U.S. Department of Transportation

Federal Railroad Administration

RESEARCH RESULTS

FRA OFFICE OF RESEARCH, DEVELOPMENT & TECHNOLOGY

RR 26-01 | **July** 2026

TRACK GEOMETRY PREDICTIONS USING POINT AND SEGMENT PROCESSING

SUMMARY

From 2019 to 2020, the Federal Railroad Administration (FRA) contracted with ENSCO, Inc. to investigate advanced methods of forecasting track geometry degradation using foot-by-foot data.

This report presents the results of the literature review and the initial investigation into the most appropriate methods for forecasting foot-by-foot track geometry data. Point processing with the SARIMAX model and segment processing with Meta's Prophet model both produced accurate forecasts for these data. Adding exogenous variables to both models to account for unexpected track maintenance between measurements significantly improved the predictions.

Using these models, along with frequent Autonomous Track Geometry Measurement System (ATGMS) measurements, can accurately predict the future behavior of track geometry and help plan preventive maintenance more effectively. Case studies applying these methods will be completed as the next step in this research.

BACKGROUND

Predicting foot-by-foot track geometry signatures and, in turn, defects can play a key role in proactively addressing safety issues before they become problematic. Previous FRA research focused on trending individual peak-value deviations from track geometry measurements (Bruzek, Stark, Sussman, Tunna, & Thompson, 2022). To improve upon this capability, research was needed to evaluate more advanced forecasting methods that (1) utilized continuous foot-by-foot geometry measurements taken directly from inspection vehicles and (2) were suitable for autonomous processing. Successful implementation of accurate foot-by-foot track geometry forecasting will allow for the expanded use of ATGMS and its associated data products.

OBJECTIVES

The objectives of this phase of the research were to (1) conduct a literature review of relevant time-series forecasting models used in other industries and (2) identify suitable models to use with track geometry and assess their performance on a revenue service data. Furthermore, the models had to account for the seasonality and maintenance effects inherent in track geometry data.

METHODS

The research efforts used two major time series forecasting models: Autoregressive Integrated Moving Average (ARIMA) and Meta's Prophet.

ARIMA models are widely used by both academia and researchers to predict various 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 p, d, and q are non-negative integers. The autoregressive parameter p indicates that the variable is regressed with its own prior or "lagged" values. The integrated parameter d is the difference between the observed values and lagged values (e.g., the value immediately prior to the current value). This parameter accounts for an overall trend in the data. The moving average parameter q represents a moving average process based on lagged error terms (Jakasa & Androcec, 2011).

SARIMAX is an extension of the ARIMA model that includes lagged values and differences to fit seasonal patterns. An exogenous variable is also added to the model to handle the effects of external factors, such as track maintenance. The exogenous variable is assigned to individual foot samples affected by these factors. Adding exogenous variables can explain the sudden and significant changes caused by external effects (Andrews, Dean, Swain, & Cole, 2013).

Prophet is open-source software developed by Meta's core data science team. It is a library for forecasting time series based on additive models where non-linear trends are fit with both time and seasonality effects. It supports time series that have strong seasonal effects and several seasons of historical data. The Prophet forecasting model also uses an additional regressor for segments affected by maintenance. According to the developers, Prophet is robust to missing data, trend shifts, and outliers (Meta, 2021).

This research used two methods for processing track geometry data as a time series: point processing and segment processing. Figure 1 explains the point processing method. It shows five successive track geometry measurements with distance in feet on the x-axis. Every foot (the vertical dashed lines) represents a time series of historical data. Separate forecasts are made for each foot and joined to form the track geometry forecast depicted by the red line.

Figure 2 illustrates the segment processing method. Track geometry data is divided into segments. In this example, the segments are 24 feet long. The five monthly surveys form a time series of shapes for each segment. These shapes can be used as inputs to seasonal time-series models or pattern recognition algorithms to give a forecast shape. The forecasts for each segment can then be joined to form an overall track geometry forecast.

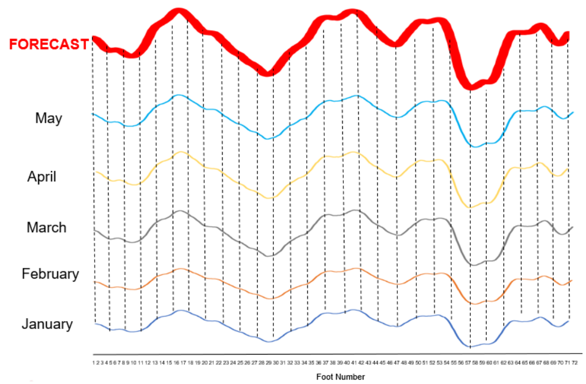

Figure 1. Point Processing Method

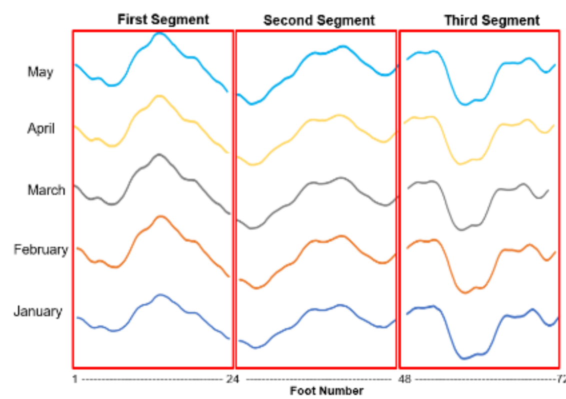

Figure 2. Segment Processing Method

RESULTS

The literature review found that time series forecasting models have been used in numerous industries and research areas. Palese et al. cited examples in the railroad industry where time series forecasting was used for operations and goods demand (Palese, Zarembski, & Attoh-Okine, 2019). In the same report they detailed their own application of time series forecasting for rail wear growth. Other research has evaluated models for railway car brake failures (Pacella & Anglani, 2008).

<u>Point Processing Results</u>

The ARIMA forecasting model was used on 62-foot chord, vertical profile track geometry data (LPROF62) for a 1,500-foot segment of track, where 83 surveys had been collected between May 2017 and January 2020 in approximately one-week intervals. The first 82 surveys were used as a training set. All 1,501 time series were then analyzed with the global ARIMA (1,1,0) model to forecast the 83rd data points. The combination of 83rd data points from each of the individual time series represents the forecast for the 83rd survey. Figure 3 overlays the measured and forecast 83rd results.

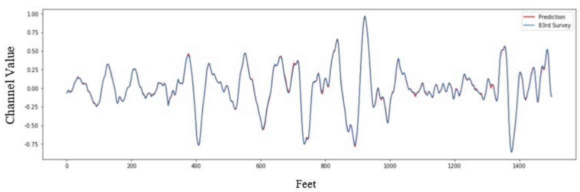

Figure 3. ARIMA Point Processed Results

Figure 3 shows the ARIMA model forecast track geometry approximately one week after the 82nd survey with a remarkable accuracy. Root mean square error (RMSE), or the square root of the average of squared errors, was selected as an evaluation measure for the models. In this example the RMSE is 0.00015 inch. This is an excellent result considering that in current practice the reproducibility limits for standard track geometry channels are 1/16 inch.

Segment Processing Results

The Prophet model was trained on 65 track geometry surveys taken at weekly intervals. Figure 4 shows good agreement between five forecasts at weekly intervals and future surveys made at those intervals. The RMSE for the forecast that was five steps in the future was 0.0121 inch. This is, again, a very good result when compared to current reproducibility limits.

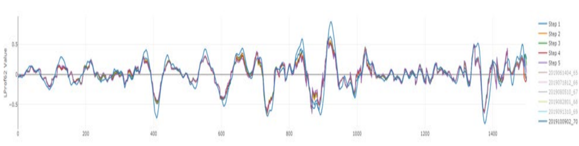

Figure 4. Prophet Segment Processed Results

Exogenous Variable Results

The effects of track maintenance were introduced as an additional variable for both point- and segment-processed track geometry data. Track maintenance activities, such as tie replacement, undercutting, and surfacing can influence the resulting geometry signature. In turn, this can affect the resultant forecasts; without these additional variables, the models would treat the time series that contain maintenance activities as normal historical data with unexplained variation. Implementation of the additional variables in the models resulted in up to 60 percent reductions in the RMSE for forecasts with known maintenance events.

CONCLUSIONS

Track geometry data can either be processed as individual time series for each surveyed foot or as a time series of shapes in short segments of track. Applying ARIMA to point-processed track geometry data and Meta's Prophet to segment-processed data can provide users with accurate forecasts of track geometry behavior.

Track maintenance can affect forecasting of both point- and segment-processed data. Using a SARIMAX model with external variables on point-processed data and adding a regressor to the Prophet model on segment-processed data can significantly improves forecasts when track maintenance has occurred. This helps improve predictability.

FUTURE ACTION

The next phase of this research is to develop a user interface to facilitate application of the forecasting methods. In addition, to demonstrate the effectiveness of the forecasting methods on foot-by-foot track geometry data, different case studies will be demonstrated.

REFERENCES

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.

Bruzek, R., Stark, T., Sussman, T., Tunna, J. T., & Thompson, H. (2022, January). Fouled Ballast Waiver Operations and Result (Report No. DOT/FRA/ORD-22/01). Federal Railroad Administration:

Meta. (2021). Forecasting at Scale. Retrieved January 20, 2021, from Meta Prophet.

Jakasa, T., & Androcec, I. (2011, May). Electricity price forecasting - ARIMA model approach. *2011 8th International Conference on the European Energy Market (EEM)*, Zagreb, Croatia, 2011, pp. 222-225.

Pacella, M., & Anglani, A. (2008, January). 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. U.S. Department of Transportation, University Transportation Centers Program.

ACKNOWLEDGMENTS

Serkan Sandikcioglu led the project team at ENSCO, which included Radim Bruzek and Daniel Einbinder.

CONTACT

Jay Baillargeon
Program Manager
Federal Railroad Administration
55500 DOT Road
Pueblo, CO 81001
(719) 584-7155
jay.baillargeon@dot.gov

Hugh Thompson
Program Manager
Federal Railroad Administration
1200 New Jersey Avenue, SE
Washington, DC 20590
(202) 573-3252
hugh.thompson@dot.gov

KEYWORDS

Prediction, forecasting, track geometry, time series, ARIMA, Prophet

CONTRACT NUMBER

693JJ6-20-F000090