# Stochastically based prediction models for track deterioration

B. Bergquist
*Quality Technology, Luleå University of Technology, Luleå, Sweden*

P. Söderholm
*Trafikverket (the Swedish Transport Administration), Luleå, Sweden &*
*Quality Technology, Luleå University of Technology, Luleå, Sweden*

ABSTRACT: The Swedish Iron ore line is subjected to sub-arctic weather conditions, high axle loads and a large yearly tonnage. These operational conditions challenge the usefulness of existing deterministic deterioration models for railway infrastructure that are validated in other contexts. However, more frequent and higher precision of condition measurements, combined with statistically based prediction methods, may offer a viable solution. Here, we study remaining useful life predictions of the geometry properties of the railway track based on recursively updated time series. We discuss how data-driven models are affected by measurement errors of track properties, such models' ability to detect seasonal effects, and how they are affected by irregular sampling. The prediction abilities and uncertainty measures for different modelling approaches are also compared. In this study, a linear regression model based on the degradation rates of the particular track segments showed the best prediction abilities, compared to competing models. More advanced models such as Cubic splines and Neural Nets had the worst prediction abilities.

## 1 INTRODUCTION

In Sweden, the maintenance of railway track geometry is primarily based on tamping. The tamping action is most often triggered by measurements performed by single measurement trains. The procedure has been to let single measurements trigger corrective maintenance if the geometry falls outside tolerable limits. However, there are also initiatives to base maintenance planning with a long time horizon on a practice using repeated measurements from the measurement trains (Vågbrink, 2019). The applied model is produced in-house by engineers at Trafikverket (the Swedish Transport Administration), and implemented in the Optram decision support system (Smith, 2016). However, the Optram system is only capable of using models of low complexity; e.g., it does not allow for iterations, and it uses a constant degradation rate for the whole track system, that is, individual differences in track stability is not accounted for (Nissen, 2019).

## 2 PURPOSE

This paper studies the prediction abilities of different analytic and data-driven methods for railway track geometry. The paper also discusses how data-driven models are affected by measurement errors of the track property, such models' ability to detect seasonal effects, and how they are affected by irregular sampling.

## 3 THEORY

Methods for prediction could be based either on physics, stochastics or on hybrids between the two. Models based on physics-of-failure are typically used to obtain deterministic point predictions and rely on that some events cause the system to react. These models of physics usually require knowledge of a certain external factor. If all variables are known, the outcome is also fully determined. These models also go under the name white-box models (Ljung, 2001) or first principles models, in that they are based on first principles such as Newtonian equations.

While physics-of-failure models can be powerful, their value for complex systems such as railway condition decay is often limited. Hybrid, or grey box, models (Ljung, 2001) may be based on some empirical experiments, for instance, a test of soil compaction, mechanical fatigue mechanisms as functions of stresses and cycles and so on. In this case, a hybrid model for track geometry degradation could be based on soil-mechanic models and measurements or calculations of the particular stresses that affect the railway. By calculating the particular load and stress that a track is subjected to, the deformation can be predicted. A substantial benefit from physics and hybrid models is that they can aid the understanding of the driving mechanisms, which can be used in the solutions to address the same principles.

Stochastic, or black-box, models (Ljung, 2001) acknowledge that there is variation and requires that some variables contain a measure of randomness. The purpose of such modelling is to determine how probable a particular outcome is, based on a set of observations that may include some element of uncertainty. Monte-Carlo simulations are examples of where uncertainties are used to calculate the probabilities of specific outcomes. A data-driven modelling approach falls into this category, and may often be suitable for modelling complex systems. By using observations from gauges of the inputs and outputs of the system, a data-driven modelling approach connects known outcomes with the variation seen in the inputs. By using the historical reactions of the system as a tool, the use for prognostics assumes that the system will behave similarly for future events. As such, the data-driven approach does not need to employ systems knowledge; the data-driven empirical model will have a predictive ability as long as the system behaves similarly to how it has behaved in the past. Downsides include that learning from such models may be limited, but also if essential variables that were constant during the model-building phase change in later phases.

## 4 METHOD

This section focuses on solutions for managing challenges of collected data and compares the usefulness of different prognostic modelling approaches.

### 4.1 *Data collection and cleaning*

The studied data is from one track section of the Swedish rail network, collected between the years 2007 and 2019 and describe track geometry properties and their location. One measurement train, the IMV 200, currently performs most of the Swedish track geometry measurements. This train is capable of measurements of up to 200 km/h. These measurements are complemented by measurements by smaller measurement trolleys with top speeds of 100 km/h. Most measurements until 2013 was performed by the now retired STRIX measurement car.

The cars use accelerometers and gyros for their geometric measurements and use those together with the recorded speeds to calculate, e.g. the height of the track by integration of the acceleration and considering the train speed. The vertical speed of the accelerometers is obtained by the first order integral of the acceleration; a second integration can obtain location if the constant speed of the train is acknowledged. The number of calculated measures has increased over the years, starting from 24 variables in 2007; each measured every 25 cm of the track length.

Of the measured properties, seven properties contain so called point-defect types, that is, defects with a limited extension in the track direction. These point defects are regulated with maintenance limits and were considered of particular interest in the present study. The regulated measures were two twist measurements (base 3 and 6 meters), the track gauge, along with the height and side deviations for both rails, measured within the shortwave (1-25 m) spectrum. In this paper, we will study the 6m twist variable, (also known as the cross-level). This property has potential safety-related consequences, since a too large twist may induce derailments. The regulations for maintenance requires maintenance in three levels. When the lowest level, UH1 (maintenance limit 1) is surpassed, maintenance planning must be done so that the next level, UH2 (maintenance limit 2), is not surpassed before the next planned maintenance period (TDOK 2013:0347). The next level, UH2, requires an immediate corrective maintenance action. Finally, there is a critical limit (KRIT), where traffic restrictions, such as reductions of maximum speed or stopped traffic are required (TDOK 2013:0347). For the 6-meter twist defects, the UH1, UH2 and KRIT limits are 11, 17 and 25 mm/m respectively (TDOK 2013:0347).

The measurement cars use dead reckoning and GPS signals, as well as manual inspections by operators passing certain objects, e.g. road passages, to link observations to a particular position. The positioning requirements of the measurement car measurements have been +/- 10 m along the track (SS-EN 13848-1:2003+A1:2008). The localisation accuracy demand has risen the accuracy demand to +/- 2 meters Easting and Northing (i.e. in Cartesian coordinates) (TDOK 2013:0347). However, the Swedish measurements have not lived up to that accuracy; in reality positioning errors of up to 50 meters have been found for older measurement cars before 2013, and the current accuracy seems to be around +/- 10m.

### 4.2 *Chosen track section*

The studied track section (119) is part of the Swedish Iron ore line, a heavy-haul line stretching between the harbours of Luleå in Sweden and Narvik in Norway with an accumulated yearly tonnage of about 34 million gross tonnes (Asplund et al., 2017). It is reasonable to assume that the condition of the Iron ore line is difficult to predict for two reasons: the particular climate and the high axle load. Models that allow for accurate predictions under such circumstances are likely useful also for railways lines operating under less severe environmental conditions with lower axle loadsaxle load, since models that work in these conditions could be assumed to be validated for among the worst possible conditions for predictive purposes.

It is well known that climatic conditions make railway maintenance more difficult in regions where climatic conditions such as ground frost change the volume and stiffness of the substructure and thus change the actual rail geometry (e.g. Larsson-Kråik, 1999; Tai et al., 2017). The Iron ore line is located both in sub-arctic and arctic conditions, with temperatures varying between over 30°C in the summer to below -40°C in the winter. The maximum allowed axle load is currently 32.5 metric tonnes. High axle loads are also a factor affecting sub-grade quality (Li & Selig, 1995), and thus geometric stability.

The studied track section 119 lies between the cities of Boden and Luleå. The track is a single track with side-tracks, has a total length of around 35 km. Besides the end stations, the track contains two stations in Gammelstad and Notviken. The track has a sub-arctic climate, with warm summers and cold winters. The highest allowed track speed for the track in good condition ranges between 120-140 km/h.

The high annual load, combined with a high axle load does make the track to qualify for the most frequent geometric measurements (inspection class 4, BVF 807.2) by measurement trains. The BVF 807.2 regulations introduced in 2005, require measurements to be performed at least six times per year for inspections class 4. In reality, measurements are not always as frequent, see Figure 1. Especially for the first four years where data are available, measurements are few. Figure 1 shows the number of measurements each year that that had produced measurement data for at least 75% of the whole track section. Note that a few years hold additional measurements or duplicate measurements. The Luleå station is an endpoint station, and the only connection from Luleå is through Boden using track section 119. Hence, sometimes the measurement train measures the same track both going back and forth, with a relatively short time in-between. Usually, the return trip only involves measurements of a few kilometres or side-tracks.

These duplicate measurements are useful for calculating the train measurement uncertainty (Bergquist and Söderholm, 2016), and it can be assumed that the average of both measurements are better than either one of them due to that the average will have lower measurement noise. However, unprecise measurements with a high degree of measurement noise close in time is not very helpful for understanding the slow degradation rates, since measurement noise will hide the much smaller degradation rate signal.

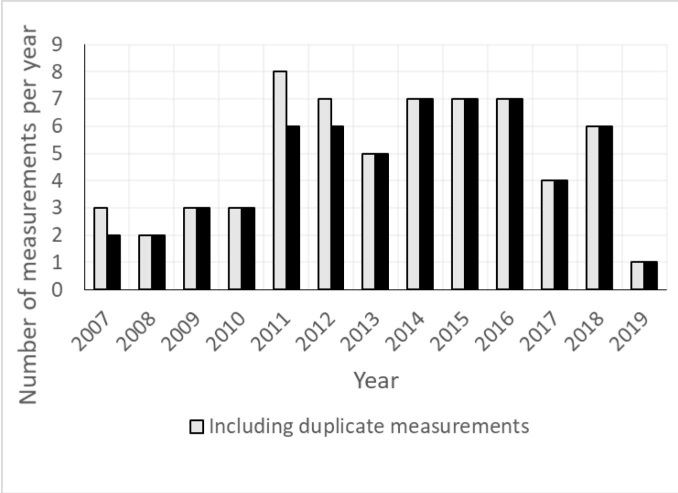


Figure 1. The number of measurements with measurement train on track section 119 per year, 2007-2018.

Many tools for time series analysis and forecasts require that data are sampled with even frequencies, and these data are not sampled in such a way. Figure 2 shows the time sequence of the used measurements.

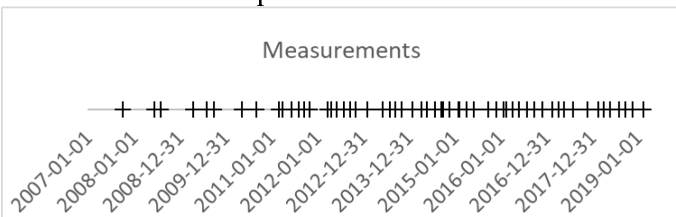


Figure 2. Measurements of more than 75% of track 119, placed on a time-scale, 2007-2018.

The data were regularised to overcome the misfit between the data with a different number of measurements each year and the method requirements of equal number and equally spaced observation. Each year was split into quarters, and measurements in that quarter were represented by the quarter date (March 31, June 30, September 30, and December 31). For example, observations from January 1 until, and including, March 31 a particular year were included in the March 31 pseudo-observation. If several measurements were obtained from a particular quarter, the pseudo-observation was represented by the average. Likewise, the measurement uncertainty was calculated for all measurement cars using the duplicate measurements using the methodology outlined in Bergquist & Söderholm (2016). The measurement uncertainty for quarters with multiple measurements was reduced according to how many measurements that constituted the base for the average calculation.

For point defects, such as twist or gauge problems, large positioning errors are problematic for the maintenance crew, but also for a data-driven prediction approach. If the growth of localised defects is to be monitored, uncertainties about their locations may cause the models to miss a specific fault at a particular location where it was located earlier.

Two basic approaches have been used to pinpoint specific point defects despite uncertainty in localisations. One approach is to calibrate the positions given in the data to known positions, such as switches that give rise to unique signatures in the measurements. The other approach is to handle uncertainty in positioning by splitting the measurements of the track into segments. These segments should be long enough for the probability, of a particular failure being counted in one passage but not in the next, to be low, unless it is near the segment border. Due to ease of calculation and challenges of the first approach, the latter approach was used here, and the segment length was chosen to be 200 m.

Another cause for irregularity is that the measurements are difficult to perform in some months. In winter, it is difficult due to cold and snowy conditions, especially in January, which prevents some of the measurements, and in July due to summer vacations of the staff, see Figure 3. The requirements (BVF 807.2) do not state how to spread out these six measurements in time, just as long as six measurements are taken during a particular year.

However, while it may be practical to measure when measurement conditions are benign, and the staff are available, the irregularities prevent some analyses. It is, for example, likely that the track geometry will deform cyclically for parts of the track due to frost heave. With more frequent and evenly spread measurements, analyses of the effect of frost heave could have been performed. The irregularities and the effect of a suspected frost component will add to the uncertainty of the measurements. However, since the size of this effect is unknown, and is expected to differ between different segments, we cannot compensate for frost. Therefore, such effects will add to the measurement noise.

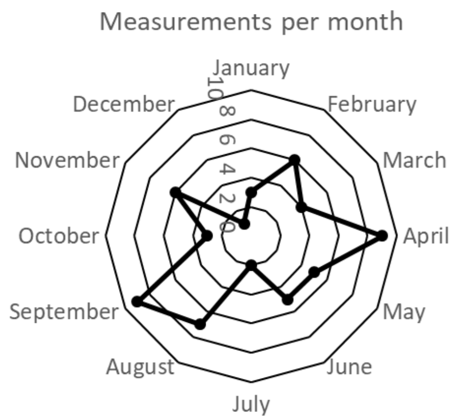


Figure 3. Distribution of measurements taken each month spread on a yearly cycle.

### 4.3 *Data preparation*

There are at least two basic options when studying the condition of a variable for a given segment; looking at how much the variable varies or looking at the maximum value. For the twist measurements, the maximum value of a segment is what triggers maintenance, but on the other hand, variation may be more important for understanding the general condition of the segment. However, here we concentrate on the maximum value seen in the segment. The reason being that the maximum value is more straightforward to connect to the maintenance limits. A study of the variation would give similar results; in fact, the two measures are highly cross-correlated (correlation coefficient 0.82).

The first step was thus to bin the data into 200 m groups for each measurement and then collect descriptive statistics for the groups. This step was performed using Microsoft Power BI Desktop® version 2.61.5192.601 64-bit.

Most methods used for prognostics require, or at least perform better if the data are independent and normally distributed. A normal probability plot of the collected data shown in Figure 4 shows some issues that need to be handled.

The data contained many zero observations and even negative values. Negative values are impossible through the definition of the twist variable, and max values of zero are, in reality, impossible for a 200 m segment. Thus, these observations were removed. With the zero and negative values removed, the distribution still needed a transformation to resemble the normal distribution more closely. A log transformation improved the fit and could be considered reasonable, given that the maximum twist has a true lower limit of zero and no upper bounds.

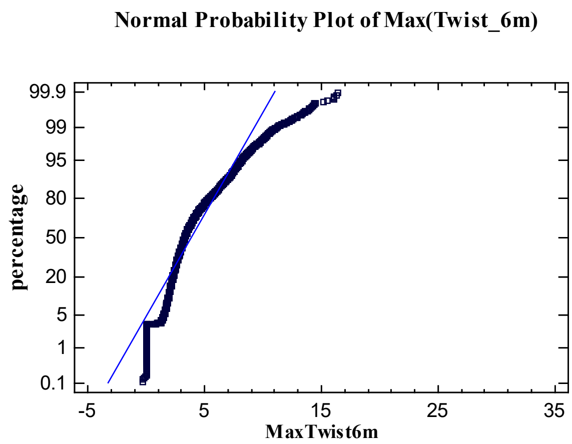


Figure 4. Normal probability plot of max 6 m twist for segments over the period. The line depicts the expected shape of data sampled from a normally distributed population, drawn from the quartiles of the plotted data.

The binned and transformed data were then imported to RStudio, version 1.1.456 for further analyses.

### 4.4 *Data analysis and prognostics*

The prediction ability for different predictive models can be analysed using parts of the data as the training set, create models and then use the remainder and most recent data as the validation set. A validation set of around 20% of the observations are often considered appropriate (Hyndman & Athanopoulos, 2018), but this also depends on the purpose of the prognoses and the length of the available time series. The models' predictive ability can then be compared using different statistical measures. Here, we used data obtained from June 2017 and onward (two years) as the validation dataset and older data as the training set.

Another analysis challenge is that the maintenance actions, in this case, tamping, are meant to change the conditions of the track, and such maintenance will, if it is efficient, change the time series as interventions that affect the observations. The maintenance is procured through a public tendering procedure, and an entrepreneur performs the maintenance of the track. The current maintenance contract states that the entrepreneur should tamp one-third of the track each year as a preventive action. The entrepreneur and Trafikverket (the Swedish Transport Administration) jointly plan the maintenance. Their joint planning is thought to secure that the tamping machine is used efficiently and that the tamped segments are those that are most likely to lead to problems shortly, i.e., that tamping is both efficient and effective. Tamping is also performed as corrective actions when the maintenance thresholds UH2 and KRIT are surpassed.

All segments of the studied track section have been tamped at some point in the twelve years' time series. We have considered models that restart when geometrical conditions improve drastically and other models that do not. We have also chosen to use segments that were not tamped during the validation period (after July 2017) for best fit, and have limited the number of tampings performed on the segment to a maximum of five, none of which was performed in the validation period. Our delimitations resulted in 21 validation segments with an average number of tampings equal to 3.6, during the training data period. We have used the empirical data represented by each studied segment for all tests; that is, we have calculated an individual estimate of the state as well as the state degradation rate.

There are several methods to evaluate the forecasts. Here we have presented the Root Mean Square Error (RMSE), a measure that uses the squared difference between observations and, therefore, its prediction will weigh large deviations more than, say the Mean Absolute Error (MAE), where only the average difference is summed. Hence, the choice between RMSE and MAE will depend on if it is more important that large prediction errors seldom are made or if the prediction errors on the average should be small. We assume that the maintenance planner would not want to miss segments that are degrading fast, and then RMSE will likely be the better choice. The Mean Absolute Percentage Error (MAPE) measures the same as MAE, but is independent of scale. Finally, Hyndman & Koehler (2006) recommends MASE for prediction errors due to its insensitivity to outliers and the ease of interpretation. Again, outliers could be interesting for the planners, so the suggested measure here is RMSE, although the others are shown for reference. Indeed, outlier predictions may not be as important as we would expect.

## 4.5 *Predictive models*

If a prognostic model is to be considered useful, it needs to outperform the benchmark models. Some prognostic models can make forecasts not only for the state and trends, but also for cyclic, seasonal behaviour. All studied models lack seasonality components since the measurement frequency of the underlying data did not allow for good estimations of seasonality components. The benchmark models and prognostic models are schematically presented below.

### 4.5.1 *Benchmark prognostic models*

A standard benchmark for prognoses is to use a random variation approach, i.e. that the last known observation to predict the future state and a draw from the variation of the process will be used to estimate the future state. Here, this approach is called Random walk.

Another common benchmark is to use the average of the training set and then predict that the future observations will be close to that. We have called that benchmark Simple averaging.

A third useful benchmark is to calculate the drift of the series and add that to the last observation, that is, a linear model based on the training set.

### 4.5.2 *More advanced prognostic models*

The first approach with some sophistication compared to the benchmark models is to use a linear model, that is, a regression of the old data and extrapolated into the future. This model is here called the Linear model.

The state-space modelling approach is another means to obtain forecasts, where also some knowledge of, not only the current state, but also what it used to be is considered. The Kalman filter is a well-known state-space model that in the steady state reduces to simple smoothing (Harvey, 1984). Here, we have studied the exponentially smoothed state space model (Hyndman et al., 2002), here called the ETS model.

The well-known ARIMA models, or Box- Jenkins models, also often make good prognoses. ARIMA stands for Auto Regressive, Integrated Moving Average. Often, it can be difficult to know the correct order of the AR, I and MA parts of the ARIMA, and tests in the current data did not indicate that the data fit was best for a particular order. The Auto-Arima approach (Hyndman & Khandakar, 2008) was used to obtain the order and constants chosen directly.

Another state-space model that combines state-space with a Box-Jenkins approach, the exponential smoothing state space model with ARMA errors and a trend is another model that was studied, here called the SAT model (De Livera et al. 2011).

Finally, we use a neural network approach, using a single-layer, feedforward network, which can generate non-linear predictions.

## 5 RESULTS

The time series of the chosen segments were analysed individually, and the errors were summed and are presented in Table 1. Hence, every prediction error presented in Table 1 is a summary of prediction errors over both the eight validation quarters (Q2 2017 – Q1 2018) and the 21 chosen validation segments.

The linear model did outperform the benchmark models and more advanced models for this data based on the error of the predictions of the validation set, see Table 1. Its performance was better, regardless of which measure that is considered (RMSE, Root Mean Square Error; MAP, Mean Absolute Error; MAPE, Mean Absolute Percentage Error). The Cubic spline was the model with the worst performance for all four prediction error measures. See Table 1.

Table 1. Prediction errors of the maximum 6 m twist variable for different predictive models

| | Prediction error measures [log(mm/m)] | | | |
|---|---|---|---|---|
| Model | RMSE | MAE | MAPE | MASE |
| Linear model | 2.77 | 2.46 | 1146 | - * |
| ETS | 3.12 | 2.74 | 1448 | 26.1 |
| Auto ARIMA | 3.35 | 3 | 1544 | 27.6 |
| Simple averaging | 3.36 | 3.02 | 1725 | 28.1 |
| SAT | 3.42 | 3.02 | 1295 | 29.8 |
| Random walk | 3.64 | 3.2 | 1373 | 29.8 |
| Neural Net | 3.74 | 3.09 | 1577 | 29.7 |
| Cubic Spline | 4.85 | 4.4 | 1983 | 38.1 |

* No MASE data was obtained for the linear model

## 6 DISCUSSION

That a linear regression model was the best model is, perhaps, surprising, given that so much effort has been put into research for more effective prognosis models. However, more research is needed to understand if linear models are sufficient when the track has been maintained more often than in the chosen segments. Likely, a forecasting approach that restarts from the latest tamping could be a better choice.

A model that uses a localised trend, restarted from the maintenance should also outperform a linear model based on a regression of the data from the first observation, as it did here. This relationship is also likely due to comparably small effects of tampings in the studied segments.

Nonetheless, a linear model with prediction errors is more powerful than just using the latest observations and more potent than not using the trend. A linear model is also useful in the sense that it is easy to implement in railway management systems that cannot handle iterations. The use of standardised software and transparent programming supports integration and implementation with current practices and it-environments.

Note that the training dataset did include maintenance, while the validation dataset did not. Any maintenance meant to improve the condition of the studied property is likely to affect it, and thus, the empirical models based on that data. A suggested future research objective is to generate data that can be used to model the effect such maintenance has on the properties. Based on studying the data, tamping effects on twist ranges from significantly lowering the twist variation of a segment, to leave conditions seemingly unaffected. Another conclusion is that there seem to be short-term effects when the maintenance was effective, such that the track needs some time to settle. The short-term effect thus seems to be an improvement followed by a rapid degradation process within a year of the maintenance, and then things settle to lower degradation rates. Studies that help predict such behaviour would be useful for modelling the data and for prediction of post-maintenance conditions.

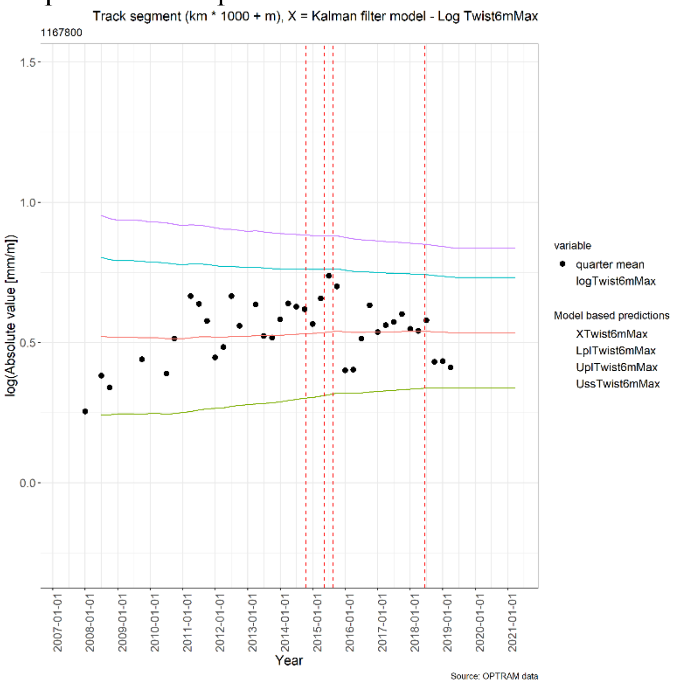


Figure 5. Maximum 6 m twist for one of the segments of track section 119. Vertical hashed lines represent known tampings. Note that some maintenance actions seem to leave the track unaffected, while others decrease the max absolute value. Large effects of the tamping sometimes are followed by rapid degradation rates.

We also expect that much of the variation seen in the data stems from that different measurement trains were used. Scrutinizing the output data, the trains seem to have different pre-processing, so some trains produce smoother curves. This will affect the observations, including the max values seen for a segment. It is not unlikely that properties such as the max value will contain different amounts of error from the trains, for instance due to different calibrations. Both of these measurement errors increases measurement noise. Since such variation may be systematic in that measurement trains are retired and replaced by others, it is not unlikely that such systematics in the data will affect the more complex models more than the simpler ones. For this reason it may be better to focus on data from just one measurement train may, but this hypothesis requires further research to be tested.

## 7 CONCLUSIONS

The paper has studied statistically based prognostic methods for the condition of the railway geometry, in this case, the log-transformed maximum twist (6m) seen on 200 m segments. The best prognoses for this data were obtained by using a linear regression model, which is useful from an implementation standpoint. The more advanced models including Neural Nets, Cubic Splines, Kalman filters, or the SAT models did worse than the simpler Linear model. The more advanced models are also more flexible, and the poorer performance of these models are likely a reflection of that the data does contain unexplained irregularities and likely a low signal to noise ratio, whereby the more flexible methods react to such noise. We expect further improvements if the behaviour of the track after maintenance can be better understood, and we also speculate that improvements can be made by only using measurements from one measurement train.

## 8 ACKNOWLEDGEMENTS

We are grateful for the financial support received from Trafikverket. We are also grateful for valuable contributions from personal at Trafikverket, especially Dr Arne Nissen, Dr Rikard Granström and Mr Kenneth Vågbrink.

## 9 REFERENCES

Asplund, M., Khan, S.A. and Nordmark, T., 2017. Improved wheel-rail system of Sweden's iron ore line. In P.J. Gräbe & R.D. Fröhling (eds), *11th International Heavy Haul Association Conference,* Cape Town, South Africa, 2-6 September 2017.

Bergquist, B. and Söderholm, P., 2016. Measurement System Analysis of Railway Track Geometry Data using Secondary Data. In U. Kumar, A. Parida, & P. Tretten (eds.), *eMaintenance 2016- Challenges in Future Decision-making*, Luleå, Sweden, 14-16 June 2016.

BVF 807.2 (2005). Säkerhetsbesiktningar av fasta anläggningar. In Swedish.

De Livera, A.M., Hyndman, R.J. and Snyder, R.D., 2011. Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106: 1513-1527.

Harvey, A.C., 1984. A unified view of statistical forecasting procedures, *Journal of Forecasting*, 3: 245-275

Hyndman, R.J. and Athanasopoulos, G., 2018. *Forecasting: principles and practice*. OTexts fpp2. Accessed on 2019-02-23.

Hyndman, R.J. and Khandakar, Y., 2007. *Automatic time series for forecasting: the forecast package for R* (No. 6/07). Clayton VIC, Australia: Monash University, Department of Econometrics and Business Statistics.

Hyndman, R.J. and Koehler, A.B., 2006. Another look at measures of forecast accuracy. *International journal of forecasting* 22: 679-688.

Hyndman, R.J., Koehler, A.B., Snyder, R.D. and Grose, S., 2002. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of forecasting*, 18: 439-454.

Larsson-Kråik, P.O., 2009. Managing the wheel–rail interface: Railway infrastructure maintenance in a severe environment: The Swedish experience. In R. Lewis & U. Olsson (eds.), *Wheel–Rail Interface Handbook*). Woodhead Publishing.

Li, D. and Selig, E.T., 1995. Evaluation of railway subgrade problems. *Transportation research record* 1489: 17-24.

Ljung, L., 2001, May. Black-box models from input-output measurements. In *IMTC 2001. Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics* (Cat. No. 01CH 37188) (Vol. 1, pp. 138-146). IEEE.

Nissen, A. 2019. *Personal communication.*2019-02-15.

Smith, A. 2016. Optram Analysis and Forecasting: A Bentley White Paper. *Accessed on-line 2019-02-22 on* https://www.bentley.com/

SS-EN 13848-1:2004+A1:2008. Railway applications – Track – Track geometry quality – Part 1: Characterisation of track geometry. SIS Förlag AB, Stockholm.

Tai, B., Liu, J., Wang, T., Shen, Y. and Li, X., 2017. Numerical modelling of anti-frost heave measures of high-speed railway subgrade in cold regions. *Cold Regions Science and Technology* 141: 28-35.

TDOK 2013:0347. Banöverbyggnad – Spårläge - krav vid byggande och underhåll. (In Swedish). Trafikverket.

Vågbrink, K., 2019, *Personal communication.*2019-02-22.