



Estimation of the Average Speed for a Railway Signaling System by Using Gaussian Process Regression Methods with Bayesian Optimization

Mehmet Taciddin AKÇAY^{*1}, Abdurrahim AKGUNDOĞDU², Hasan TIRYAKI²

¹ Department of Electrical-Elektronics Engineering, Faculty of Engineering, Halic University, Istanbul, Turkey

² Department of Electrical-Elektronics Engineering, Faculty of Engineering, Istanbul-Cerrahpasa University, Istanbul, Turkey

*mehmettaciddinakcay@halic.edu.tr

(Alınış/Received: 25.05.2021, Kabul/Accepted: 30.06.2021, Yayınlama/Published: 31.07.2021)

Abstract: Transportation systems take an essential place in traffic planning. While designing railways, transportation between access points is planned to be realized within the desired time. The average speed between access points is affected by parameters like waiting time, motion resistance, slope, curve, traction force, maximum speed, the mass of the vehicle, and distance between two stations. The motion of the vehicle is calculated with these parameters, and the system design is performed accordingly. The average speed is one of the most critical factors affecting travel time between two access points. The headway may vary depending on the average speed. In this study, different regression methods were applied to estimate the average speeds calculated between stations in rail systems, and the obtained successful results were presented comparatively. When the methods used were examined, the Gaussian process regression method, which was optimized with the Bayesian algorithm, was observed to yield the most successful result. A Gaussian process (GP) is a collection of random variables, any finite number of which have a Gaussian distribution. Following simulations, the root mean square error and mean absolute error were found to be 0.064 and 0.047, respectively, and the coefficient of determination (R^2) value was obtained as 1 when the success rate of the method was calculated.

Keywords: Average speed, Bayesian optimization, Gaussian process, Railway, Regression, Traction

Bayesian Optimizasyonu ile Gauss Proses Regresyon Yöntemlerini Kullanarak Demiryolu Sinyalizasyon Sistemi için Ortalama Hız Tahmini

Öz: Ulaşım sistemleri trafik planlaması içerisinde önemli bir yere sahiptir. Ulaşım sistemleri göz önünde bulundurulduğunda ise demiryolları tüm sistemde büyük pay kaplamaktadır. Demiryolları tasarlanırken erişim noktaları arası ulaşımın istenilen sürede gerçekleştirilmesi planlanmaktadır. Erişim noktaları arası ortalama hız; bekleme süresi, hareket direnci, eğim, kurp, cer kuvveti, maximum hız, aracın kütlesi ve iki istasyon arası mesafe gibi parametrelerden etkilenmektedir. Aracı hareketi bu parametreler ile hesaplanarak sistem tasarımı buna göre gerçekleştirilmektedir. Ortalama hız iki erişim noktası arası seyir süresini etkileyen en önemli unsurlardan biridir. Ortalama hıza bağlı olarak sefer sıklığı süresi değişebilmektedir. Bu çalışmada raylı sistemlerde istasyonlar arası hesaplanan ortalama hızların tahmini için farklı regresyon yöntemleri uygulanmış ve elde edilen başarılı sonuçlar karşılaştırılmalı olarak verilmiştir. Kullanılan yöntemler incelendiğinde Bayesian algoritması ile optimize edilen Gaussian Process Regression yönteminin en başarılı sonucu verdiği görülmüştür. Gauss proses (GP), herhangi bir sonlu sayıda Gauss dağılımına sahip rastgele değişkenlerin topluluğudur. Benzetimler sonrasında root mean square error ve mean absolute error değerleri sırasıyla 0.064 ve 0.047 olarak bulunmuş ve yöntemin başarı oranı hesaplandığında determinasyon katsayısı (R^2) değeri 1 olarak elde edilmiştir.

Anahtar Kelimeler: Bayesian optimizasyon, Cer gücü, Demiryolu, Gaussian proses,, Ortalama hız, Regresyon

Cite as: Ö. T. Akçay, A. Akgundogdu, H.Tiryaki, "Estimation of the average speed for a railway signaling system by using gaussian process regression methods with bayesian optimization," *Demiryolu Mühendisliği*, no. 14, pp. 274-286, July, 2021. doi: 10.47072/demiryolu.942730

1. Introduction

Since rail systems contain many subsystems in their structure, it keeps solutions for various system variations at its center. When designing rail systems, the passenger capacity is determined with survey studies, and simulations are made. These studies are necessary in terms of optimal design conditions and performance goals. The capacity calculated appropriately for the line, the number of vehicles, station features, warehouse area, and signaling system are revealed with this study. Efficiency conditions for rail systems are mostly the investigation of the optimum passenger transport potential. The main target to be reached in terms of the signaling system is to provide the needed headway in the calculated period. Keeping the safety conditions at the highest level while meeting this need is one of the most critical features of the signaling system. The issues related to the high efficiency envisaged for the signaling system are necessary for providing the feasibility calculations. While creating the system, design principles are established by providing the CENELEC European standards and SIL4 (Safety Integrity Level) conditions. In order to increase the performance during operation, the vehicle and its equipment are in constant communication, and they operate in an integrated manner. While designing the signaling system, the system is accepted to have a life span of about 30 years. This system is necessary for vehicles in operation to realize a safe and secure drive. Operation traffic is organized in this way. The average speed is a very critical value for the solution of the signaling system details. The time difference between the two vehicles in travel is defined as the headway [1-3]. With this duration, the passing possibilities of vehicles in the vital areas of the line are calculated [4]. The average speed and the headway time have a significant effect on the passenger load of the line [5-7]. The average speed value can be calculated by graphically extracting the speed change of the vehicle. The average speed information can be accessed from the speed profile screen with simulations of the signaling system. In the speed profile, positions with a switch, various parameters, and imposed restrictions are taken into account. Slope, curve, tunnel information, and climatic conditions are the factors that are effective in the acceleration of the vehicle. After obtaining speed distribution information, operation planning is performed in accordance with the duration of the service. The speed graphs of the vehicles in the signaling system are shown in Figure 1. It was produced with the calculations obtained from the data given in the range in Table 1. After the acceleration value calculated with the traction force, the parameters of the signaling system are created and the speed profiles loaded on the onboard equipment are designed.

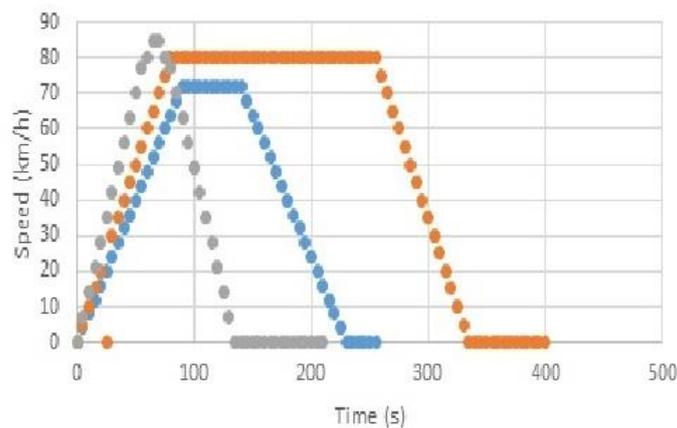


Figure 1. Speed profile versus time

The speed of the vehicle varies depending on the position of the vehicle. In the figure, three different max speed states were extracted. While the maximum speed is calculated to be 85 km/h in the gray curve, this value is 80 and 72 km/h in the orange and blue curves, respectively. The vehicles maintain their speed until the braking point after reaching the max speed. In this study, it is assumed that the safe braking distance is not a negative factor on the rail line and in interaction

with other vehicles, as indicated in the vehicle performance curves. In accordance with the jerk values and other parameters, the compliance of the safe braking distance has been confirmed at the intervals determined in Table 1. It was also checked that this value remained within the limits allowed by the signaling system. With the advancement of technology and a number of additional studies, the analysis of the signaling system can be conducted in a more effective and detailed way [8-9]. By carrying out traffic analysis, the model of the traffic flow is created [10]. The design of the electrification system must have sufficient and necessary infrastructure to ensure operational continuity and to obtain the performance values of the signaling system. When designing the electrification system, the system should provide an opportunity to be compatible with the determined headway. As a result of this, the required number of traction substations is preferred. In railways, providing the integration of all systems is essential for operational performance. Before the construction, models are created through obtaining the algorithms related to vehicle driving, and the performance calculations related to the operation are simulated in high-capacity computers with the help of these models.

The signaling system's main computer is located in the control center, and the management of the vehicles in operation is provided by this system, and the vehicles are monitored. With the speed information entered into the vehicles, the arrangement of the vehicles' travels is realized. Figure 2 shows the formation of the travel time, which is one of the most critical parameters of the signaling system, along the line. In Figure 1, the route parameter is calculated in the speed profile created by the signaling system, and the average speed is calculated by creating the route the vehicle goes. Here, the average speed is calculated by dividing the total distance by the total time. After the speed profile is loaded on the onboard equipment of the vehicles, the operating traffic is designed by placing the vehicles on the line according to the frequency of trip determined according to the feasibility factors and needs of the line.

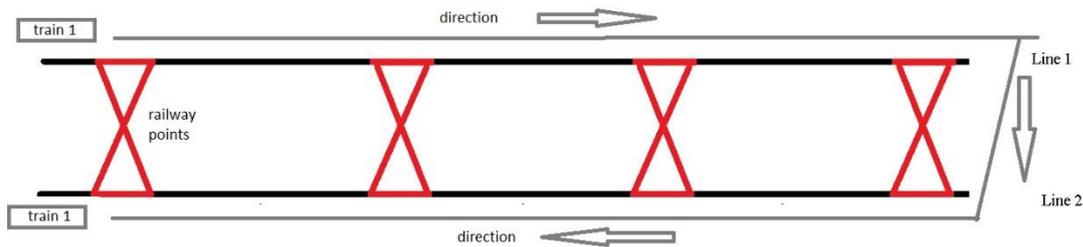


Figure 2. Formation of travel time

While the sections marked with red indicate the switch regions, the black line on the top shows line 1, the black line at the bottom shows line 2, and the gray elliptical curve shows travel time. In metro systems, the distance between the two stations does not decrease below 0.5 km. The distance between stations is determined according to the number of passengers, feeding status, and other features. There are several studies in the literature on average speed estimation. The accuracy of the traffic speed estimation was increased by using a new model named LC-RNN in the study [11]. The study [12] investigated the speed estimation models under mixed traffic conditions that occurred in urban arteries. The study [13] investigated vehicle speed estimation through the method of collecting data from mobile locations with the equipment placed in commercial vehicles. The article [14] investigated vehicle speed estimation models in the road design depending on the energy demand. The study [15] examined vehicle average speed estimation in a two-lane highway, depending on the weather and traffic characteristics. The study [16] addressed the subject of travel speed estimation by using machine learning methods.

The average speed emerges at the end of the total time elapsed from the vehicle's first take-off to the moment of stopping. For the performance of the signaling system and the all system, the efficiency of the driving profile is critical [17]. The railways vehicle speed also a necessary

parameter for monitoring and evaluating the track structure situation [18]. The speed and the general loading of a train are critical in deciding if the whole train is running in a protected and stable way [19]. The control center is the place where operation vehicle traffic is managed in rail systems. Vehicle traffic is planned at these points. With the traffic time plan obtained at these points, information regarding vehicles' positions, numbers, travel times, average speeds, headway, and the number of services is accessed. Before applying the signaling system, the operation is simulated, and the condition of the signaling system is tested. As a result of these tests, the results of the performance status of the system are obtained. With simulation, system tests are carried out by examining the operational data and vehicle features with a holistic approach. After the tests, information about the operation behavior is obtained. The integration of the parts that constitute the system is critical for operation performance in rail systems. The average speed information to be used in operation in the said tests is of great importance for determining the performance and infrastructure requirements of the system. For this reason, it is necessary to estimate the average speed information to be used in operation during the simulation phase. Thanks to this estimate, the optimal design and realization of the operation and infrastructure systems can be ensured. In this study carried out in accordance with this aim, the successful results obtained depending on the operation conditions of railway systems by applying different regression methods for the average speed estimation between stations in rail systems were presented comparatively.

In the second part of this article, the model proposed for the average speed estimation and the inputs-outputs used to create the structure of the model and the proposed method were explained. In the third part, the obtained simulation results were presented and compared with the method proposed as well as other methods commonly used in the literature for an objective evaluation. In the final part, the results of the article were evaluated in detail.

2. Materials and Method

2.1. Simulation and dataset

The traction force and acceleration status of the vehicle are calculated by using some equations for vehicle dynamics. The speed information of the vehicle is calculated by the following equations. The traction force produced by the vehicle is given by equation (1):

$$F_{\text{traction}} = F_{\text{motion}} + F_{\text{slope}} + F_{\text{curve}} + ma \quad (1)$$

F_{traction} refers to the traction force, while F_{motion} and F_{slope} express the motion and slope force. F_{curve} symbolizes the curve force. While m is the mass of the vehicle, a represents the acceleration of the vehicle.

The calculation of the total of the forces acting on the vehicle depending on the rotational mass and acceleration is given by equation (2):

$$\sum_{i=1}^n F_i = m' \times a \quad (2)$$

While m' is the rotational mass of the vehicle, F_i indicates the i^{th} force acting on the vehicle. With the help of equation (3), the distance is calculated depending on speed and time:

$$x_{(t)} = \int v_{(t)} dt \quad (3)$$

While x indicates the time-dependent position of the vehicle, v indicates the time-dependent speed of the vehicle. With equation (4), the number of vehicles to be used in operation traffic is calculated:

$$A = \Delta t/h \quad (4)$$

While the travel time is Δt , the headway is shown with h , and A indicates the number of vehicles. The minimum service frequency time is calculated with equation (5):

$$t_{min} = t_a + t_b + t_c \quad (5)$$

t_a represents the time it takes for the vehicle to travel from the switch area to the station area, t_b represents the time spent in the station area, and t_c represents the time it takes for the vehicle to travel from the station area to the switch area.

Within the scope of this study, a data set consisting of 500 samples, each with a different value, obtained using different operation data was used. From these data, Waiting Time, Motion Resistance, Slope, Curve, Traction Force, Maximum Speed, Mass of the Vehicle, and Distance Between Two Stations constitute the input parameters of the proposed method. The average speed value was chosen as the output of the system. The statistical distribution of these data is presented in Table 1. It was aimed to increase the number of sampling and increase the performance of the study by determining the speed limits randomly between the intervals used in the operation on a metro line. It has been assumed that there is no restriction on the speed targets set.

Table 1. Statistical characteristics of the data set that reveals Minimum (Min), Maximum (Max), Mean, and Standard Deviation (SD).

Influence Parameters	Description	Min	Max	Mean	SD
Waiting time (sec)	Input	10	30	20.512	5.964
Motion resistance (kN)	Input	1.1	1.3	1.199	0.058
Slope	Input	0	0.04	0.02	0.011
Curve (metre)	Input	801	10000000	6300332.098	4832444.598
Traction force (kN)	Input	200	300	250.48	28.084
Maximum speed (km/h)	Input	70	90	80.068	5.83
Vehicle mass (tons)	Input	200	220	209.534	5.731
Distance between two stations (meter)	Input	857.588	3066.826	1898.979	801.88
Average speed (km/h)	Output	27.489	67.414	47.215	9.004

The data obtained by using different operation data were obtained randomly, and the simulation image is demonstrated in Figure 3. With the help of equations (1), (2), (3) and (4), acceleration block, braking block and control block are formed. While there are traction force speed variables in the acceleration and braking block, the transitions between the two blocks are managed in the control block.

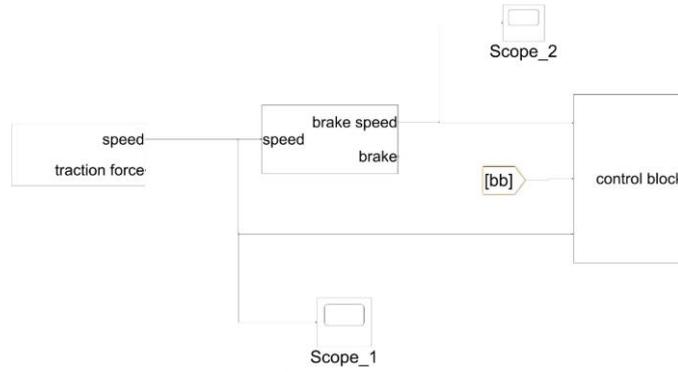


Figure 3. Simulation screen

The sampling frequency was determined as 1000 Hz to increase the simulation performance and to obtain high-accuracy results. In the simulations, the mass of the vehicle was chosen between 200 tons and 220 tons. In the solution of this problem, the regression methods were envisaged to be used due to its complicated structure and various variations being together.

In this article, the average speed of a vehicle moving between two stations was calculated by regression methods depending on different operation conditions. Among the used methods, the Gaussian process regression methods optimized with the Bayesian algorithm were observed to give the most successful results. Here, the Squared Exponential GPR, Exponential GPR, Matern 5/2 GPR, and Rational Quadratic GPR Gaussian process methods, which are among the Gaussian process regression methods, were applied with the optimization of the Bayesian algorithm, and the results were compared with the results of well-known machine learning methods.

2.2. Regression methods

Gaussian Process Regression Models: The Gaussian process regression (GPR) models are non-parametric kernel-based probability models. The training set is expressed with $(x_i, y_i); i = 1, 2, \dots, n$; where $x_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$ are drawn from an unknown distribution. When a GPR model, the new input vector x_{new} given, and exercise data are considered, y_{new} addresses the question of estimating a response variable value. A linear regression model is given in Equation 6:

$$y = x^T \beta + \varepsilon \quad (6)$$

where it was defined as $\varepsilon \sim N(0, \sigma^2)$. The error variance σ^2 and coefficients β are estimated from the data. A GPR model describes the response of the hidden variables, $f(x_i), i = 1, 2, \dots, n$, obtained from a Gaussian process (GP), and open basic functions, h . The covariance function of hidden variables captures the smoothness of the response, and the basic functions reflect the x inputs into a p -dimensional feature space [20].

A GP is a set of random variables, such that, any finite number has a common Gaussian distribution. If $\{f(x), x \in \mathbb{R}^d\}$ is a GP, then when (x_1, x_2, \dots, x_n) observations in the given number n are considered, the common distribution of random variables $f(x_1), f(x_2), \dots, f(x_n)$ is Gaussian. A GP is defined by its mean function $m(x)$ and the covariance function $k(x, x)$. If $\{f(x), x \in \mathbb{R}^d\}$ is a Gaussian process, it can be calculated as $E(f(x)) = m(x)$ and $Cov[f(x), f(x')] = E[\{f(x) - m(x)\}\{f(x') - m(x')\}] = k(x, x')$.

When the model given in Equation 7 is considered;

$$h(x)^T \beta + f(x) \quad (7)$$

where it is defined as $f(x) \sim GP(0, k(x, x'))$. In other words, $f(x)$ is obtained from a GP with zero mean with the covariance function of $k(x, x')$. $h(x)$ is a basic functions set that converts the original feature vector x in \mathbb{R}^d into $h(x)$ state, which is a new feature vector in \mathbb{R}^p . β is a p -by-1 vector of the basic function coefficients. This model represents a GPR model. A y example of the response can be modeled, as seen in Equation 8:

$$P(y_i|f(x_i), x_i) \sim N(y_i|h(x_i)^T \beta + f(x_i), \sigma^2) \quad (8)$$

Therefore, a GPR model is a probabilistic model. There is a hidden variable $f(x_i)$ introduced for each observation x_i that makes the GPR model non-parametric. In the vector form, this model is equivalent to the equation shown in Equation 9:

$$P(y|f, X) \sim N(y|H\beta + f, \sigma^2 I) \quad (9)$$

Here;

$$X = \begin{pmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, H = \begin{pmatrix} h(x_1^T) \\ h(x_2^T) \\ \vdots \\ h(x_n^T) \end{pmatrix}, f = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix} \quad (10)$$

In the GPR model, the common distribution of the hidden variables $f(x_1), f(x_2), \dots, f(x_n)$ is presented in Equation 11:

$$P(f|X) \sim N(f|0, K(X, X)) \quad (11)$$

Close to the linear regression model, $K(X, X)$ appears as in Equation 12:

$$K(X, X) = \begin{pmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_n) \\ \vdots & \vdots & \vdots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \dots & k(x_n, x_n) \end{pmatrix} \quad (12)$$

The covariance function $k(x, x')$ is usually parameterized by a series of kernel parameters or hyperparameters, θ . $k(x, x')$ is usually written as $k(x, x'|\theta)$ to clearly demonstrate dependence on θ [20].

The "fitrgp" function of MATLAB Statistics and Machine Learning Toolbox estimates the basic function coefficients, β , noise variance, σ^2 , and the hyperparameters, θ , of the kernel function that is obtained from the data, while training the GPR model [21]. The basic function, kernel (covariance) function, and initial values can be determined for the parameters.

Since a GPR model is probabilistic, it is possible to calculate the prediction intervals using the training model.

Bayesian Optimization Algorithm: Optimization, in its most general form, is the process of finding a point that minimizes a real-valued function called an objective function. Bayesian optimization is the name of such a process. Bayesian optimization internally preserves the Gaussian process model of the objective function and uses objective function evaluations to train the model. An innovation in Bayesian optimization is the use of an acquisition function that the algorithm uses to determine the next point to be evaluated. The acquisition function can balance

the sampling at points having low modeled objective functions and discover areas that are not well modeled yet [21].

The Bayesian optimization algorithm tries to minimize a scalar objective function $f(x)$ for x in a limited area. The function can be deterministic or stochastic, in other words, when x is evaluated at the same point, it can give different results. The components of x can be real names, integers, or categorical; this refers to a specific group of names [21].

The key elements in minimization are the Gaussian process model of $f(x)$, the Bayesian update procedure used to modify the Gaussian process model in each $f(x)$ evaluation, and an acquisition function $a(x)$ that is maximized in order to determine the next point for the evaluation (based on the Gaussian process model of f). Algorithm 1 shows the Bayesian optimization algorithm.

Algorithm 1. Bayesian Optimization Algorithm:

- 1: Evaluate $y_i = f(x_i)$ for x_i points taken randomly within variable limits
- 2: If there are evaluation errors, select random points until successful evaluations (The probability distribution of each component is uniform or log-scaled depending on the conversion value in the optimizable variable).
- 3: Update the $f(x)$ Gaussian process model to get a posterior distribution on the functions $Q(f | x_i, y_i)$ for $i = 1, \dots, t$.
- 4: Find a new point x , which maximizes the acquisition function $a(x)$.
- 5: Stop algorithm when;
 - 5a: When the desired number of iteration is reached.
 - 5b: When the specified time is reached.
 - 5c: When another stop criterion is provided.

Gaussian Process Regression for Fitting the Model: The underlying probability model for the objective function f is a primary Gaussian operation with added Gaussian noise as a result of observations. Therefore, the previous distribution on $f(x)$ is a Gaussian process with a mean $\mu(x; \theta)$ and covariance kernel function $k(x, x'; \theta)$. Here, θ is a vector of kernel parameters [21].

In a little more detail, $F = f_i$ shows the set of $X = x_i$ points associated with the objective function values. The common distribution of the previous function values (F) is the multivariate normal (mean $\mu(X)$ and covariance matrix $K(X, X)$, where $K_{ij} = k(x_i, x_j)$) [21]. The mean before the loss of generality is given as 0.

Furthermore, Gaussian noise with variance σ^2 was assumed to be added to the observations, and the previous distribution has the covariance of $K(X, X; \theta) + \sigma^2 I$ [21].

Fitting a Gaussian process regression model to the observations consists of finding values for noise variance σ^2 and kernel parameters θ .

2.3. Performance calculations

Statistical performance validation: In this study, the Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Correlation Coefficient (R^2) performance values were calculated. These performance values were defined by formulating below.

Mean Squared Error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (d_i - p_i)^2 \quad (20)$$

Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |d_i - p_i| \quad (21)$$

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i - p_i)^2} \quad (22)$$

Correlation Coefficient (R2)

$$R^2 = 1 - \frac{\sum_{i=1}^n (d_i - p_i)^2}{\sum_{i=1}^n (d_i - p_{avg})^2} \quad (23)$$

Here, d_i , p_i , and d_{avg} are the i^{th} desired output and estimated output, and the average of the desired output, respectively. n represents each of the samples in the dataset.

3. Results and Discussion

The data set used was simulated with the proposed GPR-Bayesian (GPR-B) model, and successful results were obtained. Figure 4 shows the block diagram of this study. Here, the 10-fold Cross-Validation (CV) method was used to avoid the overfitting of the system during the learning process. The data set was divided into ten groups during learning, and while nine groups were trained for learning, the remaining group was used in the testing process. This process was carried out ten times.

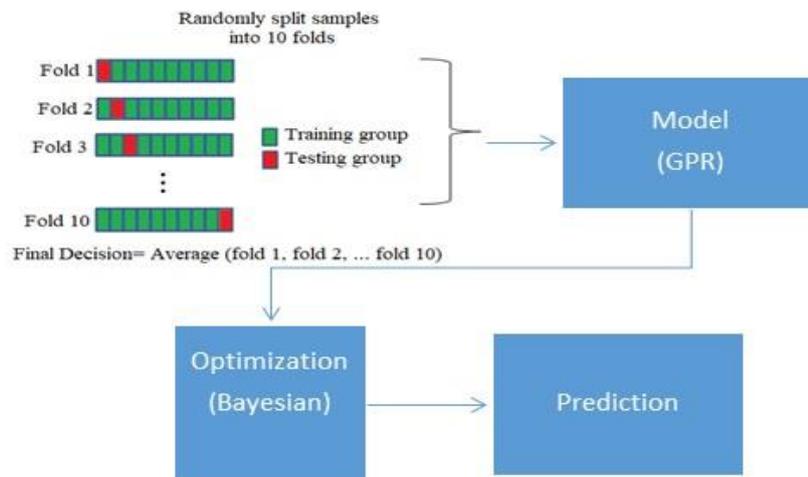


Figure 4. Block diagram of the proposed model

With the results obtained after the estimation made in Figure 5, the distribution of the actual results was shown on the graph. The closeness of the points on the graph to each other also visually proves how successful the model is.

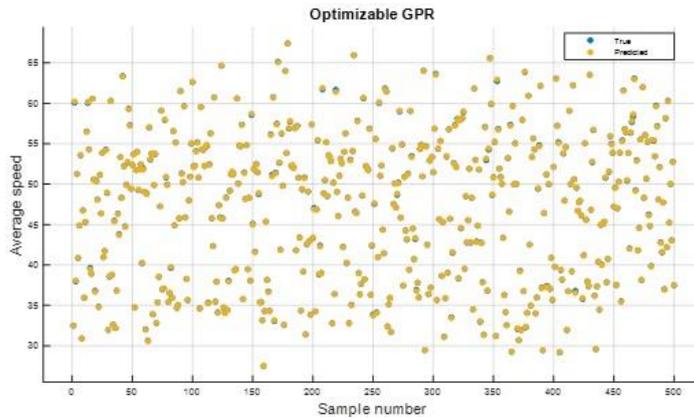


Figure 5. Distribution of the average speeds estimated with real average speeds

Upon examining Figure 6, the correlation between the desired output and the estimated output is observed. The fact that R^2 is 1 here suggests that there is a perfect correlation between the estimated average speed and the real average speed.

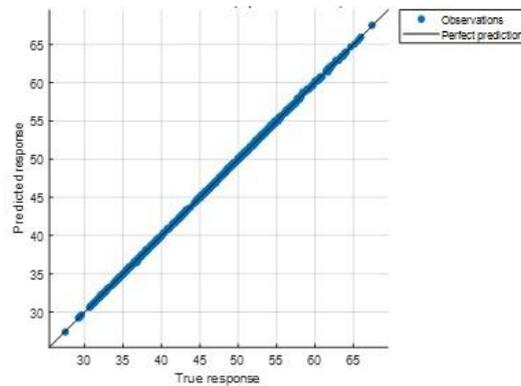


Figure 6. Regression result of the average speed estimation’s scatter plot

Figure 7 shows the decrease in error rate after GPR-B optimization during iterations. Here, the error obtained after the 5th iteration remained constant at a value of 0.00415, which is a very small value, and the iterations were stopped.

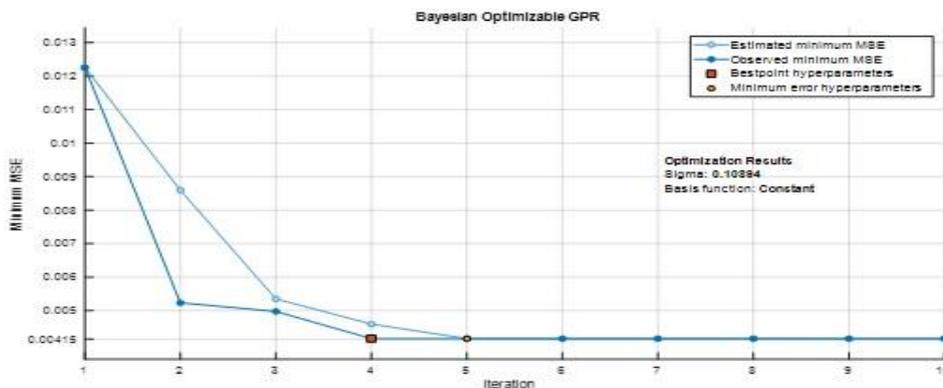


Figure 7. Displaying the error after Bayesian optimization

The comparison results of the outcomes obtained after implementing different methods are presented in Table 2. The results of the methods of Artificial Neural Network (ANN), Support Vector Machine (SVM), and Random Forest as Decision Trees, which give quite successful results in the literature, were compared with the proposed GPR-B method. Each model shown here was trained and tested under identical conditions, and a 10-fold CV was used for each method. It is clearly observed from here that GPR-B optimization is a quite successful model with the lowest error and the highest correlation when metric values such as RMSE, MAE, and R^2 are examined.

Table 2. Comparisons of prediction performances by using various approaches

Methods	Parameters	RMSE	MAE	R2
ANN	HN=5	0.4455	0.3279	0.996
	LR=0.3			
SVM	Poly Kernel	2.0433	1.5939	0.9494
	C=1			
Decision Tree	Random Forest	1.3449	1.0625	0.9812
GPR Bayesian	Kernel function=Nonisotropic	0.064552	0.047899	1
	Sigma=0.10894			
	Iteration=10			

The ANN method is the second method with the best results. The RMSE and MAE values are 0.45 and 0.33, respectively. The R^2 value is 0.996. The Decision Tree and SVM methods were observed to be the other successful methods, respectively. The R^2 value is 0.98 and 0.95 for these methods. With the LC-RNN model proposed in the study [11], the best RMSE value obtained for speed estimation is 4.68. The study [12] emphasized that the values obtained with speed estimation were found in the regression curve with a 95% success. In the study [13], while the best RMSE value for the results obtained for speed estimation was 10.84, the MAE value was 12.84. In the study [14], the R^2 value was obtained as 0.88 in the results produced in speed estimation. In the study [15], a 95% success was achieved in the results produced. In the study [16], the speed estimation results were produced by optimizing the system in a way that they remained within the limits of the ISO 14064-2 standard and by using machine learning methods.

4. Conclusion

The determination of the average speed value is very effective in increasing efficiency by capturing the line performance and the target headway in the railway operation. In this study, successful results obtained by applying different regression methods for the average speed estimation between stations in rail systems depending on the operation conditions of railways were presented comparatively. When the estimation results were examined, the Gaussian process regression method, of which optimization was carried out with the Bayesian algorithm, was observed to give the most successful result, and the GPR-B model, which is a model in which both methods are used together, was proposed. Besides the proposed method, quite effective results were also obtained with the ANN, SVM, and Decision Tree methods, among the different machine learning methods of which success is well known. The high success rates obtained show the accuracy of the model designed with the data obtained using different operation data. Here, the performance evaluation criteria such as RMSE, MAE, and R^2 were used as success criteria. Thanks to the proposed method, successful results can be obtained even if there are changes in the line parameters and features of a rail system line parameters. Therefore, it is recommended to use the proposed estimation model instead of the classical calculation methods used to determine the average speed information that causes the loss of time and labor loss due to possible errors during the design phase. As can be seen, this study is of importance in terms of the use and

dissemination of machine learning methods in rail system lines for the estimation of some desing parameters of rail system lines.

Acknowledgments

We would like to thank Istanbul Metropolitan Municipality, Rail System Department, for its support during the realization of this study. This study was supported by the Scientific Research Projects Coordination Unit of Istanbul University - Cerrahpasa. Project numbers: 23444 and 23446.

References

- [1] R., Riccardo, G., Massimiliano, “An empirical analysis of vehicle time headways on rural two-lane two-way roads,” *Procedia - Social and Behavioral Sciences*, no. 54, pp. 865 – 874, 2012.
- [2] I., W., Suweda, “Time headway analysis to determine the road capacity”, *Jurnal Spektran*, vol. 4, no. 2, pp. 71-75, 2016.
- [3] H., Nakamura, “Analysis of minimum train headway on a moving block system by genetic algorithm,” *Transactions on the Built Environment*, 34, 1014-1022, 1998.
- [4] J., Jang, C., Park, B., Kim, N., Choi, “Modeling of time headway distribution on suburban arterial:Case study from South Korea”, *Procedia - Social and Behavioral Sciences*, no. 16, pp. 240 – 247, 2011.
- [5] A., K., Maurya, S., Das, S., Dey, S., Nama, “Study on speed and time-headway distributions on two-lane B-bidirectional road in heterogeneous traffic condition,” *Transportation Research Procedia*, no. 17, pp. 428 – 437, 2016.
- [6] A., K., Maurya, S., Dey, S., Das, “Speed and time headway distribution under mixed traffic condition,” *Journal of the Eastern Asia Society for Transportation Studies*, no. 11, pp. 1774-1792, 2015.
- [7] C., C., Minh, K., Sano, S., Matsumoto, “The speed, flow and headway analyses of motorcycle traffic,” *Journal of the Eastern Asia Society for Transportation Studies*, no. 6, pp. 1496 – 1508, 2005.
- [8] D., Kong, X., Guo, “Analysis of vehicle headway distribution on multi-lane freeway considering car-truck interaction,” *Advances in Mechanical Engineering*, vol. 8, no. 4, pp. 1–12, 2016.
- [9] Y., Moriyama, M., Mistsuhashi, S., Hirai, T., Oguchi, “The effect on lane utilization and traffic capacity of adding an auxiliary lane,” *Procedia - Social and Behavioral Sciences*, no. 16, pp. 37 – 47, 2011.
- [10] H., Faheem, I., H., Hashim, “Analysis of traffic characteristics at multi-lane divided highways, case study from cairo-aswan agriculture highway,” *International Refereed Journal of Engineering and Science (IRJES)*, vol. 3 no. 1, pp. 58-65, 2014.
- [11] Z., Lv, J., Xu, K., Zheng, H., Yin, P., Zhao, X., Zhou, “LC-RNN: A deep learning model for traffic speed prediction,” *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*, Stockholm, Sweden, 2018, pp. 3470-3476.
- [12] A., Dhamaniya, S., Chandra. “Speed prediction models for urban arterials under mixed traffic conditions,” *Procedia - Social and Behavioral Sciences*, no. 104, pp. 342 – 351, 2013.
- [13] M., Gmira, M., Gendreau, A., Lodi, J., Potvin, “Travel speed prediction based on learning methods for home delivery,” *Interuniversity Research Center On Business Networks, logistics and transport, CIRRELT*, no.1, pp. 1-34, 2018.
- [14] M., Bysveen, “Vehicle speed prediction models for consideration of energy demand within road design,” *Master’s Thesis, Civil and Environmental Engineering, Norwegian University of Science and Technology*, 2017.
- [15] B., Mirbaha, M., Saffarzadeh, S., A., Beheshty, , M., Aniran, M., Yazdani, B., Shirini, “Predicting average vehicle speed in two lane highways considering weather condition and traffic characteristics”, *IOP Conference Series: Materials Science and Engineering*, no.1, pp. 1-7, 2017
- [16] M., Gmira, M., Gendreau, A., Lodi, J., Potvin, “Travel speed prediction using machine learning techniques”, *ITS World Congress*, Montreal, 2017, pp.1-10.
- [17] M., T., Akçay, “Estimation of constant speed time for railway vehicles by stochastic gradient descent algorithm”, *Sakarya University Journal of Computer and Information Sciences*, 3 (3) , 355-365, 2020.
- [18] W., Xinyue, S., Xianfeng, W., Jialiang, Y., Xun, H., Baoguo, “Train speed estimation from track structure vibration measurements,” *Applied Sciences*, no. 10, pp. 4742, 2020.
- [19] Z., Yuliang, W., Li, “Continuous monitoring of train parameters using iot sensor and edge computing,” *IEEE Sensors Journal*, no. 1, pp. 1-10, 2020.

- [20] C., E., Rasmussen, C., K., I., Williams, “Gaussian processes for machine learning”, MIT Press. Cambridge, Massachusetts, 2006.
- [21] The Mathworks, *Statistics and machine learning toolbox user’s guide*, 2019.

Özgeçmiş



Mehmet Taciddin AKÇAY

1984 yılında doğmuştur. İstanbul Teknik Üniversitesi Elektrik Mühendisliğinden mezun oldu. Yüksek Lisansını Sakarya Üniversitesi, Doktora eğitimini ise İstanbul Üniversitesi Elektrik-Elektronik Mühendisliğinde tamamlamıştır. İstanbul Büyükşehir Belediyesi Raylı Sistem Biriminde Cer Gücü, Sinyalizasyon Mühendisi olarak çalışmakta olup ve yine aynı zamanda Haliç Üniversitesi Elektrik-Elektronik Mühendisliği Bölümünde yarı zamanlı öğretim görevlisi olarak çalışarak akademik çalışmalarına devam etmektedir. Cer gücü simülasyonu, AC-DC demiryolu elektrifikasyonu, güç sistemlerinin analizi, kararlılığı, yapay zeka araştırma konuları arasındadır.

Email: mehmettaciddinakcay@halic.edu.tr



Abdurrahim AKGÜNDOĞDU

1975 yılında Bingöl’de doğmuştur. Lisans, Yüksek Lisans ve Doktora eğitimini İstanbul Üniversitesi, Mühendislik Fakültesinde tamamlamıştır. İstanbul Üniversitesi-Cerrahpaşa Elektrik-Elektronik Mühendisliği’nde Dr. Öğretim Üyesi olarak çalışmaktadır. Çalışma alanları; Biyomedikal, Yapay zeka, Örüntü tanımadır.

Email: akgundogdu@iuc.edu.tr



Hasan TİRYAKİ

1980 tarihinde doğmuştur. Lisans ve yüksek lisans eğitimlerini Kırıkkale Üniversitesi’nde tamamlamıştır. Doktora eğitimini ise İstanbul Üniversitesi’nde tamamlamıştır. İstanbul Üniversitesi-Cerrahpaşa, Elektrik-Elektronik Mühendisliği’nde Dr. Öğretim Üyesi olarak çalışmaktadır. İlgili alanına giren araştırma konuları; kontrol, yapay zekâ, elektrik makineleri ve demiryolu mühendisliğidir.

Email: hasan.tiryaki@istanbul.edu.tr

Beyanlar:

Bu makalede bilimsel araştırma ve yayın etiğine uyulmuştur. Tüm yazarların eşit oranda katkısı olmuştur.