Introduction

This report evaluates different time-series forecasting techniques for predicting wear particle concentration in gearbox lubricating oil. Since wear particles follow an exponential decay trend, choosing the right model is crucial for accurate predictions.

Why Does the Data Show a Decaying Trend?

In a well-maintained gearbox, the initial wear phase releases a significant number of large and small wear particles into the lubricant. However, over time:

- Large particles settle or get filtered out faster.
- Small particles remain in circulation longer but also decrease gradually.
- With continuous operation, the overall wear rate slows down, leading to an exponentially decreasing particle count in the lubricant.
- Any abnormal spikes in large particle concentration could indicate potential gearbox faults.

Synthetic Data Generation & Model Selection

Since real-world gearbox wear data was unavailable, I generated synthetic data that follows an exponential decay pattern with occasional anomalies (sudden wear spikes) to simulate potential faults. This allowed us to:

- Test different forecasting models under controlled conditions.
- Evaluate their effectiveness in capturing the decay trend and detecting anomalies.

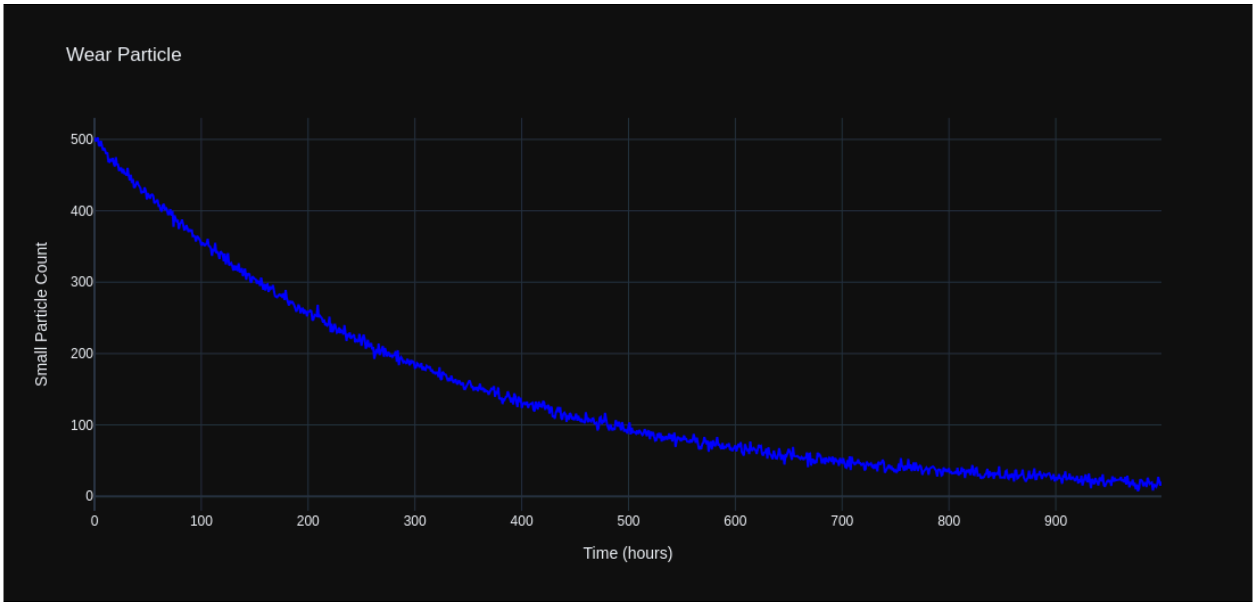
We applied three forecasting models:

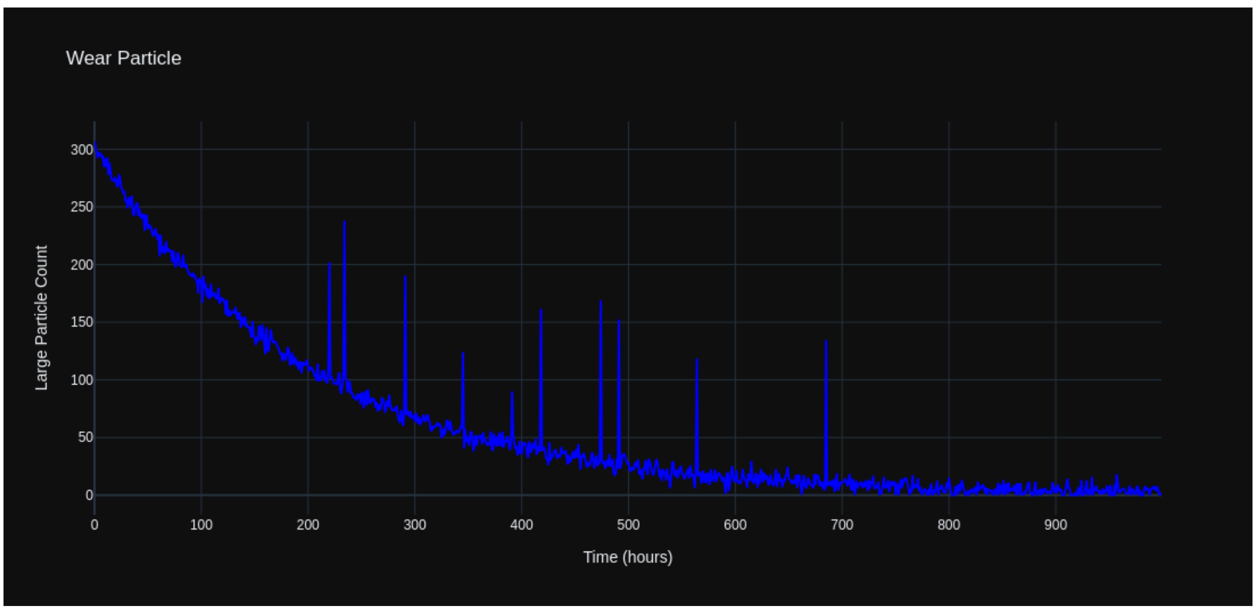
1. ARIMA (AutoRegressive Integrated Moving Average) – Suitable for stationary time-series with structured trends.
2. Exponential Smoothing (Holt-Winters Method) – Directly models exponentially decreasing trends.
3. Gaussian Process Regression (GPR) – Captures smooth, non-linear trends and provides uncertainty estimates in predictions.

These models were chosen to determine which best fits the gearbox wear decay pattern while also allowing early fault detection for proactive maintenance.

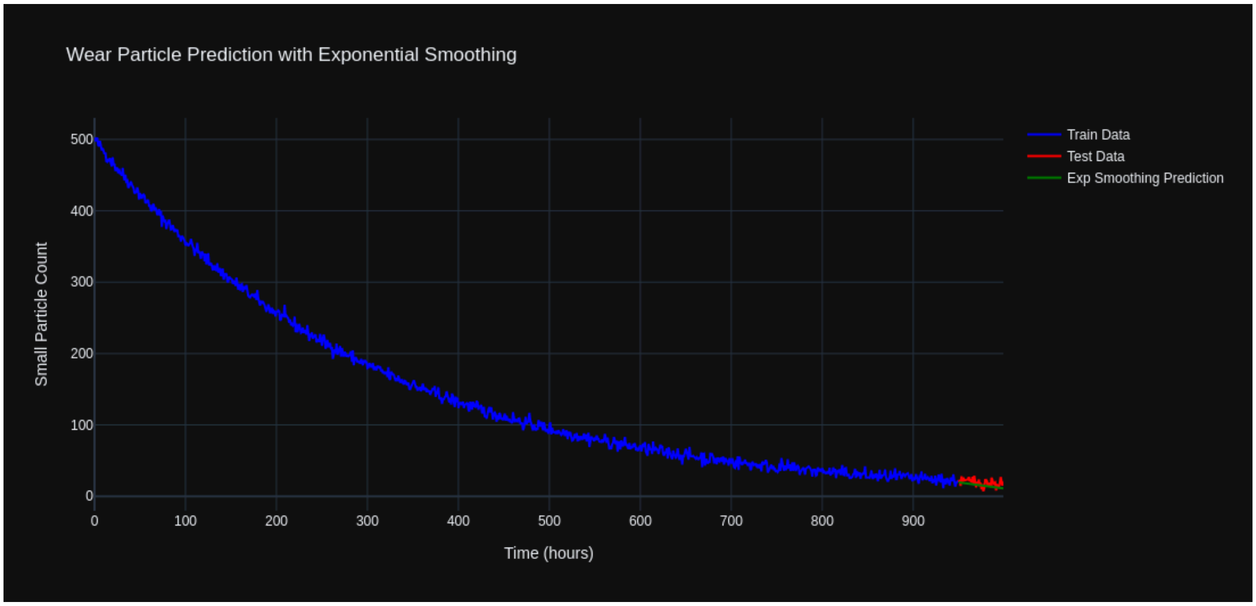
Data Visualization

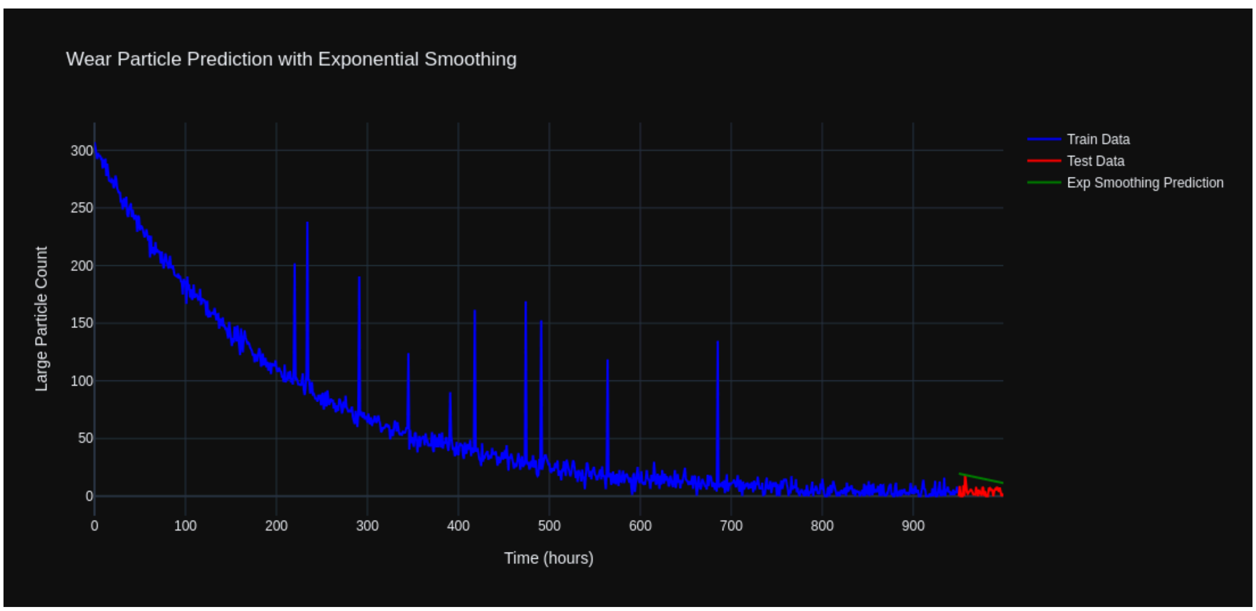
Following are the diagrams of the original data that was generated.



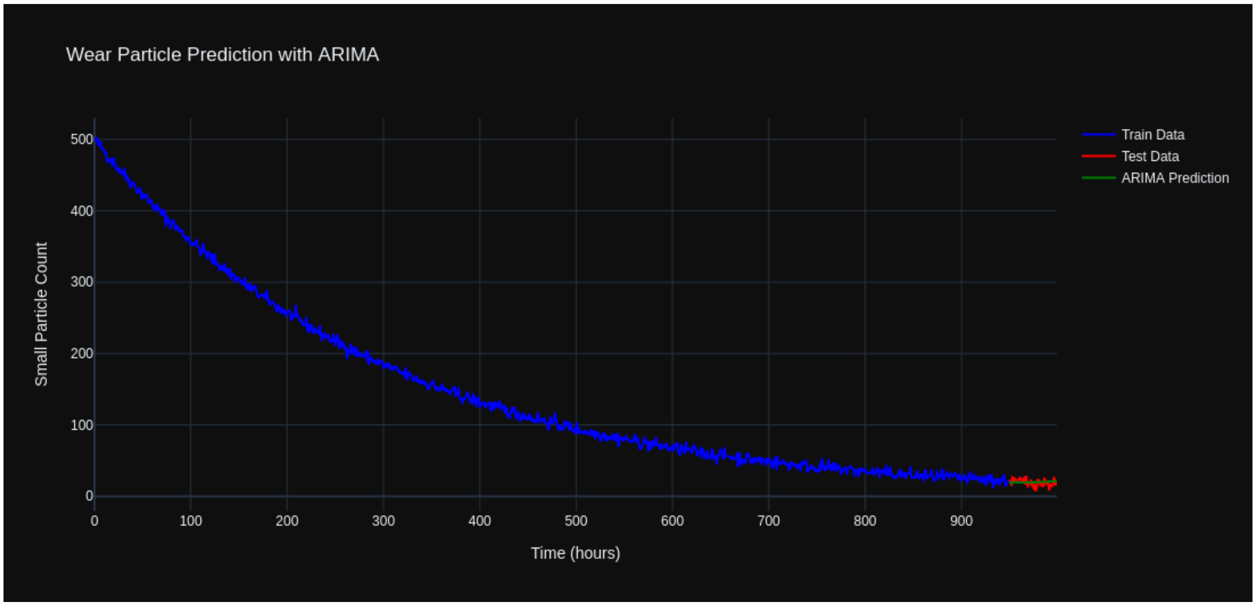


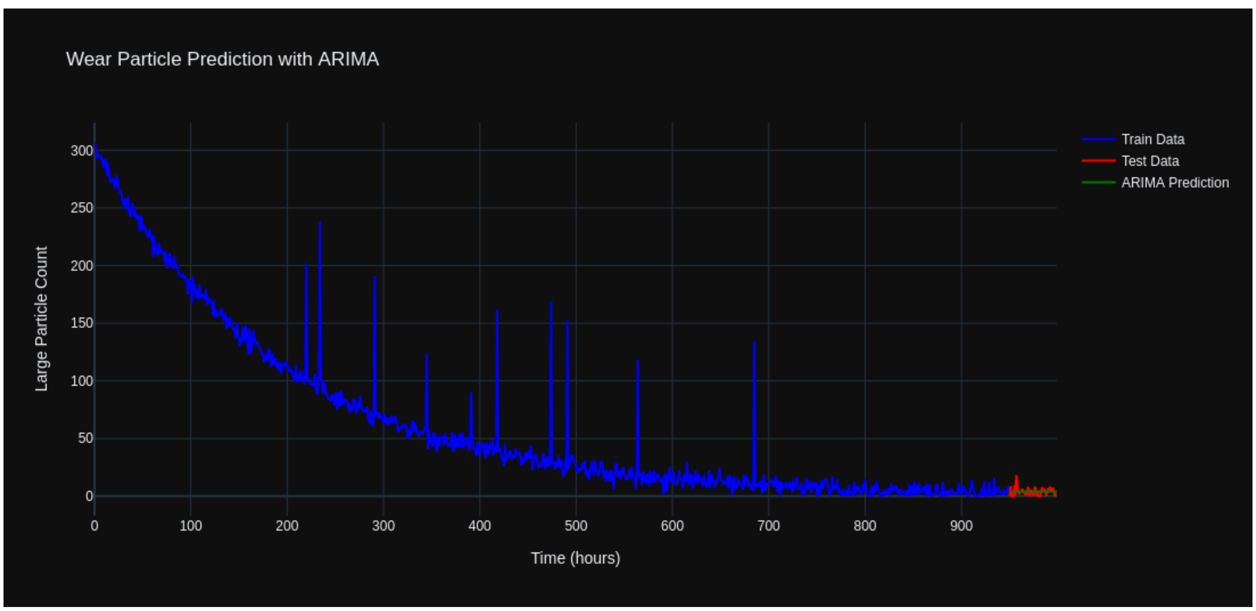
Results of Exponential Smoothing on the data



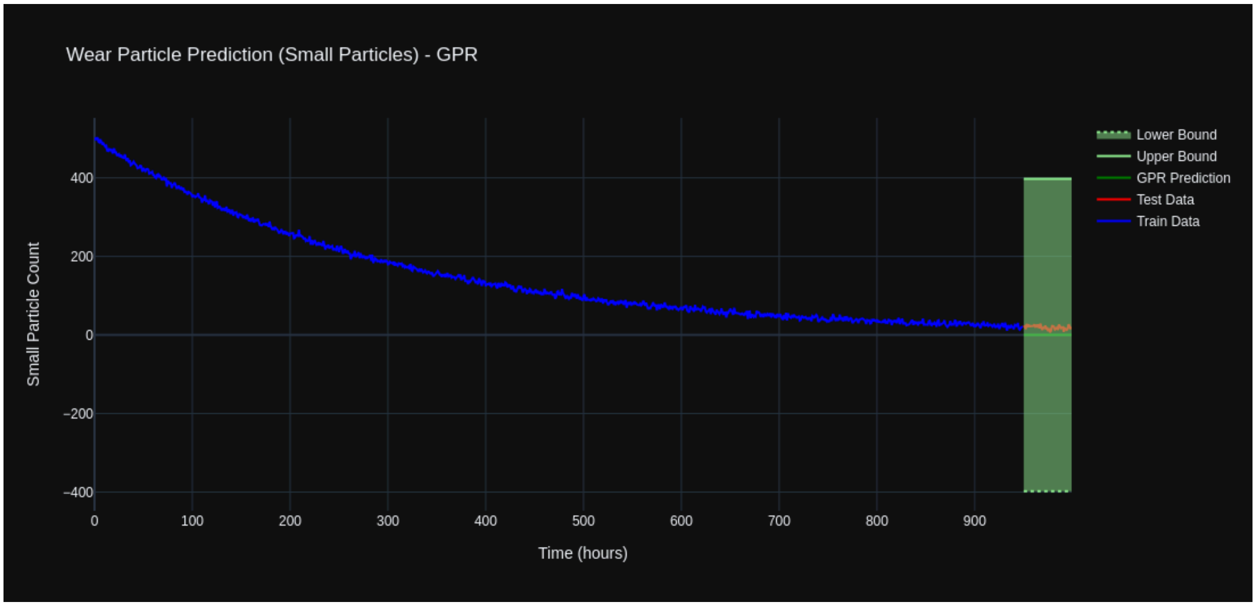


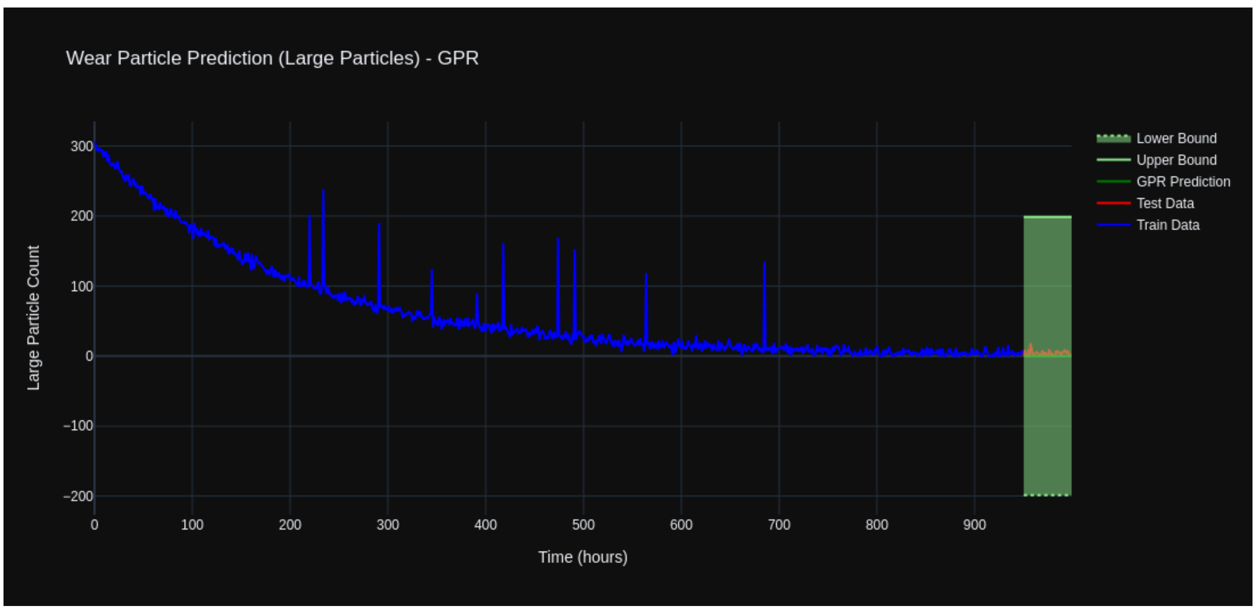
Results of ARIMA on the data





Results of GPR on data





Deep learning models like LSTMs (Long Short-Term Memory Networks) and Transformers can also be used for time-series forecasting, especially for capturing complex dependencies in long sequences. However, these models require large amounts of data and computational resources. Since the wear particle data follows a well-defined exponential decay trend, classical methods like ARIMA, Exponential Smoothing, and GPR are more interpretable, efficient, and sufficient for this task. Deep learning approaches are not necessary when simpler models can achieve accurate predictions.

Data is not normalized as this does not affect the performance of the above models.