

# 电磁频谱地图构建:稀疏高斯过程回归方法

张寿彬,王红军

(国防科技大学 电子对抗学院,合肥 230037)

E-mail: zhangshoubin19@nudt.edu.cn

**摘要:**针对大面积场景电磁频谱地图构建面临的大数据量、大计算开销的挑战,提出一种基于变分推断的稀疏高斯过程回归方法,实现电磁频谱地图构建。首先,将电磁频谱地图构建问题建模为一个回归问题,利用已采集的数据拟合地理位置与对应接收信号强度之间的映射关系。其次,将该映射关系建模为一个高斯随机过程,利用高斯过程回归的非参数化特性构建预测模型。最后,仅选取部分已采集数据作为模型的数据输入,通过最大化变分分布与后验分布之间的相似性训练得到一个稀疏的高斯过程回归预测模型。仿真实验和基于实采数据的实验表明,该方法能够在显著降低计算复杂度、加快运算速度的基础上,构建得到高精度的电磁频谱地图。

**关键词:**稀疏高斯过程回归;频谱感知;电磁频谱地图构建

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2026)05-1025-07

## Electromagnetic Spectrum Map Construction: a Sparse Gaussian Process Regression Approach

ZHANG Shoubin, WANG Hongjun

(College of Electronic Engineering, National University of Defense Technology, Hefei 230037, China)

**Abstract:** To address the challenges of massive data volume and substantial computational overhead in large area electromagnetic spectrum maps construction, a sparse Gaussian process regression method based on variational inference is proposed. First, the electromagnetic spectrum map construction problem is modeled as a regression problem, and mapping relationship between geographic location and corresponding received signal strength is fitted using collected data. Second, the mapping relationship is modeled as a Gaussian stochastic process, and the non-parametric property of Gaussian process regression is used to construct a prediction model. Finally, only part of the collected data is selected as input of the model, and a sparse Gaussian process regression prediction model is obtained by maximizing the similarity between variational distribution and posterior distribution. Simulation experiments and real data experiments show that this method can significantly reduce computational complexity and speed up operation process while obtaining high-precision electromagnetic spectrum maps.

**Keywords:** sparse Gaussian process regression; spectrum awareness; electromagnetic spectrum map construction

### 0 引言

电磁频谱地图 (Electromagnetic Spectrum Map, ESM) 是一种结合地理信息直观展示感兴趣区域 (Region of Interest, ROI) 内电磁特征分布的可视化方式。这些电磁特征包括接收信号强度、信道增益、电波衰减等<sup>[1]</sup>。随着信息技术的飞速发展,用频设备不断增多且朝着空天地海一体的广阔空间发展,ESM 被认为是辅助无线网络规划、频谱管理、辐射源定位、认知无线电、信息传输安全等应用有效且便捷的手段,因此引起了广泛关注<sup>[2-5]</sup>。现有研究提出了大量 ESM 构建的方法,这些方法可以按照对信号采集和处理方式的不同划分为正演法、反演法和混合法。

正演法首先定义辐射源参数如位置、发射功率等,然后利用电波传播模型计算出从辐射源到 ROI 内每一个位置的电波变化情况。电波传播模型包括经验传播模型如 Okumura-Hata<sup>[6]</sup> 和精确模型如射线追踪模型<sup>[7]</sup>。由于正演法仅根据

辐射源参数进行理论计算而不考虑实测数据,利用该方法构建的 ESM 并不能反应真实的电磁分布,因此常用于无线网络规划。

反演法则是在辐射源信息完全未知的情况下,根据空间稀疏分布的传感器采集的电磁数据设计相应算法,对未采集数据处的电磁特征进行推理补全。由于反演法仅依赖于实采数据而不使用电波传播模型,因此能更好反映现实场景的电磁分布,不仅可以应用到无线网络规划,还可以应用到非授权频段管理、应急救援和恶意辐射源发现与追踪等社会公共安全和反制方面。相较正演法,反演法构建 ESM 具有真实电磁环境的反映能力和更广阔的应用场景。因此,该方法是 ESM 构建领域的研究热点,本文的研究内容 and 应用背景也属于这一范畴。常用的算法包括插值类方法如克里金插值 (Kriging)<sup>[8]</sup>、反距离加权插值<sup>[9]</sup>、和样条插值<sup>[10]</sup>等,和回归类方法如随机森林 (Random Forest, RF) 回归<sup>[11]</sup> 和高斯过程回归 (Gaussian Process Regression, GPR)<sup>[12-16]</sup> 等。这两类方法具有

原理简单、易于操作的优点。此外,常用于缺失值恢复的矩阵补全(Matrix Completion, MC)<sup>[17]</sup>和张量补全<sup>[18]</sup>类算法也被应用到ESM构建中。然而现实应用中,矩阵补全类方法要获得与插值类和回归类方法相当的构建精度,往往需要数倍的采集数据作为算法的输入。因此,当ROI的面积增大时,该方法的数据采集开销过大,并不适用于大面积场景的ESM构建。随着人工智能技术的快速发展和在各个领域的广泛应用,深度学习也被应用到REM构建。考虑到深度学习类算法对图像数据处理的优越性和成熟应用,Suto等人<sup>[19,20]</sup>将ESM构建问题转化为缺失图像的补全问题,分别利用生成对抗网络和自编码器训练得到一个ESM补全模型,最终实现ESM的构建。由于深度学习类算法是数据驱动的,需要大量先验历史数据构造训练集进行训练。而这些历史数据往往是非常难以获取的。利用射线追踪模型生成的仿真数据构造的数据集训练得到的模型在现实场景的迁移效果也无法得到保证。此外,模型训练过程中输入的数据样式如图片大小等是严格限定的。模型训练结束后的输入样式须与训练时保持一致,因此无法获取任意大小区域的ESM,说明该方法缺乏应用的灵活性。

混合法是将正演法和反演法相结合实现ESM构建的方法。Wang等人<sup>[21]</sup>首先基于传播模型计算大尺度衰落,再基于观测数据利用高斯过程回归估计阴影衰落,将二者叠加实现ESM构建。混合法在正演法的基础上叠加了由反演法估计得到的衰落分量,能进一步反应实际的电波传播情况。但是该方法同样依赖于大尺度衰落正演计算的准确性,因此二者的结合应用仍需进一步研究。

近年来,由于GPR的非参数化特性和优秀的非线性数据拟合能力,不少研究工作将该方法应用到ESM构建当中。此外,GPR除了能对未观测位置的数据进行预测外,还能量化预测值的不确定性,为预测置信度和可靠性的分析提供了有效途径。尽管GPR在目标函数的形式上假设了一个高斯先验,但是可以利用似然函数有效拟合各种类型的数据分布,捕获复杂数据的变化特征,表现出优秀的通用性。这些性质进一步为GPR应用到ESM构建任务的可行性和合理性提供了理论支撑。Zhen等人<sup>[12]</sup>考虑到传感器自身位置测量存在不确定性的场景,研究了一种基于GPR且对传感器位置误差鲁棒的ESM构建方法。Konstantinos等人<sup>[13]</sup>基于GPR给出估计值的不确定性度量,并基于该度量动态指导数据的采集路径。Xu等人<sup>[14]</sup>为应对传播环境的异质性,提出了一种基于GPR的变换核学习方法。该方法在高斯过程引入一个非线性变换,将异质空间的数据输入映射到一个新的平稳空间。Wang等人<sup>[15]</sup>为提升室内指纹定位的准确性,构建了一个深度高斯过程回归网络,拟合室内复杂环境和传播效应。Nicolo等人<sup>[16]</sup>首次将GPR应用于大规模MIMO系统。

实际应用中,受限于数据采集、存储、处理和传输的开销,如何以尽量小的运算量实现高精度的大范围ESM的构建一直是一个具有重要应用价值的研究方向。综上所述,现有方法中,GPR虽然通过核函数灵活建模空间相关性并实现高精度预测,但其计算复杂度高,在大规模场景下面临严重的计算瓶颈。矩阵补全类算法虽然复杂度较低,但需数倍于其他方法的采样数据以维持同等精度,显著增加了实际部署成本。深度学

习方法虽然在图像式数据补全中表现优异,但其依赖大量历史数据训练模型,且模型泛化能力与输入格式灵活性受限,难以适应动态变化的电磁环境。上述方法的共性缺陷可以归结为两个方面:其一,高精度方法(如GPR)难以兼顾计算效率;其二,高效方法(如矩阵补全)往往以增加数据采集开销为代价。

针对上述矛盾,变分推断为平衡效率和精度提供了新思路。变分推断通过优化变分分布近似真实后验分布,避免了传统贝叶斯方法中复杂的高维积分运算,从而显著降低计算复杂度。Jiang等人<sup>[2]</sup>提出一种基于稀疏贝叶斯学习的ESM构建方案,利用变分分布和后验分布的相似性作为目标函数,优化求解传播模型中的关键待估计参数。Wang等人<sup>[22]</sup>基于稀疏贝叶斯层级模型,结合传播模型的先验,先推断辐射源的位置后实现ESM构建。而结合稀疏高斯过程回归,可以通过选取少量诱导点替代全数据集进行建模,同时保留核函数对空间相关性的刻画能力。因此,基于变分推断的稀疏高斯过程回归兼具如下优势<sup>[23]</sup>:

1)计算高效性:通过稀疏化策略显著降低存储和计算开销,适用于大范围ESM构建。

2)精度可控性:通过优化变分下界确保近似后验的准确性,避免因稀疏化导致的性能退化。

3)灵活性:不依赖于历史数据,可动态实行不同采样密度和复杂场景。

鉴于此,本文基于变分推断的稀疏高斯过程回归实现ESM构建,旨在以低计算代价实现高精度预测。

本文考虑ESM所展示的电磁特征为接收信号强度(Received Signal Strength, RSS)。假设ROI为一个二维平面,将该平面划分为大小相同的 $U$ 个网格。如果传感器在某网格内采集了电磁信号,则该网格称为采样网格,否则称为未采样网格。 $l_i = [x_i, y_i]^T$ 表示第 $i$ 个网格的坐标, $M$ 表示采样网格的数量。定义采样节点占比为采样网格与总网格之比。将采样网格对应的RSS值写成向量形式表示为 $Q_M = [Q_1, Q_2, \dots, Q_M]^T$ 。本文的任务就是根据观测向量 $Q_M$ 及其对应采样网格的位置估计未采样网格对应的RSS值。

## 1 高斯过程回归

将网格坐标和对应的RSS值之间的映射关系建模为一个回归模型,表示为:

$$Q_M = f_M(l) + \varepsilon \quad (1)$$

其中, $f_M(\cdot)$ 表示隐函数,可以用于计算得到真实的RSS值, $\varepsilon$ 表示服从高斯分布的零均值加性观测噪声,方差为 $\sigma_\varepsilon^2$ 。高斯过程回归建模认为隐函数 $f_M(\cdot)$ 来自于一个高斯过程:

$$f_M \sim GP\{\mu(l), K(l, l')\} \quad (2)$$

其中, $\mu(\cdot)$ 表示均值向量, $K(\cdot)$ 表示协方差矩阵。协方差矩阵中的元素通常通过核函数计算得到,本文选取高斯核函数,其表达式为:

$$[K]_{i,j} = \beta \exp\left\{-\frac{\|l - l'\|^2}{\gamma}\right\} \quad (3)$$

其中, $\beta, \gamma$ 表示需要拟合的超参数, $\|\cdot\|^2$ 表示两个网格之间的欧氏距离。则可以将 $Q_M$ 建模为服从高斯分布的随机变

量,即:

$$\mathbf{Q}_M \sim N(\boldsymbol{\mu}(\mathbf{I}), \mathbf{K}(\mathbf{I}, \mathbf{I}') + \sigma_\epsilon^2 \mathbf{I}) \quad (4)$$

未采样网格处的 RSS 值  $\mathbf{Q}^*$  和  $\mathbf{Q}$  之间服从联合高斯分布,表示为:

$$\begin{bmatrix} \mathbf{Q}_M \\ \mathbf{Q}^* \end{bmatrix} \sim N \left( \begin{bmatrix} \boldsymbol{\mu} \\ \boldsymbol{\mu}^* \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{MM} + \sigma_\epsilon^2 \mathbf{I} & \mathbf{K}_{M^*} \\ \mathbf{K}_{*M} & \mathbf{K}_{**} \end{bmatrix} \right) \quad (5)$$

其中,  $\mathbf{K}_{MM}$  表示采样网格之间的协方差矩阵,  $\mathbf{K}_{M^*}$ 、 $\mathbf{K}_{*M}$  表示采样网格与未采样网格之间的协方差矩阵,且  $\mathbf{K}_{M^*} = (\mathbf{K}_{*M})^T$ ,  $\mathbf{K}_{**}$  表示未采样网格之间的协方差矩阵。

假设在 ROI 内每一个网格的 RSS 值都具有相同的均值,则  $\boldsymbol{\mu}$  和  $\boldsymbol{\mu}^*$  的每一个元素值相等且表示为采样 RSS 样本的均值。则根据采样 RSS 值  $\mathbf{Q}_M$  推导  $\mathbf{Q}^*$  的后验概率分布表示为:

$$p(\mathbf{Q}^* | \mathbf{Q}_M, \mathbf{I}, \mathbf{I}') \sim N(\boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*) \quad (6)$$

其中:

$$\boldsymbol{\mu}_* = \boldsymbol{\mu}^* + \mathbf{K}_{*M} [\mathbf{K}_{MM} + \sigma_\epsilon^2 \mathbf{I}]^{-1} (\mathbf{Q}_M - \boldsymbol{\mu}) \quad (7)$$

$$\boldsymbol{\Sigma}_* = \mathbf{K}_{**} - \mathbf{K}_{*M} [\mathbf{K}_{MM} + \sigma_\epsilon^2 \mathbf{I}]^{-1} \mathbf{K}_{M^*} \quad (8)$$

则  $\boldsymbol{\mu}^*$  即为  $\mathbf{Q}^*$  的估计值。超参数  $\beta$  和  $\gamma$  通过共轭梯度法最大化边际似然函数求解:

$$\begin{aligned} \beta, \gamma &= \max_{\beta, \gamma} \{ \log p(\mathbf{Q}_M | \mathbf{I}, \beta, \gamma) \} \\ &= \max_{\beta, \gamma} \left\{ -\frac{1}{2} (\mathbf{Q}_M - \boldsymbol{\mu})^T (\mathbf{K}_{MM} + \sigma_\epsilon^2 \mathbf{I})^{-1} (\mathbf{Q}_M - \boldsymbol{\mu}) \right. \\ &\quad \left. - \frac{1}{2} \log |\mathbf{K}_{MM} + \sigma_\epsilon^2 \mathbf{I}| - \frac{M}{2} \log(2\pi) \right\} \quad (9) \end{aligned}$$

求解后验概率分布的参数时,存在一项求逆项  $[\mathbf{K}_{MM} + \sigma_\epsilon^2 \mathbf{I}]^{-1}$ , 其计算复杂度为  $O(M^3)$ 。当 ROI 面积较大,收集到的 RSS 数据过多时,其计算复杂度急剧升高,对计算单元的存储和计算开销提出了非常高的要求。为使得能以更快的速度构建 ESM,并且适应大范围的场景,本文使用稀疏高斯过程回归实现 ESM 的构建任务。期望在显著降低计算复杂度的同时保证 ESM 的构建精度。

## 2 稀疏高斯过程回归

近年来,许多研究工作提出使用稀疏高斯过程回归近似高斯过程回归的方法。这些近似方法主要分为 3 类:先验近似、后验近似和结构化近似。其中,先验近似通过引入诱导点直接对先验分布进行稀疏化,但其假设诱导点的分布独立于观测数据,可能因忽略数据特征而导致过拟合问题。结构化近似要求协方差矩阵具有特定低秩或分块结构,虽能降低计算复杂度,但需对输入空间或核函数施加强假设,限制了其在大规模复杂场景中的泛化能力。后验近似则通过变分推断直接优化后验分布的近似表达,能够在保证预测精度的同时自适应调整诱导点位置与变分参数,从而更好地平衡计算效率与模型灵活性。因此,本文选取后验近似方法中基于变分推断的稀疏高斯过程回归<sup>[24]</sup> (Variational Inference Sparse Gaussian Process Regression, VISGPR) 实现 ESM 的构建。VISGPR 通过最小化变分分布和后验分布之间的 Kullback-Leibler (KL) 散度实现,该方法可以有效处理大规模数据集,更好捕捉数据中的非线性关系,提高预测精度。其优势可以总结为:1) 计算复杂度可控:通过引入诱导点将计算复杂度从标准 GPR 的  $O(M^3)$  降低至  $O(M^2)$ ,适用于大规模数据场景;2) 泛化能力

更强:变分推断通过最大化边际似然下界优化诱导点位置与核参数,避免了先验近似对数据分布的先验假设偏差;3) 抗过拟合性:通过 KL 散度最小化约束变分分布与真实后验的差异,减少了因数据稀疏性导致的模型过拟合风险。

下面介绍 VISGPR 用于 ESM 构建的详细过程。

VISGPR 的稀疏性体现在当采样网格数量  $M$  过多而导致计算开销过大时,仅从采样网格中选取  $M$  个网格构成的子集用于训练回归模型。该子集表示为  $\mathcal{M}$ ,子集中的元素称为诱导点,则  $f_M$  表示诱导点对应的隐函数值。假设  $f_M$  和  $f_M$  服从相同的高斯先验,且  $f_M$  是  $f_M$  的一个充分统计量。那么对于任意有限函数点集  $Z$ ,有  $p(Z | f_M, f_M) = p(Z | f_M)$ 。由此可以推导如下变分分布:

$$\begin{aligned} p(Z | \mathbf{Q}_M) &= \int p(Z | f_M) p(f_M | \mathbf{Q}_M) df_M \\ &= \int p(Z | f_M, f_M) p(f_M | f_M, \mathbf{Q}_M) \times \\ &\quad p(f_M | \mathbf{Q}_M) df_M df_M \\ &= \int p(Z | f_M) p(f_M | \mathbf{Q}_M) df_M \quad (10) \end{aligned}$$

其中,  $p(Z | \mathbf{Q}_M)$  和  $p(f_M | \mathbf{Q}_M)$  表示未知的后验分布。

令  $q(Z)$  和  $\phi(f_M)$  分别表示  $p(Z | \mathbf{Q}_M)$  和  $p(f_M | \mathbf{Q}_M)$  的变分近似,可以得到:

$$p(Z | \mathbf{Q}_M) \approx q(Z) = \int p(Z | f_M) \phi(f_M) df_M \quad (11)$$

$\phi(f_M)$  的先验分布可以从任意分布簇中选取,通常假设服从高斯分布,其均值和方差分别表示为  $\mathbf{A}$  和  $\mathbf{B}$ 。当  $Z$  表示未采样网格处的 RSS 估计值  $\mathbf{Q}^*$  时,其后验分布为  $q(Z) = N(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}_*)$ 。

其中:

$$\boldsymbol{\mu}_* = \boldsymbol{\mu}^* + \mathbf{K}_{*M} \mathbf{K}_{MM}^{-1} (\mathbf{Q}_M - \mathbf{A}) \quad (12)$$

$$\boldsymbol{\Sigma}_* = \mathbf{K}_{**} - \mathbf{K}_{*M} (\mathbf{K}_{MM}^{-1} - \mathbf{K}_{MM}^{-1} \mathbf{B} \mathbf{K}_{MM}^{-1}) \mathbf{K}_{M^*} \quad (13)$$

由上述内容可知,想要利用稀疏高斯过程回归估计未采样网格处的 RSS 值,关键在于诱导点集合  $\mathcal{M}$  和变分近似  $\phi(f_M)$  的参数  $\mathbf{A}$ 、 $\mathbf{B}$ 。

$\phi(f_M)$  可以通过使得后验分布  $p(f_M | \mathbf{Q}_M)$  与变分分布  $q(f_M)$  之间的相似性最大求解得到。而概率密度函数之间的相似性通过 KL 散度衡量,即最小化  $p(f_M | \mathbf{Q}_M)$  与  $q(f_M)$  的 KL 散度。进一步考虑到诱导点的选取和获得更好的训练效果,使用增广分布  $p(f_M, f_M | \mathbf{Q}_M)$  与  $q(f_M, f_M)$  的 KL 散度作为目标函数,即:

$$M, \phi(f_M) = \min_{M, \phi(f_M)} KL(q(f_M, f_M) \| p(f_M, f_M | \mathbf{Q}_M)) \quad (14)$$

上述最小化问题等价于真实边际似然的变分下界最大化问题:

$$\begin{aligned} F_V(\mathcal{M}, \phi(f_M)) &= \int p(f_M | f_M) \phi(f_M) \times \\ &\quad \log \frac{p(\mathbf{Q}_M | f_M) p(f_M)}{\phi(f_M)} df_M \quad (15) \\ F_V(\mathcal{M}, \phi(f_M)) &\text{在 } \phi^*(f_M) \text{ 处取极大值时,得到:} \\ &F_V(\mathcal{M}, \phi^*(f_M)) \\ &= -\frac{M}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{K}_{MM} \mathbf{K}_{MM}^{-1} \mathbf{K}_{MM} + \sigma_\epsilon^2 \mathbf{I}| - \\ &\quad \frac{1}{2} (\mathbf{Q}_M - \boldsymbol{\mu})^T (\mathbf{K}_{MM} \mathbf{K}_{MM}^{-1} \mathbf{K}_{MM} + \sigma_\epsilon^2 \mathbf{I})^{-1} (\mathbf{Q}_M - \boldsymbol{\mu}) - \end{aligned}$$

$$\frac{1}{2\sigma_e^2} \text{tr}(\mathbf{K}_{MM} - \mathbf{K}_{MM} \mathbf{K}_{MM}^{-1} \mathbf{K}_{MM}) \quad (16)$$

$$\mathbf{A} = \frac{1}{\sigma_e^2} \mathbf{K}_{MM} \left( \mathbf{K}_{MM} + \frac{1}{\sigma_e^2} \mathbf{K}_{MM} \mathbf{K}_{MM} \right)^{-1} \times \mathbf{K}_{MM} (\mathbf{Q}_M - \boldsymbol{\mu}) \quad (17)$$

$$\mathbf{B} = \mathbf{K}_{MM} \left( \mathbf{K}_{MM} + \frac{1}{\sigma_e^2} \mathbf{K}_{MM} \mathbf{K}_{MM} \right)^{-1} \mathbf{K}_{MM} \quad (18)$$

基于变分推断的稀疏高斯过程回归的一大优势在于通过增广分布的处理,诱导点转化为变分参数,一方面可以结合KL散度最小化问题联合求解 $\mathcal{M}$ 和 $\phi(f_M)$ ,另一方面可以减小训练过程中的过拟合问题.对于给定的诱导点,高斯核函数的超参数通过最大化 $F_V(\mathcal{M}, \phi(f_M))$ 实现.该过程同样利用共轭梯度法完成.

由于诱导点的选取是一个离散优化的问题,本文采取贪婪选择的策略,使用交替优化的方法进行确定.将训练样本集的样本点进行编号,训练样本索引全集表示为 $S = \{1, 2, \dots, M\}$ ,诱导点索引集表示为 $I$ ,除诱导点外剩余训练样本索引集表示为 $R$ ,则有 $I \cup R = S$ .将 $I$ 初始化为空集, $R$ 初始化为 $S$ .在当前迭代的过程中,从 $S$ 中随机选取一个索引放入 $I$ , $R$ 相应减去该索引.基于当前诱导点集以最大化 $F_V(\mathcal{M}, \phi(f_M))$ 为目标,使用共轭梯度下降法确定核函数的超参数.在下次迭代过程中,保持上一次迭代得到的核函数超参数不变,再 $S$ 中随机选取一个索引放入 $I$ .若当前诱导点能保证 $F_V(\mathcal{M}, \phi(f_M))$ 递增则保持该诱导点,否则舍弃该诱导点进行重新随机选择.当 $I$ 中的元素数量达到初始设置的诱导点数量则算法停止.该策略的收敛性由Michalis等人<sup>[25]</sup>给出.

### 3 仿真实验与分析

本节内容主要介绍电磁频谱地图构建的仿真环境设置,并通过定量和定性分析的角度将所提算法与对比算法进行对比分析.

#### 3.1 仿真环境设置

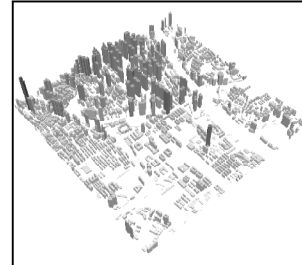
本文的基准仿真数据使用通用的电波传播和无线网络规划软件Altair WinProp生成.其中,数据生成模型选取射线追踪模型的优势路径法.为充分考虑现实条件下电波传播环境的复杂性,从广州城区内选取一片大小为 $2\text{km} \times 2\text{km}$ 的区域作为ROI.ROI内包含了该区域建筑物的真实分布.定义两个移动通信基站作为辐射源,每一个辐射源划分为3个扇区,下倾角为 $5^\circ$ ,高度为 $30\text{m}$ ,发射功率为 $43\text{dBm}$ ,载波频率为无线通信常用频率 $2.6\text{GHz}$ .城市地图建筑物分布和辐射源位置分布如图1所示.本文在计算机硬件配置为13th Gen Intel(R) Core(TM) i9-13900KF  $3.00\text{GHz}$ 和软件配置为WinProp2022、Matlab R2023b的条件下进行仿真实验.

#### 3.2 评价指标

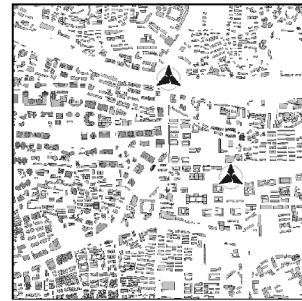
本文从ESM的构建速度和构建精度两个方面评估算法的性能.其中,构建速度通过分析算法复杂度和在相同硬件条件下的运行时间进行评估.构建精度通过均方根误差(Root Mean Square Error, RMSE)进行定量评价. RMSE的计算公式为:

$$\text{RMSE} = \sqrt{\frac{1}{U} \sum_{i=1}^U (r_i - \hat{r}_i)^2} \quad (19)$$

其中, $r_i$ 和 $\hat{r}_i$ 表示RSS的真实值和预测值.为了验证本文方法的性能,将GPR<sup>[16]</sup>、Kriging<sup>[8]</sup>、RF<sup>[11]</sup>和MC<sup>[17]</sup>算法作为基准对比方法进行对比分析.此外,为验证VISGPR算法的有效性,本文还将与VISGPR相同的诱导点输入到GPR模型中,记为GPR\*.



(a) 三维ROI示意图



(b) 二维ROI俯视图

图1 城市建筑及辐射源位置分布示意图

Fig. 1 Distribution of urban buildings and radiation source

#### 3.3 性能分析

##### 3.3.1 ESM 构建性能分析

如前所述,GPR的计算复杂度为 $O(M^3)$ ,本文选取VISGPR的诱导点数量为 $M = M/2$ ,则VISGPR的计算复杂度为 $O(M^3/8)$ .GPR\*的计算复杂度同样为 $O(M^3/8)$ .Kriging与GPR类似,涉及矩阵求逆的运算,其计算复杂度为 $O(M^3)$ .RF的计算复杂度为 $O(kfM \log M)$ .其中 $k$ 是组成随机森林的决策树的数量,本文取为100; $f$ 是特征维度,本文考虑一个二维平面,故 $f=2$ .MC法是将ROI划分的网格表示成大小为的矩阵进行计算,主要运算开销在于奇异值的分解,故其计算复杂度为 $O(U^{3/2})$ .从计算复杂度的理论分析可以得出,计算复杂度大小的排序为:

$$\text{GPR} > \text{VISGPR} = \text{GPR}^* > \text{RF} > \text{MC} \quad (20)$$

图2展示了各个构建方法在相同实验条件下ESM构建用时随采样节点占比从1%~10%的变化情况.图2(b)是图2(a)的局部放大.如前所述,GPR和Kriging的计算复杂度与采样网格数量的三次方成正比,从图2也可以看出,随着采样节点数量的不断增多,这两种方法构建ESM花费的时间快速增长.而VISGPR和GPR\*通过减少数据的输入大大减小了运算所花费的时间.从图中还可以看出,采样节点占比位于1%~5%时,GPR\*的运算时间小于VISGPR,而采样节点占比大于5%时VISGPR的运算时间更小.这说明VISGPR在大数据量的条件下,比具有相同输入数据的GPR\*更有优势,可

以取得更小的运算开销. RF 和 MC 算法的运算时间远小于其余 4 种算法,这与理论的分析保持一致.

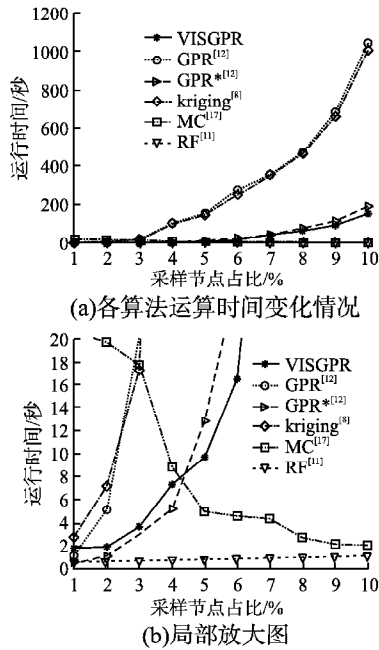


图 2 不同采样节点占比下各算法运算时间变化情况  
Fig. 2 Variation of run time with different sampling nodes

3.3.2 ESM 构建精度分析

图 3 展示了不同算法在不同采样节点占比的条件下构建得到的 ESM 的 RMSE 变化情况. 如前所述,矩阵补全类方法要获得与插值类和回归类算法相当的精度需要数倍的采样数

据输入. 因此,MC 算法的采样节点占比的变化区间为 15% ~ 60%, 对应为左侧纵坐标轴. 其余方法的采样节点变化区间为 1% ~ 10%, 对应为右侧纵坐标轴. 从图 3 可以看出,随着采样节点占比的变化, GPR 能够获得最高的构建精度, 而 VISGPR 次之.

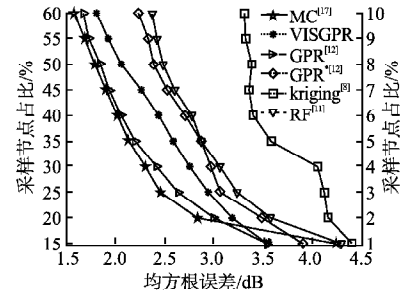


图 3 不同采样节点占比下各算法均方根误差变化情况  
Fig. 3 Variation of RMSE with different sampling nodes

图 4 展示了在采样节点占比为 5% 的条件下, 不同算法构建得到的 ESM 的可视化效果. 由于 MC 算法在该条件下并不能有效构建 ESM, 因此未展示该算法的构建结果. 从图 4 可以看出, Kriging 和 RF 方法构建得到的 ESM 只能大致反应 ROI 内的电磁分布状况, 但是对于一些细节部分, 特别是辐射源周围的电磁分布, 并不能很好地刻画. 与基于高斯过程回归的方法相比, Kriging 和 RF 是粗粒度的. 而基于高斯过程回归的方法中, GPR、GPR\* 和 VISGPR 都能较好地反应电磁分布状况. 但是从对辐射源周围的电磁分布的刻画细节来看, VISGPR 展现了更好的性能.

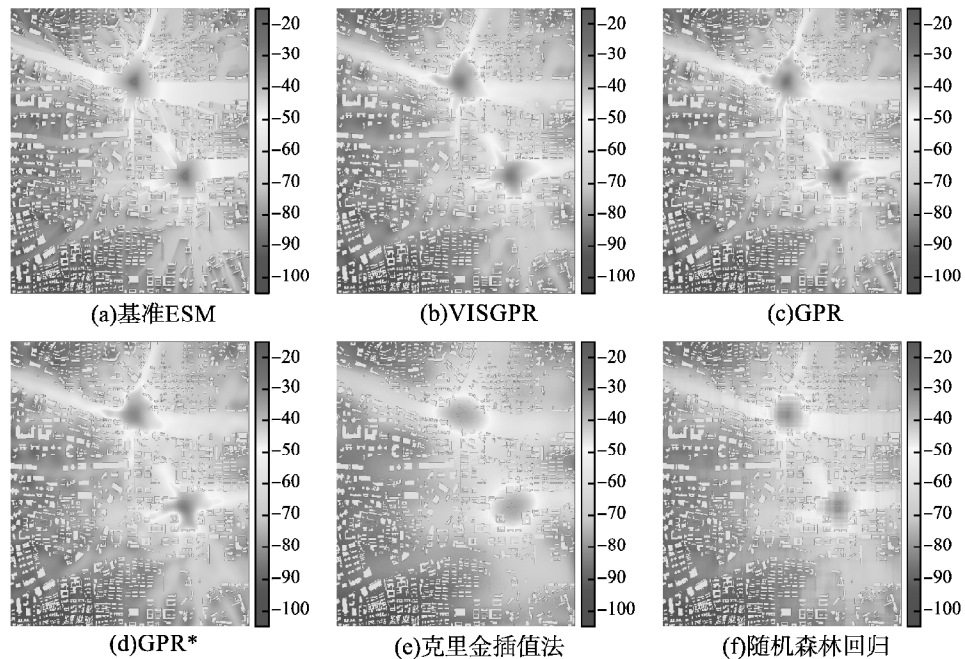


图 4 不同算法构建 ESM 可视化  
Fig. 4 Visualization of ESM via different methods

3.3.3 综合分析

从计算复杂度的理论分析来看, VISGPR 相较于 GPR, 能

明显减小计算开销, 在加速 ESM 的构建进程的同时保持相当的精度. 从构建精度来看, 虽然 MC 算法的计算复杂度和运行

时间远小于 VISGPR,但是 MC 算法要获得相当的精度,需要数倍于 VISGPR 的采集数据作为输入.即算法复杂度和构建精度需要很大的数据采集开销作为前提条件.而实际应用过程中,数据采集的难度往往是最大的.数据采集对于 ESM 构建系统的开销远大于计算资源的开销.因此,相较于 MC 算法,在大范围的 ESM 构建任务中,VISGPR 更具优势.此外,RF 的计算复杂度也低于 VISGPR,但是从 RMSE 的定量对比和可视化结果来看,RF 并不能构建高质量的 ESM.综上所述,综合考虑 ESM 的构建精度和构建速度两个方面,VISGPR 能够获得最优的性能.

#### 4 实采数据实验与分析

本节内容主要介绍电磁频谱地图构建的数据采集环境设置,并验证所提方法在真实环境下的有效性.

##### 4.1 环境设置

选择一片大小为  $400\text{m} \times 400\text{m}$  的区域固定放置两个辐射源.将该区域等面积划分为 1600 个  $5\text{m} \times 5\text{m}$  的网格.为获得完整的数据验证所提算法,实验过程中使用一个接收机对所有网格位置都进行信号的采集.辐射源和接收机都采用通用软件无线电外设 (Universal Software Radio Peripheral) 的 N210 型号配置全向天线.辐射源外接功率放大器.软件配置使用 Ubuntu18.04 操作系统配置 UHD4.0 和 GNUradio3.8.硬件和软件配置示意图如图 5 所示.

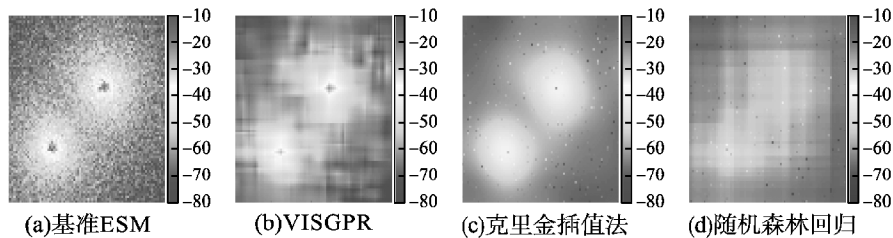


图 6 不同算法构建 ESM 可视化

Fig. 6 Visualization of ESM via different methods

#### 5 结论

本文研究了在未知辐射源参数条件下基于观测数据的电磁频谱地图构建问题.通过将电磁频谱地图构建任务转化为一个回归任务,利用已采集的数据拟合地理位置与其对应接收信号强度之间的映射关系,最终实现对未采集数据位置的接收信号强度值的预测.映射关系的拟合通过基于变分推断的稀疏高斯过程实现.仿真实验和基于实采数据实验表明,该方法能够在显著降低计算复杂度,减小计算时间的同时保持较高的构建精度,实现了构建速度和构建精度的平衡结合,将在大范围的电磁频谱地图构建的现实应用中具有更大的优势.该方法可应用于 5G/6G 动态网络规划、智能感知定位、智慧城市电磁干扰管控及空地海一体化通信等领域,为频谱资源动态分配、复杂环境监测、应急通信部署等场景提供高效支持.其低数据需求与快速计算特性尤其适用于卫星通信、工业互联网等资源受限场景,未来结合边缘计算与多模态数据融合,将进一步推动电磁环境感知技术的智能化与普适化发展,具有广泛的工程应用价值.

#### 4.2 结果分析

在采样节点占比为 5% 的条件下,基于实采数据实现 ESM 构建,可视化结果如图 6 所示.对比图 4(a) 和图 6(a) 可以看出,仿真数据模拟的 ESM 变化趋势较为平滑,而实采数



图 5 实验环境示意图

Fig. 5 Experimental environment

据相邻两个网格的接收信号强度值出现突变的位置较多.对应于图 4(f) 和图 6(d) 可以看出,随机森林回归在仿真数据中构建的 ESM 在视觉上与真实 ESM 相近,而在实采数据中表现较差.该结果说明随机森林回归方法在真实场景下的应用性欠佳.从视觉效果上看,VISGPR 构建的 ESM 与基准 ESM 最相近,相较于对比方法可以获得最佳性能.此外,VISGPR、Kriging 和随机森林回归的 RMSE 分别为 3.79、4.21 和 5.56,也从数值上进一步说明本文所提方法的有效性.

#### References:

- [1] Romero D, Kim S. Radio map estimation: a data-driven approach to spectrum cartography[J]. IEEE Signal Processing Magazine, 2022, 39(6): 53-72.
- [2] Wang J, Zhu Q M, Lin Z P, et al. Sparse Bayesian learning-based 3d radio environment map construction sampling optimization, scenario-dependent dictionary construction, and sparse recovery[J]. IEEE Transactions on Cognitive Communications and Networking, 2024, 10(1): 80-93.
- [3] Jiang H, Mukherjee M, Zhou J, et al. Channel modeling and characteristics for 6G wireless communications[J]. IEEE Network, 2021, 35(1): 296-303.
- [4] Shi W Q, Jiang H, Xiong B P, et al. RIS-Empowered V2V communications: three-dimensional beam domain channel modeling and analysis[J]. IEEE Transactions on Wireless Communications, 2024, 23(11): 15844-15857.
- [5] TU C, ZHU W L, ZHANG M, et al. A broadband intelligent direction finding method based on CNN features[J]. Information Countermeasure Technology, 2022, 1(2): 75-84.

- [ 6 ] Hata M. Empirical formula for propagation loss in land mobile radio service[J]. IEEE Transactions on Vehicular Technology, 1980, 29(3):317-325.
- [ 7 ] Straka T, Vojtech L, Neruda M. Graphical heatmap-based approach to indoor radio signal propagation: adapting advanced ray tracing and global illumination algorithms[J]. IEEE Transactions on Antennas and Propagation, 2024, 72(7):6045-6059.
- [ 8 ] Xia H Y, Zha S, Huang J J, et al. Radio environment map construction by residual kriging based on Bayesian hierarchical model [C]//Proceedings of International Symposium on Electromagnetic Compatibility (ISEMC), 2023:1-5.
- [ 9 ] TAO S F, WU Y J, LUO J, et al. Radio environment map construction method for complex scenes based on inverse obstacle distance weighted[J]. Journal of Electronics & Information Technology, 2024, 46(8):3210-3218.
- [ 10 ] ZHAO W L, TIAN X Y, CHEN C, et al. Wi-Fi Fingerprint localization uniting spline interpolation[J]. Journal of Electronics & Information Technology, 2024, 46(9):3563-3570.
- [ 11 ] Mario R C, Carla E G, Taewoong H, et al. A REM update methodology based on clustering and random forest[J]. Applied Sciences, 2023, 13(9):5362, doi:10.3390/app13095362.
- [ 12 ] Zhen P, Zhang B N, Xu Y Q, et al. Radio environment map construction based on gaussian process with positional uncertainty[J]. IEEE Wireless Communications Letters, 2022, 11(8):2162-2337.
- [ 13 ] Konstantinos D P, Alireza S, Wei Y, et al. Bayesian active learning for sample efficient 5G radio map reconstruction[J]. IEEE Transactions on Wireless Communication, 2024, 30(12):19382-19396.
- [ 14 ] Xu Y Q, Zhang B N, Zhang X K, et al. Radio environment map construction with Gaussian process and kernel transformation [C]//Proceedings of International Conference on Communication Image, and Signal Processing (CCISP), 2021:350-355.
- [ 15 ] Wang X Y, Wang X Y, Mao S W, et al. Indoor radio map construction and localization with deep gaussian processes[J]. IEEE Internet of Things Journal, 2020, 7(11):11238-11249.
- [ 16 ] Nicolo D F, Michele R, Gianluigi P, et al. Model-free radio map estimation in massive MIMO systems via semi-parametric gaussian regression[J]. IEEE Wireless Communications Letters, 2022, 11(3):473-477.
- [ 17 ] Zhang Y L, Ma L. Radio map crowd sourcing update method using sparse representation and low rank matrix recovery for WLAN indoor positioning system[J]. IEEE Wireless Communications Letters, 2021, 10(6):1188-1191.
- [ 18 ] CHEN Z B, HU J M, ZHANG B N, et al. Spectrum map construction algorithm based on tensor tucker decomposition[J]. Journal of Electronics & Information Technology, 2023, 45(11):4161-4169.
- [ 19 ] Suto K, Bannai S, Sato K, et al. Image-driven spatial interpolation with deep learning for radio map construction[J]. IEEE Wireless Communications Letters, 2021, 10(6):1222-1226.
- [ 20 ] Zhang S Y, Wijesinghe C, Ding Z. RME-GAN: a learning framework for radio map estimation based on conditional generative adversarial network[J]. IEEE Internet of Things Journal, 2023, 10(20):18016-18027.
- [ 21 ] Wang J, Zhu Q M, Lin Z P, et al. Sparse Bayesian learning-based hierarchical construction for 3D radio environment maps incorporating channel shadowing[J]. IEEE Transactions on Wireless Communication, 2024, 23(10):14560-14574.
- [ 22 ] Wang J, Zhu Q M, Lin Z P, et al. Sparse Bayesian learning-based hierarchical construction for 3D radio environment maps incorporating channel shadowing[J]. IEEE Transactions on Wireless Communication, 2024, 23(10):14560-14574.
- [ 23 ] Lyu C, Liu X, Mihaylova L. Review of recent advances in Gaussian process regression methods[C]//Proceedings of the UK Workshop on Computational Intelligence, 2022:226-237.
- [ 24 ] David R Burt, Edward R C, Wilk V D. Convergence of sparse variational inference in Gaussian processes regression[J]. Journal of Machine Learning Research, 2020, 21(119):1-63.
- [ 25 ] Michalis C S. Variational learning of introducing variables in sparse gaussian process[C]//Proceedings of International Conference on Artificial Intelligence and Statistics (AISTATS), 2009:567-574.

#### 附中文参考文献:

- [ 5 ] 屠 铖, 朱文丽, 张 旻, 等. 基于 CNN 特征的宽频段智能测向方法[J]. 信息对抗技术, 2022, 1(2):75-84.
- [ 9 ] 陶诗飞, 吴昱江, 罗 佳, 等. 基于反障碍距离加权的复杂场景电磁频谱地图构建方法[J]. 电子与信息学报, 2024, 46(8):3210-3218.
- [ 10 ] 赵万龙, 田新元, 陈 超, 等. 联合 Spline 插值的 Wi-Fi 指纹匹配定位算法[J]. 电子与信息学报, 2024, 46(9):3563-3570.
- [ 18 ] 陈智博, 胡景明, 张邦宁, 等. 基于张量 Tucker 分解的频谱地图构建算法[J]. 电子与信息学报, 2023, 45(11):4161-4169.