FOMRADIO: NONSTATIONARY FOURIER OPERATOR FLOW MATCHING FOR RADIO MAP GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Generating radio maps from scene geometry and transmitter configurations requires combining long-range propagation with local, scene-dependent interactions. We propose `FOMRadio`, a conditional flow-matching model with a geometry-conditioned nonstationary Fourier operator. Its velocity network couples distant locations spectrally, while an amplitude–phase gate and a residual spatial branch adapt the retained modes and recover structures beyond them. A condition encoder integrates building heights, analytical physical parameters, and transmitter configurations, making the effective kernel. `FOMRadio` transports a multiscale Gaussian random-field source along a straight conditional path, separating spatially structured uncertainty from environment-dependent transport without access to the target radio map at inference. Experiments on RadioMapSeer and U6G demonstrate improved pixelwise accuracy across static and dynamic scenario and variations in cross-configurations, cross-frequency bands, and beam direction.

1 INTRODUCTION

Channel knowledge maps (CKMs) associate locations and transmission configurations with electromagnetic propagation states, providing a physics-grounded information layer for wireless digital twins and environment-aware tasks such as coverage prediction and beam management (Zeng & Xu, 2021; Zeng et al., 2024). Radio maps instantiate CKMs as received-power fields whose broad coverage, sharp shadows, and multipath structures arise from interactions between transmitted beams and the environment (Zeng et al., 2024; Li et al., 2026). Because densely acquiring these fields through measurements or repeated ray tracing is costly, learned radio-map generation has emerged as a practical alternative (Levie et al., 2021; Li et al., 2026).

Learned radio-map methods have progressed from deterministic reconstruction to variational, adversarial, diffusion, and flow-matching models that capture distributions beyond a single prediction (Levie et al., 2021; Li et al., 2022; Zhang et al., 2023; Wang et al., 2025; Huang et al., 2026; Fu et al., 2026). However, most parameterize both conditions and outputs on fixed grids, coupling spatial correlations and geometry-dependent interactions to a prescribed raster. Because changes in scene or beam reshape both global coverage and local multipath structure, grid-level accuracy alone does not establish consistent behavior across discretizations.

Neural operators learn mappings between functions rather than finite-dimensional vectors. FNO provides global coupling through truncated Fourier modes and has demonstrated transfer across PDE discretizations (Li et al., 2021). However, its spatially shared spectral weights favor broad structure over layout-specific boundaries and multipath details (Liu & Tang, 2025; Murshed et al., 2026). Radio maps are nonstationary, because the same displacement may traverse free space an obstruction in another. Thus, global spectral coupling must adapt to local propagation geometry.

In contrast to treating radio maps only as fixed-grid features, we take a function-to-function view of their conditional generation. Flow Matching models transport between spatial fields with a velocity parameterized by a geometry-conditioned multiscale Fourier operator (Li et al., 2021; Lipman et al., 2023). This formulation raises two challenges:

(i) A spatially structured source. Nonstationary propagation creates a radio-map structure at several spatial ranges, whereas independent pixelwise noise has no spatial correlation. A function-valued

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

source must supply coherent stochastic variation that the conditional flow can reshape for each scene and beam.

(ii) Geometry-dependent transport. The velocity must connect distant coverage regions while retaining shadow boundaries and fine reflection- and diffraction-related patterns. Truncated Fourier modes provide the former, but their shared spectral response alone does not follow the location-dependent propagation geometry.

To address the first challenge, `FOMRadio` starts from a reference-centered multiscale Matérn Gaussian field, combining correlated perturbations at several lengths (Rasmussen & Williams, 2006). For the second, a condition encoder integrates scene geometry and analytical beam maps, and an amplitude–phase gate adapts retained Fourier modes to the evolving field and condition. Multiscale spatial features and detail-aware velocity supervision provide paths for structure beyond those modes. A straight conditional flow-matching path then transports the source to the target field (Lipman et al., 2023). The target-array radio map is not required at inference.

The main contributions are as follows:

- **Function-to-function generation.** We cast target-array CKM generation as conditional flow matching between spatial fields and use a multiscale Fourier operator to parameterize the transport velocity.
- **Correlated source and adaptive transport.** We pair reference-centered multiscale Matérn perturbations with condition-dependent amplitude–phase gating of retained Fourier modes, separating stochastic spatial variation from environment-dependent transport.
- **U6G evaluation.** We study target-array generation on the U6G benchmark (Li et al., 2026), using a fixed 8×8-to-32×32 array setting at 6.7 GHz across scenes and target beam directions.

2 Related Work

2.1 Generative Modeling for Channel Knowledge Maps

CKMs associate location and transmission configurations with channel conditions, with radio maps representing received-power fields (Zeng & Xu, 2021; Zeng et al., 2024). The generation of learned CKM has progressed from kernel function and probabilistic reconstruction (Levie et al., 2021; Li et al., 2022) to adversarial models using conditional generation, multiscale context, and attention (Zhang et al., 2023; Chen et al., 2024). Diffusion methods further model distributions over complete fields through scene- and transmitter-conditioned denoising (Wang et al., 2025; Luo et al., 2025), with subsequent extensions incorporating physics guidance, multipath structure, and volumetric generation (Jia et al., 2025a; Wang et al., 2026b;a). Flow matching instead transports a source distribution to a CKM field (Lipman et al., 2023), supporting uncertainty-aware measurement, conditional generation, observation-guided transport, and one-step reconstruction (Sun et al., 2025; Jia et al., 2025c; Huang et al., 2026; Fu et al., 2026). Although these approaches improve spatial expressiveness or sampling efficiency, their reported implementations remain tied to prescribed raster grids and do not jointly provide a spatially correlated source and a transport operator adaptive to scene geometry and the target beam. This gap motivates function-to-function transport that separates structured spatial variation from geometry-dependent field evolution.

2.2 Neural Operators for Channel Knowledge Maps

CKM fields combine broad coverage with sharp shadow boundaries and localized multipath structure, requiring both global coupling and spatial adaptation. Neural operators provide the former: FNO couples distant locations through truncated Fourier modes and supports transfer across PDE discretizations, while U-shaped operators introduce multiscale feature paths (Li et al., 2021; Rahman et al., 2023). Vision-oriented extensions further combine global spectral mixing with adaptive frequency weighting or spatial attention to recover fine details (Guibas et al., 2022; Liu & Tang, 2025); however, CKM details must be conditioned on building geometry and the target beam rather than generic image statistics. Wireless neural operators begin to incorporate such propagation context: Anchor-CKM performs layout-conditioned Fourier refinement of sparse-pilot fields, and PU-HNO

progressively models reflection, diffraction, and scattering from ray-tracing input (Jiu et al., 2026; Murshed et al., 2026). These methods primarily address measurement-supported reconstruction or simulator refinement. Conditional CKM generation instead requires a velocity operator that jointly adapts global transport and fine spatial responses to scene geometry and target excitation.

3 Preliminaries and Problem Formulation

Let $\Omega \subset \mathbb{R}^2$ denote the spatial domain of a channel knowledge map (CKM). We represent the scene geometry by E, the target transmitter configuration by ξ, and its analytical beam field by B_ξ. When an auxiliary configuration is available, we denote its power field, validity mask, beam field, and configuration by $a = (R_\mathrm{a}, M_\mathrm{a}, B_\mathrm{a}, \xi_\mathrm{a})$. The task is to model the conditional distribution of the target radio-power field $Y : \Omega \to \mathbb{R}$. On a grid $\mathcal{G}_h$, the sampling operator P_h converts each spatial field into its raster representation.

Conditional flow matching for CKMs. Flow matching transports a source field $X_0 \sim p_0(\cdot \mid c)$ to the target distribution through a time-dependent velocity field (Lipman et al., 2023). For a straight path, the intermediate field, target velocity, and training objective are

$$\mathcal{L}_{\mathrm{FM}}(\theta) = \mathbb{E}_{c,X_0,Y,t}\left[\|v_\theta(X_t, t; c) - U_t\|^2_{L^2(\Omega)}\right], \qquad t \sim \mathcal{U}[0,1]. \tag{1}$$

At inference, the learned velocity is integrated from $t = 0$ to $t = 1$ to transform a source sample into a target CKM.

From grid-wise velocities to field transport. Existing flow-matching-based CKM methods mainly study environment-guided radio dynamic generation (Jia et al., 2025c). They define velocity field as power-to-power network on a constant feature grid. Such models can generalize to unseen scenes through conditioning, but they limit to consistency across spatial discretizations. Their receptive fields Moreover, changes in geometry, frequency, array, or beam reshape both long-range coverage and localized blockage, reflection, and diffraction patterns, making a single grid-specific velocity difficult to transfer across scenes and transmitter specifications.

These limitations raise two questions. First, *how can we construct a spatially coherent source distribution that remains consistent under grid refinement and provides multiscale variations relevant to radio propagation?* Second, *how can we learn a condition-dependent velocity operator that acts on spatial fields, rather than a grid-specific velocity map, while coupling long-range transport with geometry-dependent local interactions?* Ideally, the velocity should be defined by a common field operator $\mathcal{V}_\theta$ whose discretizations satisfy

$$v_{\theta,h}(P_h X_t, t; P_h c) \approx P_h \mathcal{V}_\theta(X_t, t; c), \qquad \mathcal{V}_\theta : \mathcal{X} \times [0,1] \times \mathcal{C} \to \mathcal{X}. \tag{2}$$

The following section addresses these questions using a spatially structured source and a condition-adaptive neural-operator velocity.

4 Methodology

Figure 1 shows an overview of `FOMRadio`. The sampled radio parameters define a functional source which is transported to radio maps by a nonstationary FNO-based U-Net. An amplitude-phase-gated spectral branch captures global interactions, while a residual path preserves the local structure.

4.1 Multiscale Matérn Noise Source

We construct the initial field by perturbing the available auxiliary received-power field with a correlated Gaussian random field. For spatial separation $r = \|p - p'\|$, let $C_\ell(r) = (r/\ell)K_1(r/\ell)$ for $r > 0$, with $C_\ell(0) = 1$, where K_1 is the modified Bessel function of the second kind. This is a unit-variance Matérn covariance with smoothness $\nu = 1$ (Rasmussen & Williams, 2006) as

$$\begin{aligned} C_{\mathrm{MS}}(p, p') &= \frac{1}{3} \sum_{\ell \in \{2,8,32\}} C_\ell(\|p - p'\|), \\ z_{\mathrm{MS}} \sim \mathcal{GP}(0, C_{\mathrm{MS}}), &\qquad x_0 = M \odot R + z_{\mathrm{MS}}. \end{aligned} \tag{3}$$

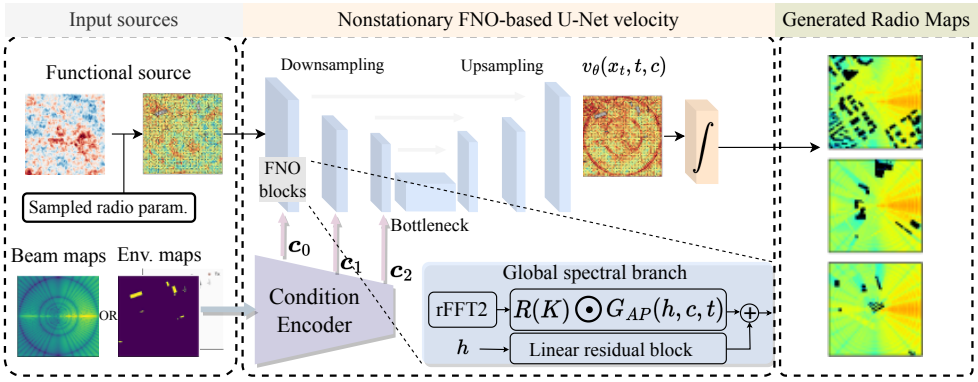


Figure 1: Overview of the proposed `FOMRadio`. Sampled radio parameters define a functional source, while beam and environment maps are encoded into multiscale conditions $\{c_0, c_1, c_2\}$. A nonstationary FNO-based U-Net predicts the velocity that transports the evolving field. Each FNO block combines a global spectral transformation, modulated by the amplitude-phase gate $G_{\mathrm{AP}}(h, c, t)$, with a linear residual path.

The length scales are measured in pixels on the training grid. We sample z_{MS} by positive-semidefinite-checked circulant embedding without per-sample normalization. The mixture provides the multi-scale variations, while $M \odot R$ anchors the source to the available field. Scene-dependent spatial responses are learned by the conditional velocity.

4.2 AP-Gated U-Shaped Fourier Velocity Network

We parameterize $v_\psi(x_t, t, c)$ with a five-scale U-shaped Fourier neural operator (Li et al. 2021; Rahman et al. 2023). Four encoder stages, a bottleneck, and four decoder stages combine multi-scale condition features with the evolving field. Each spectral stage has a truncated Fourier branch, a pointwise branch, a time embedding, and a residual update. Condition attention and skip connections carry spatial features across scales. The Fourier branch couples distant locations through shared modes and models broad coverage structure. The spatial paths retain geometric cues around buildings and shadow boundaries that the truncated branch cannot represent directly.

At stage ℓ, let h_ℓ be the latent field and $q_\ell(c)$ its condition feature. For each retained spatial mode k, the complex spectral weight $W_\ell(k)$ acts on $\widehat{h}_\ell(k)$ before an amplitude–phase (AP) gate is applied:

$$\begin{aligned} \widehat{s}_\ell(k) &= g_\ell(k; h_\ell, q_\ell, t) W_\ell(k) \widehat{h}_\ell(k), \qquad k \in \mathcal{K}_\ell, \\ g_\ell(k) &= \left[1 + \tfrac{1}{2} \tanh A_\ell(k)\right] \exp\left\{ \mathrm{i} \tfrac{\pi}{4} \tanh P_\ell(k) \right\}. \end{aligned} \tag{4}$$

The gate predicts A_ℓ and P_ℓ from normalized state and condition spectral magnitudes, their relative spatial-spectral phase, frequency coordinates, and flow time. Amplitude adjusts the contribution of each retained mode, while phase changes its spatial alignment. Both adjustments are bounded, and the zero and Nyquist frequency columns keep unit gates. The AP gate adapts the global spectral response to the current field and scene, while condition features and skip paths provide the local information needed for fine spatial structure.

4.3 Flow-Matching Loss

For a target field Y, we sample $t \sim \mathcal{U}[0, 1]$ and use the straight conditional flow-matching path (Lipman et al. 2023),

$$x_t = (1 - t)x_0 + tY, \qquad u_t = Y - x_0. \tag{5}$$

With velocity error $e = v_\psi(x_t, t, c) - u_t$, the primary objective averages squared error over valid target pixels:

$$\mathcal{L}_{\mathrm{FM}} = \frac{\sum_{p \in \Omega_h} V(p) e(p)^2}{\sum_{p \in \Omega_h} V(p)}, \qquad \mathcal{L} = \mathcal{L}_{\mathrm{FM}} + 0.05 \mathcal{L}_{\mathrm{Haar},1} + 0.05 \mathcal{L}_{\mathrm{Haar},2}. \tag{6}$$

For each Haar term, we first set invalid entries of the *velocity error* to zero and apply 2×2 average pooling. We then average the squared coefficients in the three 2D detail bands at level $j \in \{1, 2\}$, propagating the low-pass band. These terms supervise spatial changes at two scales. At inference, we draw x_0 from (3) and integrate $\dot{x}_t = v_\psi(x_t, t, c)$ from $t = 0$ to 1 with an eight-step Euler solver.

Table 1: RadioMapSeer SRM and DRM results. Baseline scores follow Jia et al. (2025c). NMSE and RMSE are lower-is-better; PSNR and SSIM are higher-is-better.

	SRM				DRM			
Method	NMSE	PSNR	RMSE	SSIM	NMSE	PSNR	RMSE	SSIM
RME-GAN	0.0115	30.54	0.0303	0.9323	0.0118	30.40	0.0307	0.9219
RadioUNet	0.0074	32.01	0.0244	0.9592	0.0089	31.75	0.0258	0.9410
UVM-Net	0.0085	30.34	0.0304	0.9320	0.0088	30.42	0.0301	0.9326
RadioDiff	0.0049	35.13	0.0190	0.9691	0.0057	34.92	0.0215	0.9536
RMDM	0.0031	34.47	0.0125	**0.9780**	0.0047	34.21	0.0146	**0.9680**
RadioFlow Lite	0.0077	34.92	0.0182	0.8897	0.0079	35.08	0.0178	0.8848
RadioFlow Large	0.0023	39.83	0.0103	0.9249	**0.0028**	39.37	0.0108	0.9236
`FOMRadio`	**0.0019**	**39.91**	**0.0101**	0.9249	0.0031	**39.41**	**0.0106**	0.9236

5 Experiments

We ask three questions: **RQ1.** How accurately does `FOMRadio` reconstruct static and vehicle-perturbed radio maps? **RQ2.** Can it preserve fine spatial structure when the target antenna array changes? **RQ3.** How does its response vary with target frequency and beam direction? More details of datasets is in Appendix A.

5.1 Experimental Setup

Tasks. We evaluate static radio maps (SRM) and dynamic radio maps (DRM) from RadioMapSeer (Yapar et al., 2022); DRM adds randomly placed vehicles to the urban scene. We also use U6G (Li et al., 2026) with 560 training, 80 validation, and 160 held-out scenes (seed 42). The U6G tasks vary the target array (8×8 to 32×32), frequency (2.6, 3.5, or 4.9 GHz), or beam direction (eight angles). The complete target radio map is unavailable to `FOMRadio` at inference.

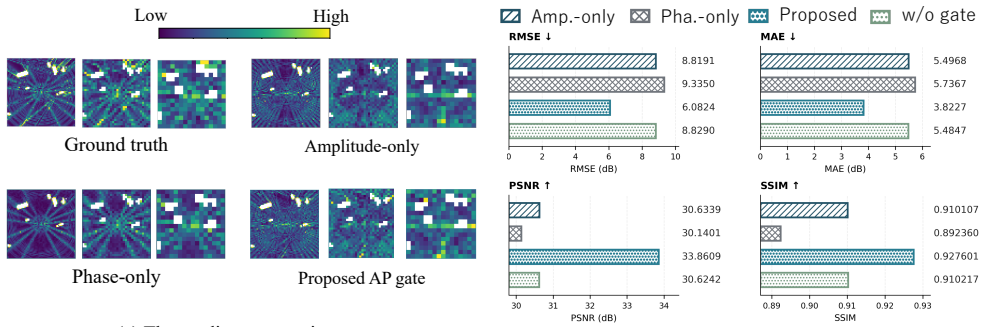
Baseline methods. We compare convolutional RadioUNet (Levie et al., 2021), adversarial RME-GAN (Zhang et al., 2023), UVM-Net (Zheng & Wu, 2024), diffusion-based RadioDiff (Wang et al., 2025) and RMDM (Jia et al., 2025b), and flow-based RadioFlow (Jia et al., 2025c). In U6G, the baselines receive the transmitter mask, building heights, and analytical target beam map. `FOMRadio` also receives the available same-scene 6.7-GHz, 0°, 8×8 auxiliary power field, its validity mask, and its beam map. The U6G RadioFlow implementation .

Metrics. For RadioMapSeer we report normalized-domain NMSE and RMSE, PSNR, and SSIM. For U6G we compute scene-mean RMSE and MAE in dB, PSNR over a 300-dB range, and SSIM (Wang et al., 2004). More details are given in Appendix **??**.

5.2 Results

Static and dynamic radio maps. Table 1 compares the methods on SRM and DRM. On SRM, `FOMRadio` attains the lowest NMSE (0.0019) and RMSE (0.0101) and the highest PSNR (39.91 dB). On DRM, it has the lowest RMSE (0.0106) and highest PSNR (39.41 dB), while RadioFlow Large has lower NMSE (0.0028 versus 0.0031). RMDM retains the highest SSIM in both settings. The model classes help explain the pattern. RadioUNet and RME-GAN rely on convolutional image synthesis, UVM-Net uses a vision state-space backbone, and RadioDiff and RMDM refine maps by denoising. RadioFlow transports an independent Gaussian source with a conditional velocity. `FOMRadio` instead transports a spatially correlated source through a Fourier operator with scene-conditioned amplitude–phase modulation. Global mode coupling supports broad coverage;

condition features and skip paths carry local obstruction cues, and the gate adapts retained spectral modes to them. This combination is consistent with its lower RMSE on both SRM and DRM. The SSIM results show that matching the finest structural patterns remains a separate challenge.

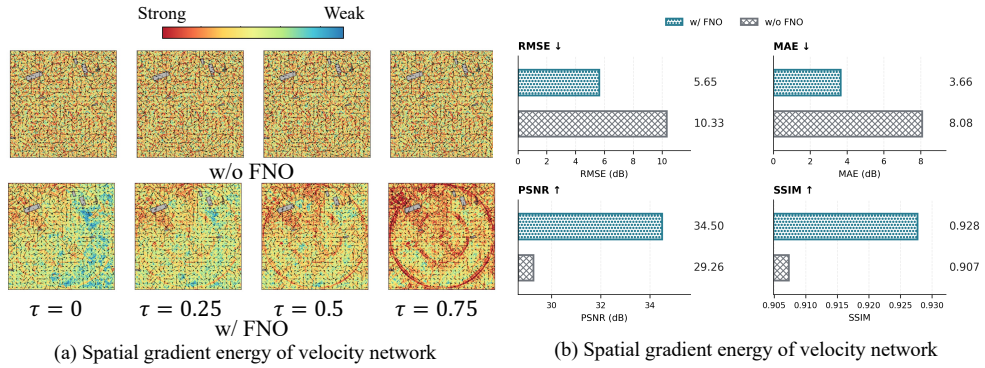


(a) The gradient comparisons

(b) The comparisons with different gate options

Figure 2: Comparison of the proposed amplitude-phase (AP) gate function. (a) Representative gradient maps produced by the amplitude-only, phase-only, and proposed AP function, where color indicates gradient magnitude. (b) Quantitative comparison against the single-component and ungated alternatives. The proposed AP gate function achieves the best performance, indicating that adaptive amplitude-phase modulation is beneficial for capturing the non-stationary radio states.

Figure 2 isolates the effect of the proposed AP gate. It reduces RMSE and MAE from 8.8290 and 5.4847 without gating to 6.0824 and 3.8227, corresponding to relative reductions of 31.1% and 30.3%. Meanwhile, PSNR increases by 3.24 dB and SSIM improves from 0.9102 to 0.9276. The amplitude-only and ungated variants perform nearly identically, while the phase-only variant is consistently weaker, showing that merely exposing the model to both components does not yield their complementary benefits. Amplitude describes the strength of individual frequency components, whereas phase determines their spatial arrangement. The AP gate turns their fixed combination into an input-dependent modulation, allowing `FOMRadio` to preserve the dominant propagation structure while selectively incorporating location-sensitive details. This interpretation is also supported by the gradient maps, where the proposed gate more closely recovers the reference patterns across different spatial scales.



(a) Spatial gradient energy of velocity network

(b) Spatial gradient energy of velocity network

Figure 3: Effect of the Fourier path on the learned transport velocity. Spatial-gradient energy of the velocity field at four flow times.

Figure 3 shows that the Fourier path changes both the structure and accuracy of the learned transport. Without FNO, the spatial gradient energy remains relatively diffuse across flow time, whereas the full model progressively organizes it into coherent, scene-dependent regions. This behavior is consistent with the global coupling induced by Fourier modes, which allows distant locations to be updated jointly rather than through predominantly local interactions. Quantitatively, enabling FNO

reduces RMSE from 10.33 to 5.65 and MAE from 8.08 to 3.66, corresponding to relative reductions of 45.3% and 54.7%, respectively. It also improves PSNR by 5.24 dB and SSIM from 0.907 to 0.928. Thus, the Fourier path does more than increase model capacity: it reorganizes the velocity field into a globally coordinated transport process that better captures scene-wide propagation structure. Since this visualization and comparison use a fixed T3 case, the result provides mechanism-level evidence rather than a dataset-wide generalization claim.

5.2.1 Transfer Across Array Sizes

Figure 4 examines whether an 8×8-conditioned model can resolve the finer propagation pattern of a 32×32 target at 6.7 GHz and $0°$ in held-out scenes. The plotted models were trained on 8×8 targets, making this a zero-shot target-array visualization.

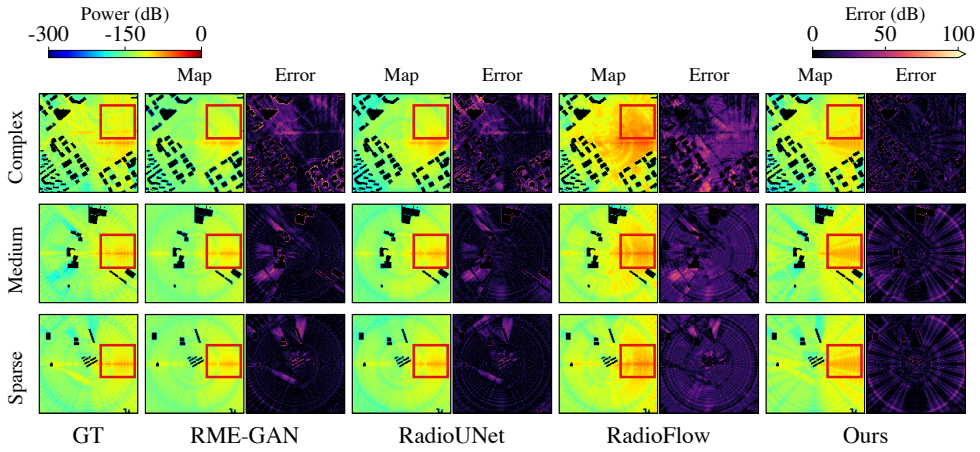


Figure 4: Zero-shot 8×8-to-32×32 transfer in complex, medium, and sparse scenes. Red boxes mark the same region in each power map.

The boxed regions in Figure 4 expose fine spatial propagation bands and shadow boundaries beyond the broad coverage lobe. RME-GAN and RadioUNet reproduce part of the smooth large-scale response but blur smaller bands, especially among the dense buildings in u785. RadioFlow shows a broad bright response displaced from the ground-truth pattern. Its independent Gaussian starting field carries no directional structure of its own, although the network receives the target beam condition. In the highlighted complex scene, `FOMRadio` follows more of the local bands and shadow boundaries. Its Matérn source supplies short-, intermediate-, and long-range spatial correlation; the beam-conditioned Fourier velocity and amplitude–phase gate reshape that variation according to the scene. The displayed medium and sparse scenes still contain fine-scale errors, so the visual evidence is strongest for the complex example.

Table 2: Array-transfer results on 32×32 targets. RMSE and MAE are in dB.

Method	RMSE	MAE	PSNR	SSIM
RME-GAN	14.230	9.933	26.809	0.665
RadioUNet	13.112	9.591	27.497	0.685
RadioFlow	22.355	19.858	22.897	0.689
Ours	**5.341**	**3.603**	**35.642**	**0.926**

Table 2 separately reports the quantitative results from the current evaluation table. `FOMRadio` obtains 5.341 dB RMSE, 3.603 dB MAE, 35.642 dB PSNR, and 0.926 SSIM over 160 scenes; RadioUNet has the lowest RMSE among the other reported rows at 13.112 dB. The auxiliary field supplies scene-level coverage, while the target beam specifies the changed array response. These numbers summarize a separately evaluated 256×256 checkpoint, whereas Figure 4 shows predictions from the zero-shot interference.

5.2.2 TARGET-FREQUENCY VARIATION

Figure 5 shows a representative 2.6-GHz 8×8, $0°$ target; Table 3 reports all three frequencies over 160 held-out scenes per frequency.

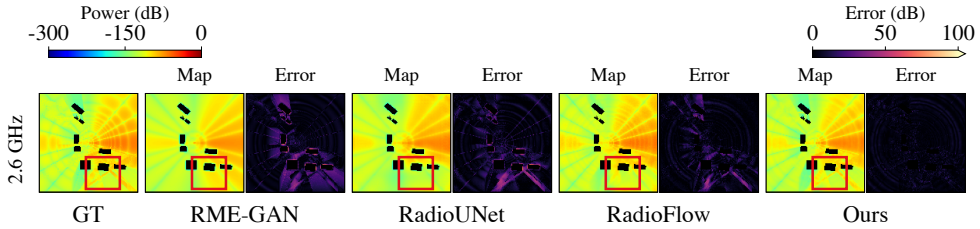


Figure 5: Representative 2.6-GHz 8×8, $0°$ target for scene u724. Each prediction is paired with its absolute-error map (dB).

Ours records 4.842, 4.528, and 5.098 dB RMSE and 0.901, 0.901, and 0.889 SSIM at 2.6, 3.5, and 4.9 GHz, respectively. Its MAE remains between 3.416 and 3.823 dB, with PSNR between 35.494 and 36.589 dB. The same-scene auxiliary field remains at 6.7 GHz while the target beam and frequency change. The similar error range suggests that the reference supports generation across these targets. RME-GAN and RadioUNet were trained at 6.7 GHz and tested zero-shot, whereas RadioFlow and Ours were trained on the three target frequencies. This setting probes how one fixed-frequency reference supports several target responses.

Table 3: Frequency results over 160 scenes per target. RMSE, MAE, and PSNR are in dB.

Method	RMSE	MAE	PSNR	SSIM
2.6 GHz				
RME-GAN	13.147	9.180	27.738	0.797
RadioUNet	11.474	7.927	28.840	0.795
RadioFlow	10.184	7.572	30.069	0.820
Ours	**4.842**	**3.823**	**35.961**	**0.901**
3.5 GHz				
RME-GAN	13.266	9.258	27.663	0.782
RadioUNet	11.362	7.875	28.941	0.780
RadioFlow	10.502	7.890	30.069	0.815
Ours	**4.528**	**3.416**	**36.589**	**0.901**
4.9 GHz				
RME-GAN	13.426	9.304	27.557	0.766
RadioUNet	11.300	7.877	28.992	0.769
RadioFlow	11.591	9.086	29.445	0.811
Ours	**5.098**	**3.801**	**35.494**	**0.889**

5.2.3 TRANSFER ACROSS BEAM DIRECTIONS

Figure 6 shows a representative $-28°$ target; Table 4 reports all eight directions at 6.7 GHz and 32×32 array size. For Ours, RMSE is lowest at $0°$ (5.341 dB), increases to 5.525 and 5.515 dB at $-7°$ and $+7°$, and reaches 6.592 and 6.272 dB at $-28°$ and $+21°$. The auxiliary field remains at $0°$ while the target beam direction changes. The training directions are $-14°$, $0°$, and $+14°$; $\pm 7°$ test interpolation, while $-28°$, $-21°$, and $+21°$ test extrapolation. The progression quantifies how accuracy changes with angular distance from the reference beam.

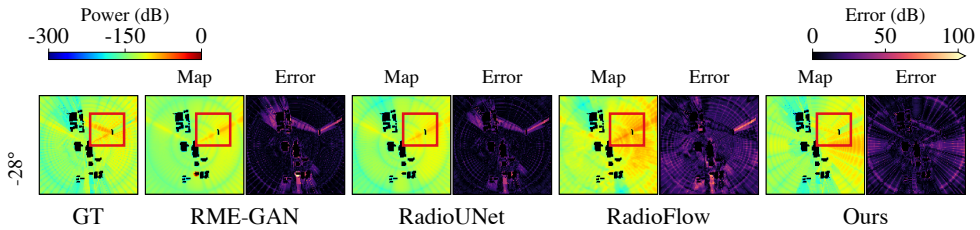
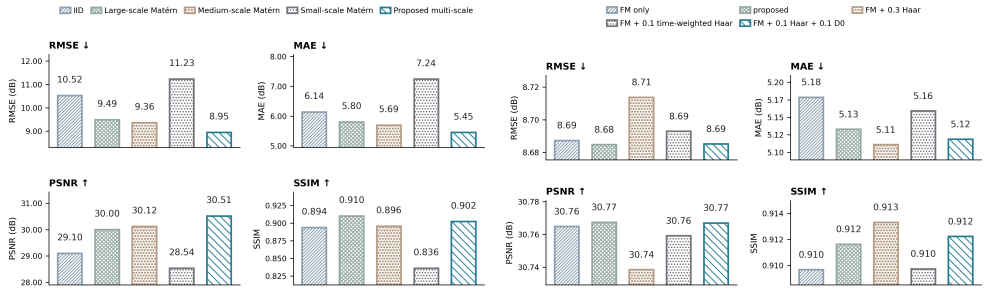


Figure 6: Representative $-28°$ target for scene u433 at 6.7 GHz and 32×32 array size. Each prediction is paired with its absolute-error map (dB).

Table 4: Direction results: scene-mean RMSE (dB) over 160 scenes per angle.

Method	$-28°$	$-21°$	$-14°$	$-7°$	$0°$	$+7°$	$+14°$	$+21°$
RME-GAN	14.106	14.108	14.148	14.239	14.230	14.249	14.241	14.207
RadioUNet	13.271	13.215	13.126	13.118	13.112	13.044	13.129	13.212
RadioFlow	21.296	21.798	22.195	22.407	22.355	22.402	22.196	21.788
Ours	**6.592**	**6.105**	**5.763**	**5.525**	**5.341**	**5.515**	**5.899**	**6.272**

5.2.4 Ablation study



(a) Ablation on different noise options

(b) Ablation on different loss options

Figure 7: Ablation of the noise construction and training objective. (a) Comparison of IID noise, single-scale Matérn noise, and the proposed multi-scale noise. (b) Comparison of the original FM objective with different Haar scales.

Figure 7 shows that noise construction has a substantially larger effect than loss design. Compared with IID noise, the proposed multi-scale noise reduces RMSE from 10.52 to 8.95 and MAE from 6.14 to 5.45 while improving PSNR by 1.41 dB; it also outperforms all single-scale variants on these metrics. Large-scale Matérn noise achieves the highest SSIM, whereas small-scale noise performs worst, indicating that the multi-scale design better balances global propagation structure and local variation. By contrast, the loss variants differ only marginally: the proposed objective achieves the lowest RMSE (8.68) and joint-highest PSNR (30.77 dB), while a larger Haar weight slightly improves MAE and SSIM at the expense of RMSE and PSNR. Overall, multi-scale noise is the primary contributor, with the loss design providing a balanced refinement.

6 Conclusion

We introduced `FOMRadio`, a conditional flow-matching framework for radio map generation based on a geometry-conditioned nonstationary Fourier operator. It separates spatially coherent source variation from environment-dependent transport: multiscale Matérn noise supplies correlations across spatial ranges, while the Fourier velocity network couples distant regions and gates retained modes according to the scene, beam, and evolving field. On RadioMapSeer, `FOMRadio` achieves the lowest reported RMSE on both static and vehicle-perturbed maps. On U6G, it maintains low error across the evaluated target array sizes, frequencies, and beam directions in held-out scenes without using the target radio map at inference. Ablations show gains from both the multiscale source and amplitude–phase gating.

Limitations. Angular extrapolation and fine multipath reconstruction remain challenging. `FOMRadio` does not achieve the highest SSIM on RadioMapSeer, and U6G assumes a same-scene auxiliary power field.

AI USE STATEMENT

We used OpenAI Codex to assist with manuscript drafting and English-language editing, literature searches, LaTeX formatting and bibliography preparation, and consistency checks between mathematical descriptions and the implementation. AI outputs were used as suggestions for writing and technical review. The authors remain responsible for the accuracy of the text, references, mathematical statements,experimental claims, and associated research artifacts, including AI-assisted material.

REFERENCES

Qi Chen, Jingjing Yang, Ming Huang, and Qiang Zhou. ACT-GAN: Radio map construction based on generative adversarial networks with ACT blocks. *IET Communications*, 18(19):1541–1550, 2024. doi: 10.1049/cmu2.12846.

You Fu, Ruyun Fu, Shengliang Fang, and Youchen Fan. RMF: One-step radio map reconstruction via mean flow matching. *Connection Science*, 38(1):2658948, 2026. doi: 10.1080/09540091.2026.2658948.

John Guibas, Morteza Mardani, Zongyi Li, Andrew Tao, Anima Anandkumar, and Bryan Catanzaro. Efficient token mixing for transformers via adaptive fourier neural operators. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=EXHG-A3jlM.

Ziyu Huang, Yong Zeng, Shen Fu, Xiaoli Xu, and Hongyang Du. Channel knowledge map construction via guided flow matching. *arXiv preprint arXiv:2601.06156*, 2026. doi: 10.48550/arXiv.2601.06156. URL https://arxiv.org/abs/2601.06156.

Haozhe Jia, Wenshuo Chen, Zhihui Huang, Lei Wang, Hongru Xiao, Nanqian Jia, Keming Wu, Songning Lai, Bowen Tian, and Yutao Yue. Physics-informed representation alignment for sparse radio-map reconstruction. In *Proceedings of the 33rd ACM International Conference on Multimedia*, MM '25, pp. 12352–12360. Association for Computing Machinery, 2025a. doi: 10.1145/3746027.3758161.

Haozhe Jia, Wenshuo Chen, Zhihui Huang, Hongru Xiao, Nanqian Jia, Keming Wu, Songning Lai, and Yutao Yue. RMDM: Radio map diffusion model with physics informed. *arXiv preprint arXiv:2501.19160*, 2025b. doi: 10.48550/arXiv.2501.19160. URL https://arxiv.org/abs/2501.19160.

Haozhe Jia, Wenshuo Chen, Xiucheng Wang, Nan Cheng, Hongbo Zhang, Kuimou Yu, Songning Lai, Nanjian Jia, Bowen Tian, Hongru Xiao, and Yutao Yue. RadioFlow: Efficient radio map construction framework with flow matching. *arXiv preprint arXiv:2510.09314*, 2025c. doi: 10.48550/arXiv.2510.09314. URL https://arxiv.org/abs/2510.09314.

Zhonghao Jiu, Fan Meng, Yongming Huang, Hang Zhan, Zening Liu, and Xiaohu You. Channel knowledge map reconstruction from sparse measurements via pilot-anchored layout-conditioned fourier refinement. *arXiv preprint arXiv:2607.01453*, 2026. doi: 10.48550/arXiv.2607.01453. URL https://arxiv.org/abs/2607.01453.

Ron Levie, Çağkan Yapar, Gitta Kutyniok, and Giuseppe Caire. RadioUNet: Fast radio map estimation with convolutional neural networks. *IEEE Transactions on Wireless Communications*, 20(6):4001–4015, 2021. doi: 10.1109/TWC.2021.3054977.

Peizheng Li, Xiaoyang Wang, Robert J. Piechocki, Shipra Kapoor, Angela Doufexi, and Arjun Parekh. Variational autoencoder assisted neural network likelihood RSRP prediction model. In *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 554–559, 2022. doi: 10.1109/PIMRC54779.2022.9978065.

Xiaojie Li, Yu Han, Zhizheng Lu, Shi Jin, and Chao-Kai Wen. U6G XL-MIMO radiomap prediction: Multi-config dataset and beam map approach. *arXiv preprint arXiv:2603.06401*, 2026. doi: 10.48550/arXiv.2603.06401. URL https://arxiv.org/abs/2603.06401.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=c8P9NQVtmnO.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Xiaoyi Liu and Hao Tang. DiffFNO: Diffusion fourier neural operator. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 150–160, 2025. URL https://openaccess.thecvf.com/content/CVPR2025/html/Liu_DiffFNO_Diffusion_Fourier_Neural_Operator_CVPR_2025_paper.html.

Xuanhao Luo, Zhizhen Li, Zhiyuan Peng, Mingzhe Chen, and Yuchen Liu. Denoising diffusion probabilistic model for radio map estimation in generative wireless networks. *IEEE Transactions on Cognitive Communications and Networking*, 11(2):751–763, 2025. doi: 10.1109/TCCN.2025.3529879.

Rafid Umayer Murshed, Saif Ur Rahman, Mingyue Tang, and Elahe Soltanaghai. Physics-unrolled neural operator for wireless field modeling. *arXiv preprint arXiv:2608.18495*, 2026. doi: 10.48550/arXiv.2608.18495. URL https://arxiv.org/abs/2608.18495.

Md Ashiqur Rahman, Zachary E. Ross, and Kamyar Azizzadenesheli. U-NO: U-shaped neural operators. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856. URL https://openreview.net/forum?id=j3oQF9coJd.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006. ISBN 978-0-262-18253-9. URL https://gaussianprocess.org/gpml/.

Hao Sun, Shicong Liu, Xianghao Yu, and Ying Sun. Flow matching-based active learning for radio map construction with low-altitude UAVs. *arXiv preprint arXiv:2509.13822*, 2025. doi: 10.48550/arXiv.2509.13822. URL https://arxiv.org/abs/2509.13822.

Xiucheng Wang, Keda Tao, Nan Cheng, Zhisheng Yin, Zan Li, Yuan Zhang, and Xuemin Shen. RadioDiff: An effective generative diffusion model for sampling-free dynamic radio map construction. *IEEE Transactions on Cognitive Communications and Networking*, 11(2):738–750, 2025. doi: 10.1109/TCCN.2024.3504489.

Xiucheng Wang, Qiming Zhang, Nan Cheng, Junting Chen, Zezhong Zhang, Zan Li, Shuguang Cui, and Xuemin Shen. RadioDiff-3D: A $3D \times 3D$ radio map dataset and generative diffusion based benchmark for 6G environment-aware communication. *IEEE Transactions on Network Science and Engineering*, 13:3773–3789, 2026a. doi: 10.1109/TNSE.2025.3590545.

Xiucheng Wang, Qiming Zhang, Nan Cheng, Ruijin Sun, Zan Li, Shuguang Cui, and Xuemin Shen. RadioDiff-k^2: Helmholtz equation informed generative diffusion model for multi-path aware radio map construction. *IEEE Journal on Selected Areas in Communications*, 44:2318–2333, 2026b. doi: 10.1109/JSAC.2025.3641105.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, 2004. doi: 10.1109/TIP.2003.819861.

Çağkan Yapar, Ron Levie, Gitta Kutyniok, and Giuseppe Caire. Dataset of pathloss and ToA radio maps with localization application. *arXiv preprint arXiv:2212.11777*, 2022. doi: 10.48550/arXiv.2212.11777. URL https://arxiv.org/abs/2212.11777.

Yong Zeng and Xiaoli Xu. Toward environment-aware 6G communications via channel knowledge map. *IEEE Wireless Communications*, 28(3):84–91, 2021. doi: 10.1109/MWC.001.2000327.

Yong Zeng, Junting Chen, Jie Xu, Di Wu, Xiaoli Xu, Shi Jin, Xiqi Gao, David Gesbert, Shuguang Cui, and Rui Zhang. A tutorial on environment-aware communications via channel knowledge map for 6G. *IEEE Communications Surveys & Tutorials*, 26(3):1478–1519, 2024. doi: 10.1109/COMST.2024.3364508.

Songyang Zhang, Achintha Wijesinghe, and Zhi Ding. RME-GAN: A learning framework for radio map estimation based on conditional generative adversarial network. *IEEE Internet of Things Journal*, 10(20):18016–18027, 2023. doi: 10.1109/JIOT.2023.3278235.

Zhuoran Zheng and Chen Wu. U-shaped vision mamba for single image dehazing. *arXiv preprint arXiv:2402.04139*, 2024. doi: 10.48550/arXiv.2402.04139. URL https://arxiv.org/abs/2402.04139.