5G Performance: A Multidimensional Variability Analysis

Varshika Srinivasavaradhan, Jiayi Liu and Elizabeth M. Belding
University of California, Santa Barbara
{varshika, jiayi979, ebelding}@ucsb.edu

Abstract. 5G networks have been broadly touted as a revolution in cellular performance. However, these networks have significant architectural, spectrum and physical layer options, such that the delivered performance can be variable. The disparity in smartphone hardware and software platforms adds another layer of performance uncertainty. Our goal in this work is to characterize the impact of these features on 5G performance. To do so, we analyze a dataset of nearly 1.75 million crowdsourced Ookla® Speedtest Intelligence® cellular network measurements over three years and eight U.S. cities. We employ a novel approach by grouping Speedtest results based on both performance metrics and their deviation, while also accounting for spatial distribution and frequency band characteristics. By using statistical distance measures, we quantify the impact of multiple PHY layer and device-specific features across these multidimensional groups. We complement our in-the-wild analysis with a controlled study to validate our findings. We observe that PHY layer parameters, such as channel quality index and signal strength are the primary drivers of performance variability within each frequency range. However, between frequency ranges, user equipment hardware emerges as the dominant factor, highlighting that the equipments themselves play a critical role in determining whether users can fully utilize 5G capabilities. This underscores the importance of advancing device hardware to keep pace with the rapid evolution of network technologies.

1 Introduction

5G technology has promised be a connectivity panacea, a dense grid of intelligent wireless devices that deliver "multi-Gbps peak rates, ultra-low latency, massive capacity, and more uniform user experience" [52]. The last four years have seen an explosion of 5G network deployment across the U.S. and worldwide [4, 5, 14]. Consistent with other advanced wireless technologies, there exists a plethora of 5G architectural and physical (PHY) layer options. There is also a wide range of 5G-capable devices and chipsets, from "5G-ready" cell phones developed in the late 2010s, to modern devices that support advanced 5G features like mmWave technology and carrier aggregation for enhanced connectivity [9, 12].

This array of deployment and usage options brings into question whether all configurations receive the multi-Gbps, high-performance experience touted by marketing campaigns. Prior work has noted wide variability in 5G performance [40, 49, 60]. We build on these findings to contextualize 5G network performance based on PHY layer and device features and quantify the impact of these features on measured performance. Specifically, we ask the question: *how variable is 5G performance, and which specific network and device features have the most significant impact on measured performance?*

To answer this question, we utilize a novel dataset of 1.75M individual crowdsourced Ookla® Speedtest® measurements from eight U.S. cities across three years. Although past work has shown that crowdsourced speed measurements may contain biases due to their uncontrolled collection [27, 33, 51], they nevertheless are a critical, rich source of "in-the-wild" network performance data [44, 50, 53, 55]. Hence, we analyze the features in this dataset and apply statistical measures to quantify their impact on 5G performance. However, because of potential biases in the crowdsourced data, we supplement this dataset through experimentation and generate an additional 3.7k Speedtest measurements on three commonly-used cellular device types and cellular carriers. We confirm our analytical findings in this more controlled setting.

We begin by analyzing three years, from 2021 to 2023, of 5G crowdsourced performance trends from three independent carriers. We find wide divergence in measured download speeds, both for individual carriers and between carriers. Notably, *the bottom 30% of 5G tests perform* <u>*worse*</u> *than the top 25% of 4G tests, even in 2023.* Motivated by this wide variability and poor performance, we employ statistical distance measures to quantify the impact of the PHY layer, device and other features on 5G download speed and latency. For each carrier in our study, we first group tests based on frequency to account for differences in wave propagation characteristics. Within each frequency range, we further categorize tests by H3 resolution 9 hexagons [22][1] to control for location-specific parameters. We then analyze each group based on both average performance (speed and latency) and variability. Our novel methodological contribution is in applying Kullback-Leibler divergence to quantify the divergence between performance distributions, allowing us to systematically identify the features that most significantly impact 5G variability. By computing the statistical distance between the consistently high-performing group and each of the other groups, we identify key features that have the most significant impact on 5G performance variability. This approach allows us to identify features that affect 5G performance both within each frequency range and across the different frequency ranges. *We find that within each frequency range, cell density and PHY layer parameters, such as channel quality index (CQI), signal strength (RSRP) and signal quality (RSRQ), unsurprisingly, are the primary contributors to performance variability. However, across frequency bands, user equipment, specifically the chipset models and software, show high divergence values for all carriers in our study.* Our results demonstrate that 5G deployment alone does not guarantee high performance. Rather, careful attention to PHY configurations, modern devices, and sufficient cell tower densities are needed to reap the benefits that 5G promises.

2 Data and experiment setup

To quantify 5G performance variability and identify key influencing factors, we analyze multiple complementary datasets. In the following section, we describe each of our data sources, including crowdsourced and controlled measurements, along with our methodologies. We also discuss the limitations of our approach.

[1] An H3 geospatial index has an average area of 0.73 km^2 [24].

Speedtest Intelligence Data. The Ookla® Speedtest® platform offers users the ability to measure Internet connection quality through either a web-based portal or a dedicated mobile application [18] using a network of over 16k measurement servers world-wide [11]. Over the past decade, Speedtest has garnered extensive usage among both consumers and policymakers [1, 3, 10]. For each Speedtest, a geographically close test server with the lowest latency is automatically selected [16] and TCP connections are used to saturate the link between the client and the test server to measure network speed and latency. We utilize the Speedtest Intelligence dataset with crowdsourced measurements collected between January 1, 2021 and December 31, 2023, which we obtained through a Data Use Agreement[2] (DUA). In contrast to the publicly available aggregated Ookla data, the DUA dataset contains additional Speedtest data points, each with multiple features, such as upload/download throughput, latency, cellular carrier, user equipment (UE) and software details, chipset, and other relevant geospatial information. Because our study analyzes cellular network performance, we focus on cellular Speedtest measurements from iOS and Android smartphone applications. Tests report additional metadata, including PHY metrics (RSRP, signal-to-noise ratio (SNR), RSRQ, CQI); type of cellular radio frequency (RF) technology (2G, 3G, 4G LTE, 5G NR); RF band frequency; channel width; and available kernel memory(for Android measurements) and 5G deployment mode (non-standalone (NSA) 5G vs standalone (SA) 5G) for iOS measurements. Our dataset contains a total of 1.75 million measurements (700k 4G and 1.04M 5G measurements) collected from eight U.S. cities from three major U.S. cellular carriers. For interested readers, we enumerate the number of data points per city and carrier in Table 2 of the Appendix. Each of the cities has a population between 400k and 700k and covers an area of 90 – 350 sq. miles.

Tower Maps data. We utilize Tower Maps [20], a proprietary dataset of cell tower locations in the U.S. that includes details such as accurate location information, tower height, and construction date, to augment our analysis. While the data does not specify the radio type for each tower, it allows us to estimate the density of cellular infrastructure around Speedtest takers. By integrating this information with our data, we can analyze how the presence and concentration of nearby towers correlates with measured 5G performance, offering insight into the relationship between infrastructure density and user experience.

Controlled experiments. To verify our results from the crowdsourced dataset, we conduct controlled Speedtest measurements for 15 days in April 2024, a duration that captures both hourly and weekly usage patterns while controlling for device and location variables. We use three popular Android phone models, which were also common models in our crowdsourced dataset: Samsung Galaxy S20+ and S23, and Google Pixel 7, each with a different chipset. We utilize three phones of each model, for a total of nine experiment phones. For consistency, we purchase nine SIM cards with identical cell plans for each of the same three cellular carriers analyzed in the crowdsourced data. We conduct Speedtests at

[2] Due to our data use agreement, we are unable to disclose carrier names, specific locations, or chipset details.

regular intervals using nine SIM cards per carrier; Speedtests are run sequentially on each phone, and the cell carrier is rotated after each test set. We conduct our experiments in eight locations identified to have 5G coverage, as indicated by the National Broadband Map [17] and the carrier official websites. In many cases, multiple test sets are conducted per location, per carrier. To analyze PHY layer and cell tower information, we utilize Accuver XCAL-M [21], a professional tool that can extract information from the Qualcomm diag [2]. Since the Google Pixel 7 chipset is not compatible with XCAL-M, we run open source Android applications CellMapper and NetMonster [7, 15] on these phones during Speedtest measurements to obtain the same granular information.

Limitations. Prior research [27, 33, 51] has recognized potential shortcomings of crowdsourced network performance metrics. Crowdsourced tests are uncontrolled and hence can introduce biases related to the test-taker, geographic location, network conditions (congestion or poor service), device type and characteristics, and the lack of cellular data plan details (e.g., speed caps or throttling limits). Additionally, our dataset lacks information on features such as beam management strategies and resource blocks available, which can impact measured performance. However, our objective is to study 5G network performance as measured by Ookla Speedtest; to identify the network and device features that contribute to 5G performance variability; and to quantify the impact of these features on measured performance. We believe the impact of inherent dataset biases on our analysis is limited. Nevertheless, *we confirm all findings derived from the crowdsourced data with the data generated by our controlled experiments.*

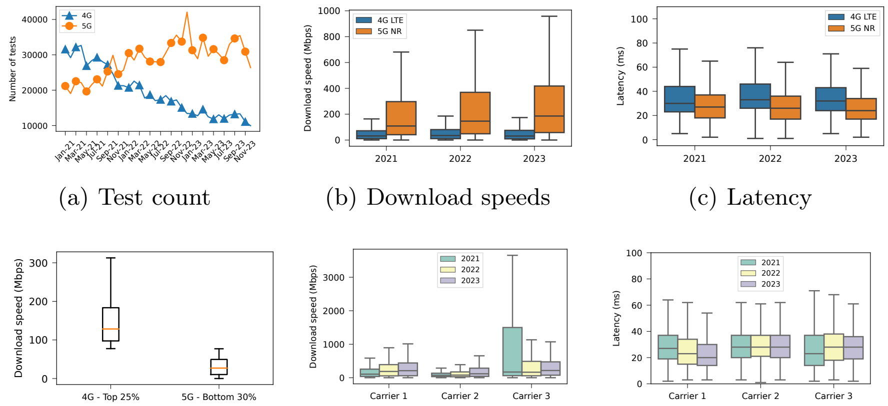
3 Dissecting 5G performance

We begin with a longitudinal analysis of 4G and 5G to characterize their evolving usage and performance in our time window, during which 5G deployment increased significantly [13, 19]. While the focus of our study is 5G, we include some initial data on 4G to contextualize the transition period from 4G to 5G. Then, we dig deeper and analyse the key factors that affect in-the-wild 5G performance and quantify the variability using statistical distance measures. We validate our results through controlled experiments. While we also analyze upload speed, we focus our presented analysis on download speed and latency. Download speed is often a key determinant of user experience [8] and the variability in download speeds is much greater than that of upload speeds. Latency also plays a crucial role, particularly for interactive applications and overall responsiveness. We confirm that the trends represented by download speeds are consistent with those in upload speeds.

3.1 Longitudinal Trends

We begin by characterizing Speedtest usage and measured performance over time to study changes that correlate with growing 5G deployment. Figure 1(a) illustrates these temporal trends, aggregated monthly over the three cellular carriers and eight cities. It indicates a general upward trajectory in the number of 5G Speedtests, while 4G Speedtests decrease by 61% during the same period.

Corresponding with the 5G test-taking trend is an increase in median 5G download speed, more than doubling over the three year period, as shown in



(a) Test count (b) Download speeds (c) Latency

(d) 2023 speed distribution (e) 5G speeds by carrier (f) 5G latency by carrier

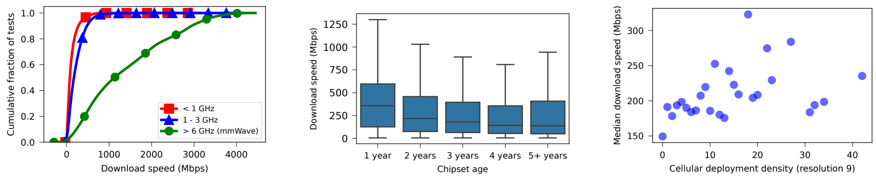
Fig. 1. Longitudinal evolution of test taking and measured download speeds.

Figure 1(b), which illustrates the range of measured download speeds annually for each cellular technology. Unlike 5G, 4G speed remains stable. A similar trend is observed in 5G latency, shown in Figure 1(c), with the median value decreasing by about 10ms from 2021 to 2023. These trends are a likely indication of increasingly widespread 5G deployment and availability, coupled with an increase in 5G-capable UEs and a movement away from 4G. Interesting, however, is the wide range in 5G download speed performance, which grows annually: in 2021, the difference between the 25^{th} and 75^{th} percentiles was 245Mbps, while in 2023 it was nearly 400Mbps. This wide, and growing, range of speeds is our first indication of the wide variability in 5G performance. Further, Figure 1(b) indicates some overlap in 4G and 5G performance, even as late as 2023. When we take a closer look, we find that, even in 2023, the bottom 30% of 5G download speeds are **worse** than the top 25% of 4G tests (Figure 1(d)). *Given this highly variable 5G performance, we ask: what key network and device features most significantly impact 5G performance?* As described earlier, the range of 5G network features, PHY layer options, and devices is wide. Our goal is to discover which of these features contribute most substantially to measured performance.

We conclude our longitudinal analysis by examining the 5G download speed and latency of each carrier individually in Figures 1(e) and 1(f), respectively. We observe a growth of over 200% in the 75^{th} percentile download speeds for carriers 1 and 2. While the latency of carrier 2 does not show huge improvements, carrier 1 exhibits a 50% decline in median latency. On the other hand, carrier 3's performance decreases significantly after 2021. While this is interesting, it is orthogonal to our analysis. We include a deeper exploration of this anomaly in the Appendix for interested readers.

3.2 Feature Analysis

In this section, we analyze measured Speedtest download speed and latency, disaggregating by test metadata to quantify performance differentials due to each



(a) Frequency range (b) Chipset age (c) Cell tower density

Fig. 2. Effect of key features in our dataset on measured downloads speeds.

feature. The test metadata we primarily focus on includes 5G RF technology, 5G PHY parameters such as RSRP, RSRQ and CQI[3]; UE features, such as device model and chipset; and cell tower density. For each feature category, we use the Speedtest datapoints that have the corresponding metadata. For brevity, we present only the findings for download speed here; we confirm our findings are consistent for latency.

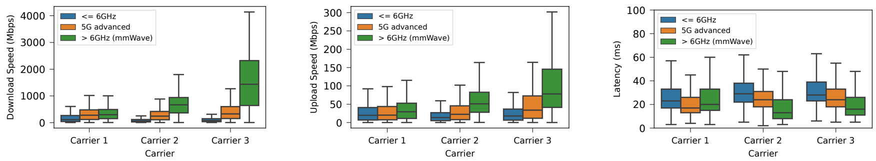
Radio frequency (RF) technology. 5G networks operate in a variety of frequency ranges: low-band (<1 GHz), mid-band (1-6 GHz), and mmWave (> 6 GHz). The characteristics of wave propagation in these bands naturally lend themselves to widely varying performance. To study the impact of these RF features, we utilize the 100k 5G tests from Android devices that contain frequency information. We have included details of the count of Speedtests across frequency bands for each year in Tables 3 and 4 of the Appendix. Figure 2(a) illustrates the significant impact radio frequency has on performance: mmWave 5G achieves median download speeds of over 1200 Mbps, a 1400% increase from the medians of sub-6 GHz speeds. On the other hand, there is minimal difference between low-band and mid-band 5G. Finally, our dataset lacked labelled measurements for NSA and SA 5G deployment modes in Android measurements. Hence, our analysis does not include these specific architectural configurations.

Device hardware. Next, we analyze the distribution of 5G download speed by chipset age[4] in Figure 2(b). We observe that the median value improves significantly with newer chipsets, with the latest models achieving speeds twice as fast as those of chipsets that are 5 years old.

Cell tower density. 726k Speedtest measurements have GPS-level location information in our dataset. We use this subset to study the relationship between cellular deployment density and measured download speeds, presented in Figure 2(c). We utilize cell tower data obtained from Tower Maps and compute the number of cell towers within the H3 resolution 9 hexagon of each Speedtest measurement; we call this cell tower density. We find that cell tower density and median download speeds have a moderate correlation (0.38). While this suggests

[3] RSRP is the average power received from a 5G reference signal. RSRQ indicates how clearly the signal can be heard over interference. CQI is a mechanism by which the UE informs the base station about channel quality.

[4] We define chipset age as the number of years since a chipset was introduced.



(a) Download speeds. (b) Upload speeds. (c) Latency.

Fig. 3. Effect of frequency on measured network performance metrics.

a positive relationship, other important factors that could impact speeds, like cellular load, are not available in our metadata, limiting further analysis.

3.3 Quantifying 5G Variability In-The-Wild

Our analysis reveals significant variability of in-the-wild 5G performance. Based on this trend, our goal in this section is to analyze this variability by estimating the impact of each feature on the measured download speed and latency to better characterize the factors that influence 5G performance. To do so, we use a subset of 370k Android measurements labelled with both PHY layer and device chipset information. Additionally, we compute cell tower density at the location of these measurements. While machine learning approaches can be valuable in analyzing large datasets such as ours, any biases in the crowdsourced Speedtest data, as well as the limited information about signal propagation and antenna beamforming characteristics, pose challenges for these methods in this context. Hence, we instead apply a statistical approach that quantifies the (dis)similarity between two distributions. This method allows us to directly compare features across performance groups and identify the ones that differ most between them.

Statistical distance. Statistical distance quantifies the distance between two statistical objects, such as two probability distributions or samples [28], and is typically used in machine learning for anomaly detection, classification, and model evaluation [31, 34, 57, 59]. We apply *Kullback-Leibler (KL) divergence* to this new context; KL divergence quantifies the dissimilarity between two probability distributions, ranges from 0 (identical) to infinity (highly dissimilar), and can be applied to both categorical and continuous features. We describe KL divergence in more detail in the Appendix. To apply the divergence model, we need more than one distribution for comparison. We utilize two key observations from our preliminary analysis. First, different carriers adopt distinct deployment strategies, leading to significant performance variations. Second, 5G operates across different frequency ranges, each with its own signal propagation characteristics, which directly influences performance; we show in Figure 3 the impact of frequency bands on network performance metrics for each carrier in our study. Based on these observations, we employ a novel methodology to group the crowdsourced dataset in a way that isolates specific test variables and controls for inherent data variations. Figure 4 gives an overview of our methodology, which we describe in detail next.

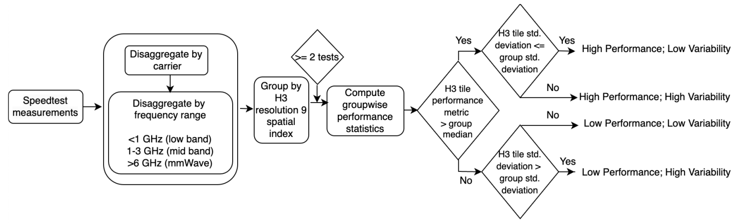


Fig. 4. Classification of tests for distance calculation.

Group	$<$ 1 GHz **(1)**	1-6 GHz **(2)**	$>$ 6 GHz **(3)**
High Performance – Low Variability **(A)**	A1	A2	A3
High Performance – High Variability **(B)**	B1	B2	B3
Low Performance – High Variability **(C)**	C1	C2	C3
Low Performance – Low Variability **(D)**	D1	D2	D3

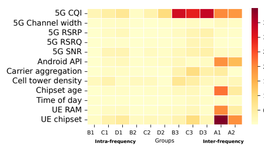
Table 1. Feature descriptions for network performance analysis.

Methodology. We begin by disaggregating the Speedtests by cellular carriers to account for differences in their deployment strategies. We then group the tests by frequency range. This allows us to isolate the effect of frequency while controlling for carrier-specific variations. In this analysis, we include only the 100k Speedtest measurements that are labelled with frequency range information. Within each frequency range, we group the tests by H3 resolution 9 hexagons. This spatial grouping serves two key purposes: it allows us to control for location-specific factors, and it aligns with the FCC's method of computing mobile broadband coverage in the U.S. [17]. To assess whether neighboring H3 hexagons exhibit similar 5G performance and thereby validate the relevance of this grouping, we leverage the Moran's I [23] statistic.[5] Our analysis reveals that all three carriers show a positive spatial autocorrelation across all eight cities for both download speed and latency. We present the average Moran's I statistic across all eight cities in Table 5 of the Appendix. This justifies our use of H3-based spatial grouping to control for location-specific factors.

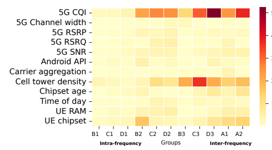
For each H3 hexagon, we compute the average and standard deviation of 5G download speed and latency to assess performance and its variability by carrier and frequency range. Using the distribution of means and standard deviations, we define the performance and variance thresholds, listed on the right side of Figure 4. This results in four categories (A-D) per frequency range (1-3) and per carrier, as outlined in Table 1. Our approach allows us to isolate the effect of frequency while examining the performance variation within each frequency range. We then compute the divergence between these groups to quantify the impact of features on 5G performance both within and between frequency ranges.

To analyze performance within each frequency range, we designate its respective "High Performance – Low Variability" group as the reference group; these

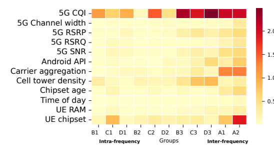
[5] Moran's I statistic has been widely used in past research [32, 41, 64] to analyze the spatial distribution of a variable of interest within a specific geographic area. It is described in more detail in the Appendix.



(a) Carrier 1

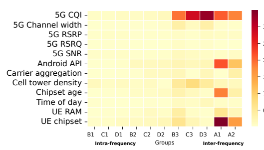


(b) Carrier 2

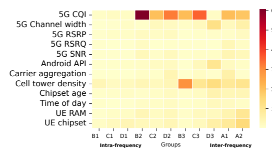


(c) Carrier 3

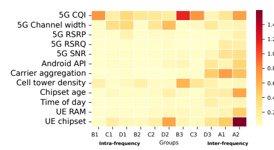
Fig. 5. KL divergence metrics - download speed.



(a) Carrier 1



(b) Carrier 2

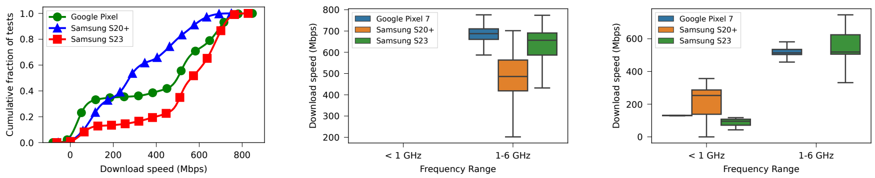


(c) Carrier 3

Fig. 6. KL divergence metrics - latency.

are groups A1, A2, A3, as labeled in the table. This allows us to identify which features diverge from optimal features within each frequency range. To analyze performance differences between frequency ranges, we also need an overall reference point. For this, we select the "High Performance – Low Variability" group in the highest frequency range (mmWave) as our overall reference group (Group A3). This enables us to identify the features that diverge when comparing optimal performance at lower frequencies to the theoretical best-case scenario of 5G technology. By using these reference groups, we can systematically analyze performance variations both within and across frequency ranges, revealing factors that influence 5G network performance. For every other group, we calculate the pairwise KL divergence for each individual feature between that group and the reference group. We include a variety of PHY layer, performance, device, time and cell tower density features; we list each feature we study in Table 6 in the Appendix. We perform Laplace smoothing [6] for any zero probability event we encounter in our reference distribution. These distance measures allow us to quantify, for each group, the dissimilarity between the distributions of individual features between itself and the reference group.

Results. We show detailed results for KL divergence from our statistical distance computation for download speed and latency in Figures 5 and 6, respectively. Our analysis reveals consistent patterns across all carriers, despite their different deployment strategies. Within each frequency range, PHY layer parameters, such as CQI (Row 1 on the y-axis), show the highest divergence, confirming that difference in signal quality leads to difference in network performance, as one would expect. As frequency increases, cell tower density becomes increasingly influential, likely due to shorter wavelengths and therefore smaller coverage areas. For inter-frequency comparisons (e.g. A1 and A2 on the x-axis), while PHY layer parameters remain important, device-related factors, particularly hardware features such as chipset model and age, and software features such as Android API,



(a) By device model (b) Carrier 1 (c) Carrier 2

Fig. 7. Download speeds by frequency and device model (each with a different chipset).

demonstrate significant divergence for all carriers. For instance, chipset model has a KL divergence of approximately 3, 1 and 2 for carriers 1, 2 and 3, respectively. This suggests that device capabilities play a crucial role in determining 5G performance across different frequency bands. Finally, we note that time of day shows low divergence, implying minimal impact on performance variability.

3.4 Experimental Validation

Our analysis in Section 3.3 yielded multiple findings about key features that affect 5G performance as measured by crowdsourced Speedtests. In this section, we confirm these findings in a controlled environment through the use of additional experiments, described in Section 2. Specifically, we analyze 3,750 5G Speedtest measurements that we collect over 15 days on nine phones. Each phone model supports the full range of frequency bands available in our measurement locations [17]. Figure 7(a) illustrates the download speeds by phone model. The Samsung S23 and Google Pixel 7 models perform well, with median download speeds reaching approximately 600 Mbps; tests on the Samsung S20+, an older phone, measure lower download speeds, with a median of 400 Mbps. On the other hand, measured latency (not shown) on all phone models is nearly identical. In the remaining analysis, we use only 5G tests from carriers 1 and 2; carrier 3, despite claims of 5G service, never connected to a 5G cell in our test locations. We leave deeper exploration of this occurrence for future work.

Measuring variability. Our goal is to ensure the findings from our crowdsourced data analysis are replicable in a controlled setup where data biases are minimized. To this end, we perform a detailed analysis of our controlled Speedtest measurements disaggregated by UE model, frequency band, and carrier. For carrier 1, measurements in the same frequency range yield similar median download speeds for the S23 and the Pixel 7 models (650 Mbps), while the S20+ shows slightly lower speeds (450 Mbps), as shown in Figure 7(b). However, for carrier 2 (Figure 7(c)), the Samsung S20+ phones consistently connect to cells operating in lower frequency bands, while the Samsung S23 and Google Pixel 7 phones connect to cells in higher frequency bands, even when running tests from the same physical location as the S20+. This suggests that while older device chipsets support higher frequency bands, newer chipsets are likely optimized to utilize these higher bands. This could explain why these newer models tend to connect to cells operating in higher frequency bands, and highlights the importance of chipset optimization for utilizing higher frequency bands effectively. We confim that our findings for latency are similar, and present them in Figure 8 of the Appendix, for completeness.

4 Related Work

Prior work has utilized speed test data to study broadband performance [44, 50, 53, 55, 56] and characterize the utility and usability of the speed tests themselves [27, 29, 54]. Other work has studied the nuances of speed test design, showing how different measurement strategies contribute to varying results [25, 33, 37, 42, 43, 51, 65]. In [42, 51], the authors highlight the importance of including metadata to improve the accuracy of performance conclusions.

In the context of mobile broadband, the authors of [26] highlight the challenges in measuring performance. Mobile access bandwidth for over 3.5 million users is analyzed in [61], highlighting the interdependence of different cellular technologies. The influence of device and PHY layer parameters on cellular performance was investigated in [30, 58]. Finally, a variety of measurement studies have focused on the identification of features that affect 5G speeds, latency, application quality of experience and power consumption, both in localized settings as well as in-the-wild through drive tests; 5G speeds are predicted using these features as well [35, 36, 38, 39, 45–49, 62, 63]. Recently, the authors of [40] conducted an in-depth analysis of 5G performance across three U.S. operators in two cities, revealing under-utilization of 5G's capabilities.

In comparison to similar studies, our research uses fine-grained, individual crowdsourced Speedtest measurements, complemented by controlled experiment data, to comprehensively assess the impact of diverse factors on cellular network performance and identify key factors that explain 5G performance variability.

5 Conclusion

Our analysis of 5G network performance variability using statistical divergence measures reveals that PHY layer parameters, particularly CQI and RSRP, show higher divergence within frequency bands, especially in mmWave frequencies. This pattern aligns with known RF signal propagation characteristics, where higher frequencies have shorter coverage ranges and are more sensitive to environmental factors. Across frequency bands, device hardware, notably chipsets, exhibit significant divergence, highlighting the need for advancements in device hardware to support the rapid evolution of network technologies. We observe that crowdsourced data with comprehensive metadata can effectively capture these nuanced performance variations, underscoring the need for speed test platforms to include such data to facilitate more accurate interpretation of mobile broadband performance. Such comprehensive data will support more informed decision-making by regulatory bodies and network operators in planning 5G infrastructure investments and also addressing potential disparities in user experience across different devices and locations.

Acknowledgements This work was funded in part by National Science Foundation (NSF) Internet Measurement Research award #2220388. We also thank Ookla for sharing their data with us.

6 Appendix – Supplemental Material

Table 2. Summary of Speedtest measurements from Ookla (City 1-8) and our own experiments. Carrier 1, 2 and 3 data are for 5G.

City	Total 4G	Total 5G	Carrier 1	Carrier 2	Carrier 3
City 1	147,443	191,274	80,464	40,994	69,816
City 2	116,738	154,498	76,960	21,468	56,070
City 3	96,904	180,672	101,727	31,737	47,208
City 4	110,766	147,512	94,496	28,451	24,565
City 5	71,530	90,759	60,765	2,108	27,886
City 6	56,321	107,078	33,932	31,701	41,445
City 7	52,233	80,362	37,390	17,552	25,420
City 8	47,555	86,572	33,950	24,037	28,585
Experiment		3,746	1,639	1,780	327
Total	699,490	1,042,473	521,323	198,828	322,322

Table 3. Annual Speedtest counts by frequency range.

Year	≤6GHz	>6GHz (mmWave)
2021	58,908	12,404
2022	76,662	11,844
2023	57,771	4,639

Table 4. Annual Speedtest counts by frequency range for the data analyzed.

Year	<1 GHz	1-6 GHz	>6 GHz (mmWave)
2021	20,125	14,054	12,020
2022	7,531	7,163	11,160
2023	6,637	10,457	4,565

Table 5. Moran's I statistic by carrier and frequency band.

Carrier	Frequency	Download		Latency	
		Moran's I	p-value	Moran's I	p-value
Carrier 1	< 1 GHz	0.326237	0.001	0.063199	0.002
	1-6 GHz	0.251701	0.001	0.111560	0.001
	> 6 GHz (mmWave)	0.398084	0.001	0.320770	0.001
Carrier 2	< 1 GHz	0.344306	0.001	0.108204	0.005
	1-6 GHz	0.352893	0.001	0.134938	0.018
	> 6 GHz (mmWave)	0.580837	0.001	0.288536	0.002
Carrier 3	< 1 GHz	0.329644	0.001	0.299311	0.001
	1-6 GHz	0.215944	0.001	0.282385	0.001
	> 6 GHz (mmWave)	0.391790	0.001	0.505062	0.001

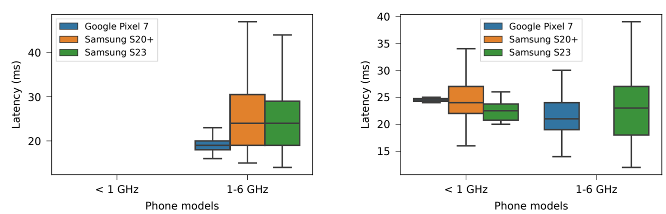
Ethics Statement. Our study does not include human subjects research. Ookla's data sharing under the DUA is fully anonymized and does not reveal full IP addresses, safeguarding the identities of individual users. Moreover, for the subset of measurements with GPS geolocation, Ookla only shares truncated coordinates, ensuring they cannot be associated with any user or residence.

Kullback-Leibler divergence. *Kullback-Leibler (KL) divergence* is a measure of how one probability distribution diverges from a second, expected probability distribution. The KL divergence can take on values in the range of $[0, \infty)$. A KL divergence of 0 indicates that the two distributions are identical. As the divergence increases, it signifies that the two distributions are increasingly dissimilar. It is important to note that KL divergence is not symmetric, and for distributions which do not have the same support, KL divergence is not bounded. The KL divergence between two discrete probability distributions P and Q in the same sample space, $\mathcal{X}$, is given by:

$$KL(P \parallel Q) = \sum_{x \in \mathcal{X}} P(x) \log \left(\frac{P(x)}{Q(x)} \right) \tag{1}$$

Table 6. Feature descriptions for network performance analysis.

Feature	Description
Download speed	Rate of data transfer from server to client in Mbps
Upload speed	Rate of data transfer from client to server in Mbps
Latency	Time delay in network communication
Frequency range	Spectrum band used for transmission
Signal strength (SS-RSRP)	Measured power of 5G reference signals
Signal quality (SS-RSRQ)	Quality of 5G reference signals
Signal strength (RSRP)	Measured power of 4G reference signals
Signal quality (RSRQ)	Quality of 4G reference signals
Signal-to-noise ratio (SS-SINR)	Ratio of signal power to noise power for 5G
Channel quality Index (CQI)	Indicator of downlink channel quality
Channel width	Width of frequency band for data transmission
Carrier aggregation	Use of multiple carriers to increase performance
Device chipset	Processor type in the mobile device
Device RAM	Memory storage of the device
Chipset age	Age of chipset
Android API version	Level of Android operating system
Hour of day	Time of measurement (0-23)
Cell tower density (resolution 9)	Cell tower density at resolution 9 H3 hexagons

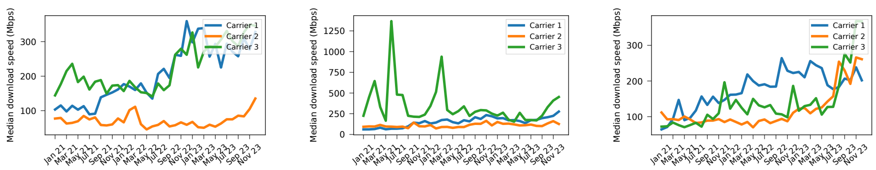


(a) Carrier 1 (b) Carrier 2

Fig. 8. Impact of frequency band and device chipset on latency.

Moran's I Statistic. Moran's I statistic is a measure of spatial autocorrelation that quantifies the degree to which similar values of a variable are clustered in a geographical area. This statistic is commonly employed to assess the spatial distribution of variables. A positive Moran's I value indicates that similar values tend to be located near one another, while a negative value implies that dissimilar values are found close together. A value of zero signifies no spatial association.

Analysis of carrier 3's performance drop. As noted in Figure 1(e), carrier 3's performance actually decreases significantly after 2021. To analyze this trend more carefully, we examine the longitudinal performance of the three carriers for illustrative cities in Figure 9. We find that performance for carrier 3 was much better in 2021 than in subsequent years in some cities (e.g. city 2), while in other cities (e.g. cities 1 and 3) it generally improves over time. This differs from the performance of carriers 1 and 2, which show either general upward trends or fairly stable performance. More analysis on carrier 3 tests shows that the majority of tests in 2021 are labelled as mmWave, while in 2022 and 2023 combined, only 1.3k of 71k tests are labelled mmWave. We hypothesize that this could be because of major mmWave deployments in 2021, followed by increased sub 6 GHz 5G deployments in subsequent years by carrier 3 that offer a better



(a) City 1 (b) City 2 (c) City 3

Fig. 9. Longitudinal performance of all three carriers in selected representative cities.

coverage radius than mmWave; however, with the available data we are unable to validate our hypothesis.

References

1. How Fast Is Your Internet? (2016), https://tinyurl.com/2yrn99z4
2. Qualcomm Linux Modems by Quectel Co. https://tinyurl.com/dtywzznh (2019)
3. The Importance Of An Internet Speed Test (2019), https://tinyurl.com/2wxaxrvc
4. U.S. Policy on 5G Technology. https://tinyurl.com/4vbszdzu (2019)
5. When was 5G introduced? https://tinyurl.com/yck2yd58 (2019)
6. Laplace smoothing. https://tinyurl.com/nhc3p4d3 (2020)
7. NetMonster: Android network monitoring app. https://netmonster.app (2020)
8. What are good download and upload speeds for home internet? https://tinyurl.com/2w96rjew (2020)
9. MediaTek Launches First mmWave Chipset for Seamless 5G Smartphone Connectivity. https://tinyurl.com/52d37k53 (2022)
10. Speedtest is the world's #1 internet utility. https://www.ookla.com/ (2022)
11. The Speedtest Server Network^{TM} (2022), https://www.ookla.com/network
12. Use your device to its greatest potential. https://tinyurl.com/a5sn4hjb (2022)
13. When did 5G come out? https://tinyurl.com/bdhurmpw (2022)
14. 5G network coverage outlook. https://tinyurl.com/vzujbku9 (2023)
15. CellMapper. https://www.cellmapper.net/ (2023)
16. How does Speedtest select a server? https://tinyurl.com/5n7cj6dc (2023)
17. National Broadband Map. https://broadbandmap.fcc.gov/home (2023)
18. Ookla's Speedtest.net (2023), https://www.speedtest.net
19. T-Mobile US builds a commanding lead over rivals ATT and Verizon. https://tinyurl.com/5n8z5397 (2023)
20. Tower Maps. https://www.towermaps.com/ (2023)
21. XCAL: World's Leading Drive Test Tool. https://tinyurl.com/nkh8mu5c (2023)
22. H3 Spatial Index (2024), https://h3geo.org/
23. Moran's I (2024), https://www.statisticshowto.com/morans-i/
24. Tables of Cell Statistics Across Resolutions (2024), https://tinyurl.com/448w9kzf
25. Bauer, S., Clark, D.D., Lehr, W.: Understanding Broadband Speed Measurements. In: *TPRC* (2010)
26. Bauer, S., Lehr, W.: Measuring Mobile Broadband Performance. *TPRC* (2018)
27. Canadi, I., Barford, P., Sommers, J.: Revisiting Broadband Performance. In: *IMC* (2012)
28. Chung, J., Kannappan, P., Ng, C., Sahoo, P.: Measures of distance between probability distributions. Journal of Mathematical Analysis and Applications (1989)

29. Clark, D.D., Wedeman, S.: Measurement, Meaning and Purpose: Exploring the M-Lab NDT Dataset. In: *TPRC* (2021)
30. Dasari, M., Vargas, S., Bhattacharya, A., Balasubramanian, A., Das, S.R., Ferdman, M.: Impact of Device Performance on Mobile Internet QoE. In: *IMC* (2018)
31. Deasy, J., Simidjievski, N., Lió, P.: Constraining Variational Inference with Geometric Jensen-Shannon Divergence. NeurIPS (2020)
32. Fan, A., Lei, F., Mancenido, M., Maceachren, A.M., Maciejewski, R.: Understanding Reader Takeaways in Thematic Maps Under Varying Text, Detail, and Spatial Autocorrelation. In: ACM CHI (2024)
33. Feamster, N., Livingood, J.: Measuring Internet Speed: Current Challenges and Future Recommendations. *Communications of the ACM* (2020)
34. Feng, L., Wang, H., Jin, B., Li, H., Xue, M., Wang, L.: Learning a Distance Metric by Balancing KL-Divergence for Imbalanced Datasets. IEEE Transactions on Systems, Man, and Cybernetics: Systems (2019)
35. Fezeu, R.A., Ramadan, E., Ye, W., Minneci, B., Xie, J., Narayanan, A., Hassan, A., Qian, F., Zhang, Z.L., Chandrashekar, J., et al.: An In-Depth measurement analysis of 5G mmwave PHY latency and its impact on End-to-End delay. In: *PAM* (2023)
36. Ghoshal, M., Khan, I., Kong, Z.J., Dinh, P., Meng, J., Hu, Y.C., Koutsonikolas, D.: Performance of Cellular Networks on the Wheels. In: IMC (2023)
37. Han, J., Lyu, M., Sivaraman, V.: On the Validity of Internet Speed Test Tools and Broadband Measurement Programs. In: AINTEC (2022)
38. Hassan, A., Narayanan, A., Zhang, A., Ye, W., Zhu, R., Jin, S., Carpenter, J., Mao, Z.M., Qian, F., Zhang, Z.L.: Vivisecting Mobility Management in 5G Cellular Networks. In: *SIGCOMM* (2022)
39. K. Fezeu, R.A., Fiandrino, C., Ramadan, E., Carpenter, J., de Freitas, L.C., Bilal, F., Ye, W., Widmer, J., Qian, F., Zhang, Z.L.: Unveiling the 5G Mid-Band Landscape: From Network Deployment to Performance and Application QoE. In: *SIGCOMM* (2024)
40. Liu, Y., Peng, C.: A Close Look at 5G in the Wild: Unrealized Potentials and Implications. In: IEEE INFOCOM (2023)
41. Locke, D.H., Ossola, A., Minor, E., Lin, B.B.: Spatial contagion structures urban vegetation from parcel to landscape. People and Nature (2021)
42. MacMillan, K., Mangla, T., Richardson, M., Feamster, N.: Best Practices for Collecting Speed Test Data. In: *TPRC* (2022)
43. MacMillan, K., Mangla, T., Saxon, J., Marwell, N.P., Feamster, N.: A Comparative Analysis of Ookla Speedtest and Measurement Labs Network Diagnostic Test (NDT7). In: *ACM SIGMETRICS* (2023)
44. Mangla, T., Paul, U., Gupta, A., Marwell, N.P., Feamster, N.: Internet Inequity in Chicago: Adoption, Affordability, and Availability. In: *TPRC* (2022)
45. Minovski, D., Ögren, N., Mitra, K., Åhlund, C.: Throughput Prediction Using Machine Learning in LTE and 5G Networks. *IEEE Transactions on Mobile Computing* (2023)
46. Narayanan, A., Ramadan, E., Carpenter, J., Liu, Q., Liu, Y., Qian, F., Zhang, Z.L.: A First Look at Commercial 5G Performance on Smartphones. In: *WWW* (2020)
47. Narayanan, A., Ramadan, E., Mehta, R., Hu, X., Liu, Q., Fezeu, R.A.K., Dayalan, U.K., Verma, S., Ji, P., Li, T., Qian, F., Zhang, Z.L.: Lumos5G: Mapping and Predicting Commercial MmWave 5G Throughput. In: *IMC* (2020)

48. Narayanan, A., Rochman, M.I., Hassan, A., Firmansyah, B.S., Sathya, V., Ghosh, M., Qian, F., Zhang, Z.L.: A Comparative Measurement Study of Commercial 5G mmWave Deployments. In: *IEEE INFOCOM* (2022)
49. Narayanan, A., Zhang, X., Zhu, R., Hassan, A., Jin, S., Zhu, X., Zhang, X., Rybkin, D., Yang, Z., Mao, Z.M., Qian, F., Zhang, Z.L.: A Variegated Look at 5G in the Wild: Performance, Power, and QoE Implications. In: *SIGCOMM* (2021)
50. Paul, U., Liu, J., Adarsh, V., Gu, M., Gupta, A., Belding, E.: Characterizing Performance Inequity Across U.S. Ookla Speedtest Users. https://arxiv.org/abs/2110.12038 (2021)
51. Paul, U., Liu, J., Gu, M., Gupta, A., Belding, E.: The Importance of Contextualization of Crowdsourced Active Speed Test Measurements. In: *IMC* (2022)
52. Qualcomm: Everything you need to know about 5G. https://www.qualcomm.com/invention/5g/what-is-5g
53. Sanchez-Arias, R., Jaimes, L.G., Taj, S., Habib, M.S.: Understanding the State of Broadband Connectivity: An Analysis of Speedtests and Emerging Technologies. IEEE Access (2023)
54. Sharma, T., Schmitt, P., Bronzino, F., Feamster, N., Marwell, N.: From point data to geographic boundaries: Regionalizing crowdsourced latency measurements (2024), https://arxiv.org/abs/2405.11138
55. Sommers, J., Barford, P.: Cell vs. WiFi: On the Performance of Metro Area Mobile Connections. In: *IMC* (2012)
56. Srinivasavaradhan, V., Park, O., Belding, E.: Mapping Cellular Network Evolution and Infrastructure Criticality: A Nationwide Analysis. In: *TPRC* (2024)
57. Stoecklin, M.: Anomaly detection by finding feature distribution outliers. In: ACM CoNEXT (2006)
58. Sui, K., Zhou, M., Liu, D., Ma, M., Pei, D., Zhao, Y., Li, Z., Moscibroda, T.: Characterizing and Improving WiFi Latency in Large-Scale Operational Networks. In: *MobiSys* (2016)
59. Sutter, T., Daunhawer, I., Vogt, J.: Multimodal Generative Learning Utilizing Jensen-Shannon-Divergence. NeurIPS (2020)
60. Xu, D., Zhou, A., Zhang, X., Wang, G., Liu, X., An, C., Shi, Y., Liu, L., Ma, H.: Understanding Operational 5G: A First Measurement Study on Its Coverage, Performance and Energy Consumption. In: *SIGCOMM* (2020)
61. Yang, X., Lin, H., Li, Z., Qian, F., Li, X., He, Z., Wu, X., Wang, X., Liu, Y., Liao, Z., Hu, D., Xu, T.: Mobile Access Bandwidth in Practice: Measurement, Analysis, and Implications. In: *SIGCOMM* (2022)
62. Ye, W., Hu, X., Sleder, S., Zhang, A., Dayalan, U.K., Hassan, A., Fezeu, R.A.K., Jajoo, A., Lee, M., Ramadan, E., Qian, F., Zhang, Z.L.: Dissecting Carrier Aggregation in 5G Networks: Measurement, QoE Implications and Prediction. In: *SIGCOMM* (2024)
63. Yuan, X., Wu, M., Wang, Z., Zhu, Y., Ma, M., Guo, J., Zhang, Z.L., Zhu, W.: Understanding 5G Performance for Real-World Services: A Content Provider's Perspective. In: *SIGCOMM* (2022)
64. Zahnd, W., Bell, N., Larson, A.: Geographic, racial/ethnic, and socioeconomic inequities in broadband access. The Journal of Rural Health (2021)
65. Zhang, Z., Shen, J., Mok, R.K.P.: Empirical Characterization of Ookla's Speed Test Platform: Analyzing Server Deployment, Policy Impact, and User Coverage. In: CCWC (2024)