Railway Track Tamping Maintenance Cycle Prediction Model Based on Power-Time-Transformed Wiener Process

Ru An [1,2], Lei Jia [2], Yuanjie Tang [3,*], Yuan Tian [4] and Zhipeng Wang [5]

1 College of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518060, China
2 Shenzhen Urban Transport Planning Center Co., Ltd., Shenzhen 518057, China
3 Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China
4 School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
5 Line Branch Company Affiliated with Beijing Mass Transit Railway Operation Co., Ltd., Beijing 100082, China
* Correspondence: tangyj@bjtu.edu.cn; Tel.: +86-18611624261

Abstract: Predicting the tamping cycles of railway track sections based on track geometry deterioration rules is necessary to reasonably allocate the limited tamping maintenance resources. Existing research on track geometry deterioration modeling for tamping cycle prediction lacks simultaneous consideration of the deterioration characteristics including heterogeneity, uncertainty, and historical dependence, thereby limiting the accuracy of the prediction results. Thus, this study considers a 200 m track segment as the basic object and uses the power-time-transformed Wiener process to develop a deterioration prediction model for the longitudinal level of a segment between two adjacent tamping operations. Moreover, it individually estimates the model parameters for each track segment to predict the tamping maintenance cycle for each segment combined with the tamping maintenance threshold of the longitudinal level index. Finally, through a case study of the Chinese Lanxin Railway line, the effectiveness of the proposed model and different parameter estimation methods is assessed.

Keywords: railway track; tamping maintenance; power-time-transformed Wiener process; maximum likelihood estimation; marine predators algorithm; adaptive MCMC

1. Introduction

1.1. Background

Ballast tamping using heavy machines is the most important and effective maintenance operation used by railway track managers in various countries to restore the geometry conditions of ballasted tracks, but it is expensive [1]. The traditional preventive maintenance strategy based on fixed time intervals can lead to "over-maintenance" and "under-maintenance", which not only affect the safety and availability of railway tracks but also increase maintenance costs [2]. The pressure on maintenance operations also continues to increase with railway network expansion. Taking China's railways as an example, from 2011 to 2021, the annual tamping kilometers increased from 103,138 km to 190,890 km, accounting for nearly 80% of the total extension length of the ballasted track [3]. Therefore, to ensure safety and reasonably control maintenance costs, the maintenance strategy for railway infrastructures in many countries has gradually shifted to condition-based, proactive maintenance [4]. Furthermore, a tamping operation requires heavy machines and dozens of workers. Therefore, it is necessary to plan tamping maintenance operations in advance because of limited maintenance resources and budgets [5]. To perform reasonable planning of track tamping operations under a condition-based maintenance strategy, it is necessary to estimate the tamping maintenance cycle (i.e., the time interval between two adjacent tamping interventions) based on the deterioration of the track geometry condition to predict the tamping maintenance needs.

1.2. Literature Review

Extensive research has been conducted on how to predict the tamping maintenance demand using methods for modeling track geometry deterioration. The types of models studied for this purpose can be divided into three categories: deterministic models, machine learning models, and probabilistic models [6–8].

Deterministic models use statistical analysis methods, such as regression analysis, to analyze the functional relationship between the evolution of track geometry and service time, accumulated traffic load, and other factors of influence [9,10]. Caetano and Teixeira [11] established linear deterioration prediction models for individual track sections that describe the change in the standard deviation of the longitudinal level over time or with the cumulative change in total traffic tonnage between two successive tamping operations. Andrade and Teixeira [12], Movaghar and Mohammadzadeh [13], and others identified multivariate linear regression relationships between the standard deviation of the longitudinal level and factors such as the average annual total traffic tonnage, service time, and track structure. Some researchers believe that the track geometry deterioration pattern between two successive tamping operations exhibits nonlinear characteristics. Based on methods such as exponential and multistage linear fitting, nonlinear prediction models have been developed for tamping maintenance cycles [14,15].

In recent years, machine learning methods have received increased attention because of their usefulness in modeling the deterioration of infrastructure conditions [16]. Artificial neural networks (ANNs), support vector machines (SVMs), and ensemble learning are machine learning techniques commonly adopted in the field of track geometry condition prediction [7,17–20]. These studies mainly sought to establish models for the relationship between the future and current states of track geometry conditions as functions of various factors for use in the short-term prediction of track geometry conditions. For long-life assets, such as tracks, tunnels, and bridges, machine learning techniques commonly use inspection and monitoring data along with data on related influencing factors as inputs for short- to medium-term condition prediction [16,21,22]. Because it is difficult to collect sufficient maintenance cycle or lifespan data to support model training, these techniques are not yet capable of supporting application scenarios, such as risk assessment and whole life-cycle maintenance cost prediction [23]. Moreover, machine learning prediction models previously established for the deterioration patterns of track geometry conditions are essentially regression models. Without sufficient consideration of all of the input factors influencing the deterioration process, they may not be able to accurately capture the uncertainty of the deterioration process, thereby failing to predict track geometry conditions with high accuracy [23].

Probabilistic models use stochastic process methods or probability distribution functions to predict the deterioration patterns of track geometry conditions, considering the uncertainty of the deterioration process. The advantage of these methods is that, on the one hand, they can predict the track geometry condition or state rate at any time during the entire life cycle, while on the other hand, they can output indicators, such as life expectancy, that support maintenance planning for a long period. A representative method is the Markov stochastic process, which discretizes the assessment of track geometry conditions into different state levels and describes the deterioration process of the transitions between these states [24,25]. Considering that track geometry condition indicators are continuous variables and using discretized state levels to describe the deterioration process may not ensure accuracy, some scholars have employed a continuous state-space stochastic process method, that is, the linear Wiener process method, to describe the deterioration of condition indicators, such as the standard deviation of the longitudinal level over a track segment [26,27]. However, the above methods all assume that the deterioration process of track geometry conditions has a Markov property—that is, that the future state only depends on the current state and is independent of past states. However, not all infrastructural deterioration processes satisfy this assumption. Whether the deterioration process of the track geometry condition satisfies this assumption requires a large amount of research

for verification [6]. In their research on the degradation of highly reliable products, Ye and Xie [28] and Xiao and Ye [29] noted that degradation rates often vary with time or stress levels, exhibiting nonstationary characteristics. To address this issue, they proposed a novel class of Wiener degradation models incorporating time-varying parameters to capture the dynamic evolution of degradation rates. These studies offer a fresh modeling perspective that potentially overcomes the limitations associated with the Markovian assumption when applied to predicting track geometry conditions.

1.3. Knowledge Gap

Based on the literature review, it can be seen that in the context of modeling the deterioration pattern of track geometry condition for the prediction of tamping maintenance demands, stochastic process methods provide more direct bases for track maintenance decisions, such as maintenance cycles and remaining useful life, compared to machine learning models, and have stronger applicability. However, existing studies based on stochastic process methods mostly adopt the Markov process, assuming that the deterioration process of track geometry conditions is memoryless, which overlooks the exponential characteristic that the deterioration process within a tamping cycle starts at a slow rate and accelerates over time [15,30,31], making it difficult to ensure the accuracy of the prediction results.

1.4. Research Goal and Contributions

To predict the demand for railway track tamping maintenance, we developed a prediction model of the longitudinal-level deterioration process within the tamping maintenance cycle for a track segment. Based on the model, the tamping maintenance cycle period of each track segment is predicted combined with the maintenance limit value of the longitudinal level indicator. This study makes the following key contributions.

(1) To our knowledge, this study is the first attempt to use the power-time-transformed Wiener process (PTT-WP) to predict the tamping maintenance cycle. Simultaneously, this study estimates the deterioration model parameters individually for each track segment, achieving simultaneous consideration of deterioration characteristics, including heterogeneity, uncertainty, and historical dependency, avoiding the limitations that the Markov property assumption may impose on the accuracy of the prediction model.

(2) Considering the characteristics of the problem, two parameter estimation methods were designed: the maximum likelihood estimation (MLE) and the adaptive Markov chain Monte Carlo (MCMC). Furthermore, under the framework of the MLE method, the effectiveness of three different parameter-solving algorithms, namely the marine predators algorithm (MPA), the genetic algorithm (GA), and the gradient descent algorithm (GD), are compared. To our knowledge, this is the first attempt to use MPA for this type of problem. Finally, by conducting a prediction error analysis based on actual cases, the effectiveness and performance of the different parameter estimation methods were assessed, providing useful guidance for similar studies.

The Methodology section (Section 2) presents the prediction model developed for the tamping cycle of a track segment. The Estimation of Model Parameters section (Section 3) describes the methods used to estimate the proposed prediction model parameters. The Case Study section (Section 4) presents a case study using actual data from the Chinese Lanxin Railway line to verify the effectiveness of the proposed method. The main conclusions and directions for future research are presented in the final section (Section 5).

2. Methodology

2.1. Problem Description

In this study, the track tamping maintenance cycle period refers to the time interval between two successive tamping maintenance operations, and its length depends on the deterioration pattern of the track geometry during the tamping cycle. Because of the influences of various factors such as the track structure, transportation organization, and maintenance history, the deterioration rules of track geometry conditions at different spatial

locations are inconsistent, exhibiting heterogeneity; at the same time, the deterioration process of the track geometry condition at a specific spatial location is also characterized by uncertainty. This study used a 200 m track segment as the basic research object, and based on historical inspection data of the longitudinal level of each track segment, the stochastic process method of the power-time-transformed Wiener process was used to develop the longitudinal-level deterioration prediction model between successive tamping maintenances. For each track segment, the deterioration model parameters were estimated individually using the historical track geometry inspection data of the track segment itself, with simultaneous consideration of the heterogeneity and uncertainty of the deterioration process. Based on the track geometry condition deterioration model with heterogeneous deterioration parameters across different segments, the time interval for the track geometry condition indicator of each segment to reach the tamping maintenance threshold, that is, the tamping maintenance cycle of each track segment, was predicted. We propose the following premises and hypotheses for the prediction model developed in this study.

(1) Railway infrastructure managers in most countries (such as the United Kingdom, France, and Sweden) adopt the longitudinal level as the main track geometry parameter to trigger preventive tamping maintenance operations [32]. As illustrated in Figure 1, the longitudinal level represents the vertical plane irregularities of each rail [33]. The measurement is typically obtained using a track geometry car equipped with an inertial reference system. Some researchers have noted that the longitudinal-level deterioration trend over time between two successive tamping maintenance events is more evident and regular [34]. Therefore, in this study, the standard deviation of the longitudinal level over a 200 m track segment was used as the track geometry condition indicator triggering tamping maintenance.
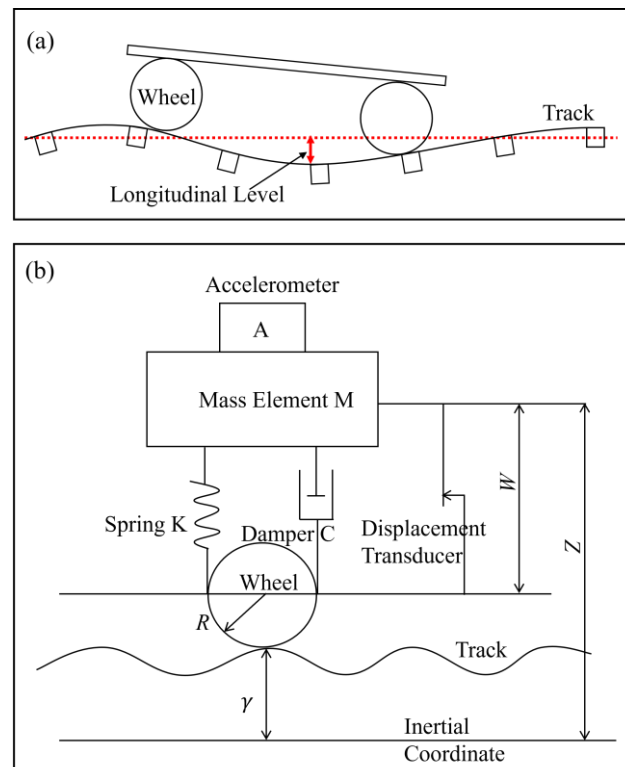

Figure 1. (**a**) Schematic diagram of the principle of track longitudinal level (red lines illustrating the deviation from the ideal geometry). (**b**) Schematic diagram depicting the principle of the inertial reference method.

(2) Assumptions regarding the initial value of the longitudinal-level indicator in the deterioration process were as follows. As shown in Figure 2, in the tamping operation process, operators employ heavy tamping machines to lift, level, and align the track based on measurement results. This operation ensures that the track geometry parameters meet the requirements specified in track design standards or maintenance codes [33,35]. Because a track dynamic stabilization car is used to stabilize a track quickly after a tamping operation [36], in this scenario, the bedding-in process (that is, the process by which the track ballast gradually stabilizes and track geometry gradually improves after the tamping operation) mentioned by Audley and Andrews [37] could be ignored. Therefore, this study assumed that the initial value of the longitudinal-level indicator during the tamping maintenance cycle was approximately equal to the first value detected after the last tamping maintenance operation.
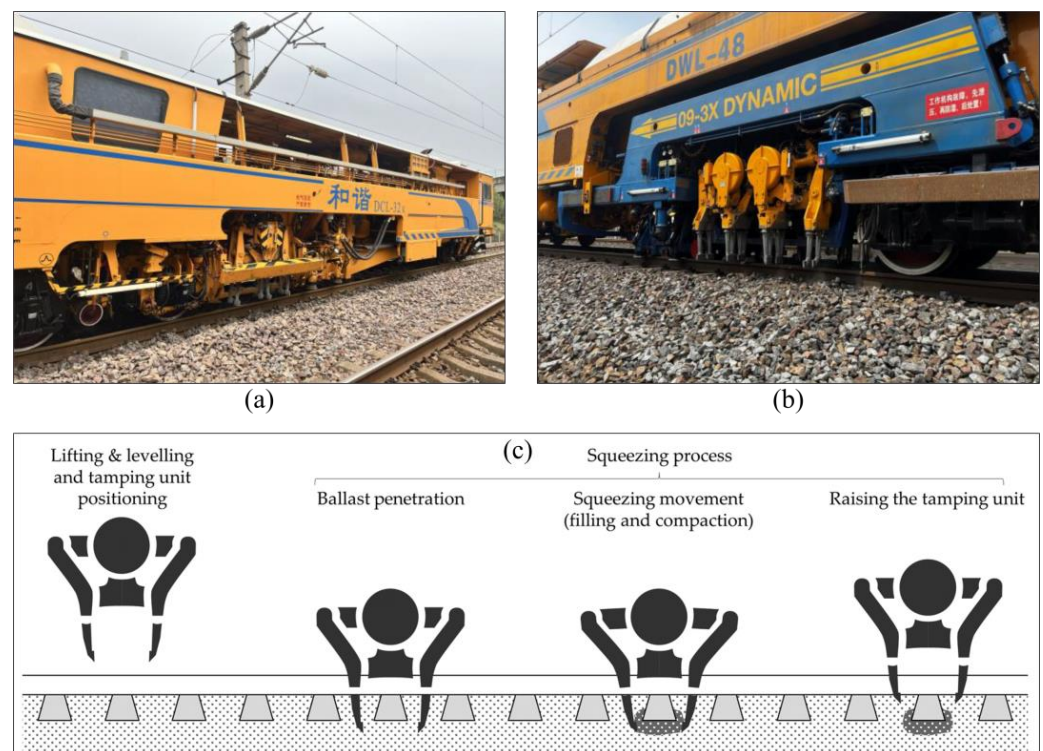

Figure 2. The main models of heavy tamping machines in China and the main phases of a tamping process: (**a**) Hexie DCL-32 tamping machine [33], (**b**) tamping machine DWL-48 that integrates tamping and stabilization [33], and (**c**) the main phases of a tamping process [35].

2.2. Longitudinal Level Deterioration Model

The power-time-transformed Wiener process (PTT-WP) method is a random method widely used in modeling research on facility performance deterioration and useful life prediction. The PTT-WP method is quite suitable for describing the long-time deterioration process with the uncertainty of continuous condition variables [38–41]. Moreover, it can capture the characteristics of the deterioration process in which the rate of deterioration accelerates over time. In this study, the PTT-WP method was used to develop a deterioration prediction model for the longitudinal level of a 200 m track segment in the tamping maintenance cycle.

According to the general form of PTT-WP, for track segment *k*, the formula for the deterioration prediction model of the longitudinal level in the tamping maintenance cycle is as follows:

$$X_k(t) = X_{0,k} + \beta_k \Lambda(t; \theta_k) + \sigma_k B(\Lambda(t; \theta_k)), \tag{1}$$

where $X_k(t)$ is the predicted value of the standard deviation of the longitudinal level (SDLL) at time t during the tamping maintenance cycle for track segment k. $X_{0,k}$ is the initial value of the SDLL during the tamping maintenance cycle. β_k is the drift coefficient. σ_k is the diffusion coefficient. The Brownian motion, denoted as $B(\cdot)$, is employed to model the uncertainty inherent in the deterioration process, particularly capturing the improvements in the SDLL attributable to local maintenance activities. $B(\cdot)$ is characterized by increments that follow a normal distribution. Specifically, in this study, the increment $B(\Lambda(t;\theta_k)) - B(\Lambda(s;\theta_k))$ follows a normal distribution with a mean of 0 and a variance of $\Lambda(t;\theta_k) - \Lambda(s;\theta_k)$. Combined with the calculation formula for $X_k(t)$ as shown in Equation (1), it can be further deduced that $\Delta X_k = X_k(t) - X_k(s)$, representing the increment of SDLL at any two time points, is normally distributed with mean $\beta_k(\Lambda(t;\theta_k) - \Lambda(s;\theta_k))$ and variance $\sigma_k^2 \cdot (\Lambda(t;\theta_k) - \Lambda(s;\theta_k))$. $\Lambda(t;\theta_k)$ is a timescale transformation function, for which the form $\Lambda(t;\theta_k) = t^{\theta_k}$ was adopted in this study, where θ_k is a parameter to be estimated. Values of $\theta_k > 1$ indicate that the trend curve of the deterioration process is concave, characterized by an increasing deterioration rate over time. When $\theta_k = 1$, the trend curve of the deterioration process is linear, and in this case, the model corresponds to the traditional linear Wiener deterioration process. Values of $\theta_k < 1$ indicate that the trend curve of the deterioration process is convex, meaning that the condition gradually deteriorates to a saturation point or an asymptotic level. Consequently, the advantage of using the timescale transformation function in the form of t^{θ_k} is that it does not necessitate the assumption that the deterioration process is linear or nonlinear. Instead, it estimates the value of θ_k using actual condition inspection data of segment k itself. Based on the estimated result for θ_k, the characteristics of the deterioration process can be further analyzed for segment k. This approach is more adaptable and flexible compared to the traditional modeling methods based on the linear Wiener process (i.e., in the case of $\Lambda(t;\theta_k) = t$). It is important to note that a substantial body of relevant literature indicates that the track geometry deterioration between two consecutive tamping operations follows a linear or exponential trend [37]. This suggests that the overall trend of the SDLL is consistently increasing. Therefore, in this study, it was posited that $\theta_k \geq 1$.

2.3. Tamping Maintenance Cycle Prediction

The concept of 'first hitting time' plays a crucial role in degradation analysis as it represents the moment when the system reaches a specified threshold. Building upon this notion, this study defines the tamping maintenance cycle of track segment k as the time interval during which the value of SDLL, $X_k(t)$, deteriorates from the initial state $X_{0,k}$ to the maintenance threshold X_{lim} for the first time:

$$T_k = \inf\{\xi_k | X_k(\xi_k) \geq X_{lim}\} = \inf\{\xi_k | \beta_k \Lambda(\xi_k;\theta_k) + \sigma_k B(\Lambda(\xi_k;\theta_k)) \geq X_{lim} - X_{0,k}\}, \quad (2)$$

where T_k is the tamping maintenance cycle. ξ_k is a random variable that represents the degradation time interval from the initial moment. The operator $\inf\{\cdot\}$ denotes the inferior limit of a variable. As previously mentioned, the increment ΔX_k follows a normal distribution. According to the principles of the Wiener process [41], $\Lambda(\xi_k;\theta_k)$ follows an inverse Gaussian distribution with a location parameter of $(X_{lim} - X_{0,k})/\beta_k$ and a scale parameter of $(X_{lim} - X_{0,k})^2/\sigma_k^2$. Therefore, the probability density function $f_k(\xi_k)$ for the tamping maintenance cycle can be expressed as follows:

$$f_k(\xi_k) = \frac{X_{lim} - X_{0,k}}{\sqrt{2\pi\sigma_k^2[\Lambda(\xi_k;\theta_k)]^3}} \exp^{-\frac{(X_{lim}-X_{0,k}-\beta_k\Lambda(\xi_k;\theta_k))^2}{2\sigma_k^2\Lambda(\xi_k;\theta_k)}} \frac{d\Lambda(\xi_k;\theta_k)}{d\xi_k}, \quad (3)$$

Substituting $\Lambda(\xi_k;\theta_k) = \xi_k^{\theta_k}$ into Equation (3) yields the following:

$$f_k(\xi_k) = \theta_k \frac{X_{lim} - X_{0,k}}{\sqrt{2\pi\sigma_k^2\left(\xi_k^{\theta_k}\right)^3}} \exp^{-\frac{\left(X_{lim} - X_{0,k} - \beta_k\xi_k^{\theta_k}\right)^2}{2\sigma_k^2\xi_k^{\theta_k}}} \xi_k^{\theta_k-1}, \tag{4}$$

In the field of probabilistic modeling for infrastructure and equipment deterioration prediction, the life expectancy calculated from the lifespan probability density function is typically used as the predicted result for the service life [39–41]. Following this approach, given the known probability density function $f_k(\xi_k)$ for the tamping maintenance cycle of track segment k, the predicted value T_k for the tamping maintenance cycle can be calculated using Equation (5):

$$T_k = \int_0^\infty \xi_k \cdot f_k(\xi_k)d\xi_k, \tag{5}$$

3. Estimation of Model Parameters

β_k, σ_k, and θ_k are the parameters to be estimated in the prediction model of the tamping maintenance cycle for track segment k. To ensure the accuracy of the prediction results, it is necessary to utilize historical inspection data and select an effective parameter estimation method to estimate the parameters. Maximum likelihood estimation (MLE) and Markov chain Monte Carlo (MCMC) methods are widely used to estimate unknown parameters in stochastic process modeling. Although some studies believe that MCMC can consider prior information to make parameter estimation results more accurate [13,23,42], there has been no comparison between the results of MLE and MCMC to verify this view. This study employed both MLE and adaptive MCMC methods to estimate the model parameters. Through a case analysis, we compared the accuracy of the estimation results produced by these two methods for the issues addressed in this research. The two parameter estimation methods are described below.

3.1. MLE with Marine Predators Algorithm

Let $\Delta x_{k,m} = x_{k,m} - x_{k,m-1}$ represent the deterioration amount of the SDLL for track segment k between two consecutive inspection times $t_{k,m-1}$ and $t_{k,m}$. In accordance with the properties of the PTT-WP, $\Delta x_{k,m}$ follows a normal distribution with the expectation $E(\Delta x_{k,m})$ and variance $\text{Var}(\Delta x_{k,m})$ given by Equations (6) and (7):

$$E(\Delta x_{k,m}) = \beta_k \cdot \Delta\Lambda(t;\theta_k)_m, \tag{6}$$

$$\text{Var}(\Delta x_{k,m}) = \sigma_k^2 \cdot \Delta\Lambda(t;\theta_k)_m, \tag{7}$$

where $\Delta\Lambda(t;\theta_k)_m = t_{k,m}^{\theta_k} - t_{k,m-1}^{\theta_k}$. Based on the principles of the MLE method, the historical inspection data at the longitudinal level for track segment k can be used to construct the log-likelihood function, given by Equation (8):

$$\ln[L(\boldsymbol{\gamma}_k|\boldsymbol{Data}_k)] = \ln[\prod_{m=2}^{M_k} \frac{1}{\sqrt{2\pi\sigma_k^2\Delta\Lambda(t;\theta_k)_m}} \exp[-\frac{(\Delta x_{k,m} - \beta_k\Delta\Lambda(t;\theta_k)_m)^2}{2\sigma_k^2\Delta\Lambda(t;\theta_k)_m}]], \tag{8}$$

where $\boldsymbol{\gamma}_k = \{\gamma_{k1}, \gamma_{k2}, \gamma_{k3}\} = \{\beta_k, \theta_k, \sigma_k\}$, is the vector composed of the model parameters to be estimated, $\boldsymbol{Data}_k = \{(t_{k,m}, x_{k,m})| m = 1, 2, \ldots, M_k\}$ is the collected dataset of the total of M_k inspections of track segment k, $t_{k,m}$ is the inspection date of inspection sample m, and $x_{k,m}$ is the SDLL value obtained for inspection sample m.

According to the theory of MLE, the values of the parameters β_k, σ_k, and θ_k that maximize the log-likelihood function, as shown in Equation (8), are the estimated results, denoted as $\hat{\beta}_k$, $\hat{\sigma}_k$, and $\hat{\theta}_k$, respectively. Therefore, this parameter estimation problem is an optimization challenge that involves determining the optimal solution in Equations (9) and (10):

$$\max \ln[L(\boldsymbol{\gamma}_k|\boldsymbol{Data}_k)], \tag{9}$$

$$\text{s.t. } \beta_k > 0, \sigma_k > 0, \theta_k \geq 1, \tag{10}$$

Considering the intricate and multimodal characteristics of the objective function as well as the nonlinear and multidimensional characteristics of the solution space, gradient-based optimization methods are susceptible to getting trapped in local optima. Consequently, in this study, we employed a high-performance heuristic global optimization solver known as the marine predators algorithm (MPA). This algorithm was proposed by Faramarzi et al. [43]. The performance of the MPA in 29 test functions and several practical engineering cases was evaluated in their study. The proposed MPA performed better than the genetic algorithm, particle swarm optimization, gravity search algorithm, and covariance matrix adaptive evolutionary strategy algorithm [43]. Algorithm 1 presents the pseudocode of MPA used to solve Equations (9) and (10).

Algorithm 1. Pseudocode of MPA

Initialization: n—population size, *max_iter*—maximum number of iterations, *lb*—lower bound of the parameter value range, *ub*—upper bound of the parameter value range, set the number of iterations $Iter = 0$, initial population $\boldsymbol{Prey}_k = \left\{ \boldsymbol{\gamma}_k^i, i = 1, \ldots, n \right\}$

Iterative process:

While *Iter < max_iter*,

    1. Calculate the fitness of the population based on Equation (8), i.e.,

        $\ln[L\left(\boldsymbol{\gamma}_k^i | \boldsymbol{Data}_k \right)$, construct the Elite matrix $\boldsymbol{Elite}_k$, and accomplish memory saving.

    2. Update the population $\boldsymbol{Prey}_k$:

        If *Iter < max_iter/3*

            $\mathbf{s}_i = \boldsymbol{R}_B \otimes \left(\boldsymbol{Elite}_k^i - \boldsymbol{R}_B \otimes \boldsymbol{\gamma}_k^i \right), \boldsymbol{\gamma}_k^i = \boldsymbol{\gamma}_k^i + P * \boldsymbol{R} \otimes \mathbf{s}_i, i = 1, \ldots, n$

        Or else if *max_iter/3 < Iter < 2 * max_iter/3*

            Divide the population into two parts: prey and predators.

            The prey perform Lévy flights and updates based on:

            $\mathbf{s}_i = \boldsymbol{R}_L \otimes \left(\boldsymbol{Elite}_k^i - \boldsymbol{R}_L \otimes \boldsymbol{\gamma}_k^i \right), \boldsymbol{\gamma}_k^i = \boldsymbol{\gamma}_k^i + P * \boldsymbol{R} \otimes \mathbf{s}_i, i = 1, \ldots, n/2$

            The predators perform Brownian motion and updates based on:

        $\mathbf{s}_i = \boldsymbol{R}_B \otimes \left(\boldsymbol{Elite}_k^i - \boldsymbol{R}_B \otimes \boldsymbol{\gamma}_k^i \right), \boldsymbol{\gamma}_k^i = \boldsymbol{Elite}_k^i + P \cdot CF \otimes \mathbf{s}_i, i = \frac{n}{2} + 1, \ldots, n$

        Or else if *Iter > 2 * max_iter/3*

            $\mathbf{s}_i = \boldsymbol{R}_L \otimes \left(\boldsymbol{Elite}_k^i - \boldsymbol{R}_L \otimes \boldsymbol{\gamma}_k^i \right), \boldsymbol{\gamma}_k^i = \boldsymbol{Elite}_k^i + P * CF \otimes \mathbf{s}_i, i = 1, \ldots, n$

        End (if)

    3. Accomplish memory saving and $\boldsymbol{Elite}_k$ update

    4. Applying Fish Aggregating Devices (FADs) effect and update $\boldsymbol{Prey}_k$ based on:

$$\boldsymbol{\gamma}_k^i = \begin{cases} \boldsymbol{\gamma}_k^i + CF[\boldsymbol{lb} + \boldsymbol{R} \otimes (\boldsymbol{ub} - \boldsymbol{lb})] \otimes \boldsymbol{U}, r \leq FADs \\ \boldsymbol{\gamma}_k^i + [FADs(1-r) + r]\left(\boldsymbol{\gamma}_k^{r1} - \boldsymbol{\gamma}_k^{r2} \right), r > FADs \end{cases}$$

End While

In the pseudocode mentioned above, $P, \boldsymbol{R}, \boldsymbol{R}_B, \boldsymbol{R}_L, CF, FADs, \boldsymbol{U}, r, r_1$, and r_2 are the algorithm parameters. Their structure and value-assigning methods were based on recommendations in the literature [43]. Appendix A provides supplementary explanations for some key calculation details in the above pseudocode. Moreover, to test the effectiveness of the MPA in solving the problem presented in this study, a comparative analysis will be conducted in the case study. The performance of the MPA will be compared to two other optimization techniques: the genetic algorithm (GA), which is a widely-used heuristic algorithm, and the gradient descent algorithm (GD), which is a popular gradient-based method.

3.2. Adaptive MCMC

As outlined above, MCMC is a quintessential parameter estimation method within a Bayesian framework. The fundamental concept involves constructing a Markov chain using sampling techniques to ensure that the steady-state distribution aligns with the posterior distribution of the target parameter. Adaptive MCMC enhances this process by dynamically fine-tuning the sampling parameters, thereby boosting sampling efficiency [44].

First, determine the posterior distribution function of the parameters to be estimated. According to the Bayesian criterion, given the observed data $\boldsymbol{Data}_k = \{(t_{k,m}, x_{k,m})\}$, the posterior distribution $p(\boldsymbol{\gamma}_k | \boldsymbol{Data}_k)$ of the parameters to be estimated $\boldsymbol{\gamma}_k = \{\beta_k, \theta_k, \sigma_k\}$ is proportional to the product of the prior distribution $p(\boldsymbol{\gamma}_k)$ and the likelihood function $p(\boldsymbol{Data}_k | \boldsymbol{\gamma}_k)$, as expressed in Equation (11):

$$p(\boldsymbol{\gamma}_k | \boldsymbol{Data}_k) = \frac{p(\boldsymbol{Data}_k | \boldsymbol{\gamma}_k) \cdot p(\boldsymbol{\gamma}_k)}{p(\boldsymbol{Data}_k)} \propto p(\boldsymbol{Data}_k | \boldsymbol{\gamma}_k) \cdot p(\boldsymbol{\gamma}_k), \tag{11}$$

where the prior distribution $p(\boldsymbol{\gamma}_k)$ in this study is a multivariate truncated normal distribution with a mean of $\boldsymbol{\mu} = (\mu_\beta, \mu_\sigma, \mu_\vartheta)$, a covariance matrix of $\boldsymbol{C}_{3\times3}$, and truncation interval $\boldsymbol{LU} = \{\beta_k > 0, \sigma_k > 0, \theta_k \geq 1\}$. This study used the MLE-based results of the parameters for the deterioration models of all segments. The mean and covariance of these results will respectively serve as the $\boldsymbol{\mu}$ and $\boldsymbol{C}_{3\times3}$ for the prior distribution. According to Equation (8), the likelihood function is as follows:

$$p(\boldsymbol{Data}_k | \boldsymbol{\gamma}_k) = \prod_{m=2}^{M_k} \frac{1}{\sqrt{2\pi\sigma_k^2 \Delta\Lambda(t;\theta_k)_m}} \exp\left[-\frac{(\Delta x_{k,m} - \beta_k \Delta\Lambda(t;\theta_k)_m)^2}{2\sigma_k^2 \Delta\Lambda(t;\theta_k)_m}\right], \tag{12}$$

Secondly, the sampling method and adaptive adjustment rules were determined. Considering the complexity of sampling from the posterior distribution and the correlation between parameters, this study adopts the more flexible Metropolis–Hastings algorithm for the sampling process. For the adaptive adjustment rules, based on the discussion in the literature [44], the study employs the optimization criterion of aiming to achieve an optimal acceptance rate range of $0.2 < \alpha^* < 0.3$ to continuously adjust the covariance matrix of the proposal distribution during the burn-in period. The adjustment coefficient λ is determined by conducting parameter estimation on randomly selected track segments and comparing the convergence under different coefficient values. The length of the burn-in period was set to 50% of the total number of samples. Samples generated during the burn-in period were not used for subsequent statistical analyses [44]. In addition, considering that the input of the initial samples has a significant impact on the convergence speed of the sampling process and the accuracy of the parameter estimation results, the sampling algorithm was executed using the MLE-based results as the initial values for the parameters to be estimated. Algorithm 2 presents the pseudocode for the adaptive MCMC algorithm used to estimate the model parameter $\boldsymbol{\gamma}_k$.

Algorithm 2. Pseudocode of Adaptive MCMC

Initialization: give the proposal distribution $p(\boldsymbol{\gamma}'|\boldsymbol{\gamma}_{n-1})$, the initial covariance matrix C_0, the initial value of the parameter $\boldsymbol{\gamma}_0$, the total number of samples $N = 20,000$, the length of the burn-in period N_b, the lower bound of parameter values lb, the lower bound of parameter values ub, and the adjustment coefficient $\lambda = 0.05$.

Iterative process:

FOR $n = 1, \ldots, N$:

    1. Draw a candidate sample $\boldsymbol{\gamma}'$ from the proposal distribution and generate a random number from a uniform distribution $u \sim U([0,1))$

    2. Calculate the acceptance probability:

$$\alpha_n = \min\left[1, \frac{p(\boldsymbol{\gamma}'|\boldsymbol{Data}_k)}{p(\boldsymbol{\gamma}_{n-1}|\boldsymbol{Data}_k)}\right]$$

    3. Decide whether to accept sample $\boldsymbol{\gamma}'$:
        If $u \leq \alpha_n$:

$$\boldsymbol{\gamma}_n = \boldsymbol{\gamma}'$$

        Otherwise:

$$\boldsymbol{\gamma}_n = \boldsymbol{\gamma}_{n-1}$$

        End (if)

    4. Update the covariance matrix when $n \leq N_b$
        If $\alpha_n < 0.2$: $C_n = (1 + \lambda) * C_{n-1}$
        Otherwise if $\alpha_n > 0.3$: $C_n = (1 - \lambda) * C_{n-1}$
        Otherwise: $C_n = C_{n-1}$
        End (if)

End FOR

Calculate the parameter estimates: $\hat{\boldsymbol{\gamma}} = \frac{1}{N-N_b}\sum_{n=N_b+1}^{N} \boldsymbol{\gamma}_n$

---

4. Case Study

4.1. Data Description

The Lanxin railway line is an electrified normal-speed railway from Lanzhou to Sinkiang with a maximum permissible speed of 160 km/h and a total length of 2423 km. It is an important part of the railway network in Northwest China. Based on dividing the line into multiple track segments with a unit length of 200 m, 2171 track segments of the downlink line within the jurisdiction of the Jiayuguan Section of Lanzhou Railway Bureau (mileage range K548+000 to K982+400) were selected and their tamping maintenance cycle analyzed using the proposed method, as shown in Figure 3. The track geometry of the Lanxin Railway is inspected at least twice a month using a track-inspection vehicle. A fixed-period-based preventive maintenance strategy was employed for the tamping activities, with the interval typically set to 1–2 years. We collected track longitudinal-level inspection data with 108,608 records from April 2015 to November 2018 for the selected 2171 track segments and tamping maintenance records. By associating the longitudinal-level inspection data with the tamping maintenance record data based on mileage and time information, the longitudinal-level inspection dataset of each track segment can be divided into multiple data subsets; that is, the inspection data within one tamping maintenance cycle is a data subset. The subset with the largest sample size from the data was used for the model construction.
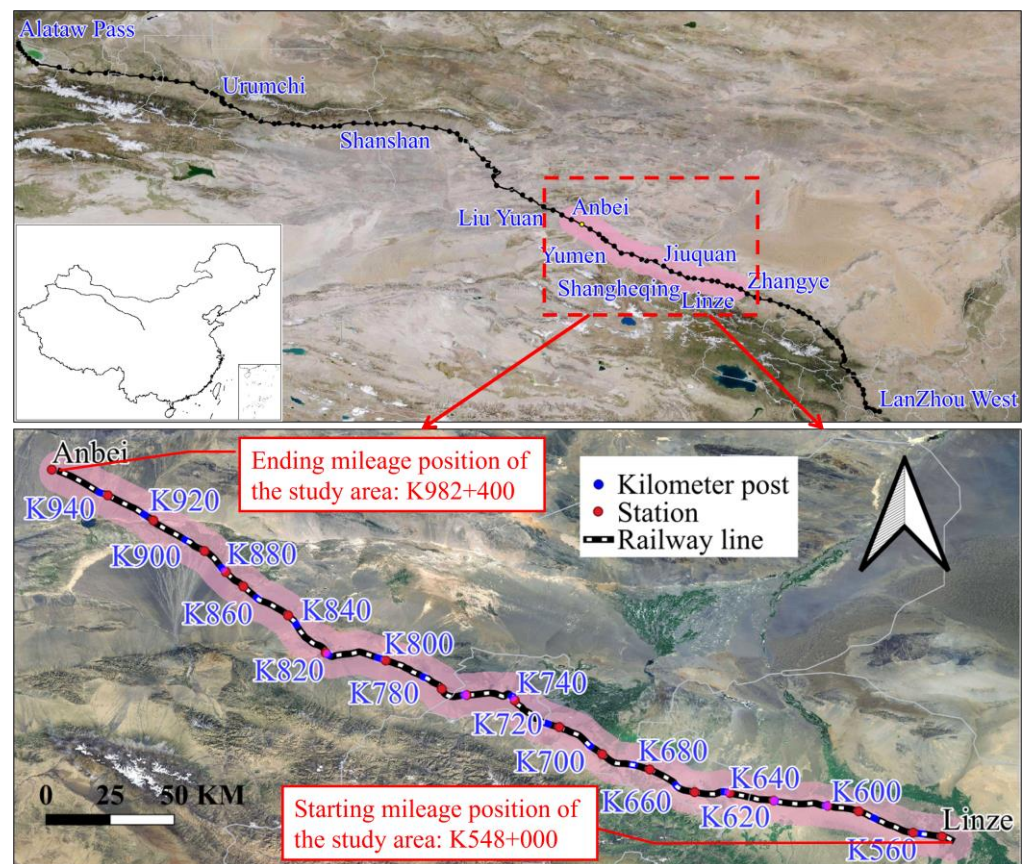
Figure 3. Study area along the Lanxin railway line in northwest China.

4.2. Analysis of the Prediction Results

This study employed four parameter estimation methods: maximum likelihood estimation with a marine predators algorithm (MLE-MPA), maximum likelihood estimation with a genetic algorithm (MLE-GA), maximum likelihood estimation with a gradient descent algorithm (MLE-GD), and adaptive Markov chain Monte Carlo (AMCMC). The findings from these estimation methods were integrated into the tamping-cycle prediction model developed in this study to forecast the time intervals until the SDLL degrades to any given value X_{lim}. The validity of the prediction model and the four parameter estimation methods was analyzed based on the statistical results for the prediction errors.

4.2.1. Results of Parameter Estimations and Tamping Cycle Prediction

The estimation results for β_k, σ_k, and θ_k for the tamping cycle prediction models for the 2171 track segments are shown in Figure 4a–c, respectively. The different line colors in the figure represent the results obtained using different parameter estimation methods. Taking X_{lim} = 3 mm as an example, the tamping maintenance cycle of each track segment was predicted using Equations (4) and (5), as shown in Figure 4d.

As Figure 4 shows, regardless of the parameter estimation method, the parameters of the longitudinal-level degradation model for each track segment were heterogeneous, and the prediction results of the tamping maintenance cycle were heterogeneous, reflecting the heterogeneity of the deterioration process of the track geometry at different spatial locations. In addition, although the estimation results obtained with the four parameter estimation methods were not significantly different, the estimation results for the tamping maintenance cycle were significantly different, as shown in the marked examples of several representative track segments in Figure 4e. This illustrates the necessity of selecting an accurate and effective parameter estimation method. The next section presents a comparison

of the prediction errors of the four methods conducted to identify the parameter estimation method with the highest accuracy.
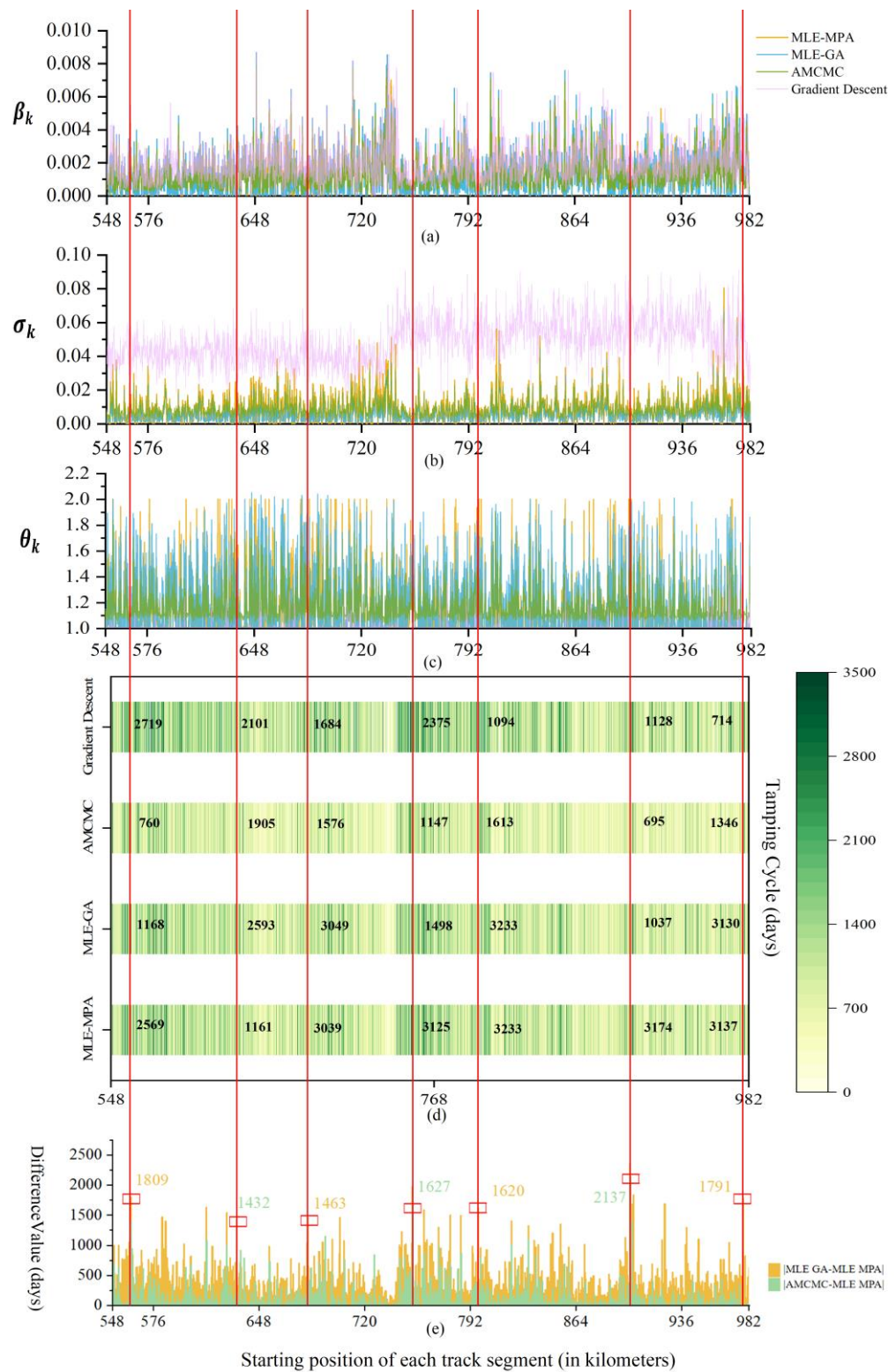

Figure 4. Results of parameter estimation and tamping maintenance cycle prediction for the 2171 track segments in the downlink line of the Lanxin railway: (**a**) estimation results for parameter β_k, (**b**) estimation results for parameter σ_k, (**c**) estimation results for parameter θ_k, (**d**) tamping maintenance cycle prediction results, and (**e**) difference in the prediction results between different methods.

4.2.2. Prediction Error Analysis

As the tamping maintenance strategy of combining "periodic preventive maintenance" with "local corrective maintenance" is still widely adopted in the Chinese railway system, most segments carried out preventive tamping maintenance operations before the SDLL degraded to the maintenance threshold (3–3.6 mm). As a result, the actual time interval of each segment during which the SDLL deteriorates from its initial value to the maintenance threshold cannot be extracted from the historical inspection data at the longitudinal level, resulting in the inability to support a prediction error analysis of the tamping maintenance cycle. The absence of benchmark data hinders prediction error analysis of tamping maintenance cycles. The tamping cycle prediction model proposed in this study essentially serves as a deterioration prediction model for the SDLL between two consecutive tamping operations. It can predict the time interval required for the SDLL to deteriorate to any specified threshold value X_{lim}. Given the situation described above, we set SDLL threshold values (i.e., X_{lim}) between 1 and 3 mm at intervals of 0.1 mm and then predicted the time required for degradation to the different given X_{lim} values to indirectly verify the predictive accuracy of the model by comparing the predicted results with the actual time intervals recorded in the field SDLL historical inspection data.

Figures 5–8 show box plots of the prediction errors corresponding to the parameter estimation results for MLE-MPA, MLE-GA, MLE-GD, and adaptive MCMC, respectively. The horizontal coordinate is the given X_{lim}. For each X_{lim} value, a box plot was generated to display the distribution characteristics of the prediction error for the time interval required for the SDLL to deteriorate to that value. In these box plots, the box itself represents the interquartile range, with the central line indicating the median. The whiskers extend to the most extreme data points that fall within 1.5 times the interquartile range. Each figure also includes two additional reference lines: a yellow dashed line representing the Inner Limit, which is the maximum value of the upper whiskers across all X_{lim} values, and a green dashed line showing the Maximum Error across all X_{lim} values. At the top of each figure, the average Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are displayed, providing overall performance metrics for each method across all scenarios. Figure 9 shows a comparison of the mean absolute error (MAE) corresponding to the three estimation methods. Table 1 shows the statistical values of the prediction errors corresponding to the four parameter estimation methods, including the MAE, the percentage of absolute error within one month ($\leq$30 days), the percentage of absolute error within two months ($\leq$60 days), and the percentage of absolute error within three months ($\leq$90 days). It should be noted that the percentages here represent cumulative frequencies, indicating that the count of samples with an absolute error of less than two months includes those with an absolute error of less than 30 days. Figure 10 visually illustrates the distribution of error ranges for different methods across various scenarios through stacked bars, providing a graphical representation of the data from Table 1.

The errors in the prediction results obtained using the four parameter estimation methods were compared and analyzed. Based on the prediction results for the parameter estimation method with the least error, the validity of the model established in this study is illustrated in combination with the tamping maintenance planning demand.

- Comparative Analysis of Four Parameter Estimation Methods

Based on the mean absolute errors (MAE) presented in Table 1, it is observed that for various X_{lim} values, the maximum MAE value (21 days) of the prediction results obtained from the MLE-MPA method exhibits reductions of 4.5%, 33.3%, and 48.8% compared to those of the MLE-GA, MLE-GD, and AMCMC methods, respectively. Furthermore, the mean MAE value (19 days) achieved by the MLE-MPA method demonstrates decreases below 5%, 26.9%, and 44.1% when contrasted with those of the MLE-GA, MLE-GD, and AMCMC methods. Additionally, Figure 9 illustrates that for each given X_{lim} value, the MAE associated with predictions using the MLE method consistently outperforms that of the adaptive MCMC method.
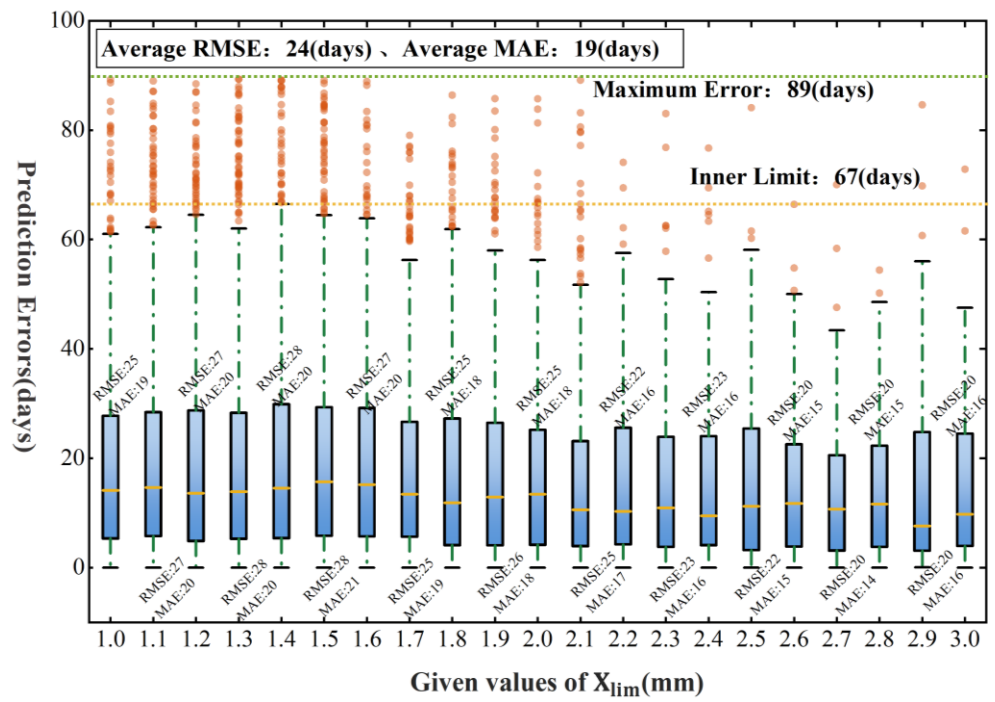
Figure 5. Absolute error distribution of prediction results obtained by MLE-MPA method for the time interval required for SDLL to deteriorate to the different given X_{lim} values.
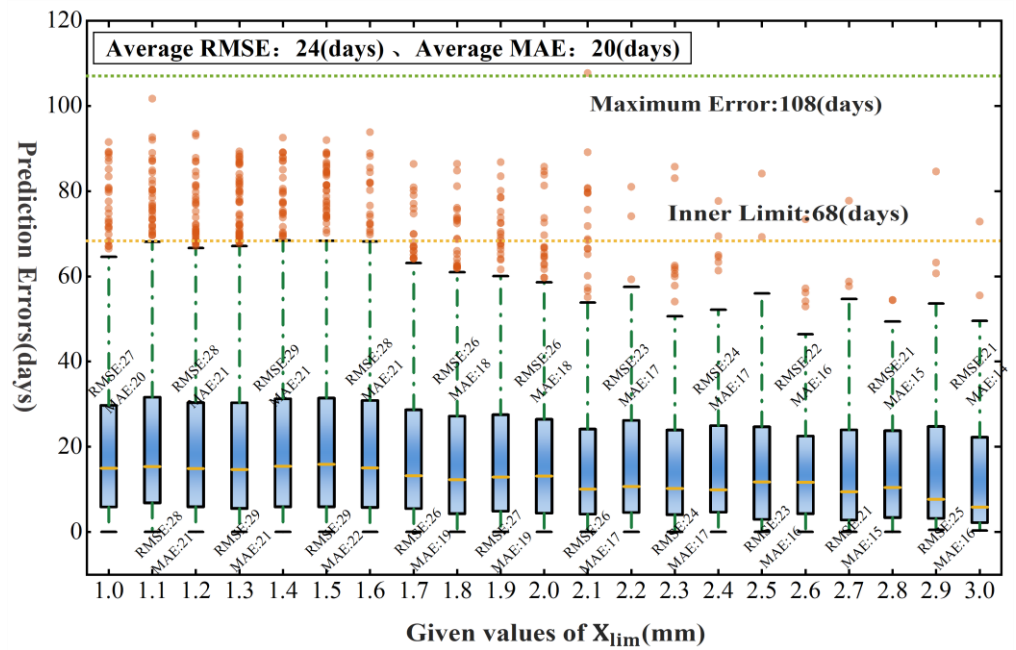

Figure 6. Absolute error distribution of prediction results obtained by MLE-GA method for the time interval required for SDLL to deteriorate to the different given X_{lim} values.
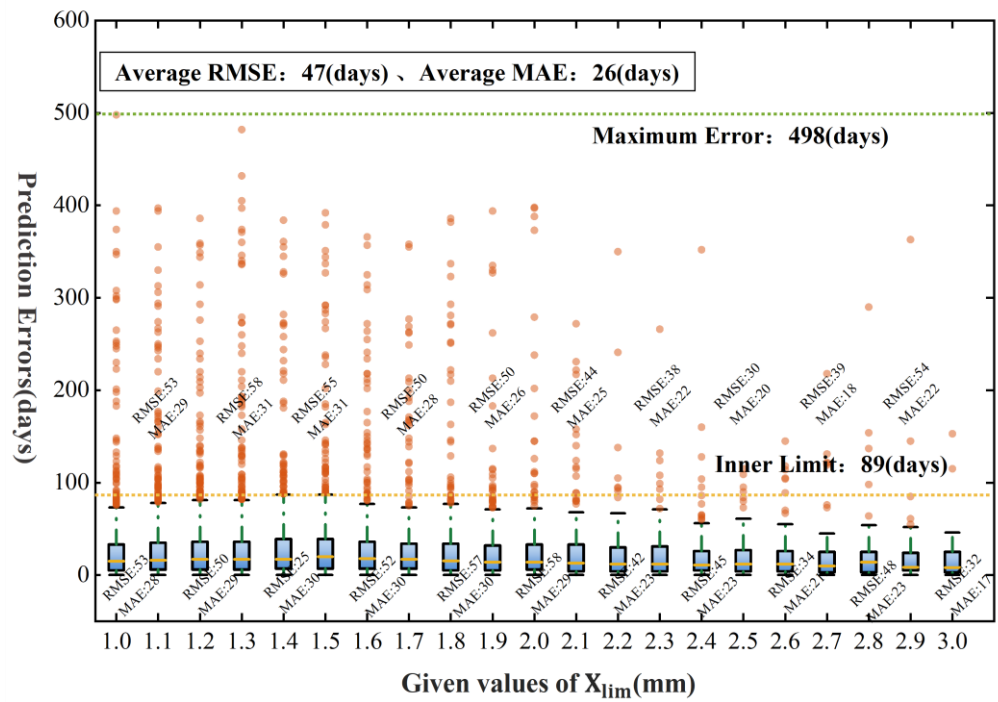
Figure 7. Absolute error distribution of prediction results obtained by MLE-GD method for the time interval required for SDLL to deteriorate to the different given X_{lim} values.
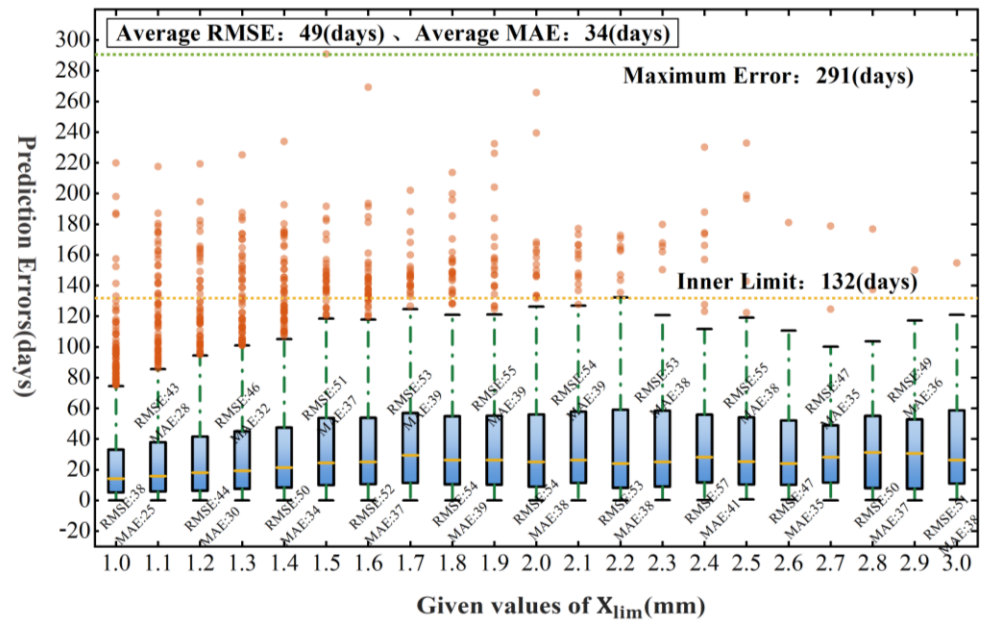

Figure 8. Absolute error distribution of prediction results obtained by adaptive MCMC method for the time interval required for SDLL to deteriorate to the different given X_{lim} values.
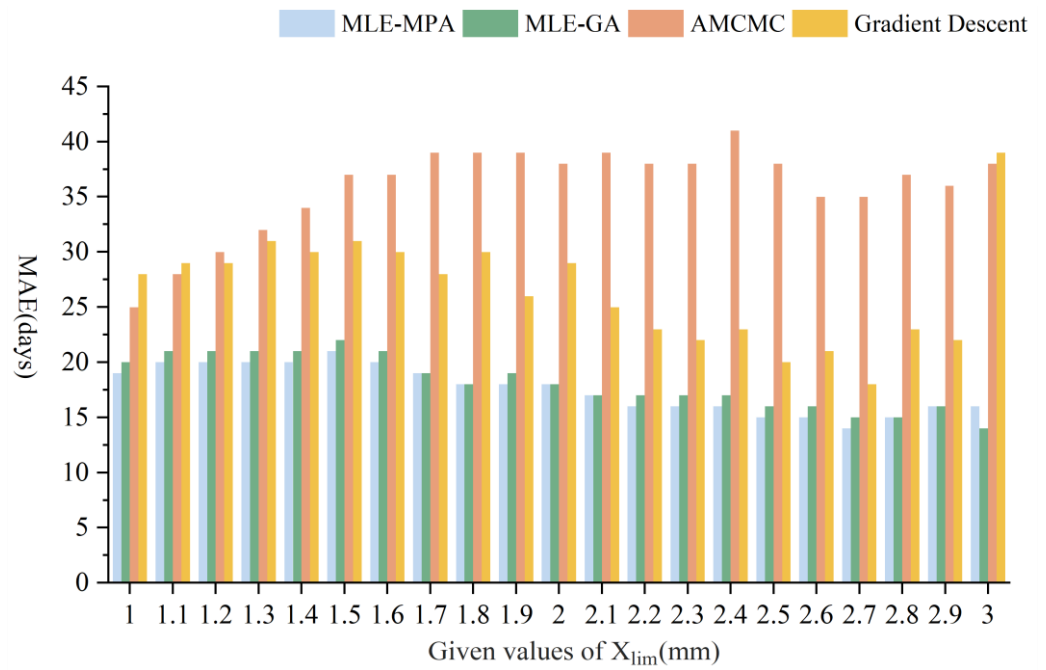
Figure 9. Comparison of prediction errors of four methods for the time interval required for SDLL to deteriorate to the different given X_{lim} values.
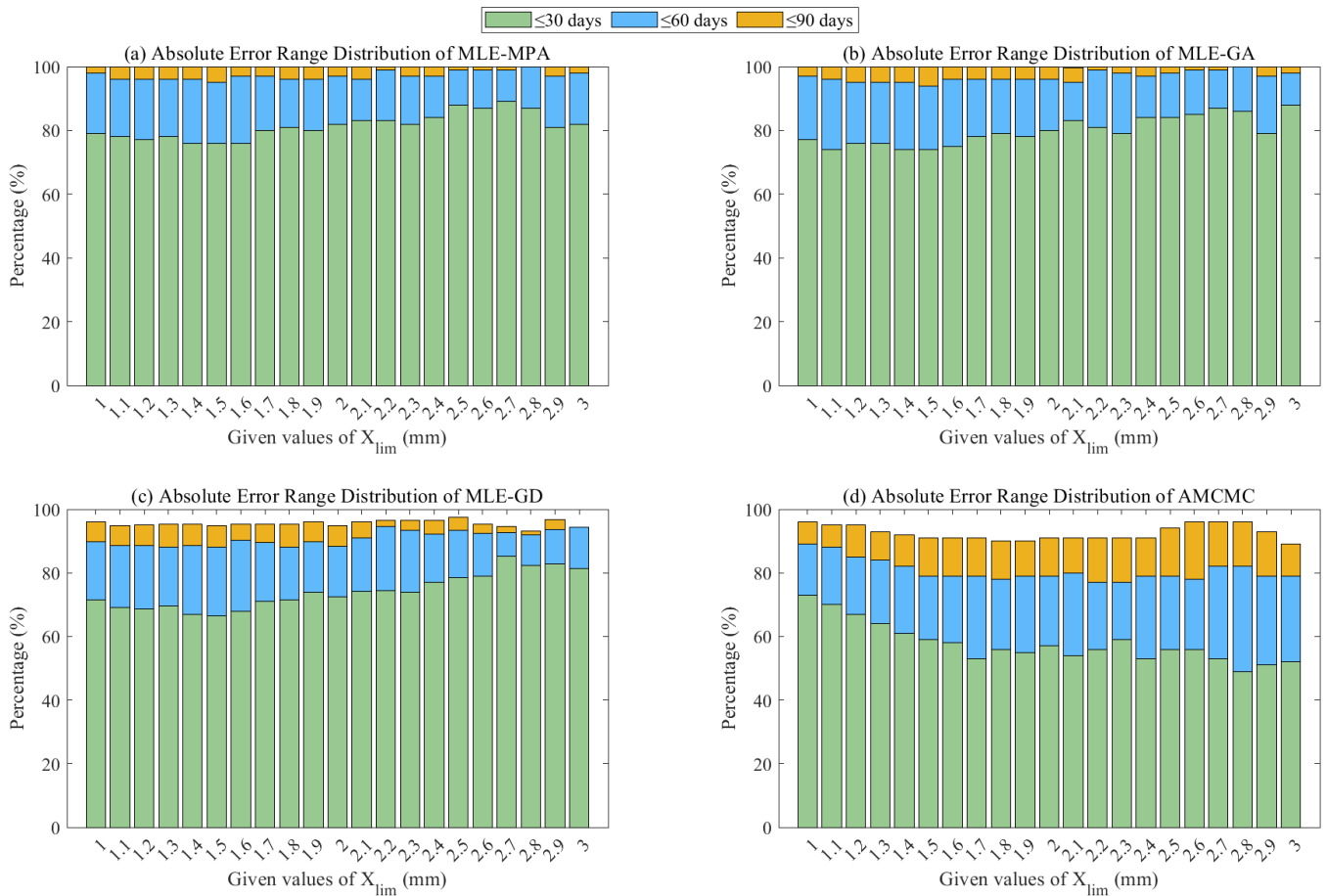

Figure 10. Stacked bar chart of the distribution of different absolute error ranges for prediction results obtained by the four methods.

Table 1. Prediction error results obtained by the four methods for the time interval required for SDLL to deteriorate to the different specified X_{lim} values.

X_{lim} (mm)	Sample Size	MAE (Days)				Percentage of Absolute Error ≤ 30 Days				Percentage of Absolute Error ≤ 60 Days				Percentage of Absolute Error ≤ 90 Days			
		MPA	GA	GD	AMCMC	MPA	GA	GD	AMCMC	MPA	GA	GD	AMCMC	MPA	GA	GD	AMCMC
1	1002	19	20	28	25	79%	77%	72%	73%	98%	97%	90%	89%	100%	100%	96%	96%
1.1	1097	20	21	29	28	78%	74%	69%	70%	96%	96%	89%	88%	100%	100%	95%	95%
1.2	1155	20	21	29	30	77%	76%	69%	67%	96%	95%	88%	85%	100%	100%	95%	95%
1.3	1135	20	21	31	32	78%	76%	69%	64%	96%	95%	88%	84%	100%	100%	95%	93%
1.4	1031	20	21	30	34	76%	74%	67%	61%	96%	95%	89%	82%	100%	100%	95%	92%
1.5	931	21	22	31	37	76%	74%	67%	59%	95%	94%	88%	79%	100%	100%	95%	91%
1.6	805	20	21	30	37	76%	75%	68%	58%	97%	96%	90%	79%	100%	100%	95%	91%
1.7	666	19	19	28	39	80%	78%	71%	53%	97%	96%	90%	79%	100%	100%	95%	91%
1.8	557	18	18	30	39	81%	79%	72%	56%	96%	96%	88%	78%	100%	100%	95%	90%
1.9	454	18	19	26	39	80%	78%	74%	55%	96%	96%	90%	79%	100%	100%	96%	90%
2	374	18	18	29	38	82%	80%	72%	57%	97%	96%	88%	79%	100%	100%	95%	91%
2.1	303	17	17	25	39	83%	83%	74%	54%	96%	95%	91%	80%	100%	100%	96%	91%
2.2	223	16	17	23	38	83%	81%	75%	56%	99%	99%	95%	77%	100%	100%	97%	91%
2.3	182	16	17	22	38	82%	79%	74%	59%	97%	98%	93%	77%	100%	100%	96%	91%
2.4	150	16	17	23	41	84%	84%	77%	53%	97%	97%	92%	79%	100%	100%	96%	91%
2.5	130	15	16	20	38	88%	84%	79%	56%	99%	98%	93%	79%	100%	100%	98%	94%
2.6	110	15	16	21	35	87%	85%	79%	56%	99%	99%	92%	78%	100%	100%	95%	96%
2.7	98	14	15	18	35	89%	87%	85%	53%	99%	99%	93%	82%	100%	100%	95%	96%
2.8	76	15	15	23	37	87%	86%	82%	49%	100%	100%	92%	82%	100%	100%	93%	96%
2.9	67	16	16	22	36	81%	79%	83%	51%	97%	97%	94%	79%	100%	100%	97%	93%
3	56	16	14	17	38	82%	88%	81%	52%	98%	98%	94%	79%	100%	100%	94%	89%
Overall results	10,602	19	20	26	34	79%	77%	74%	62%	96%	96%	91%	82%	100%	99.95%	96%	92.72%

In addition to examining MAE values, an analysis of prediction stability under different scenarios can be conducted through root mean square error (RMSE). As depicted in Figures 5–8, both MLE-GD and AMCMC methods exhibit notably higher average RMSE values, which not only indicate their inferior performance in terms of prediction errors as previously demonstrated but also reveal their instability and sensitivity to varying scenarios. This behavior may be attributed to susceptibility of the gradient descent method towards local optima when dealing with non-convex and complex objective functions as well as heavy dependence on initial distributions exhibited by the AMCMC method leading to significant discrepancies in prediction outcomes across different dataset scenarios.

Regarding the error distribution characteristics outlined in Table 1 and Figure 10, the prediction results generated by the MLE-MPA method demonstrate that all absolute errors fall within three months. A substantial proportion (79%) occurs within a span of just 30 days, while only a mere four percent exceed 60 days. The proportions of absolute errors within one and two months were both higher than those of the other three methods.

Furthermore, from Figure 9, we observe that the absolute errors do not significantly increase alongside rising X_{lim} values, contrary to general expectations. This is mainly because a larger X_{lim} value during a tamping cycle spans a longer time and provides more inspection data, resulting in an abundance of effective data for constructing the likelihood function of each track segment. As a result, prediction accuracy improves as reflected in the trend of prediction errors with the X_{lim} value for the MLE-GA and MLE-MPA methods.

It is evident that the estimation results of the MLE methods are superior to those of the adaptive MCMC method. Furthermore, in terms of computational resource requirements and solving speed, using adaptive MCMC to solve the model parameters for the 2171 segments in this case took nearly 26 h, while MLE-MPA, MLE-GA, and MLE-GD

methods took about 2 h each. MLE methods are more suitable for the prediction needs of line-level and network-level tamping maintenance cycles, especially when dynamically updating model parameter estimates such as inspection data accumulate.

Among the MLE methods, the MLE-MPA performs better than the MLE-GA and MLE-GD, confirming the general conclusions pointed out by Faramarzi et al. [43].

In addition to analyzing the prediction errors, we further conducted a comparative analysis of the uncertainty in the parameter estimation results by calculating standard deviations that reflect the confidence level of the estimators for the four methods.

For the MLE framework, we utilized the gradient-based method to calculate the standard deviation of the parameter estimates for these three methods: MLE-MPA, MLE-GA, and MLE-GD. Specifically, we approximated the covariance of the parameter estimates by computing the Hessian matrix. Subsequently, we depicted the standard deviation bars of the parameters β_k, σ_k, and θ_k for the 2171 track segments in Figures 11–13, respectively. In each figure, subplots (a), (b), and (c) correspond to the results of the MLE-MPA, MLE-GA, and MLE-GD methods, respectively. The results reveal inconsistencies in uncertainty across different methods, which underscores the importance of employing robust global optimization algorithms to avoid local optima traps.

For the adaptive MCMC method, we randomly selected a track segment ($k = 1986$, starting mileage K945+200) to record the sampling results of the posterior distribution of the parameters β_k, σ_k, and θ_k, as shown in Figure 14. The uncertainty of the parameters for this track segment is high, possibly due to its location at a slope change point of the railway line, leading to changes in track dynamics forces. Additionally, inherent track irregularities and vehicle system dynamics contribute to increased uncertainty in track geometry degradation. Interestingly, we observed that the adaptive MCMC method yields smaller prediction errors for this specific track segment with high uncertainty. Future research could explore the relationship between parameter uncertainty and factors such as track environment and structure to propose a more robust and widely applicable parameter estimation method.
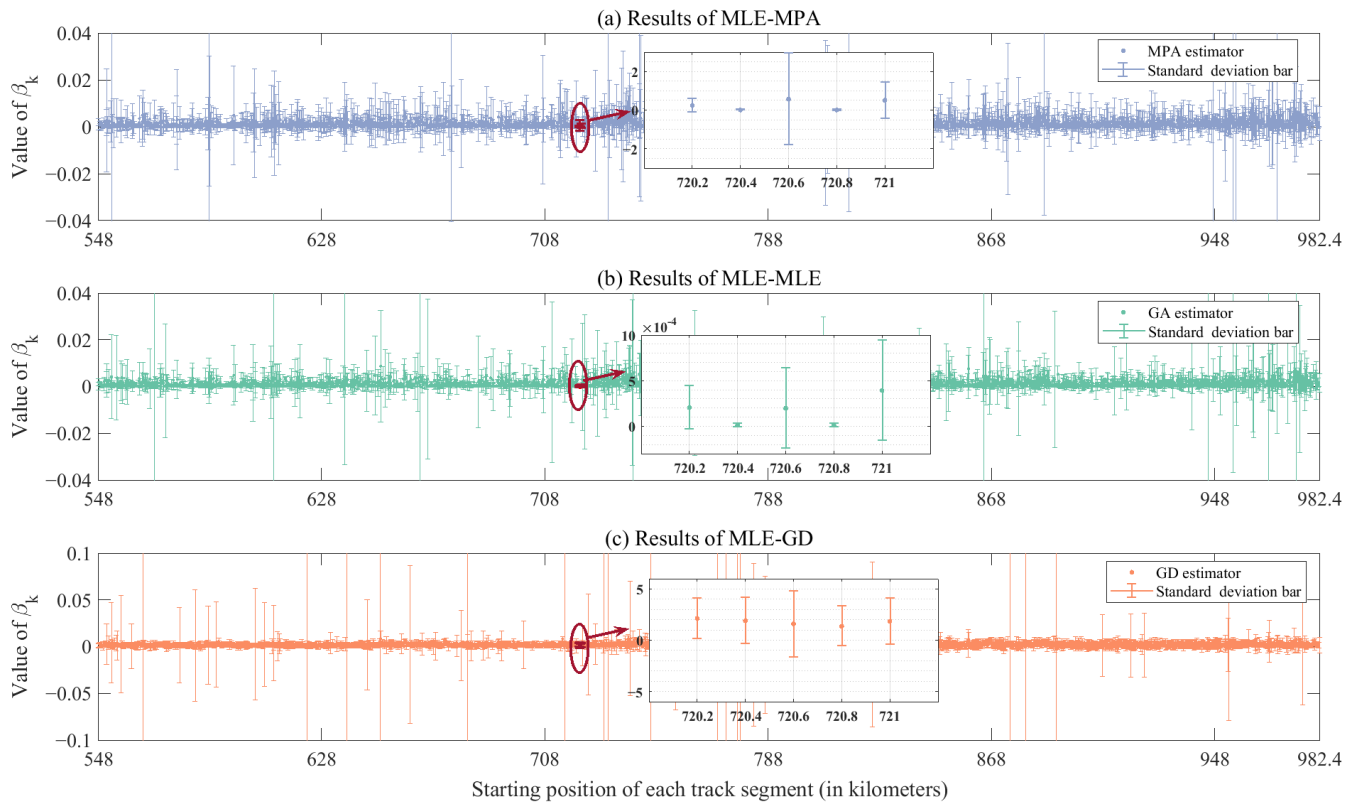

Figure 11. Uncertainty analysis of the estimated parameter β_k obtained by the three parameter estimation methods under the MLE framework.
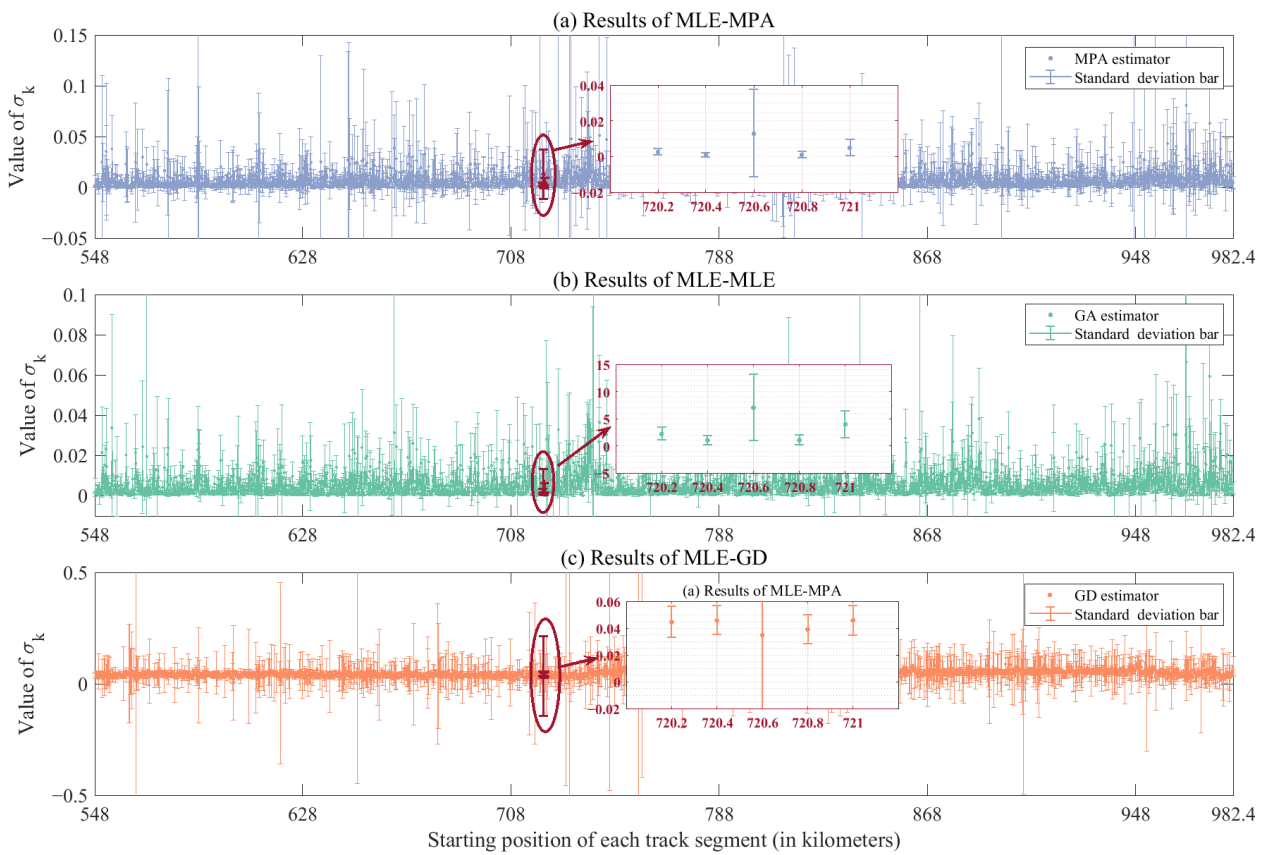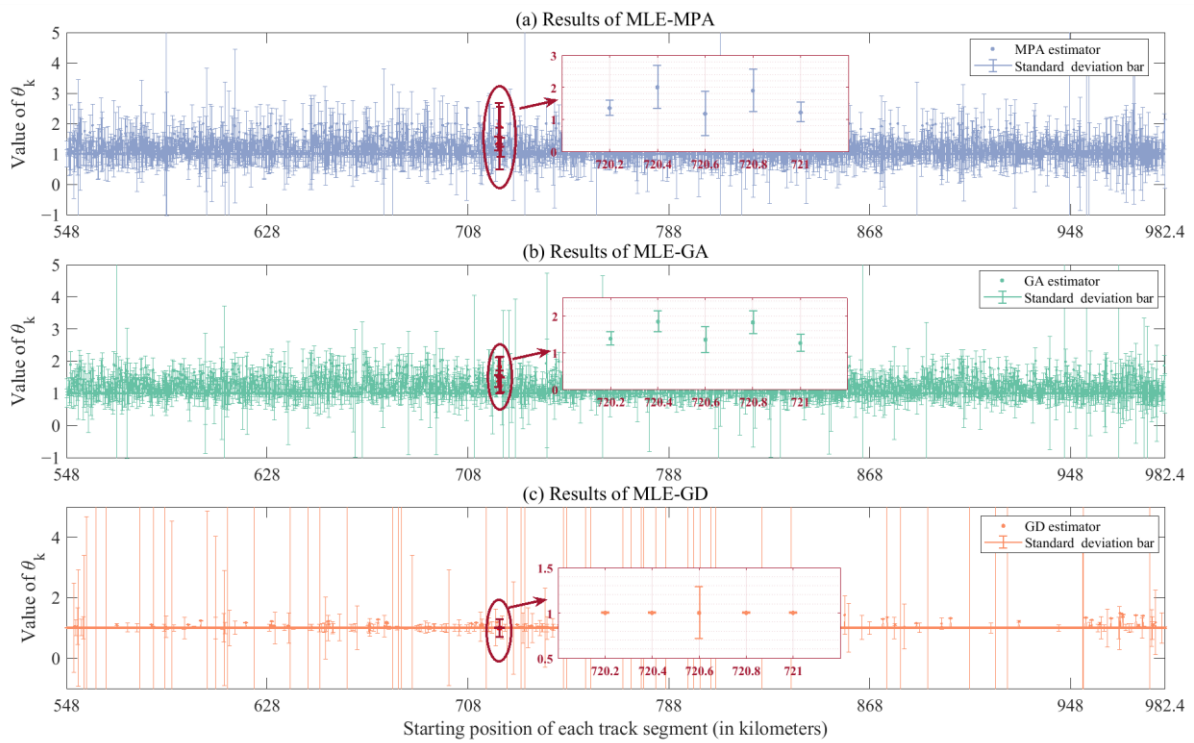
Figure 12. Uncertainty analysis of the estimated parameter σ_k obtained by the three parameter estimation methods under the MLE framework.

Figure 13. Uncertainty analysis of the estimated parameter θ_k obtained by the three parameter estimation methods under the MLE framework.
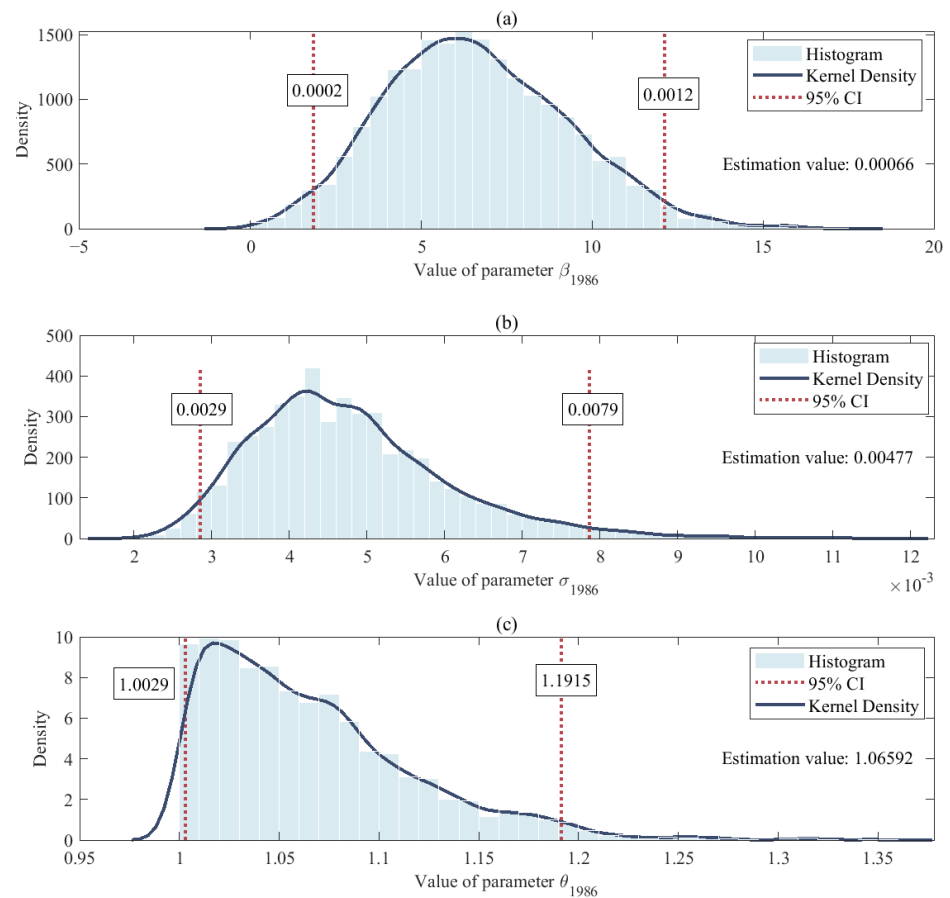
Figure 14. Posterior distribution of parameters for the track segment $k = 1986$ via adaptive MCMC sampling: (**a**) parameter β_k, (**b**) parameter σ_k, and (**c**) parameter θ_k.

- Model Validity Analysis

Based on the preceding analysis, we employed the prediction results obtained using the MLE-MPA parameter estimation method to assess the validity of the model and determine whether it met the requirements for tamping maintenance planning. We calculated the coefficient of determination R^2 of the model to be 0.98. The 80th, 85th, 90th, and 95th percentiles of the absolute error were 32, 38, 45, and 57 days, respectively.

Regarding the performance of the model, the overall prediction accuracy for the tamping maintenance cycle of the 2171 track sections was robust, as indicated by the average MAE and R^2 values.

Based on the maximum absolute error value, the overall error of the tamping maintenance cycle prediction model for the 2171 track segments was controlled within three months. Combining the MAE with the percentiles of absolute error, it is evident that the prediction error for most sections is far less than the maximum error of 89 days. For long-term (1–5 years) tamping maintenance planning of heavy machinery, the planning cycle is typically divided into multiple fixed decision-making periods, during which the segments requiring tamping maintenance in each period are determined. The length of these decision-making periods is generally set to three or six months [31], allowing tamping maintenance operations to be scheduled within the available maintenance windows in each decision-making period. This provides flexibility for opportunistic maintenance strategies and reduces the track occupation costs. Considering the experience in setting the decision-making period, the range of prediction errors for this model is acceptable and can support long-term tamping maintenance planning decisions.

5. Conclusions and Future Work

5.1. Conclusions

This study proposed a modeling method for predicting the tamping maintenance cycle of railway tracks by characterizing the deterioration pattern of track longitudinal level. The following conclusions can be drawn from the study findings.

- The proposed method employed a power-time-transformed Wiener process to construct the prediction model of each 200 m track segment, allowing simultaneous consideration of heterogeneity and uncertainty in the track geometry deterioration process.
- Both the deterioration parameters and the prediction results of the tamping maintenance cycle for the studied 200 m track segments ($n = 2171$) were inconsistent, indicating spatial heterogeneity in the deterioration pattern of track geometry.
- For the problem scenario considered in this study, the accuracy and solving efficiency of the MLE method were superior to those of the adaptive MCMC method, and the results obtained using the higher-performance optimization algorithm solver MPA when using the MLE method were more accurate.
- The overall prediction performance of all the prediction models for all segments was robust, meeting the management requirements for tamping maintenance planning over an annual or even longer time span.
- This study is helpful in assisting railway management in shifting the maintenance strategy from period-based preventive maintenance to condition-based predictive maintenance.

5.2. Future Work

In this study, the deterioration model parameters, which represent the degradation trend, were assumed to remain unchanged during different tamping maintenance cycles, making the proposed model suitable for scenarios in which the subgrade and track structure are stable, adjacent tamping cycles have a good memory of the track geometry state deterioration pattern, and the deterioration trend does not change significantly. However, for track sections that have entered the aging stage of their life cycle, the model may not adequately capture the changes in the deterioration parameters caused by structural aging. Therefore, the current model needs to be improved by further considering the influence of factors such as the amount of tamping maintenance and structural aging on the tamping maintenance cycle in the future. In other words, the influence of the track-geometry deterioration parameters β_k and θ_k during the tamping maintenance cycle must be considered for prospective prediction needs in long-term planning scenarios, such as 5 to 10 years of tamping maintenance or renewal.

Data Availability Statement: All data used during the study are confidential in nature and may only be provided with restrictions. The railway track longitudinal-level inspection data and tamping maintenance records used in this study are private data of the China Railway, and the authors do not have the right to disclose them publicly.

Conflicts of Interest: Authors Ru An and Lei Jia were employed by the company Shenzhen Urban Transport Planning Center Co., Ltd. Author Zhipeng Wang was employed by the company Line Branch Company Affiliated with Beijing Mass Transit Railway Operation Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Calculation Details of the Marine Predators Algorithm (MPA)

This appendix offers additional explanations for certain calculation details during the iterative process of the MPA as presented in Algorithm 1 in Section 3.1.

MPA divides the algorithm process into high-velocity ratio, unit-velocity ratio, and low-velocity ratio phases based on the current iteration number *Iter* and maximum iteration number *max_iter*. It incorporates the FADs effect to disrupt local optima and ultimately utilizes marine memory to re-evaluate individual fitness and update the top predator's function. In MPA, s_i denotes the step size of individuals for both prey $\textbf{\textit{Prey}}_k$ and top predator $\textbf{\textit{Elite}}_k$, which is influenced by the Lévy movement operator $\textbf{\textit{R}}_L$, Brownian motion operator $\textbf{\textit{R}}_B$, and the Kronecker product of prey and top predator at different phases.

Appendix A.1. Population Update

The following steps are executed in the while loop until the maximum number of iterations is reached:

Phase 1: In cases of high-velocity ratio or when the predator is moving faster than the prey, it falls under the first scenario for updating the position. Specifically, when the current iteration *Iter* is less than one-third of the maximum number of iterations (*max_iter*), the following mathematical model is employed to update the population:

$$\mathbf{s}_i = \textbf{\textit{R}}_B \otimes \left(\textbf{\textit{Elite}}_k^i - \textbf{\textit{R}}_B \otimes \boldsymbol{\gamma}_k^i\right), i = 1, \ldots, n$$
$$\boldsymbol{\gamma}_k^i = \boldsymbol{\gamma}_k^i + P * \textbf{\textit{R}} \otimes \mathbf{s}_i \tag{A1}$$

where R_B follows a standard normal distribution. The symbol $\otimes$ denotes element-wise multiplication of vectors. $P = 0.5$ is a predefined constant. R is a vector of uniformly distributed numbers in [0, 1]. γ_k^i represents the *i*th individual in the $Prey_k$. The multiplication of R_B by γ_k^i simulates the movement of prey. s_i determines both the distance and direction an individual moves within the search space.

Phase 2: When both predator and prey are moving at the same pace, or have a unit velocity ratio, the population is divided into two parts. The first half acts as prey and performs Lévy flights, while the second half acts as predator and undergoes Brownian motion. Specifically, when the current iteration *Iter* is between one-third and two-thirds of the maximum iterations (*max_iter*), the following mathematical model is used to update the population:

For the first half of the population (prey):

$$\mathbf{s}_i = \textbf{\textit{R}}_L \otimes \left(\textbf{\textit{Elite}}_k^i - \textbf{\textit{R}}_L \otimes \boldsymbol{\gamma}_k^i\right), i = 1, \ldots, n/2$$
$$\boldsymbol{\gamma}_k^i = \boldsymbol{\gamma}_k^i + P * \textbf{\textit{R}} \otimes \mathbf{s}_i \tag{A2}$$

where $\textbf{\textit{R}}_L$ is a vector of random numbers following the Lévy distribution to simulate the randomness of Lévy movement.

For the second half of the populations (predator:):

$$\mathbf{s}_i = \textbf{\textit{R}}_B \otimes \left(\textbf{\textit{Elite}}_k^i - \textbf{\textit{R}}_B \otimes \boldsymbol{\gamma}_k^i\right), = \frac{n}{2} + 1, \ldots, n$$
$$\boldsymbol{\gamma}_k^i = \textbf{\textit{Elite}}_k^i + P \cdot CF \otimes \mathbf{s}_i \tag{A3}$$

where $CF = \left(1 - \frac{Iter}{max\_iter}\right)^{2\frac{Iter}{max\_iter}}$, is an adaptive parameter used to regulate the step size for predator movement.

Phase 3: In cases of low-velocity ratio or when the predator moves faster than the prey, all individuals move in a Lévy manner. This situation occurs during the final phase of the optimization process, i.e., when the current iteration *Iter* is greater than two-thirds of the maximum iterations (*max_iter*). The following mathematical model is employed to update the population:

$$\mathbf{s}_i = \boldsymbol{R}_L \otimes \left(\boldsymbol{Elite}_k^i - \boldsymbol{R}_L \otimes \boldsymbol{\gamma}_k^i \right), i = 1, \ldots, n$$
$$\boldsymbol{\gamma}_k^i = \boldsymbol{Elite}_k^i + P * CF \otimes \boldsymbol{s}_i \tag{A4}$$

Appendix A.2. Application of FADs Effect

Fish Aggregating Devices (FADs) effects, a component of the Marine Predators Algorithm (MPA), operate as a mechanism that influences the optimization process by attracting prey towards promising regions within the search space. This effect is applied with a probability of 0.2 and is implemented using a binary vector $\boldsymbol{U}$.

The binary vector $\boldsymbol{U}$ is constructed by generating random numbers between 0 and 1 for each dimension of the search space. If the generated number is less than 0.2 (the probability of the FADs effect, denoted as *FADs*), the corresponding element in the binary vector is set to 0; otherwise, it is set to 1.

The FAD effect is incorporated into the position update equation as follows:

$$\boldsymbol{\gamma}_k^i = \begin{cases} \boldsymbol{\gamma}_k^i + CF[\boldsymbol{lb} + \boldsymbol{R} \otimes (\boldsymbol{ub} - \boldsymbol{lb})] \otimes \boldsymbol{U}, r \leq FADs \\ \boldsymbol{\gamma}_k^i + [FADs(1-r) + r]\left(\boldsymbol{\gamma}_k^{r1} - \boldsymbol{\gamma}_k^{r2}\right), r > FADs \end{cases} \tag{A5}$$

where $\boldsymbol{lb}$ represents the lower bound of the parameter value range, $\boldsymbol{ub}$ denotes the upper bound of the parameter value range; r is a uniform random number in [0, 1]; $r1$ and $r2$ represent random indices corresponding to elements in the prey matrix.

References

1. Wen, M.; Li, R.; Salling, K.B. Optimization of preventive condition-based tamping for railway tracks. *Eur. J. Oper. Res.* **2016**, *252*, 455–465. [CrossRef]
2. Andrews, J. A modelling approach to railway track asset management. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2013**, *227*, 56–73. [CrossRef]
3. Han, J. *China Railway Yearbook*; China Railway Publishing House Co., Ltd.: Beijing, China, 2021; pp. 246–258.
4. Su, Z.; Jamshidi, A.; Núñez, A.; Baldi, S.; De Schutter, B. Multi-level condition-based maintenance planning for railway infrastructures—A scenario-based chance-constrained approach. *Transp. Res. Part C Emerg. Technol.* **2017**, *84*, 92–123. [CrossRef]
5. Caetano, L.F.; Teixeira, P.F. Predictive Maintenance Model for Ballast Tamping. *J. Transp. Eng.* **2016**, *142*, 4016006. [CrossRef]
6. Soleimanmeigouni, I.; Ahmadi, A.; Kumar, U. *Track Geometry Degradation and Maintenance Modelling: A Review*; SAGE Publications: London, UK, 2018; Volume 232, pp. 73–102.
7. Khajehei, H.; Ahmadi, A.; Soleimanmeigouni, I.; Haddadzade, M.; Nissen, A.; Latifi Jebelli, M.J. Prediction of track geometry degradation using artificial neural network: A case study. *Int. J. Rail Transp.* **2022**, *10*, 24–43. [CrossRef]
8. Liao, Y.; Han, L.; Wang, H.; Zhang, H. Prediction Models for Railway Track Geometry Degradation Using Machine Learning Methods: A Review. *Sensors* **2022**, *22*, 7275. [CrossRef] [PubMed]
9. Giunta, M.; Bressi, S.; D'Angelo, G. Life cycle cost assessment of bitumen stabilised ballast: A novel maintenance strategy for railway track-bed. *Constr. Build. Mater.* **2018**, *172*, 751–759. [CrossRef]
10. Vale, C.; Ribeiro, I.M.; Calçada, R. Integer Programming to Optimize Tamping in Railway Tracks as Preventive Maintenance. *J. Transp. Eng.* **2012**, *138*, 123–131. [CrossRef]
11. Caetano, L.F.; Teixeira, P.F. Availability Approach to Optimizing Railway Track Renewal Operations. *J. Transp. Eng.* **2013**, *139*, 941–948. [CrossRef]
12. Andrade, A.R.; Teixeira, P.F. Hierarchical Bayesian modelling of rail track geometry degradation. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2013**, *227*, 364–375. [CrossRef]
13. Movaghar, M.; Mohammadzadeh, S. Bayesian Monte Carlo approach for developing stochastic railway track degradation model using expert-based priors. *Struct. Infrastruct. Eng.* **2022**, *18*, 145–166. [CrossRef]
14. Soleimanmeigouni, I.; Xiao, X.; Ahmadi, A.; Xie, M.; Nissen, A.; Kumar, U. Modelling the evolution of ballasted railway track geometry by a two-level piecewise model. *Struct. Infrastruct. Eng.* **2018**, *14*, 33–45. [CrossRef]
15. Famurewa, S.M.; Juntti, U.; Nissen, A.; Kumar, U. Augmented utilisation of possession time: Analysis for track geometry maintenance. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2016**, *230*, 1118–1130. [CrossRef]
16. Ahmed, M.O.; Khalef, R.; Ali, G.G.; El-Adaway, I.H. Evaluating Deterioration of Tunnels Using Computational Machine Learning Algorithms. *J. Constr. Eng. Manag.* **2021**, *147*, 04021125. [CrossRef]
17. Guler, H. Prediction of railway track geometry deterioration using artificial neural networks: A case study for Turkish state railways. *Struct. Infrastruct. Eng.* **2014**, *10*, 614–626. [CrossRef]

18. Lasisi, A.; Attoh-Okine, N. Principal components analysis and track quality index: A machine learning approach. *Transp. Res. Part C Emerg. Technol.* **2018**, *91*, 230–248. [CrossRef]

19. Lee, J.S.; Choi, I.Y.; Kim, I.K.; Hwang, S.H. Tamping and Renewal Optimization of Ballasted Track Using Track Measurement Data and Genetic Algorithm. *J. Transp. Eng. Part A-Syst.* **2018**, *144*, 04017081. [CrossRef]

20. Goodarzi, S.; Kashani, H.F.; Oke, J.; Ho, C.L. Data-driven methods to predict track degradation: A case study. *Constr. Build. Mater.* **2022**, *344*, 128166. [CrossRef]

21. Chen, S.; Feng, D.; Wang, W.; Taciroglu, E. Probabilistic Machine-Learning Methods for Performance Prediction of Structure and Infrastructures through Natural Gradient Boosting. *J. Struct. Eng.* **2022**, *148*, 4022096. [CrossRef]

22. Kim, D.; Kwon, K.; Pham, K.; Oh, J.; Choi, H. Surface settlement prediction for urban tunneling using machine learning algorithms with Bayesian optimization. *Autom. Constr.* **2022**, *140*, 104331. [CrossRef]

23. Tran, H.; Robert, D.; Gunarathna, P.; Setunge, S. Stochastic Prediction of Road Network Degradation Based on Field Monitoring Data. *J. Constr. Eng. Manag.* **2023**, *149*, 4023096. [CrossRef]

24. Bai, L.; Liu, R.K.; Sun, Q.X.; Wang, F.T.; Xu, P. Markov-based model for the prediction of railway track irregularities. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2015**, *229*, 150–159. [CrossRef]

25. Prescott, D.; Andrews, J. Investigating railway track asset management using a Markov analysis. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2015**, *229*, 402–416. [CrossRef]

26. Letot, C.; Soleimanmeigouni, I.; Ahmadi, A.; Dehombreux, P. An adaptive opportunistic maintenance model based on railway track condition prediction. *IFAC-PapersOnLine* **2016**, *49*, 120–125. [CrossRef]

27. Galván-Núñez, S.; Attoh-Okine, N. A threshold-regression model for track geometry degradation. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2018**, *232*, 2456–2465. [CrossRef]

28. Ye, Z.S.; Xie, M. Stochastic modelling and analysis of degradation for highly reliable products. *Appl. Stoch. Models. Bus. Ind.* **2015**, *31*, 16–32. [CrossRef]

29. Xiao, X.; Ye, Z. Optimal Design for Destructive Degradation Tests With Random Initial Degradation Values Using the Wiener Process. *IEEE Trans. Reliab.* **2016**, *65*, 1327–1342. [CrossRef]

30. Lee, J.S.; Hwang, S.H.; Choi, I.Y.; Kim, I.K. Prediction of Track Deterioration Using Maintenance Data and Machine Learning Schemes. *J. Transp. Eng. Part A Syst.* **2018**, *144*, 04018045. [CrossRef]

31. Khajehei, H.; Haddadzade, M.; Ahmadi, A.; Soleimanmeigouni, I.; Nissen, A. Optimal opportunistic tamping scheduling for railway track geometry. *Struct. Infrastruct. Eng.* **2021**, *17*, 1299–1314. [CrossRef]

32. Andrade, A.R.; Teixeira, P.F. Exploring Different Alert Limit Strategies in the Maintenance of Railway Track Geometry. *J. Transp. Eng.* **2016**, *142*, 04016037. [CrossRef]

33. Qu, J.; Liu, P.; Long, Y.; Xu, F. Main Factors on Effect of Precise Measurement and Precise Tamping Based on BP Neural Network. *Appl. Sci.* **2023**, *13*, 4273. [CrossRef]

34. Khouy, I.A.; Schunnesson, H.; Juntti, U.; Nissen, A.; Larsson-Kråik, P. Evaluation of track geometry maintenance for a heavy haul railroad in Sweden: A case study. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2014**, *228*, 496–503. [CrossRef]

35. Offenbacher, S.; Koczwara, C.; Landgraf, M.; Marschnig, S. A Methodology Linking Tamping Processes and Railway Track Behaviour. *Appl. Sci.* **2023**, *13*, 2137. [CrossRef]

36. China State Railway Group Co., Ltd. *General Speed Railway Line Maintenance Rules (TG/GW102—2019)*; China Railway Publishing House Co., Ltd.: Beijing, China, 2019.

37. Audley, M.; Andrews, J.D. The effects of tamping on railway track geometry degradation. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2013**, *227*, 376–391. [CrossRef]

38. Massingham, J.; Nielsen, O.; Butlin, T. An efficient method for generalised Wiener series estimation of nonlinear systems using Gaussian processes. *Mech. Syst. Signal Proc.* **2024**, *209*, 111095. [CrossRef]

39. Li, K.; Ren, L.; Li, X.; Wang, Z. Remaining useful life prediction of equipment considering dynamic thresholds under the influence of maintenance. *Eksploat. Niezawodn.* **2024**, *26*, 174903. [CrossRef]

40. Duan, F.; Wang, G.; Wei, W.; Jiang, M. Remaining useful life evaluation for accelerated Wiener degradation process model with mixed random effects and measurement errors. *Qual. Reliab. Eng. Int.* **2023**, *39*, 1334–1351. [CrossRef]

41. Yu, W.; Tu, W.; Kim, I.Y.; Mechefske, C. A nonlinear-drift-driven Wiener process model for remaining useful life estimation considering three sources of variability. *Reliab. Eng. Syst. Saf.* **2021**, *212*, 107631. [CrossRef]

42. Ashley, G.; Attoh-Okine, N. Approximate Bayesian computation for railway track geometry parameter estimation. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2021**, *235*, 1013–1021. [CrossRef]

43. Faramarzi, A.; Heidarinejad, M.; Mirjalili, S.; Gandomi, A.H. Marine Predators Algorithm: A nature-inspired metaheuristic. *Expert Syst. Appl.* **2020**, *152*, 113377. [CrossRef]

44. Luengo, D.; Martino, L.; Bugallo, M.; Elvira, V.; Särkkä, S. A survey of Monte Carlo methods for parameter estimation. *EURASIP J. Adv. Signal Process.* **2020**, *2020*, 25. [CrossRef]