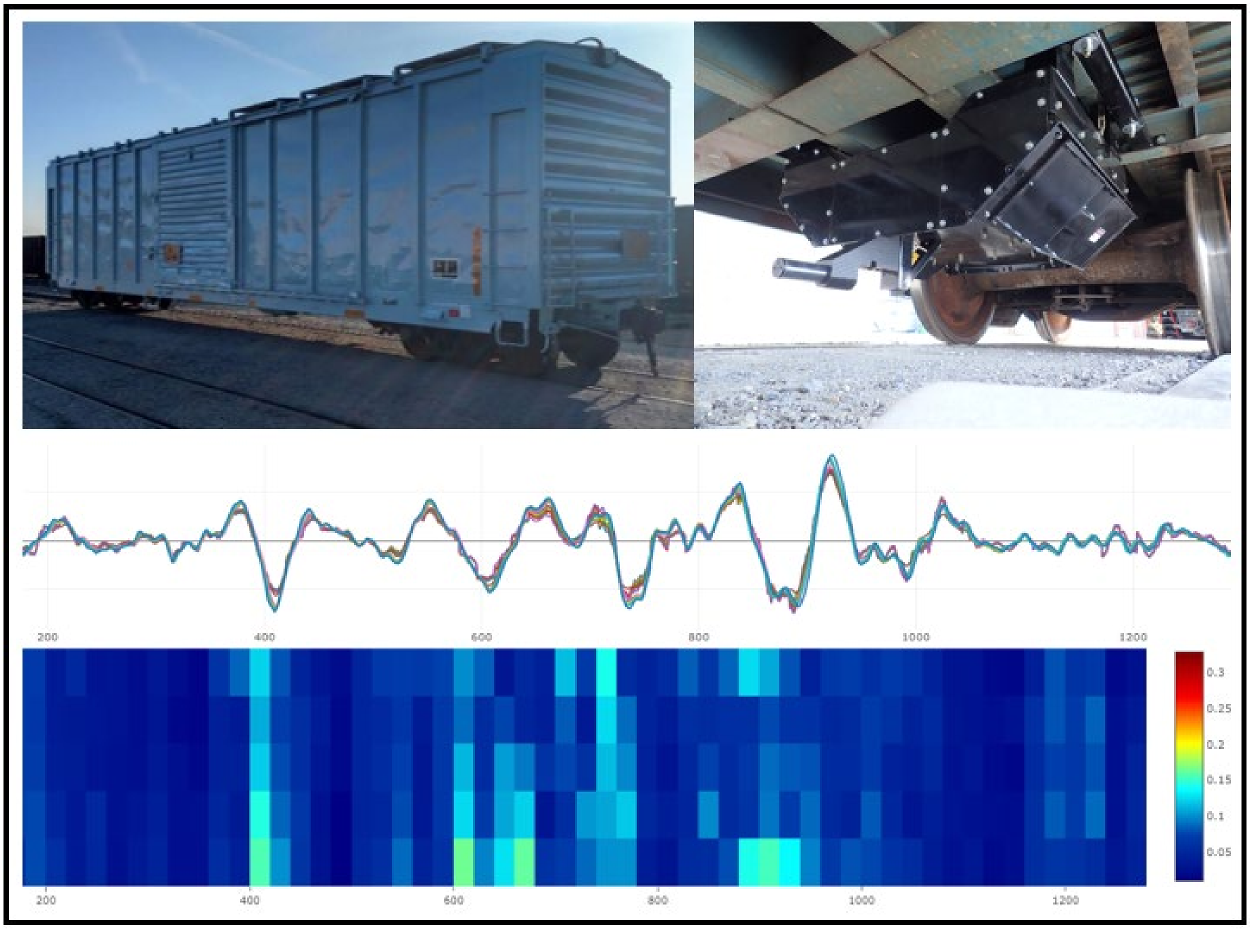
U.S. Department
of Transportation

Federal Railroad
Administration

Office of Research,
Development and Technology
Washington, DC 20590

Advanced Track Geometry Forecasting Methods



DOT/FRA/ORD-24/37

Final Report | September 2024

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED *(From - To)***
9/12/24	Technical Report	September 19, 2020 – September 18, 2021

4. TITLE AND SUBTITLE

Advanced Track Geometry Forecasting Methods

5a. CONTRACT NUMBER
693JJ618D000005

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Serkan Sandikcioglu – ORCiD: 0000-0003-4217-0836,
Radim Bruzek – ORCiD: 0000-0003-3069-9476

5d. PROJECT NUMBER

5e. TASK NUMBER
693JJ6-20-F-000090

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
ENSCO, Inc.
2211 Hydraulic Rd, Suite 301
Charlottesville, VA 22901

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Department of Transportation
Federal Railroad Administration
Office of Railroad Policy and Development
Office of Research, Development, and Technology
Washington, DC 20590

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
DOT/FRA/ORD-24/37

12. DISTRIBUTION/AVAILABILITY STATEMENT
This document is available to the public through the FRA website.

13. SUPPLEMENTARY NOTES
COR: Jay Baillargeon

14. ABSTRACT
The Federal Railroad Administration (FRA) contracted ENSCO, Inc., to apply alternative forecasting approaches to foot-by-foot track geometry data. The research team used track geometry data collected weekly by FRA's DOTX-225 and DOTX-226 autonomous inspection vehicles. This investigation also considered the effect of both seasonality and maintenance activities, which helped enhance the accuracy of the resultant models. The work began in September 2020 and was completed in September 2021.
The research team demonstrated two methods for time series-based forecasting of foot-by-foot track geometry data: 1) point data processing with a SARIMAX model and 2) segment data processing with Facebook's Prophet model. Both methods produced accurate forecasts. The team concluded that these methods, along with frequent measurements from Autonomous Track Geometry Measurement Systems (ATGMS), can accurately predict future behavior of track geometry and help railroads plan preventive maintenance.
Researchers also developed a dashboard to assess the performance of various forecasting models, which graphically showed statistical performance measures for easier interpretation.

15. SUBJECT TERMS
Prediction, forecast, track geometry, time series, degradation, Autonomous Track Geometry Measurement System, ATGMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**		38	**19b. TELEPHONE NUMBER *(Include area code)***

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in)	=	2.5 centimeters (cm)
1 foot (ft)	=	30 centimeters (cm)
1 yard (yd)	=	0.9 meter (m)
1 mile (mi)	=	1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz)	=	28 grams (gm)
1 pound (lb)	=	0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp)	=	5 milliliters (ml)
1 tablespoon (tbsp)	=	15 milliliters (ml)
1 fluid ounce (fl oz)	=	30 milliliters (ml)
1 cup (c)	=	0.24 liter (l)
1 pint (pt)	=	0.47 liter (l)
1 quart (qt)	=	0.96 liter (l)
1 gallon (gal)	=	3.8 liters (l)
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm)	=	0.04 inch (in)
1 centimeter (cm)	=	0.4 inch (in)
1 meter (m)	=	3.3 feet (ft)
1 meter (m)	=	1.1 yards (yd)
1 kilometer (km)	=	0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm)	=	0.036 ounce (oz)
1 kilogram (kg)	=	2.2 pounds (lb)
1 tonne (t)	=	1,000 kilograms (kg)
	=	1.1 short tons

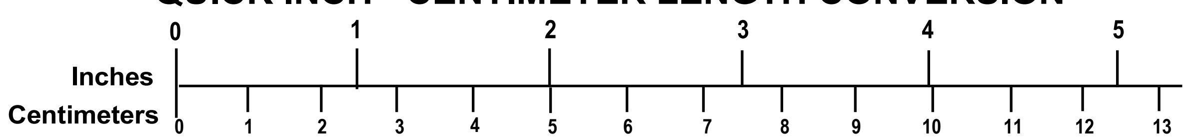
VOLUME (APPROXIMATE)

1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 liter (l)	=	2.1 pints (pt)
1 liter (l)	=	1.06 quarts (qt)
1 liter (l)	=	0.26 gallon (gal)
1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)

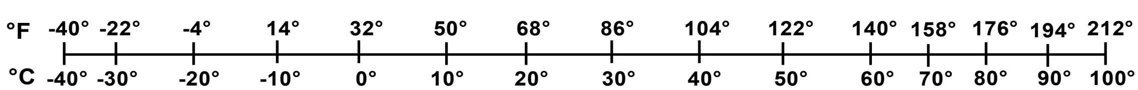
TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSIO



For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98

Contents

Illustrations

Executive Summary

The Federal Railroad Administration (FRA) contracted ENSCO, Inc., to apply alternative forecasting approaches to foot-by-foot track geometry data. The research team used track geometry data collected weekly by FRA's DOTX-225 and DOTX-226 autonomous inspection vehicles. This investigation also considered the effect of both seasonality and maintenance activities, which helped enhance the accuracy of the resultant models. The work began in September 2020 and was completed in September 2021.

The research team demonstrated two methods for time series-based forecasting of foot-by-foot track geometry data: 1) point data processing with a SARIMAX model and 2) segment data processing with Facebook's Prophet model. Both methods produced accurate forecasts. The team concluded that these methods, along with frequent measurements from Autonomous Track Geometry Measurement Systems (ATGMS), can accurately predict future behavior of track geometry and help railroads plan preventive maintenance. Researchers also developed a dashboard to assess the performance of various forecasting models, which graphically showed statistical performance measures for easier interpretation.

This research represents a significant advancement in track geometry forecasting. Previous research initiatives undertaken by FRA's Office of Research, Development, and Technology focused on identified peak-value deviations trending over time. While peak trending is a method particularly effective for identifying near-term issues at locations where identifiable deviations already exist, it is not useful for predicting issues in the long term.

Researchers found that forecasting with point-processed data (in which each surveyed foot is considered a separate time series) requires perfectly aligned track geometry surveys. On the other hand, segment processing (in which the track is divided into short sections) was more tolerant of survey misalignments. Researchers also found that a global forecasting model applied to point-processed data is computationally efficient, although it can generate erroneous forecasts for time series that do not fit the global model. Additionally, while automatically assigning forecast model parameters for each time series eliminated the errors, this also increased the computation time.

During the project, researchers found discontinuities when forecasts for separate track segments were joined into a continuous track geometry. Forecasting with overlapping segments and subsequently removing overlapping data was found to be a good solution to this issue. The team also addressed the issue of accounting for unexpected maintenance activities, as it is common for maintenance to occur between surveys. For this issue, researchers included maintenance activity as an external variable when forecasting with both point- and segment-processed data. Without this external variable, the model would treat the time series containing signature changes due to maintenance as normal historical data with an unexplained variation.

Additional effort will be necessary to deploy these models into revenue service. This report identifies further research needed to advance the maturity of the point and segment forecasting methods for industry acceptance and deployment. Unique approaches presented in this report, especially the segment forecast methodology, may also benefit studies into other railroad assets.

1. Introduction

The Federal Railroad Administration (FRA) contracted ENSCO, Inc., to apply alternative forecasting approaches to foot-by-foot track geometry data collected weekly by FRA's DOTX-225 and DOTX-226 autonomous inspection vehicles. This investigation also considered the effect of both seasonality and maintenance activities, which helped enhance the accuracy of the resultant models. The work began in September 2020 and was completed in September 2021.

1.1 Background

Predicting track geometry signatures, and, in turn, defects can play an important role in proactively addressing safety issues. Previous FRA research focused on trending individual peak-value deviations from track geometry measurements (Bruzek, Stark, Sussmann, Tunna, & Thompson, 2022). To improve upon this capability, FRA undertook research to advance forecasting methods that 1) use continuous foot-by-foot geometry measurements taken directly from inspection vehicles, and 2) are suitable for autonomous processing. Successful implementation of foot-by-foot track geometry forecasting will expand the capabilities and application of Autonomous Track Geometry Measurement Systems (ATGMS) and related data products.

1.2 Objectives

The goals of this research were to:

- Investigate advanced analytical approaches for forecasting foot-by-foot track geometry.
- Develop and evaluate an advanced predictive model that takes into consideration the effects of both seasonality and maintenance.
- Develop a preliminary user interface for selected model(s).
- Demonstrate the capability of the developed model(s) through selected case studies.

1.3 Overall Approach

The team developed a functional predictive tool that can be integrated into existing FRA data management systems. First, researchers conducted a literature review to investigate similar approaches and applications in the railroad and other industries. Researchers then investigated the performance of multiple forecasting models on track geometry and selected the best candidates for further investigation. Finally, the team developed a computer dashboard to facilitate investigation and development of forecasting models and to visualize results.

1.4 Scope

The scope of the work was restricted to data collected once per week on average within the territory of a specific Class I freight railroad by DOTX-225 and DOTX-226 ATGMS vehicles. DOTX-225 and DOTX-226 are refurbished freight boxcars with a carbody-mounted ATGMS designed for unmanned operations in revenue service. The scope did not include data collected less often or by other track geometry measurement systems, although the methods developed in this research may be useful in forecasting the condition of other railroad assets.

1.5 Organization of the Report

The rest of this report is organized as follows:

- Section 2 describes track geometry as a time series and provides results of the literature review of forecasting models.
- Section 3 explains two ways of processing track geometry: point processing and segment processing. It also describes the statistical analysis used in this research and a method of displaying results.
- Section 4 describes results from applying the forecasting models and improvements that were made to the models.
- Section 5 provides the overall conclusions of the study.
- Section 6 provides recommendations for further work.
- Appendix A includes more examples of results from the developed methods.

2. Literature Review

A literature review was conducted to find examples of forecasting with time-series data from both the railroad and other industries. Researchers focused on time-series applications because repeated measurements of track geometry can be treated in this way. Figure 1 explains how track geometry and time series are related.

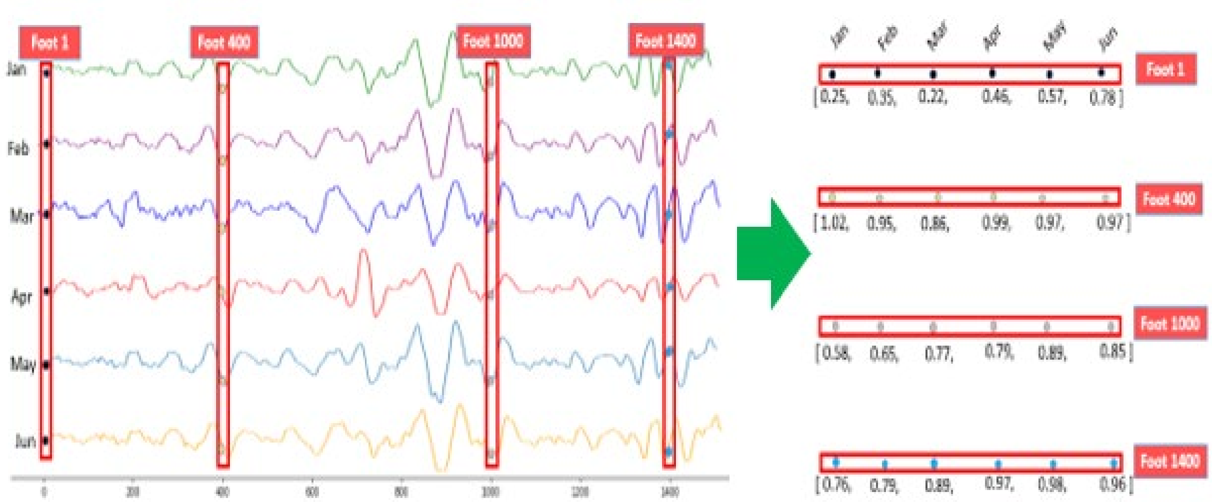


Figure 1. Track Geometry as Time Series

The left side of Figure 1 shows six successive track geometry measurements with distance in feet on the x-axis (four columns are labelled Foot 1, Foot 400, Foot 1000, and Foot 1400). The values within each column are time-ordered observations from different surveys with constant, monthly intervals. The right side of Figure 1 shows these observations as four time series with time on the x-axis.

The team found examples of time series forecasting models being used in several industries and research areas. Palese, et al., cited examples in the railroad industry of time series forecasting for operations and goods demand (Palese, Zarembski, & Attoh-Okine, 2019). In the same report they detail their own application of time series forecasting for rail wear. Researchers have also evaluated models for railway car brake failures (Pacella & Anglani, 2008).

The literature review found references to three alternative time-series forecasting models: ARIMA, SARIMAX, and Facebook's Prophet. These are described in the following sub-sections.

2.1 Autoregressive Integrated Moving Average (ARIMA)

Autoregressive Integrated Moving Average (ARIMA) time series forecasting models are widely used by both academia and researchers to predict different aspects of real-life scenarios and applications ranging from the rate of spread of Covid-19 to supply chain and demand control.

ARIMA models are denoted as ARIMA(p,d,q), where the parameters are non-negative integers. The autoregressive parameter of ARIMA, denoted by p, indicates that the variable is regressed with its own prior (i.e., "lagged") values. The integrated parameter, denoted by d, is the difference of the observation values and lagged values (i.e., the value immediately prior to the current value). The moving average parameter, denoted by q, represents a moving average process based on lagged error terms (Jakasa & Androcec, 2011). Random-walk, random-trend, autoregressive, and exponential smoothing models are all special cases of ARIMA models (Nau, 2022).

2.1.1 *Medical Focus and Covid-19 Related Research*

There is extensive research dedicated to forecasting Covid-19 related variables, such as spread rate and time to reach a critical level. Based on the AIC[1] criterion, researchers found that the best model for the incidence of Covid-19 in Russia was ARIMA(3,2,4). On the other hand, the best model for Covid-19 incidence in Brazil was ARIMA(4,2,3) (Kufel & Kufel, 2020).

In addition to studies made on a national level, there are also applications of ARIMA on a global scale. For example, ARIMA(1,0,4) was selected as the best model for the prevalence of the Covid-19 virus worldwide, while ARIMA(1,0,3) was selected as the best model for determining the worldwide incidence of Covid-19 virus (Benvenuto, Giovanetti, Vassallo, Angeletti, & Ciccozzi, 2020).

2.1.2 *Energy Industry*

Forecasting solar radiation has recently become a focus of research due to the growing interest in green energy. For example, an ARIMA model was used to predict average monthly solar radiation with seasonal lag in Seoul, South Korea. This application mimicked the seasonality and monthly cyclic nature of solar radiation with excellent performance (Alsharif & Younes, 2019).

2.1.3 *Supply Chain and Demand*

Forecasting demand-driven supply chains is important for consumer goods companies. The market has evolved into an environment with customers dictating to suppliers what products they desire and when they need them delivered. Demand forecasting is crucial for inventory management since stock levels depend on demand forecasts and inaccurate estimates of demand can incur significant costs. Consequently, many companies invest in large inventories to avoid stock shortages, although carrying excess inventory also has significant associated costs. One research effort conducted for a food company found that ARIMA(1,0,1) can be used to model and forecast future food demand. The results provide manufacturing managers with a reliable decision-making strategy (Fattah, Ezzine, Aman, El Moussami, & Lachhab, 2018).

2.2 SARIMAX

SARIMAX is an extension of the ARIMA model formed by including lags and differences to fit seasonal patterns (S). An exogenous variable (X) is also added to the model to handle the effects of external factors, such as maintenance. Adding external variables can explain the sudden and significant changes caused by external effects (Andrews, Dean, Swain, & Cole, 2013).

---

[1] The Akaike Information Criterion (AIC) is a method for evaluating the models. It compares the quality of a set of statistical models with each other and determines which one is the best fit for the data.

2.3 Facebook Prophet

Prophet, an open-source software developed by Facebook's core data science team (Taylor & Benjamin, 2017), is a library of time-series forecasting based on additive models in which non-linear trends are fit with time and seasonality effects. It supports forecasting with time series that have strong seasonal effects and several seasons of historical data. According to the developers, Prophet is robust to missing data, trend shifts, and outliers (Facebook, 2022).

3. Methodology

The research team investigated two methods for processing track geometry data: point processing and segment processing. These are described in Sections 3.1 and 3.2. Section 3.3 describes an improvement that was made to the standard segment processing approach, while Section 3.4 describes the statistical analysis methods used in this research. Researchers developed a method for displaying the results of the analysis, which is described in Section 3.5. Finally, Section 3.6 describes methods for handling misaligned track geometry time series.

3.1 Point Processing

Researchers chose to use ARIMA and SARIMAX for forecasting point-processed track geometry data. In the point processing method, every location (i.e., foot measurement) represents a time series of track geometry data. A forecast is made for each foot in the survey. The combination of forecasted values is the forecasted track geometry signature (shown in Figure 2).

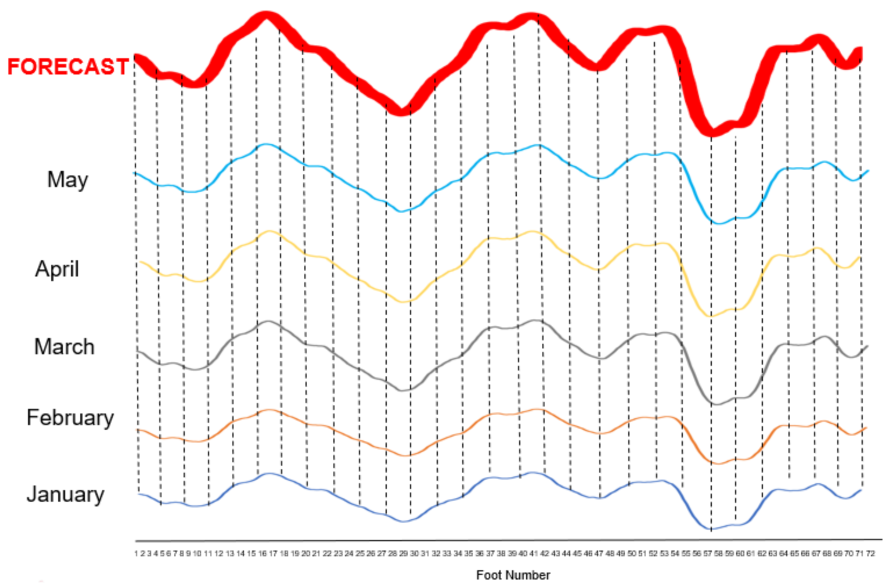


Figure 2. Point Processing Method

Each vertical dashed line in Figure 2 represents a time series of historical data used to predict the forecast depicted by the red line. The point processing method requires repeated track geometry measurements to be perfectly aligned or the forecast will be made on corrupted data. Segment processing is a potential solution to this alignment problem.

3.2 Segment Processing

Researchers chose Facebook's Prophet foreasting model for segment-processed data because it outperformed SARIMAX (see Appendix A). The segment processing method divides track geometry data into subsets of multiple successive data points. The number of data points in a segment is configurable to optimize the performance of the model. The examples given in this report use 24-foot segments (i.e., 24 track geometry data points since the last excluded).

The upper part of Figure 3 shows the data in Figure 2 divided into three segments of track geometry data, each 24 feet long. Five measurements are shown for each segment. The lower part of Figure 3 shows each segment represented as a time series, with time on the x-axis.

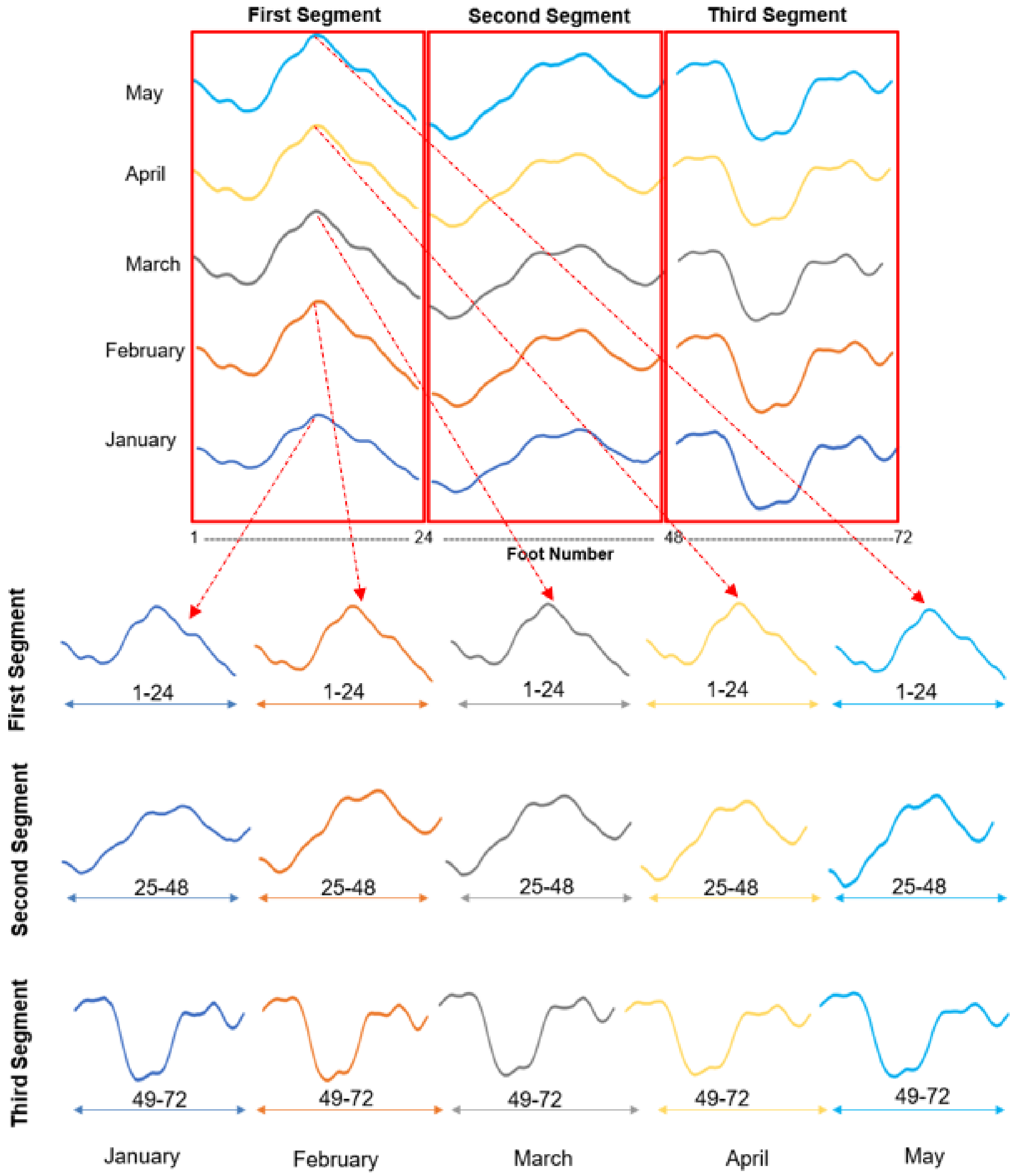


Figure 3. Segment Processing Method

Figure 3 shows the segmented data has a similar shape over time. The shapes can be used as inputs for seasonal time-series models or pattern recognition algorithms (e.g., neural networks) to forecast the next shape, as shown in Figure 4. Figure 5 shows how joining the forecasted segments form the forecasted foot-by-foot track geometry.

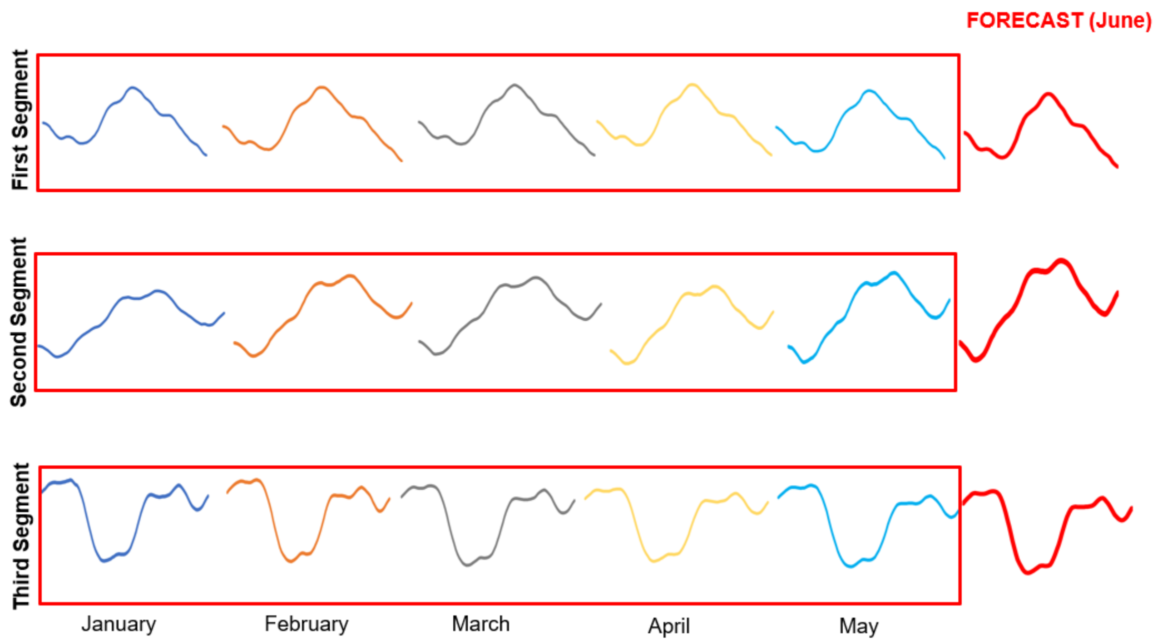


Figure 4. Forecast Segment Shapes

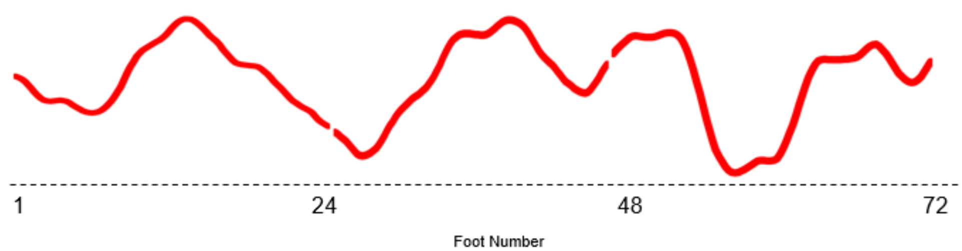


Figure 5. Combining Segments

Researchers developed a special methodology to facilitate connecting adjacent segments to avoid discontinuities in forecast foot-by-foot data. This methodology is described in Section 3.3. Segment forecasting is advantageous because it is more lenient toward data misalignment than point forecasting. In addition, it allows the use of different pattern recognition algorithms to improve robustness.

3.3 Improved Segment Processing

When forecasting with segment-processed data, researchers had to overcome the challenge illustrated in Figure 6. The black box at the top left of Figure 6 represents the overlaid data from four consecutive surveys over the first track segment. This data is input to the forecasting model one survey at a time. The Prophet forecasting model only works with continuous data, connecting the last point in one survey to the first point in the next. This creates the imaginary vertical lines marked by question marks in Figure 6. These lines corrupt the forecast of the first and last three data points in each segment, as shown in the July forecast in Figure 6. Finally, when two segments are joined, spikes can occur, as shown on the right of Figure 6.

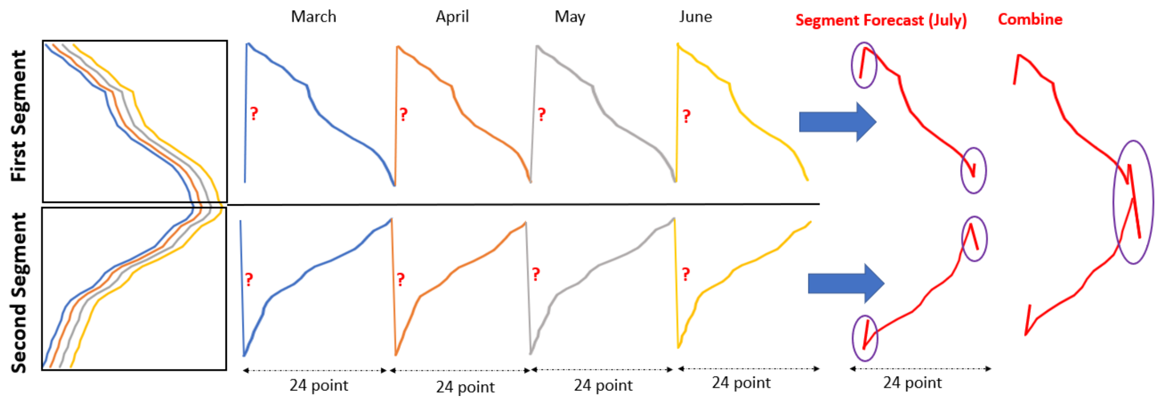


Figure 6. Discontinuities Arising at Segment Boundaries

Figure 7 illustrates the effect on a longer set of measurements. The track geometry data used in this example is the vertical mid-chord offset of the left rail on a 62-foot chord (LPROF62). The parts of the profile in Figure 7 colored red are the spikes introduced by segment-processing. This degree of corruption makes the forecast unusable.

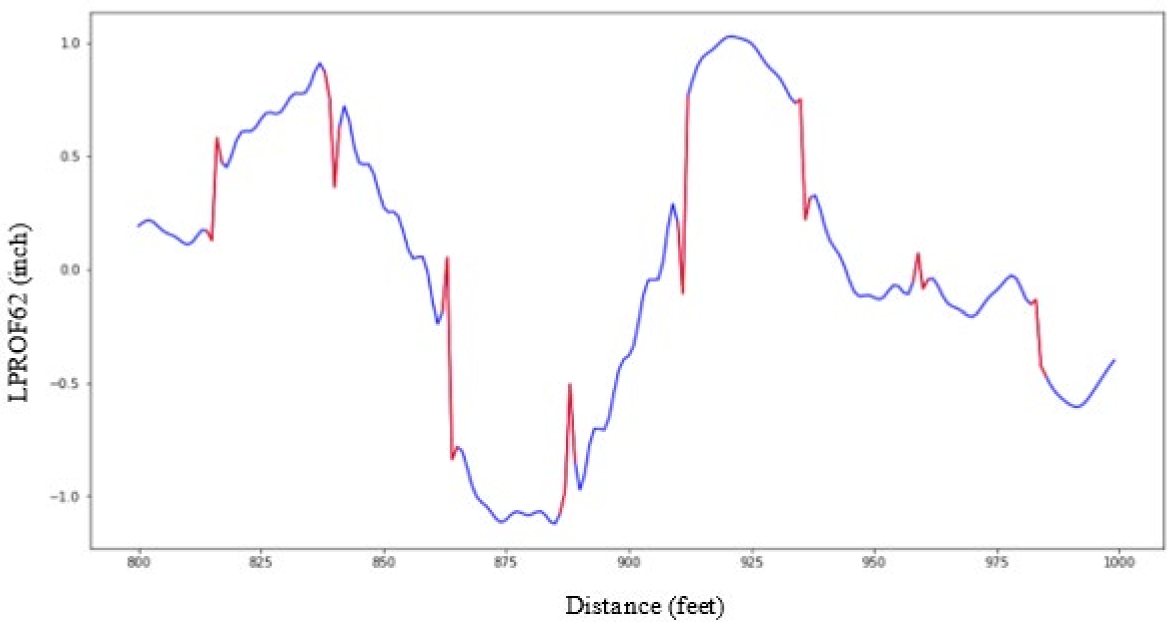


Figure 7. Segment Forecast with Spikes

To solve this problem, researchers modified the methodology to use overlapping segments, as shown in Figure 8, where each 24-foot segment overlaps the next segment by 8 feet. Separate forecasts are made for each segment and then combined, as shown in Figure 9.

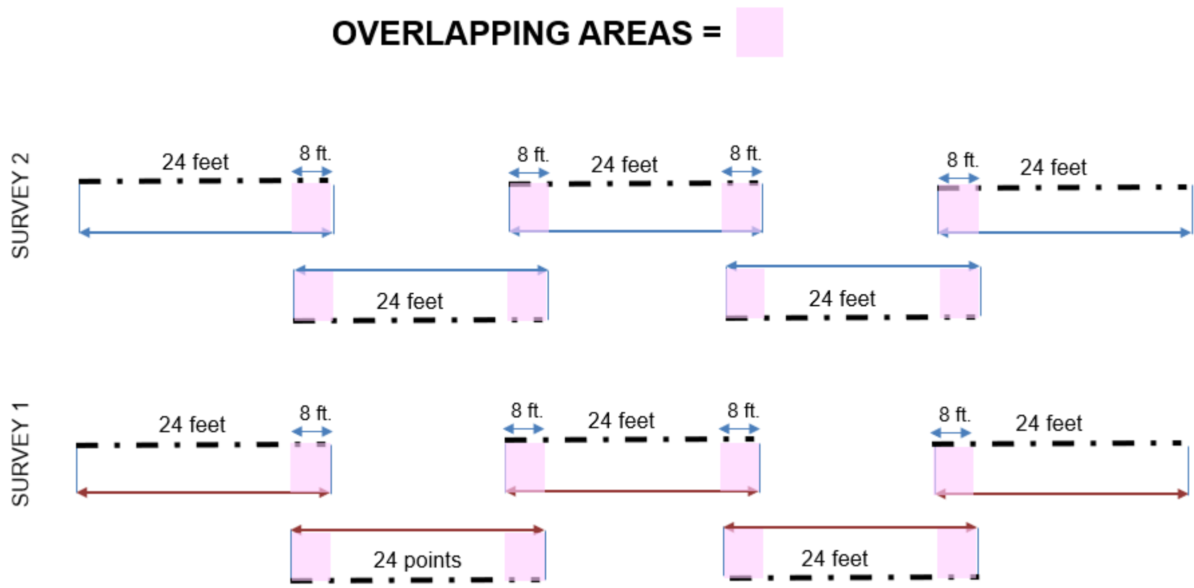


Figure 8. Overlapping Segments

Figure 9 shows that four feet at the beginning and end of each segment (in red) are removed before the forecasts are combined. Since the discontinuities were found in the first and last three feet of data, this eliminates the data spikes and results in a smooth forecast. While overlapping the segments this way eliminates the spikes, it increases the number of segments to be processed.

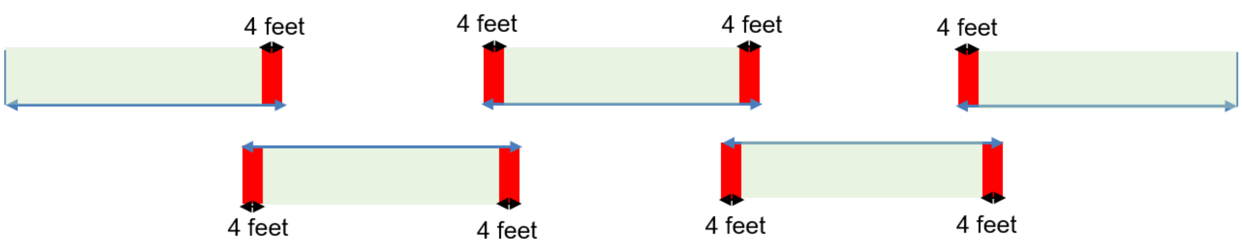


Figure 9. Removal of Data from Forecasts

3.4 Statistical Analysis

The track geometry data used in this research was tested for stationarity. A stationary time series has statistical properties (e.g., mean, variance, and autocorrelation) that are constant over time. For time-series forecasting models, stationarity is important; if the statistical properties are changing over time, the forecasting model will not be able to determine the model parameters properly. The team used the Augmented Dickey-Fuller (ADF) test for stationarity. The null hypothesis, H_0, is that the time series is nonstationary and contains a unit root, while the alternate hypothesis, H_1, supports the time series being stationary.

If a time series is found to be nonstationary, it may be converted to a stationary form by calculating the difference between its value at time t, (X_t), and its previous value (X_{t-1}). The result is called a first-differenced time series. Once a time series becomes stationary, its autocorrelation function (ACF) and partial autocorrelation function (PACF) may help identify the values of the p and q parameters for the ARIMA forecasting model.

Researchers performed a statistical analysis to compare forecast results with subsequent track geometry surveys. The team used the following measures:

- Root mean square error (RMSE) – the square root of the average of squared differences between the forecast and measured values
- Mean absolute error (MAE) – the average of absolute differences between the forecast value and measured value
- Peak absolute difference – the maximum difference between forecast and measured values

3.5 Evaluation Dashboard

Researchers developed a computer dashboard to display data and facilitate the investigation of the forecasting models. The dashboard allows users to vary model parameters and settings, and calculates and displays multiple statistical performance measures. Figure 10 shows an example dashboard display.

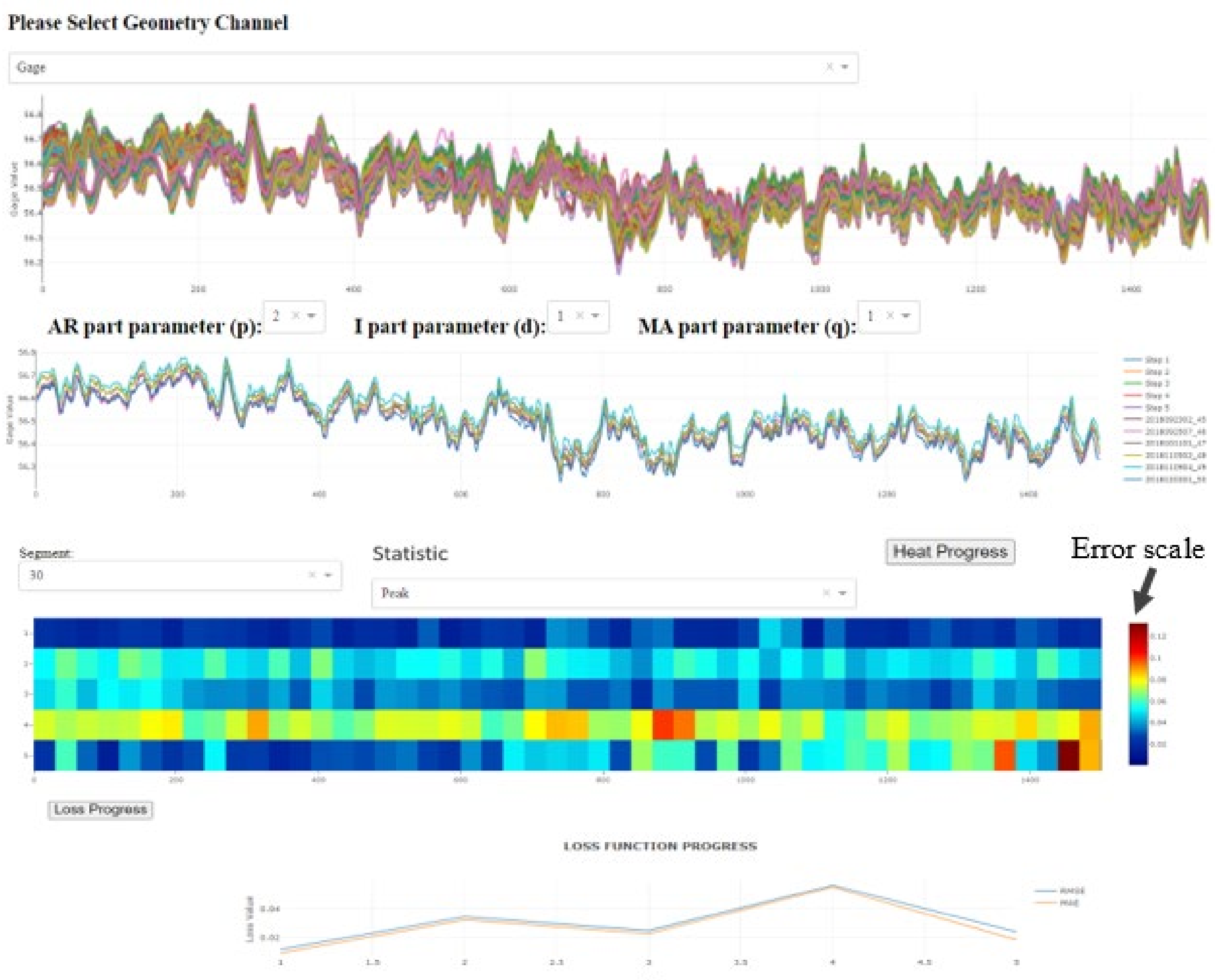


Figure 10. Example Dashboard Display

The box at the top of the dashboard lets the user select the track geometry parameter. Figure 10 shows that the track gage has been selected. The upper chart in the dashboard shows the measured track geometry data. In this example, the length of the dataset is 1,500 feet.

The boxes below the measured data let the user select the global ARIMA parameters to be used for forecasting. The dashboard can be used for ARIMA, SARIMAX, and Prophet forecasting models, and can also be used for point- and segment-processed data.

The central chart in the dashboard shows the forecast data. Forecasts are shown for five future time periods, and the surveyed data for those time periods are also plotted.

The colored block chart in the dashboard shows the results of statistical analysis on the data. The central box above this chart lets the user select the statistic to be displayed. Figure 10 shows peak absolute difference has been selected, which is then shown as a color-coded heat map. The error scale shows the range for each color. Each row in the heat map represents a forecasting step. In this example, the track geometry surveys were collected once per week, and forecasts were generated in one-week intervals. Thus, the top row in the heat map shows results for one week into the future and the bottom row five weeks into the future.

The box on the left above the colored block chart lets the user select the width for each block in the heat map. This does not have to be the same as the segment length used for segment processing, but can be from a few feet to the length of the entire survey. Statistics calculated with short segments show how accurately specific peak values are forecast. Peak values are reported by a traditional track geometry inspection system and represent direct safety concerns. Statistics calculated on long segments can help determine the accuracy of the forecasting model in calculating track quality indices for long-term maintenance planning.

Degradation rates on railroad track are highly variable. Long sections of track can remain stable with little changes in track geometry for months or even years. Other sections can degrade to safety-critical levels in the span of days or weeks. Such locations are often associated with fouled ballast, unstable subgrade, broken or otherwise compromised rail joints, ties, or other superstructure components. Accurate forecasting is very challenging in these unstable locations, so the ability to adjust the segment length is another useful feature of the dashboard.

3.6 Data Misalignment

Track geometry data is perfectly aligned if any specific location (in feet) is the same in every survey. This is ideal but sometimes difficult to achieve. Misaligned data can produce invalid forecasts. Figure 11 illustrates good and bad alignment.

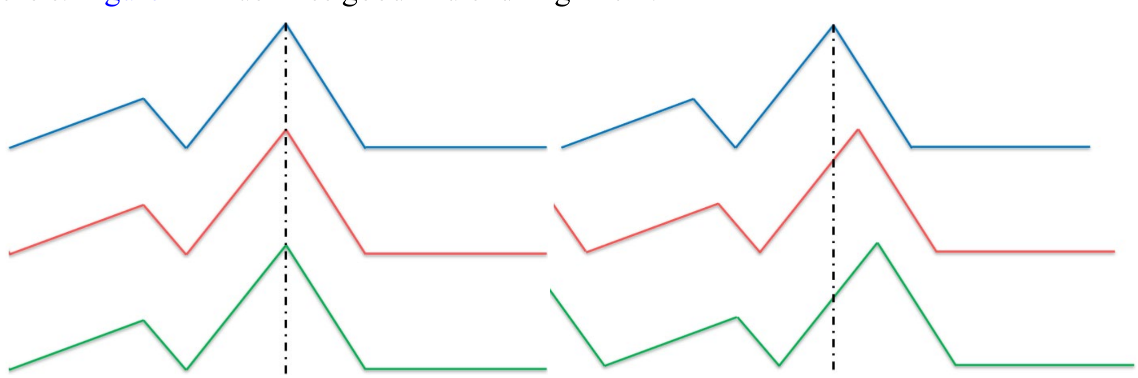


Figure 11. Aligned and Misaligned Data

The left side of Figure 11 shows three surveys in perfect alignment (i.e., the peak of the triangular shape is always at the same location). The right side of Figure 11 shows the same data misaligned. The point data at the vertical dashed line differs from one survey to the next not because of changes in the data but due to misalignment. Thus, data misalignment is a serious problem for point processing, although segment processing can still produce useful results when data is misaligned. Figure 12 explains how misaligned data can be used.

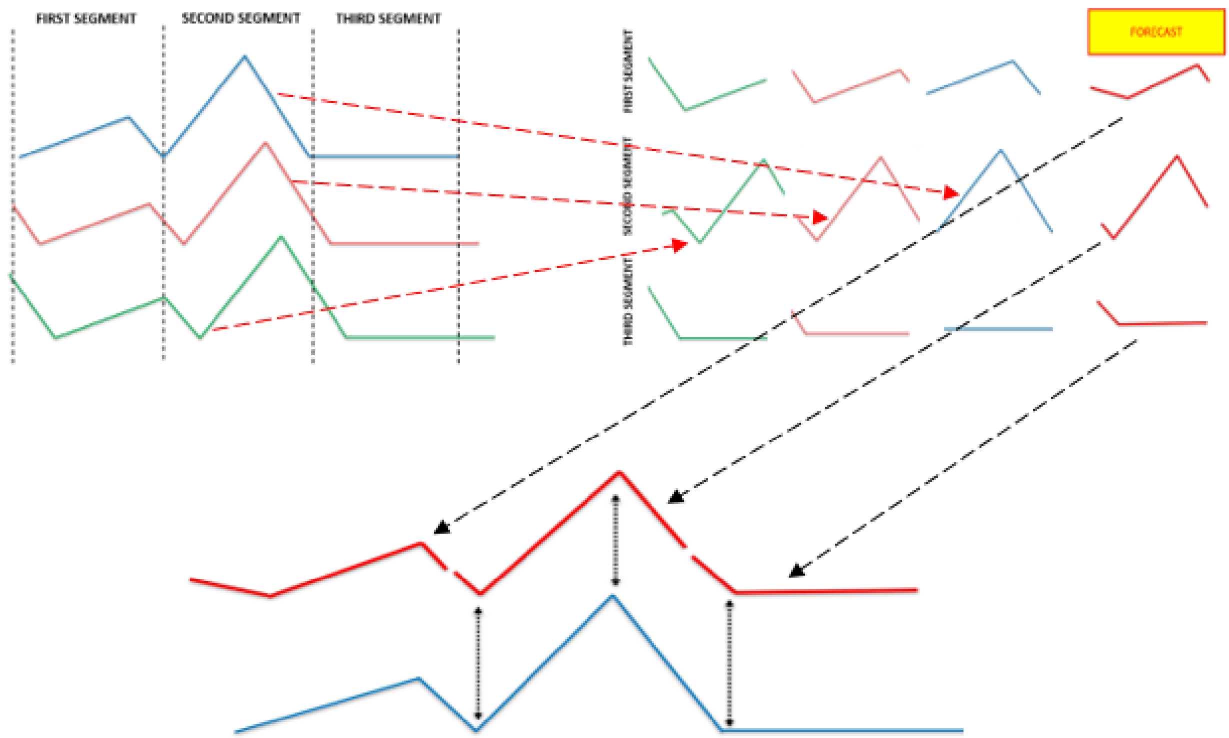


Figure 12. Segment Processing Misaligned Data

The top of Figure 12 shows the segmentation of the data into time series. The blue line is the most recent survey, and the red line is the forecast. The forecast for each segment is affected by the misalignment, as shown on the top right of Figure 12. However, as shown in the bottom of Figure 12, when the segment forecasts are combined, the resulting forecast may be acceptable.

4. Results

Section 4.1 provides an example of applying ARIMA to point-processed data. Section 4.2 shows how the results can be affected by track maintenance. Section 4.3 provides examples of SARIMAX and Prophet applied to point- and segment-processed data including track maintenance. Sections 4.4 and 4.4 describe two methods for improving forecasting accuracy.

4.1 ARIMA Results

Researchers used the ARIMA forecasting model on LPROF62 track geometry data from a 1,500-foot section of track. Eighty-three surveys were collected on this section of track between May 2017 and January 2020 in approximately one-week intervals. Researchers used the first 82 surveys as a training set and forecasted the final survey, which was compared to the forecast.

Researchers used point processing on the LPROF62 data to produce 1,501 time series. Each time series contained 82 data points from successive surveys. As an example, Figure 13 shows all data points from 82 successive surveys for the time series at 400 feet.

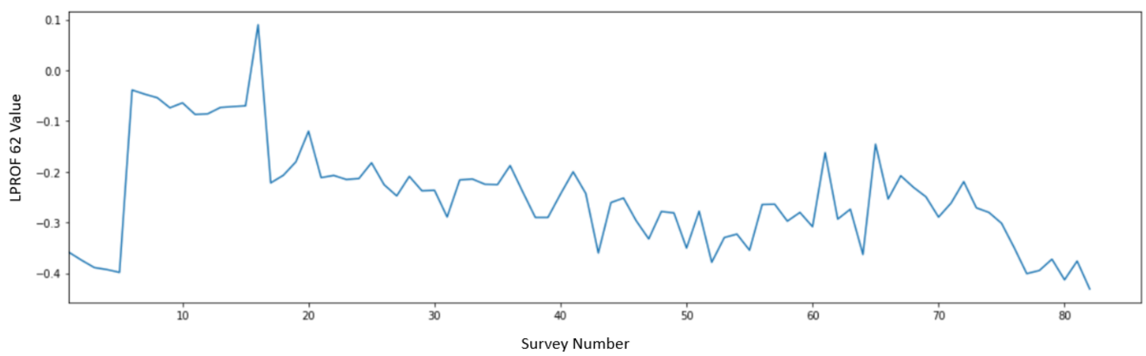


Figure 13. LPROF62 Time Series at 400 Feet from 82 Successive Surveys

Figure 14 shows the results of the ADF test described in Section 3.4 applied to the data in Figure 13. In typical statistical analysis, a null hypothesis can be rejected if the *p* value is less than 5 percent. Figure 14 shows a *p* value greater than 30 percent, so the null hypothesis that the data is nonstationary cannot be rejected. This means there is 95 percent confidence that the time series in Figure 13 has a unit root and is not stationary. Figure 15 shows the first-differenced LPROF62 data from Figure 13.

```
adf_test(foot400)

Augmented Dickey-Fuller Test:
ADF test statistic       -1.967102
p-value                   0.301181
```

Figure 14. ADF Test Results

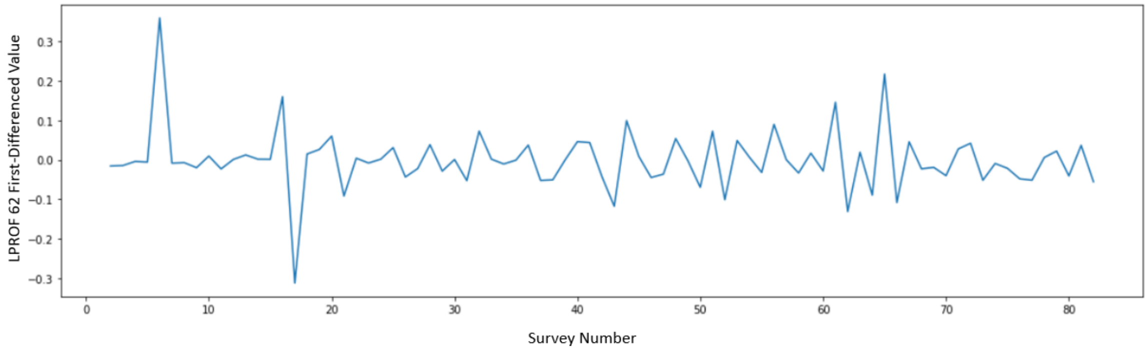


Figure 15. First-Differenced LPROF62 Time Series at 400 Feet

Figure 16 shows the results from an ADF test applied to the first-differenced data in Figure 15. Figure 16 shows a *p* value less than 5 percent. Thus, the null hypothesis can be rejected and there is 95 percent confidence that the first-differenced data is stationary. This means the *d* parameter of the ARIMA forecasting model should be set to 1.

```
1  adf_test(first_differenced_foot400)

Augmented Dickey-Fuller Test:
ADF test statistic     -7.259886e+00
p-value                 1.692388e-10
```

Figure 16. ADF Test Results of First-Differenced Data

Figure 17 shows the autocorrelation function for the first-differenced data in Figure 15. Autocorrelation values are plotted for the first 10 lags in the data. Figure 17 shows Lag-1 autocorrelation is outside the 95 percent confidence intervals represented by the blue-shaded area. Autocorrelations for the higher lags are within the confidence intervals. Thus, the *q* parameter of the ARIMA forecasting model is determined to be 1.

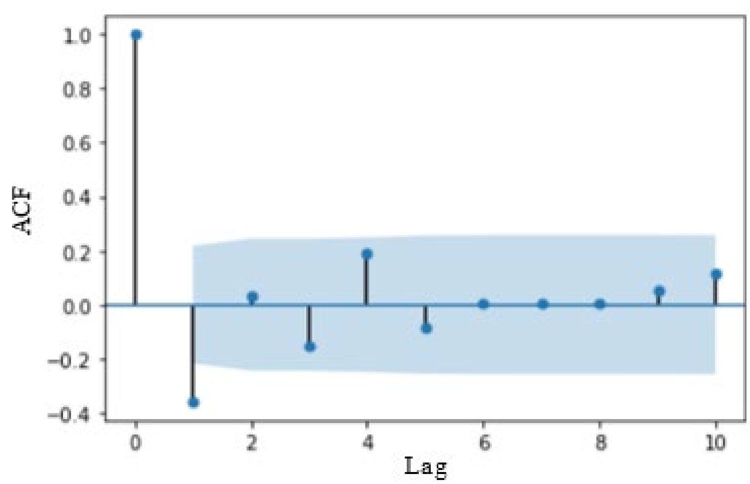


Figure 17. Autocorrelation Function of First-Differenced Data at 400 feet

Figure 18 shows the partial autocorrelation function for the first-differenced data in Figure 15. Partial autocorrelation values are plotted for the first 10 lags in the data. The PACF plot in Figure

18 has a significant spike at Lag-1, which means that all the higher-order partial autocorrelations are effectively explained by the Lag-1 result. Thus, the order of the *p* parameter in the ARIMA forecasting model is 1.

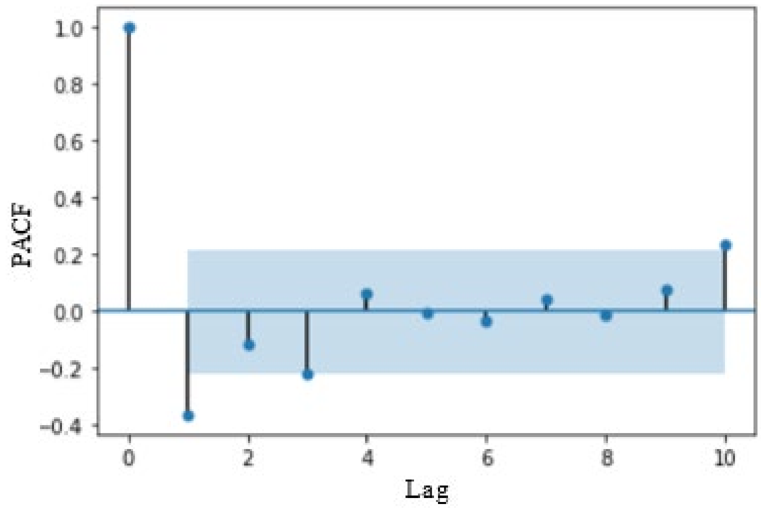


Figure 18. Partial Autocorrelation Function of First-Differenced Data at 400 feet

In summary, the spike at the first lag for both autocorrelation and partial autocorrelation functions and the fact that the series is stationary after first differencing indicate that the appropriate forecasting model for the time series at 400 feet is ARIMA(1,1,1) (Nau, 2022).

The Auto ARIMA function in the Python programming language automates the process described above for determining ARIMA parameters *p*, *d*, and *q*. After finding the best *p*, *d*, and *q* parameters for each of the 1,501 different time series, a global model was chosen. For this demonstration, ARIMA(1,1,0) was chosen because it was the most common solution found by the Auto ARIMA function, being found 630 out of 1,501 times (i.e., 42 percent of the time). For the remaining 871 time series, the Auto ARIMA function found different options, none of which exceeded the 42 percent majority.

Researchers then analyzed all 1,501 time series with the global ARIMA(1,1,0) model to forecast the final data points. The combination of the final data points from each of the individual time series represents the forecast for the final survey. Figure 19 overlays the surveyed and forecasted final results. Figure 19 shows the ARIMA model forecasted track geometry approximately one week after the 82nd survey accurately. The RMSE in this case is 0.00015 inch.

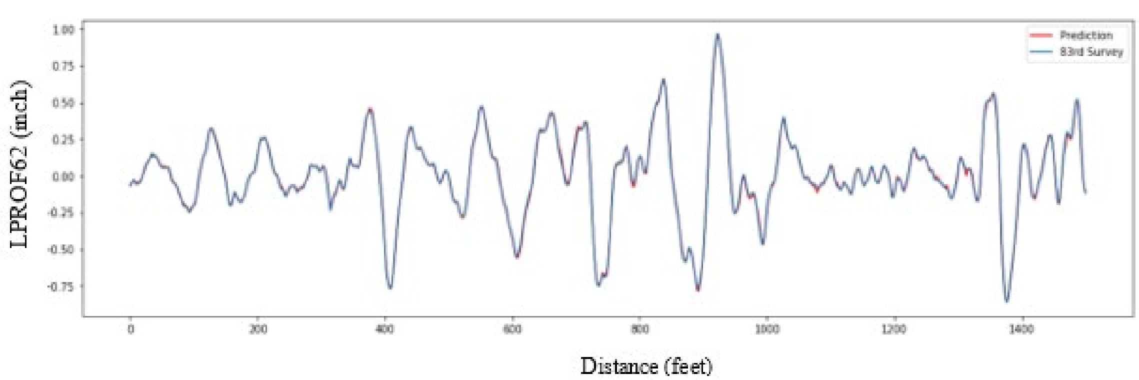


Figure 19. Point-Processed Results with ARIMA

As noted, the results shown in Figure 19 used a global ARIMA forecasting model (the same model for all time series). The challenges associated with global models are explained in Section 4.4. An improved methodology to overcome these challenges by automatic parameter assignment is defined in Section 4.4.

4.2 Effect of Maintenance Activities

Maintenance activities, such as tie replacement, undercutting, and surfacing, can influence the resulting track geometry signature and, in turn, affect the forecasts. Figure 20 illustrates this by overlaying measured LPROF62 data before and after spot surfacing.

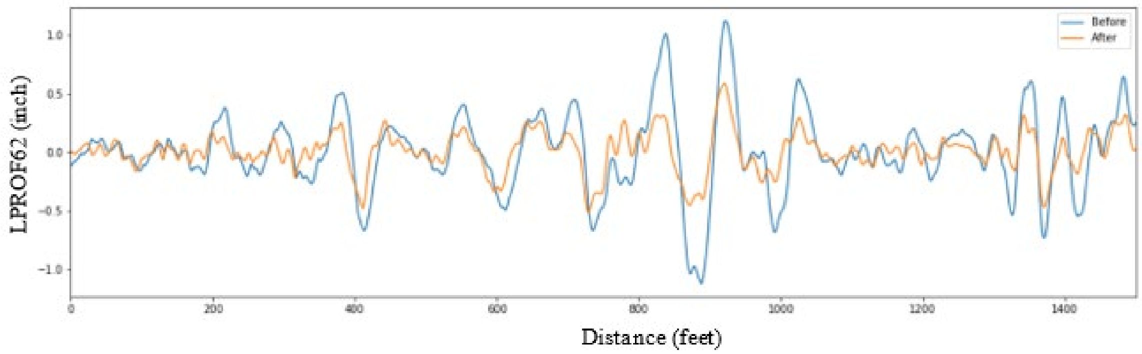


Figure 20. Maintenance Activity and Signature Change

Maintenance activities were introduced as an additional variable for both SARIMAX and Prophet. The SARIMAX forecasting model includes an exogenous variable to capture the effects of maintenance, and is introduced when predictions are affected by external factors (Arunraj & Ahrens, 2016). Without this exogenous variable, the model would treat the time series as normal historical data with an unexplained variation. Figure 21 illustrates how this maintenance variable is introduced into point-processed data.

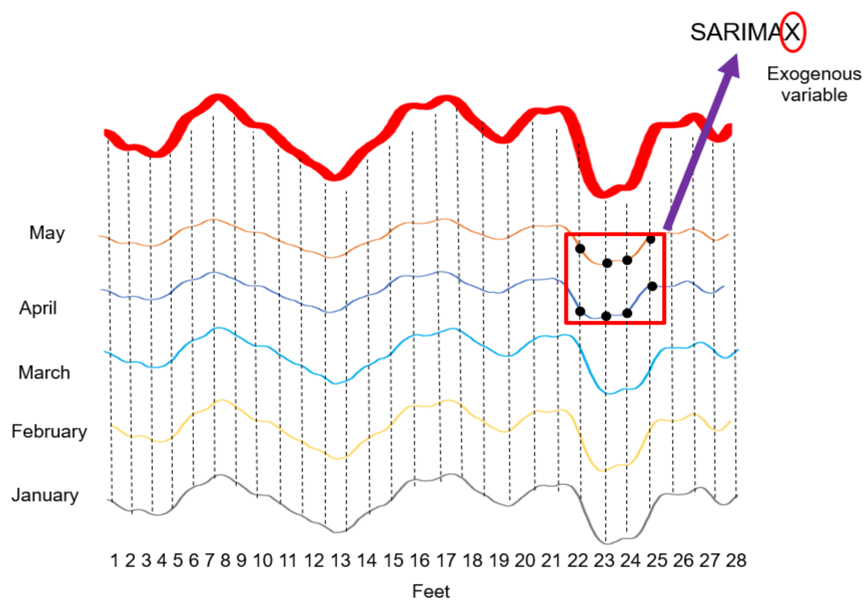


Figure 21. Implementation of Maintenance Variable with Point Processing

In the example of Figure 21, the railroad performed maintenance from foot 22 to 25 between the March and April surveys. The track geometry from foot 22 to 25 improved in April and May compared to the previous months. If the maintenance is ignored, the forecast for June will incorrectly show a worsening of track geometry at this location. Thus, the forecasting model must account for the change. Figure 22 shows the effect of track maintenance on segment-processed data.

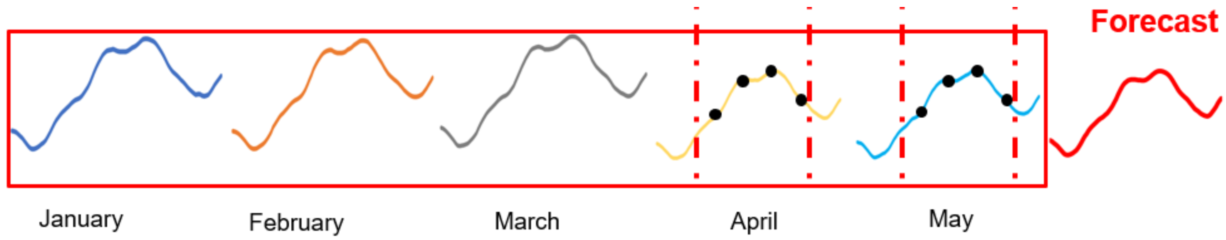


Figure 22. Implementation of Maintenance Variable with Segment Processing

The track maintenance that occurred between March and April caused a change in the shape of the track geometry. The forecasted shape on the far right of Figure 22 uses the shapes from all the surveys from January through May, but it needs to account for the change in shape due to maintenance. The Prophet forecasting model uses an additional regressor for segments affected by external events (e.g., maintenance) to handle this issue.

4.3 Performance of Forecasting Models with Maintenance Variable

Researchers produced forecasts accounting for the effect of maintenance using LPROF62 track geometry data from 65 weekly surveys. The length of the surveys was 1,500 feet. Between the 59th and 60th survey, a maintenance activity occurred from 800 to 1,000 feet, which resulted in a significant change to the track geometry. Researchers generated the forecasts in one-week intervals. Figure 23 shows results from the ARIMA forecasting model when this data was point-processed without including the maintenance variable. The heat map shows RMSE for 20-foot segments. Figure 24 shows results from the SARIMAX forecasting model when the data was point-processed and the maintenance variable was included.

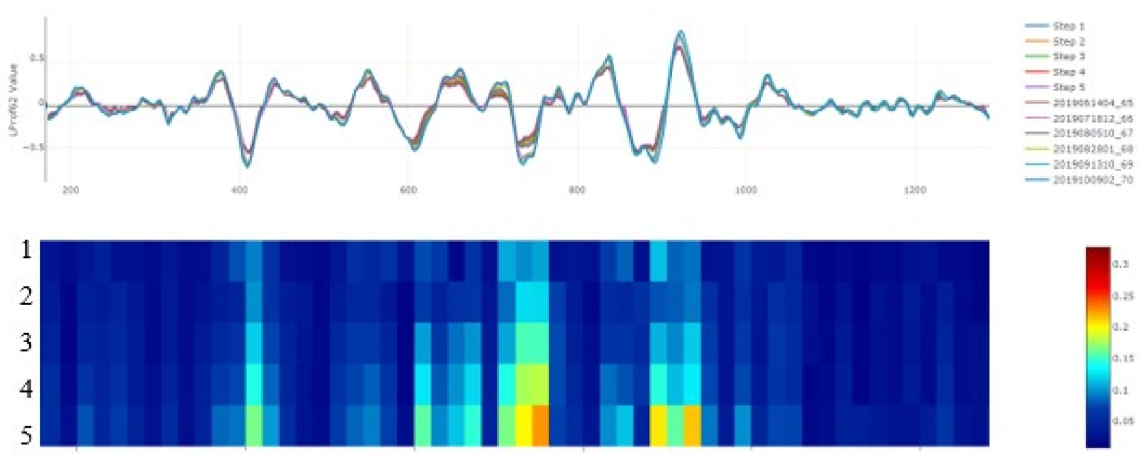


Figure 23. Point-Processed Results without Maintenance Variable

Figure 24. Point-Processed Results with Maintenance Variable

A comparison of Figure 23 and Figure 24 shows that there is a slight improvement between 800 and 1,000 feet when the maintenance variable is included. More specifically, RMSE and MAE calculated in the 20-foot segments in this zone improved by 36.3 and 34.2 percent, respectively, for the point forecast five weeks into the future. Figure 25 shows results from the Prophet forecasting model when the data was segment-processed without including the maintenance variable.

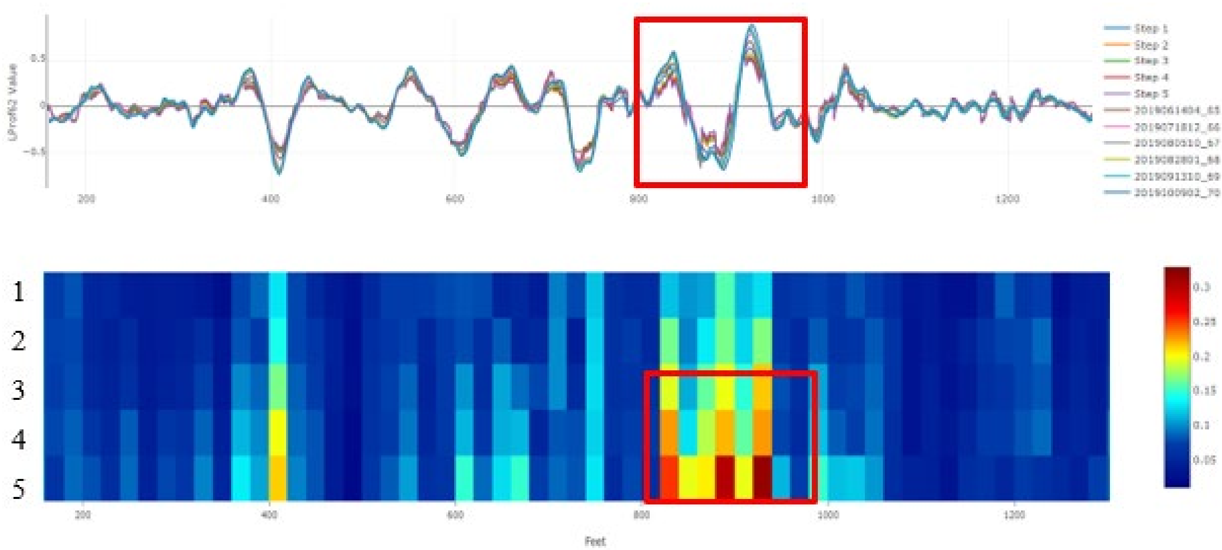


Figure 25. Segment-Processed Results without Maintenance Variable

Figure 26 shows results from the Prophet forecasting model when the data was segment-processed and the maintenance variable was included. A similar comparison between Figure 25 and Figure 26 shows that there is a significant improvement between 800 and 1,000 feet when the maintenance variable is included. In this case, the RMSE and MAE calculated in the 20-foot segments in this zone improved by 60.2 and 59.3 percent, respectively, for segment processing five weeks into the future.

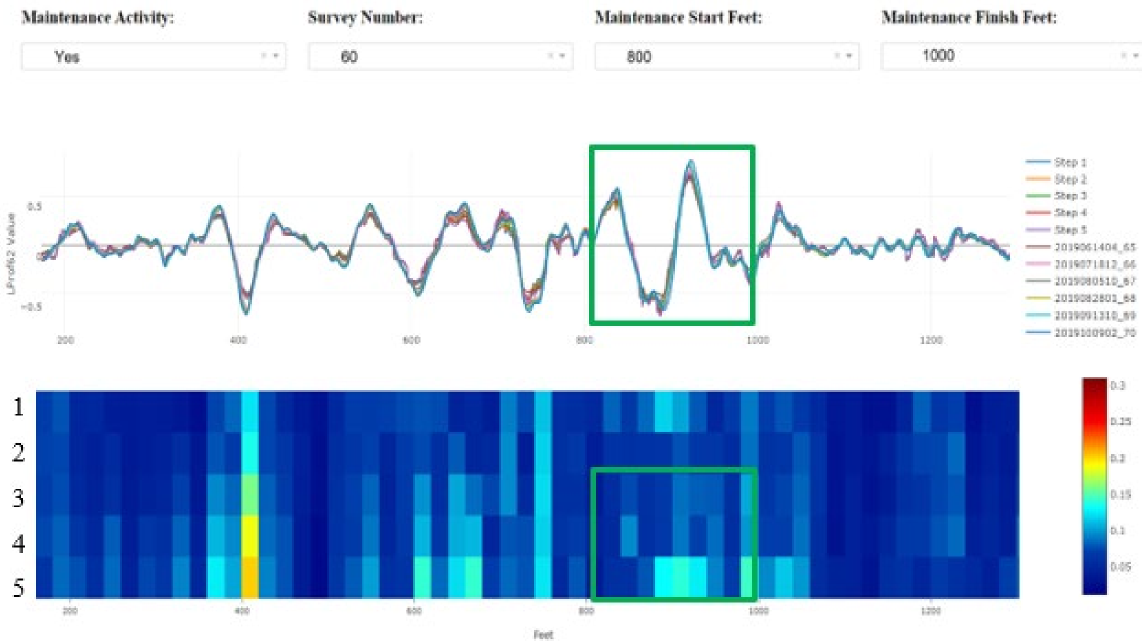


Figure 26. Segment-Processed Results with Maintenance Variable

Adding a maintenance variable to both point- and segment-processed data improved forecasting accuracy. Furthermore, the improvement with the Prophet forecasting model applied to segment-processed data was significant.

4.4 Point Processing Automatic Parameter Assignment

Parameter assignment in time series forecasting models plays a significant role in prediction accuracy. The model parameter values change according to characteristics of the time series. For example, the autoregressive component of the ARIMA(p,d,q) model, p, can be 0 in some cases, and 1 or 2 in other cases. The same is true for d (the integrated part) and q (the moving average part). There is usually no model that fits all time series.

All the results described to this point in this report with point-processed data used a global model with one set of parameters for all time series in the dataset. This means that if p was assigned as 1, d as 1, and q as 0, all time series at each individual foot location would have been trained with ARIMA(1,1,0), even though some of them did not fit those parameters. This can lead to a phenomenon where some time series do not converge or produce very large values. Figure 27 shows two examples highlighted in red where this has occurred.

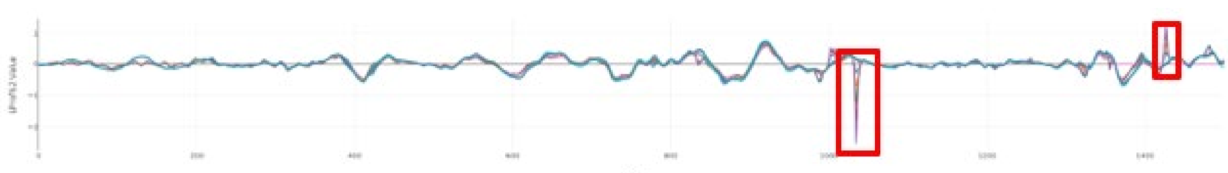


Figure 27. Forecast Using Global ARIMA Parameters

To improve forecasting reliability and accuracy, researchers developed a process that automatically selects model parameters for each time series. This process searches for the ARIMA parameters that best explain the historical data for each separate time series. Figure 28 shows the results of using global ARIMA(1,1,0) parameters on track geometry alignment data measured for 60 surveys at one week intervals. The length of the dataset is 1,300 feet. The heat map shows RMSE for five forecasts.

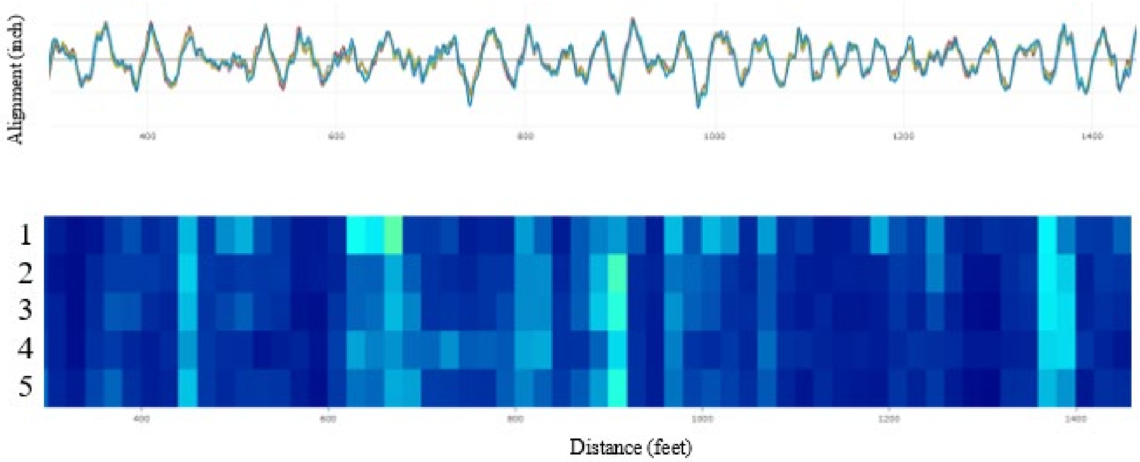


Figure 28. Forecast Using Global ARIMA(1,1,0) Parameters

Figure 29 shows the results of using automatic parameter assignment on the same data used for Figure 28. A comparison of Figure 28 and Figure 29 shows that the automatic parameter assignment improved the forecasts. Specifically, the overall RMSE for the fifth survey with global ARIMA(1,1,0) parameters (Figure 28) is 0.037 inch, which reduces to 0.028 inch with automatic parameter assignment (Figure 29) – a 24.3 percent improvement.

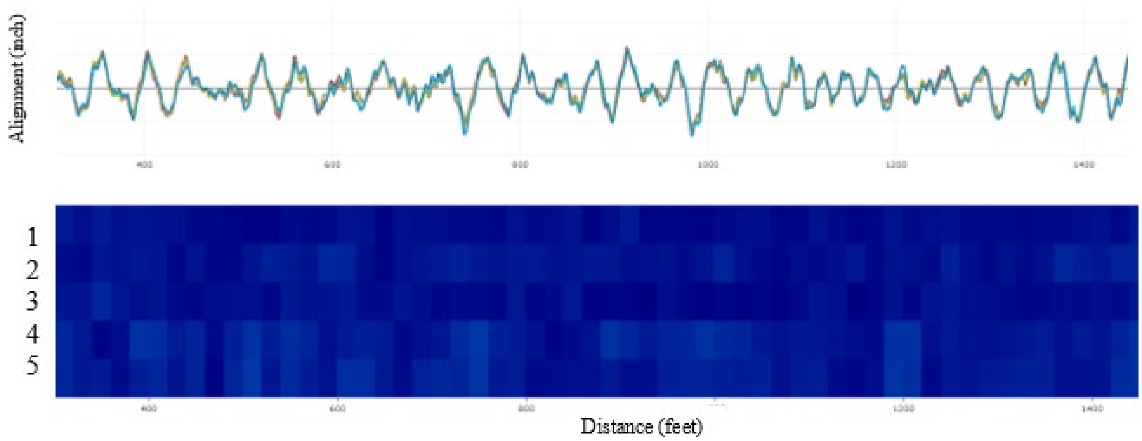


Figure 29. Forecast using Automatic Parameter Assignment

While automatic parameter assignment uses the optimal ARIMA parameters for each time series, improving the overall accuracy, it is more computationally expensive than the global approach. In practice, analysts will need to determine their priorities to decide which approach will suit their needs.

5. Conclusions

This research successfully demonstrated the efficacy of advanced forecasting methods in predicting foot-by-foot track geometry.

The research team was able to establish the viability of both point- and segment-processing techniques, each of which feature their own unique advantages and considerations. Further, the team determined that the use of global ARIMA models in point processing, while computationally efficient, was initially limited in accuracy. By introducing automatic parameter assignment, the accuracy significantly improved. The research team also found that segment processing, using the Prophet model, was very robust against misalignment issues and further benefited from overlapping segments to resolve discontinuities in the forecasted data.

By including maintenance activities as 1) an external variable in the SARIMAX model for point processing, and 2) an additional regressor in the Prophet model for segment processing, the research team successfully demonstrated that their inclusion was essential for accurate forecasting in both instances. This allowed the models to effectively account for significant, unexpected changes in track geometry due to maintenance activities, enhancing the overall predictability of the models.

Lastly, the research team found that, while there are inherent challenges in forecasting track geometry (e.g., computational intensity), the methodologies developed in this research are very promising, capable of improving both railway safety and maintenance planning. Successful implementation of these methods into revenue service could potentially advance predictive maintenance strategies significantly, offering a proactive approach to operational safety and reliability.

6. Next Steps

Additional efforts are needed to deploy the methods detailed in this report into revenue service. Researchers identified a list of next steps to advance the maturity of the forecasting methodologies for industry acceptance and operational deployment.

- Develop a method for evaluating the accuracy of forecasts. This might include a measure of how well the forecast predicts exceedances of track geometry safety limits.
- Develop logic for automatic choice of point versus segment processing. This logic should consider the accuracy of data alignment.
- Develop scalable model complexity based on how far into the future the model forecasts. The complexity of the models and their parameter assignments might increase for forecasting more than three steps ahead.
- Individual track geometry channel signatures are affected by asset types and special trackwork. Asset type could be introduced into the models as an external variable.
- Track geometry data contains anomalies and erroneous signatures that negatively affect forecasts. Advanced filtering methods based on robust statistics could be incorporated into the forecasting methodology.
- The time series-based forecasting methods are computationally intensive. Both segment-processed and automatic parameter assignment on point-processed data will require optimization to improve computational performance.

In addition, forecasting methodologies should be validated using track geometry data from different measurement systems and territories.

7. References

Alsharif, M. H., & Younes, M. K. (2019, February). *Time Series ARIMA Model for Prediction of Daily and Monthly Average Global Solar Radiation: The Case Study of Seoul, South Korea.* Retrieved from Research Gate: https://www.researchgate.net/publication/331168828_Time_Series_ARIMA_Model_for_Prediction_of_Daily_and_Monthly_Average_Global_Solar_Radiation_The_Case_Study_of_Seoul_South_Korea

Andrews, B. H., Dean, M. D., Swain, R., & Cole, C. (2013). *Building ARIMA and ARIMAX Models for Predicting Long-Term Disability Benefit Application Rates in the Public/Private Sectors.* Schaumburg, IL: Society of Actuaries. Retrieved from https://www.soa.org/globalassets/assets/files/research/projects/research-2013-arima-arimax-ben-appl-rates.pdf

Arunraj, N. S., & Ahrens, D. (2016, April). *Application of SARIMAX Model to Forecast Daily Sales in Food Retail Industry.* Retrieved from Research Gate: https://www.researchgate.net/publication/296477650_Application_of_SARIMAX_Model_to_Forecast_Daily_Sales_in_Food_Retail_Industry

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020, April). *Application of the ARIMA model on the COVID-2019 epidemic dataset.* Retrieved from Science Direct: https://www.sciencedirect.com/science/article/pii/S2352340920302341

Bruzek, R., Stark, T., Sussmann, T., Tunna, J., & Thompson, H. (2022). *Fouled Ballast Waiver Operations and Results.* Washington, DC: Federal Railroad Administration. Retrieved from https://railroads.dot.gov/elibrary/fouled-ballast-waiver-operations-and-result

Facebook. (2022, 1 20). *Forecasting at Scale*. Retrieved from Facebook Prophet: https://facebook.github.io/prophet/

Fattah, J., Ezzine, F., Aman, Z., El Moussami, H., & Lachhab, A. (2018). *Forecasting of demand using ARIMA model.* International Journal of Engineering Business Management. Retrieved from https://journals.sagepub.com/doi/full/10.1177/1847979018808673

Jakasa, T., & Androcec, I. (2011, May). *Electricity price forecasting — ARIMA model approach.* Retrieved from Research Gate: https://www.researchgate.net/publication/252019774_Electricity_price_forecasting_-_ARIMA_model_approach

Kufel, T., & Kufel, P. (2020, October 14). *Comments on "ARIMA modelling & forecasting of COVID-19 in top five affected countries"(by Sahai et al.).* Retrieved from US National Library of Medicine: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7556277/

Nau, R. (2022, 1 20). *Introduction to ARIMA: nonseasonal models*. Retrieved from Duke University: https://people.duke.edu/~rnau/411arim.htm

Pacella, M., & Anglani, A. (2008, January). *A Real-Time Condition Monitoring System by using Seasonal ARIMA Model.* Retrieved from researchgate.net: https://www.researchgate.net/publication/255569872_A_Real-Time_Condition_Monitoring_System_by_using_Seasonal_ARIMA_Model_and_Control_Charting

Palese, J., Zarembski, A. M., & Attoh-Okine, N. (2019, October 31). *A Stochastic Rail Wear Forecast Using Mixture Density Networks and the Laplace Distribution.* Retrieved from Transportation Research Board: https://trid.trb.org/view/1766932

Taylor, S. J., & Benjamin, L. (2017, September 27). *https://peerj.com/preprints/3190v2/.* Retrieved from PeerJ: https://peerj.com/preprints/3190v2/

Appendix A.
Additional Results

Figure A.1 shows an example dashboard for crosslevel track geometry data using point processing and the ARIMA forecasting model. Although there is maintenance (indicated by an apparent signal change) from 0 to 170 feet, the forecasts were made without including this external variable. RMSE and MAE for the fifth prediction in this example are 0.0169 inch and 0.0137 inch, respectively.

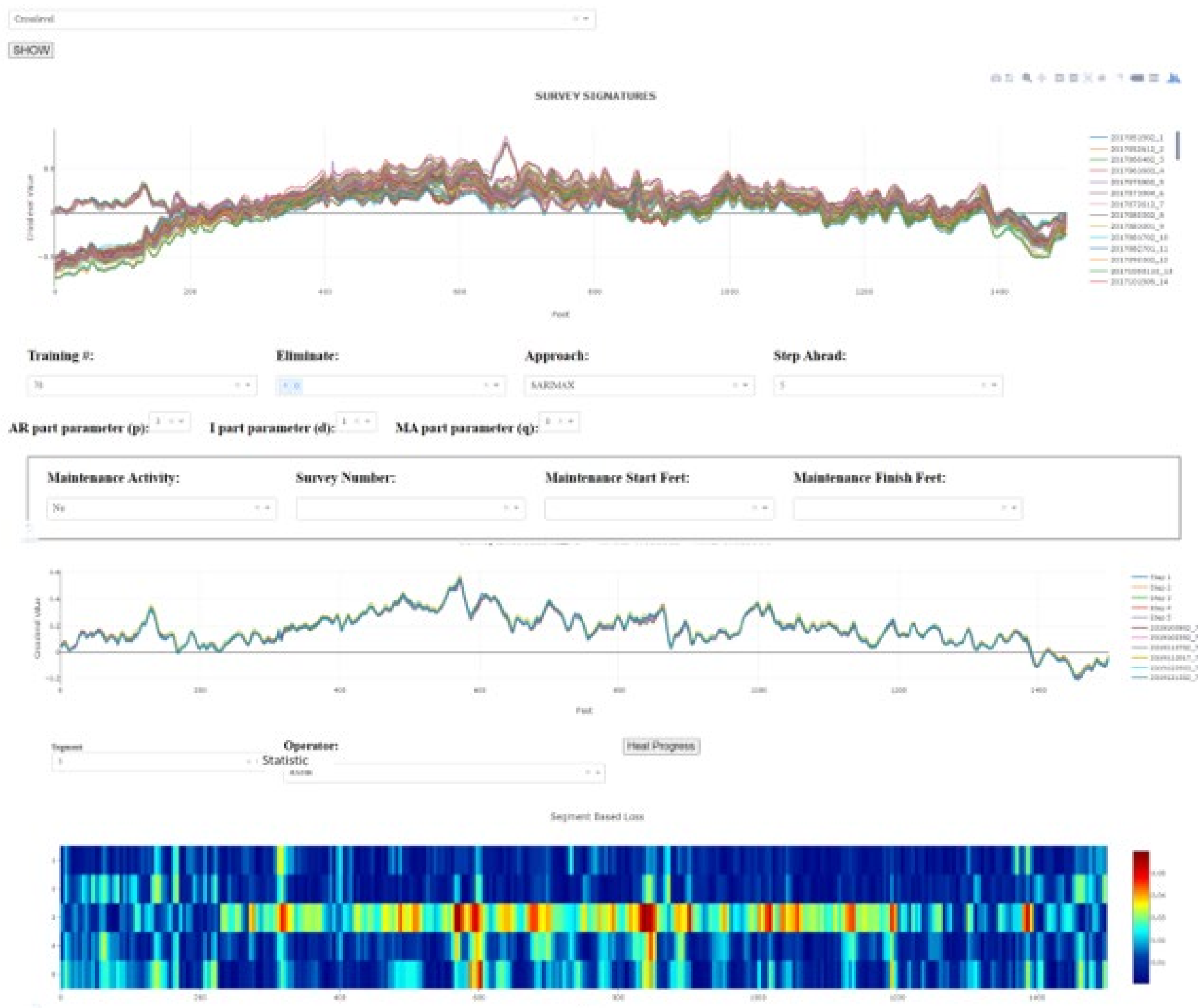


Figure A.1. SARIMAX Results for Crosslevel without Adjustment for Maintenance

Figure A.1 shows high RMSE values for the third forecast. The dashboard was used to investigate the cause, which was found to be white noise in the data.

Figure A.2 shows the same data as Figure A.1 but with maintenance between 0 and 170 feet included as an external variable.

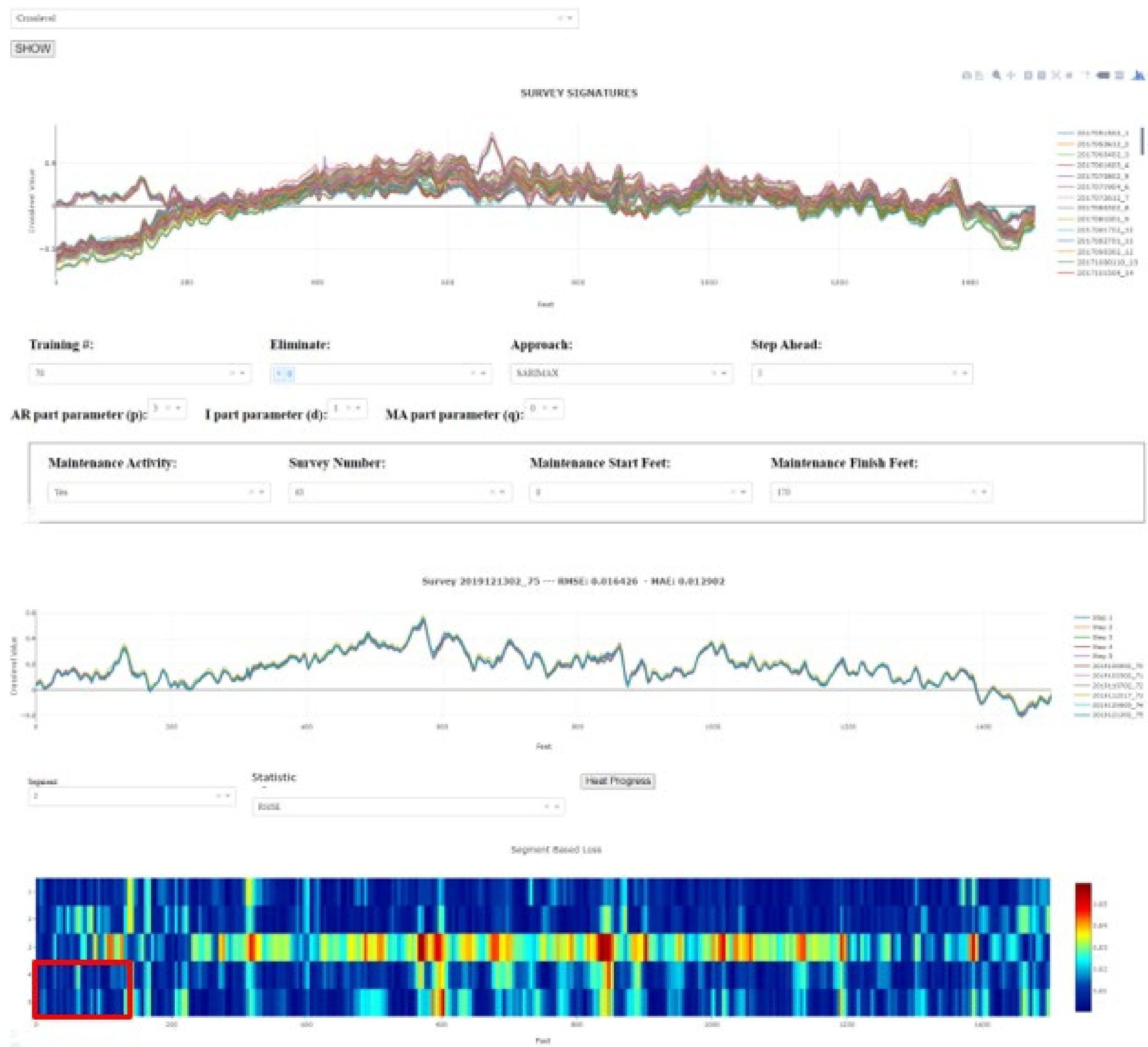


Figure A.2. SARIMAX Results for Crosslevel with Adjustment for Maintenance

Comparing Figure A.1 with Figure A.2 shows that adding external variables to account for the change in the signals due to maintenance between 0 and 170 feet improved the accuracy of the predictions. Overall RMSE decreased from 0.0169 to 0.0164, and MAE decreased from 0.0137 to 0.0129. The red box in the bottom left of Figure A.2 shows the forecasts improved for the surveys after maintenance in the area where maintenance was performed.

Figure A.3 shows results using the SARIMAX forecasting model with segment-processed LPROF62 track geometry data.

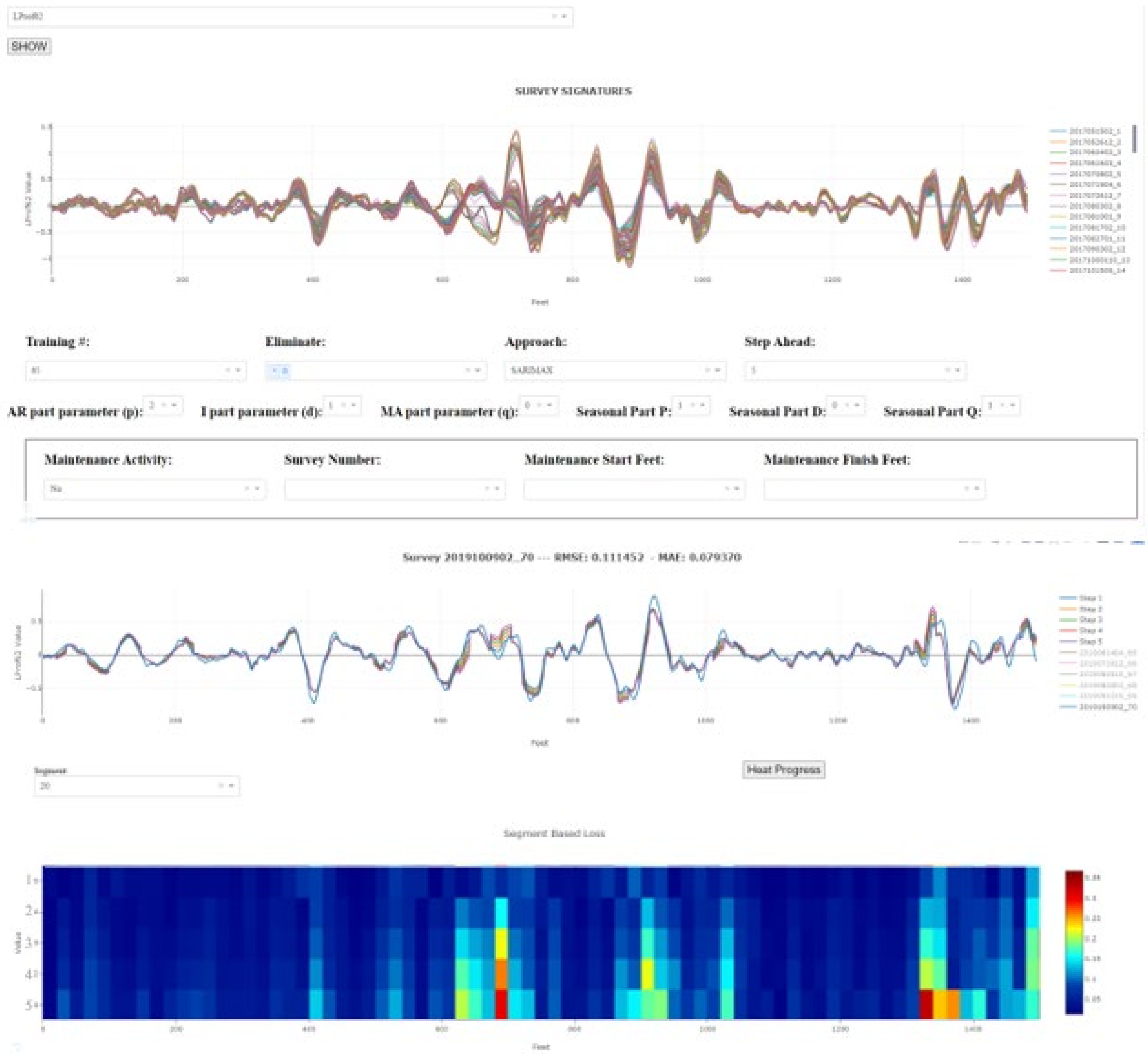


Figure A.3. SARIMAX Results on Segment-Processed Data

Although Prophet was used for segment-processed track geometry data, the dashboard allowed other forecasting models to be tested. Figure A.3 shows limited success with the SARIMAX model applied to segment-processed data with a seasonal component. Long Short Term Memory Recurrent Neural Network (LSTM-RRN) was also tested. Prophet outperformed SARIMAX and LSTM-RRN for segment-processed track geometry data.

Figure A.4 shows an example of applying the Prophet forecasting model to track warp data over a 62-foot base length (WARP62).

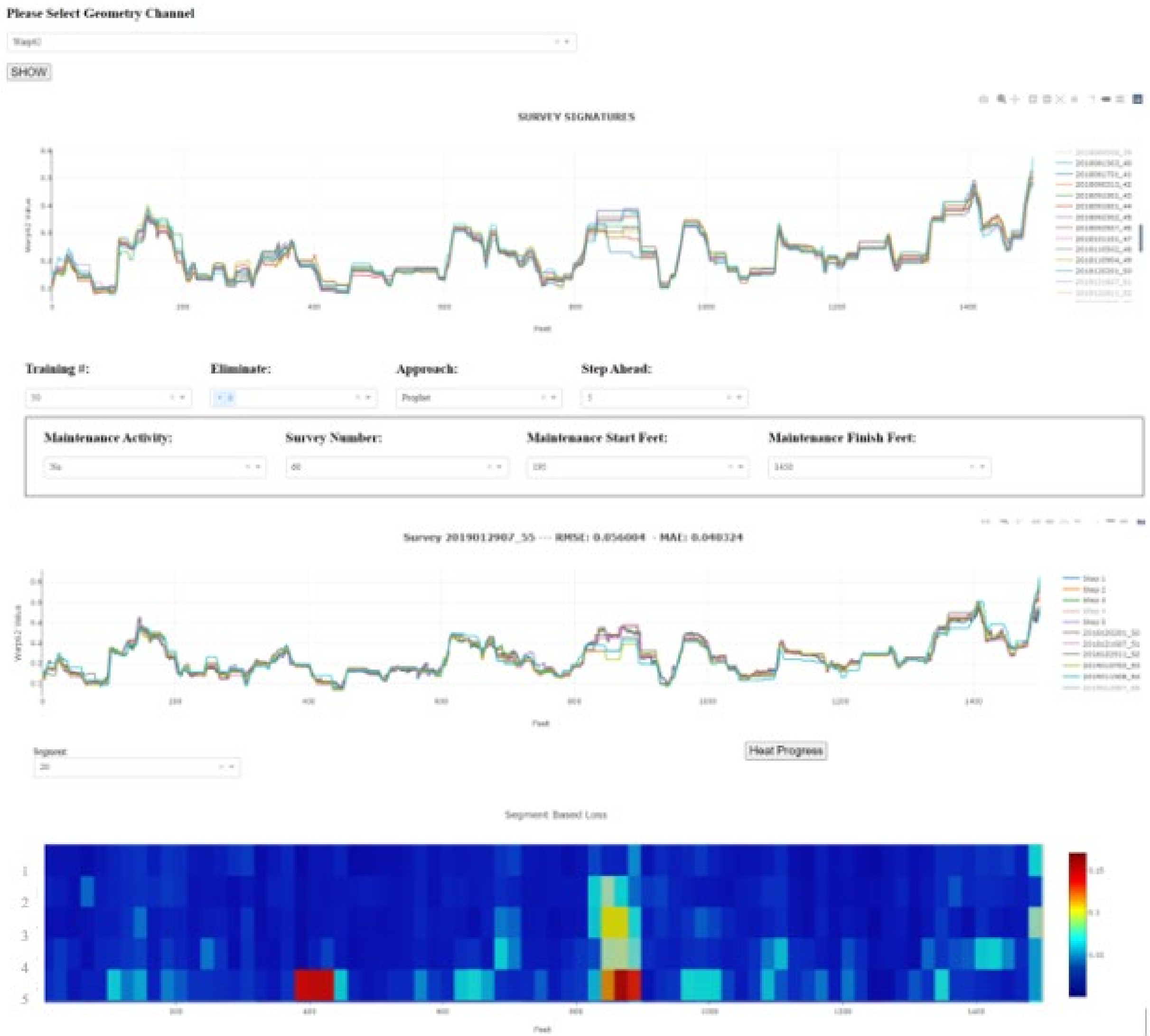


Figure A.4. Prophet Results for WARP62

Figure A.5 shows results using the SARIMAX forecasting model on the same WARP62 data used in Figure A.4.

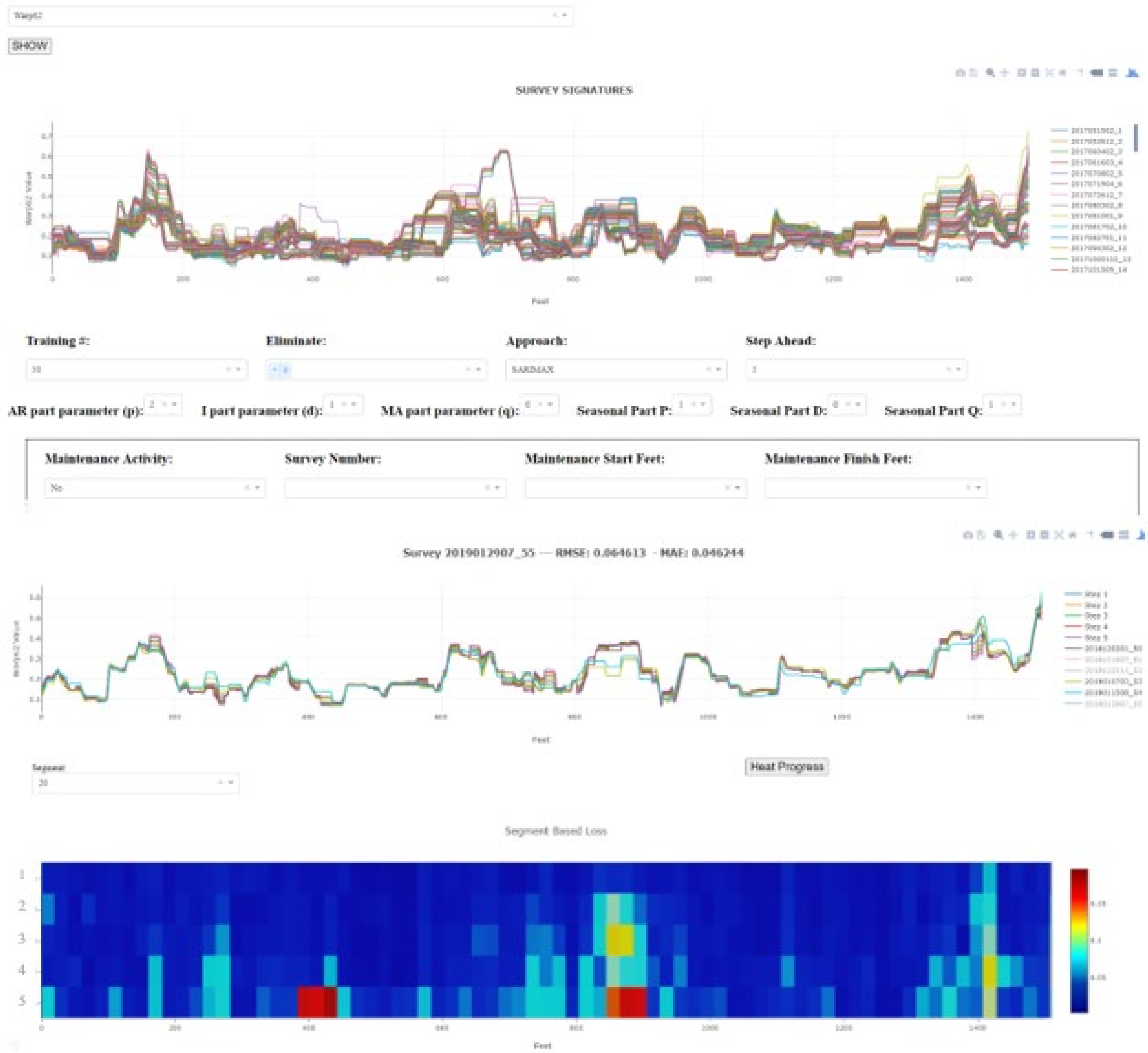


Figure A.5. SARIMAX Results for WARP62

Comparing Figure A.4 with Figure A.5 shows Prophet produced marginally better forecasts than SARIMAX.

Abbreviations and Acronyms

ACRONYM	DEFINITION
ACF	Auto Correlation Function
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
ARIMA	Autoregressive Integrated Moving Average
ATGMS	Autonomous Track Geometry Measurement System
FMCG	Fast Moving Consumer Goods
FRA	Federal Railroad Administration
LPROF	Left Profile
MAE	Mean Absolute Error
PACF	Partial Auto Correlation Function
RMSE	Root Mean Squared Error
SARIMAX	Seasonal Autoregressive Integrated Moving Average Exogenous
TGMS	Track Geometry Measurement System