https://doi.org/10.5755/j02.eie.36479 ELEKTRONIKA IR ELEKTROTECHNIKA, ISSN 1392-1215, VOL. 30, NO. 2, 2024

Weather-Based Nonlinear Regressions for Digital TV Received Signal Strength Prediction

Ivana Stefanovic[1,*], Marija Malnar[2], Snezana Mladenovic[2], Milutin Nesic[1]
[1]*School of Electrical and Computer Engineering, Academy of Technical and Art Applied Studies, Vojvode Stepe 283, 11000 Belgrade, Serbia*
[2]*Faculty of Transport and Traffic Engineering, University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia*
[*]*ivanas@gs.viser.edu.rs; m.malnar@sf.bg.ac.rs; snezanam@sf.bg.ac.rs; nesic@gs.viser.edu.rs*

***Abstract*—In this research, the impact of various weather conditions on digital television signals is investigated. Machine learning and nonlinear regression models were used to estimate the strength of the received signal. The received signal strength might vary significantly depending on the weather condition, especially in higher frequency ranges or millimetre wavelengths. Predictive analysis was performed for the radio-relay link Aval Tower-Vršac Hill, which is used for the distribution of television and radio programmes by the public company *Broadcasting Technology and Connections* in Serbia. The prediction was made using temperature, temperature index, relative humidity, and received signal strength data for the months of June, July, and August in 2022. The best results were obtained using the *RandomForest* model. Extreme variations in the strength of the received signal can be predicted by using the model mentioned above. More effective management of the broadcasting infrastructure can be done with the ability to predict sudden falls and fluctuations in received signal strength.**

***Index Terms*—Digital TV; Machine learning; Prediction methods; Propagation; Regression analysis.**

I. INTRODUCTION

Compared to analogue television, digital television has significantly increased the quality and efficiency of television transmission. The transition from analogue to digital TV has resulted in various benefits, including improved picture and sound quality, the introduction of interactive services, higher efficiency, and flexibility. Digital TV also utilises a more effective use of the radio frequency spectrum, allowing the broadcast of more channels and services with the same amount of bandwidth. The digitalisation process in Serbia was completed in June 2015. Today, the digital terrestrial television network is based on a second generation digital terrestrial television broadcasting (DVB-T2) system for transmitting and broadcasting digital television signals. DVB-T2 is capable of transmitting video in a variety of resolutions, including standard definition (SD), high definition (HD), and ultra-high definition (UHD). The public company *Broadcasting Technology and Connections* provides services for encoding, multiplexing, digital transmission, and broadcasting of national, regional, and local television programmes.

To ensure optimal system performance and successful signal reception, the received signal strength (RSS) must exceed the receiver's sensitivity. Extreme decreases in the received signal level can result in link failure, and variations in RSS can impact the quality of service. The topology of the terrain, the type of environment the signal is propagating in, the distance between the transmitter and the receiver, the frequency used, and other factors all affect the strength of the signal at the receiver. The attenuations indicated can be considered constant in the case of radio-relay links, as in the Aval Tower-Vršac Hill radio-relay link covered in this study, since neither the terrain topology nor the type of environment, nor the distance between the transmitter and receiver change. However, the quality of signal transmission is degraded due to large fluctuations in signal strength. Weather conditions are one of the elements that results in fluctuations in RSS, but they are often ignored. Therefore, we set out to show how integrating weather conditions, within the prediction model, can have a substantial impact on radio link quality.

Numerous studies have been carried out to verify the impact of various weather conditions on the strength of the received signal [1]–[8]. The aim of this study is to use machine learning to forecast RSS according to weather conditions. We started our study considering 14 weather-related factors. Temperature, temperature index, and relative humidity were chosen as input variables based on their correlation with RSS. Then, we used various nonlinear regression models to predict the exact RSS value, which is the objective variable. One of the most significant contributions of this study is that the predictive model we created is capable of predicting sudden decreases in the level of RSS. This method may be beneficial for predicting RSS levels using weather forecasts rather than just historical weather data. Compared to classification analysis, which is utilised in most works [9]–[11], and which allows the prediction of only the status of the link, we used regression analysis, which allowed us to predict the precise value of RSS. The ability to predict the precise level and sudden drops of RSS, instead of link status, allows for more effective management of the television broadcasting infrastructure through planning of transmission power, redundant radio-relay links, and increased link reliability.

Manuscript received 27 November, 2023; accepted 16 March, 2024.
This research was supported in part by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia.

The rest of this paper is organised as follows. Section II includes works relevant to this subject matter. Section III explains the materials and methods, including the adopted data sets and the evaluation metrics used to evaluate the performance of the prediction models. The most significant findings and discussions are presented in Section IV. Finally, Section V provides a brief summary of the study.

II. BACKGROUND AND RELATED WORK

Different weather conditions can have a substantial impact on the strength of the radio signal in wireless communication systems. Understanding these weather influences is critical to develop reliable communication systems. Engineers and operators must consider potential signal losses in unfavourable weather conditions while designing and optimising radio links for reliable performance to reduce the effect of weather on radio link RSS.

There are many research studies that confirm the effect of numerous weather factors, such as temperature, humidity, rain, snow, and wind, on RSS for different types of communication systems that use different frequency bands. In [1], the authors report a significant effect of humidity, wind speed, rain rate, and temperature on radio signal where the correlation values for each factor are 0.6285, 0.434, 0.3850, and 0.3339, respectively. The study was carried out in Malaysia, while the frequencies were up to 9 GHz. The authors in [2] investigate the effect of temperature on the quality of cellular signal strength. The correlation values ranged from 0.17 to 0.8 for different days and different time intervals. They also state that correlation values were high in 50 % cases when there was a strong inverse linear relation. In [3], the authors study the effect of temperature, humidity, and pressure on RSS of a digital terrestrial television broadcast station (DTTBS) in Katsina City, Nigeria. Also, in [4], they investigated the effects of air temperature, atmospheric pressure, and relative humidity on radio signals from Edo Broadcasting Service (EBS) in Benin City, Nigeria, at 743.25 MHz and concluded that radio signals have an inverse relationship with the meteorological variables mentioned, with a correlation value of -0.94, -0.92, and -0.96. In [5], the authors report a negative correlation between RSS and relative humidity for frequencies 382.5 MHz, 945 MHz, 1867.5 MHz, and 2160 MHz with correlation values ranging from -0.382 to -0.805. The authors in [6] find that the increase in atmospheric temperature will lead to a decrease in the strength of the signal. This study was also carried out in Nigeria for FM radio and TV broadcasting. There are studies, such as in [7], showing that environmental factors such as temperature and humidity have an effect on the calculation of indoor path loss as well.

A large number of studies are available that prove that weather conditions, especially temperature and relative humidity, affect RSS. Some studies indicate a lower or higher degree of correlation between temperature, relative humidity, and RSS. This can be attributed to the fact that studies include different communication systems, frequencies, regions, equipment, etc. This is what we expected when we first began our research. However, we find that some authors support a positive association between the received signal strength and temperature, while others assert a negative correlation. The scenario in which relative humidity affects RSS is the same. In our research, we found a positive correlation between temperature and RSS and a negative correlation between relative humidity and RSS. This makes sense to us because the relationship between temperature and relative humidity is generally inverse. Relative humidity typically decreases with increasing temperature and increases with decreasing temperature.

The authors in [8] also find that the results and conclusions on how atmospheric conditions affect signal strength are contradictory. So, they did the review of temperature and humidity impacts on RF signals to determine what is the relationship between temperature, relative humidity, and signal strength. Based on the results, they concluded that most related works support that temperature has a positive correlation with signal strength, while relative humidity has a negative correlation with signal strength. This is consistent with the results of the study we conducted.

There are other studies that use machine learning to investigate how the weather impacts the performance of the radio link. Using machine learning classification models, most of the works focus on predicting the status of the radio link. The gradient boosting model, with an F1 score of 0.95, is suggested by the authors in [9] for one-day and five-day link status predictions based on weather conditions. In this study, binary classification models were employed, in which the link status takes on the values 0 or 1. For a two-year period, the following conditions were taken into consideration: temperature, wind direction and speed, precipitation and precipitation coefficient, humidity and pressure. The authors in [10] employ multiple classifications in addition to binary classification to forecast the received signal strength indicator (RSSI) parameter based on weather conditions. This method allows for the prediction of the specific range, in which the value of the RSSI parameter is placed, as opposed to binary classification, which predicts whether the value of the RSSI is above or below a particular threshold. In some works, such as in [11], the opposite approach is used, where weather conditions are predicted based on the strength of the received signal. The authors in [11] propose a classification model based on neural networks for the prediction of precipitation (no rain, light rain, moderate amount of rain, and significant amount of rain) based on RSS and other parameters of the radio link.

This paper proposes the use of nonlinear regression, which makes it possible to forecast the precise value of the RSS. A similar method was used in [12], where linear regression is used to forecast the change in RSSI in wireless sensor networks based on temperature and humidity.

III. MATERIALS AND METHODOLOGY

The adopted data sets will be explained in this section, as well as data preprocessing and feature selection. The evaluation metrics used to evaluate the performance of the prediction models will be presented.

A. Data Set

Data preparation included the analysis and preparation of two data sets, a weather data set and a radio link data set. Data preparation was done using the *Python* programming language and the *Google Colab* environment. A unique data set was created for predictive analysis after the preparation and examination of separate data sets.

Weather-related information was obtained from *www.worldweatheronline.com*. The weather data set has an initial size of 4416×32 (2208 instances for each location). This set includes multiple weather conditions for Aval Tower and Vršac Hill in the months of June, July, and August 2022. The dimensions of the data set were reduced to 4416×17 during the data preparation process by eliminating 15 attributes. Since most attribute data are available in multiple units (e.g., temperature data are available in ºC and ºF), certain attributes have been removed. In addition to location, time, and date information, the data set includes the following features that describe weather conditions:

– *TempC*, temperature in ºC, from 13 ºC to 41 ºC;
– *WindSpeedKmph*, wind speed in km/h, from 0 km/h to 28 km/h;
– *WindDirdegree*, wind direction in degrees, value range from 0 º to 360 º;
– *WeatherCode*, text description of the weather, possible attribute values are 113 (clear/sunny), 116 (partly cloudy), 176 (patchy rain nearby), 119 (cloudy), 122 (overcast), 299 (moderate rain at times), 305 (heavy rain at times), and 356 (moderate or heavy rain shower);
– *PrecipMM*, precipitation in mm, from 0 mm to 6.7 mm;
– *Humidity*, relative humidity in %, from 7 % to 97 %;
– *VisibilityKm*, visibility in km, from 2 km to 10 km;
– *PressureHPa*, atmospheric pressure in hPa, from 1003 hPa to 1023 hPa;
– *CloudCover*, cloud cover amount in %, from 0 % to 100 %;
– *HeatIndexC*, heat index temperature in ºC, from 13 ºC to 42 ºC;
– *DewPointC*, dew point temperature in ºC, from 3 ºC to 20 ºC;
– *WindChillC*, wind chill temperature in ºC, from 13 ºC to 41 ºC;
– *WindGustKmph*, wind gust in km/h, from 1 km/h to 41 km/h;
– *uvIndex*, UV index, from 1 to 10.

There are eight predefined values for the *WeatherCode* attribute, each of which corresponds to a textual description of the weather. For example, 2827 hours of clear and sunny weather were recorded, compared to only 89 hours of primarily cloudy weather. The number of attributes in the data set was increased to 24 by encoding the *WeatherCode* attribute and creating dummy variables. The model will be trained using weather data from both sites. Given that those two locations are only slightly more than 80 kilometres apart, a comparison of the weather attributes at each site was made to see if there was a statistically significant difference. To determine whether there was a significant statistical difference between the weather conditions at different locations, statistical tests such as the *t* and *z* tests were used in addition to data visualisation. A 95 % confidence interval was used. The comparison revealed a significant statistical difference in seven attributes: *TempC*, *WindDirdegree*, *HeatIndexC*, *DewPointC*, *WindChillC*, *ModerateRainAtTimes*, and *Overcast*. After preprocessing, the weather related data for both sites were merged, and the data set size is 2208×31.

The second data set contains data on the radio-relay link Aval Tower-Vršac Hill, which is used for the distribution of television and radio programmes by the public company *Broadcasting Technology and Connections*, which is in charge of planning, construction, and maintenance of broadcasting infrastructure on the territory of the Republic of Serbia. Table I provides an overview of the link settings.

TABLE I. LINK SETTINGS OVERVIEW.

Parameters	**Details**
Frequency plan	ITU-R F382-8 on 4 GHz
Frequency	3940.50 MHz
Frequency band type	Licensed
Channel spacing	29 MHz
Modulation scheme	128QAM
Data rate	160 Mbps
Transmitted power	30 dBm
Sensitivity	-73 dBm
Antenna gain	37.10 dBi

The transmitter is Aval Tower and the receiver is Vršac Hill. At 204.68 metres high, the Aval Tower is the most noticeable feature of the *Broadcasting Technology and Connections* infrastructure. It is one of 11 transmitter facilities that broadcast radio and television content in Serbia [13]. The distance between two sites is 79.48 km. Figure 1 shows the topology of the terrain between the two sites.

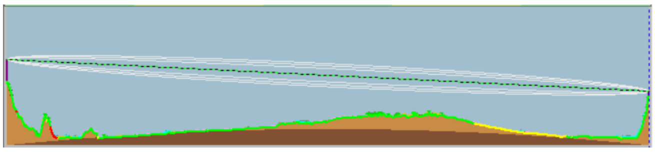

Fig. 1. Aval Tower-Vršac Hill link path.

Table II provides information about coordinates for both sites, altitude, and antenna height. The ANDREW HPX8-36 antenna, with a gain of 37.10 dBi, was used for both the transmitter and receiver. All of the radio-relay link relevant data we got from the public company *Broadcasting Technology and Connections*, for research purposes.

TABLE II. ADDITIONAL LINK SETTINGS DATA.

Parameters	**Site 1 (Aval Tower)**	**Site 2 (Vršac Hill)**
Latitude	44º 41' 45.660"	45º 07' 23.540"
Longitude	20º 30' 52.350"	21º 19' 26.390"
Altitude	439 m	338 m
Antenna height	126 m	15 m

When calculating the attenuation of the signal, the free space path loss is the significant factor. The Friis equation, defined by (1), is used to determine the free space path loss

$$FSL=\left(\frac{4\pi fd}{c}\right)^2, \tag{1}$$

where *d* is the distance between the transmitter and receiver, *f* is the frequency, and *c* is the speed of the electromagnetic wave. Equation (1) can be expressed logarithmically to calculate free space path loss in dB. The corresponding equation is presented in (2)

$$FSL[dB]=20\log(d)+20\log(f)+92.45, \tag{2}$$

where *d* is the distance between the transmitter and receiver in km, and f is the frequency in GHz. According to (2), the free space path loss for link considered in this research is 142.4 dB. However, due to processes such as signal diffraction and reflection, as well as the topology of the terrain between the transmitter and receiver, the overall attenuation during signal propagation is larger. A variety of propagation models are available to predict signal attenuation that take into account various factors during signal propagation.

The Longley-Rice propagation model is used to predict radio signal attenuation for a communication link operating in the 20 MHz–20 GHz frequency band. According to [14], [15], the Longley-Rice model is the most commonly used propagation model. To forecast signal attenuation, the Longley-Rice propagation model takes into account free space losses, as well as diffraction or scattering effects [14]. The authors in [15] compare different propagation models and software for DTV and FM broadcasting. Their results indicate that *Radio Mobile* gives overall better simulation results with a lower standard deviation. *Radio Mobile* software uses the Longley-Rice propagation model to calculate the path loss [16].

For the radio-relay link Aval Tower-Vršac Hill, calculations were made in *Radio Mobile*, and the obtained path loss value is 150.8 dB. According to the features provided by the radio link (path loss and receiver sensitivity), the RSS should be -46.6 dBm.

The second data set contains minimum and maximum received power levels in dBm for the months of June, July, and August of 2022, i.e., for the period from 2022-06-01 to 2022-08-27. 15 minute intervals were used to record the RSS level. The RSS data were filtered to match the weather data set because the weather data are available on an hourly basis. Based on the minimum and maximum RSS, the average level was also determined and shown in Fig. 2.

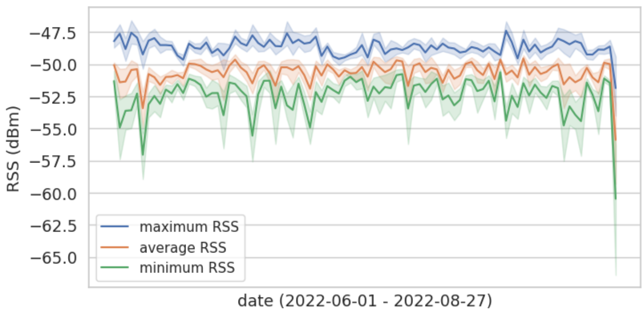

Fig. 2. Maximum, average, and minimum RSS for the period from 2022-06-01 to 2022-08-27.

Figure 2 shows that while the average received signal level straight line matches the expected value of -46.6 dBm, there are also significant oscillations in the received signal level. Table III provides information on the minimum, maximum, and mean RSS level. From Table III, it can be seen that the RSS level varies from -39 dBm to -89 dBm. Considering that the receiver's sensitivity is -73 dBm, this suggests a decrease in quality of service and the possibility of a link failure.

TABLE III. RSS DATA OVERVIEW.

RSS type	Maximum [dBm]	Minimum [dBm]	Mean [dBm]
maximum RSS	-39.00	-59.00	-48.65
minimum RSS	-46.00	-89.00	-52.42
average RSS	-44.50	-68.00	-50.63

In this research, the average RSS level was used for the objective variable. The main task of the predictive model is to identify sudden changes and drops in the RSS level.

B. Feature Selection

After the preprocessing of individual data sets, data were merged based on *date* and *time* attributes, and a unique data set for predicative analyses was created. For feature selection, a correlation matrix was determined. The correlation coefficients, defined by (3), quantitatively describe the relationship between the input features x (RSS) and the dependent feature y (weather conditions)

$$R = \frac{N\sum_{i=1}^{N} x_i y_i - \left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} y_i\right)}{\sqrt{\left[N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2\right]\left[N\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i\right)^2\right]}}. \quad (3)$$

The values of the correlation coefficients range from -1 to +1. Values close to +1 correspond to a positive correlation, i.e., they indicate that an increase in the value of one variable causes an increase in the value of another variable [17]. On the other hand, the values shining -1 correspond to a negative correlation. In the case of negative correlation, an increase in the value of one variable causes a decrease in the value of another variable [17]. Correlation coefficients, whose value is close to 0, indicate a weak correlation between variables.

Based on the correlation matrix, six features were chosen to develop a predictive model. Table IV provides information on the chosen features. The remaining attributes were eliminated due to their weak correlation with the objective variable. Feature selection is done to decrease model training time and increase prediction accuracy [18].

TABLE IV. FEATURE SELECTION OVERVIEW.

Feature	Site	Correlation coefficients
tempC	Aval Tower	0.240
tempC	Vršac Hill	0.236
heatIndexC	Aval Tower	0.233
heatIndexC	Vršac Hill	0.233
humidity	Aval Tower	-0.227
time	-	0.221

Data visualisation was done after the features were selected to properly prepare the data for machine learning. Figure 3(a) shows the variations in the received power level, Fig. 3(b) in the temperature and heat index, and Fig. 3(c) in the relative humidity variations, during a one-week period, from 2022-07-08 to 2022-07-15. A week was selected as the observation period to clearly see the daily changes in the input variables and the objective variable. Additionally, toward the end of the first and fifth days, three notable deviations of the received signal strength from the median can be seen within the observed interval. The quality of signal transmission can be greatly impacted by a sudden decrease in the RSS, and in certain situations, it can even result in link unavailability. Forecasting significant decreases in the RSS level is crucial precisely because of this. Based on the acquired correlation coefficients, Fig. 3 clearly shows a positive correlation between temperature and heat index with the target variable, as well as a negative correlation with the percentage of relative humidity. Thus, RSS increases in response to increases in temperature and heat index, and RSS decreases in response to increases in relative humidity.

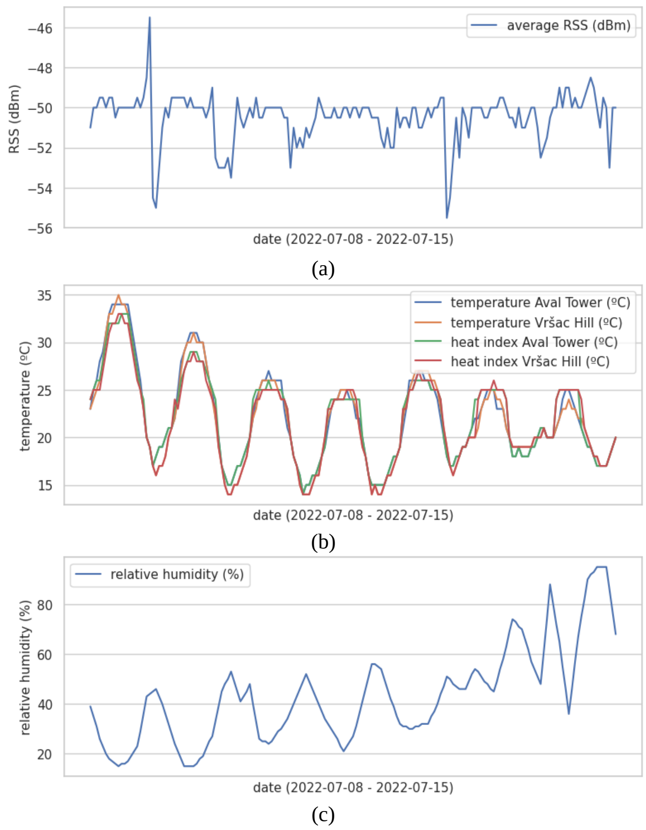

Fig. 3. (a) Change in RSS for a one-week period; (b) Change in temperature and heat index for one-week period; (c) Change in relative humidity for a one-week period.

Low correlation coefficient values were found between the precipitation-related input variables and the target variable during the correlation matrix analysis. It is important to note that measurements of the RSS level were taken in June, July, and August 2022. A significant amount of rain was only recorded for 41 hours out of the 4416 hours that were monitored at both locations, and the mean amount of precipitation was merely 0.057 mm. Precipitation has a considerable impact on the intensity of the received signal , particularly when using higher frequency bands. A large number of studies are available that confirm degradation of signal quality due to precipitation [19]–[21]. The authors in [19] show that rainfall intensities above 64 mm/h at 0.01 % in the West Africa region result in noticeable fading, squelching, and complete outages of digital television signals. The results in [20] show that the surface radio refractivity is higher during the rainy season compared to the dry season.

In this study, a more thorough examination of the influence of precipitation on RSS is not feasible within the parameters of this study due to the previously indicated limits of the data set.

C. Evaluation Metrics

As part of this research, a regression problem was considered that includes the prediction of the RSS. Figure 4 illustrates the graphic relationship between the selected features and the target variable to choose the best machine learning model. It is clear from Fig. 4 that there is no linear relationship between the input variables and the objective variable. Therefore, the training of several nonlinear regression models is required.

The machine learning models currently used for nonlinear regression problems lack explicit methodologies for model evaluation, in contrast to linear regression models that have clearly defined procedures for model validation [22]. The performance of linear regression models is often evaluated using the coefficient of determination or R^2 score. In certain studies, nonlinear regression issues are also assessed using this method. According to [22]–[24] and many other studies, the coefficient of determination is not the optimal choice for model validation in the case of nonlinear regression.

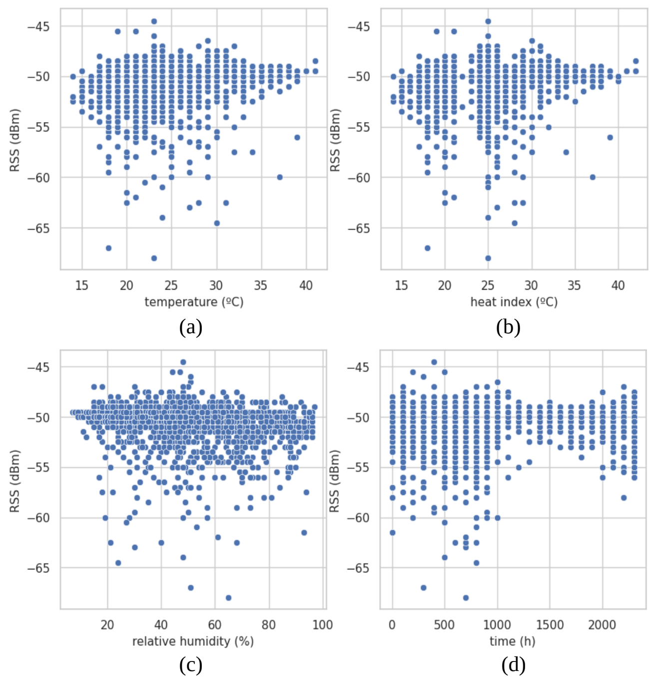

Fig. 4. Relationship between RSS and temperature; (b) Relationship between RSS and heat index; (c) Relationship between RSS and relative humidity; (d) Relationship between RSS and time.

In this research, the models were evaluated using the mean square error (MSE) and the root mean square error (RMSE). In [25], the authors investigate how a sandstorm affects RSSI in Saudi Arabia and also use RMSE to validate a nonlinear regression model. RMSE indicates the difference between the actual and predicted values and is calculated based on (4), where y_i is the exact value, $\hat{y}_i$ is the predicted value, and n is the number of observations

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}. \quad (4)$$

RMSE is easy to interpret because it takes the unit of the objective variable. The average absolute difference between the predicted and target values is represented by the standard metric, MAE, defined by (5)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right). \quad (5)$$

Furthermore, a visual interpretation of the data was carried out, which is crucial for accurately evaluating the nonlinear regression model, as stated in [22].

IV. Results and Discussion

Following the choice of the model performance metric, 80 % of the data set was used for training the model, and the remaining 20 % was utilised for testing. Five well-known machine learning models were trained, including polynomial regression, decision tree, random forest, gradient boosting, and support vector regression (SVR). By employing the *RandomizedSearchCV* approach, the hyperparameters of the chosen models were adjusted. The hyperparameters for the models that were chosen are shown in Table V.

TABLE V. OVERVIEW OF MACHINE LEARNING MODELS HYPERPARAMETERS.

Model	Hyperparameters
Polynomial regression	"poly_degree": 3
Decision Tree	"splitter": "best" "min_samples_split": 6, "min_samples_leaf": 10 "max_depth": 11
Random Forest	"n_estimators": 150 "min_samples_split": 4 "min_samples_leaf": 10 "max_depth": 16
Gradient Boosting	"n_estimators": 50 "min_samples_split": 10 "min_samples_leaf": 6 "max_depth": 3 "learning_rate": 0.1
SVR	"kernel": "rbf" "C": 100 "epsilon": 0.1

The summary of the error analysis results for all models is shown in Table VI. It must be emphasised that prior to adjusting the hyperparameters, the decision tree model achieved the worst results, with MSE 6.731 and RMSE 2.594. As can be seen in Table VI, polynomial regression and the SVR model have slightly worse results compared to other models. The RMSE obtained for decision tree, random forest, and gradient boosting models indicates a 1.7 dBm difference between the expected and actual received signal straight values.

TABLE VI. VALIDATION PERFORMANCE OF THE MODELS.

Model	MSE	RMSE	MAE
Polynomial Regression	3.269	1.808	1.119
Decision Tree	3.160	1.778	1.097
Random Forest	3.107	1.762	1.044
Gradient Boosting	3.118	1.766	1.057
SVR	3.435	1.853	1.040

To choose an appropriate model, the actual values of the received signal straight and the values obtained by prediction were visualised. Substantial variations between the models were discovered after graphical representation. Figure 5(a) shows the obtained prediction results for polynomial regression, while Fig. 5(b) shows the SVR model prediction results. The results shown in Fig. 5 show that all of the values obtained via prediction are centred on the mean value of the received signal straight, indicating that the SVR model performs weakly in forecasting extreme values of the received signal straight level. When polynomial regression was used, similar results were observed. The polynomial regression model has slightly better outcomes than the SVR model, but the results are not satisfactory. This leads us to the conclusion that the SVR and polynomial regression models are not appropriate for forecasting RSS. Regardless of the low value of RMSE, models are not able to predict sudden drops in RSS.

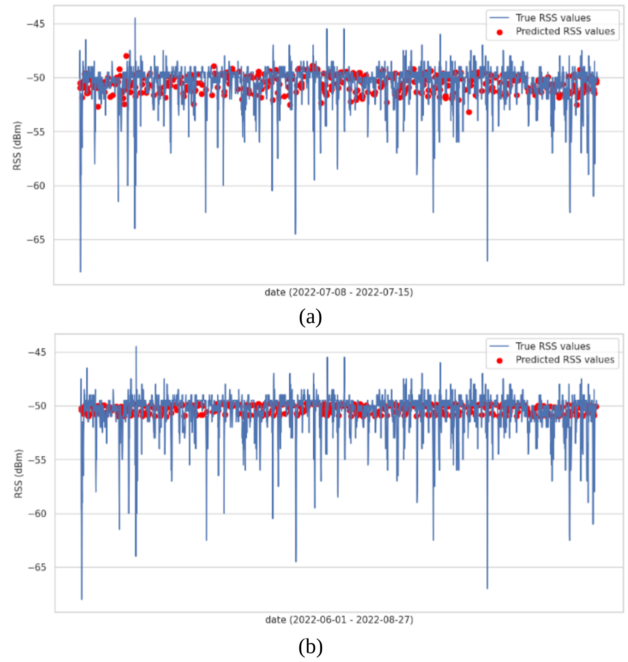

Fig. 5. Actual RSS values and prediction results while using (a) polynomial regression and (b) SVR model.

The prediction results for the decision tree, random forest, and gradient boosting models are shown in Fig. 6.

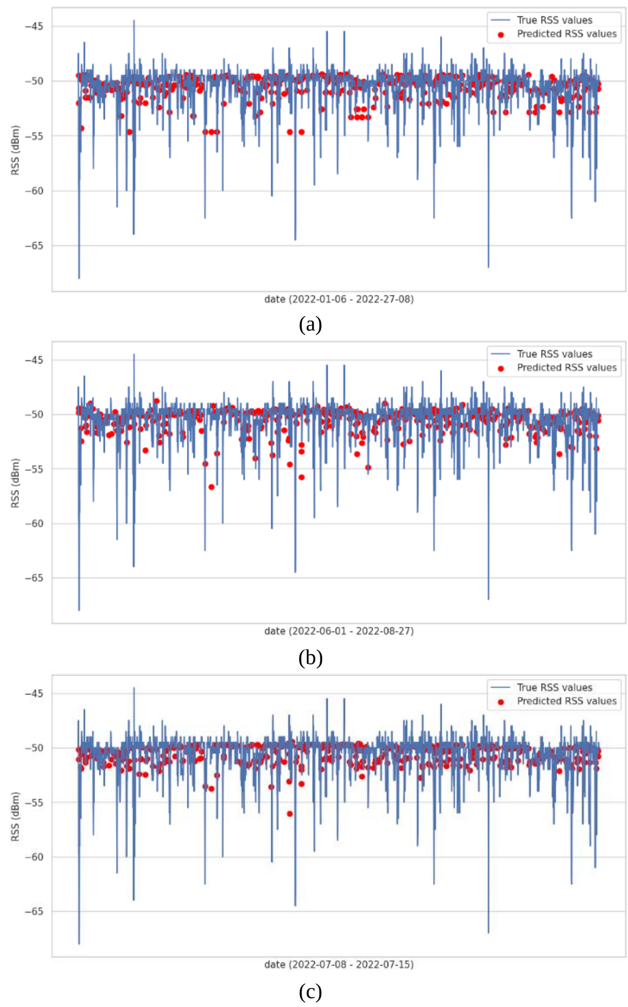

Fig. 6. Actual RSS values and prediction results while using the (a) decision tree, (b) random forest, and (c) gradient boosting model.

As can be seen from Fig. 6, the mentioned models have significantly better performance in predicting signal extreme value changes. One limitation of the developed models is evident from Fig. 6, and it has to do with the prediction of increases in the RSS. The models used do poorly when it comes to forecasting increases in the RSS, but they do well when it comes to predicting decreases in the RSS. When it comes to ensuring that the system runs effectively, the model's primary responsibility is to identify RSS declines, which can lead to link failure and a deterioration in service quality.

It is interesting to notice that prior to hyperparameter adjustment, the decision tree model could forecast increases in RSS. The performance of the decision tree model is shown in Fig. 7 prior to hyperparameter adjustment. As previously indicated, the RSME of the decision tree model was 2.594 prior to hyperparameter modification. Figure 7 illustrates the capacity of the model to forecast increases in RSS, but also shows a more noticeable prediction inaccuracy. Since the main objective of the predictive model is to identify decreases in RSS, using the decision tree model prior to hyperparameter setting was not taken into account. In that scenario, the RMSE increased significantly and the accuracy of the model in identifying RSS decreases was reduced.

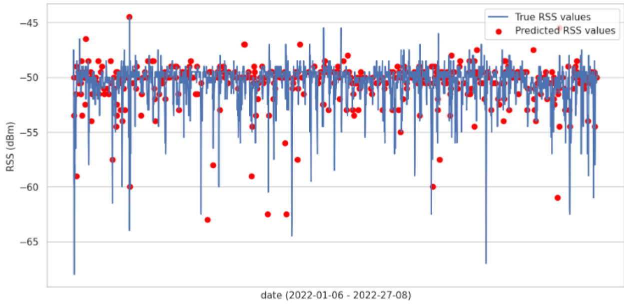

Fig. 7. Actual RSS values and the prediction results while using the decision tree model prior to hyperparameter adjustment.

Following the results of the data visualisation and analysis, further investigation was conducted to determine which prediction method would perform the best. In this part of the research, SVR and polynomial regression models were not used because the performance of the models was below expectations. Figure 8 shows the prediction results using the random forest model for a period of one week, which was also used when examining the correlation coefficients, Fig. 2. Figure 8 illustrates the ability of the model to predict sudden declines in the RSS.

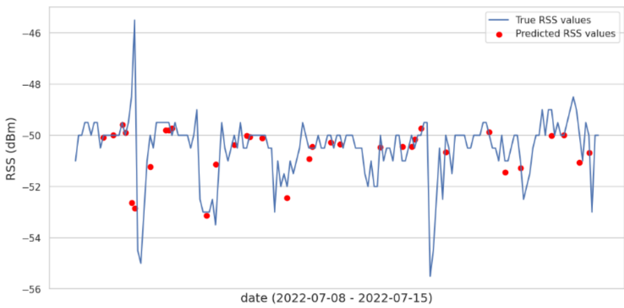

Fig. 8. Actual RSS values and prediction results while using the random forest model for a period of one week.

Similar outcomes were obtained when decision tree and gradient boosting models were used. A new train/test split of the data was carried out, and a different methodology was employed to choose the optimal prediction model. The model was tested for a one-week period, from 2022-07-08 to 2022-07-15, and the remaining data were used for training. For better visibility, the results are displayed for a three-day period in Fig. 9.

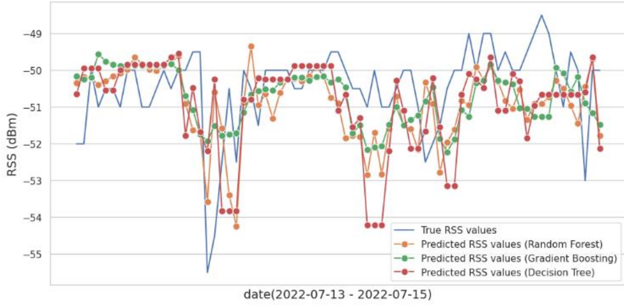

Fig. 9. Actual RSS values and the prediction results for a three-day period.

According to statistics, the decision tree model produced the worst results or the largest RSME. It was shown that when the decision tree model is applied, sudden increases in RSS are detected even if none exist. However, the RMSE values derived from the random forest and gradient boosting models differ slightly from one another. After the data are visualised, it is clear that the gradient boosting model, unlike the random forest model, hardly ever deviates noticeably from the mean value. In the case of the random forest model, there is a possibility of falls increases in RSS, but the random forest model has the capability to predict sudden increases in RSS with low RMSE values. In addition to the specified time period of one week, testing was also performed in other time periods, and after data visualisation, the specified characteristics and behaviour of the predictive models were always observed. This leads us to the conclusion that the random forest model, which solves the nonlinear regression problem of forecasting RSS based on weather conditions (temperature, heat index, and relative humidity), is the optimal model.

V. CONCLUSIONS

This study confirmed the effect of various weather conditions, including temperature, heat index, and relative humidity, on the RF signal in the case of digital TV. The results show that an increase in temperature and heat index leads to an increase in RSS, with correlation coefficients of 0.240 and 0.233, while an increase in relative humidity causes a decrease in RSS, with a correlation coefficient of -0.227. The correlation coefficients are lower compared to some other studies conducted in the tropics region, but the effects of weather conditions are significant and can be used to predict the level of RSS.

To forecast RSS based on weather conditions, we employed supervised machine learning and nonlinear regression models. According to [26], machine learning and predictive models offer cutting-edge features for state-of-the-art wireless and mobile networks. Numerous elements, including communication techniques, frequency, region, influence how the weather affects RF signal propagation. Machine learning is an excellent solution to this type of problem, as historical data can be used to train a model and then make predictions based on weather forecasts to predict and prevent potential problems such as quality of service degradation or link failure. In this study, the data from the real digital TV distribution system was utilised so that the findings could be applied to improve the level of the services that are provided.

To create a model that can precisely predict RSS and significant decreases, we chose nonlinear regression. This approach offers significantly higher flexibility and efficiency compared to other models where binary classification is used to forecast link status. We agree with the authors in [22] that accurately evaluating the nonlinear regression model requires a visual assessment of the data. The best results were obtained by the random forest model, with the lowest values of RMSE of 1.762 and MSE of 3.107. Also, the random forest model has the best performance in terms of predicting sudden decreases in RSS. The ability to predict RSS with an RMSE of 1.762 dB, while RSS ranges from -68 dBm to -44.5 dBm with a mean value of -50.63 dBm, can significantly reduce the weather effects on the radio link and improve overall system performance.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] N. H. Sabri *et al.*, "Correlation analysis of tropical rainforest climate effect on radio signal strength at KUSZA observatory, Terengganu", *Advanced Science Letters*, vol. 23, no. 2, pp. 1268–1271, 2017. DOI: 10.1166/asl.2017.8350.

[2] S. Sabu, S. Renimol, D. Abhiram, and B. Premlet, "A study on the effect of temperature on cellular signal strength quality", in *Proc. of 2017 International Conference on Nextgen Electronic Technologies: Silicon to Software (ICNETS2)*, 2017, pp. 38–41. DOI: 10.1109/ICNETS2.2017.8067893.

[3] A. Akinbolati and M. O. Ajewole, "Effect of some radio climatic factors on digital terrestrial television signal in a Sahel savannah city of Nigeria", *FUDMA Journal of Sciences*, vol. 4, no. 2, pp. 111–118, 2020. DOI: 10.33003/fjs-2020-0402-150.

[4] K. E. Ukhurebor and O. J. Umukoro, "Influence of meteorological variables on UHF radio signal: Recent findings for EBS, Benin City, South-South, Nigeria", in *Proc. of 2nd International Conference on Science and Sustainable Development (ICSSD 2018) "The Role of Science in Novel Research and Advances in Technology". IOP Conference Series: Earth and Environmental Science*, vol. 173. p. 012017, 2018. DOI: 10.1088/1755-1315/173/1/012017.

[5] R. Mat *et al.*, "Effect of humidity on tropospheric received signal strength (RSS) in ultra-high frequency (UHF) band", in *Proc. of The 2nd Joint International Conference on Emerging Computing Technology and Sports (JICETS). Journal of Physics: Conference Series*, vol. 1529, p. 042048, 2020. DOI: 10.1088/1742-6596/1529/4/042048.

[6] P. C. Iwuji and D. E. Bassey, "Investigation of the effect of atmospheric temperature on signal strength generated by FM transmitter in Imo State, Nigeria", *Intern. Journal of Science and Research*, vol. 7, no. 6, pp. 82–85, 2018. DOI: 10.21275/ART20183068.

[7] S. Lavdas, L. Zacharioudakis, A. Khalifeh, and Z. Zinonos, "The effect of temperature and humidity on indoor LoRa propagation model", in *Proc. of 2021 17th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2021, pp. 374–379. DOI: 10.1109/DCOSS52077.2021.00066.

[8] N. C. Y. Lim *et al.*, "Review of temperature and humidity impacts on RF signals", in *Proc. of 2020 13th International UNIMAS Engineering Conference (EnCon)*, 2020, pp. 1–8. DOI: 10.1109/EnCon51501.2020.9299327.

[9] Priyanshu M, V. S. I. Giri, Shachi P, G. Mishra, and Suma M N, "AI powered solution for radio link failure prediction based on link features and weather forecast", *ITU Journal on Future and Evolving Technologies*, vol. 3, no. 2, pp. 165–174, 2022. DOI: 10.52953/ODQQ8049.

[10] J. Tóth, Ľ. Ovseník, J. Turán, L. Michaeli, and M. Márton, "Classification prediction analysis of RSSI parameter in hard switching process for FSO/RF systems", *Measurement*, vol. 116, pp. 602–610, 2018. DOI: 10.1016/j.measurement.2017.11.044.

[11] F. Beritelli, G. Capizzi, G. L. Sciuto, C. Napoli, and F. Scaglione, "Rainfall estimation based on the intensity of the received signal in a LTE/4G mobile terminal by using a probabilistic neural network", *IEEE Access*, vol. 6, pp. 30865–30873, 2018. DOI: 10.1109/ACCESS.2018.2839699.

[12] L. Jari and H. Ismo, "Effects of temperature and humidity on radio signal strength in outdoor wireless sensor networks", in *Proc. of 2015 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 2015, pp. 1247–1255. DOI: 10.15439/2015F241.

[13] Public Broadcasting Technology and Connections. [Online]. Available: https://etv.rs/

[14] S. Kasampalis, P. I. Lazaridis, Z. D. Zaharis, A. Bizopoulos, S. Zettas, and J. Cosmas, "Comparison of Longley-Rice, ITM and ITWOM propagation models for DTV and FM broadcasting", in *Proc. of 2013 16th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, 2013, pp. 1–6. DOI: 10.1109/BMSB.2013.6621780.

[15] A. I. Petrariu, P.-M. Mutescu, E. Coca, and A. Lavric, "A study on LoRa signal propagation models in urban environments for large-scale networks deployment", *Advances in Electrical & Computer Engineering*, vol. 21, no. 4, pp. 61–68, 2021. DOI: 10.4316/AECE.2021.04007.

[16] *Radio Mobile Program Operating Guide* (by B. J. Henderson and P. Eng), Version 3.9, Dec. 30, 2011, pp. 7–8.

[17] S. Ozdemir, *Principles of Data Science*. Packt Publishing Ltd., 2016, pp. 151–152.

[18] M. Emeç and M. H. Özcanhan, "A hybrid deep learning approach for intrusion detection in IoT networks", *Advances in Electrical and Computer Engineering*, vol. 22, no. 1, pp. 3–12, 2022. DOI: 10.4316/AECE.2022.01001.

[19] M. E. Sanyaolu, O. F. Dairo, A. A. Willoughby, and L. B. Kolawole, "1-minute rain rate distribution for communication link design based on ground and satellite measurements in West Africa", *Telecommunications and Radio Engineering*, vol. 79, no. 6, pp. 533–543, 2020. DOI: 10.1615/TelecomRadEng.v79.i6.70.

[20] A. Akinbolati, M. O. Ajewole, A. T. Adediji, and J. S. Ojo, "Propagation curves and coverage areas of digital terrestrial television base stations in the tropical zone", *Heliyon*, vol. 6, no. 3, p. e03599, 2020. DOI: 10.1016/j.heliyon.2020.e03599.

[21] Z. J. Timothy, S. O. Samuel, L.-e. M. Frank, I. Aminu, and Y. Yunisa, "Analysis of frequency and polarization scaling on rain attenuated signal of a KU-band link in Jos, Nigeria", *International Journal of Advanced Networking and Applications*, vol. 14, no. 1, pp. 5310–5317, 2022. DOI: 10.35444/IJANA.2022.14111.

[22] F. Baty, C. Ritz, S. Charles, M. Brutsche, J.-P. Flandrois, and M.-L. Delignette-Muller, "A toolbox for nonlinear regression in R: The package nlstools", *Journal of Statistical Software*, vol. 66, no. 5, pp. 1–21, 2015. DOI: 10.18637/jss.v066.i05.

[23] A. N. Spiess and N. Neumeyer, "An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach", *BMC Pharmacology*, vol. 10, art. no. 6, pp. 1–11, 2010. DOI: 10.1186/1471-2210-10-6.

[24] V. Plevris, G. Solorzano, N. P. Bakas, and M. E. A. Ben Seghier, "Investigation of performance metrics in regression analysis and machine learning-based prediction models", in *Proc. of 8th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS Congress 2022)*, 2022, pp. 1–25. DOI: 10.23967/eccomas.2022.155.

[25] I. Aqeel, E. Iorkyase, H. Zangoti, C. Tachtatzis, R. Atkinson, and I. Andonovic, "LoRaWAN-implemented node localisation based on received signal strength indicator", *IET Wireless Sensor Systems*, vol. 13, no. 4, pp. 117–132, 2023. DOI: 10.1049/wss2.12039.

[26] S. Suljovic, D. S. Krstic, G. Nestorovic, N. N. Petrovic, S. Minic, and D. S. Gurjar, "Using level crossing rate of selection combining receiver damaged by Beaulieu-Xie fading and Rician co-channel interference with a purpose of machine learning QoS level prediction", *Elektronika ir Elektrotechnika*, vol. 29, no. 3, pp. 68–73, 2023. DOI: 10.5755/j02.eie.34018.