Coordinated, efficient railway infrastructure maintenance planning

Final delivery – project report

Svensk titel: Samordnad, effektiv planering av järnvägsinfrastrukturunderhåll

Sammanfattning

Det övergripande syftet med detta projekt har varit att optimera järnvägsunderhållsplanering genom att kombinera statistisk analys av underhållsbehov med modeller för optimerad underhållsplanering. Projektgruppen har arbetat mot syftet via två arbetspaket. Det första arbetspaketet har nyttjat ett datadrivet arbetssätt för att prediktera underhållsbehovet (eller anläggningstillståndet) baserat på tillgängliga spårgeometridata insamlade via mätvagnar. Det andra arbetspaketet har fokuserat på att optimera planering och schemaläggning av underhållsåtgärder för en given sektion och en given tidshorisont, under beaktande av kostnader och säkerhet.

De två arbetspaketen har anslutits genom att planering och schemaläggningsmodellering av underhållsåtgärder har baserats på sannolikheten för att geometriavvikelser för ett segment överskrider ett tröskelvärde, exempelvis underhållsgränsen UH1.

Detta dokument sammanfattar arbetet och resultaten av de två arbetspaketen och hur de har kombinerats och länkats för att uppnå projektsyftet. I detta projekt användes data från Stambanan genom övre Norrland (del av Malmbanan), bandel 119, dvs. sträckan Boden-Luleå som fallstudieobjekt. Det underhållsarbete som vi har nyttjat för att exemplifiera arbetssättet har varit spårriktning.

Summary

The overall aim of this project was to optimize infrastructure maintenance planning for railway by combining statistical analysis of maintenance needs with models for optimized maintenance planning. The project team has approached this aim through two work packages. The first work package has focused on a data-driven approach for maintenance need prognostics (or state of the asset) based on available track geometry data, collected by measurement wagons. The second work package has focused on optimized planning and scheduling of maintenance actions for a given section and a given time horizon considering, e.g., costs and safety.

The two work packages are linked so that the planning and scheduling modeling of maintenance actions are based on the probabilities of a segment surpassing a maintenance threshold limit, such as the UH1.

This document provides a summary of the work in the two work packages and describes how the two packages are combined to achieve the project aim. In this project, we have used a part of the Iron ore line, track section 119, Boden-Luleå as the case study object and tamping has been the maintenance action we have used to exemplify our approach.

Content

1. Project facts

Project leader: Prof. Bjarne Bergquist

Project period: 2016-12-01 - 2019-11-30.

Funding and participating organizations:

- InfraSweden2030 (VINNOVA/Formas/Energimyndigheten): 1 496 000 SEK
- Trafikverket: 1 496 000 SEK (direct cost and in-kind)
- eMaintenance365 (Now Predge): 152 520 SEK (in-kind)

1.1 Project steering group

The steering group of this project has had representatives from Trafikverket, eMaintenance365 (Now Predge) and the contractor BDX as well as the project leader (Bjarne Bergquist).

2. Participating researchers

The following researchers at Luleå University of Technology have been active in the research project and are co-authors of this document.

- Prof. Bjarne Bergquist,
- Prof. Murat Kulahci,
- Prof. Athanasios Migdalas,
- Adj. Prof. Peter Söderholm,
- Associate Prof. Erik Vanhatalo,

All researchers mentioned above are affiliated with the Quality Technology and Logistics research subject, Söderholm also with Trafikverket. Additionally, the PhD student Mahdieh Sedghi, funded by another project through Luleå Railway Research Center (JVTC) has worked in work package 2 related to the planning/scheduling modeling. The researchers have also cooperated with Dr Osmo Kauppila at Oulu University in Finland.

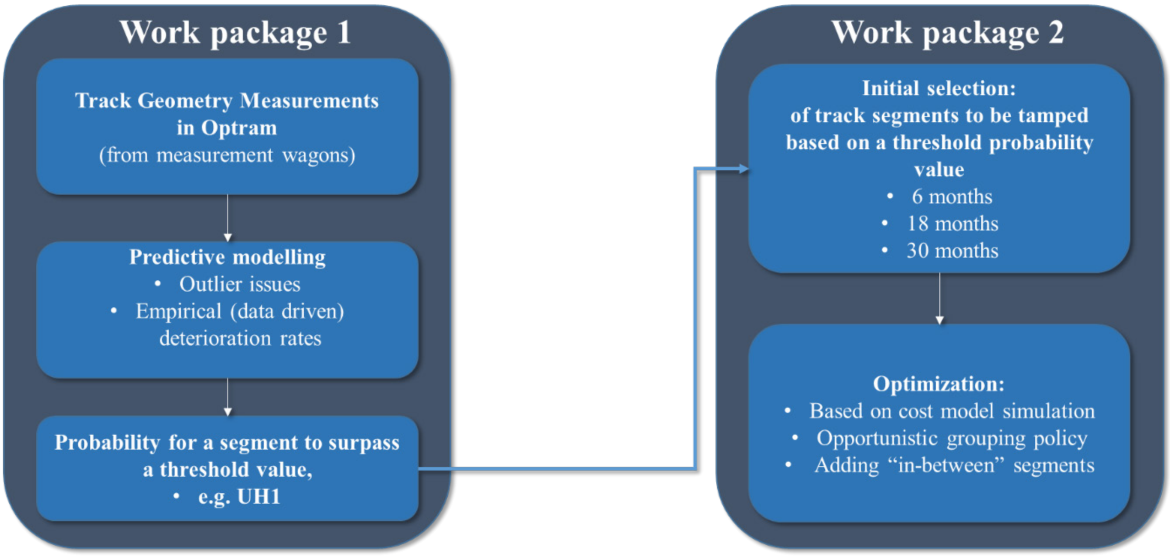
3. Method overview – general approach

3.1 Project aim

> The overall aim of this project was to optimize infrastructure maintenance planning for railway by combining statistical analysis of maintenance needs with models for optimized maintenance planning.

The project team has approached this aim by developing two main work packages. The first work package focused on a data-driven approach for prediction of the maintenance need (or state of the asset). These predictions have been based on available track geometry data collected by measurement cars. The second work package focused on optimized planning and scheduling of maintenance actions for a given section and a given time horizon considering, e.g., costs and safety.

The two work packages were linked so that the maintenance planning and scheduling modeling used the probabilities of a segment surpassing a threshold value, such as the UH1. Figure 3.1 presents an overview of this overall approach.



Software: R, R Shiny, MS Power BI, MS Excel

Software: Python, MS Excel

Figure 3.1. Overview of the main steps in the work packages.

3.2 Targets

In the initial research application, the project listed the following targets:

1. To develop a method for supporting and optimizing maintenance planning.
2. If possible, optimization should also consider costs for different maintenance choices.
3. To perform multivariate data analysis for studies of how maintenance needs for linear assets such as contact wire, track support and track correlate.
4. *(Develop stochastically based, data-driven predictive methods describing degradation rates as a function of track forces).*
5. Cost optimization of maintenance efforts from track section perspective

As the project has proceeded, the project group has not focused on the optional Target 4. While we have performed some work on Target 3, this document will summarize methods and approaches used in working with Targets 1, 2, and 5.

3.3 Specifications and delimitations

Throughout the project, the researchers have worked with data from the Iron ore line between Luleå and Boden (track section 119). Early in the project, the researchers decided to focus on tamping as the maintenance activity to connect the two work packages. This decision also meant that we have focused on track geometry measurements that we have assumed that tamping will improve. Hence, track geometry measurements of special interest in predictive modeling were Shortwave (3-25m) of the vertical level of the rails and the point defects max vertical level and the twist measurements of three and six meters.

We have based our predictions of work package 1 and track segment selection in work package 2 on 200 m segments of the track. This choice was mostly due to the poor positioning precision of older geometry datasets. We considered a length of 200 m wide enough not to miss too many point defects due to closeness to a shifting segment border, but not too wide for maintenance scheduling either. Another reason was that Trafikverket also used 200 meters as chosen segment length.

4. Work package 1 - Prediction/prognosis of the maintenance need

The main aim of work package 1 was to develop a method that would predict probabilities of "failure", here defined as the probability that the geometry measurement of interest will surpass a critical limit, e.g., the UH1 limit. Figure 4.1 illustrates the general idea of this approach as well as the detection of unknown maintenance events and/or outlier detection. For the prognostics, the probability of surpassing the threshold is estimated by calculating the cumulative probability distribution above the maintenance limit.

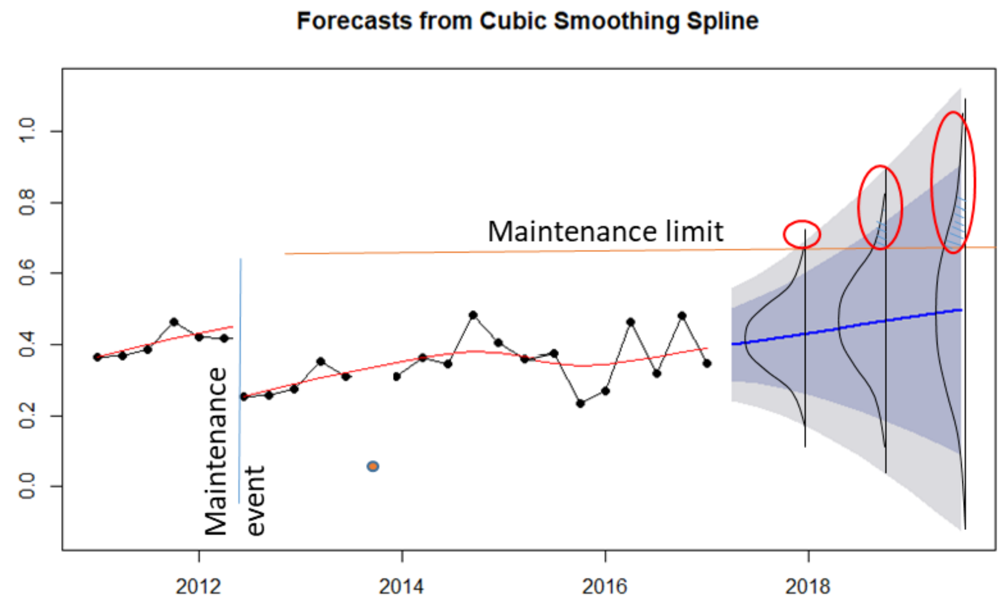


Figure 4.1. Schematic illustration of the approach of providing a predicted probability for the geometry measurement to surpass a maintenance limit at different prediction horizons. Note also the maintenance event, where neighboring observations on both sides show small variation from the new process level, and the outlier The approach needs to handle and classify both event types correctly.

4.1 Codes and packages appended in a zip file

The project has developed prediction models using the open-source software R (The R project, 2019). We have also used the R package Shiny (Shiny, 2019) for data visualization, increased interactivity, and easier manipulation of the modeling output.

The R and R Shiny code for implementation can be found in the attached supplements to this final report under the file name: **"WP1_R_codes.zip"**.

4.2 Data management, cleaning, and data issues

4.2.1 Localization accuracy

The project required dealing with the issue of varying localization accuracy of the measurement trains that record the track geometry data based on which the prediction models are built. The IMV 200 measurement cars have better localization accuracy and lower measurement errors, but these measurement cars had measured too few times to stabilize the time-based prediction models. We, therefore, decided to use all data, including the data from the IMV100N, IMV100M and the Strix measurement cars.

4.2.2 Irregular sampling

Another issue that affects modeling is that the track segments are measured irregularly. The data cleaning procedures handled this through three different methods. One method was to use a spline function to interpolate the measurements data for each segment and then sample from the splines data at a regular frequency. Another approach was to use a Kalman filter in the same way. The final approach uses quarterly averages of measurements of the preceding quarter. The difference in results among these methods was small compared to other sources of variation making a recommendation impractical at this point.

4.2.3 Faulty data

Data cleaning involves removing so-called “faulty” data which could originate from recording problems, e.g., measurements recorded as zero or at different orders of magnitude than other data measured on the track section, (see also Fig. 4.2), or simply missing. Segments where several variables were missing from a measurement point and measurement time were removed altogether. Likewise, we removed measurements from trains running at speeds slower than 50 km/h as this speed is lower than the allowed measurement standard of the wagons (SS-EN 13848-1:2004+A1:2008).

4.2.4 Non-normal data

As the assumption of normally distributed data is important for most of the tested prediction methods, non-normal data will affect the predictive ability of the models. However, in many cases, the normality assumption for the raw geometry measurement data is violated, see, for example, the illustration in Figure 4.2 using data for the maximum twist (6m) measurement.

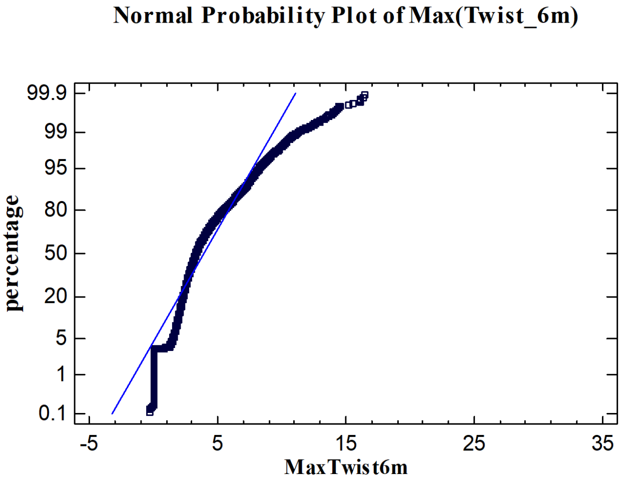


Figure 4.2. Example of a normal probability plot of track geometry data (non-normal data). This specific example comes from the Maximum twist (6m) measurement data.

Therefore, a power transformation of the data is used. The data was typically log-transformed using the base 10 logarithm. Before the transformation, zero observations were removed from variables where there cannot be zeros, for instance, the variation of a property over a 200 m segment, or the maximum of a property such as the twist of the same segment.

4.2.5 Handling unknown reasons for the improvement of track geometry data

In studying the development of track geometry data for track segments in time, there are examples of “improvements” (sudden change in the level of track geometry data) that are not recorded as tamping activities, see Figure 4.3.

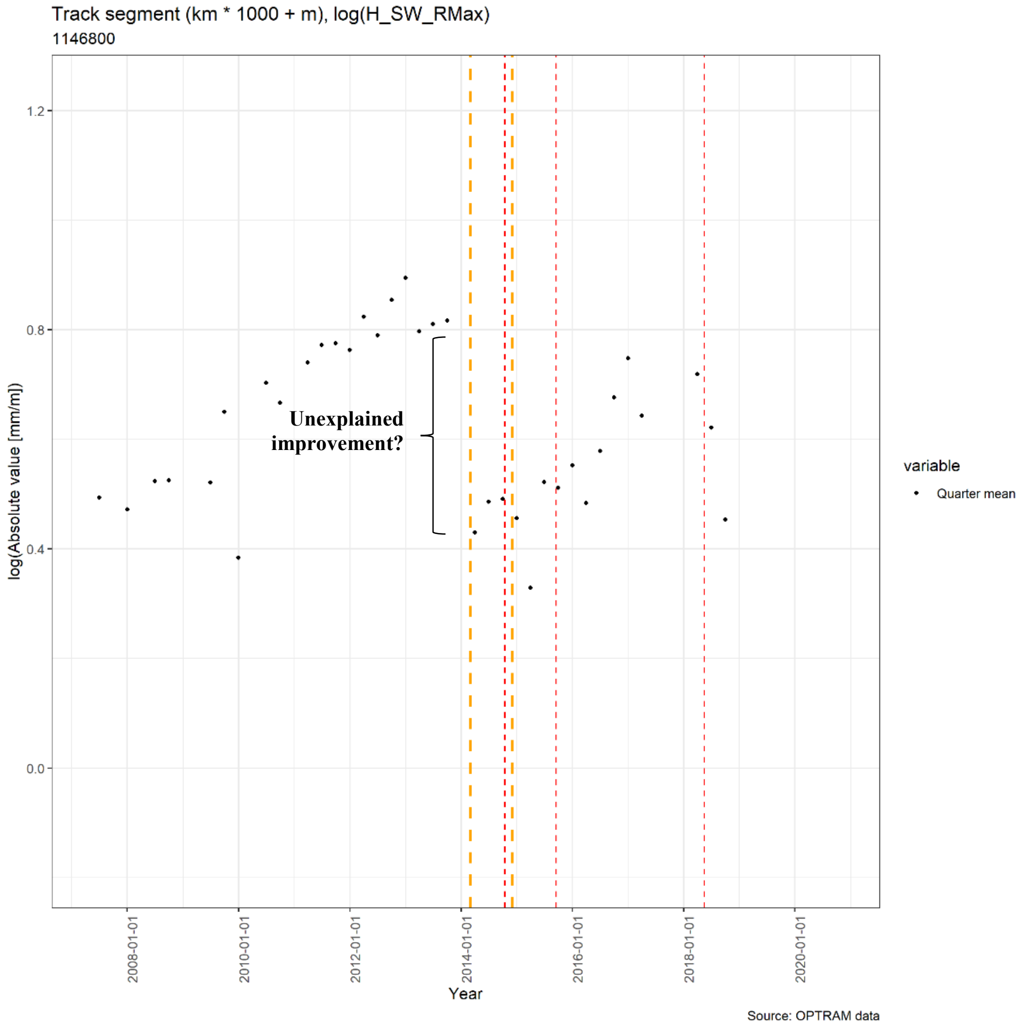


Figure 4.3. Example of unexplained improvement in track geometry, here showing the log-transformed shortwave of the right rail level. The vertical dashed lines show tamping. The red lines are known tampings found in the maintenance database. The orange dashed line (the leftmost dashed line) represents a suspected tamping based on a stable geometrical improvement of the data.

For any prediction model to work, a cleaning procedure needs to identify these maintenance actions empirically based on the behavior of the studied geometry. The modeling then needs to adjust the prediction model accordingly. In this project, two methods for empirically identifying improvements (i.e., potentially unknown tampings) were entertained. Both of these methods can be tuned to maximize the relationship between true tampings and “ghost tampings” just related to variation in the data.

Table 4.1. Methods for empirically identifying improvements of the track geometry variable due to potential unknown maintenance activities.

Method	Explanation
Method I	The test starts with defining the variation of the process based on the standard deviation measure calculated from the one lag moving range of all data. The test is then performed through two one-lag moving ranges of all data and a difference, first by comparing the one-lag moving range to detect outliers or for possible maintenance. Consider a time series of *n* observations $x_{t-n}, x_{t-n+1}, \cdots, x_{t-1}, x_t$. We are interested to know if x_{t-1} compared to earlier observations and that we also have obtained another, later observation, x_t. We can use the time series to estimate the standard deviation based on the one-lag moving range. We can also use the moving range method to successively study the time series for unusual events such as outliers and maintenance actions that have improved the status. When we have calculated the variation of the series, we introduce a test for outliers that we can perform for the data in the series:: $Outlier = \lvert x_{t-2} - x_{t-1}\rvert > k_1$ (4.1) where x_{t-2} and x_{t-1} are two consecutive measurements with at least one observation before (x_{t-3}) and one later (x_t), k_1 is a constant (in the test we used $k_1 = 2$ standard deviations). If (4.1) is true, the outlier neighbors are tested for deviations. Reduction of the property variation is classified as maintenance if the variation reduction between two successive observations is considerable, but the variation difference of the preceding and following observations are small (if there is only one improved observation in a series of poor observations, it is classified as a measurement issue rather than a maintenance action). Thus: $Maintenance = \{\lvert x_{t-3} - x_{t-2}\rvert < k_2 \ \& \ \lvert x_{t-1} - x_t\rvert < k_2 \ \& \ (x_{t-2} - x_{t-1}) > k_1\}$ (4.2) where x_{t-3} and x_t are the fourth to last and the last observations respectively, and k_2 is a constant (in the test we used $k_2 = 1$ standard deviations). If (4.1) but not (4.2) are true, the conclusion is that x_{t-1} is an outlier. If (4.1) and (4.2) are true, the conclusion is that the variation reduction is not due to an outlier but a real reduction of variation, that is, due to maintenance between the times *t-2* and *t-1*.
Method II	If a decrease of, e.g., 85%, in the variable of interest is observed, we assume that an improvement, from an unknown/unrecorded maintenance activity, has occurred on the track segment. The percentage decrease in the variable of interest can be seen as a tuning parameter that can be changed by the user. Outliers in the data are instead identified through calculating Cook's distance for each observation.

4.2.6 Automatic outlier identification through Cook's distance

The approach used for automated identification of outliers is closely related to the choice of the simple regression model for prediction, which is explained in the next section. In linear regression Cook's distance is used as a measure of how influential each observation is when fitting the model. Cook's distance measures the effect of removing a point from regression (as in leave-one-out cross-validation). Cook's distance, D_i, measures the squared distance between the least-squares estimate, $\hat{\boldsymbol{\beta}}$, including all observations and the estimate obtained by deleting the i^{th} observation, $\hat{\boldsymbol{\beta}}_{(i)}$ (Montgomery *et al.*, 2012). Observations with large values of D_i are deemed to have considerable influence on the least-squares estimates, $\hat{\boldsymbol{\beta}}$, and in our approach such observations have been considered as potential outliers. Cook's distance may be expressed as:

$$D_i = \frac{\left(\hat{\mathbf{y}}_{(i)} - \hat{\mathbf{y}}\right)^T \left(\hat{\mathbf{y}}_{(i)} - \hat{\mathbf{y}}\right)}{pMS_{\text{Res}}} \tag{4.3}$$

where $\hat{\mathbf{y}}_{(i)}$ is a vector with the estimated observation values with observation *i* removed during estimation, $\hat{\mathbf{y}}$ is a vector with the estimated observation values including all data, *p* is the number of estimated parameters (e.g. β_0, β_1) in the model and MS_{Res} is the mean square for residuals. For further details, see, e.g., Montgomery *et al.*, (2012). The threshold for Cook's distance was chosen to 2 but may be changed. That is if $D_i > 2$ then the observation is treated as an outlier.

4.3 Prediction modeling

Different prediction models have been tested and entertained throughout the project. We have tested the predictive ability of the following models:

- Benchmark models such as random walk & averaging across all times after a tamping action,
- Simple models such as linear regression and cubic splines,
- More advanced models:
  - Neural networks (feed-forward, single hidden layer),
  - ARIMA (autoregressive integrated moving average models),
  - BATS (Exponential smoothing state space, combined with Box-Cox and ARMA-time series),
  - ETS (state-space model allowing for trend and seasonality), and
  - Markov chain modeling.

The simple linear regression model based on log-transformed data was found to perform best considering root mean squared error (RMSE), mean absolute error (MAE), and mean average percentage error (MAPE). The linear regression model in the log-transformed scale thus has the following model assumption:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon \tag{4.1}$$

where y_t is the current value of the track geometry variable (after transformation), β_0 is the intercept, y_{t-1} is the past value to the track geometry variable and ε is the error which is assumed to be normally and independently distributed with mean zero and constant variance. Further details on simple linear regression can be found in Montgomery *et al.* (2012).

4.4 Calculating the probability of surpassing a maintenance limit

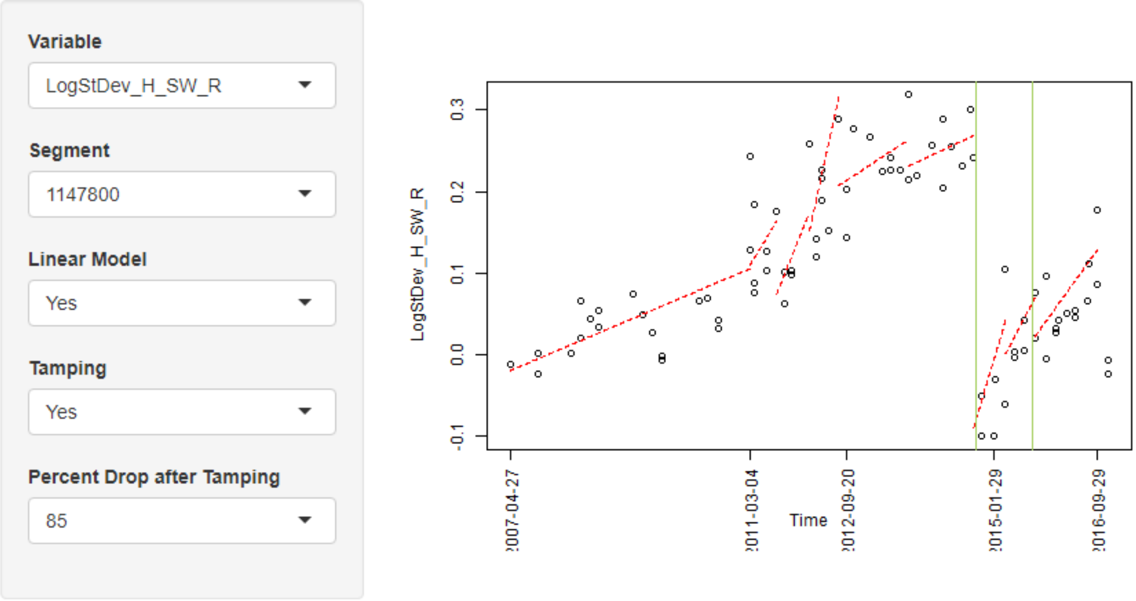
We found the probability of surpassing a given maintenance limit, e.g., UH1, by setting the upper prediction limit (or lower prediction limit if the prediction is above the threshold) and then calculate the α percentile of the *t* distribution as given in Eq. (4.3), which provides a $(1-\alpha)\cdot 100\%$ prediction interval for a new observation, $\hat{Y}_{n+1}$, when the predictor $x = x_{n+1}$.

$$\hat{Y}_{n+1} = t_{\alpha/2,n-2} \cdot \sqrt{MSE} \cdot \sqrt{1 + \frac{1}{n} + \frac{\left(x_{n+1} - \overline{x}\right)^2}{\sum\left(x_{n+1} - \overline{x}\right)^2}} \tag{4.3}$$

In Eq. (4.1) MSE is the mean squared error from the fitted linear regression model, *n* is the number of observations used in fitting the model and $\overline{x}$ is the mean of the predictors in the historical data.

4.5 Example of prediction modeling for one segment on track section 119

Figure 4.4 illustrates how the linear model in Eq. (4.1) works in R Shiny for one track segment using Method II in Table 4.1 empirical identification of improvements of track geometry due to potentially unknown maintenance activities. As illustrated in Figure 4.4 the linear prediction model restarts at these empirically identified improvement points.



Variable
LogStDev_H_SW_R

Segment
1147800

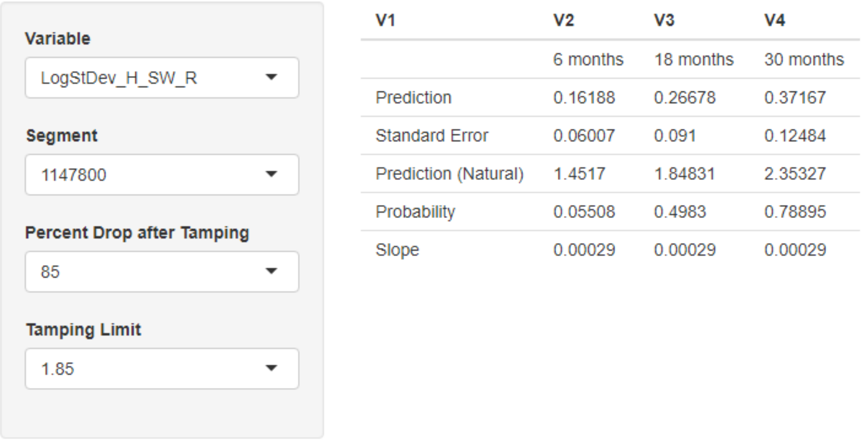
Linear Model
Yes

Tamping
Yes

Percent Drop after Tamping
85

Figure 4.4. Example from the R Shiny App: "shinylm.R". Track geometry variable: Log (base 10) of the shortwave standard deviation for the right rail. Segment: 1147800. The linear models are represented by the dashed red lines. Green vertical lines represent known tamping occasions of the segment. A drop of 85% in the variable of interest causes the model to restart.

Figure 4.5 illustrates how the linear model can be used to provide a probability of surpassing UH1 (see Eq. 4.3) through R Shiny. The R Shiny App gives the predictions, their associated standard errors, the predictions in natural variables, the probabilities of exceeding UH1, and the slopes for 6, 18, and 30 months into the future. Note how the probabilities of exceeding UH1 increase for predictions further into the future.



Variable
LogStDev_H_SW_R

Segment
1147800

Percent Drop after Tamping
85

Tamping Limit
1.85

V1	V2	V3	V4
	6 months	18 months	30 months
Prediction	0.16188	0.26678	0.37167
Standard Error	0.06007	0.091	0.12484
Prediction (Natural)	1.4517	1.84831	2.35327
Probability	0.05508	0.4983	0.78895
Slope	0.00029	0.00029	0.00029

Figure 4.5. Example from the R Shiny App: "Shyniprob.R". Track geometry variable: Log (base 10) of the shortwave standard deviation for the right rail. Segment: 1147800. UH1 limit employed: 1.85. Output: predictions, their associated standard errors, the predictions in natural variables, the probabilities of surpassing UH1, and the slopes for 6, 18, and 30 months into the future. *Note: Please disregard "V1", "V2" etc. as these have no meaning but we were not able to remove them from the visualization.*

5. Work package 2 - Optimized planning and scheduling of maintenance actions

In the second work package, we have developed a decision support tool to plan and optimize maintenance activities (here tamping) based on the predictive models from work package 1. The application of the results from this work package has the following main components and abilities:

- Integration of asset condition prediction and planning optimization,
- Visualization of planning and scheduling decisions in different scenarios,
- Analysis possibilities of the efficiency of the planning process by cost optimization,
- Analysis possibilities of the effect of contract constraints on the maintenance plan,
- Simulation possibilities of the planning process for the tamping activity, and
- Analysis possibilities of the effect of planning scenarios on maintenance costs (preventive and corrective) for the long-term horizon.

5.1 Codes and packages appended in a zip file

In work package 2 the coding of the selection of the segments to tamp is made in the free, open-source software Python, version 3.6. All codes in work package 2 are delivered in a zip file with the name **“WP2_planning.zip”**

5.2 Model inputs, variable definitions and cost calculations

5.2.1 Input parameters and variables in the planning algorithms

Input parameters to the planning algorithms are given in Table 5.1 and model variables used for calculations in the planning algorithms are presented in Table 5.2. In the planning algorithms, we have one common binary decision variable, see Table 5.3, which defines if segment n is selected to be tamped in planning period p.

Table 5.1 Input parameters to the planning algorithms with explanations.

Parameter	Explanation	Units (if applicable)
N	Number of segments.	
P	Number of planning periods, $P = \{1, 2, \ldots, p\}$. In the example below for track section 119 we have three planning periods 6, 18 and 30 month horizons. ($P = 3$)	
C_R	Renting cost for equipment and personnel. Fixed cost per day.	SEK
C_T	Cost for transportation of tamping equipment and personnel.	SEK/h
C_{CM}	Renting cost for equipment and personnel for corrective maintenance, estimated per segment.	SEK/segment
H	Maintenance window. The track is available for maintenance during this window.	Hours.
L_n	Length of each segment.	km
S^T	Tamping machine transportation speed which is assumed to be the speed between the machine depot and the first segment selected for tamping in the period.	km/h
S^{NO}	The speed at which the tamping machine moves when not performing tamping and moving between segments in the tamping plan for the specific period. This speed is assumed to be lower than S^T due to the short distance between tamping points.	km/h
S^O	Tamping machine operation speed.	km/h
T^s	Tamping setup time.	min
DG_{np}	Predicted "degradation level" of the track geometry variable. This is the point prediction of the condition of the track geometry variable of interest for segment *n* in period p.	Depends on track geometry variable
UH1	Preventive maintenance limit. (Underhåll 1). According to TDOK 2013:0347, track geometry variables exceeding this limit should be maintained as early as possible to make sure that the next maintenance limit (UH2) is not reached, with margin, before the next maintenance period.	Depends on track geometry variable
UH2	Corrective maintenance limit. (Underhåll 2). According to TDOK 2013:0347, track geometry faults should be maintained before reaching this level. For track geometry variables exceeding this level, but are below a critical limit, immediate planning of maintenance action should take place and corrective maintenance should be done without delay.	Depends on the track geometry variable. *Note that for the standard deviation for the level there is no UH2 limit defined.*
$TH1$	Probability threshold for selecting segments for preventive maintenance based on the probability of exceeding UH1.	Probability between 0 and 1.
$TH1_{CM}$	Probability threshold for selecting segments for corrective maintenance based on the probability of exceeding UH1. This threshold is only used if there is no UH2 limit defined. $TH1_{CM} > TH1_p$	Probability between 0 and 1.
$TH2_{CM}$	Probability threshold for selecting segments for corrective maintenance based on the probability of exceeding UH2.	Probability between 0 and 1.

Table 5.2 Model variables used for calculations in the planning algorithms.

Variable	Explanation	Units (if applicable)
y_p	The number of tamping days to tamp all selected segments in each period.	Day(s)
TD_{pj}	The transportation time to location *j* in each tamping day for each period.	hours
T^O	Tamping time per segment.	min
T^{NO}	The time it takes the tamping machine to pass a segment when not performing tamping. This is the inverse of S^{NO} given in minutes.	min
Z_{np}	This is a binary variable that states the need for corrective maintenance in the specific planning period *p* for segment *n*. $Z_{np} = 0$ if $P\left(DG_{np} > UH2\right) < TH2_{CM}$ and $Z_{np} = 1$ if $P\left(DG_{np} > UH2\right) \geq TH2_{CM}$. For geometry variables where no UH2 limit is defined, we have instead connected the need for corrective maintenance to the probability of exceeding the UH1 limit: $Z_{np} = 0$ if $P\left(DG_{np} > UH1\right) < TH1_{CM}$ and $Z_{np} = 1$ if $P\left(DG_{np} > UH1\right) \geq TH1_{CM}$.	Binary variable. 0 or 1.

Table 5.3 The model decision variable, which is the final output of the planning algorithms.

Output Variable	Explanation
$X_{np} = \begin{cases} 1 & \text{if segment } n \text{ is tamped in period } p \\ 0 & \text{Otherwise} \end{cases}$	Binary decision variable which defines if segment *n* is selected to be tamped in planning period *p*.

5.3.2 Cost modeling

The cost function is a sum of three cost types: transportation cost, rental cost, and late maintenance penalty cost. Table 5.4 gives some details about how preventive and corrective maintenance costs are calculated.

Table 5.4 Explanation of the preventive and corrective maintenance costs that are considered in the model.

Preventive maintenance	Transportation cost	We have used an estimate of this cost through by estimating the cost of losing/consuming maintenance window time due to transportation on the track to the position for tamping.
	Rental cost	Rental cost for equipment and personnel multiplied by the number of days needed to complete the tamping.
Corrective maintenance	Corrective cost	The probability of having to do corrective maintenance on a segment multiplied by the corrective maintenance cost per segment.

Transportation cost

Transportation cost (TC) is the cost of transportation on the track to the position for tamping.

$$TC = C_T \cdot \sum_{p=1}^{P} \sum_{j=1}^{N} TD_{pj} \tag{5.1}$$

The first part of Eq. (5.1), C_T is the estimation of the fixed cost paid for transportation per hour. The part $\sum_{p=1}^{P} \sum_{j=1}^{N} TD_{pj}$ indicates the total transportation time summed for all periods and segments. The calculation is further explained in the planning algorithms.

Rental cost

The rental cost (RC) is the daily rental cost of equipment and personnel multiplied by the number of tamping days in total in all periods.

$$RC = C_R \cdot \sum_{p=1}^{P} y_p \tag{5.2}$$

Corrective cost (Late maintenance penalty cost)

The penalty cost for late maintenance will come from the need for corrective maintenance. The corrective maintenance cost includes the cost of traffic delay or cancellation, renting equipment and personnel, and possible cost of inefficient maintenance (because of not using the proper tamping machine or due to bad weather conditions). To decide if a certain segment is selected for corrective maintenance, we need to settle the value for Z_{np}, see Table 5.2. For the probability of corrective maintenance, we can use the predicted probability of exceeding UH2 by our prediction model, $P\left(DG_{np} > UH2\right)$ or the predicted probability of exceeding UH1 (if there is no UH2 limit defined). The late maintenance penalty cost (LMC) is calculated as:

$$C_{CM} \cdot \sum_{p=1}^{P} \sum_{n=1}^{N} \left(Z_{np} \cdot (1 - \sum_{k=1}^{p} X_{nk}) \right) \tag{5.3}$$

In (5.3) the part: $Z_{np} \cdot (1 - \sum_{k=1}^{p} X_{nk})$ has the value 1 if segment *n* is predicted to need corrective maintenance at period *p* but has not been selected for tamping.

5.3 Two planning algorithms

Two planning algorithms were developed in Work package 2. Both integrate the predicted probabilities of exceeding a maintenance limit from Work package 1 with tamping planning and scheduling. The primary planning algorithm (Algorithm 1) is used to plan and schedule segments based on the predicted probability of surpassing the maintenance limit (here UH1) and calculate the tamping cost. The second algorithm for cost optimization and grouping (Algorithm 2) considers the possibility of grouping adjacent segments (moving segments to be tamped between planning periods) to reduce the total cost. The pseudo-code for the two algorithms is provided in Appendix 1.

5.4 Adding “problematic” segments to the tamping plan

Working with the scheduling modeling for the tamping maintenance activity, it came to our knowledge that there are segments that we refer to as “problematic segments” where the track geometry is deteriorating fast, i.e. that typically always have poor track geometry. These track segments may be well known by the analyst(s) at Trafikverket based on historical data and experience. We have therefore made it possible to add “problematic segments” to be tamped in more than one period in the algorithms. For example, it may beforehand be known that a certain segment will need to be tamped in all tamping campaigns, i.e. in each planning period. The addition of “problematic segments” is illustrated in the next section, where we apply our planning algorithms to track section 119.

5.5 Example for Track Section 119 – Modeling results vs Actual tamping

The maintenance planning process starts with feeding the predicted probability of surpassing the maintenance limit (UH1) for each segment as an input to the planning and scheduling algorithms. As an illustration, Figure 5.1 shows the predicted probability to surpass UH1 for all relevant segments in track section 119 in 6, 18 and 30 months.

The purpose here is to prioritize and schedule segments to be tamped in three planning periods (6, 18, and 30 months). In this project, instead of point estimates on the geometry variable in each planning period, we select segments to be tamped based on the predicted probabilities of exceeding the UH1 threshold in the planning periods. All the results in this example are based on the input parameters in Table 5.5.

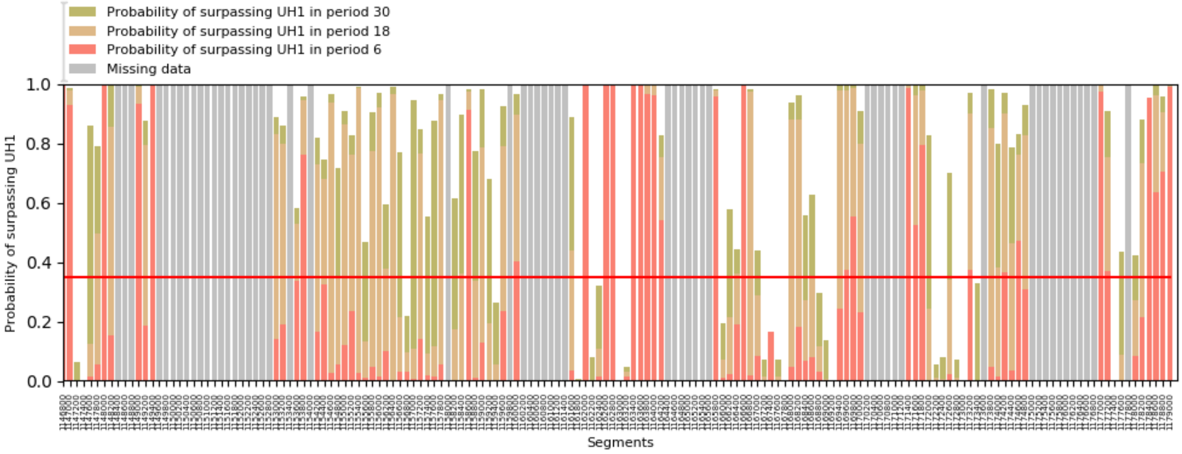


Figure 5.1. Probability prediction diagram for track section 119. Gray segments are excluded. The probabilities are the whole track section for the 6, 18, and 30 months horizons.

The procedure using the primary planning algorithm is further explained in the following three steps:

- Step 1: In the first planning period (6 months horizon), segments exhibiting a probability of surpassing UH1 higher than 0.35 in the 6 months horizon are selected for tamping.
- Step 2: In the second planning period (18 months horizon), segments exhibiting a probability higher than 0.35 in the 18 months horizon, that were not selected in Step 1, will be selected for tamping.
- Step 3: In the third planning period (30 months horizon), segments exhibiting a probability higher than 0.35 in the 30 months horizon, that were not selected in Step 1 or 2, will be selected for tamping. These segments have a lower priority, and there might be a possibility to group them with segments that are selected in period 18.

Figure 5.2 shows an example of the segment selection for the primary planning algorithm in the three planning periods based on the three steps above.

The results of employing the three steps of the model, primary planning, cost optimization and grouping, in the Python code are illustrated in Figure 5.3. The figure also includes a heatmap based on the probabilities of surpassing UH1 for each segment in the three periods. The heatmap illustrates how the probabilities of exceeding UH1 increase over the planning horizon, based on the predictions from work package 1.

Table 5.5 Input parameters relevant for the example and plots provided.

Input variable	Explanation	Value
N	Number of segments.	167
P	Number of planning periods.	3 (years)
T^s	Tamping setup time.	30 min
H	Maintenance window.	6 h
C_T *	Fixed transportation cost.	13 500 SEK/h
C_R	Tamping machine rental cost.	80 000 SEK/day
C_{CM} **	Renting cost for equipment and personnel for corrective maintenance, estimated.	40 000 SEK/segment
S^T	Tamping machine speed.	70 km/h
S^{NO}	None operation speed for crossing each segment during the operation.	4 km/h
S^O	Tamping machine operation speed.	0.5 km/h
UH1	Taken from speed class 3 for the standard deviation of the longitudinal level in TDOK 2013:0347.	1.85
$TH1$	Probability threshold for selecting segments for preventive maintenance based on the probability of crossing UH1.	0.35
$TH1_{CM}$	Probability threshold for selecting segments for corrective maintenance based on the probability of exceeding UH1.	0.5
UH2***	Not defined for the longitudinal level.	

* In this example we have estimated the fixed hourly transportation cost, C_T, as the machine rental cost divided by the maintenance window H. ** C_{CM} is here estimated as being five times higher than the “average “cost to tamp each segment during preventive tamping. ***

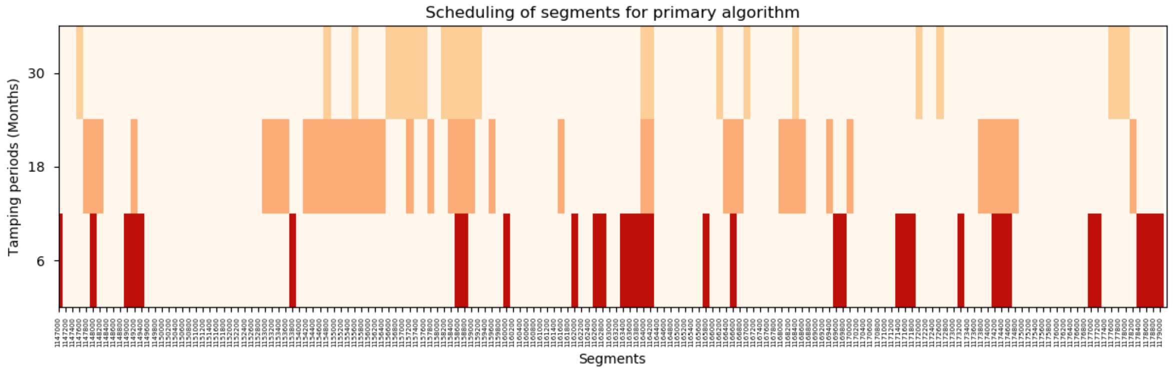


Figure 5.2. Primary segment selection for the 6, 18, and 30 months horizons (Track section 119).

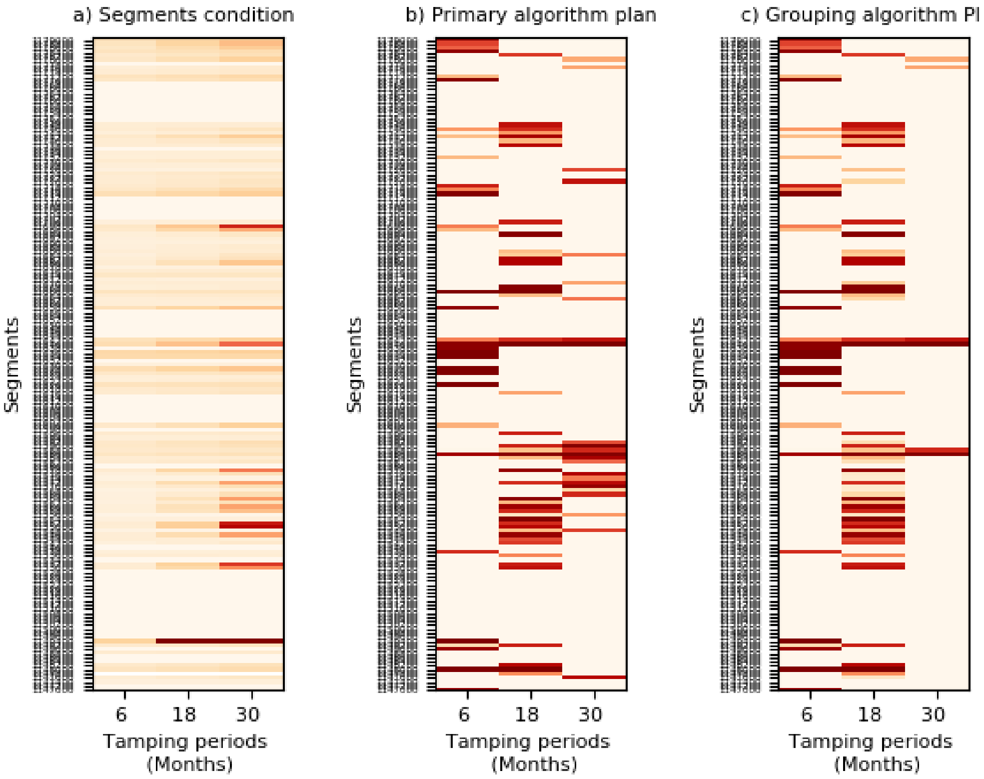


Figure 5.3. *a).* The heatmap shows the probability of surpassing the UH1 level and is an indication of the track condition. Darker colors indicate an increased probability of surpassing the UH1 limit. *b).* The primary tamping plan. *c).* Tamping plan after the grouping policy is employed and in-between segments have been added.

5.5.1 Example and comparison with actual tampings made in track section 119.

To evaluate the proposed planning model, we compared the actual tamping performed from 2017 to 2019 with the suggested tampings from our model using historical data. The probabilities of surpassing the UH1 limit for the 6, 18, and 30 months planning periods were calculated based on the track geometry measurement data up until 2016. Essentially this means that we go back to an assumed position in early 2017 and try to predict and select segments for tamping in the summers of 2017, 2018, and 2019. We then compare these selected segments with the real tampings made during these three years based on recorded tampings from Trafikverket. The results from the two planning algorithms compared with the actual tampings are illustrated in Figure 5.4.

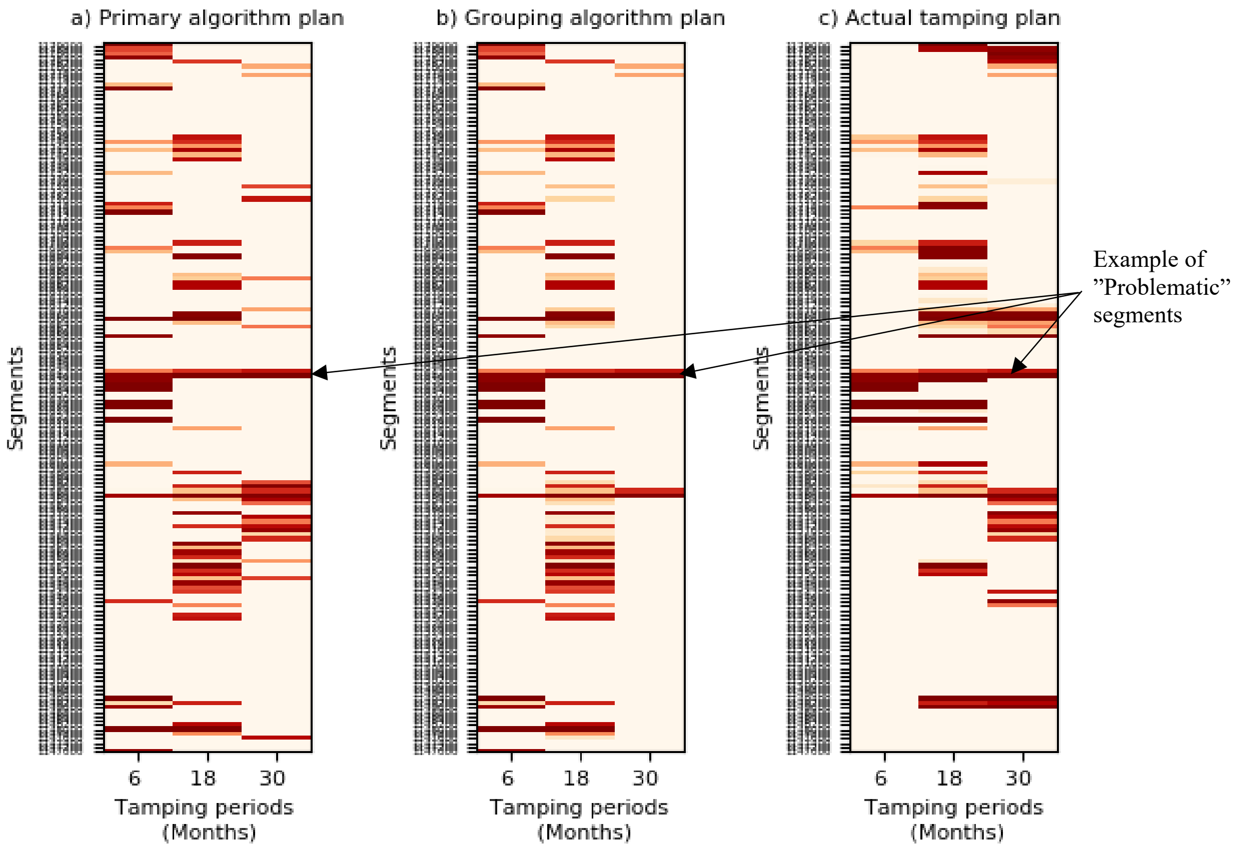


Figure 5.4. Comparison of the proposed tamping plans based on our modeling and the actual tamping plan (the figure on the far right) for the years 2017-2019.

We see from Figure 5.4 that there are some differences in the selected tampings from our model and those performed. There are indeed also similarities. Often the same segments that were tamped were also selected using our planning approach. However, we can see that using the grouping policy, our approach tends to select segments for tamping a bit earlier than the actual tamping performed. We stress that our results depend on model input parameters that may be subject to change and the actual tampings performed were based on other models and deliberations between Trafikverket and the contractor. Furthermore, we have here also illustrated the possibility to add “problematic segments” to be tamped in each period in the Python code. These “problematic” segments are seen as being selected for tamping in all three periods.

5.5.2 Comparison – the number of segments to tamp

We also evaluated the results and efficiency of our planning approach by comparing the total number of tamped segments resulting from the plan from our two algorithms with the actual number of tamped segments on track section 119 during 2017-2019, see Figure 5.5.

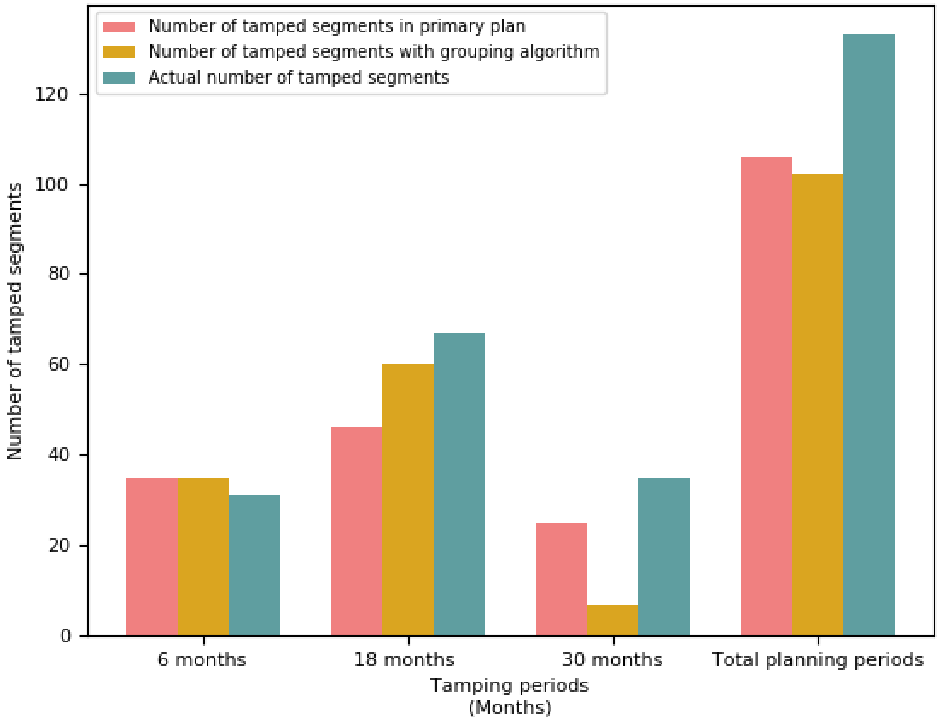


Figure 5.5. Comparison between the total number of suggested tampings in each period from our two algorithms with the actual tampings performed during 2017-2019 in track section 119.

From Figure 5.5, we can see that the total number of tamped segments, both with and without grouping applied, has decreased compared to the actual performed tamping. We also see that through grouping, there is a possibility to reduce the total number of tamped segments slightly. As stated before our planning approach in the two algorithms seem to schedule more tamping in the first 6 month period and less in the 30 month period compared to the actual tampings performed.

Our planning algorithms resulted in less scheduled segments for tamping. However, the actual tamping performed in 2017-2019 is most likely based also on track geometry data made available after 2016, while the results from the two planning algorithms are based on track geometry data up until 2016.

5.5.3 Comparison – total preventive and corrective maintenance costs

In this section, we perform a comparison of the total costs of the scheduled tamping from our two algorithms with the actual tamping activities performed. Please note that we do not have the true costs of the actual plan. This "actual" cost has been estimated through the cost functions we have for our two algorithms.

Note: Please also note that the costs that we provide as inputs to the algorithms are estimated based on discussions we had with Trafikverket and estimations that we have made. Input costs may be discussed and changed if more realistic and detailed costs can be obtained. Therefore, it is probably more interesting to look at the differences between the algorithms and the actual plan than to look at the actual specific costs and savings in SEK. Although we have tried to use realistic cost estimations, they may not be complete and fully realistic in all aspects. The results should, therefore, be viewed as a preliminary comparison of how our planning algorithms perform compared to the actual tamping performed in terms of costs.

The total costs estimated in the algorithms (with and without grouping) and based on the inputs in Table 5.5 for the two algorithms and the actual tamping performed in 2017-2019 are presented in Figure 5.6. In this specific case and for the period tested the result shows a substantial decrease, around 50 percent, in the total cost for our proposed planning approach compared to the actual performed tamping.

Although this is a rather large cost decrease keep in mind that the actual tamping performed in 2017-2019 is most likely based also on track geometry data made available after 2016 while the results from the two planning algorithms are based on track geometry data up until 2016.

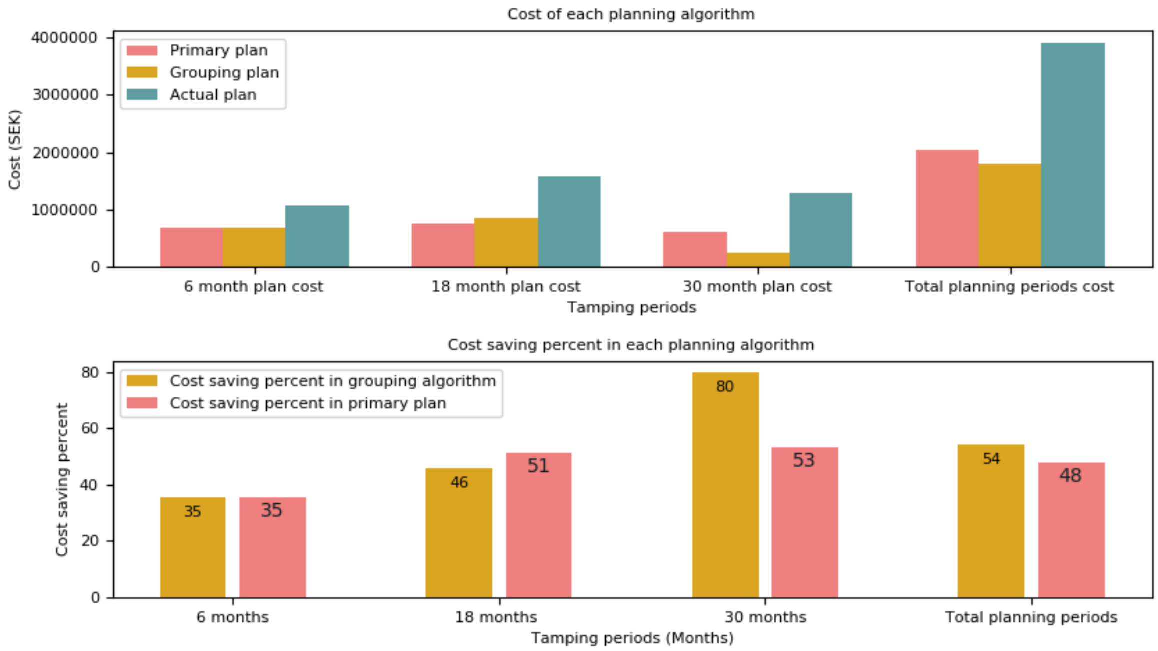


Figure 5.6. Comparison of the total costs (sum of preventive and corrective maintenance costs) for our proposed tamping plans (with and without grouping) and the actual tamping performed in 2017-2019.

For increased understanding of the reasons behind the cost savings reported above, we have separated the categories in the preventive maintenance costs and looked at the corrective maintenance costs for the two algorithms and the actual tamping plan, see Figure 5.7. We can see that the rental costs of the machine dominate the costs for preventive tamping and that the transportation costs are small in comparison.

In this example, we find that the main source of cost-saving reported above is due to a decrease in the corrective maintenance cost. As we do not have an estimation of the actual true corrective maintenance costs, we have here assumed that if we had scheduled tamping according to the actual plan, this would have resulted in the reported corrective maintenance costs in Figure 5.7. There are no reported corrective maintenance costs for our two planning algorithms in Figure 5.7, which may seem strange. However, the proposed algorithms select segments based on the probability of surpassing UH1 and are therefore rather conservative. In fact, in this example for track section 119, this approach makes sure that no track segment is left un-tamped long enough for the track condition to trigger corrective maintenance. However, when we merge the actual tamping performed with the predicted condition of the track from prediction modeling and our way of modeling corrective maintenance, we see that it results in higher corrective maintenance costs.

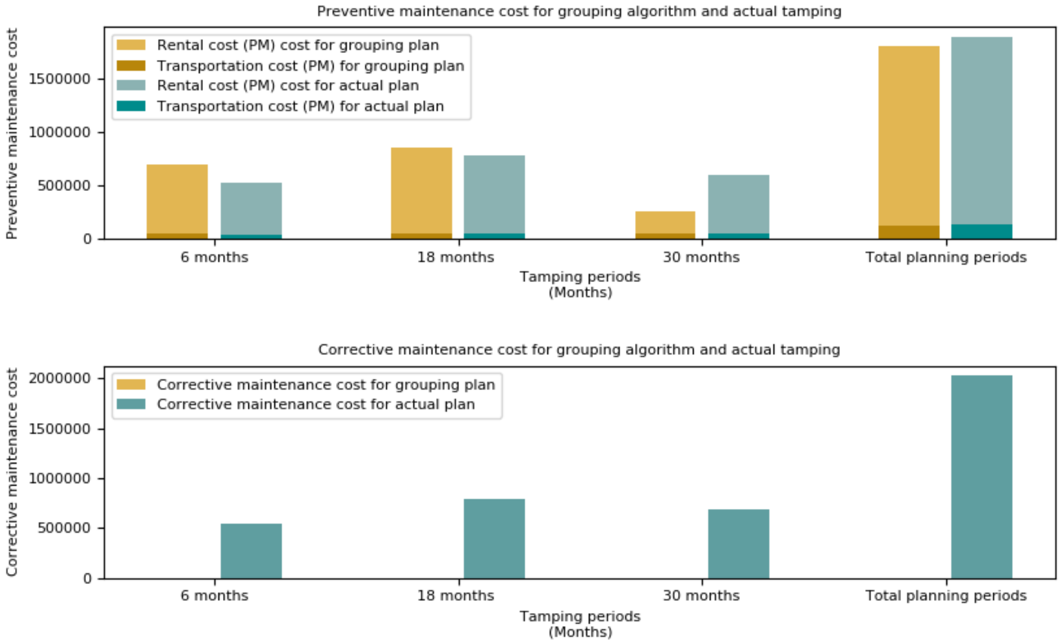


Figure 5.7. Preventive costs (rental costs and transportation costs) for algorithm 2 with grouping compared to the actual tamping performed in 2017-2019. The lower part of the figure also provides the estimated corrective maintenance costs for the actual tamping performed.

5.6 Future research and developments of the model

At the end of this project, we also want to list some future development possibilities that have been discussed related to work package 2 and optimized planning/scheduling of maintenance actions. We list these possible future developments below.

Consider multiple measurements in segment selection

Possible development of the scheduling model is to allow for consideration of multiple measured track geometry variables (multiple aspects of the track condition) in picking segments to maintain. One potential way of doing this is to feed a scheduling model with several predicted probabilities of exceeding a maintenance limit (e.g., UH1) for different geometry variables and change the algorithms and planning models to accommodate several inputs. We believe that this development may align well with how analysts at Trafikverket subjectively view the track condition – a "more complex" function of several geometry variables.

Sensitivity analysis of the model parameters

It would be valuable to perform a more extensive study of the scheduling algorithms sensitivity to model input parameters listed in Table 5.1 above. For example, it may be valuable for Trafikverket to evaluate the effect of the maintenance window length (in hours) on the results in terms of selected segments in each planning period and perhaps especially the total rental costs for the tamping machine. In the project, we have specifically discussed that one value of a decision support tool as the one we provide in these algorithms may be sensitivity analysis as in testing the effect of the maintenance window length on the maintenance plan and costs.

Cost modeling and constraints

As we have mentioned above, we believe that the costs models and the specific estimations of costs may be developed and improved further. It is easy to update specific cost estimates in the inputs (see Table

5.5) while changing and adding costs to the actual model and algorithms requires a bit more work in terms of coding.

The cost models that we used in this project mostly reflects Trafikverket's planning objectives through estimations we have made in discussions with representatives from Trafikverket. In future developments, it may be good to consider more of the contractor's objectives and constraints in the planning algorithms. Currently, the contractor's perspective is considered through the attempt to reduce the total maintenance costs, which are heavily dependent on the total number of days that the contractor needs to rent a tamping machine.

In the planning model, we have made a simplification to start tamping either from the Boden depot or from the Luleå depot. This can potentially be expanded to consider other starting points on the track section as well if they are available.

Planning horizon

The planning algorithms in this project scheduled tamping in three (3) planning periods. However, in the future, it may be valuable to increase the planning horizon and also to evaluate the effects of different planning strategies and input variables in a longer time perspective.

Adding prognoses/estimates of improvement after tamping into the planning model

We have not added predictions/estimates of improvement of track geometry after a maintenance action, such as after tamping. We have not come that far in the project. We have used predictions of the deterioration rate for individual segments. We do lack improvement estimates. However, as continued research, we believe it to interesting to try to estimate the improvement after tamping (if possible for individual segments). If so, we could rule out the use of "problematic segments" where these segments are added to the tamping plan independent of the estimated deterioration level of the segment. Estimates of track geometry improvement can also be used to study the amount of planned tamping on a long-term perspective, such as 10-15 years.

References

Montgomery, D.C., Peck, E.A., and Vining, G.G. (2012). *Introduction to Linear Regression Analysis*, 5th ed. Hoboken, NJ, Wiley.

Shiny (2019). Shiny from R Studio. URL: https://shiny.rstudio.com/

SS-EN 13848-1:2004+A1:2008. *Railway applications – Track – Track geometry quality – Part 1: Characterisation of track geometry.* SIS Swedish Standards Institute.

The R project (2019). The R Project for Statistical Computing. URL: https://www.r-project.org/

Appendix

A.1 Algorithm 1 - Primary planning

The steps of the primary planning algorithm for the tamping maintenance action is presented in Table A.1.

Table A.1 Primary planning algorithm (pseudocode).

```
Algorithm 1: Initial planning

1  Initial selection of segments to tamp based on the predicted probability of surpassing UH1:
   For p = 1, If $P(DG_{n1} > UH1) \geq TH1$ then $X_{n1}$ =1
   For p = 2, If $P(DG_{n2} > UH1) \geq TH1$ and $X_{n1} = 0$ then $X_{n2}$ =1
   For p =3, If $P(DG_{n3} > UH1) \geq TH1$ and $X_{n1} = 0$ and $X_{n2} = 0$ then $X_{n3}$ =1
   Note: $X_{np}$ is a binary indicator variable for segment n and period p. If $X_{np}$ = 1 then segment n is
   selected for tamping in the specific period.

2  Add intermediate segments: if a segment between two adjacent segments selected for tamping is left
   unselected, it is added to the tamping plan.

3  Calculate the tamping operation time based on the speed the tamping machine has when tamping and
   when moving between segments on the track:
   $T^{NO} = \frac{60 \cdot L_n}{S^{NO}}, \quad T^{O} = \frac{60 \cdot L_n}{S^{O}}, \quad \forall$ for n in N

4  Choose the closest starting point between Luleå and Boden: Divide the whole track section into two
   parts. If the first selected segment is in the first half of the track section, the tamping starts from Luleå,
   otherwise, start from Boden.

5  Calculate the required tamping days and transportation time in each period:
   For p in (1, …, P)
   j=1, $y_p = 0$
   While j is less than the total number of segments N
           i=0
           While $\sum_{n=j}^{i+j} \left[T^{NO} \cdot \left(1 - X_{np}\right)\right] + \left[T^{O} \cdot X_{np}\right] + \left[T^{S} \cdot \left(X_{np} \cdot X_{(n-1)p}\right)\right]$ is less than $60 \cdot \left(H - j \cdot \frac{L_n}{S^T}\right)$
           and $(j + i) < N$
        then i=i+1
   then:
   calculate the number of tamping days by the equation: $y_p = y_p + 1$
   Calculate transportation time in each tamping day: $TD_{p(j+i)} = \frac{L_n \cdot (j + i - 1)}{S^T}$
   $j = i - 1$

6  Calculate the total tamping costs:
   Preventive tamping + Corrective maintenance cost:
   $C = C_T \cdot \sum_{p=1}^{P} \sum_{j=1}^{N} TD_{pj} + C_R \cdot \sum_{p=1}^{P} y_p + C_{CM} \cdot \sum_{p=1}^{P} \sum_{n=1}^{N} \left( Z_{np} \cdot (1 - \sum_{k=1}^{p} X_{nk}) \right)$
```

Explanation of the steps in Algorithm 1:

Table A.2 Explanation of the steps in Algorithm 1.

Step	Explanation
1	In this step the binary indicator variable (X_{np}) for segment *n* in period *p* is set to 1 if the segment is selected for tamping in period *p*.
2	If a segment between two adjacent segments selected for tamping is left unselected, it is added to the tamping plan since it is inefficient to stop and then restart tamping leaving only one segment un-tamped.
3	In this step the time needed for tamping one segment and the time needed to transport the tamping machine over one segment are calculated.
4	Choose the closest starting point between Luleå and Boden: Divide the whole track section into two parts. If the first selected segment is in the first half of the track section, the tamping starts from Luleå, otherwise, start from Boden.
5	For each planning period this step in the algorithm calculates the total number of tamping days needed (how long time the tamping machine needs to be rented in each period). The following formula: $\sum_{n=j}^{i+j}\left[T^{NO}\cdot\left(1-X_{np}\right)\right]+\left[T^{O}\cdot X_{np}\right]+\left[T^{S}\cdot\left(X_{np}\cdot X_{(n-1)p}\right)\right]$, sums the transportation time, tamping time, and the set-up time required to perform tamping for segments *j* to *i+j*, where *j* is the starting point and *i* is the last segment that can be tamped in the available time we have for tamping each day. Available tamping time each day is: $60\cdot\left(H-j\cdot(L_n/S^T)\right)$, which is the maintenance window minus the time it takes to get from the depot to the first scheduled segment for tamping.
6	In this step of the algorithm the total cost is calculated as the sum of three costs: cost of transportation + rental cost of the tamping machine + corrective maintenance cost, which are given by the formula: $C=C_T\cdot\sum_{p=1}^{P}\sum_{j=1}^{N}TD_{pj}+C_R\cdot\sum_{p=1}^{P}y_p+C_{CM}\cdot\sum_{p=1}^{P}\sum_{n=1}^{N}\left(Z_{np}\cdot(1-\sum_{k=1}^{p}X_{nk})\right)$

A.2 Algorithm 2 - Cost optimization and grouping

In the second algorithm, we consider the situation that segments can be moved between available periods and be grouped to reduce cost and improve operation.

Table A.3 Algorithm for cost optimization and grouping (pseudocode)

```
Algorithm 2: Grouping policy

1  Initial selection of segments to tamp based on the predicted probability of surpassing UH1:
   For $p = 1$, If $P(DG_{n1} > UH1) \geq TH1$ then $X_{n1} = 1$
   For $p = 2$, If $P(DG_{n2} > UH1) \geq TH1$ and $X_{n1} = 0$ then $X_{n2} = 1$
   For $p = 3$, If $P(DG_{n3} > UH1) \geq TH1$ and $X_{n1} = 0$ and $X_{n2} = 0$ then $X_{n3} = 1$
   Note: $X_{np}$ is a binary indicator variable for segment $n$ and period $p$. If $X_{np} = 1$ then segment $n$ is selected
   for tamping in the specific period.

2  Grouping and cost optimization:
   The decision maker can decide to give a higher priority to cost reduction (default) or to prioritize
   "operational smoothness" by prioritizing grouping of sections although it may come at higher cost. See
   Table A.4 for further details.
   Phase 1: Cost optimization
   a). Start from the first selected segment and move forward until the last segment.
   b). For all selected segments in period 2 and 3 (18 and 30 months horizons), find the optimal allocation of
   the segments based on the total maintenance cost and considering the predicted probability of surpassing
   the UH1.
   Phase 2: Grouping
   a). Start from the first selected segment and move forward until the last segment.
   b). Count the number of selected segments in step 1 of the algorithm (initial selection) in a moving
   window of five (5) adjacent segments with the current entertained segment as the center. The number of
   selected segments in this window is here called the density value for the currently entertained segment.
   The density value can vary in integer steps between 0 and 4 selected segments in the moving window.
   c). Assign the segment to the period with the highest density value and the lowest cost given the following
   rules:
   If period 2 has the lowest cost:
       If period 2 has the highest density value: → The segment is assigned to period 2.
       If period 3 has the highest density value: If the cost reduction is prioritized the segment is assigned to
       period 2; otherwise it is assigned to period 3.
   If period 3 has the lowest cost:
       If period 2 has the highest density value:
           If the predicted probability of surpassing the UH1 in period 3 is less than 0.5: → If the cost
           reduction is prioritized. segment is assigned to period 3
           Otherwise, the segment is assigned to period 2.
       If period 3 has the highest density value:
           If the predicted probability of surpassing the UH1 in period 3 is less than 0.5: → The segment is
           assigned to period 3.
           Otherwise, assign it to period 2.

3  Add intermediate segments (see Algorithm 1)

4  Calculate the tamping operation time based on the speed the tamping machine has when tamping and
   when moving between segments on the track:
   $T^{NO} = \frac{60 \cdot L_n}{S^{NO}}, \quad T^{O} = \frac{60 \cdot L_n}{S^{O}}, \quad \forall$ for $n$ in $N$

5  Choose the closest starting point between Luleå and Boden (see Algorithm 1)
```

```
6   Calculate the required tamping days and transportation time in each period:
    For $p$ in $(1, \ldots, P)$
    $j=1,\ y_p = 0$
    While $j$ is less than the total number of segments $N$
            $i=0$
            While $\sum_{n=j}^{i+j}\left[T^{NO}\cdot\left(1-X_{np}\right)\right]+\left[T^{O}\cdot X_{np}\right]+\left[T^{S}\cdot\left(X_{np}\cdot X_{(n-1)p}\right)\right]$ is less than $60\cdot\left(H-j\cdot\frac{L_n}{S^T}\right)$
            and $(j+i)<N$
        then $i=i+1$
    then:
    calculate the number of tamping days by the equation: $y_p = y_p + 1$
    Calculate transportation time in each tamping day: $TD_{p(j+i)} = \frac{L_n\cdot(j+i-1)}{S^T}$
    $j=i-1$
7   Calculate the total tamping costs:
    Preventive tamping + Corrective maintenance cost:
    $C = C_T\cdot\sum_{p=1}^{P}\sum_{j=1}^{N}TD_{pj}+C_R\cdot\sum_{p=1}^{P}y_p+C_{CM}\cdot\sum_{p=1}^{P}\sum_{n=1}^{N}\left(Z_{np}\cdot(1-\sum_{i=1}^{p}X_{ni})\right)$
```

Explanation of the steps in Algorithm 2:

Table A.4 Explanation of the steps in Algorithm 2.

Step	Explanation
1	In this step the binary indicator variable (X_{np}) for segment *n* in period *p* is set to 1 if the segment is selected for tamping in period *p*.
2	In our model cost reduction is the default perspective for the decision maker. However, in the grouping algorithm we have offered a possibility to prioritize grouping of segments to have a smooth tamping operation. Before starting the algorithm, the decision-maker can decide if grouping or cost reduction has higher priority. This can be implemented by an extra input to the algorithm through an Excel file. In this step the initial selection of the segments is rescheduled in the phase 1 and 2 of the defined cost optimization and grouping policy.
3	If a segment between two adjacent segments selected for tamping is left unselected, it is added to the tamping plan since it is inefficient to stop and then restart tamping leaving only one segment untamped.
4	In this step the time needed for tamping one segment and the time needed to transport the tamping machine over one segment are calculated.
5	Choose the closest starting point between Luleå and Boden: Divide the whole track section into two parts. If the first selected segment is in the first half of the track section, the tamping starts from Luleå, otherwise, start from Boden.
6	For each planning period this step in the algorithm calculates the total number of tamping days needed (how long time the tamping machine needs to be rented in each period). The following formula: $\sum_{n=j}^{i+j}\left[T^{NO}\cdot\left(1-X_{np}\right)\right]+\left[T^{O}\cdot X_{np}\right]+\left[T^{S}\cdot\left(X_{np}\cdot X_{(n-1)p}\right)\right]$, sums the transportation time, tamping time, and the set-up time required to perform tamping for segments *j* to *i*+*j*, where *j* is the starting point and *i* is the last segment that can be tamped in the available time we have for tamping each day. Available tamping time each day is: $60\cdot\left(H-j\cdot(L_n/S^T)\right)$, which is the maintenance window minus the time it takes to get from the depot to the first scheduled segment for tamping.
7	In this step of the algorithm the total cost is calculated as the sum of three costs: cost of transportation + rental cost of the tamping machine + corrective maintenance cost, which are given by the formula: $C=C_T\cdot\sum_{p=1}^{P}\sum_{j=1}^{N}TD_{pj}+C_R\cdot\sum_{p=1}^{P}y_p+C_{CM}\cdot\sum_{p=1}^{P}\sum_{n=1}^{N}\left(Z_{np}\cdot(1-\sum_{i=1}^{p}X_{ni})\right)$