DEPARTMENT OF COMPUTER SCIENCE

MSC IN DATA SCIENCE

# CM50268 Coursework Two

Part 2C

## Bayesian Machine Learning Mini-Project

Submitted By

**KAJOL PAWAR**

**(Student Id: 249168369)**

Lecturer:

**Michael Tipping**

**April 29, 2025**

## INTRODUCTION

This investigation addresses the challenge of forecasting the future integrity of mechanical components based on limited and noisy early-life data. The client's component "K" plays a critical role in machinery and is monitored using an integrity metric (0–100). While some components have well-populated time-series measurements, others, particularly 25 newer cases, contain only sparse early data, making long-term prediction more difficult.

Traditional deterministic models offer point predictions without accounting for uncertainty, making them unsuitable for high-stakes environments where reliable risk estimates are crucial.

To solve this, a Bayesian hierarchical modelling was adopted. Unlike traditional deterministic models, it captures uncertainty at all levels, shares information across components, and handles data sparsity effectively. Degradation is modelled as exponential decay, with each component's decay rate influenced by both individual behaviours and five known characteristics (X1–X5). Modern inference techniques, including Hamiltonian Monte Carlo with NUTS, were used for efficient posterior sampling.

## PREDICTIVE MODELLING

### 1. Baseline Model

#### 1.A. Model Structure

Based on the outlined requirements, the baseline model assumes that each component's integrity degrades exponentially over time, described by:

$$f_i(t) = u_i \, exp(-v_i \, t/100)$$

where, $f_i(t)$ represents the predicted integrity of component i at time t, $u_i$ is the initial integrity level and $v_i$ is the degradation rate for component i.

Uncertainty is incorporated at multiple levels. The parameters $u_i$ and $v_i$ are drawn from shared population distributions:

$$u_i \sim N(\mu_u, \sigma_u), \quad v_i \sim N(\mu_v, \sigma_v)$$

with weakly informative priors:
$\mu_u \sim N(90,10), \quad \sigma_u \sim HalfNormal(5),$
$\mu_v \sim N(90,10), \quad \sigma_v \sim HalfNormal(5),$

Observation noise is modelled as:

$$yi(t) \sim N(fi(t), \sigma y), \qquad \sigma y \sim HalfNormal(5)$$

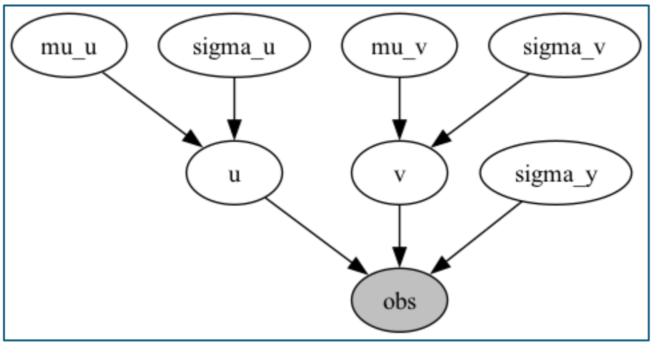


**Figure 1. Baseline Model Structure**

The model's hierarchical structure, depicted in Figure 1 using `numpyro.render_model`, illustrates that component-specific parameters are conditionally sampled based on shared hyperparameters across the population.

## 1.B. Model Incorporation within Bayesian Framework

The baseline model was embedded in a Bayesian hierarchical framework by modelling component integrity as an exponential decay over time, governed by latent parameters $u_i$ and $v_i$. These were drawn from shared population-level distributions with weakly informative priors $\mu_u$, $\sigma_u$, and $\mu_v$, $\sigma_v$, allowing for partial pooling across components. The observed integrity data was linked through a Gaussian likelihood with noise $\sigma_y$ capturing uncertainty. This structure enabled robust predictions even with sparse data, while transparently quantifying uncertainty at all levels.

## 1.C. Baseline Model Performance and Diagnostics

The baseline model's performance was assessed using MCMC diagnostics after sampling via the No-U-Turn Sampler (NUTS) across four chains, collecting 2000 samples after 2000 warmups. All parameters achieved perfect convergence with Gelman-Rubin statistics ($\hat{r}$) and high effective sample sizes ($n_{eff}$), often exceeding 8000, ensuring stable posterior estimates.

Trace plots as depicted in Figure 2 showed good mixing with no divergence or multimodality. Posterior predictive checks aligned closely with observed integrity values, displaying narrow credible intervals for well-observed components and wider uncertainty for sparse cases, accurately reflecting epistemic uncertainty.

The model proved both reliable and robust predictive performance, effectively adapting to data sparsity and delivering well-calibrated uncertainty estimates.

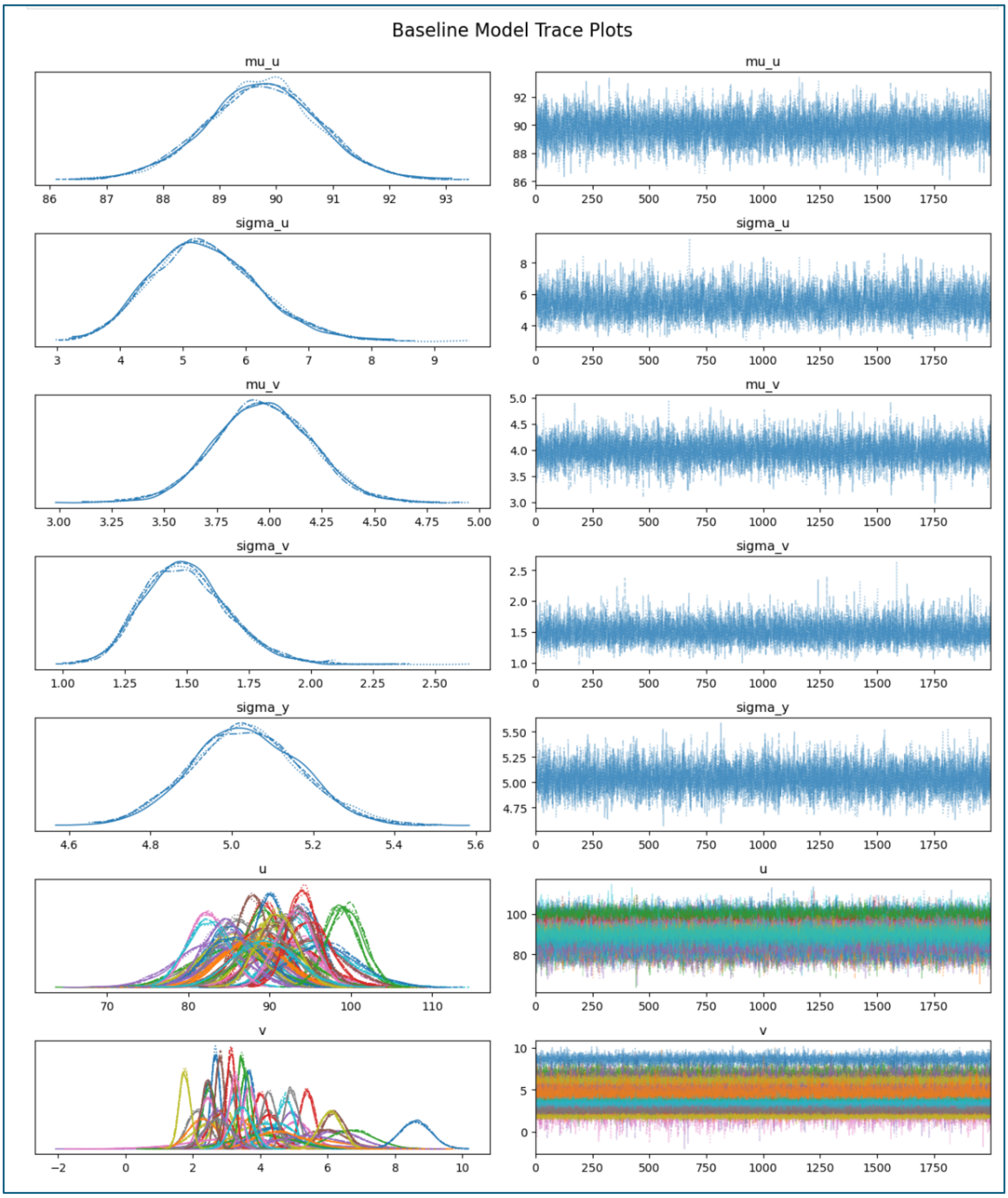


**Figure 2**: **Trace plots for key parameters of the baseline Bayesian hierarchical model**

## 2. Enhanced model

### 2.A. Model Structure

To improve predictive performance, the baseline model was extended to incorporate component-specific features. In the enhanced model, the degradation rate depends both on an individual-specific parameter and a linear combination of five measurable characteristics X1 to X5. The degradation function was modified as:

$$f_i(t) = u_i \, \boldsymbol{exp}(-\frac{v_i + \sum_{j=1}^{5} w_j x_{ji}}{100} t)$$

$w_j$ represents global feature weights that quantify how each characteristic affects the degradation process.

As in the baseline model, $u_i$ and $v_i$ were drawn hierarchically from shared population normal distributions with hyperparameters $\mu_u$, $\sigma_u$, and $\mu_v$, $\sigma_v$ respectively, using weakly informative priors. The global feature weights $w_j$ were assigned standard normal priors N(0,1), allowing the data to reveal whether features have positive, negative, or negligible effects without strong prior bias. Observation noise $\sigma_y$ remained modelled using a HalfNormal prior to capture measurement variability.

This framework naturally integrates feature-driven variability into the hierarchical structure, enabling the model to learn from both early integrity measurements and associated component characteristics. By doing so, it facilitates enhancing prediction especially when early-life data is sparse.

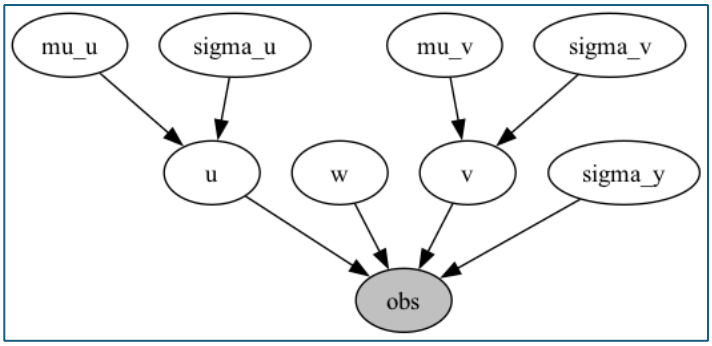


**Figure 3. Enhanced Model Structure**

A graphical representation of this extended model is provided in Figure 3, illustrating the hierarchical relationships among hyperparameters, component-specific variables, feature weights, and observations.

### 2.B. Model Incorporation within Bayesian Framework

The enhanced model builds on the baseline hierarchical structure by incorporating feature-driven variability into the degradation process. Each component's integrity follows an exponential decay, where the degradation rate is influenced by both a component-specific latent variable $v_i$ and a weighted sum of its characteristics via global weights $w_j$.

The model maintains the hierarchical structure for $u_i$ and $v_i$, drawn from population-level priors. Additionally, the feature weights $w_j$ are treated as global latent variables with their own priors, capturing uncertainty in the influence of each feature on degradation. This allows the model to jointly infer the population behavior, individual component tendencies, and systematic effects of physical features within a unified Bayesian framework.

Observed integrity values are modelled with a Gaussian likelihood centred on the predicted $f_i(t)$, with shared observation noise $\sigma_y$. This integration of component features improves predictions, especially for components with limited historical integrity measurements.

Overall, the enhanced hierarchical Bayesian model improves prediction accuracy, while maintaining principled uncertainty quantification through full probabilistic modelling.

## 2.C. Model Performance and Diagnostics

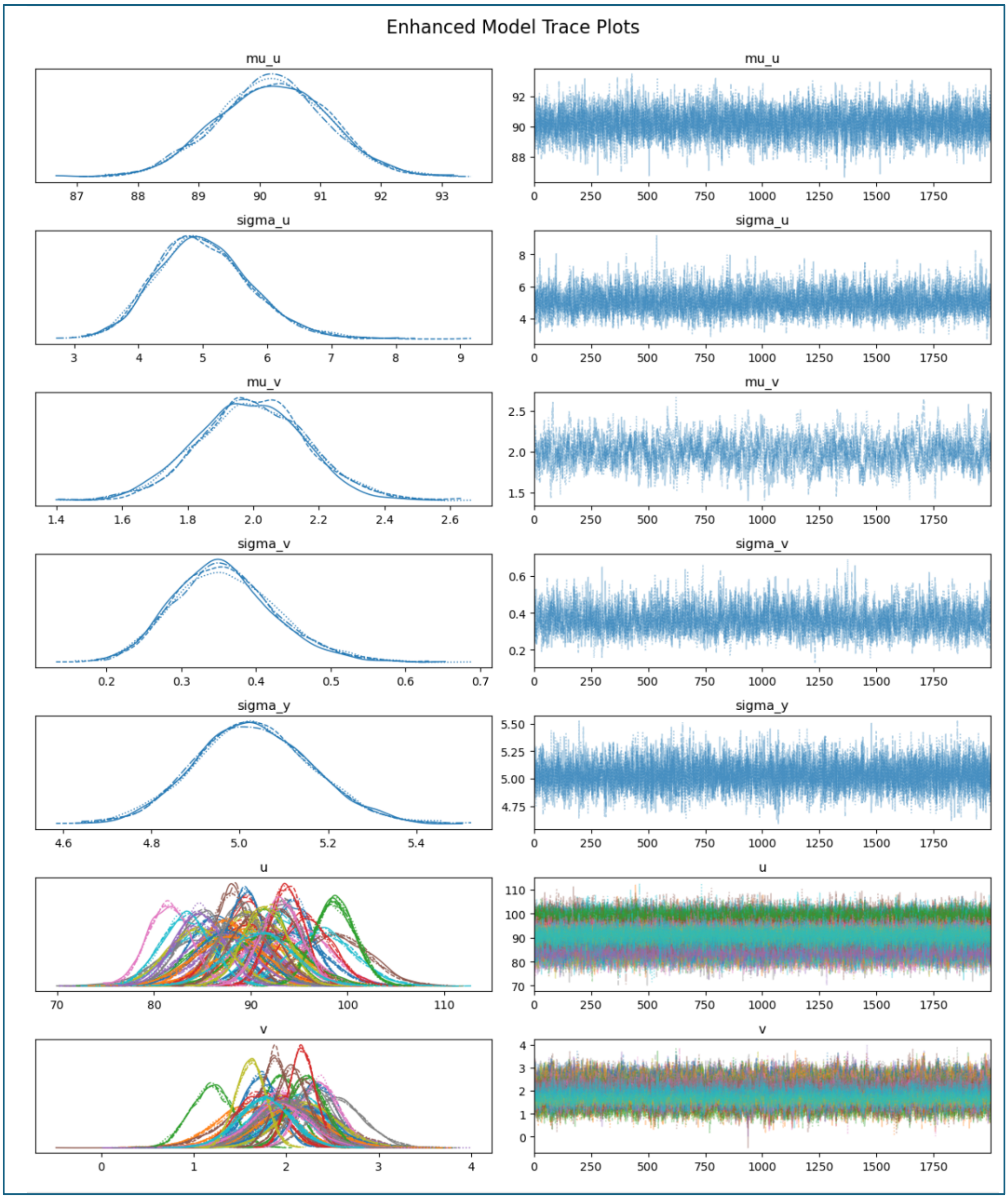


**Figure 4: Trace plots for key parameters of the Enhanced Bayesian hierarchical model**

MCMC diagnostics for the enhanced model indicated strong and reliable performance. No divergences were recorded across any of the chains, and all Gelman-Rubin statistics ($\hat{r}$) were exactly 1.00across all sampled parameters, confirming full convergence. Effective sample sizes ($n_{eff}$) were high across all parameters, including the global feature weights $w_j$.

Trace plots as shown in Figure 4 for hyperparameters $\mu_u$, $\sigma_u$, and $\mu_v$, $\sigma_v$, observation noise $\sigma_y$ , and component-specific parameters $u_i$ and $v_i$ ,showed stable chain mixing, no multimodality, and minimal autocorrelation. Chain behaviours was stable, and samples thoroughly explored the posterior distribution.

These results confirm that the enhanced model reliably inferred both component-level and feature-driven effects, providing a strong foundation for predictive performance and feature importance analysis.

## 3. Model Fit Comparison and Uncertainty Interpretation

Model fits for four representative components, Component 10 (dense data) and Components 53, 56, and 72 (sparse data),are shown side by side for the baseline and enhanced Bayesian models. Each plot displays the posterior predictive mean (solid line) with two uncertainty bands. Function uncertainty (darker band) reflects variability from posterior estimates of latent parameters (such as $u_i$, $v_i$, and, in the enhanced model, $w_j$), capturing epistemic uncertainty in the degradation process. Total uncertainty (lighter band): combines function uncertainty with observation noise $\sigma_y$, representing the full posterior predictive spread.

Both intervals represent 95% credible regions, calculated as ±1.96 times the relevant posterior predictive standard deviation. This dual-band visualization offers insight into the model's confidence in predictions and its handling of uncertainty due to limited or noisy data.

Across all components, both models produce similar overall degradation shapes, especially in the early phase. For Component 10 (with dense data), the predicted trajectories are tightly bounded in both models, reflecting high confidence. In this case, the inclusion of feature effects in the enhanced model has limited impact, as the data itself strongly informs the parameters.

More noticeable differences emerge in Components 53, 56, and 72, which have sparse observations. The enhanced model produces narrower function uncertainty bands than the baseline, indicating improved parameter estimation due to the inclusion of component features via the global weights $w_j$. While the baseline model relies solely on hierarchical priors for these cases, the enhanced model leverages shared patterns across the dataset to better inform its predictions.

Crucially, the enhanced model's total uncertainty remains well-calibrated, balancing confidence and caution. This highlights a key strength of hierarchical Bayesian modelling: its ability to integrate sparse data, structured priors, and relevant features to produce interpretable, probabilistic forecasts.

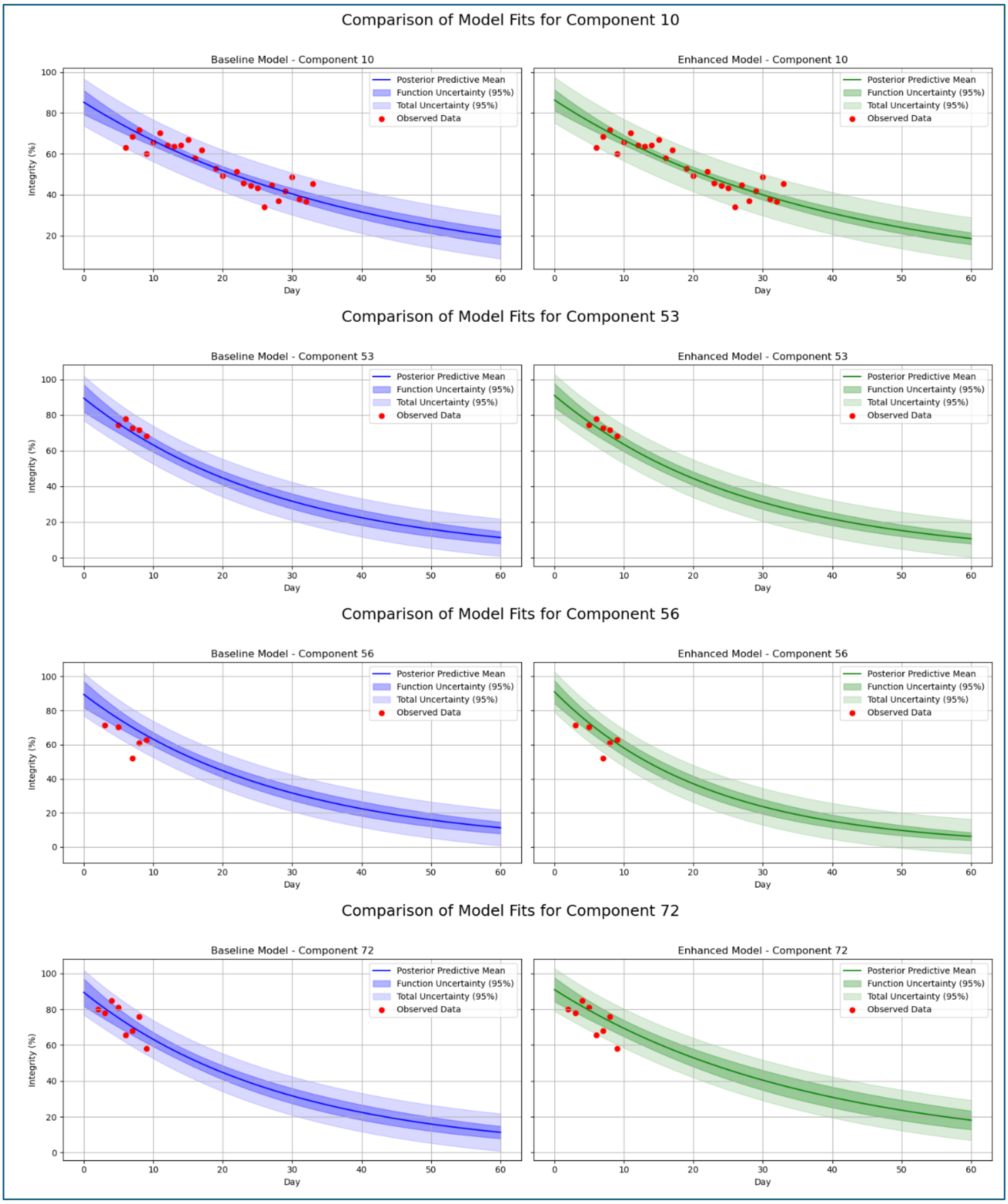

**Figure 5: Comparison of model fits for Baseline and Enhanced Models**

# 4. POTENTIAL CAUSE OF DEGRADATION

To assess the influence of component characteristics on degradation, we analysed the posterior distributions of the feature weight parameters $w_j$ in the enhanced Bayesian model. Figure 6 represents trace plots and posterior distributions for the five weights associated with features X1 to X5.

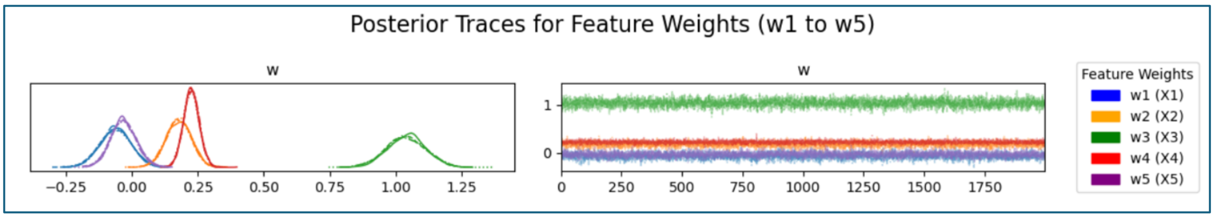


**Figure 6: Posterior distributions and sampling traces for feature weights**

Figure 6 shows the posterior trace plots for the five feature weights after model convergence.

From the figure, it is evident that certain feature weights exhibit significantly non-zero posterior means. Notably, the posterior for $w_3$ is centred distinctly away from zero with a positive mean and low variance, suggesting strong, reliable effect on the degradation rate. This indicates higher X3 values correspond to faster degradation, as $w_3$ contributes positively to the effective decay parameter. Similarly, $w_2$ and $w_4$ also show moderate positive influence, though with slightly greater uncertainty.

In contrast, $w_1$ and $w_5$ are centred near zero with wider distribution, suggesting minimal or no consistent effect on the degradation behaviours.

When predicting component integrity or prioritizing components for maintenance or monitoring, the characteristics associated with X3, and to a lesser extent X2 and X4, should be given primary attention. Meanwhile, efforts focused on X1 and X5 would likely yield minimal returns in terms of predictive power or actionable maintenance strategies.

Overall, the enhanced model successfully identifies and quantifies the relative importance of different component characteristics, providing valuable insight into the underlying drivers of degradation that were not accessible through the baseline model alone.

## 5. ADDRESSING QUERIES

### 5.A. Black-Box Model

To address the engineering query on alternative modelling approaches, a non-Bayesian black-box regression model was developed to predict integrity at t=30 days using early-time data. The feature set included mean integrity, mean log-integrity, average slope from $t \in [0,10]$ and component features X1 to X5.

Only components with available observations up to t=30 were included. A **linear regression model** was trained on these inputs.

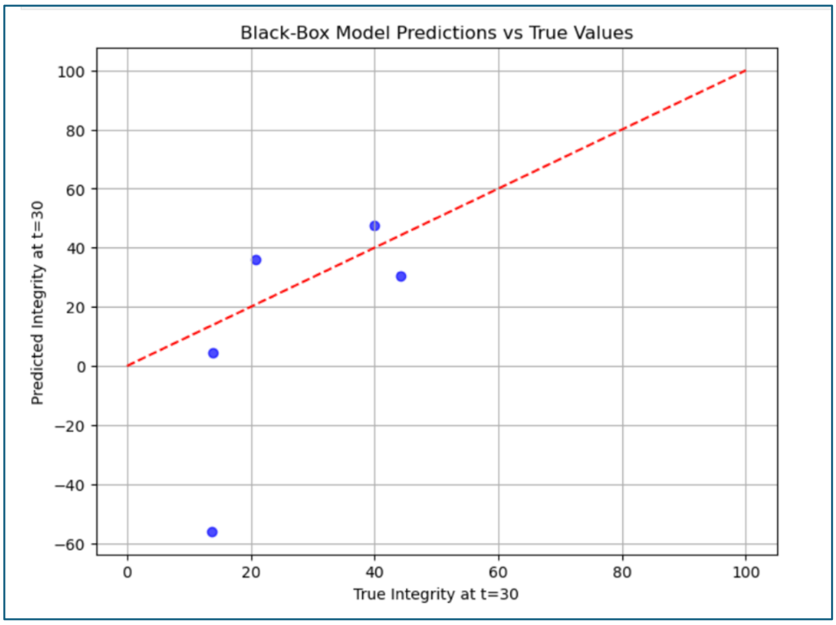


**Figure 7: Black-box model predictions versus true integrity values at t=30**

Figure 7 displays the predicted vs. actual integrity values on the test set, with the red dashed line representing perfect predictions.

While the black-box model captured general trends, it produced notable prediction errors, including some implausibly negative predictions. Predictions for low-integrity components were especially unreliable, highlighting its limitations in extrapolating degradation under uncertainty.

In contrast, the Bayesian hierarchical models incorporated physically realistic behaviours, shared structural information across components, and provided full uncertainty quantification. These advantages make the Bayesian approach far better suited for critical integrity forecasting tasks.

### 5.B. Support Vector Machine

Support Vector Machines (SVMs), while powerful in certain classification and regression tasks, are not well-suited for this application. SVMs generate point estimates without any measure of uncertainty, an essential limitation when forecasting integrity in risk-sensitive engineering contexts. They also lack the flexibility to model hierarchical or structured data, where components share common behaviours but vary individually.

In contrast, hierarchical Bayesian models offer full uncertainty quantification and allow information to be shared across components through partial pooling. This approach accommodates sparse data, respects the physical nature of degradation, and produces interpretable probabilistic forecasts. These capabilities make Bayesian modelling more appropriate for this setting than traditional black-box methods like SVMs.

## 6. BLIND TEST PREDICTION

The enhanced Bayesian model was used to estimate the probability that each of the 25 test components (indices 50–74) would degrade to 30% integrity or lower by day 30. These predictions are based on posterior samples and incorporate both component features and uncertainty.

| ID | index | probability |
|---|---|---|
| K#0050 | 50 | 0.0 |
| K#0051 | 51 | 1.0 |
| K#0052 | 52 | 1.0 |
| K#0053 | 53 | 0.30075002 |
| K#0054 | 54 | 0.0 |
| K#0055 | 55 | 0.947125 |
| K#0056 | 56 | 0.99500006 |
| K#0057 | 57 | 0.0 |
| K#0058 | 58 | 1.0 |
| K#0059 | 59 | 0.591125 |
| K#0060 | 60 | 0.002 |
| K#0061 | 61 | 0.0 |
| K#0062 | 62 | 0.0005 |
| K#0063 | 63 | 0.00175 |
| K#0064 | 64 | 0.0 |
| K#0065 | 65 | 0.99975 |
| K#0066 | 66 | 1.0 |
| K#0067 | 67 | 1.0 |
| K#0068 | 68 | 0.012375001 |
| K#0069 | 69 | 1.0 |
| K#0070 | 70 | 0.993125 |
| K#0071 | 71 | 0.786875 |
| K#0072 | 72 | 0.0 |
| K#0073 | 73 | 0.99112505 |
| K#0074 | 74 | 0.0 |

**Table 1: Predicted Probability of Integrity ≤ 30% at t = 30 for Test Components**

Table 1 summarizes the results, highlighting varying risk levels across components. Full prediction values are provided in the attached **predictions.csv**. This probabilistic approach offers a more informative and risk-aware alternative to deterministic methods.

## 7. CONCLUSION

This study demonstrated the effectiveness of hierarchical Bayesian modelling for predicting component degradation under uncertainty. The enhanced model, which incorporates both population-level structure and component-specific features, outperformed simpler alternatives by providing realistic, interpretable, and risk-aware forecasts. Its blind test predictions offer practical value for maintenance planning, especially where early data is limited.