Journal of Information Hiding and Multimedia Signal Processing ©2024 ISSN 2073-4212
Ubiquitous International Volume **15**, Number **2**, June **2024**

RFSTC-HIL: Generating Dataset of Relevant Features Signal Type Classification Based on Hardware–in–the-Loop

Mohamed A. Aboelazm[1], Mark Aziz Anis[1], Sherif A. Elgamel[1], Mohamed A. Elshafey[2,*]

[1]Electronic Warfare Department
[2]Computer Engineering and Artificial Intelligence Department
Military Technical College, Cairo, Egypt
{mohamed.a.aboelazm, marcazizanis, elgamel, m.shafey}@mtc.edu.eg

Corresponding author: Mohamed A. Elshafey

Received May, 2024, revised June, 2024, accepted June, 2024.

Abstract. *Wireless communication systems heavily depend on signal classification for applications including spectrum monitoring, cognitive radio, and signal intelligence. Accuracy of artificial intelligence based signal classification models are deeply dependent on the used dataset for training these models. This paper explores a signal dataset synthesizing framework, where the selected signal types have different common communication standards in various applications. These signals have relevant features that modulation recognition is not sufficient for classifying them. For analysis and comparison purposes, multiple dataset versions are synthesized based on software defined radio (SDR) that vary in inclusion of communication propagation channel conditions and hardware effects. The proposed framework incorporates signal impairment sources to precisely simulate real-world RF propagation environment scenarios. A neural network model ResNet is used for training and testing several versions of the dataset to examine its performance in several case studies. Simulation results and analysis demonstrate that the proposed dataset that considers real-world communication scenario, for SNR values greater than 0 dB, achieves an average rise in accuracy of 3% to 4% w.r.t. results of training on data sets that ignore inclusion of large and small channel fading effects and 6% to 12% w.r.t. results trained on data set of the simplest conditions.*
Keywords: Deep Learning, Synthesizing dataset , SDR, Signal classification, ResNet.

1. **Introduction.** Wireless communication systems utilize signal classification as a fundamental task, where it plays a crucial role in various applications such as spectrum monitoring, cognitive radio, and signal intelligence. The process of signal classification involves identifying the modulation type, the coding scheme, or the bandwidth of a received signal, which is essential for adaptive modern communication systems. Traditional methods for signal classification often rely on handcrafted features and heuristic algorithms, which may lack robustness and adaptability to changes in signal conditions [1]. Recently, deep learning techniques have emerged as promising solutions for modulation recognition and signal classification tasks [1]. Deep learning models can automatically extract relevant features from raw signal data, eliminating the need for manual feature engineering. This approach offers several advantages, including improved accuracy, scalability, and adaptability of various modulation schemes and signal characteristics variations in modern communication systems [2].

For relevant characteristics standard signals classification task, there are limited number of open-sourced datasets for that task and only for small number of types [3, 4]. Conversely, for the modulation recognition application, it is facilitated in research and benchmarking with several standardized datasets that have been developed and widely adopted by the research community. Among these datasets, the RadioML 2016 (RML 2016) [5], RadioML 2018 (RML 2018) [6], and Signal 53 (Sig 53) [7] datasets are prominently featured, where each dataset offers a diverse collection of modulation types and signal characteristics, catering to different research requirements and scenarios. Although modulation recognition process is crucial in various systems, yet it may not provide as much detail as signal type classification. This is because multiple signal standards can share the same modulation technique but differ in bandwidth, pulse shaping filter, baud rate, and coding scheme, resulting in them being classified under the same modulation category. In essence, signal classification defines the inherent properties of the signal, while modulation recognition extracts only the modulation technique [8].

Given that communication systems primarily adhere to specific well-established standards, it would be more effective to classify them based on these standards [9]. Another aspect to be considered is the signal degradation type that have to be considered in the used dataset to enhance the trained model's resilience to various distortions the signal may encounter in the propagation channel [2]. In signal processing based on AI, datasets can be generated in three main methods: simulating the dataset using software such as MATLAB®, using software-defined radio (SDR) to create a more authentic dataset with hardware imperfections (HWI), or recording real-world RF signals. Each method has its challenges and considerations [8]. This work explores the creation of a dataset based on SDR and analysis of the implications of incorporating various distorting factors into this dataset. This paper is organized as follows: Section 2 presents the system model. Section 3 explains the proposed framework Section 4 evaluates the classification performance of the proposed dataset generation framework with the classification accuracy analysis against the different signal conditioning in the generated datasets. Finally, Section 5 concludes the paper.

2. **System Model.** The system model comprises two main components: Standard signals of interest and signal degradation modelling.

2.1. **Standard signals of interest.** Wireless digital communication standards vary to enable efficient and reliable voice and data transfer, providing interoperability, compatibility, and performance across various communication systems and networks. This study aims to generate a dataset that generalizes across various applications by incorporating a wide range of technologies and standards. The selection of standards is intended to cater to the diverse requirements of multiple applications in the telecommunications, public safety, and military communication sectors.

This study includes 11 communication standard signals [10, 11, 12, 13, 14] as indicated in Table 1 and the relevance between the selected standard signals is noticed, where many modulation schemes across various standards are either identical or closely correlated. subsequently, modulation classifier is not sufficient to analyze these signals,where most of them are identified as differential QPSK modulation, despite having distinct characteristics such as symbol rate, coding scheme, or bandwidth for transmission.

2.2. **Signal Degradation Modelling.** Signal modelling or conditioning is crucial in simulating communication systems due to the distortions caused by the propagation channel and imperfections in the receiver hardware. Communication systems utilize several modeling approaches to represent fading phenomena, including deterministic, statistical, and

Table 1. List of Signal Standards Generated [10, 11, 12, 13, 14]

<table>
<tr><th>Standard</th><th>Modulation</th><th>Symbol Rate [ksps]</th><th>Application</th></tr>
<tr><td>APCO Phase 1C4FM</td><td>4FSK,
Deviation 1.8kHz</td><td rowspan="4">4.8</td><td rowspan="7">Public-Safety and Military Communications</td></tr>
<tr><td>APCO Phase1 CQPSK</td><td rowspan="3">$\pi/4$-DQPSK</td></tr>
<tr><td>APCO Phase1 LSM</td></tr>
<tr><td>APCO Phase1 WCQPSK</td></tr>
<tr><td>APCO Phase2 H-CPM</td><td>4FSK,
Deviation 3kHz</td><td rowspan="2">6</td></tr>
<tr><td>APCO Phase2 H-DQPSK</td><td>$\pi/4$-DQPSK</td></tr>
<tr><td>APCO Phase2 H-D8PSK Wide/Narrow</td><td>$\pi/8$-DQPSK</td><td>4</td></tr>
<tr><td>NADC</td><td rowspan="3">$\pi/4$-DQPSK</td><td>21</td><td rowspan="2">Digital cellular communications</td></tr>
<tr><td>PDC</td><td rowspan="2">18</td></tr>
<tr><td>TETRA</td><td>Public-Safety and Military Communications</td></tr>
</table>

empirical models. Fading occurs on a large scale due to movement across a wide area, leading to average path loss and variance from this average, and on a small scale due to reflection, diffraction, and scattering causing multipath fading [15].

In this study, a readily constructed empirical model to simulate the large-scale fading in the VHF and UHF frequency bands is ITU-R P.1546 [16]. It uses interpolation and extrapolation of field curves generated from empirical data [16]. This standard describes location variability as the relation to the variation in excess path loss across the full-service area of a transmitter, encompassing all terrain influences as well as local shadowing. Thus, for a land receiving fixed/mobile antenna location the field strength E in dB(μV/m), which will be exceeded for q% of locations can be defined by [16]:

$$E(q) = E(md) + Q_i(q/100)\sigma_L \tag{1}$$

where $E(md)$ is the median electric field value interpolated or extrapolated from the electric field curves. $Q_i(x)$ is the inverse complementary cumulative normal distribution as a function of probability, and σ_L is the standard deviation of the Gaussian distribution of the local means in the area under study, where representative values of σ_L are 8, 10, and 12 dB for urban, suburban, and open areas respectively [16].

Rayleigh and Rician statistical distributions are used in wireless communications to simulate fading channels modeling the impact of multipath propagation, which occurs when signals traverse many routes as a result of reflections, diffraction, and scattering. The received envelope amplitude would then follow Rician distribution, which can be expressed as [15]:

$$p(x) = \frac{x}{\sigma^2} \exp\left[-\frac{x^2 + A^2}{2\sigma^2}\right] I_0\left(\frac{xA}{\sigma^2}\right) \tag{2}$$

where x is the magnitude of the faded signal, σ^2 is the variance, and A is the specular component, where $A > 0$.

A parameter K is often used to describe the Rician distribution as [15]:

$$K = A^2/\left(2\sigma^2\right) \tag{3}$$

As the specular component approaches zero, the Rician pdf approximates Rayleigh distribution, described as [15]:

$$p\left(x_0\right) = \frac{x}{\sigma^2} \exp\left[-\frac{x^2}{2\sigma^2}\right], \text{ for } x \geq 0 \tag{4}$$

From the receiver side of view, Additive White Gaussian Noise (AWGN) exists and it is accompanied by other HWI such as frequency offset, phase offset, sample rate offset, and IQ imbalance. The received signal in the form in-phase and quadrature terms can be expressed as:

$$r[n] = r_I[n] + j \cdot \overrightarrow{r_0}[n] \tag{5}$$

where $r[n]$ is the received complex signal and $r_I[n]$, and $r_Q[n]$ are the in-phase and quadrature components of the signals, where $r[n]$ when subjected to the imperfection mentioned can be described as:

$$\begin{aligned} r[n] = (\ & m\left[n + \Delta T_s\right]) \\ & \cdot e^{j(2\pi \Delta f [n+\Delta T_s] + \Delta\phi)} \\ & + n_w[n] \end{aligned} \tag{6}$$

where $m\left[nT_s\right]$ is the front-end received signal, ΔT_s is the sampling rate offset, Δf is the frequency offset, $\Delta\phi$ is the phase offset, and $n_w[n]$ is the AWGN with unit variance. IQ imbalance, as a type of HWI, will alter the perfect reception denoted in equation (5) to be expressed as:

$$\begin{aligned} & r[n] = kI \cdot r_I[n] + j \\ & \cdot (kQ \cdot \cos(\theta) \cdot r_Q[n] + kQ \\ & \cdot \sin(\theta) \cdot r_I[n]) \end{aligned} \tag{7}$$

where kI, and kQ are the linear in-phase and quadrature gains, while θ is the phase difference between the two paths. It should be noted that the range of values for these HWI is hardware-specific and will be defined in the next section.

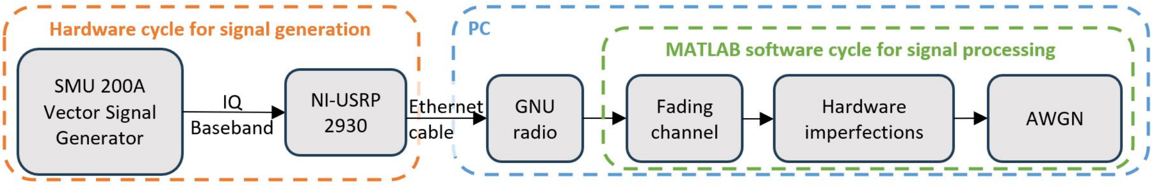

FIGURE 1. Block diagram of the dataset generation.

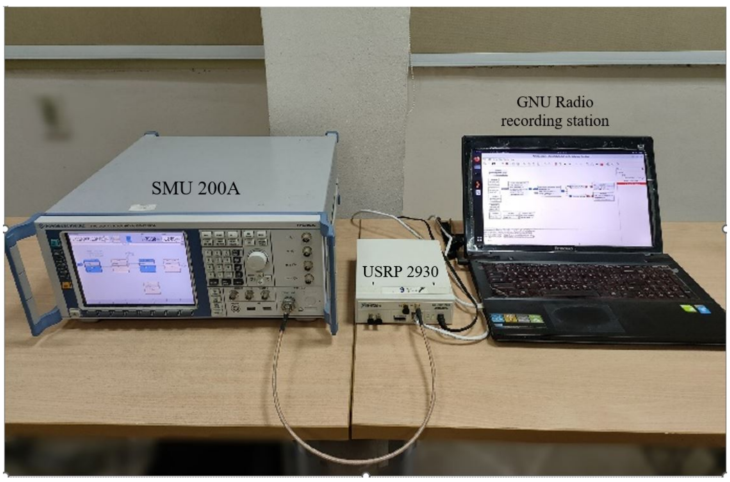

FIGURE 2. Recording the dataset setup.

3. **Proposed Framework.** The datasets with the highest number of citations in the field of deep-learning-based signal processing include RML 2016 [3], RML 2018 [4], and Sig 53 [5], specifically dedicated to the modulation recognition job, where it focuses on identifying the modulation scheme, while communication system reconnaissance explores parameters like transmission capacity, baud rate, and signal coding. A dataset meeting the task's requirements must be developed.

The proposed dataset consists of digitally modulated signals, with 106,000 examples for each signal type, totaling 1,272,000 examples in the dataset. The dataset is equally distributed across the signal types and covers a range of SNR values from -20 dB to 20 dB with a step of 2 dB. Each signal example consists of 1024 complex IQ samples. The sampling rate was set at 35 kHz to meet the Nyquist rate for all desired signals, resulting in a signal duration of around 30 milliseconds for each example.

The signal types listed in Table 1 are generated as illustrated on 1, where the signal vector generator SMU 200A is used for generating baseband signals in IQ format. SDR-based receiver, NI-USRP-2930, is connected directly to SMU 200A signal generator through coaxial cable and SDR-based receiver is controlled by Personal Computer (PC) via Ethernet LAN cable, as shown in Fig. 2. GNU Radio and MATLAB soft-wares are used for recording, controlling and signal processing operations. In order to synthesize dataset example, the generated signal is subjected to various data augmentations as detailed in section 2.2, referring to as the fading channel and hardware imperfection blocks in Fig. 1. The SC_2024 dataset will be accessible on Google Drive[1] shortly after it is published.

4. **Simulation Results and Analysis.** The proposed work is divided into three key sections: dataset creation, signal processing, and model training and testing with analysis. Five dataset versions are synthesized, where each of them includes signal examples that exhibit a combination of shadowing effects in rural regions, small-scale fading with an equal probability of being either Rayleigh or Rician fading, AWGN, and HWI. Table 2 lists the randomly selected ranges of values for each effect along with their distribution.

Table 2. Dynamic Range of Values Utilized

Dynamic Parameter	**Distribution**
Vehicle velocity [km/h]	$U(0, 60)$
K-factor [dB]	$N(3, 1)$
Path delay [usec]	$U(0.05, 10)$
Average path fading attenuation [dB]	$U(-3, 15)$
SNR [dB]	$[-20, 20]$
Frequency accuracy [ppb]	$U(-25, 25)$
Phase offset [degree]	$U(0, 2\pi)$
Sample rate offset [ppb]	$U(-25, 25)$
Inphase gain (IQ imbalance)	$U(0.0583, 0.1818)$
Quadrature gain (IQ imbalance)	$U(0.0583, 0.1818)$

The first dataset contains signal examples affected by AWGN only, the second dataset signals are affected by both AWGN and HWI, the third dataset has the effects of AWGN, HWI, and small-scale fading (SSF), and the fourth dataset consists of signals impacted by AWGN, HWI, and shadowing (SH). The fifth dataset has signals incorporated with AWGN, HWI, SSF, and SH.

[1] https://drive.google.com/drive/folders/1X7fg0dOEX-oYm70PMN8KamnhA3GdQqbR

In our simulation, the performance measure between the synthesized five dataset versions is the accuracy of a unified classifier ResNet model [6], which has the network structure as illustrated in table 3. Thus, Five trained models will correspond to each version of the synthesized dataset that represent variation of signal conditioning.

Table 3. ResNet Network Layout [6]

Layer	Output dimensions
Input	2×1024
Residual Stack	32×512
Residual Stack	32×256
Residual Stack	32×128
Residual Stack	32×64
Residual Stack	32×32
Residual Stack	32×16
Dense + SeLU	128
Dense + SeLU	128
Dense + softmax	24

For evaluation and analysis, testing ResNet models will be conducted on fifteen case study studies involving various combinations of testing five different trained models on five different datasets to analyze the proposed dataset synthesizing frameworkated in Table 4.

Table 4. Simulated case study studies

		ResNet model Trained on Dataset version :				
		AWGN	**AWGN + HWI**	**AWGN + HWI + SSF**	**AWGN + HWI + SH**	**AWGN + HW + SH +SSF**
Testing dataset version:	**AWGN**	case study 1	case study 2			
	AWGN + HWI	case study 3	case study 4			
	AWGN + HWI+ SSF			case study 5	case study 6	case study 7
	AWGN + HWI + SH			case study 8	case study 9	case study 10
	AWGN + HWI + SH +SSF	case study 11	case study 12	case study 13	case study 14	case study 15

The performance evaluation starts with measuring the accuracy of the trained model when tested on a the proposed dataset , case study 15 ,representing the most challenging scenario with all propagation channel effects and hardware imperfections, as shown in Fig.3. case study 15 has the highest accuracy compared to case studies 13 and 14,that considers testing trained models on HWI with shadowing effect or small scale fading separately, with an average rise in accuracy of 3% to 4%, and 6% to 12% over case studies 11 and 12,that considers testing trained models on only SNR effects only or SNR effects with HWI, for SNR values greater than 0 dB.

Due to significant disparities in the trained model's accuracy results, as shown in Fig. 3, between case studies (11, 12) and case studies (13, 14, 15), each group was subsequently analyzed in greater depth separately. case study studies 11 and 12 results were produced by a model trained on a data set that considers only AWGN effect, and another on AWGN and HWI.

Case studies 1, 2, 3, and 4 are dedicated to compare between trained classifiers on data sets that considers only SNR effect and that considers both SNR and HWI,as shown in Fig. 5. A notable remark in the results of the classifiers' accuracy is that the AWGN has a more obvious impact on accuracy compared to the HWI effect. The difference in the classifiers' accuracy in case studies 1, 2, 3, and 4 is particularly evident in the

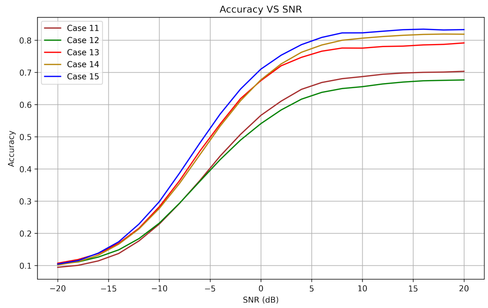

FIGURE 3. Accuracy Vs SNR of case study 11 to 15.

Signal-to-Noise Ratio (SNR) range of -5 dB to 2 dB, where case study 3 outperforms case study 2 by an accuracy margin ranging from 3% to 10%. case study 1 and case study 4 have the highest performance, as expected where the trained models are tested on the datasets used to train their respective classifiers. Fig. 5 compares case studies 5, 6, 8,

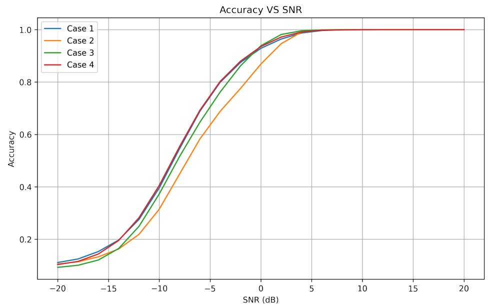

FIGURE 4. Accuracy Vs SNR of case study 1 to case study 4.

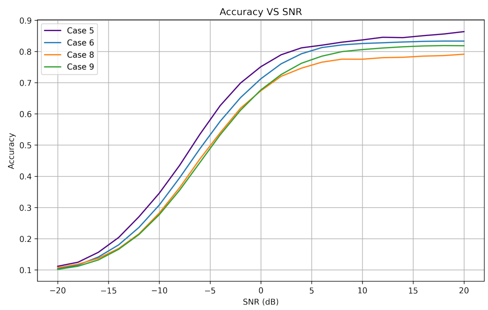

FIGURE 5. Accuracy Vs SNR of case studys 5, 6, 8, 9.

and 9 to analyze the significance of the fading effects , including large-scale fading like the shadowing effect, and small-scale fading such as multipath fading. case study 5 and 9 tested their corresponding classifiers on the same dataset used to train them. case study 5 accuracy results are the highest among the case studies, while case study 9 classifier accuracy ranks third. This shows the significant impact of the shadowing effect when included in the dataset, reducing the average accuracy by 4.8% across the whole SNR range compared to case study 5. case study 6 has the lowest accuracy, dropping by 1% compared to case study 9, while case study 8 had an average accuracy decline of 2.5% compared to case study 5. Fig. 6 displays the classifier's accuracy for case studies 5, 7, 13, and 15 versus SNR. Results indicate the classifier's accuracy in case study 5 outperforms case study 7 with a difference of 1.72% in accuracy for SNR more than 0 dB, while at the same time classifier's accuracy in case study 15 exceeds over classifier's accuracy in case study 13 in a more realistic fading environment. case study 5 was 6.36% higher for SNR more than 0 dB, showing the robustness of the model trained on shadowing.

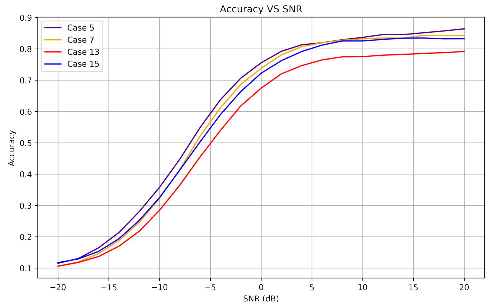

FIGURE 6. Accuracy Vs SNR of case study 5, 7, 13, 15.

5. **Conclusions.** In this paper, we proposed a robust framework for synthesizing a dataset of communication standard signals that are used in civilian and military applications and have related characteristics, which represent limitations for applying modulation recognition to classify between them. The ResNet-based neural network model, which is trained on the proposed dataset for signal classification, which considered real-world signal environment conditions, achieved notable enhancements in accuracy over the results of trained ResNet models on other datasets that considered separately shadowing effect, as a large scale fading type, or small scale fading effect and / or HWI effects with variable SNR operational range scenarios.

The comparative evaluations emphasized the strength of the proposed dataset synthesizing framework, where simulation results and analysis demonstrated that the proposed dataset that considers real-world communication scenario, for SNR values greater than 0 dB, achieved an average rise in accuracy of 3% to 4% w.r.t. results of training on data sets that ignored inclusion of large and small channel fading effects and 6% to 12% w.r.t. results trained on data set of the simplest conditions. Future work will be extended to signal classification of wideband signals and mixed signals scenarios that are challenging signal processing tasks in wireless communication systems.

References.

[1] Quoc-Viet Pham et al. "Intelligent Radio Signal Processing: A Survey". In: *IEEE Access* 9 (2021), pp. 83818–83850. DOI: 10.1109/ACCESS.2021.3087136.
[2] Noureddine El-Haryqy, Zhour Madini, and Zouine. "A Survey on Automatic Signal Detection Using Deep Learning". In: vol. 0. 0. 2023, pp. 1–6. DOI: 10.1109/ICOA58279.2023.10308833.
[3] Huaming Wu et al. *Deep learning driven wireless communications and mobile computing*. 2019. DOI: https://doi.org/10.1155/2019/5629572.
[4] Merima Kulin et al. "A Survey on Machine Learning-Based Performance Improvement of Wireless Networks: PHY, MAC and Network Layer". In: *Electronics* 10.3 (2021). ISSN: 2079-9292. DOI: 10.3390/electronics10030318. URL: https://www.mdpi.com/2079-9292/10/3/318.
[5] Timothy O'Shea and Nathan West. "Radio Machine Learning Dataset Generation with GNU Radio". In: *Proceedings of the GNU Radio Conference* 1.1 (2016). URL: https://pubs.gnuradio.org/index.php/grcon/article/view/11.

[6] Timothy James O'Shea, Tamoghna Roy, and T. Charles Clancy. "Over-the-Air Deep Learning Based Radio Signal Classification". In: *IEEE Journal of Selected Topics in Signal Processing* 12.1 (2018), pp. 168–179. DOI: 10.1109/JSTSP.2018.2797022.

[7] Luke Boegner et al. "Large Scale Radio Frequency Signal Classification". In: *arXiv preprint arXiv:2207.09918* (2022). URL: https://arxiv.org/pdf/2207.09918.

[8] R. David Badger, Kristopher H. Jung, and Minje Kim. "An Open-Sourced Time-Frequency Domain RF Classification Framework". In: *2021 29th European Signal Processing Conference (EUSIPCO)*. 2021, pp. 1701–1705. DOI: 10.23919/EUSIPCO54536.2021.9615929.

[9] R. H. Coase. "The Federal Communications Commission". In: *The Journal of Law and Economics* 2 (1959), pp. 1–40. DOI: 10.1086/466549. eprint: https://doi.org/10.1086/466549. URL: https://doi.org/10.1086/466549.

[10] April 2000 T. I. A. E. I. A. (TIA/EIA). *TIA/EIA-102.BAAC (April 2000)*. Available: https://store.accuristech.com/standards/tia-ansi-tia-102-baac-d?product_id=2594307.

[11] Rudi Bekkers. *Mobile Telecommunications Standards: Gsm, Umts, Tetra, and Ermes*. Artech House, 2001.

[12] *TECHNICAL TBR 023, TBR 023, E. T. S. I. (ETSI), October 1996 1996.* https://www.etsi.org/deliver/etsi_tbr/001_099/023/01_60/tbr_023e01p.pdf.

[13] T. I. Association. *Mobile station-base station compatibility standard for dual-mode wideband spread spectrum cellular system"*. https://standards.globalspec.com/std/945351/tia-95. TIA/EIA/IS-95, 1993.

[14] R. STD. "Research & Development Center for Radio Systems (RCR)". In: *Personal digital cellular telecommunication system RCR standard* (1991). URL: http://www.arib.or.jp/english/html/overview/doc/5-STD-27_L-1p3-E.pdf.

[15] Bernard Sklar. *Digital communications: fundamentals and applications*. Pearson, 2021.

[16] I. T. U. (ITU). *Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3000 MHz*. https://www.itu.int/rec/R-REC-P.1546/en. 2001.