Speech Quality Classifier Model based on DBN that Considers Atmospheric Phenomena

Marielle Jordane da Silva, Dick Carrillo Melgarejo, Renata Lopes Rosa, and Demóstenes Zegarra Rodríguez

Original scientific paper

***Abstract*—Current implementations of 5G networks consider higher frequency range of operation than previous telecommunication networks, and it is possible to offer higher data rates for different applications. On the other hand, atmospheric phenomena could have a more negative impact on the transmission quality. Thus, the study of the transmitted signal quality at high frequencies is relevant to guaranty the user´s quality of experience. In this research, the recommendations ITU-R P.838-3 and ITU-R P.676-11 are implemented in a network scenario, which are methodologies to estimate the signal degradations originated by rainfall and atmospheric gases, respectively. Thus, speech signals are encoded by the Adaptive Multi-Rate Wideband (AMR-WB) codec, transmitted and the perceptual speech quality is evaluated using the algorithm described in ITU-T Rec. P.863, mostly known as POLQA. In this work, a novel non-intrusive speech quality classifier that considers atmospheric phenomena is proposed. This classifier is based on Deep Belief Networks (DBN) that uses Support Vector Machine (SVM) with radial basis function kernel (RBF-SVM) as classifier, to identify five predefined speech quality classes. Experimental results show that the proposed speech quality classifier reached an accuracy between 92% and 95% for each quality class overcoming the results obtained by the sole non-intrusive standard described in ITU-T Recommendation P.563. Furthermore, subjective tests are carried out to validate the proposed classifier performance, and it reached an accuracy of 94.8%.**

***Index Terms*—Wireless communications, speech quality, atmospheric phenomena, rain, atmospheric gases.**

I. INTRODUCTION

Nowadays, the demand for services with high data rates is being increasingly demanded, due to the emergence applications such as video streaming, online gaming and virtual reality [1], [2]. In mobile communication systems, four generations (1 to 4G) were implemented. These generations use bandwidth up to 780 MHz. However, these frequency spectra are no longer sufficient to meet the new needs of mobile service providers. Therefore, wireless communication systems are advancing in the use of the millimeter wave frequency spectrum (*mmWave*) [3].

Manuscript received March 17, 2020; revised March 28, 2020. Date of current version March 31, 2020. The associate editor prof. Nikola Rožić has been coordinating the review of this manuscript and approved it for publication.

M. J. da Silva, R. L. Rosa and D. Z. Rodríguez are with the Department of Computer Science, Federal University of Lavras, MG, Brazil, e-mail: marielle-jordane@hotmail.com, renata.rosa@ufla.br, demostenes.zegarra@ufla.br.

D. Carrillo is with the LUT School of Energy Systems, Finland, e-mail: dick.carrillo.melgarejo@lut.fi

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) under Grant 2015/24496-0 and Grant 2018/26455-8.

Digital Object Identifier (DOI): 10.24138/jcomss.v16i1.1033

It is expected that the use of *mmWave* in communication systems will bring a considerable advance in solving congestion problems, limited bandwidth and restricted channel capacity of wireless systems. In this context, the fifth-generation systems (5G) are expected to meet these demands, due to the use of the (*mmWave*) frequency spectrum. This frequency waveband is being considered very important for the new generation of wireless systems, due to wider bandwidth, low latencies, less than 1ms, and data rates higher than 10Gbps. Thus, the 5G networks promise to improve Quality of Service (QoS) in telecommunications [1], [2], [4].

Frequency spectrum availability is a fundamental requirement to allow the testing and deployment of 5G in 2020 [5]. Therefore, according to [6], a larger spectrum band, such as the (*mmWave*) frequency band, will be used by 5G systems, since this band has a large available bandwidth. Several countries have used similar frequency bands for testing and deploying 5G systems. Despite of 5G network advantages, the use of very high carrier frequencies is associated with serious propagation losses [3], [7]–[9]. This problem occurs mainly because of the rain. According to [3], at high frequencies (*mmWave*), the rain causes random fluctuations in the refractive index of the air. These fluctuations cause random changes in the intensity of a propagation signal, causing the phenomenon called *fading,* [10], [11]. This phenomenon is directly related to the frequency of operation, channel conditions, path length, as well as rainfall rate, among others [12]. Considering previous generations (1 to 4G), rain does not cause significant degradation in the transmitted signals, since the wavelength of the carrier frequency is different from the physical dimensions of the raindrops.

To quantify the influence of rain on the attenuation of the transmitted signal, the Radio Communications Sector of the International Telecommunication Union (ITU-R) has produced a global standard named ITU-R P.838-3. This recommendation provides a specific attenuation model for the signal transmitted due to the influence of rain. In addition to this recommendation, ITU-R has another recommendation that quantifies the impact that atmospheric gases in the transmitted signal, which is called ITU-R P.676-11. In both recommendations, the operating frequency range varies from 1 to 100 GHz.

Speech quality in a telecommunications system is considered an important parameter for assessing the user satisfaction [13]. It depends on different factors, such as network conditions, speech codec, environment noise, among others. There are different speech codec characteristics, one of the most used in current cellular networks is the Adaptive Multi-Rate Wideband (AMR-WB) codec [14]. Speech quality assessment is complex, since it is a subjective concept, as it is determined by the listener's perception. The Mean Opinion Score (MOS) test, defined in ITU-T Rec. P.800, is widely accepted as a standard for the subjective classification of speech quality [15]. This test is carried out in a controlled laboratory. At the end of the tests, the scores of all subjects are collected and the average is calculated, and named Mean Opinion Score (MOS). This subjective test is the most accurate, however, it is considered time-consuming and expensive.

Objective methods can be classified into two categories, intrusive or non-intrusive. Intrusive metrics assess speech quality by comparing reference samples (sent) and degraded (received), that is, there is a need for a reference signal to verify the actual degradation of the transmitted signal. Non-intrusive metrics, on the other hand, use the signal in service to make predictions of speech quality, without the need for a reference signal. The algorithms described in the ITU recommendations, P.862 [16], P.863 [17] and P.563 [18] are examples of objective measures. The two first are intrusive algorithms; thus, they need both a reference and degraded speech signals. The third one is a non-intrusive metric, and it only uses a degraded speech signal to estimate a MOS score. For this reason, the P.862 and P.863 algorithms present more accurate results. In addition, it is worth noting that the P.863 algorithm also presents additional features related to modern communication systems and works from narrow-band to full-band networks. However, the P.563 algorithm is more recommended for real-time applications, such as VoIP communications [19].

Another approach to estimate speech quality is the parametric models [20]. They use network parameters to estimate speech quality at the receptor, such as the algorithm described in ITU-T Rec. G.107 [21], [22], mostly known as E-model.

Nowadays, several algorithms, such as, the Artificial Neural Networks (ANN), has been employed for speech analysis and recognition. The Deep Convolutional Neural Network (DCNN) [23], the Restricted Boltzmann Machine (RBM) [24] among others, are used in speech and image applications The RBM is a generative stochastic ANN. It works using a supervised or unsupervised approach. In unsupervised tasks [25], [26] that need a classification step, a supervised learning algorithm need to be added, classifying the samples based on the features extracted by the RBM. Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel (RBF-SVM) associated to DBN presents reliable results in several tasks, such as speech signal pattern recognition [27].

In this arena, this research presents additional contributions regarding [28], which can be summarized as follows:

- A wireless network simulator that considers the implementation of the AMR-WB speech codec using two different operation modes. Thus, more realistic results are obtained. The implementation of atmospheric phenomena in the transmission channel is also considered.
- The P.863 algorithm is used to evaluate the speech signal quality, which gives more realistic results because it considers modern network characteristics.
- A proposed speech quality classifier model that considers atmospheric phenomena parameters and speech codec characteristics, which is based on Deep Belief Networks (DBN) with RBF-SVM. It is important to note that this proposed model analyze the speech signal itself, different that the solution proposed in [28] that considered a network parametric approach.

It is worth noting that the proposed model estimates a perceptual speech quality class, and its main goal is to be useful in wireless network transmission monitoring tasks. Experimental validation results show that the proposed classifier model results are highly correlated to the results obtained by subjective tests.

The article is divided into the following sections. Section II summarizes the recommendations ITU-R P.838-3 and R P.676-11 related to the attenuation caused by rain and atmospheric gases, respectively. Section III presents an overview of deep belief network and speech signal features. The overview of AMR-WB speech codec is presents in Section IV. The implemented simulation scenario is described in Section V. Section VI presents the proposed speech quality assessment model based on DBN. The results obtained are presented in the Section VII. Finally, the conclusion is discussed in Section VIII.

II. Impact of Atmospheric Phenomena on Wireless Transmission

The use of millimeter wave frequencies has gained prominence in recent years, since it was proposed for 5G systems. The millimeter wave bands allow the use of simpler interfaces to achieve high data rates. Also, the large bandwidth available is an advantage of *mmWave*. According to [29], *mmWave* has the following advantages:

- Large bandwidth capacity.
- Highly directive beams.
- Relatively small antennas.
- Low transmitter power requirement.

In recent years, due to the emergence of a new mobile generation system, several types of research are focused on main technologies for 5G, in addition to issues related to network architecture, resource allocation and spectrum management [30].

According to [31] the modeling of radio channels and the propagation prediction for 5G mobile communication systems is one of the most important questions to technically assess whether this new technology will work properly. Also, the propagation of electromagnetic waves is a fundamental factor to understand the design of the transmitter and receiver, the antenna requirements, the power transmission and, the interference levels.

Despite the numerous benefits of using *mmWave* in 5G systems, this frequency range has a disadvantage. The use of very high carrier frequencies involves serious propagation losses, mainly due to rain. This interference in the transmitted signals is more common in regions of intense rain.

In this context, when radio waves propagate in a rainy area, the absorption and dispersion of electromagnetic waves result in significant path-loss. The effect of scattering occurs because the physical dimensions of the raindrops are the same order as the wavelength of the carrier frequency, above 28GHz. In addition to the frequency, the attenuation caused by the rain depends on the size distribution of the raindrop, as well as the polarization of the waves.

[30] states that the attenuation of the transmitted signal due to rain is considered very small at frequencies below 5 GHz. However, considering frequencies above 5 GHz, rain falls, in the form of absorption and dispersion, become more evident, contributing to transmission losses. Although the effect of the rain are small, at frequencies below 5GHz, the absorption effect is evident for frequencies below roughly 1 GHz and the scattering effect is more considered for frequencies above 1 GHz.

High rainfall can cause interruption of *mmWave* length links, which could lead to the disconnection of part of the mobile network. Thus, it is necessary to apply mathematical models to verify the effects of rain on wireless transmissions by *mmWave*, since this atmospheric phenomenon limits the availability and performance of the system. There are several models in the literature that estimate the attenuation caused by rain. These models are based on the same equation to calculate attenuation due to precipitation, as shown in (1).

$$A[dB] = \int_0^d kR^{\alpha}(l)dl \tag{1}$$

where k and α are empirical coefficients and depend on frequency and polarization. $R(l)$ corresponds to the point rain intensity in mm/h, along the path at distance l and d is the path length of the link.

According to [32], the rainfall rate and the attenuation caused can vary considerably, along longer paths. Thus, in practice, the average value on the way is considered.

The most commonly used rain attenuation prediction model is the ITU-R P.530-15. This model does not use the complete distribution of the rain rate, but only one parameter, called $_0.01$. This parameter represents the rainfall rate, obtained for 0.01% of an average year (with an integration time of 1 min). ITU-R P.530-15 determines that if these values are not available from local measurement sources, it is possible to obtain an estimate according to recommendation ITU-R P.837. After determining the rain rate, specific attenuation is calculated (γ_R). The relationship between the attenuation suffered by the signal (γ_R), in dB/km, and the rate of rainfall (R), in mm/h, can be calculated as presented in (2). This attenuation calculation is described in the recommendation ITU-R P.838 [33].

$$\gamma_R = kR^{\alpha} \tag{2}$$

where the coefficients k and α are defined according to variable related to frequency, rainfall, temperature, refractive indexes, elevation angles and polarization state (horizontal or vertical) of the system, as shown in (3) and (4).

$$log_{10}k = \sum_{j=1}^{4}\left(a_j exp\left[-\left(\frac{log_{10}f - b_j}{c_j}\right)^2\right]\right) + m_k log_{10}f + c_k \tag{3}$$

$$\alpha = \sum_{j=1}^{5}\left(a_j exp\left[-\left(\frac{log_{10}f - b_j}{c_j}\right)^2\right]\right) + m_{\alpha} log_{10}f + c_{\alpha} \tag{4}$$

The values of the constants $a_j, b_j, c_j, m_k, c_k, c_{\alpha}$, and m_{α} are given in [33]. The operating frequency of this model varies from 1 to 1000 GHz.

Then, to set the attenuation value along the way, just multiply the attenuation value γ_R by the effective length of the link path.

As mentioned, this model makes predicts rain attenuation based on only the rainfall rate for 0.01% of an average year. However, there are methods depend upon the full rainfall rate distribution, such as UK (2003 RAL) and the Brazil models [32]. However, the model presented in Recommendation ITU-R P.530-15 is the most used, and is therefore used in this work.

In addition to rain, atmospheric gases are also responsible for causing attenuation in the signals transmitted in wireless network systems. When dealing with this phenomenon, the ITU-R produces the recommendation ITU-R P.676-1. This recommendation provides methods to estimate the attenuation by atmospheric gases for electromagnetic waves, in the 1 a 1000 GHz frequency range [34].

According to [35], ITU calculations are generated from a number of equations applied to a dataset of absorption lines water vapor. This dataset is fixed and includes 34 water vapor absorption lines, in different frequency bands.

Thus, the attenuation due to dry air and water vapor can be evaluated for any pressure, temperature and humidity value through a sum of the individual oxygen and water vapor resonance lines. This gas attenuation, according to ITU-R Recommendation P.676-11, is given by [34]:

$$\gamma = \gamma_0 + \gamma_W \tag{5}$$

$$\gamma = 0.1820f(N''_{Oxygen}(f) + N''_{WaterVapour}(f)) \tag{6}$$

where γ_0 and γ_W correspond the specific attenuation, in dB/km, of dry air and water vapor, respectively. f corresponds to the operating frequency which may range from 2 to 1000 GHz. $N''_{Oxygen}(f)$ and $N''_{WaterVapour}(f)$ consists of the imaginary parts of the refractive dependent frequencies related to air pressure and water vapor pressure, respectively.

III. Deep Belief Network and Speech Signal Features

Several speech signals features are used to determine the speech signal characteristics used for different applications [36], [37]. Zero-crossing rate (ZCR) parameter represent fast changes on the speech signal that is composed by vowel and consonant sequence in the temporal domain. There are a high number of parameters based on the frequency domain analysis, such as, the Mel-Frequency Cepstrum Coefficients (MFCC) [38] that gives a speech representation using the mel-frequency scale.

In a speech signal recognition, the hidden Markov models (HMMs) is very used for recognizing the temporal variability of the speech and the Gaussian mixture models (GMMs), which is used to model the density of the states in the HMM because the speech signal can be observed such as a piecewise stationary signal [39]. The area of speech signal recognition in the majority involves a reference signal to be manipulated [40], however the adaptability of new characteristics is more difficult. On the other hand, a reduced-reference method for speech recognition permits to measure the accuracy of classification.

Speech signal parameters are used for several applications applying different methodologies, for instances, in [41], MFCC, amplitude and ZCR parameters are used with GMM for speech discrimination. In [42], authors proposed a model for music and speech classification based on MFCC, ZCR, Linear Predictive Coding (LPC), the spectral centroid, rolloff and flux parameters.

In addition, speech recognition can be accomplished by unsupervised learning. This technique builds representations of the input, which are useful for data classification. There are different techniques that can be used for this purpose, such as density estimation, clustering and Principal Component Analysis (PCA). In addition to these techniques, there is vector quantization (VQ). VQ provides discrete inputs, being considered an initial application for audio analysis. Unsupervised training models work with initial models. Thus, small amounts of transcribed data are represented and the model is used to decode large amounts of un-transcribed data. In this way, new models are trained using part or all of this automatic labeled data.

In recent years, due to the high processing power combined with the expansion of computer memories, it favored the development of complex learning algorithms, such as DNN [43]. This algorithm can be composed of a large number of layers containing non-linear hidden units, as well as many output layers. It should be noted that the DNN can be implemented using unsupervised, as well as supervised machine learning techniques. Another existing technique is RBM. This technique can learn more discriminatory characteristics for a given problem [44]. Thus, the fundamental idea of this technique is to feed the network with unlabeled examples and then rebuild the input data. Thus, this technique can provide an improvement computational cost and, consequently, in the time necessary to complete the training process.

The structure of the RBM is basically composed of visible and hidden units, and the adjacent layers are connected by weights. Thus, RBMs are similar to classic Boltzmann Machines. However, in RBM connections between neurons in the same layer are not allowed. Among the existing methods for training, there is a Contrastive Divergence (CD) [45]. According to [46], this method commonly used in RBMs, due to its efficiency, as well as its reliable results. The CD aims to adjust the input values into the model, in order to work with the approach of maximum likelihood learning. Thus, in this work, it was used as a learning rule.

According to [47], RBM can be used to model fragments of a speech signal. The structure of a DBN is composed by many RBMs. In this structure, the first RBM is trained and, its output is used as an input to the second RBM and, successively. Thus, a hierarchical model learns low-level resources, in order to obtain a high-level representation. The output of a DBN is used as an input for supervised learning methods, such as SVM. In [47], RBM is used for a better representation of speech sound waves. According to these authors, the performance of phoneme recognition using the proposed RBM model is better than solutions based on MFCC.

IV. Overview of AMR-WB Speech Codec

AMR-WB speech codec is described in ITU-T Rec. G.722.2 [14]. It works with nine bit rates from 6.60 kbps and 23.85 kbps, and a low rate background noise encoding mode. The AMR-WB codec can change its bit-rate every 20 ms speech, and it is based on ACELP algorithm.

The number of bits of header and supplementary information of the AMR-WB frame structure are the same of its predecessor the AMR narrow-band (NB) codec. The core frame length–in bits–is different for each operation mode as can be observed in Table I.

TABLE I
Bits of AMR-WB considering each operation mode.

Frame Type	*Number bits*	*Bit-rate (kbps)*
0	132	6.60
1	177	8.85
2	181	12.65
3	213	14.25
4	245	15.85
5	293	18.25
6	325	19.85
7	389	23.05
8	405	23.85

The AMR-WB speech codec works with a sample rate of 16 kHz reaching an improved speech signal quality in relation to the AMR-NB codec. This codec is used in modern communication networks because its reasonable performance in adverse network conditions [48], providing high quality phone calls. The operation modes that use high bit-rates are used for high quality such as music, and its lower bit-rates presents a better performance considering speech quality than other NB codecs.

V. Implementation of the Network Test Scenario

The influence of atmospheric phenomena was analyzed in several audio signals. For this, 20 voice files were extracted, with telephone conversation characteristics. The files were extracted from Rec. ITU-T P.862, with an average duration of around 8 seconds [49]. All of these files have similar characteristics, the total length each audio file is composed by 60% of voice segments and 40% of silence segments. These signals were transmitted in a wireless channel that considers atmospheric factors related to rain and gases.

For the analysis of the degradation of these voice files a simulator was implemented, developed in the Matlab ® software, version 2017b. This software has a function package in which it is possible to simulate the rain and atmospheric gases degradation model, by the *rainpl* and *gaspl* functions, respectively.

The *rainpl* function was developed in accordance with Rec. ITU-R P. 838-3. According to this recommendation, the variables that can influence the degradation of the transmitted signal are the operating frequency carrier of the system, which has values between 1 and 1000 GHz; the distance between the transmitter and receiver (m); the rainfall rate (mm/h); the elevation angles and polarization state of the system, both ranging from -90 to 90°.

The evaluation of the degradation caused by atmospheric gases was verified by the *gaspl* function, which was implemented according to Rec. ITU-R P.676-11. According to this recommendation, the parameters that influence signal degradation consist of operating frequency; atmospheric pressure (Pa), air temperature (°C); and relative humidity (g/cm^3).

In this work, the range of 10 to 100 GHz for the operating frequency was stipulated because it largely covers the frequencies of 5G networks. Others parameters were fixed to specific values to restrict the number of test scenarios. The distance between the transmitter and receiver was fixed to 1000 m, to enable attenuation assessment at each 1 Km. The atmospheric pressure was fixed to $101,300$ Pa, because it corresponds to sea-level pressure. The values of the elevation angles and polarization state used were equivalent to 0°. Preliminary tests were performed with the angle varying from -90 to 90°. However, in all tests, it was found that such parameters did not represent significant changes in the MOS and BER index. Thus, the value of 0° for the variables of the elevation and polarization angles were used in all test. The temperatures of 12, 14, 16, 18, 20 and 25°C were selected. According to [50] the relative humidity of the air equivalent to these temperatures are: 10.68, 12.09, 13.65, 15.4, 17.31 and 23.07 g/cm^3, therefore, these values are adopted. In the transmission system, only the QPSK modulation scheme was implemented.

The speech signals were encoded using the AMR-WB speech codec that is used in current telecommunication networks, specifically the mode operations that represent the maximum (Mode 8) and minimum (Mode 0) bit-rates are implemented to obtain different speech perceptual qualities.

Table II presents an overview of the input parameters used in the simulator.

TABLE II
CONFIGURATION PARAMETERS FOR DIFFERENT TEST SCENARIOS

Parameters	Options/values
Speech Codec	AMR-WB codec - Modes 0 and 8
Modulation scheme	BPSK, QPSK, QAM (16, 64, 256)
Frequency of Operation	10 a 100 GHz - steps of 10 GHz
Rainfall	0, 5, 25, 50, 75, 100 and 200 mm/h
Temperature	12, 14, 16, 16, 18 and 20°C
Relative Humidity of Air	10.68, 12.09, 13.65, 15.4, 17.31 and 23.07 g/cm^3

Fig. 1 presents the block diagram used for the development of the simulator. As previously stated, the simulator aims to quantify the level of degradation of the transmitted audio signals, in the GHz frequency bands. This degradation analysis is done using the MOS index given by the POLQA algorithm.

VI. Proposed Speech Classifier Based on DBN

The proposed speech quality classifier model that considers the presence of atmospheric phenomena during signal transmission is based on an DBN with RBF-SVM. This proposed model uses the signal information of different speech samples.

The network and atmospheric parameters introduced in Table II are used to create different transmission scenarios. Thus, in total six parameters, each one with different values or operation modes, are utilized in the simulation tests. The speech samples at the end side are analyzed considering their speech signal features. In addition, the speech quality of each sample is evaluate by the P.863 algorithm (POLQA), given as a result a MOS score. Therefore, a data-set of speech samples with different impairment types is obtained, and a speech quality score is determined for each sample. Then, a DBN model can be determined using this data-set as shown in Fig. 2.

As can be observed in Fig. 2, the first step in the training process is the speech signal feature extraction. These features are extracted using Matlab tools and they are utilized to build the DBN structure. In order to permit a fast processing, those features are stored in a vector. Additionally, the feature processing of each sample is associated with the corresponding POLQA score that is considered as the learning model output. Because the AMR-WB codec characteristics, a 20 ms speech segment is considered to evaluate the speech features. This period of time is valid for every AMR-WB operation mode.

The speech signal features considered in the tests are the inputs of the DBN that gave the estimated values for every of the impaired speech segments. In this study, 64 features were used, which are: ZCR, spectral flux, spectral roll-off, pitch, spectral centroid, 13 MFCC static features and the first and second derivatives of the static features, and 20 FFT Power Spectrum.

As previously stated, the proposed speech quality assessment model does not give a MOS value, but gives a speech quality class. In this work, five quality classes or categories are proposed. It is important to note that these classes are based on the ITU-T Recommendation P.800, specifically in the five-point MOS scale. Also, it is expected that in real application to know only the speech class be useful. Table III presents the quality scales used in this research.

TABLE III
SPEECH QUALITY CLASSES AND THEIR MOS VALUES RANGE

Speech Quality Class	Perceived quality (ACR Scale)	MOS index values
Class-A	Excellent	5.00-4.00
Class-B	Good	3.99-3.00
Class-C	Fair	2.99-2.00
Class-D	Poor	1.99 and 1.30
Class-E	Very Poor	1.29 and lower

It is important to highlight that two classifiers were tested in initial tests, specifically, the Softmax function and the

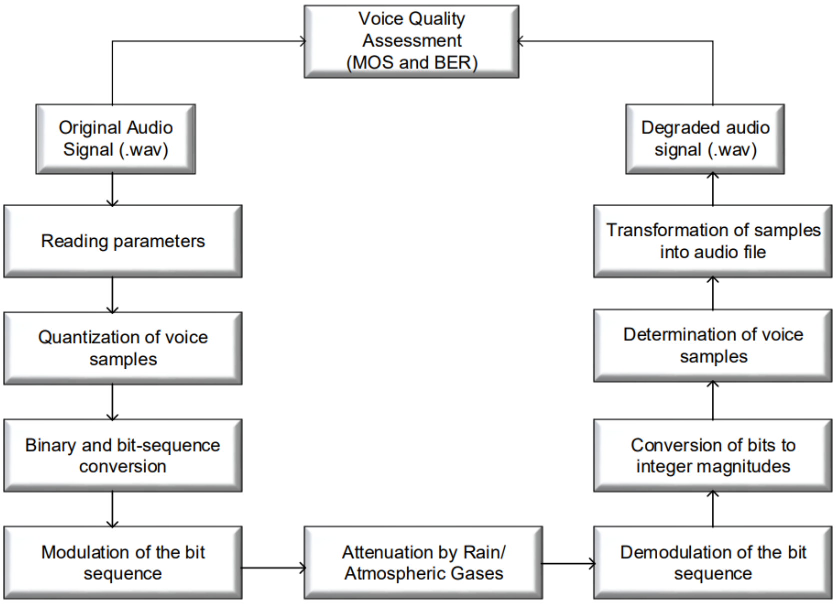

Fig. 1. Block diagram of the network simulator to generate impaired speech samples due to atmospheric phenomena.

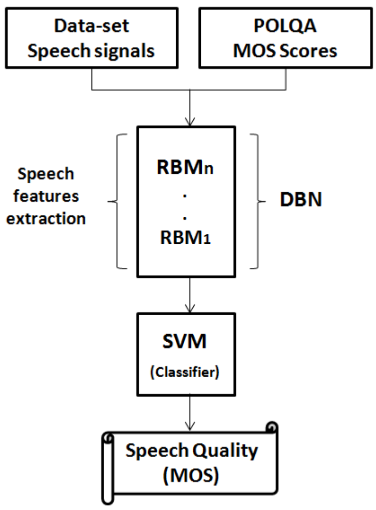

Fig. 2. Flowchart of the DBN training process.

RBF-SVM. In preliminary results the DBN with RBF-SVM classifier reached the best accuracy, then, it was used in all experiments.

VII. Results and Discussion

In this section, firstly, the impact of the rain and atmospheric gases on speech signal quality are presented, for both AMR-WB mode operation used. Later, the performance validation of the proposed model is shown.

A. *Impact of Rain and Atmospheric Gases on Speech Quality*

As stated before, to evaluate the impact of rainfall rates on speech quality, the model presented in Rec. ITU-R P. 838-3 was considered in the network scenario, and POLQA algorithm evaluated impaired speech samples. The results are shown in Tables IV to VIII, that correspond to the radio frequencies of: 10, 30, 50, 70 and 90 GHz, respectively. In these test scenarios, the QPSK modulation scheme was used, and the AMR-WB operation modes 0 (AMR-WB M-0) and 8 (AMR-WB M-8) were implemented.

TABLE IV
Signal Impairment using a Rainfall rate (R) parameter at 10 GHz

f (GHz)	R (mm/h)	AMR-WB M-8	AMR-WB M-0
10	0	4.23	3.52
	5	4.23	3.52
	25	4.23	3.52
	50	4.23	3.52
	75	4.23	3.51
	100	4.22	3.50
	200	4.21	3.50

TABLE V
Signal Impairment using a Rainfall rate (R) parameter at 30 GHz

f (GHz)	R (mm/h)	AMR-WB M-8	AMR-WB M-0
30	0	4.23	3.52
	5	4.23	3.52
	25	4.23	3.51
	50	4.21	3.48
	75	4.19	3.32
	100	3.82	3.02
	200	3.11	2.49

As can be observed in the Tables IV to VIII, in ideal transmission conditions (R=0), AMR-WB M-8 and M-8 obtained the MOS scores of 3.52 and 4.23, respectively, which are the highest speech quality reached in the test scenario.

It is important to note that in each case, as the rainfall rate increases (R), the MOS index decreases that is according to (1), in which the degradation suffered by the signal, γ_R

TABLE VI
SIGNAL IMPAIRMENT USING A RAINFALL RATE (R) PARAMETER AT 50 GHZ

f (GHz)	R (mm/h)	AMR-WB M-8	AMR-WB M-0
50	0	4.23	3,52
	5	4.23	3.52
	25	4.23	3.52
	50	4.19	3.47
	75	3.33	2.64
	100	2.74	2.16
	200	1.78	1.41

TABLE VII
SIGNAL IMPAIRMENT USING A RAINFALL RATE (R) PARAMETER AT 70 GHZ

f (GHz)	R (mm/h)	AMR-WB M-8	AMR-WB M-0
70	0	4.23	3.52
	5	4.22	3.50
	25	4.20	3.19
	50	3.63	2.92
	75	3.11	2.43
	100	2.05	1.38
	200	1.57	1.13

TABLE VIII
SIGNAL IMPAIRMENT USING A RAINFALL RATE (R) PARAMETER AT 90 GHZ

f (GHz)	R (mm/h)	AMR-WB M-8	AMR-WB M-0
90	0	4.23	3.52
	5	4.21	3.51
	25	4.20	3.51
	50	3.55	2.85
	75	2.94	2.34
	100	2.22	1.68
	200	1.51	1.11

increases with the higher the rainfall rate. Additionally, MOS values do not present a significant variation for frequencies lower than 30 GHz. This is because, according to [3], rainfall impact is more evident at frequencies higher than 28 GHz. Thus, in the frequency range from 30 and 100 GHz, the MOS value is negatively affected.

In order to evaluate the impact of atmospheric gases, the test scenario was implemented according to the ITU-R recommendation P.676-11. Similarly, the frequency range considered in the test was from 10 to 100 GHz, with steps of 10 GHz. The experimental results are shown in Tables IX to XIII, corresponding to carrier frequencies of 20, 40, 60, 80 and 100 GHz, respectively.

TABLE IX
SIGNAL IMPAIRMENT USING ATMOSPHERIC GASES PARAMETERS AT 20 GHZ

f (GHz)	Temp. (°C)	Humidity (g/cm^3)	AMR-WB M-8	AMR-WB M-0
20	12	10.68	4.23	3.52
	14	12.09	4.23	3.52
	16	13.65	4.23	3.52
	18	15.4	4.23	3.52
	20	17.31	4.23	3.52

TABLE X
SIGNAL IMPAIRMENT USING ATMOSPHERIC GASES PARAMETERS AT 40 GHZ

f (GHz)	Temp. (°C)	Humidity (g/cm3)	AMR-WB M-8	AMR-WB M-0
40	12	10.68	4.23	3.52
	14	12.09	4.23	3.52
	16	13.65	4.23	3.52
	18	15.4	4.23	3.52
	20	17.31	4.23	3.52

TABLE XI
SIGNAL IMPAIRMENT USING ATMOSPHERIC GASES PARAMETERS AT 60 GHZ

f (GHz)	Temp. (°C)	Humidity (g/cm3)	AMR-WB M-8	AMR-WB M-0
60	12	10.68	4.02	3.24
	14	12.09	3.96	3.21
	16	13.65	3.94	3.18
	18	15.4	3.92	3.16
	20	17.31	3.91	3.12

TABLE XII
SIGNAL IMPAIRMENT USING ATMOSPHERIC GASES PARAMETERS AT 80 GHZ

f (GHz)	Temp. (°C)	Humidity (g/cm3)	AMR-WB M-8	AMR-WB M-0
80	12	10.68	4.23	3.52
	14	12.09	4.23	3.52
	16	13.65	4.23	3.52
	18	15.4	4.23	3.52
	20	17.31	4.23	3.52

TABLE XIII
SIGNAL IMPAIRMENT USING ATMOSPHERIC GASES PARAMETERS AT 100 GHZ

f (GHz)	Temp. (°C)	]@c@Humidity (g/cm3)	AMR-WB M-8	AMR-WB M-0
100	12	10.68	4.23	3.52
	14	12.09	4.23	3.52
	16	13.65	4.23	3.52
	18	15.4	4.23	3.52
	20	17.31	4.23	3.52

As can be observed, the MOS values are very stable, always being close to the highest quality of AMR-WB -8 and AMR-WB-0.Therefore, with the exception of frequencies around 60GHz, the impact of atmospheric gases parameters on speech quality is almost negligible. This is because the water vapor first resonant line appears at frequencies above 100 GHz, as well as the influence of pressure-induced nitrogen attenuation.

In [28], the behavior of signal attenuation (dB/Km) in the frequency range of 10 to 100 GHz due to atmospheric gases is presented as can be observed in Fig. 3.

It is worth noting that that there is a greater attenuation at 60 GHz, and this is the only scenario in which the speech quality is negatively affected. In [51], authors stated that frequency range has not been used in current network generations because the production of RF equipment that support this frequency has a high cost.

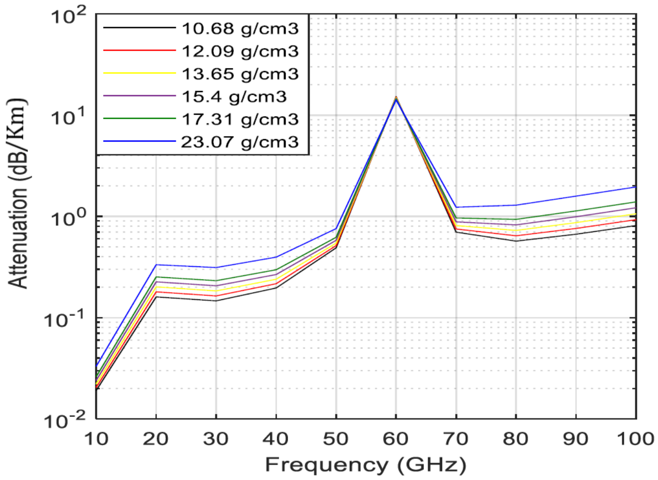

Fig. 3. Signal attenuation originated by atmospheric gases [28].

At this frequency band, many oxygen absorption lines merge together at sea-level pressures to form a single, broad absorption band. Thus, this fact causes a greater attenuation at 60 GHz, since the attenuation of dry air vapor is higher [34]. This fact is worthy of note, since the 60 GHz band is considered promising due to its large availability of unlicensed bandwidth around the world [52]–[54]. In addition, the development of *mmWave* frequency standards, such as IEEE 802.11ad, IEEE 802.15.3c and IEEE 802.11ay demonstrates that 60 GHz band technology can be used in new network generations, such as 5G technology.

B. Performance Assessment of the Proposed Model

The parameters presented in Table II were used in the test scenarios to obtain impaired speech samples, which are used to determine the proposed DBN model. In the DBN training process, the topology of the network used a learning rate of 0.0015, a Dropout Fraction value of 0.1, CD steps to 1, and momentum 0.8. Each network of the algorithm was trained using 500 epochs. The DBN is based on three hidden layers, with RBM of 100 neurons each one. This topology is used because reached the best results in relation to other configuration that were previously tested.

As first step, the unimpaired speech samples from the dataset are randomly divided for training and validation phases considering 80% and 20%, respectively. Then, they are used in the all network simulation scenarios, and the MOS values obtained by POLQA are used as reference values.

Table XIV presents DBN Model performance assessment results for speech quality class estimation in the validation tests, using the confusion matrix format, considering in both cases the POLQA results as ground truth. For comparison purposes, Table XV present the results obtained by the non-intrusive P.563 algorithm, in which each MOS value is attributed to a quality class according to Table II.

From Tables XIV and XV can be observed the superior performance results obtained by the proposed DBN classifier model. The proosed DBN model can be used as a non-intrusive quality metric that considers the atmospheric phenomena.

As the last step, subjective tests of speech quality assessment were performed in a controlled environment. The ITU-T P.563 algorithm was used for comparison purposes because is the sole standardized non-intrusive algorithm in the current literature. In the subjective test, 34 volunteers participated. 4 speech files that correspond to scenarios presented in Tables IV to XIII were considering; thus, in total 480 speech files were evaluated by subjective tests. The classification accuracy is used to compare proposed DBN model and ITU-T P.563 algorithm considering subective test results. These results are presented in Table XVI.

As can be observed in Table XVI, the proposed speech quality classifier model presented a better correlation with subjective test results than the P.563 algorithm. Thus, the proposed model can be useful in modern communication networks that operate in high frequency bands, such as those used in 5G networks. It is important to note that P.563 is a non-intrusive algorithm which results are not reliable in lossy channel transmissions [55], but there is not another standardized non-intrusive algorithm.

VIII. Conclusions

In this work, a network transmission scenario is implemented that considers different intensities of rainfall and atmospheric gases according to ITU-R recommendation P.530-15 and ITU-R recommendation P.676-1, respectively. Our preliminary experimental results show the impact of these atmospheric phenomena on signal quality transmission can be significant at high frequency range, such as 10 to 100 GHz. It is important to note that current 5G network implementations consider this frequency band. Experimental results show that rainfall has a considerable negative impact on the signal transmission for frequencies higher than 30 GHz. In the case of the atmospheric gases model, the impact on the signal quality is almost negligible, except for the 60 GHz frequency that causes a high signal attenuation. The low signal degradation shown in these frequency ranges is due to the influence of oxygen. That is because the water vapor first resonant line is above 100 GHz.

In the test scenario, actual speech samples are used, which are encoded by the AMR-WB speech codec that is used in current cellular networks. Additionally, to evaluate the speech quality is used the POLQA algorithm that is the latest ITU-T standard regarding the intrusive speech quality algorithms, and also it incorporates modern telecommunication network characteristics. The resulting impaired speech samples are assessed by the POLQA algorithm and classified in one of the five quality classes previously determined.

The proposed speech quality classifier is based on the RBM that extracts features from speech sample signals, and the RBF-SVM classifier. The results obtained show the high performance of the proposed DNC classifier, reaching 92.46% to 95.76% classification accuracy in the validation test overcoming the results obtained by the algorithm described in the ITU-T recommendation P.563. It is important to stress that te proposed model can be used as non-intrusive method. Additionally, subjective test were performed to evaluate 480

TABLE XIV
CONFUSION MATRIX FOR DBN MODEL CLASSIFICATION RESULTS (IN PERCENTAGE)

Speech Qual. Class	DBN / P.563 Class-A	DBN / P.563 Class-B	DBN / P.563 Class-C	DBN / P.563 Class-D	DBN / P.563 Class-E
Class-A	95.76	4.24	0.0	0.0	0.0
Class-B	3.85	94.58	1.57	0.0	0.0
Class-C	0.0	3.59	93.27	3,14	0.0
Class-D	0.0	0.0	2,26	92.46	5.28
Class-E	0.0	0.0	1.08	6.11	92,81

TABLE XV
CONFUSION MATRIX FOR P.563 RESULTS CONSIDERING SPEECH QUALITY CLASSES (IN PERCENTAGE)

Speech Qual. Class	DBN / P.563 Class-A	DBN / P.563 Class-B	DBN / P.563 Class-C	DBN / P.563 Class-D	DBN / P.563 Class-E
Class-A	46.15	23.85	28.44	1.56	0.0
Class-B	1.19	59.21	31.28	7.11	1.21
Class-C	0.43	4.26	82.65	10.44	2.22
Class-D	0.0	0.27	4.53	89.88	5.32
Class-E	0.0	0.02	2.15	6.58	91.25

TABLE XVI
PERFORMANCE ASSESSMENT VALIDATION OF THE PROPOSED MODEL IN RELATION TO SUBJECTIVE TEST RESULTS

	Accuracy
DBN Model vs Subjective tests	0.948
ITU-T P.563 vs Subjective tests	0.693

additional impairment samples, in which the proposed DBN model reached 94.8% of accuracy.

REFERENCES

[1] G. W. et al., "5gnow: non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 97–105, February 2014.

[2] D. Soldani and A. Manzalini, "Horizon 2020 and beyond: On the 5g operating system for a true digital society," *IEEE Vehicular Technology Magazine*, vol. 10, no. 1, pp. 32–42, March 2015.

[3] H. Y. Lam, L. Luini, J. Din, M. J. Alhilali, S. L. Jong, and F. Cuervo, "Impact of rain attenuation on 5g millimeter wave communication systems in equatorial malaysia investigated through disdrometer data," in *11th European Conference on Antennas and Propagation (EUCAP)*, March 2017, pp. 1793–1797.

[4] D. Z. Rodríguez and S. Möller, "Speech quality parametric model that considers wireless network characteristics," in *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*, June 2019, pp. 1–6.

[5] *5G Spectrum Bands*, https://gsacom.com/5g-spectrum-bands/, 2017.

[6] A. Tamayo–Dominguez, X. Sun, and J. Fernández–González, "New manufacturing technologies for 5g millimeter wave antennas," in *2018 11th Global Symposium on Millimeter Waves (GSMM)*, May 2018, pp. 1–4.

[7] O. Adetan and T. Afullo, "The influence of disdrometer channels on specific attenuation due to rain over microwave links in southern africa," *SAIEE Africa Research Journal*, vol. 105, no. 1, pp. 31–40, March 2014.

[8] C. Capsoni, L. Luini, A. Paraboni, C. Riva, and A. Martellucci, "A new prediction model of rain attenuation that separately accounts for stratiform and convective rain," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 1, pp. 196–204, Jan 2009.

[9] E. Matricciani, "Physical-mathematical model of the dynamics of rain attenuation based on rain rate time series and a two-layer vertical structure of precipitation," *Radio Science*, vol. 31, no. 02, pp. 281–295, March 1996.

[10] H. Y. Lam, L. Luini, J. Din, C. Capsoni, and A. D. Panagopoulos, "Investigation of rain attenuation in equatorial Kuala Lumpur," *IEEE Antennas and Wireless Propag. Letters*, vol. 11, pp. 102–105, 2012.

[11] E. T. Affonso, D. Z. Rodríguez, R. L. Rosa, T. Andrade, and G. Bressan, "Voice quality assessment in mobile devices considering different fading models," in *Proc. IEEE Int. Symposium on Consumer Electronics*, Sao Paulo, Brazil, Sep. 2016, pp. 21–22.

[12] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. Maccartney, T. S. Rappaport, and A. Alsanie, "Radio propagation path loss models for 5g cellular networks in the 28 ghz and 38 ghz millimeter-wave bands," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 78–86, Sep. 2014.

[13] C. A. B. Mello and R. Q. Albuquerque, "Reference-free speech quality assessment for mobile phones based on audio perception," in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, Oct 2015, pp. 2413–2417.

[14] ITU-T Rec. G.722.2, "Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB) ," Jul 2013. [Online]. Available: https://www.itu.int/rec/T-REC-G.722.2-200307-I/en

[15] ITU-T Rec. P.800, "Methods for subjective determination of transmission quality," Jun. 1996. [Online]. Available: www.itu.int/rec/T-REC-P.800/en

[16] ITU-T Rec. P.862, "An objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs," Set. 2001. [Online]. Available: www.itu.int/rec/R-REC-P.862/en

[17] ITU-T Rec. P.863, "Perceptual objective listening quality prediction," Mar. 2018. [Online]. Available: http://www.itu.int/rec/T-REC-P.863-201803-I/en

[18] ITU-T Rec. P.563, "Single-ended method for objective speech quality assessment in narrow-band telephony applications," May. 2004. [Online]. Available: www.itu.int/rec/T-REC-P.563/en

[19] R. Dantas Nunes, C. H. Pereira, R. L. Rosa, and D. Z. Rodríguez, "Real-time evaluation of speech quality in mobile communication services," in *IEEE International Conference on Consumer Electronics (ICCE)*, Jan 2016, pp. 389–390.

[20] D. Z. Rodríguez, G. F. Pívaro, R. L. Rosa, G. Mittag, and S. Möller, "Improving a parametric model for speech quality assessment in wireless communication systems," in *26th Intern. Conf. on Software, Telecommunications and Computer Networks (SoftCOM)*, Sep. 2018, pp. 1–5.

[21] ITU-T Rec. G.107, "The e-model: a computational model for use in transmission planning," Jun. 2014. [Online]. Available: www.itu.int/rec/T-REC-G.107.

[22] D. Z. Rodríguez, R. L. Rosa, F. L. Almeida, G. Mittag, and S. Möller, "Speech quality assessment in wireless communications with mimo systems using a parametric model," *IEEE Access*, vol. 7, pp. 35 719–35 730, 2019.

[23] Y. Qian, M. Bi, T. Tan, and K. Yu, "Very deep convolutional neural networks for noise robust speech recognition," *IEEE/ACM Trans. on Audio, Speech, and Language Processing*, vol. 24, no. 12, pp. 2263–2276, Dec. 2016.

[24] Y. Wang, S. Zhao, J. Li, and J. Kuang, "Speech bandwidth extension using recurrent temporal restricted boltzmann machines," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1877–1881, Dec. 2016.

[25] C. L. P. Chen, C. Y. Zhang, L. Chen, and M. Gan, "Fuzzy restricted

boltzmann machine for the enhancement of deep learning," *IEEE Trans. on Fuzzy Systems*, vol. 23, no. 6, pp. 2163–2173, Dec. 2015.

[26] Y. Qian and P. C. Woodland, "Very deep convolutional neural networks for robust speech recognition," *IEEE/ACM Trans. on Audio, Speech, and Language Processing*, vol. 24, no. 12, pp. 481–488, Dec. 2016.

[27] G. Wen, H. Li, J. Huang, D. Li, and E. Xun, "Random deep belief networks for recognizing emotions from speech signals," *Computational Intelligence and Neuroscience*, vol. 2017, no. 2, pp. 1–9, Mar. 2017.

[28] M. J. da Silva, D. C. Begazo, and D. Z. Rodríguez, "Evaluation of speech quality degradation due to atmospheric phenomena," in *2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, Sep. 2019, pp. 1–6.

[29] K. Ulaganathen, I. M. Rafiqul, K. Abdullah, and T. A. Rahman, "Specific rain attenuation analysis and modeling for 5g communication," in *2018 7th International Conference on Computer and Communication Engineering (ICCCE)*, Sep. 2018, pp. 1–4.

[30] K. Ulaganathen, A. R. Tharek, R. M. Islam, and K. Abdullah, "Rain attenuation for 5g network in tropical region (malaysia) for terrestrial link," in *2017 IEEE 13th Malaysia International Conference on Communications (MICC)*, Nov 2017, pp. 35–38.

[31] C. Kourogiorgas, S. Sagkriotis, and A. D. Panagopoulos, "Coverage and outage capacity evaluation in 5g millimeter wave cellular systems: impact of rain attenuation," in *2015 9th European Conference on Antennas and Propagation (EuCAP)*, April 2015, pp. 1–5.

[32] P. Kántor, J. Bitó, and Drozdy, "Characteristics of 5g wireless millimeter wave propagation: Transformation of rain attenuation applying different prediction models," in *2016 10th European Conference on Antennas and Propagation (EuCAP)*, April 2016, pp. 1–5.

[33] ITU Rec. P.838.3, "Specific attenuation model for rain for use in prediction methods," Mar. 2005. [Online]. Available: https://www.itu.int/rec/R-REC-P.838-3-200503-I

[34] ITU-T Rec. P.676, "Attenuation by atmospheric gases," Set. 2006. [Online]. Available: https://www.itu.int/rec/R-REC-P.676/en

[35] J. F. O'Hara and D. R. Grischkowsky, "Comment on the veracity of the itu-r recommendation for atmospheric attenuation at terahertz frequencies," *IEEE Transactions on Terahertz Science and Technology*, vol. 8, no. 3, pp. 372–375, May 2018.

[36] Y. J. Hu and Z. H. Ling, "DBN-based spectral feature representation for statistical parametric speech synthesis," *IEEE Signal Processing Letters*, vol. 23, no. 3, pp. 321–325, Mar. 2016.

[37] J. Fredes, J. Novoa, S. King, R. M. Stern, and N. B. Yoma, "Locally normalized filter banks applied to deep neural-network-based robust speech recognition," *IEEE Signal Processing Letters*, vol. 24, no. 4, pp. 377–381, Apr. 2017.

[38] D. Chazan, R. Hoory, G. Cohen, and M. Zibulski, "Speech reconstruction from mel frequency cepstral coefficients and pitch frequency," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing*, Istanbul, Turkey, Aug. 2000, pp. 1299–1302.

[39] Y.-S. Yun and Y.-H. Oh, "A segmental-feature HMM for speech pattern modeling," *IEEE Signal Processing Letters*, vol. 7, no. 6, pp. 135–137, Jun. 2000.

[40] A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*, ser. Studies in Computational Intelligence. Springer, 2012, vol. 385.

[41] M. Carey, E. Parris, and H. Lloyd-Thomas, "A comparison of features for speech, music discrimination," in *Proc. of IEEE Int. Conf. on Acoustics, Speech, Signal Processing*, Arizona, USA, Mar. 1999, pp. 191–199.

[42] C. Liu, L. Xie, and H. Meng, "Classification of music and speech in mandarin news broadcasts," in *National Conf. on Man-Machine Speech Communication*, Anhui, China, Oct. 2007, pp. 17–20.

[43] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

[44] G. Pan, J. Qiao, W. Chai, and N. Dimopoulos, "An improved RBM based on bayesian regularization," in *Proc. Int. Joint Conf. on Neural Networks*, Beijing, China, Jul. 2014, pp. 2935–2939.

[45] E. T. Affonso, R. D. Nunes, R. L. Rosa, G. F. Pivaro, and D. Z. Rodríguez, "Speech quality assessment in wireless voip communication using deep belief network," *IEEE Access*, vol. 6, pp. 77 022–77 032, 2018.

[46] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computing*, vol. 18, no. 7, pp. 1527–1554, 2006.

[47] N. Jaitly and G. E. Hinton, "Learning a better representation of speech soundwaves using restricted boltzmann machines." in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal*, Prague, Czech Republic, May 2011, pp. 5884–5887.

[48] D. Z. Rodríguez, M. Arjona Ramírez, L. F. Bernardes, G. Mittag, and S. Möller, "Impact of fec codes on speech communication quality using wb e-model algorithm," in *2019 Wireless Days (WD)*, 2019, pp. 1–4.

[49] D. Z. Rodríguez, G. F. Pívaro, R. L. Rosa, G. Mittag, and S. Möller, "Quantifying the quality improvement of MIMO transmission systems in voip communication," in *International Conference on Software, Telecomm. and Comp. Networks (SoftCOM)*, Sep. 2018, pp. 1–5.

[50] *Relative humidity*, http://www.etec.com.br/muda3.html, 2019.

[51] R. Gomes, L. Sismeiro, C. Ribeiro, T. R. Fernandes, M. G. Sanchéz, A. Hammoudeh, and R. F. S. Caldeirinha, "Will COTS RF Front-Ends Really Cope With 5G Requirements at mmWave?" *IEEE Access*, vol. 6, pp. 38 745–38 769, 2018.

[52] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-ghz integrated circuits and systems for wireless communications," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1390–1436, Aug 2011.

[53] R. C. Daniels and R. W. Heath, "60 ghz wireless communications: Emerging requirements and design recommendations," *IEEE Vehicular Technology Magazine*, vol. 2, no. 3, pp. 41–50, Sep. 2007.

[54] J. Kim, J.-J. Lee, and W. Lee, "Strategic control of 60 ghz millimeter-wave high-speed wireless links for distributed virtual reality platforms," *Mobile Information Systems*, vol. 2017, pp. 1–10, 03 2017.

[55] E. T. Affonso, R. L. Rosa, and D. Z. Rodríguez, "Speech quality assessment over lossy transmission channels using deep belief networks," *IEEE Signal Processing Letters*, vol. 25, no. 1, pp. 70–74, 2018.

M. J. da Silva received her Master's degree in Systems Engineering and Automation from the Federal University of Lavras, 2019. She received her graduation in Electrical Engineering from the Federal Institute of Education, Science and Technology of Minas Gerais, 2016. She is computer technician from the Federal Center for Technological Education of Minas Gerais, 2011. She is currently a substitute professor in the Mechatronic Engineering Department at CEFET-MG, Divinópolis campus.

D. C. Melgarejo (M'08) received the B.Eng. degree (Hons.) in electronics and electrical engineering from San Marcos National University, Lima, Perú, and the M.Sc. degree in electrical engineering from Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil, in 2004 and 2008, respectively. Between 2008 and 2010, he contributed to WIMAX (IEEE 802.16m) standardization. From 2010 to 2018, he worked with the design and implementation of cognitive radio networks and projects based on 3GPP technologies. Since 2018 he is a researcher at LUT University, where he is also pursuing the Ph.D degree in electrical engineering. His research interests are mobile technologies beyond 5G, energy harvesting, intelligent meta-surfaces, and Cell-free mMIMO.

R. L. Rosa received her M.S. degree from the University of São Paulo in 2009 and her Ph.D. degree from the Polytechnic School of the University of São Paulo, in 2015 (EPUSP). She is currently an Adjunct Professor with Department of Computer Science, Federal University of Lavras, Brazil. Her current research interests include computer networks, artificial intelligence algorithms, recommendation systems, telecommunication systems, wireless networks, and quality of service and quality of experience in multimedia services.

D. Z. Rodríguez (M'12-SM'15) received the B.S. degree in electronic engineering from the Pontifical Catholic University of Peru, the M.S. degree and Ph.D. degree from the University of São Paulo in 2009 and 2013. He is currently an Adjunct Professor with the Department of Computer Science, Federal University of Lavras, Brazil. He has a solid knowledge in Telecommunication Systems and Computer Science based on 15 years of Professional experience in major companies. His research interest includes QoS and QoE in Multimedia services, artificial intelligence algorithms, and architect solutions in Telecommunication Systems.