High-Efficiency Urban 3D Radio Map Estimation Based on Sparse Measurements

Xinwei Chen, Xiaofeng Zhong, Zijian Zhang, *Graduate Student Member, IEEE*, Linglong Dai, *Fellow, IEEE*, and Shidong Zhou, *Member, IEEE*

***Abstract*—Recent widespread applications for unmanned aerial vehicles (UAVs) – from infrastructure inspection to urban logistics – have prompted an urgent need for high-accuracy three-dimensional (3D) radio maps. However, existing methods designed for two-dimensional radio maps face challenges of high measurement costs and limited data availability when extended to 3D scenarios. To tackle these challenges, we first build a real-world large-scale 3D radio map dataset, covering over 4.2 million m^3 and over 4 thousand data points in complex urban environments. We propose a Gaussian Process Regression-based scheme for 3D radio map estimation, allowing us to realize more accurate map recovery with a lower RMSE than state-of-the-art schemes by over 2.5 dB. To further enhance data efficiency, we propose two methods for training point selection, including an offline clustering-based method and an online maximum a posterior (MAP)-based method. Extensive experiments demonstrate that the proposed scheme not only achieves full-map recovery with only 2% of UAV measurements, but also sheds light on future studies on 3D radio maps.**

***Index Terms*—3D radio map, Unmanned aerial vehicle (UAV), Gaussian Process Regression (GPR), sparse measurements.**

I. Introduction

Driven by the burgeoning demand for smart city solutions, unmanned aerial vehicle (UAV) technology has become an integral component [1]. Due to their inherent flexibility, cost-effectiveness, and eco-friendliness, UAVs have demonstrated remarkable growth in applications such as inspection, disaster relief, and logistics. In the logistics field, leading companies such as Meituan and SF Express have widely adopted UAVs, thereby significantly boosting delivery capacity and efficiency across diverse terrains and scenarios.

However, the deployment of UAVs in urban areas encounters considerable challenges in maintaining robust communication during missions, a critical factor for effective flight trajectory planning [2], [3]. First, navigating low-altitude urban areas is complex due to the impact of dense buildings on communication signal strength and spectrum resources, which vary significantly in three-dimensional (3D) spaces. Second, the high cost of data measurements and the restricted flight accessibility of UAVs result in limited data availability [4]. Third, there is an urgent need for efficient updates to adapt to the dynamic urban development. Consequently, recovering comprehensive 3D radio maps from sparse UAV measurements becomes indispensable.

Existing works have shown that altitude significantly impacts signal strength and throughput of transceivers, with differences at various heights reaching up to a factor of two [5]. Current open-source radio map datasets [6], [7] and estimation schemes [8], [9], [10] are usually two-dimensional (2D), and there are currently no open-source 3D real-world datasets available, which is insufficient for research on UAV trajectory planning. Most existing methods for radio map estimation are data-driven, which are high-cost and difficult to achieve comprehensive coverage, especially in 3D scenarios [11], [12], [13], [14]. However, model-driven approaches, while avoiding extensive data collection, often rely on simplified assumptions, yielding model errors of approximately 13.5 dB [15], [16]. Due to their inherent limitations in capturing the complex signal variability of real-world scenarios, these approaches are typically hard to generalize in new environments or represent signal dead zones [17].

To achieve high-efficiency 3D radio map recovery, this paper develops an urban 3D radio map estimation scheme, which utilizes the sparse data collected by UAVs. In urban areas, inadequate base station coverage often leads to disconnections. To mitigate this, we propose an urban 3D radio map estimation platform that leverages base station information, sparse data selection and measurement, and urban architectural information. The estimated 3D radio Reference Signal Received Power (RSRP) can then be used to guide UAV trajectory for better data collection. We summarize our contributions as follows:

- An urban 3D radio signal measurement dataset collected via UAVs in densely-built areas increases complexity over 2D datasets for deeper analysis and provides greater realism than 3D simulated datasets, enhancing urban radio mapping research for UAV applications.
- To estimate urban 3D radio maps, based on Gaussian Process Regression (GPR), a joint model-and-data-driven scheme is proposed. Utilizing prior information provided by simulations, our scheme achieves higher estimation accuracy than the state-of-the-art schemes.
- Selecting proper measurement/training points is essential for map estimation. To this end, we propose both an online MAP-based method and an offline clustering-based method for training point selection, which are parallel approaches and can significantly reduce the measurement cost of 3D radio map recovery.

The rest of the paper is organized as follows. Section II introduces an urban 3D radio signal measurement dataset collected by UAVs in densely-built urban environments. Section III proposes a GPR-based model-and-data-driven 3D radio map estimation scheme for sparse measurements and provide both offline and online methods to guide UAVs in selecting measurement points. In Section IV, results are provided for quantifying the performance of our proposed scheme. Finally, conclusions are drawn in Section V.

II. Urban 3D Radio Signal Measurement Dataset

In this section, we introduce a 3D radio signal measurement dataset for densely-built urban environments, built through the data collected by UAVs. To provide a comprehensive overview, we detail the measurement scenario selection, the UAVs, measurement devices, and the characteristics of the dataset as follows.

Received 30 July 2024; revised 12 December 2024 and 16 March 2025; accepted 9 May 2025. Date of publication 26 May 2025; date of current version 17 October 2025. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62394294, and Grant 62394290 and in part by the National 2011 Collaborative Innovation Center of Wireless Communication Technologies under Grant 2242022k60006. The review of this article was coordinated by Prof. Yuanwei Liu. *(Corresponding author: Xiaofeng Zhong.)*

The authors are with the State Key Laboratory of Space Network and Communications, the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, and also with the Beijing National Research Center for Information Science and Technology (BNRist), Beijing 100084, China (e-mail: cxw22@mails.tsinghua.edu.cn; zhongxf@mail.tsinghua.edu.cn; zhangzj20@mails.tsinghua.edu.cn; daill@mail.tsinghua.edu.cn; zhousd@mail.tsinghua.edu.cn).

Digital Object Identifier 10.1109/TVT.2025.3573595

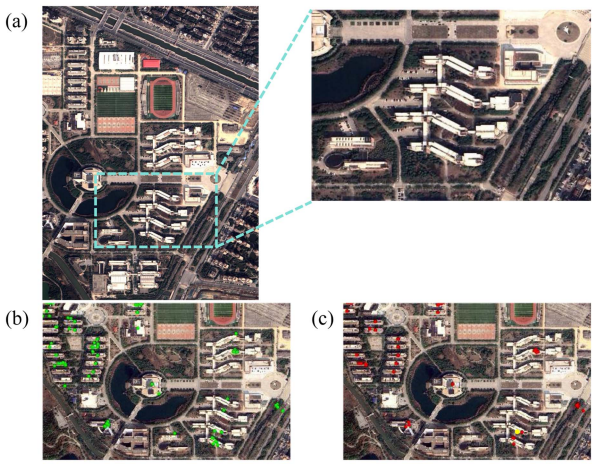

Fig. 1. Satellite images: (a) Overview of the campus; (b) Distribution of 5G base stations; and (c) Distribution of 4G base stations.

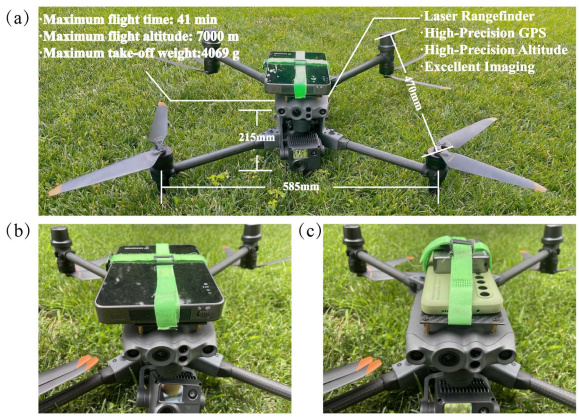

Fig. 2. UAV and measurement devices: (a) UAV equipped with measurement devices; (b) Base station switchable device; and (c) Base station fixed device.

A. Measurement Scenario Selection

Considering the application scenarios of UAVs in logistics field, which often involve densely-built areas, we select the teaching building area of a campus in Nanchang for radio map data collection due to its typical urban characteristics. As illustrated in Fig. 1, the blue dashed rectangle on the left highlights the data collection region, with a magnified view on the right. This region provides a representative urban environment with various obstacles, both man-made (buildings, cars, walls) and natural (trees and vegetation), that affect radio signal propagation. Additionally, as shown in Fig. 1(a), we obtain the latitude and longitude of base stations from the operator. Fig. 1(b) displays the distribution of 5G base stations within the campus marked by green dots, while Fig. 1(c) illustrates the distribution of 4G base stations marked by red dots.

B. UAV and Measurement Devices

1) UAV: As illustrated in Fig. 2(a), the UAV, DJI Matrice 30 RTK, equipped with the measurement devices used in our experiments, is shown along with the relevant parameters.

2) Measurement Devices: The measurement devices used in our experiments, both equipped with Qualcomm modules, are presented in Fig. 2(b) and (c). The black device in Fig. 2(b) can record the UAV's handovers between base stations, tracking connectivity changes as it navigates through the urban environment. The gray device in Fig. 2(c) is set to concentrate on a specific base station and can continuously record RSRP.

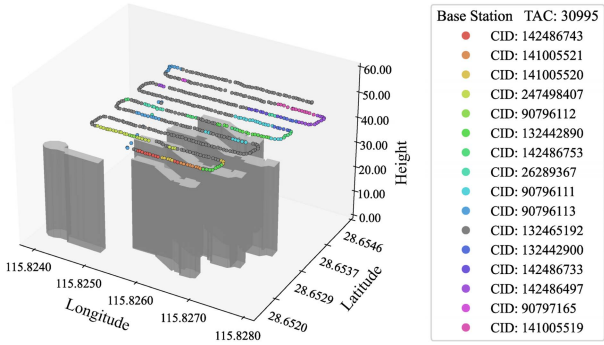

Fig. 3. Base station handovers while UAV's flying over the data collection region.

C. Dataset Composition

In the data collection region, we use the switchable device to simulate a mobile phone, capturing base station handovers and confirming the set of receivable base stations. The results, shown in Fig. 3, demonstrate frequent base station handovers during continuous UAV flights over the region. Using the distribution of the base station distribution provided by the operator, we identify the positions of these receivable base stations and select a 4G base station that indicated by the yellow dot in Fig. 1(c) as the transmitter for measurement. The UAV, equipped with the fixed device, then navigates the region, continuously collecting the RSRP of the base station. The spatial distribution of the RSRP in the 3D radio signal measurement dataset is shown in Fig. 4. The measurement device provides real-time latitude, longitude, and RSRP, while the UAV's log provides altitude information.

To enhance the dataset, we incorporate *simulated RSRP* for a comprehensive comparison with real-world measurements. Using Wireless Insite[1], we simulate RSRP based on 3D building models from Open Street Map[2] in the data collection region. To evaluate the accuracy of the *simulated RSRP*, we calculate the RMSE between the simulated and measured RSRP, obtaining an RMSE of 24.57 dB. The simulation provides a theoretical model of signal propagation, accounting for the physical structures and environmental conditions of our measurements. The parameters for measurements and simulations are summarized in Table I.

III. GPR-Based 3D Radio Map Estimation Scheme

In this section, the GPR-based map estimation scheme is proposed to recover a full 3D radio map from a small number of UAV measurements, i.e., sparse data.

A. Preliminary

We consider the downlink transmission of a cellular network to focus on the coverage of a single transmitter. The RSRP reflects path loss laws within the solid angle visible to the user equipment. Since

[1]https://www.remcom.com/wireless-insite-propagation-software
[2]https://www.openstreetmap.org

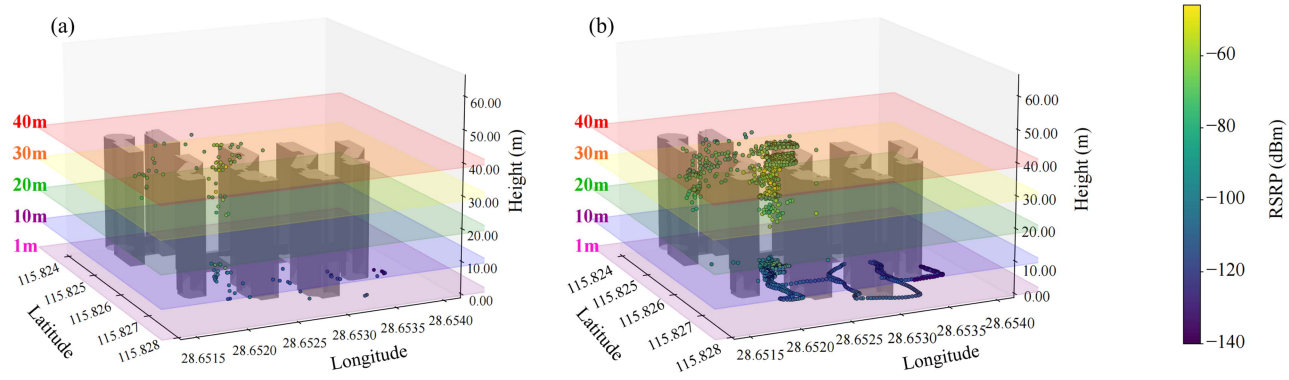

Fig. 4. 3D radio maps obtained by UAV measurements: (a) Measurements for training (10%); (b) Overall measurements (100%).

TABLE I
PARAMETERS OF THE 3D RADIO MAP DATASET

Categories	Parameters	Value
Parameters for Radio Map	Area (length × width)	300 [m] × 280 [m]
	Height span	0 [m] - 50 [m]
	Height range of the buildings	6 [m] - 38.6 [m]
	Tx height	2 [m] above the rooftop
	Center carrier frequency	2.645 [GHz]
Simulation Parameters	Channel bandwidth	20 [MHz]
	Propagation model	Full 3-D
	Ray tracing method	Shooting and bouncing rays
	Waveform	Sinusoid
Measurement Parameters	Measurement time window	3 seconds
	Measurement granularity	Slot-level RSRP
	Data volume	4274

fast-fading effects have been averaged out over time by the receivers during measurements, environmental effects, such as diffraction, reflection, and absorption, are slow-fading effects that contribute to signal variations [18]. For N positions $\mathcal{P}^N := \{\mathbf{p}_1, \ldots, \mathbf{p}_N\}$, where $\mathbf{p}_n := [p_x, p_y, p_z]^{\mathrm{T}} \in \mathbb{R}^3$, the RSRP at position $\mathbf{p}_n$, denoted as $\gamma(\mathbf{p}_n) \in \mathbb{R}$, can be modeled as

$$\gamma(\mathbf{p}_n) = \hat{\gamma}(\mathbf{p}_n) + r(\mathbf{p}_n) + z(\mathbf{p}_n), \tag{1}$$

where: $\hat{\gamma}(\mathbf{p}_n)$ represents the RSRP prediction from ray-tracing (simulated RSRP, in dBm), $r(\mathbf{p}_n)$ is the residual (in dB), which accounts for the prediction errors due to factors not captured by ray-tracing, such as slow fading caused by unmodeled environmental effects, $z(\mathbf{p}_n)$ represents the additive zero-mean noise (in dB) at position $\mathbf{p}_n$.

Our goal is to estimate the RSRP of the 3D radio map for a base station. For the ease of expression, we provide the following definitions:

- *Base station information:* The location, center carrier frequency, and channel bandwidth of the base station.
- *Simulated RSRP:* The dataset comprises M training points $\mathcal{P}^M := \{\mathbf{p}_1, \ldots, \mathbf{p}_M\}$ and $N - M$ testing points $\bar{\mathcal{P}}^{N-M} := \{\mathbf{p}_{M+1}, \ldots, \mathbf{p}_N\}$. The *simulated RSRP* at position $\mathbf{p}_n$ is denoted as $\hat{\gamma}(\mathbf{p}_n)$, where $n \in \{1, \ldots, N\}$ and $N \gg M$.
- *Measured RSRP:* The RSRP $\gamma(\mathbf{p}_m)$ measured at the M training points $\mathcal{P}^M$, where $m \in \{1, \ldots, M\}$.

For example, Fig. 4(a) shows a map of RSRP measurements, where 10% of the measurements are available for training. Fig. 4(b) displays the actual 3D RSRP map obtained from UAV measurements. RSRP, represented by $\hat{\gamma}(\mathbf{p}_n)$ in (1), is derived from ray-tracing using Wireless InSite, with unknown transmission power. Note that, although ray tracing accurately captures path loss, it cannot fully simulate the nuances of real-world slow fading, leading to inaccuracies in $\hat{\gamma}(\mathbf{p}_n)$. We aim to use M RSRP measurements in Fig. 4(a) and $N - M$ *simulated RSRP* values to recover the overall RSRP map in Fig. 4(b). The five planes with different colors planes in subfigures (a) and (b) correspond to heights of 1 m, 10 m, 20 m, 30 m, and 40 m, respectively, and are used to facilitate the 3D visualization of the measurement points.

Thus, our problem becomes estimating the slow-fading components of a 3D RSRP map. They are influenced by environmental details that ray-tracing methods cannot capture and should be determined based on sparse measurement data. Slow-fading effects depend on the local environment's topological structure, such as land use type, nearby obstructions like trees and buildings, and line of sight conditions. Determined by environmental effects, there is a spatial correlation between the RSRP at close-by locations. Therefore, modeling the residual $r(\mathbf{p})$ as a stochastic process offers an elegant means of specifying function properties. The residual $r(\mathbf{p})$ is defined as the difference between the measured RSRP $\gamma(\mathbf{p})$ and the *simulated RSRP* $\hat{\gamma}(\mathbf{p})$, given by

$$r(\mathbf{p}) = \gamma(\mathbf{p}) - \hat{\gamma}(\mathbf{p}). \tag{2}$$

B. GPR-Based 3D Radio Map Estimation Scheme

A Gaussian process (GP), denoted as $\mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$, is a collection of random variables where any finite subset follows a multivariate Gaussian distribution. It is fully determined by its mean function $\mu(\mathbf{x})$, often assumed to be zero, and its covariance kernel $k(\mathbf{x}, \mathbf{x}')$, which dictates the smoothness characteristics of the function it models. In our proposed GPR-based scheme, the residual $r(\mathbf{x})$ accounts for prediction errors due to factors not captured by the ray-tracing models. These residuals exhibit spatial correlation, meaning that the errors depend on the position in the measurement space. Such spatial correlation is a key characteristic of the residuals that makes GPs a natural and effective modeling choice for $r(\mathbf{x})$. Since the close-by locations may exhibit significant differences in signal strength due to diverse building shapes, the input $\mathbf{x}$ includes both the locations and the *simulated RSRP*, i.e., $\mathbf{x} = [p_x, p_y, p_z, \hat{\gamma}(p_x, p_y, p_z)]^{\mathrm{T}}$. This dimension extension ensures that spatially adjacent points with similar environmental characteristics have high correlations, enabling localized GPR to effectively handle non-stationary data.

Consider a prior $\mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$ over $r(\mathbf{x})$. Let $y(\mathbf{x})$ denote the RSRP measurement at points in $\mathcal{X}^M := \{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$, and $\hat{y}(\mathbf{x})$ denote the simulated RSRP at those points. Thus, the residual $r(\mathbf{x})$ is given by

$$r(\mathbf{x}) = y(\mathbf{x}) - \hat{y}(\mathbf{x}). \tag{3}$$

The joint probability distribution of the residuals at measured (training) points $\mathcal{X}^M := \{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$ and unmeasured (testing) points

$\mathcal{X}^{N-M} := \{\mathbf{x}_{M+1}, \ldots, \mathbf{x}_N\}$ is

$$\begin{bmatrix} r(\mathbf{x}_*) \\ \boldsymbol{r} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu(\mathbf{x}_*) \\ \boldsymbol{\mu} \end{bmatrix}, \begin{bmatrix} \Sigma(\mathbf{x}_*) & \mathbf{k}(\mathbf{x}_*)^{\mathrm{T}} \\ \mathbf{k}(\mathbf{x}_*) & \mathbf{K} \end{bmatrix} \right), \tag{4}$$

where $\mathbf{x}_* \in \mathcal{X}^{N-M}$, $\mu(\mathbf{x}_*)$ denotes the priori mean of the residual at $\mathbf{x}_*$, $\mathbf{r} := [r(\mathbf{x}_1), \ldots, r(\mathbf{x}_M)]^{\mathrm{T}}$, $\mathbf{k}(\mathbf{x}_*) := [k(\mathbf{x}_1, \mathbf{x}_*), \ldots, k(\mathbf{x}_M, \mathbf{x}_*)]^{\mathrm{T}}$, $\boldsymbol{\mu} := [\mu(\mathbf{x}_1), \ldots, \mu(\mathbf{x}_M)]^{\mathrm{T}}$, and the (i, j)-th entry of $\mathbf{K} \in \mathbb{C}^{M \times M}$ is $k(\mathbf{x}_i, \mathbf{x}_j)$.

Thus, the posterior over $r(\mathbf{x}_*)$ is also a Gaussian process, with mean and covariance can be derived from (4) as

$$\mu_*(\mathbf{x}_*) = \mu(\mathbf{x}_*) + \mathbf{k}(\mathbf{x}_*)^{\mathrm{T}} \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}), \tag{5}$$

$$\Sigma(\mathbf{x}_*) = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}(\mathbf{x}_*)^{\mathrm{T}} \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}_*). \tag{6}$$

With sufficient training points, $\mu_*(\mathbf{x}_*)$ can be viewed as the Bayesian estimator of $r(\mathbf{x}_*)$, which can achieve our goal.

In GPR, the wide-sense stationarity of data significantly impacts model performance. Non-stationary data require a covariance function that accounts for varying statistical properties, which increases model complexity and poses challenges for parameter estimation. These factors hinder GPR's generalization and prediction accuracy. To mitigate these issues, we employ a composite kernel function. This function combines the Constant Kernel (k_{const}), Matérn Kernel ($k_{\text{Mat\érn}}$), and White Noise Kernel (k_{WN}), and is expressed as follows:

$$k_{\text{com}}(\mathbf{x}, \mathbf{x}') = k_{\text{const}}(\mathbf{x}, \mathbf{x}') \times k_{\text{Mat\érn}}(\mathbf{x}, \mathbf{x}') + k_{\text{WN}}(\mathbf{x}, \mathbf{x}'), \tag{7}$$

as detailed in [19, Eq. (4.14)].

Before employing Gaussian processes for 3D radio map estimation, initial global RSRP predictions are simulated based on building structures. This provides a foundational estimate of the radio environment. As measurement points are strategically selected and their actual RSRPs are obtained, the Gaussian process kernel is updated, incrementally refining the map's accuracy. Proper measurement point selection is crucial as it directly impacts the model's accuracy, representativeness, resource efficiency, and robustness. We introduce two methods for selecting measurement points: an online method that updates strategies in real-time based on feedback data, and an offline method that relies solely on pre-existing knowledge without real-time feedback.

1) Online Point Selection Method: In GPR, the posterior variance at each point quantifies the uncertainty of the model's predictions. High posterior variance indicates greater uncertainty and more potential information. Thus, selecting the next measurement point associated with the MAP variance can maximize the information acquisition. To this end, the candidate point $\mathbf{x}_{t+1}$ can be chosen according to

$$\mathbf{x}_{t+1} = \underset{\mathbf{x} \in \mathcal{X}^N / \mathcal{X}^t}{\operatorname{argmax}} \Sigma^t(\mathbf{x}), \tag{8}$$

where $\mathcal{X}^N$ represents all potential measurement points, $\mathcal{X}^t$ represents measured point, and $\Sigma^t(\mathbf{x})$ is the posterior covariance derived from the measured points $\mathcal{X}^t$ by (6). The online point selection algorithm is detailed in Algorithm 1. In practice, UAVs implement this online MAP-based approach by continuously updating the posterior covariance as new measurement points are collected. At each step, the UAV recalculates the kernel function to refine the covariance estimates and selects the point with the highest posterior variance, maximizing information gain. The UAV then navigates to the selected point, measures the actual RSRP and iteratively updates the map.

2) Offline Point Selection Method: Although the MAP method is effective, it is constrained by its requirement for real-time computation and its exclusive focus on GPR. Given the spatial correlation of environmental factors in 3D radio maps, we propose an offline KMeans-based method to select a series of measurement points simultaneously. By clustering the points by $\mathbf{x} = [p_x, p_y, p_z, \hat{\gamma}]$, and employing cluster centroids as measurement points, we maximize sampling diversity and enhance the efficiency and applicability of 3D radio map estimation. The specific algorithm is detailed in Algorithm 2. In this approach, UAVs use pre-simulated RSRP values to select measurement points based on clustering. By applying KMeans clustering to the candidate set, the UAV identifies a fixed number of optimal points as cluster centroids. Since this method doesn't rely on real-time feedback, all measurement points can be pre-determined before the UAV's flight. This allows for efficient route planning and batch data collection without needing a specific order, enabling a streamlined and resource-efficient mapping process.

Algorithm 1: Online Point Selection Using GPR with MAP.

Input: The candidate set $\mathcal{X}^N := \{\mathbf{x}_n\}_{n=1}^N$, the target number M for the training set.

1: Randomly select $\mathbf{x}_1$ from $\mathcal{X}^N$ and initialize $\mathcal{X}_{\text{train}} := \{\mathbf{x}_1\}$.
2: Initialize the remaining candidate set $\mathcal{X}_{\text{cand}} := \mathcal{X}^N / \mathcal{X}_{\text{train}}$.
3: **while** $t \in \{1, \ldots, M-1\}$ **do**
4: Using the training set $\mathcal{X}_{\text{train}}$, calculate the posterior covariance $\Sigma^t(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}_{\text{cand}}$ via (6).
5: Select the candidate point: $\mathbf{x}_{t+1} = \operatorname{argmax}_{\mathbf{x} \in \mathcal{X}_{\text{cand}}} \Sigma^t(\mathbf{x})$.
6: Update the training set: $\mathcal{X}_{\text{train}} \leftarrow \mathcal{X}_{\text{train}} \cup \{\mathbf{x}^{t+1}\}$.
7: Update the candidate set: $\mathcal{X}_{\text{cand}} \leftarrow \mathcal{X}_{\text{cand}} / \{\mathbf{x}^{t+1}\}$.
8: **end while**
9: Obtain training set $\mathcal{X}^M := \mathcal{X}_{\text{train}}$.

Output: The designed training set $\mathcal{X}^M := \{\mathbf{x}_i\}_{i=1}^M$.

Algorithm 2: Offline Point Selection Using KMeans Clustering.

Input: The candidate set $\mathcal{X}^N := \{\mathbf{x}_n\}_{n=1}^N$, the target number M for the training set.

1: Initialize KMeans with M clusters and fit it to candidate set $\mathcal{X}^N$.
2: Assign all candidate points to M clusters via KMeans.
3: **for** each cluster $l \in \{1, 2, \ldots, M\}$ **do**
4: Get the centroid of the l-th clusters by $\mathbf{c}_l = (\sum_{\mathbf{x} \in \mathcal{X}_l} \mathbf{x}) / |\mathcal{X}_l|$ where $\mathcal{X}_l$ is the set of candidate points in cluster l.
5: Find the candidate point: $\mathbf{x}_{n_l} = \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}_l} \|\mathbf{x} - \mathbf{c}_l\|$.
6: **end for**
7: Obtain training set $\mathcal{X}^M = \{\mathbf{x}_{n_1}, \mathbf{x}_{n_2}, \ldots, \mathbf{x}_{n_M}\}$.

Output: The designed training set $\mathcal{X}^M := \{\mathbf{x}_i\}_{i=1}^M$.

IV. Performance Evaluation

In this section, we evaluate the performance of the proposed GPR-based scheme on our built dataset.

We first evaluate the effect of various kernel combinations on modeling environmental factors and slow-fading effects influenced by local topological structures. Radial Basis Function (k_{RBF}), Matérn ($k_{\text{Mat\érn}}$), and Rational Quadratic (k_{RQ}) kernels are used to address smoothness, roughness, and multiple variation scales. Additionally, Constant (k_{const}) and White Noise (k_{WN}) kernels are included to account for constant bias and inherent measurement noise, enhancing the robustness and accuracy of the GPR model. As shown in Table II, the best performance is achieved with a combination of k_{const}, $k_{\text{Mat\érn}}$, and k_{WN} due to the Matérn kernel's tunable smoothness parameter.

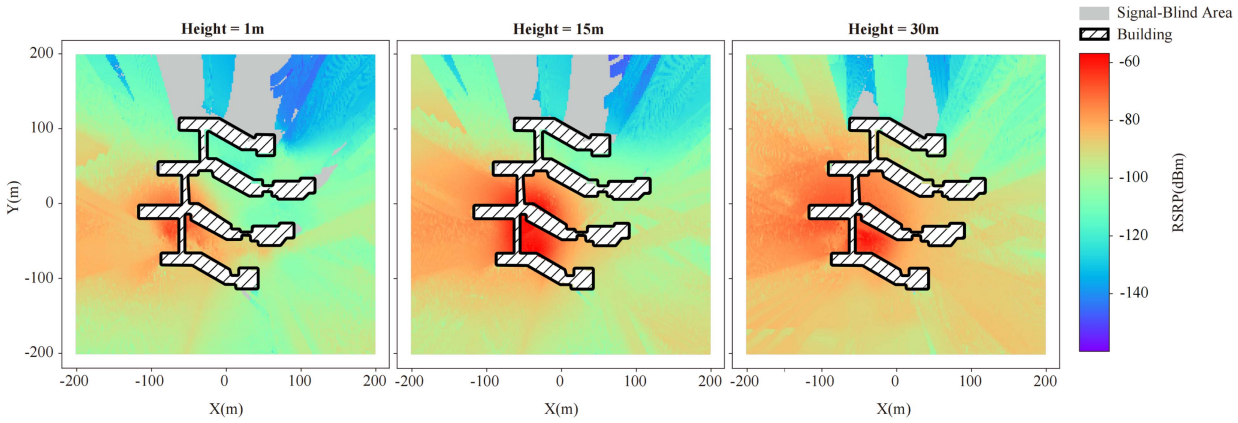

Fig. 5. 3D radio maps obtained by our proposed estimation method at 1 m, 15 m, and 30 m height.

TABLE II
COMPARISON OF KERNEL FUNCTION PERFORMANCE

Kernel Type	Kernel Selection	RMSE (dB)
Radial Basis Function	$k_{\text{RBF}} + k_{\text{WN}}$	23.9106
	k_{RBF}	17.5451
	$k_{\text{const}} \times k_{\text{RBF}}$	14.4392
	$k_{\text{const}} \times k_{\text{RBF}} + k_{\text{WN}}$	8.2593
Rational Quadratic	$k_{\text{RQ}} + k_{\text{WN}}$	23.9106
	k_{RQ}	10.3063
	$k_{\text{const}} \times k_{\text{RQ}}$	9.1089
	$k_{\text{const}} \times k_{\text{RQ}} + k_{\text{WN}}$	8.0131
Matérn	$k_{\text{Matérn}} + k_{\text{WN}}$	23.9107
	$k_{\text{Matérn}}$	8.8130
	$k_{\text{const}} \times k_{\text{Matérn}}$	8.0265
	$\mathbf{k_{const} \times k_{Matérn} + k_{WN}}$	**7.9286**

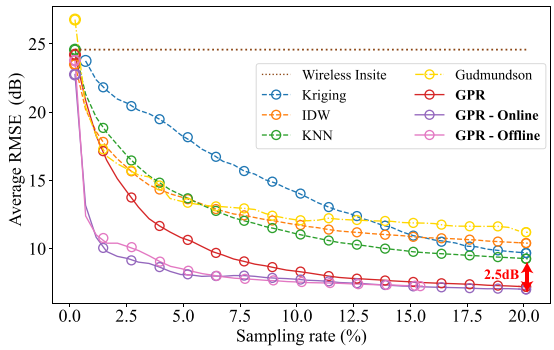

Fig. 6. Comparison of average RMSE as a function of sampling rate for different radio map estimation schemes. GPR schemes based on random point selection, online method, and offline method are compared.

Thus, in the following simulations, we use this setup to enable our proposed scheme.

The following four baselines are considered:

- *Inverse Distance Weighting (IDW):* A local neighborhood method that interpolates spatial points based on their influence within a certain distance [20].
- *K-Nearest Neighbors (KNN):* A classical machine learning algorithm for intensity prediction. It interpolates unknown values by identifying the k nearest training samples [18].
- *Linear Ordinary Kriging:* A geostatistical technique that interpolates unknown signal power based on spatial distances and correlation modeled through a variogram [21]. For 3D inputs, a linear semivariogram model is employed. For 4D inputs, our hybrid approach first clusters the data into 3 groups based on the fourth dimension (*simulated RSRP*) and then applies 3D Ordinary Kriging independently within each cluster to predict the residual.
- *Exponential Ordinary Kriging (Gudmundson):* An ordinary Kriging estimator based on the well-known Gudmundson Shadowing Model corresponding to the exponential model. The algorithmic input is the spatial coordinates and the output is the predicted RSRP [17].

In Fig. 5, we present the 3D radio maps estimated by our approach at heights of 1 m, 15 m, and 30 m, demonstrating the spatial distribution of the radio signal estimation at different altitudes. Following this, Fig. 6 illustrates the average RMSE for different estimation schemes. Specifically, Ordinary Kriging, IDW, and KNN schemes employ random point selection with *simulated RSRP* included for comparison. The brown dashed line represents the simulation results provided by Wireless Insite, which does not utilize residuals. The Gudmundson Shadowing Model (Exponential Ordinary Kriging) employs an online MAP-based point selection method and directly uses the spatial coordinates to predict the RSRP, as described in [17]. In particular, GPR employs the online MAP-based point selection method (purple), the offline clustering-based (pink) point selection method, and random selection (red) for simulations, respectively. One can find that our GPR-based scheme achieves the lowest RMSE across all sampling rates. Notably, the mathematical essence of Kriging and GPR is indeed aligned, as both methods rely on similar covariance structures for spatial interpolation; however, their performance differences stem from distinct design choices in input features and kernel/semivariogram formulations. Particularly, the RMSEs of the proposed schemes can be at least 2.5 dB lower than those of other baseline schemes, including Kriging, KNN, IDW, and Gudmundson, with this performance observed at a 20% sampling rate, where the best baseline (KNN) is compared to our method. Remarkably, both the offline and online methods achieve an RMSE of 10 dB with only 2% of the sampling points, while other schemes require over 14% of the sampling points to achieve the same performance. Additionally, the offline method shows performance comparable to the best-performing online method, significantly surpassing random selection. This phenomenon highlights that the clustering-based method enables batch point selection using pre-existing knowledge and data, incurring minimal performance loss compared to the MAP-based method.

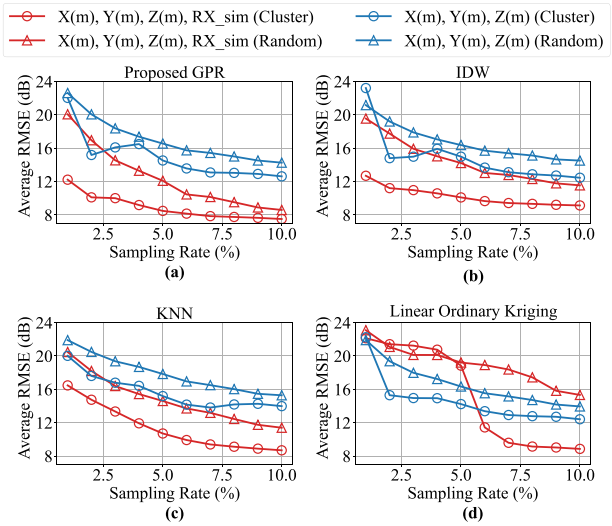

Fig. 7. Comparisons of average RMSE as a function of sampling rates using different estimation schemes, input features, and point selection methods.

We conduct additional experiments to show the generality of our two proposed designs: The incorporation of *simulated RSRP* alongside spatial coordinates, and the use of clustering-based point selection, across various schemes. As shown in Fig. 7, the inclusion of *simulated RSRP* consistently enhances estimation performance across all four schemes, with RMSE reductions of up to 10 dB. The clustering-based method shows a marked improvement in performance, which reduces RMSE by up to 8 dB compared to the random selection method. These findings underscore the generality of these two designs in integrating with other 3D radio map estimation schemes.

V. Conclusion

In this paper, we develop a 3D UAV-collected dataset in urban scenarios and propose a GPR-based platform for 3D radio map estimation. Posterior variance calculation is incorporated into the GPR framework for an online MAP point selection method. This integration enables more efficient and accurate estimations with sparse measurements by prioritizing points with higher uncertainty, resulting in the best performance observed. Additionally, we propose an offline clustering-based point selection method to ensure representative training points, which performs as well as its online counterparts. This efficient platform is crucial for logistics UAVs, as accurate 3D radio maps have become essential for reliable navigation and communication in densely-built urban environments. We also prove the generality of our designs, thereby providing a framework for future works on 3D radio map estimation.

References

[1] S. Zhang, B. Choi, F. Ouyang, and Z. Ding, "Physics-inspired machine learning for radiomap estimation: Integration of radio propagation models and artificial intelligence," *IEEE Commun. Mag.*, vol. 62, no. 8, pp. 155–161, Aug. 2024.

[2] T. Hu, Y. Huang, J. Chen, Q. Wu, and Z. Gong, "3D radio map reconstruction based on generative adversarial networks under constrained aircraft trajectories," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8250–8255, Jun. 2023.

[3] D. Romero and S.-J. Kim, "Radio map estimation: A data-driven approach to spectrum cartography," *IEEE Signal Process. Mag.*, vol. 39, no. 6, pp. 53–72, Nov. 2022.

[4] A. Ivanov, B. Muhammad, K. Tonchev, A. Mihovska, and V. Poulkov, "UAV-based volumetric measurements toward radio environment map construction and analysis," *Sensors*, vol. 22, no. 24, Dec. 2022, Art. no. 9705.

[5] S. Horsmanheimo et al., "5G communication QoS measurements for smart city UAV services," in *Proc. 16th Eur. Conf. Antennas Propag.* (EuCAP), Apr. 2022, pp. 1–5.

[6] C. Yapar, R. Levie, G. Kutyniok, and G. Caire, "Dataset of pathloss and ToA radio maps with localization application," Nov. 2022. [Online]. Available: https://dx.doi.org/10.21227/0gtx-6v30

[7] N. Leonor et al., "Site-specific radio wave measurements for 5G macrocell coverage at 750 MHz, 2.5 GHz and 3.5 GHz signal frequencies," 2021. [Online]. Available: https://dx.doi.org/10.21227/wmks-4475

[8] M. Hoffmann, P. Kryszkiewicz, and A. Kliks, "Frequency selection for platoon communications in secondary spectrum using radio environment maps," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2637–2650, Mar. 2022.

[9] J. Wang, Y. Hao, Z. Wu, Y. Shi, and C. Yang, "A broadcast map constructing method based on the LSTM and assimilation theory," *IEEE Trans. Broadcast.*, vol. 70, no. 3, pp. 924–934, Sep. 2024.

[10] J. Wang, Z. Wu, Y. Hao, and C. Yang, "Broadcasting maps construction method based on particle swarm optimization assisted support vector machines integrated model," *IEEE Trans. Antennas Propag.*, vol. 72, no. 8, pp. 6638–6651, Aug. 2024.

[11] Y.-Q. Xu, B. Zhang, G. Ding, B. Zhao, S. Li, and D. Guo, "Radio environment map construction based on spatial statistics and Bayesian hierarchical model," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 3, pp. 767–779, Sep. 2021.

[12] S. Shrestha, X. Fu, and M. Hong, "Deep spectrum cartography: Completing radio map tensors using learned neural models," *IEEE Trans. Signal Process.*, vol. 70, pp. 1170–1184, Jun. 2022.

[13] K. Sato, K. Suto, K. Inage, K. Adachi, and T. Fujii, "Space-frequency-interpolated radio map," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 714–725, Jan. 2021.

[14] Y. Hu and R. Zhang, "A spatiotemporal approach for secure crowdsourced radio environment map construction," *IEEE/ACM Trans. Netw.*, vol. 28, no. 4, pp. 1790–1803, Aug. 2020.

[15] J. Wang et al., "Sparse Bayesian learning-based 3D radio environment map construction—sampling optimization, scenario-dependent dictionary construction and sparse recovery," *IEEE Trans. Cognit. Commun. Netw.*, vol. 10, no. 1, pp. 80–93, Feb. 2024.

[16] J. Thrane, D. Zibar, and H. L. Christiansen, "Comparison of empirical and ray-tracing models for mobile communication systems at 2.6 GHz," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC)*, Sep. 2019, pp. 1–5.

[17] R. Shrestha, D. Romero, and S. P. Chepuri, "Spectrum surveying: Active radio map estimation with autonomous UAVs," *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 627–641, Jan. 2022.

[18] F. Sohrabi and E. Kuehn, "Construction of the RSRP map using sparse MDT measurements by regression clustering," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[19] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006, vol. 38.

[20] D. Denkovski, V. Atanasovski, L. Gavrilovska, J. Riihijärvi, and P. Mähönen, "Reliability of a radio environment map: Case of spatial interpolation techniques," in *Proc. 7th Int. ICST Conf. Cong. Radio Oriented Wireless Netw. Commun. (CROWNCOM)*, Jun. 2012, pp. 248–253.

[21] A. Ivanov, K. Tonchev, V. Poulkov, A. Manolova, and A. Vlahov, "Limited sampling spatial interpolation evaluation for 3D radio environment mapping," *Sensors*, vol. 23, no. 22, Nov. 2023, Art. no. 9110.