The special case of Egli and Hata model optimization using least-square approximation method

Thaisa Jawhly¹ · Ramesh Chandra Tiwari¹

Received: 19 November 2019 / Accepted: 12 June 2020 / Published online: 27 June 2020
© Springer Nature Switzerland AG 2020

Abstract

Terrestrial signal loss prediction for the VHF band is limited to a few models. With perpetual prominence in its application, understanding VHF signal variation in a different scenario stimulates interests and demand for ruling out a decent performing model applicable in the VHF band. This paper aims to optimize the selected model using measurement data for the precise deployment of cellular mobile signal over the study area. We chose the least-square regression fitting method and applied it to Egli and Hata model for improved prediction with measurement data made over a vast hilly forested area. We observed that optimized models of Egli and Hata model predicted equal loss, although framed differently. We validated the adjusted Egli and Hata's model with the original models and the FCC-based Perez Vega model using a coefficient of determination, which showed improved performance for the North and the South. This study will help in optimizing the existing path loss model using the least-square approximation method and obtain the optimal prediction model for the study area or other areas with similar topography. This will aid communication service enhancement, the proper planning, and the future deployment of wireless communication system across the study area.

Keywords Least-square fitting · Egli model · Hata model · Propagation losses · Statistical analysis

1 Introduction

The very high-frequency (VHF) transmission has varied prominence in civilian and military application. The characteristics of the VHF frequency band have the potential to cover large areas and support reliable communication systems [1]. VHF wireless propagation is far less complicated compared to the higher band due to reduced scattering and reflection [2]. Over the decade, numerous works have stressed the prospect of exploiting the vacant band using cognitive radio access (CR) and television white space (TVWS) [3–9].

The low VHF band and the upper VHF band are underutilized in most of the countries. [10–13], and it is considered a potential solution for bridging digitally divided regions. In India, digital switchover is still underway [14],

and this switchover will generate a vacant band, creating a prospect for providing low-cost wireless facility to the remote rural areas. In this light, understanding signal variability of the same frequency spectrum in the study area poses a vital aspect.

Terrestrial mobile radio propagation can roughly be characterized as the variation of signal strength due to propagation loss through distance, shadowing, and multipath fading. A diverse primary physical principle causes each of these nearly independent phenomena. Propagation path loss is the weakening of power radiated as it spreads out in space and the counterproductive effects of the propagation channel. The base station and the mobile station antenna height greatly influenced the propagation mechanism and the losses involved, but the primary contributor to propagation loss is its propagation

✉ Ramesh Chandra Tiwari, ramesh_mzu@rediffmail.com; Thaisa Jawhly, thaisjly@gmail.com | ¹Department of Physics, Mizoram University, Tanhril, Aizawl, Mizoram 796004, India.

environment. This attenuation limits the rate of the data transmission which needs to be accounted for when designing a communication network.

In wireless communication, path loss models aim to forecast the mean deviation of signal strength received at a particular setting and formulate the parameters or the variables that affect the signal variations. To model propagation loss, conversely, is challenging, which requires vast data over diverse channel parameters and apprehend their impact on various signal frequencies. The other substitute is to exploring existing empirical propagation loss model. In practice, due to attenuation and fading (depending on the environment), even a decent performing model for one location does not work for another [15] and hence requires adjustment. Furthermore, our measurement data have indicated model imprecision when traversing two different cluttered environments with varying topography. Therefore, we need to fine-tune the selected path loss models to apprehend the loss based on the variation in the propagation channel.

The vast majority of loss reports for a diverse propagation scenario consists of fine-tuning the propagation loss models. Techniques like genetic algorithm optimization method of path loss model are discussed in [16], while the literature like [17] has reported the success of particle swarm over the genetic algorithm optimization of Hata model with improved results. Authors in [18] have presented the success of an RMSE-based tuning method for the Hata path loss model. Notwithstanding these methods, the least-square approximation method is also another indispensable tool for minimizing errors due to its simplistic design, which has been incorporated in a number of studies for the optimization of path loss models [19–22]. In the literature, the majority of the tuning method lean toward the Hata model, and little attention is given to the Egli model. Nonetheless, it is worthwhile to note that in all of the technique mentioned above, paramount concepts are to diminish the errors associated with the model prediction. In this paper, we employ path loss models to consider the terrestrial VHF signal variations of an elevated base antenna propagating over a hilly forested scenario in an outdoor environment. An initial small sample test of measured data [23] has shown that Egli, Hata, and Perez perform decently. We focus this paper on correcting the model prediction for the different scenarios using the least-square regression fitting method. We estimated the loss using the Egli model [24], Hata model [25], and Perez Vega et al. model [26] using the parameters specified by our study area in Sect. 2. We performed the least-square regression analysis for the Egli and Hata model and applied a modification for the two models based on this analysis in Sect. 4. We learned that though expressed differently, the optimized model of Egli

and Hata model predicts the equal loss with improved performance and we evaluated these findings by performing *R*-squared (r^2) test for all the models. This study will help in tuning existing propagation models using the least-square approximation method for different scenarios over a large signal coverage area.

2 Path loss models

2.1 Egli loss model

To understand the characteristics of electromagnetic propagation over irregular terrain, Egli [27] developed graphs and correcting curves based on a statistical analysis of wide-range measurements made between 40 and 1000 MHz in a different region of the USA. This model fundamentally is the theoretical plane earth propagation model with added terrain factor (dependent on frequency) and an evaluation of position [28]. Some of the literature works [24, 29, 30] provide a distinctive mathematical expression of Egli's graphs, which approximately gives the same result when compared. Egli model is expressed in the form [24]:

$$L_p(\text{dB}) = 20 \log_{10}(f_c) + 40 \log_{10}(R) - 20 \log_{10}(h_b) + \begin{cases} (76.3 - 10 \log_{10}(h_m)) & h_m < 10 \\ (85.9 - 20 \log_{10}(h_m)) & h_m \geq 10 \end{cases} \quad (1)$$

f_c is the frequency (MHz), h_b is the base antenna (Tx) height in meter (m), h_m is mobile antenna (Mx) height (m), and R is Tx and Mx separation in kilometers (km).

2.2 Okumura–Hata loss model

The first comprehensive field strength measurement was done in 1968 by Okumura et al. for the Japanese environments [31]. The Okumura prediction method is one of the most referenced models applicable within the frequency – 150 to 2000 MHz. The underlying principle of the model is that path loss consists of free space loss plus the attenuation factor relative to this open space, which is a function of distance and frequency [32]. The model provides sets of curves [30, 33, 34] for estimating the propagation losses. The Okumura method requires complex graphical interpretations to determine the propagation loss. To reduce this complication, Hata [25] developed a model from Okumura's curves applicable between 150 and 1500 MHz. Hata model offers three terrain-based mathematical expressions with urban area propagation loss as the standard expression in each terrain type and also provided a correction factor to account for the different terrain types.

$$L_u(\text{dB}) = 69.55 + 26.16 \log_{10} f_c + (44.9 - 6.55 \log_{10} h_t) \log_{10} R - 13.82 \log_{10} h_t - \underbrace{(1.1 \log_{10} f_c - 0.7) h_m - (1.56 \log_{10} f_c - 0.8)}_{\text{medium-small city correction}} \quad (2)$$

$$L_{\text{suburb}}(\text{dB}) = L_u - 2 \log_{10}(f_c/28)^2 - 5.4$$

$$L_{\text{open}}(\text{dB}) = L_u - 4.78(\log_{10}(f_c))^2 + 18.33 \log_{10}(f_c) - 40.94 \quad (3)$$

where f_c is the frequency (MHz), h_t is the Tx height (m), h_m is the Mx height (1 m to 10 m), R is Tx and Mx separation (1 km to 20 km) and the 'medium-small city correction' factor is the mobile antenna height correction for the concerned area.

In (2) and (3) L_u is the urban area propagation loss with medium-small area correction, L_{suburb} and L_{open} are the respective Suburban and Open area loss. Comparisons of the three area correction factors show insignificant differences for 203.25 MHz frequency and 1.8-m mobile antenna, and the Suburban and the Open area model with medium-small city correction gives the best fit prediction with our measured data.

2.3 Perez Vega and Zamanillo loss model

In theoretical and experimental-based studies [31, 33–35] there are clear indications of the logarithmic variation of the average received signal with distance.

$$L_p(\text{dB}) = 10n \log_{10}(R) + L_0 \quad (4)$$

R is the Tx and Mx separation (m), L_0 is the propagation loss at 1 m in free space, and n is the path loss exponent.

This relationship is one consistent model for significant area propagation path loss where the power could vary as a function of the transmission factors. Based on this Perez-Vega [26] developed a simple propagation model for estimating the average propagation loss in the VHF and UHF band where the exponent of distance n is extracted from the $F(50, 50)$ graph and given in SI unit as graphs [32].

2.4 Signal loss

To calculate the loss (dB) associated with our measured signal strength ($\text{dB}\mu\text{V}/\text{m}$) we employ the relation given in [25].

$$L(\text{dB}) = P_t - E_f - 10 \log_{10}(\lambda^2/4\pi) + 145.8 \quad (5)$$

Where P_t is the power transmitted (dBw) and E_f is the received signal strength ($\text{dB}\mu\text{V}/\text{m}$), and λ is the wavelength of the 203.25 MHz signal.

3 Signal measurement survey

3.1 Study area

We performed signal measurements in Aizawl, Mizoram India. Mizoram is one of the North-Eastern states of India comprising of rugged valleys and steep hills with dense forest cover, which accounts for 86.27 % of its total area. The base antenna (Tx) location is Aizawl, the capital of Mizoram (Fig. 1). The measurement site topography of the North (Kolasib) and the South (Thenzawl) is a hilly terrain region with a blend of both open and moderate-dense forest. For both the northern and southern part, forest cover ranges from low to moderate within an approximate 10 km away from Tx. Beyond 10 km, we have moderate-heavy forest cover in the majority of the sites with low forest cover in some locations. Both forest ecosystem appears almost the same; however, the topology of both measurement direction (North and South) is entirely different.

3.2 Signal strength measurement

In Aizawl Mizoram, there is only two frequency 182.25 MHz and 203.25 MHz in the Band III channel propagating across the city. At the time of the survey, we measured a continuous wave propagation for 203.25 MHz only as there are irregularities in the 182.25 MHz measurements. Table 1 provides the experimental parameters of the survey details.

We noted down the coordinates of Tx location and using Google Map. We mark points at a regular interval of an approximate 2 km aerial distance away from Tx. We navigate to these points using the Google navigation system, available in Google map application, and in each of these sites, different sets of signal strength measurements were taken using a spectrum analyzer (Anritsu Sitemaster) coupled with the standard dipole antenna. We mount the dipole antenna on a tripod stand on clear ground and set the dipole antenna length accordingly to match our

Table 1 Experimental details

Parameters	Model, scale, and units
Spectrum analyzer	Anritsu Sitemaster S332E
Dipole antenna (mobile antenna)	Anritsu MP543B
Base antenna (Tx) height	40 m above ground
Mobile antenna (Mx) height	1.8 m above ground
Frequency measured	203.25 MHz
Power transmitted	40 dBw
Measurement sites	32 sites (North) and 28 sites (South)

Fig. 1 Map showing the location of Aizawl, Mizoram (North-East India), and the extended view of the North (Kolasib) and South (Thenzawl) direction with each measurement sites

measured frequency (203.25 MHz). Following this procedure, we took signal strength measurements (in $\text{dB}\mu\text{V}/\text{m}$) at 32 sites (64 km aerial distance approximate) and 28 places (56 km aerial distance approximate) for the Northern and Southern regions of Tx, respectively, as shown in Fig. 1.

4 Methodology

4.1 Graphical analysis

A scatter plot provides a useful insight into the data [36]. After evaluating the loss associated with different sets of signal measurements taken at each point using (5), we obtained a scatter plot of this as a function of transmitter (Tx)–receiver (Mx) separation for both directions. We observed that the southern data showed a higher loss, as shown in Fig. 2, and that both the losses showed a positive correlation with the Tx–Mx separation. From the averaged loss, we calculated a linear fit for both. We noted that the

Fig. 2 Plot showing the loss dataset as a function of Tx–Mx separation (km) for the two directions

fit lines are almost parallel to one another, which signifies that the signal propagation loss, although different, shows a similar trend.

4.2 Least-square optimization

Different studies have widely referenced the Hata model and performed optimization for different region. In our study, we presented a correction to the Egli and the Hata model based on least-square regression analysis, while we neglected the Perez Vega model as there is no simple optimization process.

We express the error between the measured field strength and prediction made by the model as

$$\text{Erro}(\epsilon) = \text{Measured}(Y_i) - \text{Predicted}(Z_i). \quad (6)$$

Using a basic regression analogy, the predictor variable X_i and the dependent variable Z_i (Predicted Model) are related as

$$(Z_i) = (A + BX_i) \quad (7)$$

where A and B are the regression parameters. We can subsequently write the errors in the form

$$\begin{aligned} \sum_{i=1}^n \epsilon^2 &= \sum_{i=1}^n (Y_i - (A + BX_i))^2 \\ \sum_{i=1}^n \epsilon^2 &= \sum_{i=1}^n (Y_i - A - BX_i)^2 \end{aligned} \quad (8)$$

According to the least-square method, the estimators A and B are those respective values of α and β , which minimizes the error ϵ for the given observation sample [37]. The minimum value of ϵ (minimum error between the measured and predicted model) can be obtained by equating the first-order derivative of ϵ concerning for α and β to zero.

$$\begin{aligned} \frac{\partial}{\partial \alpha} \sum_{i=0}^n \epsilon^2 &= 2 \sum_{i=0}^n (Y_i - \alpha - \beta X_i)(-1) = 0 \\ - \sum_{i=0}^n Y_i + \sum_{i=0}^n \alpha + \beta \sum_{i=0}^n X_i &= 0 \end{aligned} \quad (9)$$

$$\begin{aligned} \sum_{i=0}^n Y_i &= n\alpha + \beta \sum_{i=0}^n X_i \\ \text{where } n\alpha &= \sum_{i=0}^n \alpha \end{aligned} \quad (10)$$

Again taking derivatives with respect to the β term and equating it to zero, we obtain:

$$\frac{\partial}{\partial \beta} \sum_{i=0}^n \epsilon^2 = 2 \sum_{i=0}^n (Y_i - \alpha - \beta X_i)(-X_i) \quad (11)$$

$$\sum_{i=0}^n X_i Y_i = \alpha \sum_{i=0}^n X_i + \beta \sum_{i=0}^n X_i^2 \quad (12)$$

(10) and (12) are called normal equations, and α and β are the respective point estimators of A and B .

$$\begin{aligned} \sum_{i=0}^n Y_i &= n\alpha + \beta \sum_{i=0}^n X_i \\ \sum_{i=0}^n X_i Y_i &= \alpha \sum_{i=0}^n X_i + \beta \sum_{i=0}^n X_i^2 \end{aligned} \quad (13)$$

Rearranging (10) and (12), we can write them as a set of equation (13) and solve the α and β associated using a matrix method.

$$\begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} \quad (14)$$

$$\alpha = \frac{\sum Y_i - \beta \sum X_i}{n} \quad (15)$$

$$\beta = \frac{n \sum (X_i Y_i) - \sum X_i \sum Y_i}{n \sum (X_i)^2 - (\sum X_i)^2} \quad (16)$$

In the predicted value (Z_i) if we assume α to be constant and β as the variable that is changing with some parameters, then optimization of the models could be made. Based on the methods discussed, we corrected the errors associated with Egli and Hata models.

4.3 Egli model optimization

Egli model for ($h_m < 10$ m) is

$$\begin{aligned} L_{\text{Egli}}(\text{dB}) &= 20 \log_{10}(f_c) - 20 \log_{10}(h_b) + 76.3 \\ &\quad - 10 \log_{10}(h_m) + 40 \log_{10}(R) \end{aligned} \quad (17)$$

Separating the term for the intercept and the slope, we can write

$$L_{\text{Egli}} = E_a + E_b X_i \quad (18)$$

Here X_i term is the ($\log_{10} R$) term, where

$$\begin{aligned} E_a &= 20 \log_{10}(f_c) - 20 \log_{10}(h_b) + 76.3 - 10 \log_{10}(h_m) \\ E_b &= 40 \end{aligned} \quad (19)$$

Comparing this with (7) in α and β term, we can write

$$\begin{aligned} L_{\text{Egli}} &= E_a + E_b \log_{10} R \\ L_{\text{Egli}}(Z_E) &= \alpha + \beta X_i \end{aligned} \quad (20)$$

$$L_{\text{Egli}}(Z_E) = \underbrace{(E_a \pm C_\alpha)}_{\alpha} + \underbrace{(E_b \pm C_\beta)}_{\beta} \underbrace{(\log_{10} R)}_{X_i} \quad (21)$$

where C_α and C_β are the respective correction needed for E_a and E_b

Comparing with the normal equations (10,12)

$$\begin{aligned} X_i &= \log_{10} R \\ Y_i &= L_m \text{ (Measured loss)} \end{aligned} \quad (22)$$

and substituting in (15) and (16), we obtain:

$$\alpha = \frac{\sum L_m - \beta \sum (\log_{10} R)}{n} \quad (23)$$

$$\beta = \frac{n \sum (L_m \log_{10} R) - \sum (L_m) \sum (\log_{10} R)}{n \sum (\log_{10} R)^2 - (\sum \log_{10} R)^2} \quad (24)$$

When we solve (23) and (24) (i.e., the α and β equation), a condition arises that $E_a \pm C_\alpha$ cannot be higher than α and $E_b \pm C_\beta$ cannot be higher than β , i.e.,

$$\begin{aligned} E_a \pm C_\alpha &< \alpha \\ E_b \pm C_\beta &< \beta \end{aligned} \quad (25)$$

4.4 Hata model optimization

In the Hata model, we have used the Suburban area loss model for the Southern region and the Open area model for the North. In both of these, we have used 'medium-small city' correction factors. Therefore, we estimated a separate fit model for the North and Southern regions in the case of Hata model analysis. To find the fit for the Hata Suburban and Open area we rearranged (3) as

$$\begin{aligned} H_{a(\text{Suburb})} &= 64.15 + 26.16 \log_{10} f_c - [(1.1 \log_{10} f_c - 0.7)h_m \\ &\quad - (1.56 \log_{10} f_c - 0.8)] - 2 \log_{10}(f_c/28)^2 - 13.8 \log_{10} h_b \\ H_{a(\text{Open})} &= 29.05 + 26.16 \log_{10} f_c - [(1.1 \log_{10} f_c - 0.7)h_m \\ &\quad - (1.56 \log_{10} f_c - 0.8)] - 4.78(\log_{10}(f_c))^2 + 18.33 \log_{10}(f_c) \\ &\quad - 13.8 \log_{10} h_b \end{aligned} \quad (26)$$

$$H_b = (44.9 - 6.55 \log_{10} h_b) \quad (27)$$

$$\begin{aligned} L_{\text{Hata}} &= H_a + H_b \log_{10} R \\ L_{\text{Hata}}(Z_H) &= \alpha + \beta X_i \end{aligned} \quad (28)$$

$$L_{\text{Hata}}(Z_H) = \underbrace{(H_a \pm C_\alpha)}_{\alpha} + \underbrace{(H_b \pm C_\beta)}_{\beta} \underbrace{(\log_{10} R)}_{X_i} \quad (29)$$

where C_α and C_β are the respective correction needed for H_a and H_b . Following the Egli model optimization concept, we evaluated the fit for the Hata Suburban and Open area model separately. As in Egli calculations, we have

$$\alpha = \frac{\sum L_m - \beta \sum (\log_{10} R)}{n} \quad (30)$$

$$\beta = \frac{n \sum (L_m \log_{10} R) - \sum (L_m) \sum (\log_{10} R)}{n \sum (\log_{10} R)^2 - (\sum \log_{10} R)^2} \quad (31)$$

The condition of α and β as we ensured in the Egli model (25) correspondingly applies here, and the values of α and β are calculated separately for both directions.

4.5 Statistical evaluation

To compare the measured loss with the optimized model and the models presented, we used the coefficient of determination (r^2) for the evaluation of each model's performance. We have to note that the models shown are logarithmic, and we cannot evaluate the correlation coefficient, which identifies the strength of linear association [38] between the models and the measured data, without transforming the variables. There are different transformation methods to linearize a dataset [39], but a data transformation can introduce a bias leading to erroneous reports [40]. Therefore, instead we used the modified coefficient of determination for evaluating the nonlinear dataset which is given by the relation [41, 42]

$$r^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \quad (32)$$

$$\begin{aligned} SS_{\text{res}} &= \sum (\kappa - \hat{\kappa})^2 \\ SS_{\text{total}} &= \sum (\kappa - \bar{\kappa})^2 \end{aligned} \quad (33)$$

where $\hat{\kappa}$ and $\bar{\kappa}$ represent the predicted data and the mean of the measured loss data, respectively, while κ represents the measured loss data. The result r^2 analysis is given in Table 2, the value of r^2 ranges from 0 to 1, where a value closer to 1 suggests a stronger relationship.

5 Result

Based on the least-square fitting method discussed, we calculated α , β , C_α , and C_β in the Egli and Hata model individually for the North and South, as shown in Table 3.

Table 2 Coefficient of determination (r^2) of the Egli, Perez, Hata models (L_u , L_{suburb} , L_{open}), and the optimized model (Egli or Hata) for the North and the South

r^2	Egli	Perez	L_u	L_{suburb}	L_{open}	Optimized
North	0.44	0.34	-0.04	0.36	0.04	0.47
South	0.36	0.02	0.49	0.52	-0.67	0.55

5.1 Optimized Egli and Hata model

As explained in Sect. 4.4, we adjusted the Hata Suburban area model for the Southern area and the Hata Open area model for the Northern region of our study area while we used the same model for the North and the South in the Egli model optimization.

Using the correction factors calculated and provided in Table 3, we estimated the least-square fitted Egli and Hata model, respectively, for both North and South.

Optimized model for the North

$$L_{\text{hata open}} = 42.16 + 44.49 \log_{10}(f_c) - 13.82 \log_{10}(h_b) - \gamma - 4.78(\log_{10}(f_c))^2 + (43.47 - 6.55 \log_{10}(h_b)) \log_{10} R \quad (34)$$

$$L_{\text{Egli}} = 85.16 + 20 \log_{10}(f_c) - 20 \log_{10}(h_b) - 10 \log_{10}(h_m) + 32.98 \log_{10} R \quad (35)$$

Optimized model for the South

$$L_{\text{hata suburban}} = 68.55 + 26.16 \log_{10}(f_c) - 13.82 \log_{10}(h_b) - \gamma - 2(\log_{10}(f_c/28))^2 + (47.65 - 6.55 \log_{10}(h_b)) \log_{10} R \quad (36)$$

$$L_{\text{Egli}} = 87.83 + 20 \log_{10}(f_c) - 20 \log_{10}(h_b) - 10 \log_{10}(h_m) + 37.161 \log_{10} R \quad (37)$$

where $\gamma = (1.1 \log_{10}(f_c) - 0.7)h_m - (1.56 \log_{10}(f_c) - 0.8)$ is the Hata model medium-small city correction factor

In arrears to the same least-square correction method employed, we observed that, although Egli and the Hata model expressions are entirely different from one another, they provide equal prediction value, as presented in Fig. 3. This equality attributes from the fact that we used the same least-square correction factor (α and β) for the Egli and Hata models w.r.t the same measurement loss data. We denote this Egli and Hata corrected model as the Optimized model. Although we did not obtain the plot, to further assess the performance of the Optimized model, we conduct an r^2 test for all the Hata model (open area, suburban area, and an urban area) for North and South w.r.t to the loss data. A plot of the optimized model compared

Fig. 3 Plot of Egli and Hata optimized loss estimate. Here straight line indicates equal loss estimated from the two different corrected model

Fig. 4 Plot showing comparison of Egli, Hata, Perez, and the optimized Egli/Hata model for the North. Optimized fit line $r^2 = 0.47$

with the measured loss, Egli, Hata, and Perez model is shown in Figs. 4 and 5.

Our results showed that the optimized model of Egli and Hata showed an equal prediction, and therefore, we present a single optimized model for both the models.

Fig. 5 Plot showing comparison of Egli, Hata, Perez, and the optimized Egli/Hata model for the South. Optimized fit line $r^2 = 0.55$

The coefficient of determination (r^2) showed that 47% of the loss data accounted for by the optimized model in the North while accounting for 55% of the loss data in the South. Thus, the rest 53% and 45% of the respective North and Southern region is still unexplained. Since the propagation models presented do not consider the environmental factor, we could attribute this to the fading, scattering, diffraction, and reflection [43] that the radio signal tends to suffer along the propagation channel. The original models with Perez model presented a lower performance compared to this tuned model. This coefficient of determination (r^2) is the explanatory power of the model, and from this analysis, we learned that the model before and after optimization has higher explanatory power compared to their original models.

6 Discussion

The global analog-to-digital television migration practice has led many countries to look into the cognitive radio access or the so-called white space vacancy opportunity in the TV's VHF and UHF band [7–9]. With the rising demand for bandwidth, there is a prospect of utilizing this VHF spectrum as this band is underutilized across India [4]. Though wireless technology and optical fiber technology can coexist, the idea of wireless services is preferable to

high-speed optical fiber services in terms of the cost and the location settings. Furthermore, taking into account the coverage area and the cost-effectiveness, lower frequency transmission could still play an important role, parallel to those high-frequency transmissions.

From the model analysis, we learned that the empirical irregular terrain Egli model [27] and quasi smooth terrain Hata model [25] gives a different extrapolation since both are modeled from different measurement regions (the USA and Japan). Both models acquire the probability of prediction through the fittings of the specific-region measurement data [44]. Although the Hata model provides three different area prediction curves, both models did not incorporate the impact of geodata factors explicitly. As such, our data taken in the hilly forested region contrast with these empirical reproductions to a certain extent. From this examination, we learned the importance of considering the impact of geodata to get a complete picture of signal propagation loss.

In wireless communication, to model a loss, one requires a vast amount of data and their dependence on different propagation parameters (frequency, Tx and Mx height, transmitted power, gain, etc.). And the generalization of a loss model pertinent to every region would require exploiting the effect of the different region-specific channel influencing factors like terrain diffraction, line of sight, vegetation attenuation, etc., to be more deterministic [45]. This method, however accurate, is again inefficient. In this light, tuning of the existing empirical loss model tends to be one practical approach alternative to modeling. It is worthwhile mentioning that the optimized model does not pose to be better than the original models. Instead, it is the same model with the least-square approximated minimum errors-based model to consider the additional loss or gain due to the differences in the attenuation of each measurement region.

Our study aims at understanding the VHF signal variation across the hilly terrain region of one of the North-Eastern states of India and test the existing path loss models. We learned that our study area propagation channel differs from the selected empirical model measurement region. Furthermore, we learned that the least-square optimization correction method is a simple and effective method for loss model tuning. This optimized model is based on an actual in-site measurement survey, with a minimized error. Therefore, the adjusted model could be

Table 3 Least-squared fit correction factors for Egli and Hata model

	Egli				Hata			
	α	β	C_α	C_β	α	β	C_α	C_β
North	96.73	32.98	8.86	− 7.02	96.73	32.98	13.55	− 1.42
South	99.40	37.16	11.53	− 2.84	99.40	37.16	− 0.10	2.75

used as an optimal prediction model for future loss estimation across the study area or other areas with similar topographical backdrops. This will benefit broadcasters in the proper forecasting, performance analysis, and the deployment of wireless service across the region.

7 Conclusion

In this paper, we use Egli and Hata model to consider the terrestrial VHF signal variations of an elevated base antenna propagating over a hilly forested scenario in an outdoor environment. An initial small sample test of measured data has shown that Egli, Hata, and Perez perform decently for our study area. Our primary focus is on correcting the model prediction variance for the different scenarios using the least-square regression fitting method and obtaining the best-fitted model for future signal prediction across the study area or areas with similar topography. Based on this least-square regression analysis, we modified the Egli and Hata original model and presented a correction factor for the respective North and South composed of different topography. We observed that the optimized Egli and Hata model predicts the same result although expressed differently. A coefficient of determination (r^2) analysis showed that the optimized model explained 47% of loss variation for the North and 55% for the South. Finally, we presented a separate optimized model for the Northern and Southern regions. This method of model tuning is not region-specific or model-specific. However, we need to recalculate the correction coefficients for different areas composed of changeable attenuation parameters to obtain a minimal error for a given expanse.

Acknowledgements The authors thanked the Ministry of Tribal Affairs for their financial support and acknowledged the assistance of Matthias Hneisa Jawhly, Ravindra Naithani, KC Jain, and Doordarshan staffs, Aizawl Mizoram, for their aid and support during the measurement campaign. This research was supported by National Fellowship and Scholarship for Higher Education of ST Students (NFST) (Award No: 201718-NFST-MIZ-00476), Ministry of Tribal Affairs, Government of India.

Funding This research was funded by National Fellowship and Scholarship for Higher Education of ST Students (NFST) (Award No: 201718-NFST-MIZ-00476), Ministry of Tribal Affairs, Government of India. A PhD fellowship was awarded to the first author.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

Ethical standard This study does not involve any human or animal participant.

Informed consent Not applicable as the study does not involve any human or animal participant.

References

1. Rahman M, Saifullah A (2019) A comprehensive survey on networking over TV white spaces. *Pervasive Mobile Comput* 59:101072
2. Dagefu FT, Verma G, Rao CR, Yu PL, Fink JR, Sadler BM, Sarabandi K (2015) Short-range low-VHF channel characterization in cluttered environments. *IEEE Trans Antennas Propag*. <https://doi.org/10.1109/TAP.2015.2418346>
3. Prasad NC, Deb S, Karandikar A (2016) In: IEEE international symposium on personal, indoor and mobile radio communications. <https://doi.org/10.1109/PIMRC.2016.7794856>
4. Kumar A, Karandikar A, Naik G, Khaturia M, Saha S, Arora M, Singh J (2016) Toward enabling broadband for a billion plus population with TV white spaces. *IEEE Commun Mag*. <https://doi.org/10.1109/MCOM.2016.7509375>
5. Agarwal A, Sengar AS, Gangopadhyay R, Debnath S (2016) In: International conference on wireless communications, signal processing and networking (WiSPNET). IEEE, pp 2035–2039
6. Chavez A, Littman-Quinn R, Ndlovu K, Kovarik CL (2016) Using TV white space spectrum to practise telemedicine: a promising technology to enhance broadband internet connectivity within healthcare facilities in rural regions of developing countries. *J Telemed Telecare*. <https://doi.org/10.1177/1357633X15595324>
7. Nekovee M (2010) In: IEEE symposium on new frontiers in dynamic spectrum. DySPAN. <https://doi.org/10.1109/DYSPA.2010.5457890>
8. Kumar P, Rakheja N, Sarswat A, Varshney H, Bhatia P, Goli SR, Ribeiro VJ, Sharma M (2013) In: 2013 IEEE 14th international symposium on a world of wireless, mobile and multimedia networks. WoWMoM. <https://doi.org/10.1109/WoWMoM.2013.6583462>
9. Naik G, Singhal S, Kumar A, Karandikar A (2014) In: Twentieth national conference on communications (NCC). <https://doi.org/10.1109/NCC.2014.6811306>
10. Zhao Z, Vuran MC, Batur D, Ekici E (2019) Shades of white: impacts of population dynamics and tv viewership on available tv spectrum. *IEEE Trans Veh Technol*. <https://doi.org/10.1109/TVT.2019.2892867>
11. Khamlichi BE, Abdelaali C, Ahmed L, Abbadi JE (2016) In: SITA 2016—11th international conference on intelligent systems: theories and applications. <https://doi.org/10.1109/SITA.2016.7772293>
12. Elshafie HEA, Faisal N, Syed-Yusof SK, Mohamad H, Ramli NB, Jayavalan S, Hassan WA, Zubair S (2015) VHF band utilization measurement for cognitive radio application in Malaysia. *Wirel Pers Commun* 85(4):2727
13. Ufoaroh SU, Abu KR (2019) Assessment of TV white spaces availability in southern Nigeria (A case study of Ugbowo, Benin City). *Electrosc J* 10(10):22
14. Kumar S (2019) Global digital cultures: perspectives from south asia. University of Michigan Press, Ann Arbor, p 53
15. Hoomod HK, Al-Mejibli I, Jabboory AI (2018) Analyzing study of path loss propagation models in wireless communications at 0.8 GHz. *J Phys Conf Ser*. <https://doi.org/10.1088/1742-6596/1003/1/012028>
16. Garah M, Djouane L, Oudira H, Hamdiken N (2016) Path loss models optimization for mobile communication in different areas. *Indones J Electric Eng Comput Sci*. <https://doi.org/10.11591/ijeecs.v3.i1.pp126-135>

17. Banimelhem O, Al-Zu'bi MM, Al Salameh MS (2015) In: Proceedings—15th IEEE international conference on computer and information technology, CIT 2015, 14th IEEE international conference on ubiquitous computing and communications, IUCC 2015, 13th IEEE international conference on dependable, autonomous and se. <https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.248>
18. Udofia KM, Friday N, Jimoh AJ (2016) Okumura–Hata propagation model tuning through composite function of prediction residual. *Math Softw Eng* 2(2):93
19. Cota N, Serrador A, Vieira P, Beire AR, Rodrigues A (2017) On the use of Okumura–Hata propagation model on railway communications. *Wirel Person Commun*. <https://doi.org/10.1007/s11277-014-2224-y>
20. Saidi R, Saidi L, Regai ZEA (2017) Contribution to the performance of mobile radio systems by optimizing the Okumura–Hata model by linear regression: application to the city of Annaba in Algeria. *Int J Electric Eng Inf*. <https://doi.org/10.15676/ijeei.2017.9.4.3>
21. Mardeni R, Kwan KF (2010) Optimization of Hata propagation prediction model in suburban area in Malaysia. *Progr Electromag Res C*. <https://doi.org/10.2528/PIERC10011804>
22. Faruk N, Popoola SI, Surajudeen-Bakinde NT, Oloyede AA, Abdulkarim A, Olawoyin LA, Ali M, Calafate CT, Atayero AA (2019) Path loss predictions in the VHF and UHF bands within urban environments: experimental investigation of empirical, heuristics and geospatial models. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2019.2921411>
23. Jawhly T, Tiwari RC (2019) Lecture notes in electrical engineering. Springer, New York, pp 53–63. <https://doi.org/10.1007/978-981-13-1642-5>
24. Delisle G, Lefevre JP, Lecours M, Chouinard JY (1985) Propagation loss prediction: a comparative study with application to the mobile radio channel. *IEEE Trans Veh Technol*. <https://doi.org/10.1109/T-VT.1985.24041>
25. Hata M (1980) Empirical formula for propagation loss in land mobile radio services. *IEEE Trans Veh Technol*. <https://doi.org/10.1109/T-VT.1980.23859>
26. Pérez-Vega C, Zamanillo JM (2002) Path-loss model for broadcasting applications and outdoor communication systems in the VHF and UHF bands. *IEEE Trans Broadcast*. <https://doi.org/10.1109/TBC.2002.1021273>
27. Egli JJ (1957) Radio propagation above 40 MC over irregular terrain. *Proc IRE*. <https://doi.org/10.1109/JRPROC.1957.278224>
28. Longley A (1978) In: 28th IEEE vehicular technology conference. <https://doi.org/10.1109/VTC.1978.1622591>
29. Shepherd NH (1988) Coverage prediction for mobile radio systems operating in the 800/900 MHz frequency range. *IEEE Trans Veh Technol* 37(1):3
30. Schmid HF (1970) A prediction model for multipath propagation of pulse signals at VHF and UHF over irregular terrain. *IEEE Trans Antennas Propag*. <https://doi.org/10.1109/TAP.1970.1139661>
31. Andersen JB, Rappaport TS, Yoshida S (1995) Propagation measurements and models for wireless communications channels. *IEEE Commun Mag*. <https://doi.org/10.1109/35.339880>
32. Salous S (2013) Radio propagation measurement and channel modelling. Wiley, New York. <https://doi.org/10.1002/9781118502280>
33. Erceg V, Greenstein LJ, Tjandra SY, Parkoff SR, Gupta A, Kulic B, Julius AA, Bianchi R (1999) An empirically based path loss model for wireless channels in suburban environments. *IEEE J Select Areas Commun*. <https://doi.org/10.1109/49.778178>
34. Otermat DT, Kostanic I, Otero CE (2016) Analysis of the FM radio spectrum for secondary licensing of low-power short-range cognitive internet of things devices. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2016.2616113>
35. Karttunen A, Gustafson C, Molisch AF, Wang R, Hur S, Zhang J, Park J (2016) Path loss models with distance-dependent weighted fitting and estimation of censored path loss data. *Microw IET Antennas Propag*. <https://doi.org/10.1049/iet-map.2016.0042>
36. Han J, Pei J, Kamber M (2011) Data mining: concepts and techniques. Elsevier, Amsterdam
37. Kutner MH, Nachtsheim CJ, Neter J, Li W (1996) Applied linear statistical models. *J R Stat Soc Ser A (General)*. <https://doi.org/10.2307/2984653>
38. Kasuya E (2019) On the use of r and r squared in correlation and regression. *Ecol Res* 34(1):235. <https://doi.org/10.1111/1440-1703.1011>
39. Kantardzic M (2011) Data mining: concepts, models, methods, and algorithms. Wiley, New York
40. Galetto M, Verna E, Genta G (2020) Accurate estimation of prediction models for operator-induced defects in assembly manufacturing processes. *Qual Eng*. <https://doi.org/10.1080/08982112.2019.1700274>
41. Kvalseth TO (1983) Note on the R^2 measure of goodness of fit for nonlinear models. *Bull Psychon Soc* 21(1):79. <https://doi.org/10.3758/BF03329960>
42. Cornell JA (1987) Factors that influence the value of the coefficient of determination in simple linear and nonlinear regression models. *Phytopathology* 77(1):63. <https://doi.org/10.1094/phyto-77-63>
43. Liaskos C, Nie S, Tsioliaridou A, Pitsillides A, Ioannidis S, Akyildiz I (2018) A new wireless communication paradigm through software-controlled metasurfaces. *IEEE Commun Mag* 56(9):162
44. Cui Z, Guan K, Briso-Rodríguez C, Ai B, Zhong Z (2019) A unified design of massive access for cellular internet of things. *IEEE Internet Things J* 6:3934–3937
45. Zhang J, Gentile C, Garey W (2019) On the cross-application of calibrated pathloss models using area features: finding a way to determine similarity between areas. *IEEE Antennas Propag Mag* 62(1):40

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.