Methods for Detecting and Predicting Localized Rapid Deterioration of Track Irregularity Based on Data Measured with High Frequency

Hirofumi TANAKA
Track Geometry & Maintenance Laboratory, Track Technology Division

Shuhei YAMAMOTO
Track Geometry & Maintenance Laboratory, Track Technology Division

Takashi OSHIMA
Track Geometry & Maintenance Laboratory, Track Technology Division (Former)

Masashi MIWA
Track Geometry & Maintenance Laboratory, Track Technology Division

This study describes the development of methods to detect and predict localized rapid deterioration of track irregularity based on data measured with high frequency. First of all, a highly accurate position correction technique was developed. This technique seeks out phases where the correlation coefficient between waveforms of the different measured data reaches a maximum, and corrects these phase gaps. The automatic extraction of localized rapidly deteriorating track irregularity is made possible from the difference in measured data which has already had its position errors corrected. Secondly, a method for predicting track irregularity was developed. This technique predicts the track irregularity stochastically through updates using new measurement data, applying the Bayesian theorem. Finally, these techniques were applied to field data, confirming their effectiveness.

Keywords: commercial train monitoring, high frequency measurement, position correction, rapidly detection, prediction of track irregularity deterioration

1. Introduction

In recent years, compact track geometry measurement devices, such as inertial mid-chord offset track inspection devices [1], have been developed for practical use in Japan and in overseas countries. By installing this type of compact device on commercial trains, it is possible to carry out track measurements more frequently than when using conventional track geometry vehicles [2]. Given the large volumes of data being collected, the issues facing researches now is how to develop a method to efficiently process this track geometry data to make full use of it for track maintenance.

The common slipping and sliding of wheels on track-geometry measuring vehicles means that even for sections that have been measured multiple times there can be fluctuations in the recorded track geometry data, because of inaccurate positioning on the track during the measurement process. Data storage ground coils are therefore installed along tracks to allow track geometry measuring vehicles to detect the exact position of each storage point to correct measurement positioning errors. The data storage coils are often installed at intervals of about 1 km to obtain an approximate correction of position at these points. However, this method is not applied to correct position across different track geometry data sets, which means that when data sets are collected on different dates or at different times small positional errors are introduced. In addition, deterioration of track irregularity cannot be determined simply from the difference between waveforms in two different measurements, therefore the level of track deterioration has until now usually been determined using the standard deviation of the waveforms (i.e. data curves) obtained from a certain section [3]. Although this method can be used to determine the track deterioration along a specific section, i.e. it may be useful for determining locations for tamping maintenance, as the sensitivity to localized rapidly deteriorating track irregularity is relatively low when this indicator is used.

In view of the above, this study proposes a method that aims to enable more effective maintenance by determining the points with localized rapidly deteriorating track irregularity at an early stage by using very frequently measured track geometry data. This was achieved first, by aiming to automatically detect localized rapidly deteriorating track irregularity. A method was then developed to obtain high-precision correction of positional errors, by maximum phase cross-correlation between the waveforms of two separate measurements. Another method was developed to automatically determine localized rapidly deteriorating track irregularity was occurring based on the difference between sets of track geometry data measured on different dates. Finally, a method was developed to predict future deterioration of track irregularity at these identified points, by successively updating track irregularity deterioration using Bayes estimation each time track measurements were made. These methods were then applied to actual data to verify their effectiveness.

2. Method to obtain high-precision position correction by cross-correlation

2.1 Position correction algorithm based on cross-correlation

In the position correction method developed to pick out localized rapidly deteriorating track irregularity using dif-
ferences between two separate measurements, rather than simply aligning the phases of two waveforms alone, correlations were also made between a number of other data points, as well as trying to match entire sampling intervals between the two sets of measurements. Figure 1 shows an outline of the concept underlying the proposed position correction method. This method consists of two steps: phase alignment and resampling.

First, two data sets are specified. One is the reference measurement data (hereinafter referred to as “reference data”) to be used as the base reference, and the other the data to be corrected (hereinafter referred to as “correction target data”). Then, as shown in Fig. 1 (a), while shifting the phase of the correction target data in small increments, (1) was used to calculate the cross-correlation between the two data curves to obtain the phase shift that gives the maximum cross-correlation between the two waveforms. The phase of the correction target data is corrected to the shift in phase.

\[
\hat{r}(\tau) = \frac{1}{\sigma(x)\sigma(y)} \frac{1}{N} \sum_{n=1}^{N} [x(n)y(n+\tau)]
\]

\(\tau\) : Phase shift
\(\sigma(x)\) : Standard deviation of reference data for section under examination
\(\sigma(y)\) : Standard deviation of the correction target data in the section under examination
\(N\) : Number of reference data points in the section
\(n\) : Data number
\(x(n)\) : Reference data
\(y(n)\) : Correction target data

Next, as shown in Fig. 1 (b), in order to have exactly the same number of data points in both data sets for the phase-aligned section, the correction target data is resampled at each of the sampling points corresponding to the sampling points in the reference data.

In this study, this position correction method is referred to as the “cross-correlation method,” which makes use of cross-correlation coefficients. Theoretically, this method can correct track position errors to within the sampling interval, or ±0.25 meters. It should be noted that the conventional approximate position correction method to use data storage coils can be used prior to applying the cross-correlation method to reduce the search range of cross-correlation for phase alignment, for more efficient calculation.

2.2 Generation of waveforms based on position correction

According to the algorithm described above, the measurement data waveform after position correction can be generated as shown in Fig. 2. First, using the sequence of procedures described above, position correction is made for a certain section (hereinafter referred to as a “lot”). The position data at the end point of this lot is inherited as the position data at the starting point of the next lot, for further position correction for the next lot according to the same procedure. By repeating this procedure, detailed position correction is made to all the lots.

If a position fails to be corrected due to track maintenance activities resulting in significant changes in the waveform, the uncorrected lot is corrected using the position data of the next successfully adjusted lot. In order to determine whether the position correction was successful or not, two additional parameters were introduced to this method.

One is the lower limit value of the cross-correlation coefficient calculated in a lot. When is below the lower limit, it is considered that position correction of the lot failed. The other parameter is the lower limit value of the percentage of successfully position-corrected lots among those subjected to position correction using the cross-correlation method. When is below the lower limit, the position correction through cross-correlation method is deemed to have been unsuccessful for the measurement data as a whole.

It should be noted that the lot length used to calculate the cross-correlation coefficient depends on various factors including the frequency characteristics of the waveform used for position correction. In this, in the case of typical track geometry data, it has been confirmed that the success rate of position correction is high for the lot length of 10 to 100 meters [4].

![Fig. 2 Outline of position correction for a measurement data set](image)

3. Method to determine rapidly deteriorating track irregularity

3.1 Method to determine rapidly deteriorating track irregularity based on data difference

For any two geometry data sets with different mea-
There are two types of significant change in track irregularity that can be determined using the difference of data sets, first, significant changes resulting from actual deterioration of track irregularity, and second, changes due to improved track conditions after track maintenance. In particular, if the measurement is made using a 10 m mid-chord offset method, given that a pseudo waveform will be generated on both sides of the deteriorated point, it is not possible to rely only on the increase or decrease in the amplitude to arrive at a conclusion or decision. While it may be considered desirable to omit maintained sections by cross-referencing with track work records, it would be impractical to do so because of the high frequency measurements on service trains. Therefore, it is necessary to have a method to automatically distinguish actual track irregularity changes from differences found because of maintenance, without using work records.

Figure 3 is a flow chart showing how the presence or absence of rapidly deteriorating track irregularity is determined, in two steps. First, even if some deterioration is observed during a certain period of time, it will not immediately affect the safety of train operation if changes are very small. Therefore, rapid deterioration is considered not to be occurring if the level of deterioration in the first step is equal or below a certain threshold (threshold A) as shown in (3).

\[
\Delta y > \alpha : \text{Points with suspected rapidly deteriorating track irregularity}
\]

\[
\Delta y \leq \alpha : \text{Points where there is no rapidly deteriorating track irregularity}
\]

Next, if threshold A is exceeded, the standard deviation of the measurements from the section around the point is calculated. If the standard deviation value exceeds a threshold (threshold B), as shown in (4), the point is determined to have actually experienced localized rapidly deteriorating track irregularity. Otherwise, it is considered that the point is simply either in a section of track that has been maintained, or that it is not rapid deterioration. Note that it is important to determine threshold B to properly reflect actual conditions taking into account measurement errors that may occur in actual measurements, and other relevant factors. The lot length used for the calculation of the standard deviation should also be properly selected because longer lot lengths will decrease the sensitivity to locally observed deterioration in track irregularity. Also note that the proper lot length for calculation is discussed and verified in Section 3.2.2.

\[
\sigma_{\text{after}} - \sigma_{\text{before}} > \beta : \text{Points found to have rapidly deteriorating track irregularity}
\]

\[
\sigma_{\text{after}} - \sigma_{\text{before}} \leq \beta : \text{Points found not to have rapidly deteriorating track irregularity}
\]

\[
\Delta y = y_{\text{after}}(x) - y_{\text{before}}(x)
\]

\[
\alpha \Delta \leq
\]

\[
\beta
\]

3.2 Example application to high frequency measurements

3.2.1 Example of detection of localized rapidly deteriorating track irregularity

Figure 4 is an example of high frequency track irregularity measurements, giving two sets of data measured 7 days apart, one for level irregularity and the other showing the difference between them (hereinafter referred to as “difference in longitudinal level”). There are two sets of results for difference in longitudinal level: one where the conventional error correction method is applied, and the other where the cross-correlation method is applied, in addition to the conventional method.

From Fig. 4 (a), it is not possible to see any significant difference between the two sets of measurement waveforms. However, Figure 4 (b), shows that the level irregularity difference measured by applying the conventional position correction method is relatively large over the entire section, and reveals that the deterioration of track irregularity was miscalculated due to small positional errors as previously mentioned. Figure 4 (c) meanwhile illustrates application of the cross-correlation method to correct positions, and the level irregularity difference is generally low except at one point adjacent to a bridge section without a ballast floor where the difference is about 3 mm and confirms that localized rapidly deteriorating track irregularity has been accurately detected. This confirms therefore that by applying the cross-correlation method to correct small positional errors between measurement data sets, it is now possible to calculate and determine deterioration in track irregularity from waveform levels.
4. Method for predicting rapidly deteriorating track irregularity

4.1 Outline of method for predicting rapidly deteriorating track irregularity

4.1.1 Application of Bayes’ theorem to the prediction model

A prediction method was studied using statistical processing of historical data, to find a means to forecast rapidly deteriorating track irregularity using historical track geometry data collected on a very frequent basis. The Bayes’ theorem was applied to predict deterioration in track irregularity: updating the historical data with newly acquired measurement data improves the tracking of more recent trends, which is called Bayes estimation.

As shown in (5), the specific feature of Bayes’ theorem is in the structure where the probability of occurrence distribution of an event (prior distribution) is assumed beforehand and the prior distribution is updated to obtain a new probability distribution (posterior distribution) when new information is acquired later in relation to the prior distribution. In other words, Bayes’ theorem is a kind of machine learning to produce posterior information by adding new information.

\[
P(B | A) = \frac{P(A | B) P(B)}{P(A)} \tag{5}
\]

\(P(B | A)\) : Probability of occurrence of event B after occurrence of event A (posterior distribution)
\(P(A | B)\) : Probability of occurrence of event A after occurrence of event B
\(P(A)\) : Probability of occurrence of event A
$P(B)$: Probability of occurrence of event A (prior distribution)

Here, deterioration in track irregularity is considered as a probability distribution function; by applying Bayes’ method to predict track irregularity when new data is acquired, the posterior distribution that reflects the latest variation in the track irregularity is obtained as shown in Figure 6. Note that this method can be applied not only to localized rapidly deteriorating track irregularity but also to any point where track irregularity has deteriorated.

Figure 7 is a flow chart showing how to predict deterioration in track irregularity. Three processing steps are applied to the historical track irregularity data for a specific point: 1) processing of historical data to produce a prior distribution, 2) Updating with newly acquired data to produce a posterior distribution, and 3) further processing to forecast the date on which an approximate target value in irregularity will be reached.

![Diagram outlining track irregularity deterioration prediction using Bayes’ theorem](image)

**Fig. 6** Diagram outlining track irregularity deterioration prediction using Bayes’ theorem

**Fig. 7** Flow chart to forecast track irregularity deterioration

### 4.1.2 Processing of historical data

The historical track geometry data can fluctuate due to accidental inclusion of abnormal values and/or due to the effect of any measurement errors. Calculating track irregularity deterioration using data that contains this type of fluctuation, can make accurate prediction difficult if the variation is significant. This method therefore applies exponential smoothing to the historical data as shown in Figure 8 to remove any significant variation before calculating trends in deterioration. After that, the smoothed historical data is used to calculate the amount of deterioration on each measurement date, the distribution of the amount of irregularity deterioration, or the prior distribution of the track irregularity deterioration. Note that normal distribution is assumed for the amount of deterioration and its average values.

![Exponential smoothing of track geometry historical data](image)

**Fig. 8** Exponential smoothing of track geometry historical data

### 4.1.3 Updating with newly acquired data

Based on the premise that trends in the newly acquired data will affect the average of the track irregularity as previously noted, the distribution of the average value is updated with a Bayes estimation. Equation (6) shows the average value and standard deviation of track irregularity deterioration after the update. Note that this method can work with either single or multiple newly acquired data sets.

$$
\mu_a = \frac{\sum a^2}{n} + \frac{\zeta_a \sigma_a^2}{\sigma_a^2 + \sigma_0^2}
$$

$\mu_a$: Average value of track irregularity deterioration after the update

$\mu_b$: Average value of track irregularity deterioration before the update

$\sigma_a$: Standard deviation after the update, $\sigma_b$: Standard deviation before the update

$\sigma_0$: Standard deviation of track irregularity deterioration obtained from the history data

$\zeta_a$: Average value of track irregularity deterioration based on newly acquired data

$n$: Number of track irregularity deterioration data sets based on newly acquired data

$$
\left(1 + \frac{\zeta_a \sigma_a^2}{\sigma_a^2 + \sigma_0^2}\right)
$$

### 4.1.4 Future forecasting process

The next step is to forecast track irregularities using the posterior distribution of track irregularity deterioration. The forecast value after days can be given by (7). Here, because the amount of deterioration after the update

$$
\left(1 + \frac{\zeta_a \sigma_a^2}{\sigma_a^2 + \sigma_0^2}\right)
$$
is a probability variable, is also a probability variable. Thus, a probability distribution can be calculated for. As a result, it is possible to determine probabilistically the time when the value is expected to exceed a target value specified beforehand. For example, if the average value of the amount of deterioration is used, the day on which the target value is expected to reach can be calculated with (8).

\[ x_n = x_0 + n \cdot \Delta x \]  \hspace{1cm} (7)

\[ t_m = \frac{m - x_0}{\mu_x} \]  \hspace{1cm} (8)

\( x_n \): Forecast of track irregularity after \( n \) days
\( x_0 \): Track irregularity measured on the last measurement date
\( \Delta x \): Deterioration per day
\( t_m \): Number of days before reaching the target value, \( m \): Target value

Also, over a period of days, if the distribution of the amount of deterioration with average value and standard deviation appears repeatedly for days, assuming that the distribution of the amount of deterioration is a normal distribution and the distributions are independent between days, the average of the irregularity deterioration distribution is calculated to be and the standard deviation. Then, the distribution of the forecast values will widen with time. Therefore, it is also possible to estimate “date of reaching the target value” as the day when a percentile value of the distribution reaches the target value.

4.2 Example showing application of method to high frequency measurements and accuracy of prediction

4.2.1 Verification of prediction accuracy

Figure 9 shows the result of applying this prediction method to historical longitudinal level measurements made very frequently using service trains on a conventional line. The figure also shows other forecast results that are based on linear regression and exponential regression of the historical data. The track section shown in Fig. 9 experienced significant deterioration (level change) after maintenance was carried out reducing deterioration, which produces a pattern similar to ballast settlement. For point such as this one, where the historical data period shows a change in pattern, a linear regression model can result in unreasonably high predicted levels of deterioration compared reality, due to the effect of maintenance work. However, the exponential regression model and the model proposed in this paper can produce predictions that are closer to the actual trend.

To verify the prediction accuracy of this method, the prediction error was defined as the difference between predicted and measured values on the 15th day after starting the prediction. Table 1 shows the resulting prediction errors when prediction was made for 30 randomly selected track locations. The table shows that the prediction model produced the smallest prediction error, within 1 mm, demonstrating a high level of accuracy. As such, this prediction model is considered to be more suitable to practical applications than the exponential regression and linear regression models.

4.2.2 Example of calculations based on successive updates

Figure 10 shows example predictions by successive updates for longitudinal level where the track irregularity deterioration trends change frequently. As shown in the figure, even when the trends are changing, this method can make predictions that follow closely the changing patterns by updating the prediction each time new measurement data is acquired. This is considered to be the biggest advantage of applying Bayes estimation. The resulting predictions clearly follow the changing trends with each successive update.

It should be noted that it is known that prediction accuracy varies with the period of available historical data used to calculate the prior distribution, as well as with the number of newly acquired data sets used for successive updates. With this in mind, the prediction accuracy for various types of line sections will be examined further, using historical track condition data, and/or measurement items other than longitudinal level.

4.2.3 Example calculations to find expected date for reaching a target value

As stated in Section 4.1.4, this prediction model can
predict future track irregularities in the form of a probability distribution. Figure 11 shows example results of probabilistically predicted track irregularity (for longitudinal level) and the number of days to reach an assumed target irregularity value of 10 mm together with the expected probabilities. In this case, the number of days to reach the target irregularity value was calculated to be 23 ± 2 days with 90% probability. By predicting the day on which a target irregularity is expected to be reached using a probabilistic method, it is possible to locate areas where target irregularity limits are likely to be exceeded within a short period of time, allowing for preventive maintenance planning at an earlier stage. Consequently, this can be applied to prioritize maintenance work.

$$\text{Fig. 11 Example calculation of days to reach the target}$$

5. Conclusions

The findings of this study can be summarized as follows:

1. A new position correction method, or cross-correlation method, was developed to calculate deterioration in track irregularity through waveform levels, as a way to overcome the difficulty in calculating this through conventional position correction methods. The resulting theoretical accuracy of position correction was kept within the sampling interval of the measurement data, or within ±0.25 m. This method makes it possible to automatically determine localized rapidly deteriorating track irregularity from differences in track irregularity between two data sets.

2. A method was developed to predict deterioration of track irregularity by applying Bayes estimation to the historical track irregularity data that has been already corrected for position errors using the cross-correlation method. The observed prediction error using this method in this study was within ±1 mm over 15 days for a measurement frequency of once per day. Also, as this method predicts the deterioration of track irregularity probabilistically, the probability of exceeding a target irregularity value on a specific date can be calculated, allowing maintenance to be prioritized.

References

[1] Tsubokawa, Y., Yazawa, E., Ogiso, K. and Nanmoku, T., “Development of the Car Body Mounted Track Measuring Device with the Inertial Mid-chord Offset Method,” *QR of RTRI*, Vol.53, No.4, pp.216-222, 2012.

[2] Kasai, R., Saito, Y., Komatsu, Y., Ogiso, K., Yahagi, H. and Konishi, S., “Development outline of track facility monitoring device and future prospects,” *JR East Technical Review*, No.55, pp.21-24, 2016 (in Japanese).

[3] Sano, K., Miwa, M., Yamaguchi, T., Yoshida, N., Yasaka, K. and Sakaguchi, K., “Developing an Efficient Method of Applying High-Frequency Measured Track Inspection Data to Diagnosis of Track Condition and Establishment of Optimal Track Maintenance Strategy,” *RTRI REPORT*, Vol.29, No.8, pp.47-52, 2015 (in Japanese).

[4] Tanaka, H., Yamamoto, S., Ohshima, T., Mori, T. and Saito, Y., “Development of detection method of track irregularity of quick growth using high frequency track measurement data that applies the cross correlation method,” *Journal of Railway Engineering*, Vol.21, pp.1-8, 2017 (in Japanese).

[5] Yamamoto, S., Miwa, M., Tanaka, H. and Kashima, T., “Building a track irregularity prediction model considering characteristics of high frequency measurement data,” *Journal of Railway Engineering*, Vol.21, pp.9-16, 2017 (in Japanese).

Authors

**Hirofumi TANAKA**
Senior Researcher, Track Geometry & Maintenance Laboratory, Track Technology Division
Research Areas: Track Geometry Management, Signal Processing of Track Geometry Data, On-board Monitoring

**Takashi OSHIMA**
Researcher, Track Geometry & Maintenance Laboratory, Track Technology Division
(Former)
Research Areas: Track Geometry Management

**Shuhei YAMAMOTO**
Researcher, Track Geometry & Maintenance Laboratory, Track Technology Division
(Former)
Research Areas: Track Geometry Management

**Masashi MIWA**
Senior Chief Researcher, Laboratory Head, Track Geometry & Maintenance Laboratory, Track Technology Division
Research Areas: Track Geometry Management, Maintenance Support System