Spatio-Temporal and Context-Aware Dataset for Mobile Network Forecasting

Mathias Gerstner, Sebastian Fischer, Rudolf Hackenberg
Dept. Informatics and Mathematics, OTH Regensburg
Regensburg, Germany
e-mail: {mathias.gerstner, | sebastian.fischer, | rudolf.hackenberg}@oth-regensburg.de

***Abstract*—As autonomous driving becomes increasingly feasible, the German government has introduced a legal framework to enable the operation of vehicles with Level 4 automated driving functionality. A key requirement is the maintenance of a continuous connection between such vehicles and a remote technical supervisor. If this link is lost, the vehicle must transition into a safe state by bringing itself to a controlled stop. To mitigate the risk of connection loss, accurate forecasting of mobile network availability along routes is essential. This paper describes a spatio-temporal analysis of mobile network signal quality metrics along a fixed rural route based on 48 repeated measurement drives. The route was segmented into 320 spatial reference sections to enable consistent cross-trip comparison and stability assessment. A classification into usable and non-usable connectivity states reveals that more than half of the sections show pronounced trip-to-trip variability. The results show that mobile network quality is not only dependent on geographic location but is also influenced by environmental conditions. While global correlations between weather parameters and raw signal metrics remain weak, moderate relationships can be seen when analyzing the share of functionally usable connectivity in spatially unstable sections. A walk-forward forecasting model demonstrates that the inclusion of temperature in the model reduces prediction error compared to a purely historical baseline in 73.7% of evaluated trips. The findings show the importance of feature selection and also the limitations of linear modeling approaches. Rather than acting as deterministic predictors of signal strength, contextual parameters primarily modulate connectivity uncertainty in spatially unstable regions. These findings underscore the importance of contextual information and localized modeling to predict network availability for safety-critical systems, such as autonomous vehicles.**

***Keywords*-mobile network stability; connectivity forecasting; vehicular communication; autonomous driving; signal quality indicators.**

I. INTRODUCTION

This extended study builds upon a previous contribution [1], presented at The Second International Conference on Vehicular Systems (VEHICULAR ANALYTICS 2025). While the previous study focuses on the context-aware prediction of mobile networks, this study extends the data analysis and presents one additional forecasting model approach. It also investigates the spatial stability and trip-to-trip variability of mobile network quality along the selected rural route.

Even though recent forecasts are more cautious regarding the future proliferation of automated driving vehicles, it is evident that they will gradually become part of everyday life as technology evolves [2]. In this context, emerging automated driving functions and Vehicle-to-Everything (V2X) communication systems highlight the growing importance of robust and reliable vehicular connectivity [3].

Automated driving functions provided by future vehicles at Society of Automotive Engineers (SAE) Level 4 depend on reliable mobile network connectivity. While the core driving decisions must be performed locally inside the vehicle, the mobile network is considered a requirement to enable remote supervision and instructions in case of uncertainties. Therefore, it is important to assess the connectivity along a route before such an automated vehicle drives into an area without connection. While static cellular coverage maps provide a coarse indication of network availability, they fail to capture variations in signal quality that may occur at many spatial locations [4]. Although mobile networks are generally planned for wide geographical coverage, this must be distinguished from functional connectivity. A geographically covered location may still have insufficient signal quality under real driving conditions. Therefore, this study focuses on the stability and functional usability of connectivity rather than on theoretical coverage stated by mobile network operators. Variations in mobile network quality are particularly critical for safety-relevant automated driving applications, where transient connectivity loss or severe degradation may trigger fallback behaviors, such as bringing the vehicle to a safe stop [5].

Network connection quality is influenced by various dynamic factors, including the environment and time of day [6][7]. Open areas near base stations usually offer strong coverage, while remote or forested areas do not. To investigate this variability, 48 test drives along a predefined 64 km rural route, recording passive mobile network parameters and Global Navigation Satellite System (GNSS) data, were conducted. During post-processing, additional context, such as weather conditions and metadata, was integrated.

From a system-level perspective, such pre-trip connectivity information can be integrated into the route planning and data transmission schedules of automated vehicles. By identifying route segments with unstable or unreliable connectivity in advance, vehicles may adapt operational strategies, adjust communication-intensive data transmission, or initiate alternative routes.

The main extensions of this contribution are:

- A data cleaning and preprocessing methodology tailored to vehicular mobile network measurements.
- An approach for the functional classification of mobile network quality into usable and non-usable connectivity states.
- A spatial stability analysis revealing route segments with persistent connectivity, persistent coverage holes, and pronounced trip-to-trip variability.

- An investigation of contextual influences on connectivity variability.

This study is based on measurements collected along a fixed route over the year 2025. The data is restricted to linear correlations between contextual and network parameters. Urban environments and larger datasets, beyond the 48 measurement drives used here, remain for future work and are expected to further improve prediction performance.

A. Motivation

The key motivation for this work comes from recent regulatory developments in Germany. According to legislation, automated vehicles classified under SAE Level 4 (high driving automation [8]) must maintain a permanent connection to a technical supervisor (§ 1e(2) no. 10 StVG [5]). This supervisor must receive real-time status information and can intervene, for example, by enforcing a safety stop or assisting in complex scenarios [5]. One example of such a scenario is near construction zones, where the vehicle may need to cross continuous lane markings, an action typically prohibited by internal rules. If the connection to the technical supervisor is lost, the vehicle has to bring itself to a controlled stop. Driving under Level 4 automation must not resume until a network connection is reestablished. According to the law, the technical supervisor has to be a natural person with specific knowledge about vehicular systems (§ 1d(3) no. 3 StVG [5]. The qualification of the supervisor must be demonstrated by a diploma, bachelor's, master's, or technical degree in mechanical engineering, automotive engineering, electrical engineering, or a related field (§ 14(1) no. 1 AFGBV [9].

Another motivation is the observation that static coverage representations are insufficient for characterizing mobile network availability in safety-critical automated driving scenarios. Even if average signal metrics indicate acceptable coverage at a given location, individual trips may experience severe degradation or complete loss of connectivity (shown in Section VI). Such trip-to-trip variability cannot be captured by static maps or historical averages, yet it directly impacts the operational reliability of Level 4 automated vehicles.

Unlike prior studies that focused on static coverage maps, short-term prediction, or urban areas, this work presents a long-term measurement campaign on a rural route and also integrates contextual factors (e.g., time of day, ambient temperature). To our knowledge, this is among the first studies to target pre-trip forecasting of Long-Term-Evolution (LTE)/ 5G quality for Level 4 driving scenarios.

B. Structure

The remainder of this paper is organized as follows. Section II discusses the state of the art and related work. Section III introduces the fundamentals of mobile networks and describes the target parameters. Section IV outlines the data collection followed by the data cleaning and preprocessing methodology in Section V. Next, Section VI presents a spatial stability and trip-to-trip variability analysis of mobile network connectivity. Section VII investigates the influence of contextual factors. Section VIII introduces an initial connectivity forecasting model. Finally, Sections IX and X discuss the results and limitations of the study. In the end, Section XI concludes the paper and outlines directions for future work.

II. Related Work

Several studies have addressed the prediction and analysis of mobile network quality, to some extent, in vehicular contexts. These efforts vary in scope, methodology, and focus, often targeting real-time diagnostics, network-side optimization, or static coverage mapping rather than pre-trip forecasting.

Torres et al. [10] propose a data-driven approach to forecast congestion in LTE networks using analytics and Machine Learning (ML) techniques. Their model is designed to support Self-Organizing Network (SON) optimization and real-time traffic management from a network operator perspective. However, it does not incorporate contextual factors such as weather or time of day and is not intended for vehicle-centric or route-specific pre-trip predictions.

Madariaga et al. [6] adopt a time-series perspective and demonstrate that meteorological conditions significantly affect mobile Quality of Service (QoS). By combining ML models with temperature, humidity, and rainfall data derived from crowdsourced Android measurements, they show that context-aware approaches outperform purely historical averages, particularly in dense urban environments. In a related study, they address the spatial aggregation of signal strength measurements using advanced interpolation techniques [11]. While these works provide important insights into context-aware modeling and spatial estimation, they are restricted to stationary urban settings and do not investigate long-term variability or repeated measurements along fixed routes.

Sahin and Sathya [12] present a lightweight classification approach that estimates mobile network quality in a given area using only Global Positioning System (GPS) data. Although computationally efficient, their method omits dynamic environmental influences such as weather or time and does not explicitly address predictive uncertainty. Similarly, Rahman et al. [13] employ reinforcement learning in combination with Unmanned Aerial Vehicles (UAVs) to detect mobile coverage holes in urban environments. Their work focuses on infrastructure-centric optimization and coverage assessment rather than vehicle-specific planning or pre-trip prediction.

Other studies have also investigated connectivity-aware navigation and planning. Hultman et al. [14] introduce a route planning approach based on static mobile signal coverage maps derived from publicly available data. While such methods are useful for coarse connectivity-aware routing, they lack the temporal and contextual sensitivity required to capture the trip-to-trip variability experienced during real-world driving.

More recently, large-scale measurement campaigns and open-access datasets have become available. Schippers et al. [7] introduced the DoNext dataset, a comprehensive multi-region 5G measurement campaign designed to support data-driven mobile network research. The dataset provides extensive spatial and temporal coverage and enables the analysis of general network behavior across different contexts. However, its primary focus lies on general-purpose analysis rather than route-specific forecasting or the investigation of connectivity stability along repeated vehicular trajectories. Their results nevertheless confirm that signal quality at a given location can vary significantly over time, for example, due to congestion effects during peak hours.

Beyond dataset-centric contributions, recent work has emphasized methodological challenges in predictive mobile networking. Blasco et al. [4] discuss the concept of predictive QoS and highlight fundamental issues such as non-stationarity, feature selection, and evaluation bias in cellular networks. Similarly, Palaios et al. [15] provide an extensive analysis of machine learning-based QoS prediction for vehicular communication, focusing on robustness, data splitting strategies, and explainability. While these studies offer valuable guidance for model design and evaluation, they primarily treat variability as a modeling challenge rather than as an inherent spatial and temporal property of the environment.

Overall, prior work has addressed many important aspects of mobile network analysis, including data aggregation, spatial modeling, and urban QoS prediction. However, most existing approaches implicitly assume that mobile network conditions at a given location can be sufficiently described by average or aggregated metrics. Such assumptions are insufficient for autonomous driving applications, where transient connectivity loss or severe degradation can have immediate operational consequences.

The present study focuses on real-world measurements along a rural route and explicitly investigates the stability of mobile network connectivity across multiple trips. Rather than treating variability solely as noise or uncertainty to be mitigated by predictive models, it is additionally analyzed as a characteristic property of specific route segments. Furthermore, a functional distinction between usable and non-usable connectivity states is introduced that directly reflects the operational requirements of Level 4 autonomous driving and supports pre-trip connectivity assessment under uncertainty.

III. Background and Fundamentals

This section outlines the technical background necessary to understand the mobile network measurements analyzed in this study. In addition to a general overview of LTE and 5G principles, it explains the radio signal metrics that are particularly relevant for vehicular communication. The relationships between these Key Performance Indicators (KPIs) and their impact on the functional connectivity of autonomous driving systems are also described. Furthermore, the physical propagation effects caused by atmospheric conditions are briefly discussed to motivate the investigation of contextual influences in later sections.

A. Principles of Mobile Networks

Mobile communication networks consist of wireless links between User Equipments (UEs) and base stations, which connect to a core network. The transition from LTE to 5G enhances bandwidth, latency, and reliability [16], which are critical for autonomous driving applications [17]. LTE provides high data rates and reduced latency through techniques such as Multiple Input Multiple Output (MIMO), handovers, and advanced radio resource management [18]. 5G builds on LTE by adding capabilities such as beamforming and network slicing. It also specifies dedicated QoS mechanisms, including profiles for specific service classes.

5G can be deployed in Non-Standalone (NSA) or Standalone (SA) architectures. In NSA, 5G is operated with existing LTE infrastructure, whereas in SA deployments use a dedicated 5G core [19]. While German mobile network operators are introducing an increasing amount of 5G SA services, the public network still involves LTE and 5G NSA, depending on location and operator. As a result, robustness against environmental influences and coverage gaps is still strongly tied to LTE radio characteristics, particularly in rural areas where dense 5G deployments are uncommon.

5G operates across low-, mid-, and high-frequency bands, including millimeter waves [20]. Research has also investigated the fundamental performance limits of millimeter-wave systems for cooperative localization, underlining the relevance of advanced 5G [21]. Moreover, 5G offers even lower latency and increased capacity supporting safety-critical automotive applications [17]. In both LTE and 5G networks, signal quality plays a central role in determining the performance and reliability of a connection [22].

In the context of remote supervision for automated driving applications, the data volume of status messages like the Cooperative Awareness Message (CAM) is generally low (a size of 200 to 800 bytes and a frequency between 1 Hertz (Hz) and 10 Hz [23]). For this use case, connection robustness is more important than peak data throughput. This motivates a closer examination of signal stability rather than instantaneous performance metrics.

B. Mobile Network Signal Quality Metrics

Radio signal propagation in vehicular environments is influenced by various physical and atmospheric factors. Temperature and humidity affect air density, while fog, rain, and precipitation can introduce additional attenuation and scattering effects, particularly at higher carrier frequencies [24]. Although such effects are often secondary compared to geometric factors and network load, they can contribute to signal variability and are therefore considered as contextual parameters in this study.

Signal quality in LTE and 5G networks is assessed using a set of key radio frequency indicators. These metrics help to evaluate the reliability and performance of the wireless link between the UE and a cellular base station [25]. The most commonly used parameters are as follows:

- Received Signal Strength Indicator (RSSI) represents the total received power observed by the device within the channel bandwidth [22]. This value includes not only the desired signal but also background noise and interference from neighboring cells. Measured in decibel-milliwatt (dBm), RSSI serves as an indicator of signal presence but lacks specificity for signal quality [26][27].

- Reference Signal Received Power (RSRP) quantifies the average power of reference signals transmitted by a serving cell [22]. Unlike RSSI, it excludes interference and noise, offering a more accurate measure of usable signal strength. RSRP is expressed in dBm and is used by UEs for tasks such as cell selection, handover decisions, and radio link monitoring [26].
- Reference Signal Received Quality (RSRQ) is the ratio of RSRP to RSSI, normalized by the number of resource blocks [22]. It reflects both the signal strength and the level of interference within the channel. It is particularly useful in scenarios with heavy network load or high interference levels, where a strong signal may still result in poor quality due to congestion. RSRQ is commonly measured in decibel (dB) [26].
- Signal-to-Interference-plus-Noise Ratio (SINR) indicates the quality of the wireless communication channel by comparing the power of the desired signal to the sum of interference and background noise [22]. It directly impacts achievable data rates and overall link performance. SINR is also measured in dB and is a parameter for evaluating network efficiency and optimizing system performance [26].

TABLE I. SIGNAL QUALITY METRICS FOR LTE AND 5G NETWORKS [22]

Parameter	Unit	Range	Quality Interpretation
RSSI	dBm	-120 to -30	> -65: Excellent -65 to -75: Good -75 to -85: Fair < -85: Poor
RSRP	dBm	-140 to -60	> -80: Excellent -80 to -90: Good -90 to -100: Fair < -100: Poor
RSRQ	dB	-20 to -3	> -10: Excellent -10 to -15: Good -15 to -20: Fair < -20: Poor
SINR	dB	-20 to +30	> 15: Excellent 10 to 15: Good 5 to 10: Fair < 5: Poor

Table I summarizes the value ranges for each of the four metrics based on industry standards and field experiments, including recommendations provided by the applied router's manufacturer [22].

In general, higher values of these metrics correspond to better network quality and more stable connectivity. However, for autonomous driving applications, the relevance of these metrics lies not only in their absolute values but also in their temporal stability and their ability to indicate whether a connection is functionally usable.

IV. DATASET

This section describes the measurement campaign, the data acquisition setup, and the resulting dataset used for the subsequent analysis.

A. Measurement Campaign

The dataset used in this study was collected during a series of 48 real-world measurement drives along a 64 km segment of the road B16 in southern Germany. The route was selected to cover different environmental conditions. It includes suburban areas near the cities of Ingolstadt and Regensburg, smaller villages, extended rural sections, and forested segments. In addition, the route contains various topological properties such as small hills, bridges, and roadside structures, which are known to influence mobile radio propagation [28]. This diversity was expected to induce a broad variance in mobile network conditions and connectivity behavior, making the route suitable for studying spatial variability and stability of network quality. The measurement campaign extended over 10 months (March to December 2025) to capture a broad range of temporal, seasonal, and environmental variability. In total, 92,210 measurement points along the route were collected.

For the data acquisition, a custom-built mobile measurement unit was used. An overview of the setup, including its hardware components and their interconnections, is illustrated in Figure 1.

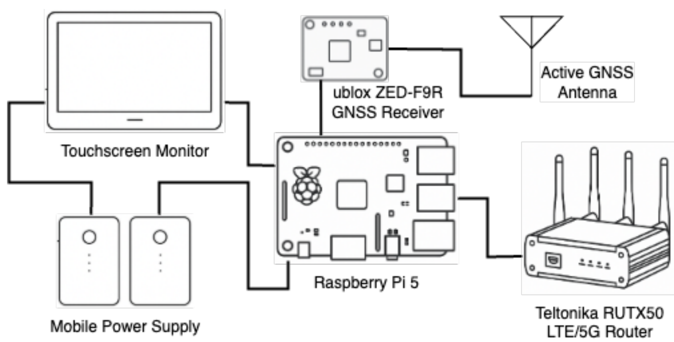

Figure 1. Setup of the custom-built mobile measurement unit

In detail, the setup consists of the following components:

- Raspberry Pi: As the central processing unit, a Raspberry Pi 5 with 8 GB of Random-Access Memory (RAM) is used. It is running Raspberry Pi OS in the Graphical User Interface (GUI) version. For data processing, Python 3.12 is used on this device.
- Router: For the connection to the mobile network, a Teltonika RUTX50 LTE/5G router is used. It is connected to the Raspberry Pi via a Local Area Network (LAN) cable. The router is capable of measuring the passive mobile network parameters needed for the study. Those parameters are transferred to the Raspberry Pi over Simple Network Management Protocol (SNMP).

The router does not use dedicated 5G QoS profiles for this measurement campaign.

- GNSS receiver: A u-blox ZED-F9R GNSS receiver is included in the setup for determining a precise geographic location while driving along the route. The data is transferred to the Raspberry Pi via a serial connection using a Universal Serial Bus (USB) port. The receiver uses the National Marine Electronics Association (NMEA) standard commonly used in navigation devices. These messages are in a human-readable format and contain not only the location but also information about time, date, current speed, heading direction, visible satellites, and precision of location determination like Position Dilution of Precision (PDOP).
- GNSS antenna: The active antenna u-blox ANN-MB-00-00 is connected to the receiver. During measurement drives, it is placed in a fixed position under the vehicle's windshield.
- Monitor: A 7-inch display from Waveshare is built into the setup. It shows the current state of the GNSS receiver and the LTE/5G router, as well as the currently measured values. With its capacitive touchscreen, the measurements can be started and stopped via two buttons. The screen is connected to the Raspberry Pi via High-Definition Multimedia Interface (HDMI).
- Power supply: Two rechargeable mobile batteries are included in the setup to provide the power needed. One battery powers the Raspberry Pi while the other powers the display. Both batteries provide 3 Ampere (A) at 5 Volt (V). The router is powered by the vehicle's electrical system to ensure stable power availability for this critical device.

The collected mobile network parameters, as presented in Section III-B were logged continuously throughout each drive. In addition, the GNSS receiver provided geographical coordinates and motion-related parameters, enabling precise spatial alignment of network measurements along the route.

B. Dataset Description

The initial dataset comprises time-stamped mobile network measurements, geographic location information, and motion-related parameters. Recorded network metrics include RSRP, RSRQ, RSSI, and SINR, as well as cell identifiers and connection state information. Spatial context is provided by latitude, longitude, altitude, and GNSS-derived quality indicators such as PDOP, Horizontal Dilution of Precision (HDOP), and number of visible satellites. Vehicle speed and heading were recorded to capture the motion context of each measurement.

In total, 92,210 measurement points were collected with a temporal resolution of approximately 2–3 seconds. Not all measurements are located on the target route, as recordings also include short segments at the beginning and end of each drive. These off-route measurements are excluded during preprocessing, as described in Section V.

C. Relationship Between Signal Metrics and Data Rate

To empirically motivate the relevance of the selected passive mobile network parameters, we analyze an independent open-access dataset collected in the Salzburg region, which provides both passive radio signal metrics and measured data rates [29].

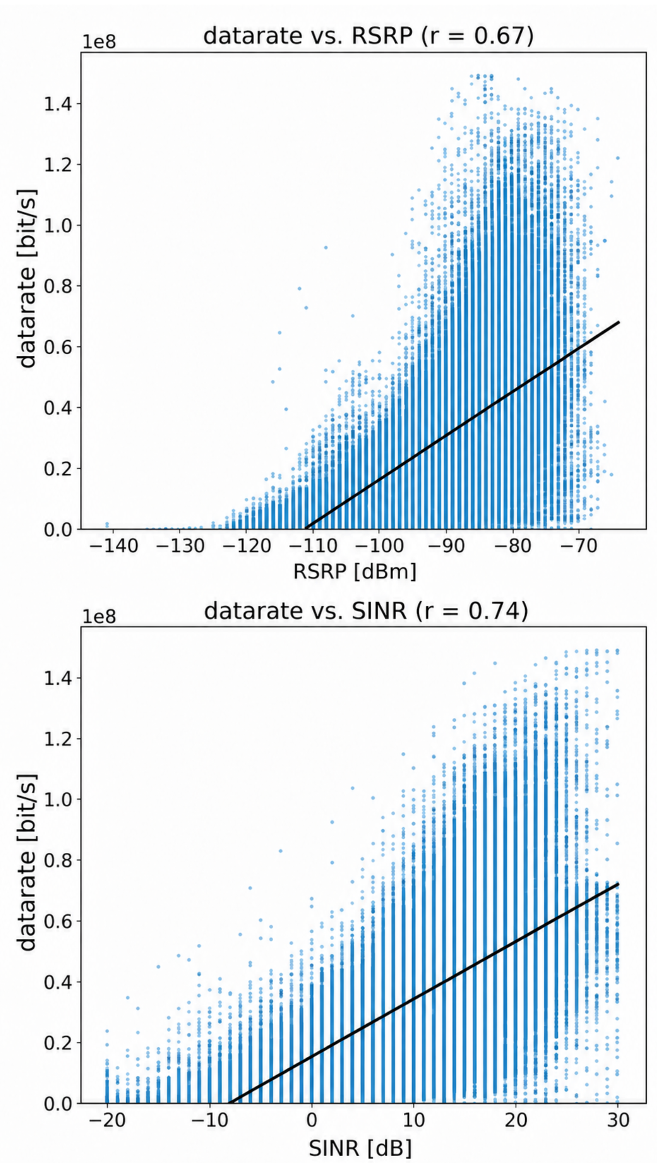

Figure 2. Relationship between measured data rate and passive mobile network parameters based on an open-access dataset [29]. Linear correlation coefficients are indicated for each parameter.

Figure 2 illustrates the relationship between data rate and RSRP and SINR. The corresponding Pearson correlation coefficients are $R = 0.67$ for RSRP and $R = 0.74$ for SINR, indicating a positive linear association between signal quality and achievable data rate. The correlation R-value of the two other metrics is $R = 0.58$ for RSRQ and $R = 0.58$ for RSSI as well. The Pearson correlation coefficient (R-value) quantifies the strength and direction of a linear relationship between two variables on a scale from -1 (perfect negative correlation) to

$+1$ (perfect positive correlation), with values closer to zero indicating weaker linear dependence.

Across all parameters, a clear positive correlation trend can be observed, with SINR exhibiting the strongest linear correlation. These results are consistent with prior studies on QoS prediction in cellular networks and confirm that passive signal metrics provide meaningful information about the achievable data rate.

V. Data Cleaning and Preprocessing

While the dataset provides rich spatial and temporal information, the raw measurements include missing values, default placeholders, and off-route samples. The following section describes the data cleaning and preprocessing pipeline applied prior to analysis. All steps were implemented in Python using the libraries pandas and numpy and are designed to preserve variability while removing invalid measurements.

The first preprocessing step was the validation of the GNSS data. A total of 339 measurements (0.37%) contained invalid location information, either due to missing coordinates or fallback values with latitude or longitude equal to zero. Notably, 215 of the zero-valued GNSS samples originated from a single measurement run, where the GNSS receiver did not yet obtain a valid fix at the beginning of the drive. As these samples occurred in extended consecutive blocks, no spatial interpolation was applied. Instead, all measurements without valid GNSS coordinates were excluded from spatial analysis. This issue could be avoided in future campaigns by explicitly verifying a valid GNSS fix prior to initiating measurements.

The next cleaning step focused on the identification of placeholders used by the router when no network connectivity is available at all. Radio quality indicators were first analyzed without applying any filtering or correction. By comparing the observed minimum, maximum, and average values of RSSI, RSRP, RSRQ, and SINR with the plausible ranges defined in Table I, fallback values were identified. Specifically, the values -32768 (corresponding to the minimum signed 16-bit integer), -255, and 0 were found. Based on this observation, explicit placeholder values (-32768 and -255) were consistently replaced by Not a Number (NaN) for all radio quality indicators. This served as the basis for all subsequent cleaning steps. An additional column was included in the dataset that marks the locations at which placeholders were replaced.

Handling zero-valued SINR measurements required special attention. Because it lies within the normal range of SINR, a value of 0 can also represent a valid signal condition. But it was also used by the device as a fallback value when no SINR calculation was possible. To distinguish between valid and invalid zero-values, the occurrence of SINR values equal to zero and anomalies in other radio KPIs were analyzed. Many SINR samples with a value of $0\,\mathrm{dB}$ were found to coincide with missing or implausible values in the other radio quality indicators. These samples were therefore classified as invalid values and excluded from the cleaned SINR distribution. Figures 3 and 4 illustrate the SINR value distributions before and after removing those invalid fallback values. The comparison shows the impact of poorly chosen device fallbacks on the observed SINR statistics.

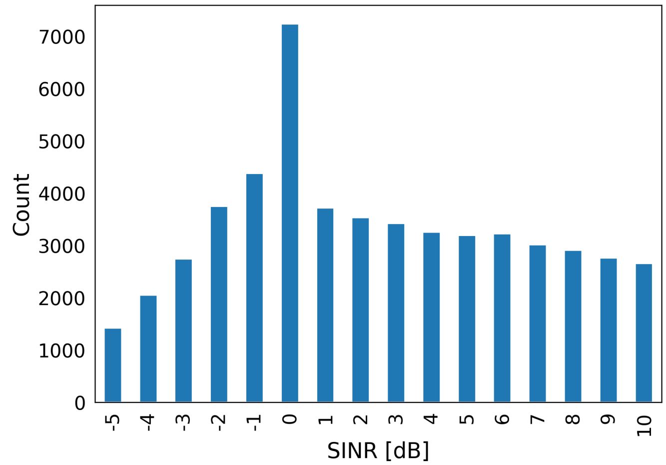

Figure 3. Frequency distribution of SINR values after removing explicit sentinel values. Zero-valued SINR samples include both physical measurements and implicit device fallback values.

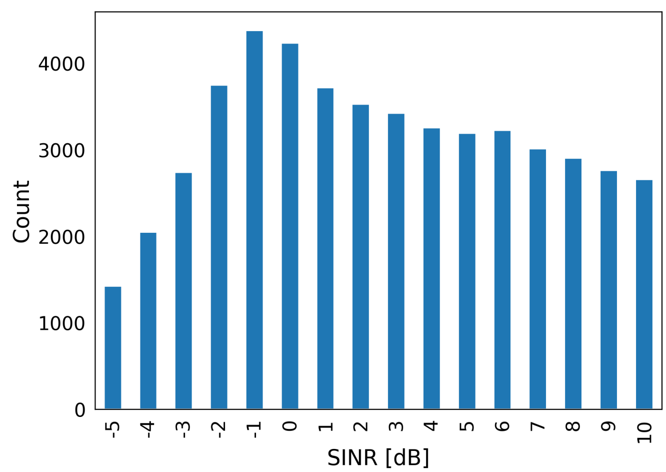

Figure 4. Frequency distribution of SINR values after removing both explicit sentinel values and implicit zero-valued fallback measurements.

After removing the placeholders, all mobile network KPIs follow physically plausible ranges and statistically consistent distributions. Missing radio measurements were not treated as noise but explicitly encoded as a separate indicator variable. This allows the prediction models to distinguish between absent connectivity and weak but valid signal conditions. There are no indications of further systematic invalid values. Figure 5 shows an example of the distribution of the frequency of RSRP values. The other KPIs follow a similar pattern.

Table II summarizes basic descriptive statistics of the cleaned measurements. All KPIs exhibit value ranges and central tendencies consistent with expected cellular radio conditions in vehicular environments, indicating that no further systematic invalid values remain after the applied filtering steps.

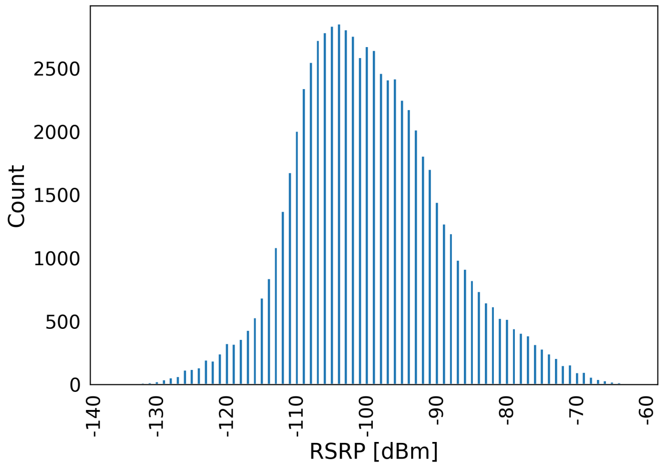

Figure 5. Frequency distribution of RSRP after cleaning the dataset.

TABLE II. Summary statistics of cleaned radio KPI measurements after removal of explicit and implicit placeholder values.

KPI	Count	Min	Max	Mean	Median
RSRP [dBm]	72 606	-140	-59	-99.6	-100
RSRQ [dB]	72 627	-20	-3	-13.7	-13
RSSI [dBm]	72 594	-115	-26	-65.8	-67
SINR [dB]	72 538	-20	30	4.7	4

In addition to cleaning, the dataset was enriched with contextual information describing the environment. Base station proximity was incorporated by computing the distance between each measurement point and the serving base station. The base station locations were obtained from the open-source platform CellMapper [30], and distances were computed using haversine-based spatial calculations. Also, historical weather data were added to each measurement point using the Open-Meteo Application Programming Interface (API) [31]. The weather attributes (e.g., temperature, wind speed, precipitation, ground fog) were matched based on timestamp and geographical coordinates, allowing for the inclusion of environmental context in the analysis process.

As a result of this process, a spatiotemporally tagged dataset that combines passive network parameters, location, and environmental conditions was created. This dataset forms the basis for the subsequent spatial stability and variability analysis.

VI. Spatial Network Variability

To gain insights into the structure and variability of the collected dataset, an Exploratory Data Analysis (EDA) was conducted. This analysis aims to find spatial and contextual patterns in the signal measurements and to identify the most relevant features for subsequent modeling tasks.

A. Route Segmentation

As shown in Figure 6, the measurement route was segmented into Reference Points (RPs) along the road. This resulted in a total of 320 RPs covering the test route. The segmentation enables consistent spatial aggregation by grouping measurements recorded at the same physical location across repeated drives. For each RP, the mean, median, minimum, and maximum values of the target parameters were calculated. This helps to mitigate the influence of measurement noise and improves the robustness of comparisons between different measurement runs.

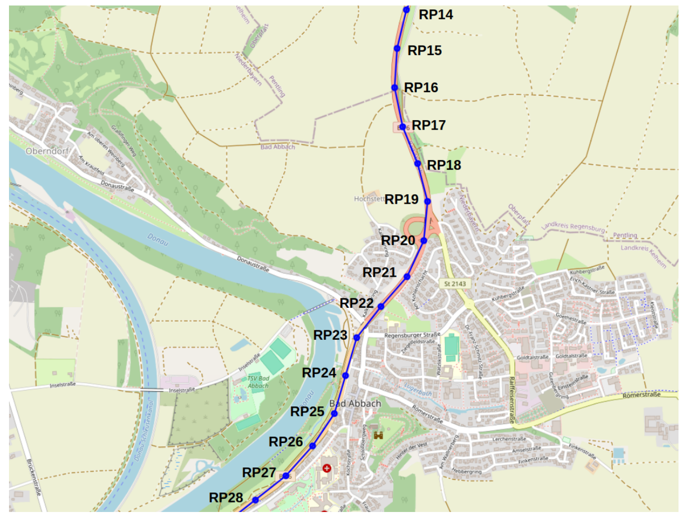

Figure 6. Extract of the route clustered into reference points with a distance of 200 m.

To limit the analysis to measurements recorded on the predefined test route, each sample was assigned to its nearest RP. The RPs were placed on the road centerline with a fixed spacing of 200 m. Consequently, measurements located less than half of the spacing away from their nearest RP originate from the interior of the route. Measurements that exceed the 100 m distance to a RP are likely associated with boundary regions or could also be off-route segments.

TABLE III. Validation of distance cutoff values based on reference point exceedances.

Cutoff [m]	Reference Point ID	Count
100	1	13 932
	11	1
	19	1
	97	1
	123	1
	152	2
	319	1
	320	1 883
101	1	13 931
	152	2
	320	1 882
102	1	13 930
	320	1 880

To determine an appropriate distance threshold, increasing cutoff values were tested, and the occurrence of threshold exceedances across the RPs was inspected. Table III summarizes the number of samples that exceed different distance thresholds and their associated RP identifiers. For cutoff values of 100 m and 101 m, a small number of on-route measurements were still observed. This could plausibly be due to GNSS jitter. At a cutoff of 102 m, all exceedances occurred only at the first and last RPs, which are at the beginning and the end of the route. No interior RPs exhibited distances exceeding this threshold. Based on this observation, a distance cutoff of 102 m was selected as the smallest value that reliably excludes off-route measurements while preserving all valid interior route samples.

B. Variability Across Repeated Drives

After associating the measurements with their nearest RPs, the next step is to analyze how mobile network connectivity differs across repeated drives. While the vehicle's location is a dominant factor influencing signal quality, an initial analysis confirms that connectivity conditions at the same location can vary substantially from trip to trip, as also suggested by prior work discussed in Section II.

Figure 7 illustrates the relationship between the mean probability of usable connectivity and its trip-to-trip variability for each RP.

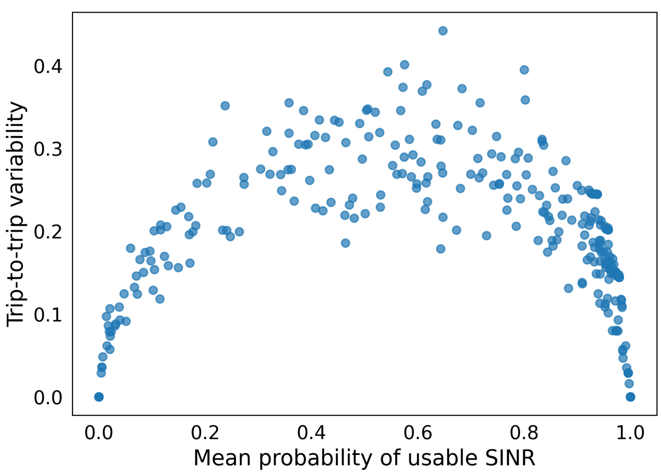

Figure 7. Trip-to-trip variability of functional connectivity based on the SINR. Every point shows one reference point of the route.

For this plot, connectivity was classified as usable if the measured SINR was available and higher than 0. Accordingly, the mean value shows the measurements with functionally usable connectivity across the drives. The standard deviation quantifies how this functional state fluctuates between trips.

Route sections with consistently good connectivity (x-axis close to 1) and sections with persistently poor connectivity (x-axis close to 0) showed very low variability (y-axis), indicating stable connectivity conditions across drives. In contrast, sections with intermediate mean values consistently show trip-to-trip variability. At these locations, the same geographic position on the route alternates between usable and non-usable connectivity depending on the individual drive.

The resulting inverted U-shaped distribution shows that connectivity cannot be sufficiently characterized by static coverage maps or average signal metrics alone [4][15]. The same analysis was conducted for the RSRP, showing the same distribution. These findings underline the necessity of incorporating additional contextual information when assessing mobile network conditions before starting a trip.

C. Stability Classification

Based on the per-trip aggregation of functional connectivity, each route section was classified according to its stability across repeated drives. Four distinct stability classes were defined.

- *Stable good* sections show consistently usable connectivity across almost all trips. These segments represent areas where a good mobile connection can be expected, independent of contextual variations.
- *Stable hole* sections are characterized by persistently non-usable connectivity. Across repeated drives, these segments rarely reach sufficient signal quality for data transmission, indicating stable coverage gaps.
- *Unstable* sections are those in which functional connectivity alternates between usable and non-usable states across different measurement drives. As a result, the same physical location may support reliable data transmission during one drive while being unusable during another. This behavior cannot be captured by static coverage representations or average signal metrics.
- *Mixed* sections denote transitional segments that do not exhibit pronounced trip-to-trip state changes, yet cannot be classified as consistently usable or consistently non-usable. These sections reflect moderate variability without clear state transitions.

TABLE IV. CLASSIFICATION OF ROUTE SECTIONS ACCORDING TO THEIR FUNCTIONAL CONNECTIVITY STABILITY ACROSS REPEATED DRIVES.

Stability Class	Number of Sections	Share [%]
Stable good	25	7.8
Stable hole	102	31.9
Unstable	111	34.7
Mixed	82	25.6
Total	320	100.0

Table IV shows the distribution of those classes along the route. Notably, only a small fraction of the route sections (7.8%) exhibit consistently good connectivity, whereas about one third (31.9%) are classified as stable holes. The remaining 60.3% of the route cannot be uniquely characterized by a single functional connectivity state.

Figure 8 illustrates the spatial distribution of the stability classes along a part of the route. Stable hole sections mostly appear in rural and near forested areas, whereas stable good connectivity is mainly observed near villages. Unstable sections frequently occur in clusters between these extremes. This finding indicates that functional connectivity uncertainty is not randomly distributed but is tied to specific environmental and infrastructural characteristics.

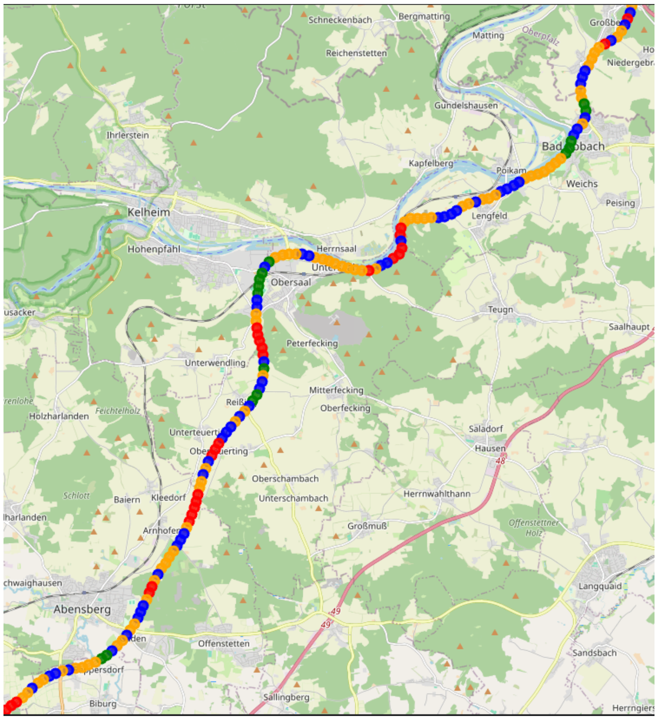

Figure 8. Spatial distribution of route sections classified by functional connectivity stability across repeated drives. The colors show the four classes: stable good (green), stable hole (red), mixed (blue) and unstable (orange).

The stability analysis shows that a significant part of the route cannot be characterized by a single functional connectivity state. Instead, several route sections exhibit pronounced trip-to-trip variability, alternating between usable and non-usable connectivity despite identical spatial locations.

This observation underlines the research question of whether such variability can be partially attributed to contextual factors that differ between individual drives. Environmental conditions, temporal aspects, and geometric relationships to the cellular infrastructure may influence the instantaneous radio conditions and thereby affect the functional connectivity state observed along a given route section.

In the following section, the role of contextual parameters and their relationship with the observed connectivity variability is analyzed. The goal is to assess whether contextual information provides systematic indications that can support pre-trip connectivity prediction under uncertainty.

VII. Influences of Contextual Factors

To improve the prediction of mobile network KPIs, particularly in areas with weak or inconsistent coverage, it is important to account for environmental and temporal influences. Contextual parameters are assumed to influence radio propagation and therefore the network performance [24][28].

The previous study [1], based on 38 measurement drives, investigated correlations between contextual features and the four passive radio quality indicators RSRP, RSRQ, RSSI, and SINR. The analyzed context variables include weather conditions, temporal features, distance to the serving base station, GNSS precision metrics (e.g., PDOP and the number of visible satellites), and vehicle speed. Figure 9 summarizes the correlations reported in that study.

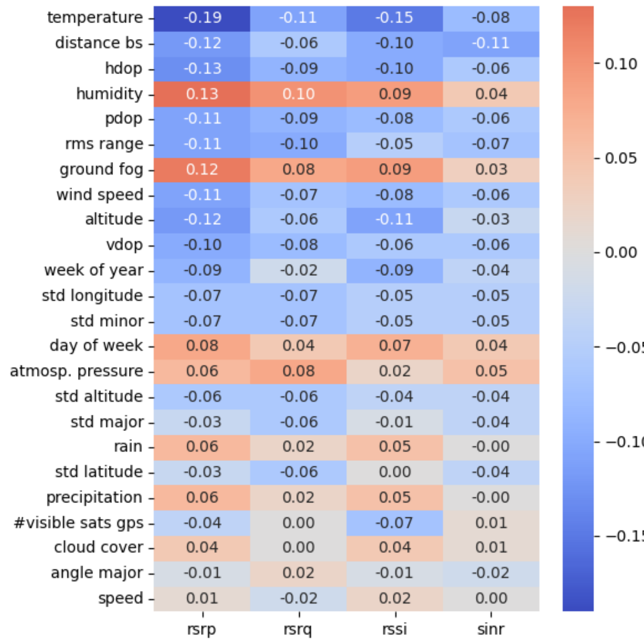

Figure 9. Heatmap showing the correlations between environmental parameters and received signal quality metrics. Based on 38 test drives of the previous study [1].

These first results indicate weak to moderate correlations for the parameters. In particular, ambient temperature showed a negative correlation with RSRP and RSSI, while humidity and ground fog exhibited small positive correlations with multiple signal metrics. Temporal features such as the day of the week and the week of the year also showed minor correlations, potentially reflecting small time-dependent network load. Vehicle speed, in contrast, showed no relevant influence on radio signal quality.

For the pre-trip connectivity forecast assessment in autonomous driving scenarios, it is essential to verify whether such contextual parameters provide explanatory or predictive value beyond spatial location. Therefore, a similar correlation analysis with an extended dataset of 48 measurement drives was conducted. Because weather parameters appeared to be the most promising, the analysis focused on them. As a first step, correlations between weather and passive mobile network indicators were computed over the entire dataset. Missing radio measurements (NaN due to no signal at all) were replaced by worst-case reception values to preserve the distinction between weak signal and missing connectivity. Both Pearson correlation coefficients, which capture linear relationships, and Spearman correlations, which capture monotonic dependencies, were evaluated.

Across the complete route, all correlations between weather parameters and radio KPIs remain close to zero. This observation is consistent for temperature, relative humidity, precipitation, wind speed, and surface pressure. Table V reports the Spearman correlation coefficients for the global dataset.

TABLE V. SPEARMAN CORRELATION BETWEEN WEATHER PARAMETERS AND RADIO KPIS (ENTIRE DATASET, NANS REPLACED BY WORST-CASE VALUES).

Context	RSRP	RSRQ	RSSI	SINR
Temperature	-0.06	-0.00	-0.07	0.00
Relative humidity	0.04	-0.00	0.04	-0.02
Precipitation	0.02	0.01	0.02	-0.01
Wind speed	-0.02	0.00	-0.02	0.01
Surface pressure	0.03	0.02	0.04	0.04

This result indicates that environmental context alone does not explain the dominant spatial structure of mobile network quality. Instead, infrastructure-related factors and geographic location remain the primary determinants on a global scale.

To isolate potentially context-sensitive behavior, the analysis was repeated for route sections classified as *unstable* or *mixed*, as described in Section VI-C. These sections showed high trip-to-trip variability and therefore represent critical candidates for contextual influence. In these selected sections, correlations increase slightly in magnitude compared to the global analysis but remain weak overall. This suggests that weather-related context does not directly determine instantaneous signal strength.

As the next step, the analysis was shifted from numeric signal values to a functional connectivity perspective. Radio KPIs are mapped to binary usability indicators based on the thresholds for poor connectivity as described in Subsection III-B. If the KPI is above the respective threshold, it is defined as usable; otherwise, it is defined as not usable. This transformation reflects whether mobile connectivity is sufficient for data transmission rather than absolute signal strength. The analysis was again restricted to unstable and mixed route sections. For each measurement drive, the share of usable signal conditions was computed and correlated with aggregated contextual parameters.

When evaluating the share of functional connectivity instead of raw signal values, moderate correlations with contextual parameters can be observed. In particular, higher ambient temperatures are associated with a reduced probability of usable connectivity. Additionally, higher relative humidity and precipitation show a small positive association with usable signal conditions. Table VI summarizes these Spearman correlations between the weather parameters and the share of usable connectivity states in unstable and mixed route sections.

The effects are clearer for RSRP and RSSI than for RSRQ and SINR. The observed correlations suggest that environmental factors affect the likelihood of functional connectivity, particularly in spatially unstable regions. However, they are also insufficient to predict connectivity state transitions deterministically on their own. Contextual information provides

TABLE VI. SPEARMAN CORRELATION BETWEEN CONTEXTUAL PARAMETERS AND THE SHARE OF USABLE SIGNAL CONDITIONS IN UNSTABLE AND MIXED ROUTE SECTIONS.

Context	RSRP	RSRQ	RSSI	SINR
Temperature mean	**-0.53**	-0.10	**-0.62**	-0.01
Humidity mean	**0.32**	0.03	**0.35**	-0.21
Wind speed mean	-0.29	-0.08	-0.27	-0.06
Surface pressure	0.24	0.16	0.16	-0.06
Rain sum	0.12	0.21	0.04	-0.15

complementary value by quantifying uncertainty tendencies. This motivates the integration of contextual parameters into higher-level probabilistic models rather than relying on direct signal prediction, as shown in Section VIII-C.

A. Correlation with Ambient Temperature

To further investigate the influence of temperature, Figure 10 illustrates the average RSRP values across different temperature bins, separated by connection type. The diagram originates from the previous study with a total of 38 measurement drives. A gradual degradation in signal strength with increasing ambient temperature can be observed. LTE connections appear to be more sensitive to temperature increases than 5G-NSA, which seems to be a little more robust across the temperature range.

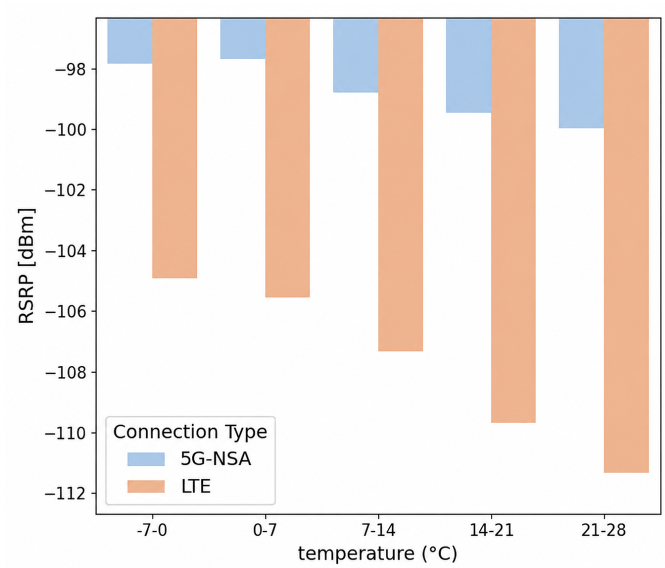

Figure 10. Influence of the ambient temperature on the RSRP value separated into 5G-NSA and LTE measurements. Lower RSRP corresponds with lower network availability.

To show another perspective, Figure 11 evaluates temperature effects from a functional connectivity standpoint, as also used in Section VII. In this case, all 48 measurement drives were included. Instead of analyzing absolute signal values, the figure shows the share of functionally usable RSRP measurements per trip, aggregated over repeated drives and grouped by mean ambient temperature.

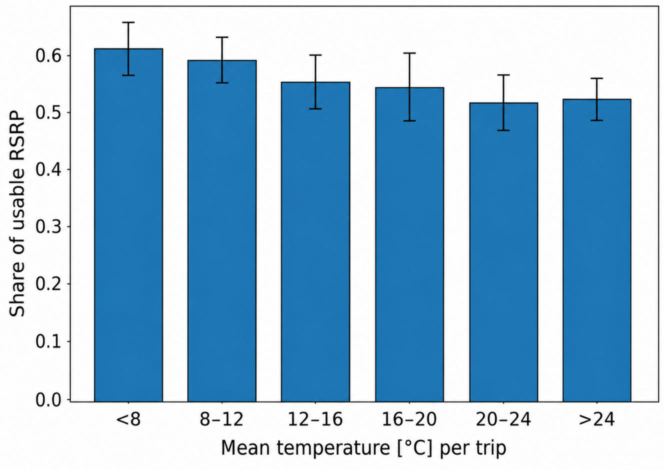

Figure 11. Influence of the ambient temperature on the share of functionally usable RSRP values aggregated per trip. The black lines indicate the value range of the respective temperature bin.

A gradual decrease in the probability of usable RSRP can be observed with increasing temperature up to approximately 24°C. This trend is consistent with the correlation analysis presented in Section VII, which revealed moderate relationships between temperature and functional connectivity metrics. The highest temperature bin (>24°C) contains only three measurement drives and is therefore not considered statistically representative.

The corresponding statistics across all KPIs further support this observation. While the average share of usable connectivity remains above 0.45 for all metrics, the minimum observed shares indicate that higher temperatures coincide with more frequent connectivity degradation events, particularly for RSRP and RSSI. In contrast, SINR shows a smaller sensitivity, suggesting that interference and noise conditions partially compensate for temperature-related signal attenuation.

VIII. Connectivity Forecasting Model

The discussed correlations provide a foundation for feature selection and motivate the inclusion of contextual parameters into a predictive model, which is described in this section. The Mean Absolute Error (MAE) was used to assess the influence of contextual features on the target KPIs. MAE is chosen for its interpretability and robustness to outliers. It provides a direct average of absolute errors in the same units as the target variable, making it suitable for this purpose. The MAE is defined as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

where y_i denotes the true value, $\hat{y}_i$ the predicted value, and n the number of samples [32].

As mentioned in Section III-B, the target KPIs are expressed in dB, respectively dBm, which are both logarithmically scaled. Therefore, even small changes in the MAE can have a significant impact on the effective signal quality [26].

A. Context Parameter Scaling

To investigate how external context parameters affect mobile network signal quality, a modeling approach that adjusts the section-wise signal predictions based on environmental deviations was developed. The baseline model predicts the expected target parameters for every RP along the route using only the mean value from historical measurements. This section explores whether incorporating individual context parameters can improve the prediction by reducing the MAE.

For each context parameter, a linear model was trained using the deviation between the parameter's current value and its historical section average. This deviation was multiplied by a scale factor to model the resulting shift in signal behavior. The prediction was then computed by adding this adjustment to the baseline section average. To identify the optimal scale factor, values ranging from –5.0 to 15.0 in increments of 0.1 for each context parameter were tested. The scale factor that minimized the MAE on the test measurement was selected. The goal was to quantify how strongly each context parameter contributes to prediction accuracy and whether any of them provided consistent improvements over the baseline.

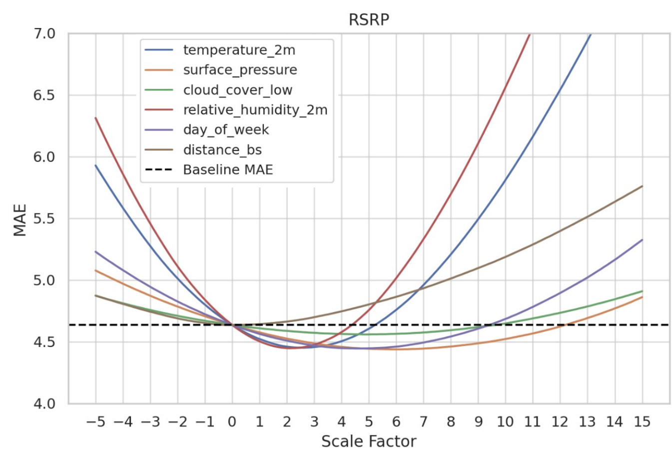

Figure 12. Influence of the context parameters with different scale factors on the MAE compared to the baseline MAE without context

Figure 12 shows the results of the scale factor optimization for the signal parameter RSRP. Each curve in the plot represents the MAE progression as a function of the scale factor applied to the respective context parameter. In all cases, the MAE initially decreases with increasing scale factor, reaches a minimum, and then increases again. The findings show that most context parameters can improve prediction accuracy compared to the baseline. However, the improvement varies between the parameters. The distance to the base station could not improve the MAE. This behavior was expected because the base station for every RP is the same for most drives, and therefore, the distance is a fixed parameter.

B. Combining Context Parameters

The aim of this subsection is to investigate whether the prediction accuracy of mobile network parameters could be further improved by simultaneously incorporating multiple context parameters. Building on the results of the individual analysis, it was examined whether combining two context parameters, each with its own scale factor, could reduce the MAE even further. Therefore, all possible pairs of the five most relevant context parameters (temperature, surface pressure, relative humidity, ground fog, and day of week) were systematically evaluated. For each pair, a linear model was trained using the deviation of each context value from its corresponding section average. A grid search over scale factor combinations in the range -5 to 15 (in steps of 0.1) was conducted. The goal was to identify the combination that minimized the MAE for a separate test measurement.

However, the results in Table VII show that none of the tested parameter combinations yielded significantly better performance than the single context parameter. For RSRP and SINR, the MAE even increased when using two parameters instead of one.

TABLE VII. MAE COMPARISON OF BASELINE, BEST SINGLE FEATURE, AND BEST FEATURE COMBINATION

Metric	Baseline	Best Feature	Best Combination
RSSI	3.684	3.556 Surface Pressure	3.548 Surface Pressure Humidity
RSRP	4.635	4.437 Surface Pressure	4.440 Surface Pressure Humidity
RSRQ	2.158	2.065 Surface Pressure	2.063 Surface Pressure Ground Fog
SINR	3.041	3.017 Temperature	3.024 Temperature Surface Pressure

This finding suggests that the influence of context parameters on mobile network measurements cannot be adequately described by a simple linear superposition. Another explanation may be the multicollinearity among weather variables, which adds redundancy and inflates linear model variance. For example, temperature and relative humidity are often negatively correlated [24]. It is likely that complex interactions or nonlinear relationships exist between the context features that are not captured by the linear model.

C. Pre-Trip Temperature-Conditioned Forecasting

Since the previous subsections used 38 measurement drives, another approach was conducted using data from 48 drives. The extended experiment focuses on the influence of environmental temperature on the RSRP because this combination is the most examined in Section VII. Under real-world conditions, only historical measurements are available before the start of a trip. Therefore, a walk-forward approach was implemented in which each measurement campaign was predicted using only past trips. A minimum of ten historical trips was required before a forecast was generated.

The baseline model in this experiment predicts the expected signal level for each route section using only the historical section mean value. Building on the correlation analysis in Section VII, temperature was taken as a global adjustment factor. Instead of performing scale-factor optimization, a linear regression [33] was trained on historical averages per trip:

$$\bar{y}_{trip} = \alpha + \gamma \bar{T}_{trip} \tag{2}$$

where $\bar{y}_{trip}$ denotes the mean signal level of a historical trip, and $\bar{T}_{trip}$ denotes its corresponding mean temperature. The resulting temperature shift applied to a future trip is then defined as:

$$\Delta y = \gamma(\bar{T}_{trip} - \bar{T}) \tag{3}$$

where $\bar{T}$ is the historical mean trip temperature. The final prediction becomes:

$$\hat{y}_s^{\text{temp}} = \mu_s + \Delta y \tag{4}$$

where μ_s denotes the historical mean signal level of section s. The results of this simple model using the extended set of measurements are as follows. For the RSRP, the baseline achieved the following MAE:

$$\text{MAE}_{\text{baseline}} = 4.443 \pm 0.233 \text{ dB}$$

Incorporating the global temperature adjustment reduced the error to:

$$\text{MAE}_{\text{temp}} = 4.415 \pm 0.211 \text{ dB}$$

corresponding to an average improvement of:

$$\Delta\text{MAE} = 0.028 \text{ dB}$$

The absolute improvement is small, but it must be interpreted in the context of logarithmic signal scaling. The model, taking into account the ambient temperature, reached a lower error in 73.7% of all trips. This shows that temperature can help predict signal metrics, but the influences seem to be more complex than a simple linear model can represent. The estimated temperature sensitivity was:

$$\gamma = -0.109 \pm 0.021 \text{ dB/°C}$$

indicating that higher temperatures are associated with systematically reduced expected RSRP levels. The uncertainty term ± 0.021 dB/°C is the standard error of the estimated coefficient.

An additional analysis examined only route sections with strong trip-to-trip variability. Sections exhibiting a standard deviation of 6 dB across historical trips were classified as variable. In total, 37 of the 320 sections met this criterion. When the forecasting model was restricted only to these variable sections, the temperature adjustment produced an absolute improvement of:

$$\Delta\text{MAE}_{\text{variable}} = 0.071 \text{ dB}$$

This confirms that contextual information is particularly relevant in dynamically fluctuating route segments, whereas permanently stable sections are already well captured by the historical mean.

IX. Discussion

The study demonstrates that location alone is insufficient to characterize mobile network availability along a route. The stability analysis in Section VI revealed that more than half of the sections on the specific route cannot be described by a functional connectivity state (stable good or stable hole). In particular, sections classified as unstable show high trip-to-trip variability, indicating that connectivity uncertainty is a property of certain spatial environments.

The forecasting experiments show that including contextual information yields moderate but consistent reductions in prediction error. Although the MAE improvements are small in magnitude, the win-rate across trips suggests that contextual conditioning captures variability rather than noise. The results further indicate that contextual effects are not globally uniform but are concentrated in spatially unstable sections. This suggests that context influences mobile network uncertainty rather than shifting signal levels in a linear way.

The weak global correlations between weather parameters and radio KPIs contrast with the moderate correlations observed when analyzing functional usability in unstable sections. This indicates that contextual variables do not directly determine instantaneous signal metrics but influence the likelihood of a route section's current state between usable and non-usable. Rather than treating variability as noise, the findings suggest that spatial instability zones represent structurally sensitive regions where environmental conditions influence existing propagation weaknesses.

The limited improvements achieved by the linear integration of multiple context parameters indicate that the relationship between environmental factors and radio propagation may be nonlinear. Multicollinearity among weather variables further limits interpretability within linear models. The findings support the need for nonlinear modeling approaches capable of capturing conditional dependencies and interaction effects, particularly within spatially unstable sections.

A. Implications for Vehicle Connectivity

The findings of the analysis have different implications for the operation of modern (automated) vehicles. Pre-trip forecasts of route-section-wise connectivity quality could be used to reschedule the path between start and finish for cars for which a continuous network connection is required (SAE Level 4). In addition, such forecasts can enable strategic scheduling of communication-intensive but non-safety-critical data transfers. For example, the upload of high-resolution sensor logs, map updates, or cloud-based analytics can be scheduled accordingly. By allocating large data packets to route segments with predicted stable and high-quality connectivity, vehicles can reduce the probability of transmission interruption, preserve bandwidth for safety-relevant remote supervision, and minimize fallback maneuvers triggered by temporary degradation. Conversely, predicted unstable or low-quality segments could be internally marked as communication-restricted zones in which only safety-critical data streams are maintained. These use cases can make connectivity forecasts an important component of active communication management for Level 4 automated vehicles.

The practical integration of the proposed forecasting framework into automated vehicle architectures will be part of future work. In such an implementation, the connectivity forecasting module could act as an additional network planning and communication-management layer. Before a trip, the module would combine the planned route, historical route-section-wise connectivity data, map information, network operator information, and contextual parameters such as weather conditions. During operation, it could be updated with live radio measurements from the vehicle's communication unit. The resulting outputs could include route-section-wise connectivity risk scores, predicted signal quality, and warnings for unstable or communication-restricted segments.

X. Limitations

The dataset is restricted to a single rural route and one mobile network operator. Urban environments were not considered but are often analyzed in other studies [6][7][11]. All measurements were conducted using a single hardware configuration, which limits generalizability across device types. Temporal coverage is limited to the measurement period of one year, and some meteorological events are not represented.

Functional usability was defined using fixed KPI thresholds derived from industry recommendations. No direct throughput, latency, or packet-loss measurements were incorporated into the usability classification. Therefore, the functional connectivity state represents a proxy for application-level performance rather than a validated QoS.

The forecasting models rely on linear mechanisms and section-wise historical averaging. Nonlinear relationships and temporal dependencies between consecutive route sections were not modeled. The walk-forward validation approach requires a minimum number of historical trips before generating forecasts. While this ensures temporal realism, it limits applicability in early deployment stages where only limited historical data is available.

XI. Conclusion and Future Work

This paper presents a spatio-temporal and context-aware analysis of mobile network stability along a fixed route based on 48 repeated measurement drives. The stability classification revealed that a substantial portion of the route exhibits trip-to-trip variability that cannot be captured by static coverage representations. Contextual conditioning yields consistent reductions in prediction error and improves forecasting performance, especially in unstable geographic locations. The results indicate that context parameters act as uncertainty modulators rather than deterministic predictors of signal quality. The study demonstrates that connectivity forecasting can support not only connectivity-aware route planning but also vehicle communication management strategies.

Future work will focus on extending the modeling framework. Nonlinear ML approaches (e.g., gradient boosting, random forests, or neural networks) will be investigated to capture more complex conditional dependencies. The dataset will also be expanded to include an additional route and another measurement platform to improve generalizability and robustness. Future measurement campaigns will include connectivity performance indicators, such as throughput and latency, alongside passive radio KPIs. Through these extensions, an advanced framework for context-aware mobile network prediction for automated driving environments will be created. Such capabilities represent important improvements for meeting the safety and reliability requirements of modern and future vehicles, as well as for mobile data exchange in general.

References

[1] M. Gerstner and R. Hackenberg, "Context-aware forecasting of mobile network quality for autonomous vehicle connectivity", in *Proc. Vehicular Analytics 2025, The Second International Conference on Vehicular Systems*, IARIA, 2025.

[2] Goldman Sachs, *Partially autonomous cars forecast to comprise 10% of new vehicle sales by 2030*, [Online]. Available: https://www.goldmansachs.com/insights/articles/partially-autonomous-cars-forecast-to-comprise-10-percent-of-new-vehicle-sales-by-2030. Accessed: Feb. 19, 2026.

[3] A. Takacs and T. Haidegger, "A method for mapping V2X communication requirements to highly automated and autonomous vehicle functions", *Future Internet*, vol. 16, no. 4, p. 108, 2024. DOI: 10.3390/fi16040108.

[4] R. Blasco, G. C. Ferrante, C. Watermann, and A. Palaios, "Predictive quality of service in cellular networks: Challenges, framework, and application in vehicular communications", *IEEE Commun. Mag.*, vol. 61, no. 3, pp. 44–49, 2023. DOI: 10.1109/MCOM.004.2200180.

[5] Federal Ministry of Justice and Consumer Protection, *Act to amend the road traffic act and the compulsory insurance act – act on autonomous driving*, [Online]. Available: http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl121s3108.pdf. Accessed: Mai. 21, 2026.

[6] D. Madariaga, M. Panza, and J. Bustos-Jimenez, "I'm only unhappy when it rains: Forecasting mobile QoS with weather conditions", in *Proc. 2018 Network Traffic Measurement and Analysis Conf. (TMA)*, 2018, pp. 1–6. DOI: 10.23919/tma.2018.8506509.

[7] H. Schippers, M. Geis, S. Böcker, and C. Wietfeld, "DoNext: An open-access measurement dataset for machine learning-driven 5G mobile network analysis", *IEEE Trans. Mach. Learn. Commun. Netw.*, 2025. DOI: 10.1109/TMLCN.2025.3564239.

[8] Federal Motor Transport Authority, *Levels of automation*, [Online]. Available: https://www.kba.de/EN/Themen_en/Marktueberwachung_en/Produktpruefungen_en/AutomatisiertesAutonomesFahren_en/Automatisierungsstufen_en/Automatisierungsstufen_node_en.html. Accessed: Mai. 21, 2026.

[9] Federal Republic of Germany, *Ordinance on the approval and operation of motor vehicles with autonomous driving functions in defined operating areas – AFGBV*, [Online]. Available: https://www.buzer.de/AFGBV.htm. Accessed: Mai. 20, 2026.

[10] P. Torres et al., "Data analytics for forecasting cell congestion on LTE networks", in *Proc. 2017 Network Traffic Measurement and Analysis Conf. (TMA)*, 2017, pp. 1–6. DOI: 10.23919/tma.2017.8002917.

[11] D. Madariaga, J. Madariaga, J. Bustos-Jiménez, and B. Bustos, "Improving signal-strength aggregation for mobile crowdsourcing scenarios", *Sensors*, vol. 21, no. 4, p. 1084, 2021. DOI: 10.3390/s21041084.

[12] O. Sahin and V. Sathya, *Network quality prediction in a designated area using gps data*, SSRN preprint, 2023. DOI: 10.2139/ssrn.4525388.

[13] M. Rahman, I. Guvenc, D. Ramirez, and C.-W. Wong, *Uav-assisted coverage hole detection using reinforcement learning in urban cellular networks*, 2025. DOI: 10.48550/arXiv.2503.06494.

[14] T. Hultman, A. Boudjadar, and M. Asplund, "Connectivity-optimal shortest paths using crowdsourced data", in *Proc. IEEE Int. Conf. on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 2016, pp. 1–6. DOI: 10.1109/percomw.2016.7457106.

[15] A. Palaios et al., "Machine learning for QoS prediction in vehicular communication: Challenges and solution approaches", *IEEE Access*, vol. 11, pp. 92 459–92 477, 2023. DOI: 10.1109/ACCESS.2023.3303528.

[16] G. Soos, D. Ficzere, P. Varga, and Z. Szalay, "Practical 5G KPI measurement results on a non-standalone architecture", in *Proc. NOMS 2020*, 2020, pp. 1–5. DOI: 10.1109/noms47738.2020.9110457.

[17] Z. Szalay et al., "5G-enabled autonomous driving demonstration with a V2X scenario-in-the-loop approach", *Sensors*, vol. 20, no. 24, p. 7344, 2020. DOI: 10.3390/s20247344.

[18] C. Inc., *LTE category guide: Enhancing wireless communication*, [Online]. Available: https://www.cavliwireless.com/blog/nerdiest-of-things/lte-category-wireless-communication. Accessed: Mai. 21, 2026.

[19] H. Fehmi, M. F. Amr, A. Bahnasse, and M. Talea, "5G network: Analysis and compare 5G NSA / 5G SA", *Procedia Computer Science*, vol. 203, pp. 594–598, 2022. DOI: 10.1016/j.procs.2022.07.085.

[20] I. Maskulainen et al., "Performance evaluation of adaptive beamforming in 5G-V2X networks", in *Proc. European Conf. on Networks and Communications (EuCNC)*, 2017, pp. 1–5. DOI: 10.1109/EuCNC.2017.7980728.

[21] V. A. Reddy et al., "Fundamental performance limits of mm-wave cooperative localization in linear topologies", *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 1899–1903, 2020. DOI: 10.1109/LWC.2020.3007417.

[22] Teltonika Networks, *Mobile signal strength recommendations*, [Online]. Available: https://wiki.teltonika-networks.com/view/Mobile_Signal_Strength_Recommendations. Accessed: Mai. 21, 2026.

[23] R. Molina-Masegosa, M. Sepulcre, J. Gozalvez, F. Berens, and V. Martinez, "Empirical models for the realistic generation of cooperative awareness messages in vehicular networks", *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5713–5717, 2020. DOI: 10.1109/TVT.2020.2979232.

[24] R. H. Ratul et al., *Atmospheric influence on the path loss at high frequencies for deployment of 5g cellular communication networks*, arXiv:2306.01894, 2023.

[25] A. L. Imoize, S. O. Tofade, G. U. Ughegbe, F. I. Anyasi, and J. Isabona, "Updating analysis of key performance indicators of 4G LTE network with the prediction of missing values of critical network parameters based on experimental data from a

dense urban environment", *Data in Brief*, vol. 42, p. 108 240, 2022. DOI: 10.1016/j.dib.2022.108240.
[26] F. Afroz, R. Subramanian, R. Heidary, K. Sandrasegaran, and S. Ahmed, "SINR, RSRP, RSSI and RSRQ measurements in long term evolution networks", *Int. J. Wireless & Mobile Networks*, vol. 7, no. 4, pp. 113–123, 2015. DOI: 10.5121/ijwmn.2015.7409.
[27] Techlteworld, *RSSI (received signal strength indicator) in LTE*, [Online]. Available: https://techlteworld.com/rssi-received-signal-strength-indicator-in-lte/. Accessed: Mai. 21, 2026.
[28] A. Molisch, F. Tufvesson, J. Karedal, and C. Mecklenbrauker, "A survey on vehicle-to-vehicle propagation channels", *IEEE Wireless Commun.*, vol. 16, no. 6, pp. 12–22, 2009. DOI: 10.1109/MWC.2009.5361174.
[29] S. Farthofer et al., "An open mobile communications drive test data set and its use for machine learning", *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1688–1701, 2022. DOI: 10.1109/OJCOMS.2022.3210289.
[30] CellMapper, *Cellular coverage and tower map*, [Online]. Available: https://www.cellmapper.net/. Accessed: Jan. 18, 2026.
[31] Open-Meteo, *Free open-source weather API*, [Online]. Available: https://open-meteo.com/. Accessed: Sep. 18, 2025.
[32] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance", *Climate Research*, vol. 30, no. 1, pp. 79–82, 2005. DOI: 10.3354/cr030079.
[33] O. Zeigermann and C. N. Nguyen, *Machine Learning: A Brief Introduction with Python, Scikit-Learn, and TensorFlow*, 3rd. Heidelberg, Germany: O'Reilly, 2024.