Original scientific paper

Comparative Performance of Mean-Shift and Mode-Shift Outlier Detection of Wireless Channel Multipaths

Jojo Blanza

***Abstract* — Wireless channel models generate multipaths clustered according to delay, angle of departure, and angle of arrival. The multipaths serve as datasets in the performance evaluation of a proposed wireless communication system. These datasets contain outliers in clustering the multipaths. Outlier removal improves clustering performance. This study compares the performance of mean-shift and mode-shift detectors in identifying outliers in wireless channel model datasets. Mode-shift detected more outliers in the IMT-2020 dataset, whereas mean-shift detected more outliers in the other channel-model datasets. The 2·*SD* (Standard Deviation) threshold yields the most outliers across all channel scenarios in the datasets. k-medoids performs best on outlier-free datasets, while DBSCAN achieves similar performance on the outlier-free IMT-2020 dataset, after pruning with both outlier methods.**

***Keywords* — channel model, clustering algorithm, multipath clustering, outlier detection**

I. Introduction

THE performance of a proposed wireless communication system is analyzed using wireless channel models before its implementation. Among popular channel models are European Cooperation in Science and Technology (COST 2100) [1], International Mobile Telecommunications-2020 (IMT-2020) [2], Quasi-Deterministic Radio Channel Generator (QuaDRiGa) [3], and Wireless World Initiative New Radio II (WINNER II) [4]. These channel models generate multipaths that are grouped as clusters based on delay, angle of departure, and angle of arrival. They served as datasets for multipath clustering to determine which clustering algorithm best suits the data generated by measurement campaigns. However, the datasets generated by the channel models contain outliers. Clustering performance can be improved by removing these outliers. In [5], the outliers in the COST 2100 indoor scenarios were analyzed using a mean-shift outlier detector, while [6] examined the IMT-2020 dataset using the same method. This paper applies both the mean-shift and medoid-shift outlier detectors [7] to identify outliers in wireless channel model datasets. Their performance in detecting and removing outliers is compared. The datasets, with and without outliers, are clustered to examine the effects of the two outlier detection methods on clustering performance.

Paper received April 11, 2026; accepted July 29, 2026. Date of publication: August 29, 2026. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Nataša Nešković.

Jojo Blanza is with the Faculty of Engineering, University of Santo Tomas, Philippines,
(phone: +63-2-8731-4041; e-mail: jfblanza@ust.edu.ph).

The paper is organized as follows: Section II describes the methodology, Section III discusses the results, and Section IV concludes the study.

II. Methodology

The methodology of the study is shown in Fig. 1. The datasets were generated using the COST 2100, IMT-2020, QuaDRiGa, and WINNER II channel models. The datasets are available via IEEE DataPort [8], [9]. COST 2100 has eight channel scenarios, namely: Indoor B1 LOS Single Link (Indo B1), Indoor B2 LOS Single Link (Indo B2), Semi-Urban B1 LOS Single Link (SUB1LSL), Semi-Urban B2 LOS Single Link (SUB2LSL), Semi-Urban B1 NLOS Single Link (SUB1NSL), Semi-Urban B2 NLOS Single Link (SUB2NSL), Semi-Urban B1 LOS Multiple Links (SUB1LML), and Semi-Urban B2 LOS Multiple Links (SUB2LML). The following are the nine channel scenarios of IMT-2020: InH A LOS (InHL), InH A NLOS (InHN), RMa A LOS (RMaL), RMA A NLOS (RMaN), RMa A O2I (RMaO), UMa A LOS (UMaAL), UMa A NLOS (UMaAN), UMi A LOS (UMiAL), and UMi A NLOS (UMiAN). On the other hand, QuaDRiGa has eight channel scenarios: BERLIN UMa LOS (BUMaL), BERLIN UMa NLOS (BUMaN), BERLIN UMi Campus LOS (UMiCL), BERLIN UMi Campus NLOS (UMiCN), BERLIN UMi Square LOS (UMiSL), BERLIN UMi Square NLOS (UMiSN), Industrial LOS (IndusL), and Industrial NLOS (IndusN). Lastly, WINNER II has six channel scenarios: Indoor A1 LOS (IndoL), Indoor A1 NLOS (IndoN), UMa C2 LOS (UMaL), UMa C2 NLOS (UMaN), UMi B1 LOS (UMiL), and UMi B1 NLOS (UMiN).

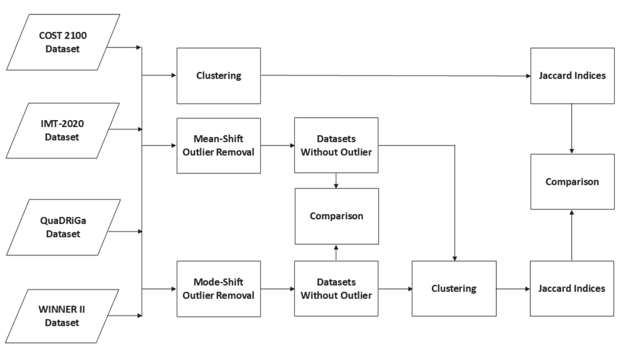

Fig. 1. Methodology flowchart.

K-means [10], k-medoids [11], Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [12], AC-Single [13], and spectral clustering (SC) [14] cluster the datasets with outliers. The Jaccard index measures clustering accuracy. The datasets are then processed for outlier detection and removal using mean-shift and mode-shift outlier detectors. The study uses three well-known outlier thresholds: $2 \cdot SD$, $2.5 \cdot SD$, and $3 \cdot SD$, where SD is the standard deviation. This means that any data with an outlier score above the outlier threshold is considered an outlier. The outlier-free datasets are clustered again using the same clustering algorithms. The performance of the two outlier detectors is compared, along with the clustering performance of the algorithms, on datasets with and without outliers.

III. Results and Discussion

The MATLAB implementation [15] of the clustering algorithms is used to cluster the original datasets with outliers. Their clustering performance is reported in [16] together with the number of outliers detected by the mean-shift method. The performance of the clustering algorithms on outlier-free datasets is also analyzed in the paper.

Outliers in the dataset are detected and removed using the two outlier detectors. Table 1 presents the number of outliers per outlier threshold for each channel scenario using the mode-shift method. Numbers in bold indicate a greater number of outliers generated by mode-shift compared to the mean-shift, while italicized numbers denote a lesser number of outliers. The regular font represents the same number of outliers generated by the two methods.

Table 1: Number of outliers using mode-shift outlier detector.

Channel Model	*Channel Scenario*	*2*SD*	*2.5*SD*	*3*SD*
	Indo B1	**65**	**48**	**32**
	Indo B2	*55*	*38*	*30*
	SUB1LSL	*1,057*	**577**	*311*
COST 2100	SUB2LSL	*1,160*	*632*	**335**
	SUB1NSL	*1,930*	**1,107**	*575*
	SUB2NSL	*1,920*	1,077	*597*
	SUB1LML	*2,213*	*1,376*	*830*
	SUB2LML	*2,153*	**1,332**	793
	InHL	**506**	**71**	**4**
	InHN	**2,465**	**544**	**100**
	RMaL	*33,229*	**10,730**	*1,435*
IMT-2020	RMaN	*37,365*	*17,496*	*4,506*
	RMaO	*68,822*	**17,969**	**116**
	UMaAL	**3,134**	*194*	*11*
	UMaAN	**63,719**	**21,619**	**5,044**
	UMiAL	**9,212**	**1,678**	274
	UMiAN	**23,860**	**6,465**	**1,551**
	BUMaL	*28,437*	*16,046*	*7,597*
	BUMaN	*43,599*	*24,773*	*13,782*
	UMiCL	*9,944*	*3,181*	*963*
QuaDRiGa	UMiCN	*18,883*	*7,606*	*2,526*
	UMiSL	*11,544*	*6,284*	*2,673*
	UMiSN	*20,059*	*11,646*	*5,989*
	IndusL	*10,608*	*6,639*	*4,152*
	IndusN	*11,086*	*6,817*	*4,189*
	IndoL	*3,833*	*1,016*	*71*
	IndoN	*6,579*	*3,650*	*1,963*
WINNER II	UMaL	*14,699*	*6,036*	*1,479*
	UMaN	*35,078*	*20,713*	*11,949*
	UMiL	*8,429*	*5,307*	*3,117*
	UMiN	*11,215*	*6,662*	*4,164*

The number of outliers decreases as the outlier threshold increases across all channel scenarios. This shows that the $3 \cdot SD$ threshold is less prone to outliers than the $2 \cdot SD$ threshold. At the $2 \cdot SD$ threshold, the higher the total number of multipaths for a given channel scenario, the higher the number of outliers. Mode-shift detected more outliers in the IMT-2020 dataset, while mean-shift detected more outliers in the other datasets. Across all channel scenarios of QuaDRiGa and WINNER II datasets, the mean-shift method detected more outliers.

Using mode-shift as the outlier detector, the outlier-free datasets are clustered again using the same clustering algorithms. Table 2 gives clustering performance on the outlier-free datasets using the $2 \cdot SD$ threshold. Jaccard indices in bold indicate better clustering performance when mode-shift is used for outlier detection, whereas those in regular font denote better performance with mean-shift as the outlier detector. Italicized values suggest the same clustering performance for both outlier methods. There is no trend to infer, as indices vary. But it is worth noting that clustering performance across all algorithms is better with the mean-shift outlier detector in COST 2100 multiple links. Also, across the IMT-2020 dataset, DBSCAN performs the same for both outlier detectors.

Table 2: Performance of algorithms in clustering the datasets free of outliers using the $2 \cdot SD$ threshold of the mode-shift method. No data means that the computational memory requirement exceeds the computer memory.

Channel Model	*Channel Scenario*	*k-means*	*k-medoids*	*DBSCAN*	*AC-Single*	*SC*
	Indo B1	**0.5392**	**0.7572**	0.3444	**0.6433**	0.2619
	Indo B2	**0.5184**	**0.6221**	0.2005	**0.5264**	0.2070
	SUB1LSL	**0.0231**	0.0235	**0.0219**	**0.0273**	**0.0282**
COST 2100	SUB2LSL	**0.0242**	0.0209	0.0212	**0.0256**	**0.0237**
	SUB1NSL	0.0229	**0.0220**	0.0235	**0.0232**	0.0238
	SUB2NSL	**0.0258**	**0.0241**	0.0227	**0.0233**	**0.0239**
	SUB1LML	0.0106	0.0110	0.0108	0.0107	0.0130
	SUB2LML	0.0133	0.0129	0.0111	0.0113	0.0139
	InHL	0.0468	**0.0490**	*0.0345*	**0.0362**	**0.0458**
	InHN	**0.0341**	**0.0331**	*0.0270*	0.0274	0.0315
	RMaL	0.0557	0.0557	*0.0476*	0.0476	0.0556
IMT-2020	RMaN	0.0526	**0.0528**	*0.0526*	*0.0527*	0.0528
	RMaO	**0.0531**	**0.0532**	*0.0526*	0.0526	-
	UMaAL	0.0529	0.0527	*0.0435*	*0.0436*	0.0497
	UMaAN	0.0281	0.0283	*0.0256*	-	-
	UMiAL	0.0494	0.0495	*0.0435*	0.0435	0.0488
	UMiAN	0.0287	0.0290	*0.0270*	*0.0271*	0.0284
	BUMaL	**0.0632**	0.0624	**0.0426**	*0.0346*	0.0453
	BUMaN	**0.0356**	**0.0358**	0.0305	0.0204	**0.0288**
	UMiCL	**0.0826**	**0.0818**	*0.0809*	*0.0437*	**0.0628**
QuaDRiGa	UMiCN	**0.0491**	0.0482	0.0438	*0.0258*	0.0358
	UMiSL	0.0772	**0.0774**	0.0724	*0.0437*	**0.0598**
	UMiSN	**0.0479**	**0.0480**	0.0379	*0.0258*	**0.0352**
	IndusL	0.0252	**0.0247**	0.0279	**0.0208**	**0.0252**
	IndusN	0.0231	0.0225	*0.0278*	**0.0199**	0.0239
	IndoL	**0.0747**	0.0740	0.0772	0.0440	**0.0590**
	IndoN	0.0593	**0.0608**	*0.0531*	0.0326	**0.0458**
WINNER II	UMaL	*0.1118*	*0.1114*	**0.1003**	*0.0668*	**0.0873**
	UMaN	0.0442	0.0429	0.0353	*0.0257*	**0.0336**
	UMiL	**0.0989**	0.0951	0.0901	**0.0670**	*0.0780*
	UMiN	**0.0431**	**0.0420**	*0.0363*	*0.0324*	**0.0412**

Table 3 shows the mean Jaccard indices of the algorithms on outlier-free clustering datasets using the $2 \cdot SD$ threshold for both outlier detectors. k-medoids has the best clustering performance for both outlier methods.

Table 3: Mean Jaccard indices of algorithms in clustering the outlier-free datasets using the $2 \cdot SD$ threshold.

Outlier Detection	*k-means*	*k-medoids*	*DBSCAN*	*AC-Single*	*SC*
Mean-shift	0.0728	0.0831	0.0576	0.0654	0.0554
Mode-shift	0.0779	0.0879	0.0570	0.0708	0.0541

The use of mode-shift in outlier detection enhanced the clustering performance of k-means, k-medoids, and AC-Single, while mean-shift improves the clustering performance of DBSCAN and SC.

Table 4 presents the clustering performance on the outlier-free datasets. The datasets were pruned using the $2.5{\cdot}SD$ threshold of the mode-shift detector. As with the $2{\cdot}SD$ threshold, there is no trend in the indices, as they vary across channel scenarios. However, the clustering performance of all algorithms improves when the mean-shift detector is used in COST 2100 multiple links. Also, DBSCAN performs the same on the IMT-2020 dataset pruned by both detectors. Lastly, DBSCAN performs better in QuaDRiGa when mean-shift is used as the outlier detector, except in IndusN, where the same clustering accuracy is achieved by DBSCAN clustering the dataset pruned by both detectors.

TABLE 4: PERFORMANCE OF ALGORITHMS IN CLUSTERING THE DATASETS FREE OF OUTLIERS USING THE $2.5{\cdot}SD$ THRESHOLD OF THE MODE-SHIFT METHOD.

Channel Model	*Channel Scenario*	*k-means*	*k-medoids*	*DBSCAN*	*AC-Single*	*SC*
	Indo B1	**0.5868**	**0.6935**	0.3444	**0.5943**	**0.2696**
	Indo B2	**0.5262**	**0.5766**	0.2005	**0.4641**	0.1977
	SUB1LSL	0.0247	0.0234	**0.0223**	**0.0255**	0.0254
COST	SUB2LSL	**0.0229**	0.0213	0.0211	0.0250	0.0234
2100	SUB1NSL	0.0216	**0.0228**	0.0232	**0.0229**	**0.0250**
	SUB2NSL	**0.0224**	**0.0235**	**0.0227**	0.0224	**0.0279**
	SUB1LML	0.0104	0.0125	0.0105	0.0113	0.0127
	SUB2LML	0.0123	0.0132	0.0112	0.0111	0.0149
	InHL	0.0455	0.0474	*0.0345*	0.0361	**0.0454**
	InHN	0.0330	0.0334	*0.0270*	**0.0274**	0.0306
	RMaL	0.0559	**0.0565**	*0.0476*	*0.0477*	0.0573
IMT-	RMaN	**0.0530**	**0.0531**	*0.0526*	*0.0527*	0.0524
2020	RMaO	**0.0529**	**0.0528**	*0.0526*	*0.0526*	-
	UMaAL	0.0529	0.0531	*0.0435*	*0.0436*	**0.0498**
	UMaAN	0.0312	0.0314	*0.0256*	-	-
	UMiAL	**0.0500**	0.0499	*0.0435*	0.0435	0.0500
	UMiAN	0.0294	0.0293	*0.0270*	*0.0271*	0.0290
	BUMaL	**0.0633**	0.0631	0.0387	*0.0346*	0.0458
	BUMaN	**0.0358**	**0.0362**	0.0307	0.0204	**0.0276**
	UMiCL	**0.0827**	**0.0819**	0.0812	0.0436	**0.0577**
QuaDRiGa	UMiCN	*0.0482*	**0.0485**	0.0440	*0.0258*	0.0334
	UMiSL	**0.0775**	0.0763	0.0733	*0.0437*	**0.0597**
	UMiSN	**0.0473**	0.0470	0.0380	0.0257	**0.0343**
	IndusL	**0.0255**	**0.0259**	0.0281	**0.0208**	**0.0248**
	IndusN	*0.0227*	**0.0229**	*0.0281*	*0.0199*	**0.0238**
	IndoL	**0.0730**	**0.0752**	0.0779	**0.0441**	0.0577
	IndoN	0.0605	0.0609	**0.0536**	0.0325	**0.0458**
WINNER	UMaL	**0.1125**	**0.1124**	*0.0667*	*0.0668*	**0.0879**
II	UMaN	0.0433	0.0418	*0.0356*	*0.0257*	**0.0324**
	UMiL	0.0965	0.0944	0.0913	*0.0670*	0.0765
	UMiN	0.0424	**0.0430**	*0.0365*	*0.0324*	**0.0407**

Fig. 2 illustrates the mean Jaccard indices of the clustering algorithms across all datasets free of outliers.

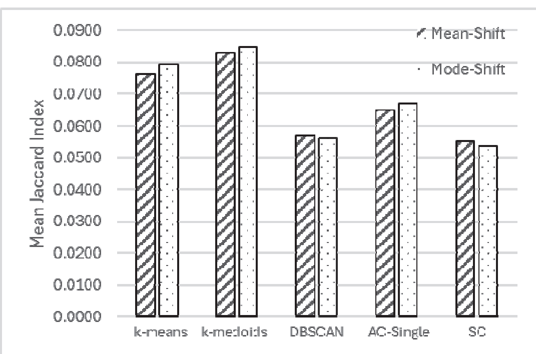

Fig. 2. Mean Jaccard indices of algorithms in clustering the outlier-free datasets using the $2.5{\cdot}SD$ threshold.

The outliers are removed using the $2.5{\cdot}SD$ threshold for both outlier methods. k-medoids still has the best clustering performance. When mode-shift is used for outlier detection, k-means, k-medoids, and AC-Single have better clustering performance, whereas DBSCAN and SC fared better with the mean-shift.

Table 5 presents the clustering performance on datasets free of outliers using the $3{\cdot}SD$ threshold of the mode-shift method. Mode-shift improves the clustering performance of all algorithms in COST 2100 Indo B2, including k-medoids, when clustering WINNER II. On the other hand, mean-shift improves the clustering performance of all algorithms in COST 2100 SUB1NSL and multiple links. The clustering performance of DBSCAN is the same when it clusters the IMT-2020 dataset pruned by both outlier methods.

TABLE 5: PERFORMANCE OF ALGORITHMS IN CLUSTERING THE DATASETS FREE OF OUTLIERS USING THE $3{\cdot}SD$ THRESHOLD OF THE MODE-SHIFT METHOD.

Channel Model	*Channel Scenario*	*k-means*	*k-medoids*	*DBSCAN*	*AC-Single*	*SC*
	Indo B1	**0.5996**	**0.6731**	**0.3526**	**0.5478**	0.2181
	Indo B2	**0.4926**	**0.5838**	**0.2005**	**0.4620**	**0.2394**
	SUB1LSL	**0.0236**	0.0231	**0.0220**	0.0236	**0.0277**
COST	SUB2LSL	**0.0232**	**0.0218**	0.0213	**0.0246**	**0.0251**
2100	SUB1NSL	0.0222	0.0226	0.0234	0.0227	0.0232
	SUB2NSL	0.0216	0.0235	*0.0226*	*0.0228*	0.0244
	SUB1LML	0.0122	0.0130	0.0106	0.0117	0.0141
	SUB2LML	0.0119	0.0139	0.0111	0.0114	0.0167
	InHL	**0.0476**	0.0470	*0.0345*	*0.0361*	0.0453
	InHN	0.0331	**0.0335**	*0.0270*	0.0273	0.0313
	RMaL	0.0562	**0.0559**	*0.0476*	*0.0477*	0.0570
IMT-	RMaN	**0.0532**	0.0526	*0.0526*	*0.0527*	0.0527
2020	RMaO	**0.0525**	*0.0526*	*0.0526*	*0.0526*	-
	UMaAL	0.0525	**0.0533**	*0.0435*	0.0436	0.0493
	UMaAN	0.0315	0.0312	*0.0256*	-	-
	UMiAL	**0.0499**	0.0499	*0.0435*	0.0435	0.0500
	UMiAN	*0.0297*	**0.0298**	*0.0270*	*0.0271*	**0.0291**
	BUMaL	*0.0638*	**0.0635**	*0.0361*	*0.0346*	0.0426
	BUMaN	0.0349	**0.0357**	**0.0309**	*0.0205*	**0.0257**
	UMiCL	**0.0822**	**0.0819**	*0.0815*	*0.0436*	0.0555
QuaDRiGa	UMiCN	**0.0485**	0.0484	*0.0441*	*0.0258*	0.0321
	UMiSL	**0.0778**	**0.0764**	*0.0739*	*0.0437*	**0.0584**
	UMiSN	**0.0470**	**0.0474**	*0.0382*	*0.0258*	*0.0325*
	IndusL	**0.0266**	0.0252	*0.0282*	**0.0208**	0.0235
	IndusN	0.0230	**0.0235**	0.0283	**0.0200**	**0.0237**
	IndoL	**0.0750**	**0.0749**	**0.0782**	**0.0440**	0.0568
	IndoN	**0.0615**	**0.0612**	**0.0539**	*0.0325*	*0.0445*
WINNER	UMaL	**0.1147**	**0.1139**	*0.0667*	*0.0668*	0.0796
II	UMaN	**0.0447**	**0.0440**	*0.0356*	*0.0257*	**0.0309**
	UMiL	**0.0966**	**0.0954**	0.0896	**0.0670**	**0.0798**
	UMiN	0.0433	**0.0434**	*0.0366*	*0.0324*	**0.0397**

Fig. 3 presents the mean Jaccard indices of the clustering algorithms across all datasets free of outliers using the $3{\cdot}SD$ threshold. k-medoids remains the best clustering algorithm. k-means, k-medoids, DBSCAN, and AC-Single yield better clustering performance when paired with mode-shift, whereas SC performs better with mean-shift.

In general, k-medoids clusters best across all channel scenarios, applying all outlier thresholds, and using both outlier detectors. The highest clustering accuracy is achieved when k-medoids is applied to datasets free of outliers using mode-shift with a $2{\cdot}SD$ threshold. It is 3.85% better than the second-best k-medoids at a $2.5{\cdot}SD$ mode-shift threshold.

The $2{\cdot}SD$ threshold yields the best clustering performance because it detects and removes more outliers. The $3{\cdot}SD$ threshold yields the lowest clustering accuracy because it removes fewer outliers. Among the clustering algorithms, k-medoids performs best, followed by k-means, while SC has the lowest clustering accuracy.

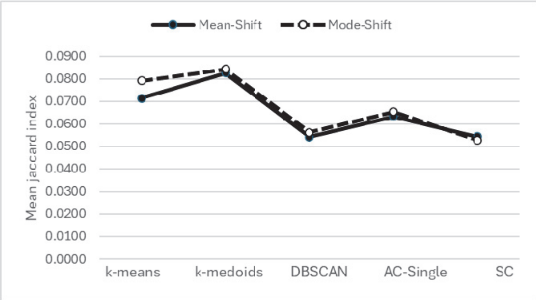

Fig. 3. Mean Jaccard indices of algorithms in clustering the outlier-free datasets using the $3{\cdot}SD$ threshold.

IV. Conclusion

The study compares the performance of mean-shift and mode-shift outlier detectors when applied to wireless multipath datasets. Mode-shift detected more outliers in the IMT-2020 dataset, while mean-shift detected more in the other datasets. $2{\cdot}SD$ outlier threshold yields the most outliers, while the $3{\cdot}SD$ threshold produces the fewest. The performance of clustering algorithms is also compared on datasets with and without outliers. Across all channel scenarios, outlier detectors, and outlier thresholds, k-medoids performs best while SC is the least accurate. k-medoids in mode-shift $2{\cdot}SD$ threshold achieves the highest clustering accuracy and is 3.85% better than the next-best, k-medoids in mode-shift $2.5{\cdot}SD$ threshold. DBSCAN achieves the same performance when clustering the IMT-2020 outlier-free dataset after pruning with both outlier detectors. For future studies, deep learning-based outlier detectors can be explored.

Acknowledgment

The author thanks the University of Santo Tomas Research Center for the Natural and Applied Sciences for its support of this work.

References

[1] R. Verdone and A. Zanella, *Pervasive Mobile and Ambient Wireless Communications: COST Action 2100 (Signals and Communication Technology)*. London: Springer, 2012.

[2] *Guidelines for Evaluation of Radio Interface Technologies for IMT-2020*, ITU-R M2412-0, 2017.

[3] F. Burkhardt, S. Jaeckel, E. Eberlein and R. Prieto-Cerdeira, "QuaDRiGa: a MIMO channel model for land mobile satellite," in *8th Eur. Conf. Antennas Propag.*, 2014, pp. 1274–1278.

[4] J. Meinilä, P. Kyösti, T. Jämsä, and L. Hentilä, "WINNER II channel models," in *Radio Technologies and Concepts for IMT-Advanced*, M. Döttling, W. Mohr, and A. Osseiran, Eds. West Sussex, England: Wiley, 2009, pp. 39–92.

[5] J. Blanza, X. Cabasal, J. Cipriano, G. Guerrero, R. Pescador, and E. Rivera, "Indoor wireless multipaths outlier detection and clustering," *J. Phys.: Conf. Ser.*, vol. 2356, 012037, pp. 1–12, 2022.

[6] A.J. Domingo, J.G. Gallardo, J.U. Nacino, J. Sevilla, P.L. Tresmaria, and J. Blanza, "Outlier detection and clustering of IMT-2020 dataset," *J. Comput. Sci. Comput. Math.*, vol. 14, Issue 02, pp. 7–11, June 2024.

[7] J. Yang, S. Rahardja, and P. Fränti, "Mean-shift outlier detection and filtering," *Pattern Recognit.*, vol. 115, 107874, pp. 1–11, 2021.

[8] J. Blanza, A. J. Teologo, and L. Materum (2019). Datasets for multipath clustering at 285 MHz and 5.3 GHz bands based on COST 2100 MIMO channel model. Available: https://dx.doi.org/10.21227/4cb9-hf81

[9] J. Blanza, E. Trinidad, and L. Materum (2023). Scedasticity descriptor of terrestrial wireless communications channels for multipath clustering datasets. Available: https://dx.doi.org/10.21227/4khe-gx08

[10] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, 1982.

[11] L. Kaufman and P. Rousseeuw, "Partitioning around medoids (Program PAM)," in *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 1990, pp. 68–125.

[12] M. Ester, H.-P. Kriegel, J. Sander, and X. Xiaowei, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. Int. Conf. Knowl. Discov. Data Min.*, 1996, pp. 226–231.

[13] R. Sibson, "SLINK: an optimally efficient algorithm for the single-link cluster method," *Comput. J.*, vol. 16, no. 1, pp. 30–34, 1973.

[14] A. Ng, Y. Weiss, and M. Jordan, "On spectral clustering: analysis and an algorithm," in *Advances in Neural Information Processing Systems*, vol. 14, T. Dietterich, S. Becker, and Z. Ghahramani, Eds. Cambridge, MA, USA: MIT Press, 2001, pp. 849–856.

[15] https://www.mathworks.com/help/stats/cluster-analysis.html?s_tid=CRUX_lftnav

[16] J. Blanza and J.B. Cipriano, "Outlier detection and clustering of fifth-generation wireless channel model datasets," *Telkomnika*, vol. 23, no. 4, pp. 871–881, 2025.