informative priors for RUL prediction in power modules. By continuously updating the model parameters as new data points are collected, our approach maintains high accuracy and adapts to changing conditions. The transition to linear polynomial regression ensures

computational efficiency for accurate predictions. Overall, our proposed method offers a robust and effective solution for RUL estimation. It requires few computational resources and can be used at edge devices.

This method is developed specifically for the degradation of power modules. It can be further developed and evaluated for other electrical and mechanical systems as well.

We still see potential for optimization in refining the linear regression model. We see that the response variable shows heteroscedasticity, which even increases by the logarithmic transformation. This imbalanced variation of the error term puts more weight to the regression parameters at the beginning of the degradation. Nevertheless, the prediction of RUL becomes more accurate with the BLM model. Alternative transformations could be assessed to further stabilize the regression over time while optimizing the accuracy of the RUL prediction. Also, alternative prior distributions could be investigated.

The structure how the prior information changes from informative to non-informative prior could further be refined by using advanced machine learning methods. This can be a further lever for more accurate predictions.

Acknowledgement

ARCHIMEDES is supported by the Chips Joint Undertaking and its members, including the top-up funding by National Authorities under Grant Agreement No 101112295, funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the granting authority. Neither the European Union nor the granting authority can be held responsible for them.

References

Ahmad, R. and S. Kamaruddin (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering 63*, 135–149.

Cheng, C., G. Ma, Y. Zhang, M. Sun, F. Teng, H. Ding, and Y. Yuan (2020). A deep learning-based remaining useful life prediction approach for bearings. *IEEE/ASME Transactions on Mechatronics 25*, 1243–1254.

Di Nuzzo, G., H. Lewitschnig, M. Tuellmann, S. Rzepka, and A. Otto (2023). A data-driven condition monitoring method to predict the remaining useful life of SiC power modules for traction inverters. In *2023 IEEE International Conference on Prognostics and Health Management (ICPHM)*, pp. 312–319. IEEE.

Dong, S., S. Yin, B. Tang, L. Chen, and T. Luo (2014). Bearing degradation process prediction based on the support vector machine and markov model. *Shock and Vibration 2014*, 1–15.

Erickson, C. (2025). *GauPro: Gaussian Process Fitting*. R package version 0.2.13.9000.

Heng, A., S. Zhang, A. C. Tan, and J. Mathew (2009). Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical Systems and Signal Processing 23*, 724–739.

Hyndman, R. J. and Y. Khandakar (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software 27*(3), 1–22.

Jaw, L. C. and W. Merrill (2008). CBM+ research environment - facilitating technology development, experimentation, and maturation. In *2008 IEEE Aerospace Conference*, pp. 1–6. IEEE.

Jaw, L. C. and W. Wang (2006). Mathematical formulation of model-based methods for diagnostics and prognostics. In *Volume 2: Aircraft Engine; Ceramics; Coal, Biomass and Alternative Fuels; Controls, Diagnostics and Instrumentation; Environmental and Regulatory Affairs*, pp. 691–697. ASMEDC.

Kordestani, M., M. Saif, M. E. Orchard, R. Razavi-Far, and K. Khorasani (2021). Failure prognosis and applications—a survey of recent literature. *IEEE Transactions on Reliability 70*, 728–748.

Liao, L. and F. Kottig (2014). Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. *IEEE Transactions on Reliability 63*, 191–207.

Medjaher, K., D. A. Tobon-Mejia, and N. Zer-

houni (2012). Remaining useful life estimation of critical components with application to bearings. *IEEE Transactions on Reliability 61*, 292–302.

Navarra, K., R. Lawton, S. McCusker, and N. Hearrell (2007). An enterprise strategy for implementing conditioned-based maintenance plus (CBM+) research in the USAF. In *2007 IEEE Aerospace Conference*, pp. 1–7. IEEE.

Rehab, A., I. Ali, W. Gomaa, and M. N. Fors (2021). Bearings fault detection using hidden markov models and principal component analysis enhanced features. *PHM Society European Conference 6*, 11.

Si, X.-S., W. Wang, C.-H. Hu, and D.-H. Zhou (2011). Remaining useful life estimation – a review on the statistical data driven approaches. *European Journal of Operational Research 213*, 1–14.

Taylor, S. and B. Letham (2021). *prophet: Automatic Forecasting Procedure*. R package version 1.0.

Wang, L., H. Cao, Z. Ye, and H. Xu (2023). Bayesian large-kernel attention network for bearing remaining useful life prediction and uncertainty quantification. *Reliability Engineering & System Safety 238*, 1–16.

Wang, Y., C. Deng, J. Wu, Y. Wang, and Y. Xiong (2014). A corrective maintenance scheme for engineering equipment. *Engineering Failure Analysis 36*, 269–283.