International Journal of Technology 13(4) 848-859 (2022)
Received August 2021 / Revised January 2022 / Accepted April 2022

International Journal of Technology

http://ijtech.eng.ui.ac.id

Path Loss Modelling for High-speed Rail in 5G Communication System

Selvi Lukman[1*], Yul Yunazwin Nazaruddin[2], Bo Ai[3], Endra Joelianto[2,4]

[1]*Doctoral Program of Engineering Physics, Faculty of Industrial Technology, Institut Teknologi Bandung, Bandung, Indonesia*
[2]*Instrumentation and Control Research Group, Institut Teknologi Bandung, Bandung, Indonesia*
[3]*State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China*
[4]*University Center of Excellence on Artificial Intelligence for Vision, NLP and Big Data Analytics, Institut Teknologi Bandung, Bandung, Indonesia*

Abstract. A new path loss model for high-speed rail (HSR) in the 5G communication system is constructed in this paper. The model is identified to obtain an accurate mathematical representation of path loss multipath propagation in line of sight of HSR scenarios. The grey box modelling utilization of Generalized Reduced Gradient (GRG) and Genetic algorithm (GA) is applied to find the unknown parameters of the constructed path loss model since some uncertainties in obtaining the corresponded parameters are unavoidable to be collected in the field. Both algorithms achieve excellent results in finding the unknown parameter values with RMSE and MAPE evaluation which are converging finally to 2.779 and 1.701 %. The visualization of fitting plots is also presented, and GA provides a better-adjusted agreement with the measurement dataset of HSR. Accordingly, the constructed path loss model is successfully validated since it is capable of following the dynamic characteristic of the original HSR path loss measurement. The path loss model can then be utilized for the future dense deployment of HSR infrastructures for the 5G communication network.

Keywords: Generalized reduced gradient algorithm; Genetic algorithm; Grey-box modelling; Parameter estimation; Path loss model

1. Introduction

High-speed rail (HSR) refers to passenger rail systems utilizing a specialized rolling stock integrated system in the dedicated tracks. The deployment of FRCMS (Future Railway Communication System) is expected to fulfill the significantly increased demand for railway signaling systems. Some major communication functions in FRMCS are strictly related to railway operations with safety implications for the critical applications of the similar 5G technologies as radio communication cellular systems (Monserrat et al., 2020). Over the last decades, many researchers have been focusing on wireless communication technology that will be applied to HSR to ensure data transmission in the 5G framework (Suryanegara, 2018). A satisfying investigation of millimetre-wave propagation characteristics for HSR on field measurement in viaduct and tunnel scenarios has yielded the reliability of wireless transmission (Park et al., 2020).

The train backbone wireless networking is implemented based on point-to-point links

*Corresponding author's email: sylvia_lukman@students.itb.ac.id; Tel.: +62-22-2504252
doi: 10.14716/ijtech.v13i4.5058

devices, and the study of path loss in multipath propagation of HSR is stepping into new challenges when dealing with large-scale fading and shadowing. This types of propagation are the most likely to occur in railway scenarios. Since practically, HSR runs over 300 km/hr, it suffers from severe fading, vehicle penetration losses (VPL) and unavoidable Doppler effect. Accordingly, it is important to understand a chosen path loss model that can be utilized in line of sight of HSR propagation for a 5G communication system. At the same time, still revealing the stability of parameter, accuracy, and functionality of the limited measurement dataset.

It brings out some new challenges in obtaining some path loss parameters value in the field because of the combination of high velocity and spectrum allocation, particularly for the future 5G-HSR wireless system level. Those uncertainties parameters issues are almost existed not only in 5G cellular systems but also in the 5G-HSR scenario therefore a path loss parameterization scheme related to HSR environment scenarios from surrounding physical factors to the model variables must accommodate causal functions of associated 5G-HSR particularly, in line-of-sight variables which in this study, a new path loss model for 5G-HSR is constructed. The accuracy is validated by using a different approach of grey box modelling to yield a comprehensive knowledge of path loss for 5G-HSR that allows network designers to plan the most optimal infrastructures for HSR.

Furthermore, the major limitation of the existing research is based on particular scenarios, whether empirical or deterministic models. (He et al., 2018) was motivated to observe the path loss model using key parameters such as coherence time, polarization ratios, and Doppler shifts. A simulation was demonstrated based on channel measurement for HSR communications in a 5G Millimeter-Wave Band. As the future 5G technology requires many supporting technologies such as base station infrastructures and fiber optics to be laid on the tracks (Suryanegara, 2016), a local standard emerges as another solution to mainstream technologies. It led to another challenge for a requirement of an accurate path loss that can be utilized world-widely.

The early studies of the path loss model are majorly conducted in cellular networks. The models are derived from electromagnetic propagation theory (MacCartney et al., 2013), which are not very accurate but easy to implement. Therefore, some correction mechanisms must be constructed in a definite environment to achieve desired accuracy results. (Phillips et al., 2013) investigated additional parameters such as carrier frequency, distance, transmitter, receiver heights and carrier frequency for these cellular path loss models. The research yielded a more accurate path loss model for a 5G cellular network. Accordingly, a prior knowledge or an explicit measurement must be combined for a special path loss model for 5G-HSR (Zhong et al., 2021).

Several studies concerning 5G coverage path loss prediction were evaluated in recent years with the development of promising stochastic path loss models with the combination of antenna configuration and beamforming in cellular networks (Sousa et al., 2021). Other researchers introduced some key parameters of a line-of-sight characteristics (Sun et al., 2015). This work provided important key parameters of large-scale path loss scenarios and shadow fading for the future 5G communication system in urban macro cellular. The comparison was presented as well at the frequency of 2, 10, 18, and 28 Ghz in Aalborg, Denmark. Other works in realizing 5G stochastic path loss models were studied by (Rappaport et al., 2017) by investigating large-scale path loss models in wireless communication channels such as mm-MAGIC, NYUSIM, and 3GPP TR38,901. The study concluded that additional random variables in a path loss model must account for supplementary fading due to scattering and multipath effects, which are dominant but difficult to obtain in most stochastic path loss models.

One of the revolutionary technologies to develop a critical signaling 5G-HSR is millimeter wave technology. It is accessible to a massive capacity and bandwidth in frequency bands above 24 GHz. Since millimeter-wave suffers from higher propagation loss, MIMO directional antenna is widely accepted for designing a wireless communication system. However, 5G-HSR employs different characteristics from traditional cellular scenarios. Some specific characteristics in the 5G-HSR propagation environment, such as line of sight dominance, Doppler shift, high velocity, and multiple scenarios, must be carefully considered to obtain the optimal design and performance (Zhang et al., 2018). These considerations were related to diversity effects and the Doppler shift. Tuned free space path loss model was analyzed, and in this term a path loss model for HSR was estimated theoretically by gaining its diversity effect and Doppler shift performed by Maximal Ratio Combining (MRC) scheme (Roy & Fortier, 2004).

The capability of machine learning in analyzing the existing system performance to be more accurate is undoubtedly more resourceful for performing prediction tasks. It has been conducted several times by performing various training and testing on path loss datasets where an estimated model contains necessary predictions to be compared with the actual dataset. It produced the best prediction model for path loss, majorly in the cellular environment. In a single scenario, a requirement of evenly distributed data must be sufficient to be fed to the model for given prediction accuracy. However, the process becomes more complicated when incremental learning algorithms are involved because of gradual model constructions are performed without retraining accomplishment (Zhang et al., 2019).

The path loss models utilized in most existing research do not contemplate physical factors. In accordance with it, this work investigates an alternative approach to achieving an accurate path loss model for 5G-HSR. A grey box modelling is initiated as the mathematical representation of the path loss model for HSR in 5G communication system. The provision of grey box modelling implementation with appropriate algorithms will allow an optimal or almost optimal model that adjusts to the given path loss measurement. The idea is to find the global minimum cost function in a search space direction. The objective is to minimize the mean square error between the prediction dataset from the optimized model with the real measurement established in the field of study (Řehoř & Havlena, 2011).

The utilization of grey box modelling to obtain mathematical representation had been investigated to find the unknown parameters for AMPS (Automated People Mover System) train by using the Generalized Reduced Gradient Method (Suryana et al., 2020). A near-optimal solution was also achieved by maximizing the sum-rate capacity of a dynamic beam strategy to fulfill the critical quality of high-speed rail requirements through a problem decomposition using GA (Gao et al., 2018). Garah et al. (2016) investigated the Genetic Algorithm (GA) to produce a near-optimal solution for the GSM path loss model. The comparison has yielded a good agreement with the measurement result of the SUI model, COST-231 empirical path loss model, and COST-231 Hata.

Numerous path loss predictions in 5G scopes with different methods have been analyzed mostly for the case of cellular networks with recent contributions of machine learning implementation (Wu et al., 2010). However, to the best of author's knowledge, a grey box modelling approach to validate path loss models for 5G-HSR has not been found in any literatures because not only unavoidable measurement difficulties to be taken in the field but also a comprehensive knowledge about causative relationship between path loss parameters must be well constructed.

For a typical deployment of HSR wireless communication infrastructures, the line of Sight (LOS) scenario is usually referred to minimize radio waves reflection after traveling over a large area (Kanhere & Rapapport, 2021). The terminal equipment in HSR can add signal interferences either constructively or destructively. Random and rapid fluctuations in the received amplitude on a running HSR will cause a situation where signals spread in the frequency domain. It leads to one of the measurement difficulties, which copes with the path loss value $\overline{PL}(d_0)$. This parameter denotes a close-in measurement or a free space assumption from the transmitter where the signal starts to attenuate (Vahidi., 2021). The values of fading X_σ under HSR scenario is perhaps the most difficult parameter to achieve because of multipath propagation in a high mobility environment; consequently, an alternative solution must be considered for this parameter. The performance of wireless system transmission under increased mobility of high-speed rail is dependent on sub-carrier signal frequency shift and Orthogonal frequency division multiplexing [OFDM] due to Doppler Effect. This parameter explains a functional relationship with distance, and when the millimeter -Wave is taken into account for data transmission, a higher Doppler effect will be emerged (Xiong et al., 2021).

In this study, the investigation of GRG and GA for a grey box model identification is utilized to find some missing parameters value of the constructed path loss model for HSR in a 5G communication network. In this regard, the error between the output of the optimized path loss model and path loss original measurement data will be considered as objective functions with the visualization of fitting plots. The rest of the paper is organized as follows. The grey box model is introduced in Section 2. The optimization method and the constructed path loss model for 5G-HSR are investigated in Section 3. Section 4 displayed the simulation results, and finally, the paper is concluded in Section 5.

2. The Grey Box Model

In mathematics, statistics, and computational modelling, a grey box model is a well-known combination of partial theoretical structures and the real output system to gain an approximate model. The accuracy of parameter values as the representative of real dynamics behavior in a real system can be achieved by using optimization techniques (Bohlin, 2006; Hauth, 2008). Figure 1 illustrates the modelling procedure applied in this investigation.

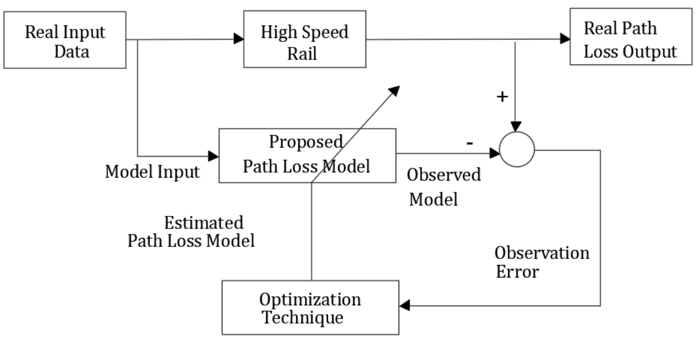


Figure 1 Modelling procedure using identification technique

Until now, the existing studies of resolving path loss problems are situated only at particular scenarios of high-speed rail without any physical factors considered. Subsequently, this work introduces a new path loss model for HSR with 26 Ghz in 5G spectrum allocation which is represented by:

$$\overline{PL}(dB) = 92.45 + 20\log|26Ghz| + \overline{PL}(d_o) + 10\gamma\log\left(\frac{d}{d_o}\right) + X_\sigma + f_d^{ab} + VPL - A_G \quad (1)$$

Where 92.45 is a constant value when f is measured in units of Ghz, The parameter d is in km units, and the utilized frequency carrier f_c is 26 Ghz. Figure 2 displays a typical HSR communication link under a line of sight scenario.

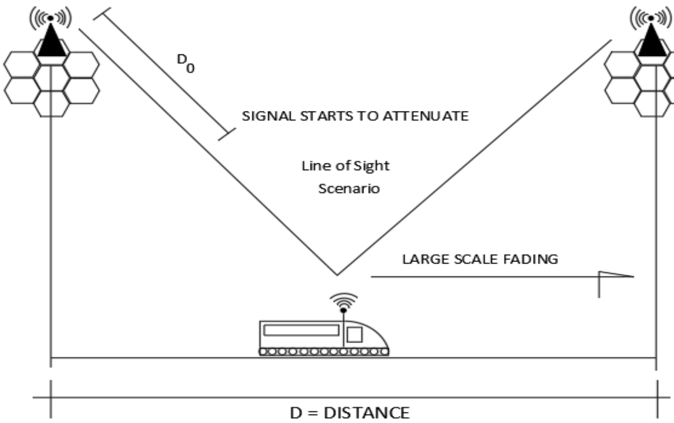


Figure 2 The illustration of high-speed rail communication link

Assuming $\overline{PL}(do)$ is denoted as X_1 is which is the first unknown parameter that describes the path loss values from close-in reference distance and typically determines from measurements where the signal starts to attenuate. The bars in the equation denote the ensemble average of all possible path loss values for a given value (d). When a log scale is plotted, the modeled path loss is a straight line with a slope equal to 10γ dB per decade. The value of γ depends on the specific propagation environment of HSR, as in free space γ is equal to 2, but larger obstructions will cause larger values of γ.

X_σ is the second unknown parameter which is denoted as X_2. It refers to large-scale fading when signal power is attenuated and fluctuated due to obstacles and interferences between the transmitter and receiver over a long distance. The last unknown parameter is the Doppler Effect, denoted as f_d^{ab}. The parameter of the Doppler Effect is defined as a nonlinear equation. It denotes the function of distance to time: $f_d^{ab}(t) = x(t).X_3$.

Particularly, Vehicular penetration loss (VPL) must be considered in the path loss parameterization because it severely affects signal propagation handovers due to the rapid change in the environment and multiple intersections, especially for high mobility scenarios of HSR. VPL can be put in the range of 15-25 dB depending on the operating frequency and vehicle types (Laiyemo, 2018). It is quantified by the ratio of received power immediately outside the HSR to the receiver inside the vehicle. Moreover, the signal attenuation in uplink and downlink is significantly affected by antenna configuration, carrier frequency, and the material used in the construction of high-speed rail. In the constructed path loss model, VPL is calculated 25 dB for 5G railway frequency band occupation. It is one of the critical challenges to perform an accurate path loss measurement since a large VPL will lead to the unreliable achievement of broadband communication for high-speed rail.

The antenna gain of 14.89 dB is added in the path loss model as the simulation result of 26 Ghz of MIMO 2x2 micro-strip antenna configuration for the 5G railway communication system. This gain antenna achievement is adequate than the existed gain of an integrated antenna system for 4G and mm-wave 5G (Naqvi et al., 2019). Data acquisition was carried out by measuring power transmission level when high speed approached and retreated the BTS (Base Station). Data acquisition was conducted when high-speed rail is in an idle state condition, which means no communications such as computers information, automatic data processing, information, or control system exchanges were being transmitted or received, and since most HSRs are built on viaducts scenario so consequently, additional undesirable random noises no longer existed. Moreover, the achievement of simulated antenna gain in the model is more than sufficient to cover those unwanted interferences. As easy as it is to represent the unknown parameters during the iteration process, the constructed path loss model can be written:

$$\overline{PL}(dB) = 92.45 + 20\log|26Ghz| + X_1 + 10\gamma\log\left(\frac{d}{d_o}\right) + X_2 + x(t).X_3 + 25dB - 14.89 \quad (2)$$

The objective function for the optimization process is represented by the root mean square error (RMSE) and Mean Absolute Percentage Error which are given as below formulae.

$$RMSE = \sqrt{\frac{1}{T}\sum_{i=1}^{N}\left(\left(y(dB) - \hat{y}(dB)\right)\right)^2} \quad (3)$$

$$MAPE = \sum_{t=1}^{n}\left|\frac{y_i - \hat{y}_i}{\hat{y}_i}\right| \times 100\% \quad (4)$$

GRG and GA are utilized to provide physical knowledge for model structures, combined with system identification, adaptation of the parameters and measurement dataset. The constraints to the optimization are defined by minimizing Equation 3 and Equation 4 where: $-500 \leq X_1 \leq 100, -500 \leq X_2 \leq 500, X_3 \geq 0$. The constraints are chosen large enough, assuming the anticipation of the high velocity and high frequency of 5G. By modifying X_1, X_2, X_3 desirable results of the optimized path loss model for high-speed rail, which fits adjustably to the given measurements, are achieved.

3. Optimization Method

3.1. GRG Algorithm

Solving a nonlinear equation system is perhaps the most difficult problem in numerical computations because of range in engineering application diversities. The convergence and performance characteristics are dominantly sensitive to the initial guess of the solution for most numerical methods such as metaheuristic algorithms.

Nevertheless, a former GRG (Generalized Reduced Gradient) algorithm offers some advantages of reliability and robustness. GRG describes a first-order optimization algorithm with the first derivative when updating the parameters. It was applied for solving large sparse nonlinear programs and the complete description of how to code the GRG algorithm was conducted efficiently. In its most basic form, the GRG algorithm focuses on the gradient or slope of the objective function as the change of decision variables or input values by regulating them until the optimum solution is achieved. The inequality constraints are converted into equality forms by utilizing the slack variables. GRG is proven to be robust for a former local search algorithm.

Basically, it linearizes nonlinear objective functions with some constraints for each candidate solution which in this term is the path loss of unknown parameter values X_1, X_2, X_3 by applying Taylor expansion in the search space. The candidate solutions X_1, X_2, X_3 are driven to find the most optimal points while avoiding some appointed constraints. The candidate solutions X_1, X_2, X_3 will automatically reverse themselves whenever they collide with the constraints, as depicted in Figure 3.

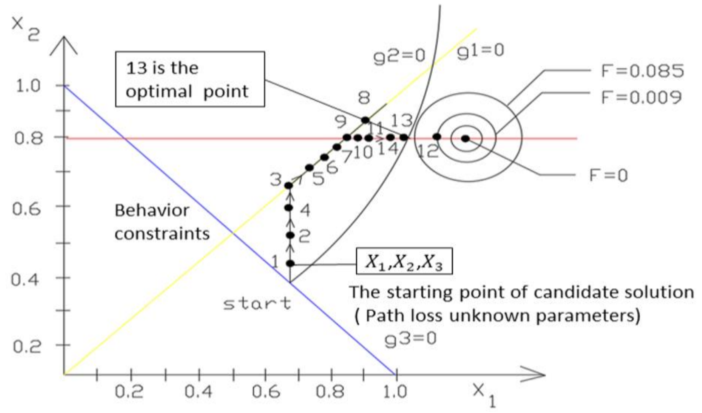


Figure 3 The procedure of GRG in finding the unknown parameters in the path loss model

A general form of a nonlinear equation where $F(c_d)$ and $g(c_d)$ are denoted as continuous functions that can be derived in the domain $\{c_d | l \leq c_d \leq u\}$. The problem is defined as:

$$\text{Minimize } F(c_d) \text{ with constraints } g(c_d) = b_g \text{ and } l \leq c_d \leq u \tag{5}$$

It is added to the decision variable of all constraints by setting a lower bound of zero, so it can start as a natural basic feasible solution. The procedure to complete the GRG algorithm can be written as follows:

1. Initialize c_d as an initial value and Find a feasible solution which must be less than the error value ε where $\varepsilon > 0$.
2. If the value is less than error value ε, the process must be terminated. If else, proceed to step 3.
3. Compute Jacobian (J) of the constraints.
4. Select a group of basic variables c_d as the basic solution for non-zero values. The matrix value B_p from the base, column J becomes non-singular, followed by the factorization of B_p.The remaining variable, which is namely the non-base variable C_{dnb} and the variable in the basic solution, which has zero value, can be written as $B_p = \frac{\partial_g(C_d)}{\partial_g(C_d)}$ and calculate The multiplier value $B_P^T = \frac{\partial F(C_d)}{\partial(C_{db})}$ is calculated.
5. Solve the value of reduced gradient $r = \frac{\partial F(C_d)}{\partial(C_{db})} - J^T p$, If the r value is projected on the small constraints, the calculation is terminated due to the optimal value achievement.
6. Find and construct search directions d for the non-base variables r and construct tangential directions. For each step, adjust the value C_{db} which fulfills $g(C_{db}, C_{dnb}) = b_g$ by utilizing B_p
7. Return to step 2

3.2. Genetic Algorithm

The Genetic Algorithm (GA) is a random-based evolutionary algorithm inspired by Charles Darwin's theory of natural evolution. The GA is inspired by the process of natural genetics, which comprises reproduction of an original population, crossover performances, mutation, and selection of the best (D'Angelo, 2021). GA, as the utilized optimization technique, works on some solutions in the population size. However, unlike the GRG algorithm, every candidate solution is individually driven to find optimal solutions while strictly considering some constraints. In GA, candidate solutions of X_1, X_2, X_3 are dispersed in the search space as a single chromosome representing genes to be converted into the binary system. Figure 4 illustrates the searching process of the unknown path loss parameters for the most optimal solution.

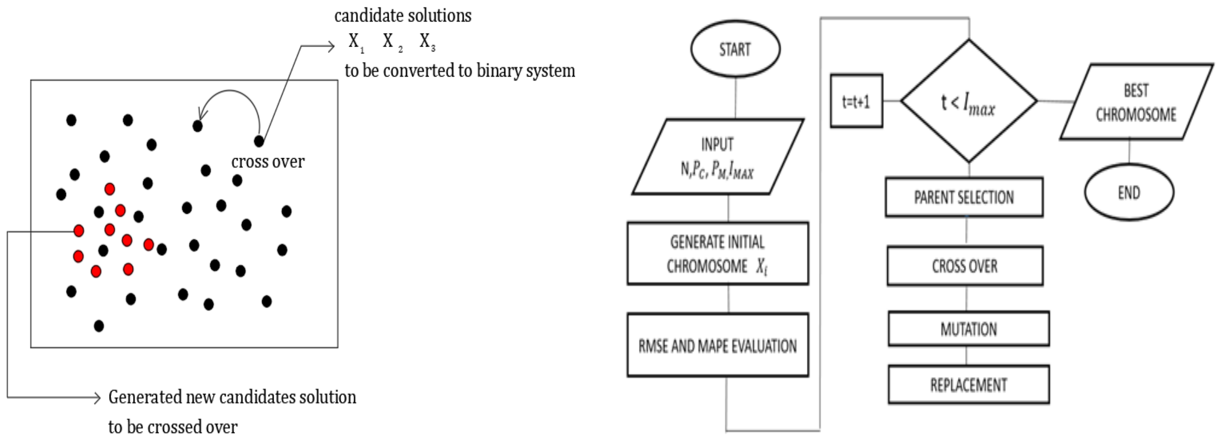


Figure 4 The procedure of the Genetic Algorithm in finding the best solution for the path loss model

Figure 4 shows that each individual who is this term is the unknown parameters of the path loss model in a population has a fitness value. The best individuals that represent the quality of the solutions are selected in a fitness function. Each individual's fitness score is given based on their ability to compete with others the best individual selection based on quality is called a mating pool, where the genes are exchanged. The higher probability is achieved by a higher quality individual as well. Mating high-quality individuals are expected to attain a better quality offspring than their parents, which will prevent bad individuals from generating more bad individuals. So accordingly, there shall be higher opportunities to keep only good properties of the individuals.

This process will end up with the desired optimal solution. The most significant phase in GA is crossover. The process of cross-over is selected randomly from within the genes for each mated pair of parents. The offspring have the characteristics of their parents and thus the similar drawbacks in their parents will actually endure in the new offspring. Some genes can be subjected to a mutation of a low random probability for particular new offspring formed. The diversity in the population is maintained, and the premature convergence can be avoided as well. The process is slow and gradual until the best solution is achieved. The iteration of GA is terminated when the population has reached its convergence. It implies that significant differences from the previous generation of offspring are no longer produced. The sequence of phases is repeated to obtain individuals in each new generation who are better than the previous generation.

4. Simulation Results

GRG and GA require initial values before starting the iteration process to solve the unknown parameters of path loss for 5G-HSR (Koshikawa et al., 2007; He et al., 2020). Equation 2 shows there are three parameters that will be approximated: X_1, X_2 and X_3. These candidate solutions must be fulfilled the initial constraints, which are shown is Table 1 and Table 2. The dataset was separated into trainingand testing datasets in a ratio of 80:20, where 80% is used for data training and the rest 20% is for data testing. The objectivity of doing this procedure is to evaluate the model's consistency and achieve better model robustness.

Table 1 The initial $X_{1,}X_2$,X_3 approximate new values and performance evaluation of GRG

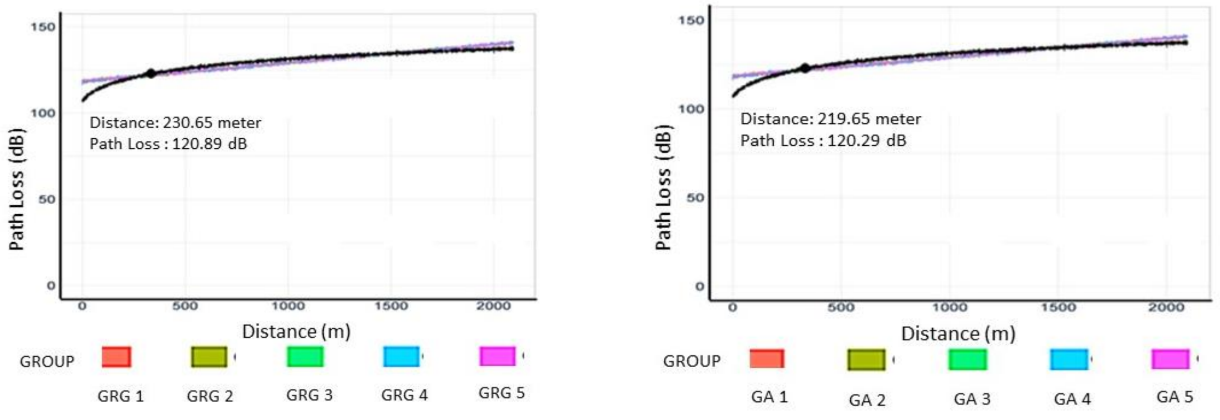
Trial		1	2	3	4	5
Initial X_1		47	57	67	77	87
Initial X_2		-100	-200	-300	-400	-500
Initial X_3		0.01	0.02	0.03	0.04	0.05
New X_1		0.0017	37.745	59.10	74.29	87.44
New X_2		-399.41	-437.15	-458.51	-473.70	-486.86
New X_3		0.01075	0.01075	0.01075	0.01075	0.01075
Training Set	RMSE	2.708	2.708	2.708	2.708	2.708
	MAPE	1.676%	1.676%	1.676%	1.676%	1.675%
Testing Set	RMSE	2.979	2.979	2.979	2.979	2.979
	MAPE	1.775%	1.775%	1.775%	1.775%	1.775%
Fully Optimized Set	RMSE	2.7790	2.7791	2.7790	2.7790	2.7791
	MAPE	1.7015%	1.7014%	1.7014%	1.7015%	1.7015%

Table 2 The initial $X_{1,}X_2$,X_3 approximate new values and performance evaluation of GA

Trial		1	2	3	4	5
Initial X_1		47	57	67	77	87
Initial X_2		-100	-200	-300	-400	-500
Initial X_3		0.01	0.02	0.03	0.04	0.05
New X_1		50.40	38.67	64.79	53.04	74.80
New X_2		-449.79	-438.12	-462.22	-452.45	-474.22
New X_3		0.01074	0.01079	0.01075	0.01074	0.01075
Training Set	RMSE	2.7089	2.7090	2.7089	2.7089	2.7089
	MAPE	1.674%	1.678%	1.677%	1.676%	1.677%
Testing Set	RMSE	2.9799	2.9770	2.9786	2.9793	2.9789
	MAPE	1.7738%	1.7793%	1.7770%	1.7755%	1.7766%
Fully Optimized Set	RMSE	2.7793	2.7786	2.7789	2.7791	2.7790
	MAPE	1.6993%	1.7056%	1.7027%	1.7010%	1.7022%

Choosing the initial parameter values that can provide a fast convergence of RMSE and MAPE value is challenging because incorrect ones will trap the values in the local minima. As in definition, RMSE shows how accurate the model is in performing prediction tasks even though the criterion depends on individual objectives, so more than one indicator might be very useful. Additionally, mean absolute percentage error (MAPE) describes the average of the absolute percentage errors. It simply indicates how many errors are yielded in performing prediction tasks with the real value of measurements. Table 1 and Table 2 show similar satisfying RMSE and MAPE values that eventually converge to RMSE = 2.779 and MAPE = 1.701 %. Even though GA suffers from a slow convergence right start from the beginning of the iteration process, GA has already proven its robustness in performing approximation tasks. For a better understanding of the performance evaluation from both algorithms, an additional evaluation performance is provided to show agreements with the measurement dataset presented in Figure 5 below.

Figure 5 displays the almost similar visualization fitting plots. Both optimization techniques are confident to follow the dynamic behavior of HSR at the distance of 230.65 meters and path loss value at 120.29 dB for GRG, followed by GA at the distance of 219.90 meters and path loss value of 120.53 dB. Although, slight discrepancies exist in the fitting plot of both algorithms. GA provides a better-adjusted agreement with the measurement dataset result which is faster than GRG. The validation performance shows that the constructed path loss model for 5G-HSR is accurate.



[a] [b]

Figure 5 The visualization of fitting plot performance: [a] GRG; and [b] GA

Hence, based on the satisfactory achievement from GA, the new parameter values are put back into Equation 2, and the resulting path loss model for high-speed rail in a 5G communication system is as below.

$$\overline{PL}(dB) = 92.45 + 20\log|26Ghz| + 74.80(dB) + 10\gamma\log\left(\frac{5000}{240}\right) + (-474.2)(dB) + 21.9(dB) + 385dB \tag{6}$$

The representation of GRG and GA path loss model complies with the best fit approximate solutions to the original path loss measurement of high-speed rail even though slowness in converging to the solution was performed in GA. The satisfactory result indicates that the construction path loss model of 5G-HSR has successfully dealt with the original data measurement, and accordingly, the accuracy validation of the path loss model is effectively implemented by utilizing grey box modelling

5. Conclusions

As the significance of 5G wireless network planning continues to grow, so will the requirements for better methods of measuring wireless signal propagation and modelling a path loss prediction for high speed rail. This paper gives a broad overview of approaches given in a grey box modelling to validate a newly constructed path loss model in the 5G communication system for HSR. The grey box modelling with the application of GRG and GA has shown excellent results in finding the unknown parameters value of the newly constructed path loss model with satisfying results of RMSE convergence approximately to 2.779 and MAPE value 1.701 %, respectively. The results revealed that the new path loss model is successfully validated. The framework in this study had shown that the created path loss model had a good adjusted agreement with the dynamic characteristic of the original path loss measurement which GA ultimately achieves. In future works, many possible directions in this area with promising great impacts in high-speed rail crucial applications are widely open for investigation. Comparative validation techniques and measurement-based approaches are required, so the validated path loss model can be utilized to design the future dense wireless communication infrastructures for high-speed rail in a 5G communication network.

Acknowledgements

The Indonesian Ministry of Research partially funds this research, Technology and Higher Education under WCU Program managed by Institute Technology Bandung and Institute Technology Bandung Research Program 2022. The authors would also like to acknowledge BJTU, Beijing, China, for providing dataset measurement of a running bullet train.

References

Bohlin, T.P., 2006. *Practical Grey Box Identification: Theory and Applications.* London: Springer-Verlag

D'Angelo, G., Palmieri, F., 2021. A Modified Genetic Algorithm with Gradient Based Local Search for Solving Constrained Optimization Problems. *Information Sciences,* Volume 547, pp. 136–162

Gao, M., Ai, B., Zhong, Z., Liu, Y., Ma, G., Zhang, Z., Li, D., 2018. Dynamic mmWave Beam Tracking for High Speed Rail Communications. *In*; IEEE Wireless Communication and Networking Conference Workshop (WCNCW), pp. 278–283

Garah, M., Djouane, L., Oudira, H., Hamdiken, N., 2016. Path Loss Models Optimization for Mobile Communication in Different Areas. *Indonesian Journal of Electrical Engineering and Computer Science*, Volume 3(1), pp. 126–135

Hauth, J., 2008. *Grey Box Modelling for Non Linear Systems.* Dissertation, Fachbereich Mathematik der Technischen Universitat Kaiserslautern

He, D., Ai, B., Guan, K., Zhong, Z., Hui, B., Kim, J., Chung, H., Kim, I., 2018. Channel Measurement Simulation and Analysis for High Speed Railway. Communications in 5G Millimetre-Wave Band. *IEEE Transactions on Intelligent Transportation Systems,* Volume 19, pp. 3144–3158

He, D., Wang, Y., Chen, Y., Yuan, Y., Zhen, T., Liu, J., 2020. Topology Optimization of Train Communication Network Based on Improved Adaptive Genetic Algorithm. *In*: 2020 International Conference in Communication Technology (ICCT), pp. 182–189

Kanhere, O., Rappaport, T.S., 2021. Position and Location of Futuristic Cellular Communications: 5G and Beyond. *IEEE Communication Magazine*, Volume 59(1), pp. 70–75

Koshikawa, R., Terui, A., Mikawa, M., 2007. *Solving Systems of Non Linear Equations with Genetic Algorithm and Newton's Method.* University of Tsubuka, Tsubuka, Japan, https://doi.org/10.48550/arXiv.2007.05159

Laiyemo, A.O., 2018. *High Speed Moving Networks in Future Wireless System*. Doctoral Dissertation, Universitatis Ouluensis, Oulu, Finlad

MacCartney., J, Zhang., S, Nie., Rappaport, T.S., 2013. Path Loss Models for 5G Millimeter Wave Propagation Channels in Urban Microcells'. *In*: 2013 IEEE Global Communications Conference (GLOBECOM), pp. 3948–3953

Monserrat, J.F., Adam, D., Lamas, C.B., Sultan, S., 2020. Envisioning 5G Enabled Transport. Report for World Bank Group, Washington

Naqvi, S.I., Naqvi, A.H., Arshad, F., 2019. An Integrated Antenna System for 4G and Millimetre-Wave 5G for Future Handheld Devices. *IEEE Access,* Volume 7, pp. 116555-116566

Park, J.J., Lee, J., Kim K.W., Kim, M.D., 2020. 28 Ghz High Speed Train Measurements and Propagation Characteristic Analysis. *In*: 14 th European Conference on Antennas and Propagation (EuCAP). pp. 1–5

Phillips, C., Sicker, D., Grunwald, D., 2013. A Survey of Wireless Path Loss Prediction and Coverage Mapping Methods. *IEEE Communications Surveys & Tutorials,* Volume 15(1), pp. 255–270

Rappaport, T.S., Sun, S., Shafi, M., 2017. Investigation and Comparison of 3GPP and NYUSIM Channel Models for 5G Wireless Communications. *In*: 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall), pp. 1–5

Řehoř, J., Havlena, V., 2011. A Practical Approach to Grey-box Model Identification. *In*: IFAC Proceedings, Volume 44(1), pp. 10776–10781

Roy, S., Fortier, P., 2004. Maximal-Ratio Combining Architectures and Performance with Channel Estimation Based on a Training Sequence. *In*: IEEE Transactions on Wireless Communications, Volume 3(4), pp. 1154–1164

Suryana, L.E, Joelianto, E., Hidayat, Y.A., 2020. Grey Box Modelling and Validation of AMPS Train. *International Journal on Electrical Engineering and Informatics,* Volume 12(4), pp. 956–965

Suryanegara, M., 2016. 5G as Disruptive Innovation: Standard and Regulatory Challenges at a Country Level. *International Journal of Technology*, Volume 7(4), pp. 635–642

Suryanegara, M., Asvial, M., 2018. The Patterns of Innovation Agendas on 5G Mobile Technology. *International Journal of Technology*, Volume 9(5), pp. 876–887

Sousa, M., Alves, A., Vieira, P., Queluz, M.P., Rodrigues, A., 2021. Analysis and Optimization of 5G Coverage Predictions Using a Beamforming Antenna Model and Real Drive Test Measurements. *IEEE Access,* Volume 9, pp. 101787–101808

Sun, S., Thomas, T.A., Rappaport, T.S., Nguyen, H.C., Kovacs, I.Z., Rodrigues, I., 2015. Path Loss, Shadow Fading, and Line-of-Sight Probability Models for 5G Urban Macro-Cellular Scenarios. *In*: 2015 IEEE Globecom Workshops (GC Wkshps), pp. 1–7

Vahidi, V., 2021. High Speed Trains Communication Systems in 5G Cellular Network. *Digital Signal Processing*, Volume 115, p. 103075

Wu, D., Zhu, G., A, Bo., 2010. Application of Artificial Neural Networks for Path Loss Prediction in Railway Environments. *In*: 2010 5th International ICST Conference on Communications and Networking in China, pp. 1–5

Xiong, L., Zhang, Z., Yao, D., 2021. Dynamic Doppler Prediction in High Speed Rail Using Long Short Term Memory Neural Network. *Transactions on Emerging Telecommunication Technologies,* Volume 32(9), p. e4269

Zhang, C., Wang, G.P., Jia, M.Z., He, R., Zhou, Li., Ai, B., 2018. Doppler Shift Estimator for Millimeter-Wave Communication Systems on High-Speed Railways. *In*: 2018 IEEE/CIC International Conference on Communications in China (ICCC), pp. 227–231

Zhang, Y., Wen, J.X., Yang, G.S., He, Z.W., Wang, J., 2019. Path Loss Prediction Based on Machine Learning: Principle, Method, and Data Expansion. *Applied Sciences*, Volume 9(9), p. 1908

Zhong, G., Xiong, K., Zhong, Z., Ai, B., 2021. Internet of Things for High-speed Railways. *Intelligent and Converged Networks,* Volume 2(2), pp. 115–132