

学 号	2017151613
密 级	公 开

哈尔滨工程大学学士学位论文/毕业设计

## 空间辐射场三维重构方法

院（系）名 称：核科学与技术学院

专 业 名 称：核工程与核技术

学 生 姓 名：刘铭

指 导 教 师：宋玉收 教授

哈尔滨工程大学

2021 年 6 月

学 号	2017151613
密 级	公 开

# 空间辐射场三维重构方法

## Three-dimensional Deconstruction Method of Space Radiation Field

学 生 姓 名：刘铭

所 在 学 院：核科学与技术学院

所 在 专 业：核工程与核技术

指 导 教 师：宋玉收

职 称：教授

所 在 单 位：哈尔滨工程大学

论文提交日期：2021 年 6 月 20 日

论文答辩日期：2021 年 6 月 13 日

学位授予单位：哈尔滨工程大学

## 摘 要

辐射场可视化仿真技术是辐射防护中的一项重要技术，其应用场景十分广泛，例如在核设施退役工程中评估工作人员照射的辐射剂量、在核辐射安全科普工作中向公众展示辐射分布信息、在矿井窑洞等辐射剂量高的地域监测放射性浓度等等。空间辐射场重构方法是辐射场可视化仿真技术中最核心的技术，因此找寻一种重构速度快、精度高、适用性好的辐射场重构方法就显得尤为重要。

针对上述问题，本文对插值重构方法进行深入研究，提出一种基于多层 B 样条插值方法和克里金插值方法的空间辐射场插值重构方法，主要研究工作如下：

- 1.提出并实现了基于多层 B 样条插值方法和克里金插值方法的辐射场重构方法，详细地分析了多层 B 样条插值原理和克里金插值原理，通过将两种插值方法结合，提出一种精度更高、适用性更好的辐射场重构方法；
- 2.利用 Geant4 设计并构建了四种不同类型的辐射场模型，用于验证辐射场重构方法的可行性，分别探究了源项数量、空间状况以及测点数量对辐射场重构方法的影响；
- 3.将本论文提出的辐射场重构方法与多层 B 样条插值重构方法、克里金插值重构方法进行辐射场重构效果对比，进一步说明该辐射场重构方法的可行性。

**关键词：**辐射场重构；多层 B 样条插值；克里金插值；辐射场可视化

## ABSTRACT

Radiation field visualization simulation technology is an important technology in radiation protection, which has a variety of real-world applications, such as evaluating the radiation dose of workers in nuclear facility decommissioning projects, displaying radiation distribution information to the public in nuclear radiation safety science work, Monitoring the concentration of radioactivity in areas with high radiation doses such as mines and cave dwellings, etc. The space radiation field reconstruction method is the core technology of the radiation field visualization simulation technology. Therefore, it is particularly important to find a radiation field reconstruction method with fast reconstruction speed, high accuracy and good applicability.

Aiming at the above problems, this paper makes an in-depth study on the single interpolation and reconstruction method, and proposes a spatial interpolation and reconstruction method of radiation field based on Multilevel B-splines interpolation method and Kriging interpolation method. The main research contents are as follows:

1. The radiation field reconstruction method based on the Multilevel B-splines interpolation method and the Kriging interpolation method is proposed and implemented. The principle of the Multilevel B-splines interpolation and the Kriging interpolation principle are analyzed in detail. By combining the two interpolation methods, a radiation field reconstruction method with higher accuracy and better applicability is proposed;
2. Four different types of radiation field models were designed and constructed using Geant4 to verify the feasibility of the radiation field reconstruction method. The influence of the number of radioactive sources, space conditions and the number of measuring points on the method of radiation field reconstruction are respectively explored.
3. The radiation field reconstruction method proposed in this paper is compared with the Multilevel B-splines interpolation reconstruction method and the Kriging interpolation reconstruction method to compare the radiation field reconstruction effect, and further illustrate the feasibility of the radiation field reconstruction method.

**Keywords:** Radiation field reconstruction; Multilevel B-splines interpolation; Kriging interpolation; Radiation field visualization

## 目 录

第 1 章	绪论 .....	1
1.1	研究背景与意义 .....	1
1.2	研究方法和内容 .....	2
1.3	国内外研究现状 .....	2
1.4	本论文章节安排 .....	3
第 2 章	样条插值重构方法 .....	4
2.1	样条插值介绍 .....	4
2.1.1	插值与逼近 .....	4
2.1.2	多层 B 样条插值 .....	4
2.2	样条插值原理 .....	5
2.2.1	B 样条插值 .....	5
2.2.2	多层 B 样条插值 .....	8
2.3	本章小结 .....	10
第 3 章	克里金插值重构方法 .....	11
3.1	克里金插值介绍 .....	11
3.2	克里金插值原理 .....	11
3.3	区域化变量理论 .....	14
3.4	变异函数与模型 .....	15
3.4.1	半方差有序集 .....	15
3.4.2	变异函数模型 .....	16
3.4.3	选取变异函数 .....	17
3.5	本章小结 .....	17
第 4 章	辐射场重构算法设计与实现 .....	18
4.1	辐射场数据获取及存储结构 .....	18
4.2	辐射场重构算法设计 .....	19
4.3	空间辐射场可视化 .....	22
4.4	本章小结 .....	22
第 5 章	辐射场重构方法评价 .....	23

5.1	辐射场重构方法影响因素研究 .....	23
5.1.1	源项数量对辐射场重构效果的影响 .....	23
5.1.2	空间状况对辐射场重构效果的影响 .....	25
5.1.3	测量点数据对辐射场重构效果影响 .....	27
5.2	辐射场重构方法比较 .....	28
5.2.1	辐射场重构相对偏差比较 .....	29
5.2.2	辐射场重构时间比较 .....	30
5.3	本章小结 .....	31
总 结	.....	32
参考文献	.....	33
致 谢	.....	36
附录 克里金法的多变量扩展	.....	37

## 第1章 绪论

### 1.1 研究背景与意义

空间辐射场三维重构方法是辐射场可视化仿真技术的一项重要技术。辐射场可视化仿真技术不仅能够应用在核设施退役工程,将辐射场的分布以可视化的形式显示在虚拟场景中,评估施工人员所接受的辐射剂量<sup>[1]</sup>;还能在核与辐射安全科普工作中发挥积极的作用,将辐射场的剂量值、分布等定量信息展现给公众,易于科普工作者与公众进行互动沟通,提高公众对核科学的认识<sup>[2]</sup>。随着计算机硬件及相关技术的发展,可视化和虚拟现实技术在辐射防护领域得到了迅速发展<sup>[3]</sup>,进而辐射场重构技术的方法研究就显得格外迫切。

总所周知,核辐射是一种弱致癌物,例如氡(Rn)作为天然气态放射性核素,氡及其衰变子体已成为肺癌的第二大诱因<sup>[4]</sup>。根据联合国原子辐射影响科学委员会(UNSCEAR)对氡及其子体的流行病学和计量学审查研究,矿工的肺癌终生风险率每工作一个月提高 $2.4 \sim 7.5 \times 10^{-4}$ (用 WLM 表示, Working Level Month)<sup>[5]</sup>。2015 年德国萨克森辐射防护局检察员 Jörg Dehnert 对矿井内 252 名矿工进行个人剂量监测,其中最高有效年剂量高达  $14.4 \text{ mSv}$ ,并且改进矿井内辐射防护措施:通过对矿井内放射性气体浓度进行监测,根据浓度调节矿井内风机转速<sup>[6]</sup>。放射性浓度高的地区除了矿山矿井,还有地下建筑、窑洞、工业废渣建筑室内以及随着核工业技术的发展,核设施周围都存在不同强度的辐射场。这些辐射场都存在着一些特点:放射性物质分布不明、源项分布或场景结构复杂。

对于复杂的辐射场,目前常用的重构方法有正演方法和反演方法。正演方法是通过辐射传输方程进行求解,在了解放射源基本信息的基础上,构造准确的系统模型进行粒子运输模拟,从而将空间辐射场进行重构;反演方法是在未知放射源基本信息的情况下,通过实际测量获得有限、离散的采样数据进行分析和空间重构,从而获得完整的辐射场分布数据。正演方法目前常用的一些算法有蒙特卡洛法<sup>[7]</sup>和点核积分法<sup>[8]</sup>等,其中蒙特卡洛法是通过随机性方法对辐射传输方程进行求解计算,点核积分法是通过确定论方法对辐射传输方程进行求解。反演方法也指散乱数据重构方法,目前主要有插值<sup>[9]</sup>和逼近,常用的插值算法包括多项式插值、径向基插值、反距离权重插值等,常用的逼近算法包括最小二乘法、最小立方法等。

本文研究的重构方法基于反演方法,在未知源项信息的空间辐射场中,基于测量的有限个离散数据点,利用插值重构方法快速重构出空间辐射场的剂量分布情况。本文研

究方法能够应用于监测工作场所辐射场分布、预估未知空间场源项位置、辅助核设施现场辐射防护计算等，拥有广阔的使用场景和应用领域。

## 1.2 研究方法和内容

空间插值重构方法有多种分类方式<sup>[10]</sup>：按照插值区域范围分，有整体插值、局部插值、边界内插法等。其中整体插值是利用研究区域的所有散乱数据点进行全局特征拟合，采用整体插值方法时，整个区域的数值会影响单个插值点的数值，同理单个散乱数据点的数值增加、减少或删除对整个区域的特征拟合都会造成影响，代表性的整体插值方法有趋势面分析插值方法等；局部插值是利用临近数据点来预测插值点的值：首先定义邻域或搜索范围，然后在该区域内搜索散乱数据点，再对该区域内数据点选择插值函数进行拟合，最后通过计算插值函数得到预测点的值，代表性的局部插值方法有样条插值法、反距离权重插值法和克里金插值法等；边界内插法假设任何数值变化都发生在区域边界上，并且边界的变化是均匀的、同质的，代表性的边界内插方法有泰森多边形法等。

按照插值的标准分，可以分为确定性插值、地统计插值：确定性插值法主要采用数学工具，利用计算插值函数的方法来进行插值，这种方式用来研究某区域内部的相似性，其代表插值法有样条插值法、反距离加权插值法等；地统计插值是基于空间自相关性的，由观测数据产生具有统计关系的曲面，代表插值法有克里金插值法等。按照插值的精度分，可以分为精确插值、近似插值。精确插值重构出包括所有散乱数据点的辐射场；近似插值重构出不包含所有散乱数据点的辐射场。

本论文采用的空间插值重构方法是基于样条插值法和克里金插值法，对于给定的辐射场离散数据，首先分别通过样条插值方法和克里金插值方法重构出相应的辐射场数据；然后对比两种方法重构出的辐射场数据，选取重构结果数据相差较大的区域进行重新采样测量，直到两种方法重构出的辐射场结果大致相同；最后将两种方法重构的辐射场数据进行组合，得到重构空间辐射场分布。

## 1.3 国内外研究现状

相比于基于散乱数据插值方法重构辐射场，国内外当前对辐射场重构研究大部分研究都是基于正演方法。法国原子能委员会(CEA)开发基于虚拟现实技术的辐射剂量评估软件 NARVEOS<sup>[11]</sup>，该软件基于点核积分法对辐射场进行重构；比利时核能研究中心(SCK · CEN)开发基于点核积分法和蒙卡抽样的辐射防护最优化工具 VISIPLAN<sup>[12]</sup>；中科院核安所(FDS)开发了基于蒙特卡洛方法的虚拟仿真平台 SuperMC/RVIS<sup>[13]</sup>。



由于正演法重构辐射场的点核积分法和蒙特卡洛法都必须在了解源项信息以及空间场结构信息后才能求解,而且点核积分法不适用于复杂的源项和空间场,蒙特卡洛方法对屏蔽较厚的场景无法得出可靠的结果以及计算时间过长,近年来,国内外开始研究基于散乱数据插值法的空间辐射场重构方法。俄罗斯科学院 Krasovskii 数学与力学研究所 Aleksey M. Grigoryev 基于径向基函数插值对辐射场进行插值重构并基于重构辐射场计算路线最优化问题<sup>[14]</sup>;英国布里斯托尔大学 Samuel R. White 基于投影线性重建(PLR)算法对源项进行定位<sup>[15]</sup>;中国工程物理研究院赛雪基于 Multiquadric 散乱数据插值方法对辐射场可视化进行研究,证明了基于 Multiquadric 方法对辐射场重构是可行的<sup>[16]</sup>;华南理工大学电力学院王状提出一种基于网格函数插值方法,通过点源辐射场数据进行重构,验证了其方法的可行性<sup>[9]</sup>。

尽管近些年来国内外都有对散乱数据插值方法的辐射场重构进行研究,但是他们都是基于单一的插值方法并且仅在二维空间辐射场上对其方法验证其可行性,而实际情况下辐射场为三维空间,采用三维空间辐射场数据对插值重构方法验证更能体现其可行性。

## 1.4 本论文章节安排

本文通过综合分析空间插值重构方法在各个领域的应用以及国内外研究现状,在传统单一插值重构方法上进行改进,提出结合优化样条插值重构法和克里金插值重构法的新型插值重构方法,通过对两种重构方法的结合,获得更符合实际辐射场的插值重构方法。结果表明,该方法重构出的辐射场相比于传统方法重构出的辐射场更接近实际辐射场。同时,在采用新型插值重构方法时,选取离散采样点时更有趋向性。

第一章主要介绍本论文的研究背景和研究意义以及空间辐射场重构的插值方法,分析国内外空间辐射场重构研究现状,提出本论文的主要研究内容和章节安排。

第二章主要介绍样条插值重构方法,分别对 B 样条插值和多层 B 样条插值的原理进行阐述,给出 BA 算法和 MBA 算法的具体计算流程。

第三章主要介绍克里金插值重构方法,具体阐述了克里金插值方法的具体理论,基于普通克里金模型推导出克里金插值方法的矩阵形式。

第四章提出基于样条插值方法和克里金插值方法的新型插值重构方法,通过程序框图具体说明新型辐射场重构方法的过程。

第五章基于蒙特卡洛应用工具框架 Geant4,模拟几组不同应用场景下的空间辐射场数据,用于验证空间辐射场重构方法的合理可行性,对比其与传统重构方法的重构效果。

最后总结全文,对空间辐射场重构方法的进一步研究提出展望和建设性意见。

## 第2章 样条插值重构方法

### 2.1 样条插值介绍

样条插值(Spline Interpolation)是计算机辅助几何设计(Computer Aided Geometric Design, 简称 CAGD; 亦称计算几何 Computational Geometry<sup>[17]</sup>)中用于解决插值问题的方法<sup>[17]</sup>。相比于其他插值方法, 样条插值更容易在构造整体上达到参数连续阶(指可微性)。

#### 2.1.1 插值与逼近

对于一组散乱数据点 $s_i(i = 1, 2, \dots, n)$ , 要求构造一个函数通过这些数据点, 称为散乱数据插值(Sattered Data Interpolation)<sup>[18]</sup>, 所构造的函数称为插值函数。若这些数据点原来位于某函数上, 则原函数称为被插函数。若所构造的函数为一条曲线, 则称该方法为曲线插值法。同理, 若所构造的函数为曲面, 则称为曲面插值法。

但是在某些情况下, 测量所得的数据点数据存在较大误差, 那么要求构造一个函数严格通过数据点的函数则会增加其误差。因此, 提出一种与插值类似的方法: 构造一个函数使其在某种意义下最接近测量的数据点, 这种方法称为逼近(Approximation), 所构造的函数称为逼近函数。与插值相同, 对于曲线逼近的方法称为曲线逼近法; 对于曲面逼近的方法称为曲面逼近法。

插值和逼近统称为拟合(Fitting)。

#### 2.1.2 多层 B 样条插值

多层 B 样条(Multilevel B-Splines)是由 Seungyong Lee 于 1997 年提出的一种插值方法<sup>[19]</sup>, 该方法利用控制晶格的粗略到精细层次结构来生成一系列双三次 B 样条函数, 这些函数的和等效为一个 B 样条函数, 从而实现插值计算效率和精度的提高。

基于多层 B 样条方法的散乱数据插值在计算机图形学、声学、勘探地球物理学、生物医学工程、地震学中有非常广泛的应用。2017 年澳大利亚国立大学的张浩阳将多层 B 样条方法用于对象掩码注册的深度自由形式变形网络<sup>[20]</sup>; 2019 年澳大利亚联邦大学的 Linh Nguyen 将多层 B 样条方法应用到声源定位中, 并且定位精度和计算成本方面在模拟实验和实际环境中得到验证<sup>[21]</sup>; 2013 年英国伦敦帝国理工学院的姜玉乐和张楠将多层 B 样条插值应用于自然资源勘探的地球物理采集中磁场异常数据的插值并成功实现

[22]; 2011 年美国路易斯维尔大学王慧提出使用包含相位信息的多层 B 样条插值模型来分析标记核磁共振(MR)图像的方法[23]; 2010 年英国伦敦帝国理工学院的张楠和王阳华将多层 B 样条插值法对地震数据进行 2D 和 3D 插值, 取得了良好的准确性和计算效率[24]。

## 2.2 样条插值原理

### 2.2.1 B 样条插值

B 样条方法(Basic Spline)是由 Gordon 与 Riesenfeld 在研究贝齐尔方法的基础上引入的, 它在计算上具有递推性、规范性、局部支承性(非负性)、可微性等优点。

#### 基本思想

若  $\Omega = \{(x, y, z) | 0 \leq x < m, 0 \leq y < n, 0 \leq z < l\}$  为  $xyz$  空间上的定义域, 设三维空间场中有若干离散数据点  $P = \{(x_c, y_c, z_c, v_c)\}$ , 其中  $(x_c, y_c, z_c)$  为定义域  $\Omega$  中的一点。为了通过散乱数据点集  $P$  拟合未知空间场, 构造一个均匀 B 样条函数  $f$ , 样条函数的参数由该定义域上的控制栅格  $\Phi$  求得。通常情况下, 控制栅格由  $(m+3) \times (n+3) \times (l+3)$  个控制点组成。选择不同的控制栅格会影响插值函数的计算。

设  $\phi_{ijk}$  为控制栅格  $\Phi$  上第  $i \times j \times k$  个控制点, 其中  $(i, j, k)$  有  $i = -1, 0, \dots, m+1$ ;  $j = -1, 0, \dots, n+1$ ;  $k = -1, 0, \dots, l+1$ 。插值函数由这些控制点被定义为:

$$f(x, y, z) = \sum_{i=0}^3 \sum_{j=0}^3 \sum_{k=0}^3 B_i(r) B_j(s) B_k(t) \phi_{(a+i)(b+j)(c+k)} \quad (2-1)$$

其中,  $a = [x] - 1, b = [y] - 1, c = [z] - 1, r = x - [x], s = y - [y], t = z - [z]$ 。

$B_i(r)$ 、 $B_j(s)$  和  $B_k(t)$  为均匀三次 B 样条基函数:

$$\begin{aligned} B_0(t) &= (1-t)^3/6 \\ B_1(t) &= (3t^3 - 6t^2 + 4)/6 \\ B_2(t) &= (-3t^3 + 3t^2 + 3t + 1)/6 \\ B_3(t) &= t^3/6 \end{aligned}$$

其中  $t \in [0, 1)$ 。它们根据每个控制点到  $(x, y, z)$  的距离来衡量每个控制点对  $f(x, y, z)$  的影响。

利用三次 B 样条基函数, 插值函数的求解被简化为寻找最接近散乱数据集  $P$  中的控制栅格  $\Phi$  的控制点。

为了确定控制栅格  $\Phi$ , 首先考虑离散数据点集  $P$  中的一个数据点  $(x_c, y_c, z_c, v_c)$ , 由三次样条插值公式 2-1 可知, 样条函数值  $f(x_c, y_c, z_c)$  与其相邻的 64 个控制点  $(x_c, y_c, z_c)$  相

关。例如：假设  $1 \leq x_c, y_c, z_c < 2$ ，则控制点  $\phi_{ijk}(i, j, k = 0, 1, 2, 3)$  决定位置在  $(x_c, y_c, z_c)$  的值  $f$ 。由于函数  $f$  在  $(x_c, y_c, z_c)$  的值为  $v_c$ ，因此控制点  $\phi_{ijk}$  必须满足：

$$v_c = \sum_{i=0}^3 \sum_{j=0}^3 \sum_{k=0}^3 w_{ijk} \phi_{ijk} \quad (2-2)$$

其中  $w_{ijk} = B_i(r)B_j(s)B_k(t)$ ； $r = x_c - 1$ ； $s = y_c - 1$ ； $t = z_c - 1$ 。

存在许多组控制点数据  $\phi_{ijk}$  满足公式 2-2，为了使插值函数  $f$  在定义域上偏差最小化，选取一组最小二乘最小的数  $\sum_{i=0}^3 \sum_{j=0}^3 \sum_{k=0}^3 \phi_{ijk}^2$  作为控制点（这将对多层 B 样条插值有利）。可推导出控制点方程为：

$$\phi_{ijk} = \frac{w_{ijk} v_c}{\sum_{d=0}^3 \sum_{e=0}^3 \sum_{g=0}^3 w_{deg}^2} \quad (2-3)$$

从公式 2-3 可以看出，距离插值点  $(x_c, y_c, z_c)$  越近的控制点越能决定插值点的值，因为离插值点越近的控制点权重  $w_{ijk}$  越大。计算得样条插值函数  $f$  在  $(x_c, y_c, z_c)$  的值为  $v_c$ ，在插值点周围函数变化平滑。

现在，考虑散乱数据点集  $P$  中所有数据点，对于每个数据点  $p_i$ ，可以通过公式 2-3 确定其领域内  $4 \times 4 \times 4$  个控制点的值。若存在两个数据点距离较近可能存在控制栅格重叠的部分，对于这种情况可以对其共享的几个控制点分配不同的值。通常情况下，综合考虑控制点  $\phi$  的  $4 \times 4 \times 4$  领域内的所有数据点来解决对控制点  $\phi$  的多重分配，只有这些点可以通过公式 2-3 来影响  $\phi$  的值，将这组点称为  $\phi$  的临近数据集。

设  $P_{ijk}$  为控制点  $\phi_{ijk}$  的临近数据集，则：

$$P_{ijk} = \{(x_c, y_c, z_c, v_c) \in P | i-2 \leq x_c < i+2, j-2 \leq y_c < j+2, k-2 \leq z_c < k+2\}$$

对于每个散乱数据点  $P_{ijk} = (x_c, y_c, z_c, v_c)$ ，由公式 2-3 可以得出控制点  $\phi_{ijk}$  的值  $\phi_c$ ：

$$\phi_c = \frac{w_c v_c}{\sum_{d=0}^3 \sum_{e=0}^3 \sum_{g=0}^3 w_{deg}^2} \quad (2-4)$$

其中  $w_c = w_{ijk} = B_i(r)B_j(s)B_k(t)$ ； $i = (a+1) - [x_c]$ ； $j = (b+1) - [y_c]$ ； $k = (b+1) - [z_c]$ ； $r = x_c - [x_c]$ ； $s = y_c - [y_c]$ ； $t = z_c - [z_c]$ 。最终计算选取误差  $e(\phi_{ijk}) = \sum_c (w_c \phi_{ijk} - w_c \phi_c)^2$  最小的值作为控制点值  $\phi_{ijk}$ ，其中  $(w_c \phi_{ijk} - w_c \phi_c)$  是控制点  $\phi_{ijk}$  对插值函数  $f$  在  $(x_c, y_c, z_c)$  实际权重与真实权重之差。

换句话说， $e_{\phi_{ijk}}$  是控制点  $\phi_{ijk}$  的近似误差，用不同数据点对该控制点的权重来计算控制点值，可以得到：

$$\phi_{ijk} = \frac{\sum_c w_c^2 \phi_c}{\sum_c w_c^2} \quad (2-5)$$

临近数据集  $P_{ijk}$  是数据集  $P$  中在控制点  $\phi_{ijk}$  附近对插值函数  $f$  有影响的点的集合。当  $P_{ijk}$  包含多个点时，公式 2-5 通过每个数据点权重的最小平方解来求得局部逼近误差最

小的控制点值  $\phi_{ijk}$ 。当临近数据集中仅有一个数据点时，采用公式 2-3 计算控制点值，其不存在近似误差。但是，当临近数据集中没有数据点时，数据点对控制点的值起不到影响，则控制点  $\phi_{ijk}$  可以设置为任意值，通常设置为 0 或者  $v_c$  的平均值。

### 算法

利用公式 2-5，由离散数据点集  $P$  中的数据点确定控制栅格  $\Phi$ ，可知不需要显式地识别每个控制点地临近数据集。因为  $P$  中的每个数据点所影响  $\Phi$  中  $4 \times 4 \times 4$  个相邻控制点的集合，所以它只属于这些控制点的临近数据集。因此，可以对每个控制点依次考虑每个数据点，对公式 2-5 的分子分母进行有效累加。如果分母不为零，则直接通过公式计算得出，只有当控制点的临近数据集为空时，才会出现分母为零，这种情况下将控制点值设为零。图为 B 样条插值方法的伪代码，称之为 BA 算法。

表 2-1 BA 算法

---

**Algorithm 1** BA 算法

---

**Require:** 散点数据集  $P = \{x_c, y_c, z_c, v_c\}$

**Ensure:** 控制栅格  $\Phi = \{\phi_{ijk}\}$

---

```

1: for all  $i, j$  do
2:    $\delta_{ijk} = 0$  and  $\omega_{ijk} = 0$ 
3: end for
4: for each point  $(x_c, y_c, z_c, v_c)$  in  $P$  do
5:   let  $i = [x_c] - 1$  and  $j = [y_c] - 1$  and  $k = [z_c] - 1$ 
6:   let  $r = x_c - [x_c]$  and  $s = y_c - [y_c]$  and  $t = z_c - [z_c]$ 
7:   compute  $w_{ijk}$ 's and  $\sum_{d=0}^3 \sum_{e=0}^3 \sum_{g=0}^3 w_{deg}^2$ 
8:   for  $i, j, k = 0, 1, 2, 3$  do
9:     compute  $\phi_{ijk}$  with Formula 2-5
10:    add  $w_{ijk}^2 \phi_{ijk}$  to  $\delta_{(a+i)(b+j)(c+k)}$ 
11:    add  $w_{ijk}^2$  to  $\omega_{(a+i)(b+j)(c+k)}$ 
12:   end for
13: end for
14: for all  $i, j$  do
15:   if  $\omega_{ijk} \neq 0$  then
16:     compute  $\phi_{ijk} = \delta_{ijk} / \omega_{ijk}$ 
17:   else let  $\phi_{ijk} = 0$ 
18:   end if
19: end for
    
```

---

BA 算法的时间复杂度和空间复杂度为  $O\{p + mnl\}$ ，其中  $p$  为散乱数据点集中散乱点的个数， $(m + 3) \times (n + 3) \times (l + 3)$  为控制栅格的大小。尽管控制点的值是局部确定的，但是要尽量减小插值误差，使结果函数尽可能地重构出分散的数据。因此，选取控制栅格  $\Phi$  的疏密程度，直接影响到插值函数的重构效果，对于控制点稀疏的控制栅格，

插值函数重构更加平滑；对于控制点密集的控制栅格，插值函数重构更加陡峭。

由于 BA 算法时间复杂度较低，即使在数据点集很大的情况下，通过 BA 算法也能很快得重构出插值函数。此外，由于 B 样条函数具有规范性，所以在添加或删除数据点时，只需要更新控制栅格中一个小的领域。由 BA 算法生成的插值函数  $f$  是  $C^2$  连续的，因为它是由控制栅格  $\Phi$  生成的三次 B 样条空间。因为  $\Phi$  可以很快地从散乱数据点中计算得出，所以计算函数  $f$  的大部分时间是用来在定义域  $\Omega$  上从控制栅格  $\Phi$  计算插值函数  $f$ 。

### 2.2.2 多层 B 样条插值

通过 BA 算法生成的插值函数在形状平滑度和精度之间存在折衷，因此，提出多层 B 样条插值算法来避免这种折衷。多层 B 样条算法通过控制栅格的多层结构来生成插值函数  $f_k$  序列，对插值函数序列  $f_k$  每个函数进行加和，得到最终插值函数。在该序列中，来自稀疏栅格的函数提供粗略的近似，该近似在精度上被来自更精细栅格的函数进一步细化，最终将这些函数的和减少到一个等价的 B 样条函数。

#### 基本算法

考虑到定义域  $\Omega$  上存在多层控制栅格  $\Phi_0, \Phi_1, \dots, \Phi_h$ ，假设  $\Phi_0$  的控制点之间的间隔已知，并且从一个栅格到另一个栅格的间隔减半。因此，若  $\Phi_k$  栅格存在  $(m+3) \times (n+3) \times (l+3)$  个控制点，则下一层控制栅格  $\Phi_{k+1}$  则有  $(2m+3) \times (2n+3) \times (2l+3)$  个控制点。 $\Phi_k$  中第  $ijk$  个控制点位置与  $\Phi_{k+1}$  中第  $(2i, 2j, 2k)$  个控制点位置重合。

多层 B 样条插值方法首先将 BA 算法应用于最粗的控制栅格  $\Phi_0$ ，所得到的函数  $f_0$  作为平滑的初始近似函数。 $f_0$  作为近似函数在散乱数据点集中每个数据点上有一个偏差  $\Delta^1 v_c = v_c - f_0(x_c, y_c, z_c)$ 。在下一层控制栅格  $\Phi_1$  中，这个偏差作为近似差  $\mathbf{P}_1 = \{x_c, y_c, z_c, \Delta^1 v_c\}$ ，对改近似差进行插值，求得近似函数  $f_1$ ，得到较为精确的重构场  $f_0 + f_1$  以及  $P$  中每个数据点的一个更小偏差  $\Delta^2 v_c = v_c - f_0(x_c, y_c, z_c) - f_1(x_c, y_c, z_c)$ 。

对于第  $k$  层，使用近似差数据集  $\mathbf{P}_k = \{(x_c, y_c, z_c, \Delta^k v_c)\}$  计算出第  $k$  层控制栅格  $\Phi_k$ ，从而推导近似函数  $f_k$ ，其中  $\Delta^k v_c = v_c - \sum_{i=0}^{k-1} f_i(x_c, y_c, z_c) = \Delta^{k-1} v_c - f_{k-1}(x_c, y_c, z_c)$ ， $\Delta^0 v_c = v_c$ 。该计算过程从初始最粗控制栅格  $\Phi_0$  开始，一直计算到最精细栅格  $\Phi_h$ 。最后得到的近似函数  $f$  为序列函数  $f_k$  之和，即  $f = \sum_{k=0}^h f_k$ 。通过多层 B 样条这种方式，得到的是一个平滑且接近于  $P$  的近似解。以下为多层 B 样条拟合的基本算法，称之为 MBA 算法。

表 2-2 MBA 算法

**Algorithm 1** MBA 算法

**Require:** 散点数据集  $\mathbf{P} = \{x_c, y_c, z_c, v_c\}$

**Ensure:** 多层控制栅格  $\Phi_0, \Phi_1, \dots, \Phi_h$

```

1: let  $k = 0$ 
2: while  $k \leq h$  do
3:   let  $\mathbf{P}_k = \{(x_c, y_c, z_c, \Delta^k v_c)\}$ 
4:   compute  $\Phi_k$  from  $\mathbf{P}_k$  by BA algorithm
5:   compute  $\Delta^{k+1} v_c = \Delta^k v_c - f_k(x_c, y_c, z_c)$  for each data point
6:   let  $k = k + 1$ 
7: end while

```

设散乱数据点集  $P$  中有  $p$  个数据点，最精细控制栅格  $\Phi_h$  的大小为  $(m+3) \times (n+3) \times (l+3)$ ，第  $k$  个控制栅格控制点数量为第  $k+1$  个控制栅格的八分之一。因此，MBA 算法的时间复杂度为  $O(p + mnl) + O(p + \frac{1}{8}mnl) + \dots + O(p + \frac{1}{8^h}mnl) = O(hp + \frac{8}{7}mnl)$ ，空间复杂度为  $O(p + \frac{8}{7}mnl)$ 。通过 MBA 算法得到的近似函数是  $C^2$  连续的，因为它是通过多个  $C^2$  连续的样条函数加和得到的。

### 优化算法

多层 B 样条插值算法(MBA 算法)通过生成多个控制栅格层数，来求得近似重构函数  $f$ 。为了求得  $f$ ，必须求出每层控制栅格  $\Phi_k$  的控制点值，从而确定该层的近似函数  $f_k$ ，再将它们的值相加从而得到重构函数  $f$ 。如果  $f$  在定义域  $\Omega$  中有大量的散乱数据点，则会出现大量的计算时间开销。为此，提出一种在控制栅格中逐步应用 B 样条优化来解决这个问题，使得  $f$  可以用一个 B 样条函数来表示，而不是几个 B 样条函数之和。因此， $f_k$  的计算仅在  $\Phi_k$  中少量的控制点，而不是  $\Omega$  中的所有点。

设  $F(\Phi)$  为控制栅格  $\Phi$  生成的 B 样条函数，设  $|\Phi|$  的大小，则通过 B 样条优化可以从最粗的控制栅格  $\Phi_0$  推导出控制栅格  $\Phi'_0$ ，如  $F(\Phi'_0) = f_0, |\Phi'_0| = |\Phi_1|$ ，那么，函数  $f_0$  和  $f_1$  的和可以由控制栅格  $\Psi_1$  计算得出，该控制栅格是由  $\Phi_0$  和  $\Phi_1$  中的每对对应的控制点相加得到的。也就是说， $F(\Psi_1) = g_1 = f_0 + f_1$ ，其中  $\Psi_1 = \Phi_0 + \Phi_1$ 。

设  $g_k = \sum_{i=0}^k f_i$  是层次中直到第  $k$  层的函数  $f_i$  的部分和，假设函数  $g_{k-1}$  是由控制栅格  $\Psi_{k-1}$  计算得出， $\Psi_{k-1} = \Phi_{k-1}$ 。和上面计算  $\Psi_1$  一样，可以通过优化  $\Psi_{k-1}$  得到  $\Psi'_{k-1}$ ，并将  $\Psi'_{k-1}$  加上  $\Phi_k$  得到  $\Psi_k$ ，得到  $F(\Psi_k) = g_k, |\Psi_k| = |\Phi_k|$ ，也就是  $\Psi_k = \Psi'_{k-1} + \Phi_k$ 。因此，从  $g_0 = f_0, \Psi_0 = \Phi_0$  可以计算出一系列的控制栅格  $\Psi_k$ ，最终计算出最后控制栅格

格  $\Psi_k$  来得到最终的近似函数  $f = g_k$ 。下面给出 MBA 优化算法的伪代码。

表 2-3 MBA 优化算法

---

**Algorithm 1** MBA 优化算法

---

**Require:** 散点数据集  $\mathbf{P} = \{x_c, y_c, z_c, v_c\}$

**Ensure:** 控制栅格  $\Psi$

```

1: let  $\Phi = \Phi_0$ 
2: let  $\Psi' = 0$ 
3: while  $\Phi \neq \Phi_h$  do
4:   compute  $\Phi$  from  $\mathbf{P}$  by the BA algorithm
5:   compute  $\mathbf{P} = \mathbf{P} - F(\Phi)$ 
6:   compute  $\Psi = \Psi' + \Phi$ 
7:   let  $\Phi_k = \Phi_{k+1}$ 
8:   refine  $\Psi$  into  $\Psi'$  whereby  $F(\Psi') = F(\Psi)$  and  $|\Psi'| = |\Phi|$ 
9: end while
    
```

---

设散乱数据点集  $P$  中有  $p$  个数据点, 最终控制栅格的大小为  $(m+3) \times (n+3) \times (l+3)$ , 由于优化 MBA 算法所需的计算数量与控制点的数量成线性关系, 因此 MBA 优化算法的时间复杂度为  $O\left(lp + \frac{8}{7}mnl\right)$ , 空间复杂度为  $O(p + mnl)$ 。如果层级结构的深度比控制点数量少得多, 那么 MBA 优化算法运行时间复杂度和空间复杂度与最稠密控制栅格的 BA 算法相同, 但是 MBA 优化算法生成的函数要平滑得多。

## 2.3 本章小结

本章首先介绍了样条插值方法的由来以及优势, 给出插值和逼近的区别进而介绍多层 B 样条插值的由来以及该方法在诸多领域的应用; 然后详细介绍了样条函数的插值原理, 给出 BA 算法的具体思想和伪代码; 最后在 BA 算法的基础上介绍 MBA 算法和 MBA 优化算法, 并分别给出其伪代码。



## 第3章 克里金插值重构方法

### 3.1 克里金插值介绍

“Kriging”一词在空间统计中的使用已经成为“最佳预测”的同义词<sup>[25]</sup>，应用于已知测量点附近位置的预测。20世纪50年代，南非金矿开采工程师 D.G. Krige 提出了几种统计方法来预测平均金矿品位。受他方法的启发，法国数学家马瑟顿（G. Matheron）在20世纪60年代为随机场数据首次提出了一种基于回归的空间预测方法。马瑟顿为感激 Krige 提出的想法创造 Kriging 一词，以表彰 Krige 的开拓性工作<sup>[26]</sup>。

克里金方法是一种地质统计学的预测方法，地质统计学最初是在采矿业中发展起来的，现在广泛应用于环境科学——陆地、大气和海洋<sup>[27]</sup>。环境科学研究者需要绘制出带有属性(降雨量、土壤营养等)的地图，这些地图上有或多或少的连续区域，但是通常只有稀疏的样本数据。在这样的情况下，研究者们使用地质统计学方法来预测更大的区域(一维、二维和三维)空间上样本点的值。

克里金方法在地质统计学中有非常广泛的应用。正如刚才所提到的，克里金方法的提出就是在南非威特沃特斯兰德金矿的采矿工程师 Krige 在金矿开采方面的应用，用于估计当地金属浓度来获取更高的利润。在英法海底隧道建设时，Chilès 就使用 Kriging 插值方法预测地层顶部和底部的深度，绘制泥灰岩的形态及其上下限，随着隧道的钻凿，工程师们发现地层上下限实际情况与预测精度相一致<sup>[28]</sup>。

克里金方法不仅仅应用于地质统计学，在污染生态学、精准农业以及渔业学也都有比较多的应用。2014年中国科学院地理科学与资源研究所钟卜青、陶亮使用指标克里金法(IK)和多变量克里金法(MVIK)估算中国西南部喀斯特地区重金属超过有限混合物分布模型(FMDM)阈值的概率<sup>[29]</sup>；2012年 A.Castrignanò 利用共同克里金法和因子克里金法对橄榄种植区上橄榄果蝇种群密度的多变量空间(海拔)和时间(周期)数据进行分析，生成专题图描绘监测区域<sup>[30]</sup>；2010年 Charles F.Adams 收集1999年至2007年在美国 Nantucket 岛附近的两个区域扇贝量，通过 Kriging 方法对物种密度进行估计，并绘制该区域内的扇贝分布数量，确保该区域的扇贝种群不会枯竭<sup>[31]</sup>。

### 3.2 克里金插值原理

Kriging 方法是对于连续空间变化的随机模型，通过稀疏样本数据来预测未测量站点值的一种方法，其插值由先验协方差控制的高斯过程建模。在先验的适当假设下，克

里金法可为中间值提供最佳的线性无偏预测。普通克里格法不需要过多额外信息，只需要这些信息加上测量值和它们的地理坐标。这是目前为止最流行的克里金方法；它在大多数情况下都能够应用，因为它的应用条件很容易得到。

普通克里金方法是基于这样的假设：连续空间变化是随机的并且与空间相关。由于随机过程具有恒定的平均值与方差，因此预测点的值取决于相对位置，而与绝对位置无关。于是，我们把目标值分为确定性趋势值以及随机的自相关函数值：

$$V(s) = \mu(s) + \varepsilon(s) \quad (3-1)$$

其中 $V(s)$ 是目标值， $\mu(s)$ 是确定性趋势值， $\varepsilon(s)$ 是随机的自相关函数值， $s$ 为标识位置，在三维空间中可视为空间 $x$  (经度)、 $y$  (纬度)、 $z$  (高度)坐标。

假设有 $N$ 个随机变量 $V$ 的值已经记录在采样点 $s_1, s_2, s_3, \dots, s_N$ 作为已知点，对于克里金插值，我们通过以下公式预测其插值点 $s_0$ 的值：

$$\hat{V}(s_0) = \sum_{i=1}^N \lambda_i V(s_i) \quad (3-2)$$

其中， $\lambda_i$ 为权重。为了确保估计值无偏以及权重之和为1，有：

$$E[\hat{V}(s_0) - V(s_0)] = 0 \quad (3-3)$$

$$\sum_{i=1}^N \lambda_i = 1 \quad (3-4)$$

预测方差为：

$$\begin{aligned} Var[\hat{V}(s_0)] &= E[\hat{V}(s_0) - V(s_0)]^2 \\ &= 2 \sum_{i=1}^N \lambda_i \gamma(s_i - s_0) - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma(s_i - s_j) \end{aligned} \quad (3-5)$$

其中，函数 $\gamma(s_i - s_0)$ 代表采样点 $s_i$ 和目标预测点 $s_0$ 之间的半方差函数(也称半变异函数)； $\gamma(s_i - s_j)$ 是采样点 $s_i$ 和 $s_j$ 之间的半方差函数。

半方差函数是从变异函数模型中导出的，一方面是因为没有观测值的数据点和目标点之间没有半方差的度量，另一方面原因是只有这样才能保证方差不是负的。如果一个目标点碰巧也是一个采样点，那么克里金方法直接返回那里的观测值，估计方差为零。

前面所提到的克里金技术都是在预测特定未采样位置的变量值，这些位置可以被视为空间点，因此，这种克里金法也被称为点克里金法。当不确定性相对较大时，可能需要通过比单个点更大的区域上执行克里金法来得到平滑插值结果，这样的方法称为块克里金法。相比于点克里金法，块克里金法具有降低空间预测误差的优势，但是它可能丢

失一些有用信息。但是为了得到更好的插值重构效果，块克里金表现的更好。

与点克里金法插值公式相似，块克里金法的任意块估计值也是测量值的加权平均值：

$$\hat{V}(B) = \sum_{i=1}^N \lambda_i V(s_i) \quad (3-6)$$

与点克里金一样，块克里金公式中的 $\lambda_i$ 之和为1，其预测值方差为：

$$\begin{aligned} \text{Var}[\hat{V}(\mathbf{B})] &= E[\hat{V}(\mathbf{B}) - V(\mathbf{B})^2] \\ &= 2 \sum_{i=1}^N \lambda_i \gamma(\mathbf{s}_i, \mathbf{B}) - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma(\mathbf{s}_i - \mathbf{s}_j) - \bar{\gamma}(\mathbf{B}, \mathbf{B}) \end{aligned} \quad (3-7)$$

其中， $\gamma(s_i, B)$ 是数据点 $s_i$ 和目标块 $B$ 之间的平均半方差函数； $\bar{\gamma}(B, B)$ 是目标块 $B$ 内的平均半方差，即块内方差。

下一步，要在权重之和为1的约束下，寻找使得预测值方差最小的权重。通过公式 3-2 减去公式 3-5，可推导出  $N+1$  个公式，其中包含  $N+1$  个未知数：

$$\sum_{i=1}^N \lambda_i \gamma(\mathbf{s}_i - \mathbf{s}_j) + \psi(\mathbf{s}_0) = \gamma(\mathbf{s}_j - \mathbf{s}_0) \quad \text{for all } j \quad (3-8)$$

$$\sum_{i=0}^N \lambda_i = 1 \quad (3-9)$$

其中， $\psi(s_0)$ 为拉格朗日数乘。

展开作为方程组形式为：

$$\begin{cases} \gamma(\mathbf{s}_1 - \mathbf{s}_1)\lambda_1 + \gamma(\mathbf{s}_1 - \mathbf{s}_2)\lambda_2 + \cdots + \gamma(\mathbf{s}_1 - \mathbf{s}_n)\lambda_n - \psi(\mathbf{s}_0) = \gamma(\mathbf{s}_1 - \mathbf{s}_0) \\ \gamma(\mathbf{s}_2 - \mathbf{s}_1)\lambda_1 + \gamma(\mathbf{s}_2 - \mathbf{s}_2)\lambda_2 + \cdots + \gamma(\mathbf{s}_2 - \mathbf{s}_n)\lambda_n - \psi(\mathbf{s}_0) = \gamma(\mathbf{s}_2 - \mathbf{s}_0) \\ \cdots \\ \gamma(\mathbf{s}_n - \mathbf{s}_1)\lambda_1 + \gamma(\mathbf{s}_n - \mathbf{s}_2)\lambda_2 + \cdots + \gamma(\mathbf{s}_n - \mathbf{s}_n)\lambda_n - \psi(\mathbf{s}_0) = \gamma(\mathbf{s}_n - \mathbf{s}_0) \\ \lambda_1 + \lambda_2 + \cdots + \lambda_n = 1 \end{cases}$$

由克里金方程中每个测量点所占的权重，可将预测方差公式变为：

$$\sigma^2(\mathbf{s}_0) = \sum_{i=1}^N \lambda_i \gamma(\mathbf{s}_i - \mathbf{s}_0) + \psi(\mathbf{s}_0) \quad (3-10)$$

克里金方程也可以写成矩阵形式：

$$A\lambda = b \quad (3-11)$$

其中，矩阵 $A$ 表示第 $i$ 个取样点和第 $j$ 个取样点之间的半方差； $\lambda$ 表示权重和拉格朗日乘子向量； $b$ 表示每个采样点和目标点之间的半方差向量。

矩阵的具体形式为：

$$\begin{bmatrix} \gamma(\mathbf{s}_1 - \mathbf{s}_1) & \gamma(\mathbf{s}_1 - \mathbf{s}_2) & \cdots & \gamma(\mathbf{s}_1 - \mathbf{s}_n) & 1 \\ \gamma(\mathbf{s}_2 - \mathbf{s}_1) & \gamma(\mathbf{s}_2 - \mathbf{s}_2) & \cdots & \gamma(\mathbf{s}_2 - \mathbf{s}_n) & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma(\mathbf{s}_n - \mathbf{s}_1) & \gamma(\mathbf{s}_n - \mathbf{s}_2) & \cdots & \gamma(\mathbf{s}_n - \mathbf{s}_n) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \\ -\psi(\mathbf{s}_0) \end{bmatrix} = \begin{bmatrix} \gamma(\mathbf{s}_1 - \mathbf{s}_0) \\ \gamma(\mathbf{s}_2 - \mathbf{s}_0) \\ \cdots \\ \gamma(\mathbf{s}_n - \mathbf{s}_0) \\ 1 \end{bmatrix}$$

矩阵 $A$ 可逆，权重和拉格朗日乘子向量可表示为：

$$\lambda = A^{-1}b \quad (3-12)$$

矩阵形式计算预测方差为：

$$\widehat{\sigma^2}(s_0) = b^T \quad (3-13)$$

相同的，块克里金法也可按照以上推导，可得出计算公式：

$$\sum_{i=1}^N \lambda_i \gamma(s_i - s_j) + \psi(B) = \bar{\gamma}(s_j, B) \quad \text{for all } j \quad (3-14)$$

$$\sum_{i=1}^N \lambda_i = 1 \quad (3-15)$$

块克里金预测方差为：

$$\sigma^2(\mathbf{B}) = \sum_{i=1}^N \lambda_i \bar{\gamma}(\mathbf{s}_i, \mathbf{B}) + \psi(\mathbf{B}) - \bar{\gamma}(\mathbf{B}, \mathbf{B}) \quad (3-16)$$

在块克里金法的矩阵表示中， $b$ 表示每个采样点和目标块之间的半方差向量，块克里金预测方差可表示为：

$$\widehat{\sigma^2}(B) = b^T \lambda - \hat{\gamma}(B, B) \quad (3-17)$$

通常情况下，块克里金法的预测方差都小于点克里金法，因为任何块克里金方差都完全包含在块内方差中。块克里金预测之间的波动也比点克里金预测之间的波动小一些，因此块克里金法生成的辐射场要比点克里金法生成的辐射场更加平滑。

上面描述的克里金方程中充分的表现出了区域化变量理论以及变异函数与模型的重要性，因此，接下来的部分仔细讲述区域化变量理论和变异函数与模型的内容。

### 3.3 区域化变量理论

平稳性是克里金方法实用性的基础，假设插值区域数据视为有一定趋势的确定性变化，可以将插值空间场的数值分为两个部分，即确定性趋势函数值和随机自相关函数值：

$$V(s) = \mu(s) + \varepsilon(s) \quad (3-18)$$

其中,  $\mu(s)$  为确定性函数变量;  $\varepsilon(s)$  为均值为0, 协方差为  $C(h)$  的随机自相关变量, 其中  $h$  是空间中的相对距离, 称为滞后。

协方差表达式为:

$$C(h) = E[\varepsilon(s)\varepsilon(s+h)] \quad (3-19)$$

当  $\mu(s)$  为常数时(也就是普通克里金方法,  $\mu(s) = \mu(\text{常数})$  时), 协方差函数可以表达为:

$$C(h) = E[\{V(s) - \mu\}\{V(s+h) - \mu\}] = E[\{V(s)\}\{V(s+h)\} - \mu^2] \quad (3-20)$$

其中,  $V(s)$  和  $V(s+h)$  表示空间场物理变量在  $s$  和  $s+h$  的值,  $E$  表示期望。

从公式 3-20 看出, 协方差仅与  $h$  有关, 即样本之间在距离和方向上的间隔。这是基于二阶平稳性的假设, 在实际空间场插值过程中, 通常不能假设平均值是常数, 否则方差就不存在。实际情况中, 通常将平稳性假设改为 Matheron 所提出的内在平稳性假设<sup>[32]</sup>, 即期望差为0:

$$E[V(s) - V(s+h)] = 0 \quad (3-21)$$

从而可以推导出变异函数与预测值方差的关系:

$$\text{Var}[V(s) - V(s+h)] = E[\{V(s) - V(s+h)\}^2] = 2\gamma(h) \quad (3-22)$$

其中,  $\gamma(h)$  是滞后  $h$  的半方差, 它是与相对位置  $h$  有关的变异函数。

对于二阶平稳状态, 协方差函数和变异函数是等价的:

$$\gamma(h) = C(0) - C(h) \quad (3-23)$$

其中,  $C(0) = \sigma^2$  表示方差。

因此, 对于普通克里金模型, 半方差函数可以看成:

$$\begin{aligned} \gamma(h) &= \frac{1}{2} E[\{\varepsilon(s) - \varepsilon(s+h)\}^2] \\ &= \frac{1}{2} E[\{V(s) - V(s+h)\}^2] \end{aligned} \quad (3-24)$$

### 3.4 变异函数与模型

变异函数是许多地质统计学应用的基石。在统计学应用中, 变异函数和任何与之适应的模型都应该是准确的, 只有这样, 重构出的数据才能更好地符合实际。克里金法需要使用一个变异函数, 来得到最小的克里金预测方差。

#### 3.4.1 半方差有序集

获得变异函数地第一步是通过已知的测量点数据  $V(s_1), V(s_2), \dots$ , (其中  $s_1, s_2, \dots$  表示

样本在三维空间中的位置)来估算变异函数。在估算变异函数之前, 需要保证选取的这些样本点是随机的, 因为在定义变异函数时我们认为变量为随机过程的结果。

估算变异函数的常用方法有 Matheron 矩量法估计(MoM):

$$\hat{\gamma}(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} \{v(s_i) - v(s_i + h)\}^2 \quad (3-25)$$

其中,  $v(s_i)$ 和 $v(s_i + h)$ 为位置在 $s_i$ 和 $s_i + h$ 的测量值;  $m(h)$ 为在滞后 $h$ 处的成对比较次数。

通过改变 $h$ , 可以得到半方差的有序集, 这些构成估计变异函数的数据。

### 3.4.2 变异函数模型

变异函数模型主要可以分为两类——有界函数和无界函数。最常用的三个变异函数模型为: 幂函数(无界)、球函数(有界)和指数函数(渐进有界)。如果常用的变异函数不能够符合实际值, 也可以拟合更加复杂的函数。

下面介绍四种常见的变异函数:

**指数函数** 指数函数是一种无界函数:

$$\gamma(h) = c_0 + gh^\beta \quad (0 < \beta < 2) \quad (3-26)$$

**球函数** 球函数是一种有界函数(分段函数):

$$\gamma(h) = \begin{cases} c_0 + c \left\{ \frac{3h}{2r} - \frac{1}{2} \left( \frac{h}{r} \right)^3 \right\} & (0 < h \leq r) \\ c_0 + c & (h > r) \\ c & (h = 0) \end{cases} \quad (3-27)$$

其中,  $c_0$ 为块金方差;  $c$ 为空间相关程度的方差;  $r$ 为空间范围。

**指数函数** 指数函数是一种渐进有界函数:

$$\gamma(h) = \begin{cases} c_0 + c \left\{ 1 - \exp\left(-\frac{h}{a}\right) \right\} & (0 < h) \\ 0 & (h = 0) \end{cases} \quad (3-28)$$

其中,  $a$ 为距离参数。该函数模型渐进地接近平滑, 它的定义域为无穷, 但是为了接近实际, 通常定义一个有效距离范围。

**嵌套球形函数** 嵌套球形函数是在不同范围采用不同的球函数:

$$\gamma(h) = \begin{cases} c_0 + c_1 \frac{3h}{2r_1} - \frac{1}{2} \left( \frac{h}{r_1} \right)^3 + c_2 \frac{3h}{2r_2} - \frac{1}{2} \left( \frac{h}{r_2} \right)^3 & (0 < h \leq r_1) \\ c_0 + c_1 + c_2 \frac{3h}{2r_2} - \frac{1}{2} \left( \frac{h}{r_2} \right)^3 & (r_1 < h \leq r_2) \\ c_0 + c_1 + c_2 & (h > r_2) \\ 0 & (h = 0) \end{cases} \quad (3-29)$$

其中,  $c_1, r_1$  为近距离分量的空间相关程度方差和范围;  $c_2, r_2$  为远距离分量的空间相关程度方差和范围。

### 3.4.3 选取变异函数

在地质统计学中, 变异函数拟合模型仍然存在争议, 尽管它是克里金插值最重要的步骤之一。一些工程师们凭借肉眼来选取拟合模型, 这可能导致半方差在点与点之间波动过大, 并且它的准确性是不稳定的。因此, 常常会有些选取指标来辅助我们来选取变异函数模型, 例如 RSS(残差平方和)、AIC(Akaike 信息准则)。选取变异函数模型通常经过以下几个步骤:

1. 将公式 3-25 计算出的半方差有序集, 作为散点绘制在坐标图中;
2. 选择几个形状相似的模型, 用加权最小二乘法依次进行拟合, 绘制拟合曲线;
3. 评估选取的变异函数模型是否合理。若选取的模型都拟合得很好, 则选择 RSS 最小的模型; 若模型参数个数不相等, 则选择 AIC 最少的模型。

AIC 估计公式为:

$$AIC = \{n \ln \left( \frac{2\pi}{n} \right) + n + 2\} + n \ln R + 2p \quad (3-30)$$

其中,  $n$  为半方差有序集中点的个数;  $p$  为模型参数的个数;  $R$  为残差的均方。大括号里的表达式, 对于任何选取模型都是常数, 因此, 判断 AIC 只需要计算:

$$\widehat{AIC} = n \ln R + 2p \quad (3-31)$$

## 3.5 本章小结

本文详细介绍了克里金插值重构方法的主要内容, 首先阐述了克里金方法的由来以及现应用的主要领域; 然后详细推导了克里金插值方法的插值原理; 再分别说明了区域化变量理论和变异函数和变异函数模型的具体内容; 最后列举了几种常用的克里金多变量扩展模型(附录)。

## 第4章 辐射场重构算法设计与实现

### 4.1 辐射场数据获取及存储结构

辐射场数据获取方式包括实验测量和计算机模拟。对于  $\gamma$  空间辐射场测量, 通常采用便携式  $\gamma$  剂量率检测仪, 但实验测量获取辐射场数据在考虑测量人员辐照量的情况下, 实验空间内距离放射源较远的位置剂量率难以与本底区分, 因此本论文用于验证空间辐射场重构方法所获取的辐射场数据, 全部采用计算机模拟方法获得。

目前, 计算机模拟方法中使用最多的是蒙特卡洛法(Monte Carlo method)<sup>[33]</sup>, 该方法也称为统计模拟方法, 其方法是以概率统计理论为基础, 获得模拟参数近似解的数值计算方法。该方法最早在核科学领域的应用是第二次世界大战时期, 美国洛斯阿拉莫斯国家实验室(Los Alamos National Laboratory, 简称 LANL)使用蒙特卡洛方法解决核武器研发和制造过程中核燃料中子随机扩散概率问题。经过多年的发展, 目前 LANL 已经开发出了第六代蒙特卡洛粒子输运代码(Monte Carlo N-Particle, 简称 MCNP), 该仿真工具为辐射防护和剂量测定、辐射屏蔽、射线照相、医学物理学、核临界安全性、探测器设计和分析、核油测井、加速器目标设计、裂变和聚变反应堆的设计、净化及退役等应用领域提供了必不可少的数值模拟。同样, 欧洲核子研究组织(CERN)也在 20 世纪 90 年代开发蒙特卡洛应用软件框架 Geant4, 其用于模拟粒子在物质中输运的物理过程, 由于其开源的特点, Geant4 发展至今已开发出 10.7 版本。Geant4 在涉及微观粒子与物质相互作用的诸多领域获得了广泛应用, 例如空间应用、微电子学、辐射医学、屏蔽计算。

本论文辐射场数据通过 Geant4 数值模拟获取, 基于 C++语言, 利用面向对象程序设计对 Geant4 蒙特卡洛模拟框架分别在几何、跟踪、探测器相应、运行管理等方面进行重新构造, 模拟出本论文所需的辐射场各个位置剂量数据。

对于辐射场模拟数据, 采用 csv 格式进行存储并存储为三个文件:

1. data.csv: 用于存储测量点数据(x, y, z, value)
2. points.csv: 用于存储重构点位置数据(x, y, z)
3. values.csv: 用于存储重构点辐射场剂量数据(value)



## 4.2 辐射场重构算法设计

空间辐射场三维重构程序采用模块化设计，其主要由四部分组成，分别为数据模块、插值重构模块、偏差运算模块以及可视化模块，如图 4-1 所示。

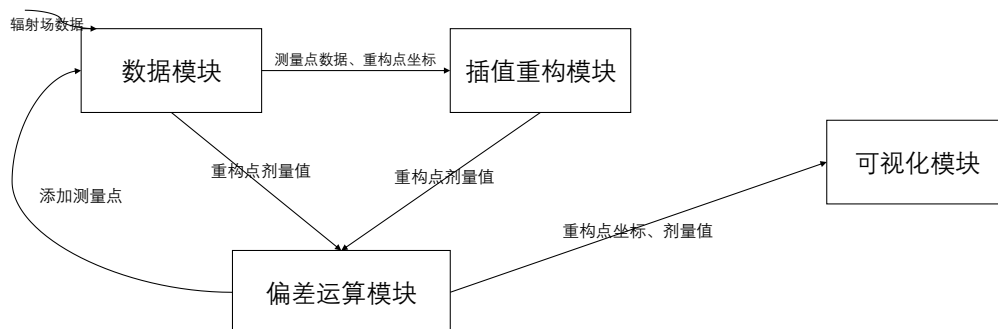


图 4-1 辐射场重构方法程序模块化设计

### 1. 数据模块

数据模块对辐射场整体数据进行处理以及对重构后偏差较大的点进行判断，该模块具体功能包括：

- 处理 Geant4 输出数据，将数据按照 csv 格式进行存储；
- 添加测量点数据，对于插值偏差较大的区域添加测量点，将其存储于测量点数据文件；
- 整合重构点数据，将重构后的数据进行整理，存储为 ROOT 绘制所需文件类型；

### 2. 插值重构模块

插值重构模块是空间辐射场三维重构程序中的核心模块，其中包含了两种插值算法——样条插值算法和克里金插值算法，其功能是通过两种重构算法对辐射场数据进行插值重构，得到插值重构场，并按权重对两种辐射场数据进行组合，最终得到合适的辐射场；

### 3. 偏差运算模块

偏差运算模块为辐射场三维重构程序的重要模块，该模块为本论文提出的插值重构方法中最关键的创新点实现的模块。该模块实现的功能为计算重构辐射场与 Geant4 模拟辐射场中各个坐标点的偏差，对于样条插值重构辐射场与克里金插值重构辐射场偏差较大的区域，选择一个点作为测量点；

### 4. 可视化模块

可视化模块是基于计算机的图像处理技术，将辐射场插值重构数据以更直观的

形式展现出来，从而使得重构辐射场能够在 ALARA 设计分析、辐射剂量估算等领域得到更好的应用。

数据模块先从 Geant4 模拟中得到辐射场数据，通过数据处理，转换为与插值重构算法模块能够耦合的数据类型，然后存储为 csv 格式；插值重构模块将处理后的数据进行读取、插值重构，分别将样条插值法和克里金插值法重构的辐射场数据进行输出；偏差运算模块得到两个辐射场重构数据后，将其进行偏差计算，若偏差值大于设定偏差量，则将插值区域输出至数据模块，让其进行再次测量，直至两种算法重构的辐射场偏差值在设定范围内，将其辐射场数值进行输出；可视化模块根据重构辐射场数据，将辐射场进行图像绘制。

在辐射场重构方法程序各模块设计基础上，考虑其实现的方法论及相应算法，总程序采用高内聚、低耦合的设计思想，尽可能将内容内聚、数据耦合，形成模块内功能内部连续，外部通过数据接口进行连接。本论文采用的辐射场重构算法的程序框图如图 4-2 所示。

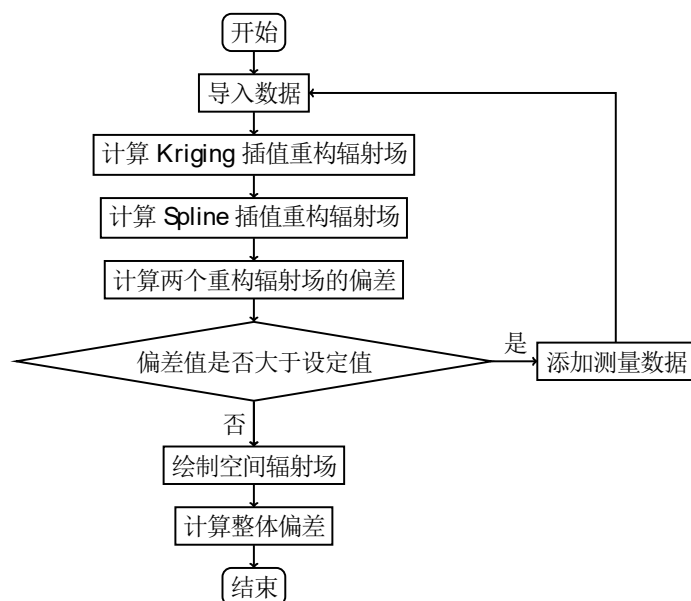


图 4-2 辐射场重构方法流程图

下面介绍该程序每一步骤所实现的方式：

#### 1) 导入数据：

辐射场可视化数据包括测量点数据文件、插值点坐标数据文件以及插值点辐射场剂量数据文件，导入数据步骤将三个文件分别按照各文件存储格式进行导入指定目录下；

#### 2) 计算 Kriging 插值重构辐射场：

Kriging 插值算法采用普通克里金模型, 变异函数采用指数函数, 具体计算方法原理及算法见第三章。本论文 Kriging 插值算法基于日本 GIS 公司开发的开源框架 Polatory 开发, Polatory 框架是基于径向基函数(RBF)插值的快速、高效框架;

3) 计算 Spline 插值重构辐射场:

Spline 插值算法具体原理以及算法见第二章, 本论文采用的样条插值算法为多层 B 样条插值算法(MBA), 本论文的辐射场样条插值算法是基于俄罗斯科学院研究员 Denis Demidov 在 Github 开源平台开源的 mba 库, mba 库是基于 Seungyong Lee 发表的《使用多层 B 样条进行散乱数据插值》论文中多层 B 样条插值算法<sup>[19]</sup>编写;

4) 计算两个重构辐射场的偏差:

通过 Kriging 插值和 Spine 插值后, 得到两个不同的辐射场数据, 通过计算两个辐射场的相对偏差(relative deviation), 来决定是否需要添加测量点来优化辐射场重构效果。相对偏差计算公式如下:

$$RD = \frac{|v_k - v_s|}{\frac{1}{2}(v_k + v_s)} \times 100\% \quad (4-1)$$

其中 RD 表示辐射场相对偏差;  $v_k$  代表克里金插值算法重构出的插值;  $v_s$  代表多层 B 样条插值算法重构出的插值;

5) 判断偏差值是否大于设定值:

通过上一步骤计算得到整个辐射场两种重构方法重构的偏差值之后, 对其偏差较大的区域(大于设定偏差值), 选取该区域内最大的偏差点进行再次测量;

6) 添加测量数据:

将上一步骤中得到的偏差最大的点的数据添加到测量点数据文件中, 保存为相应格式;

7) 绘制空间辐射场:

获得到重构辐射场数据后, 将数据保存为 root 格式, 使用 ROOT 进行绘制, 具体步骤在下一节中进行详细介绍;

8) 计算整体偏差:

在获得重构辐射场数据后, 将辐射场插值数据与模拟数据进行计算, 比较辐射场重构方法的重构效果。

### 4.3 空间辐射场可视化

辐射场可视化是对空间辐射场重构方法直观性的验证。相比于辐射场重构点数据，以图形、图像形式展现空间辐射场更直观。对于辐射场重构后的应用场景，例如核设施退役检修，辐射场可视化更容易对维修人员进行 ALARA 设计；矿井内重构辐射分布，以可视化的形式展示更加有利于矿工做出应对措施；公众进行核科普宣传时，以图形、图像甚至 VR、AR 形式对公众进行展示辐射场，更能提高公众对辐射防护、核安全的学习兴趣。

三维空间辐射场数据为 4D 数据(x,y,z,value)，4D 数据可视化可通过现有的虚拟现实编辑软件实现<sup>[34]</sup>，或者利用一些工具框架进行实现，例如：基于 MATLAB 程序设计语言的 scatter3 等函数；基于 Python 程序设计语言中 matplotlib 库中的相关函数；基于 C++ 程序设计语言的 OpenGL 框架、ROOT 框架等等。本论文采用的可视化工具为 ROOT 数据处理框架，ROOT 是 CERN 开发的数据处理框架，可以用于数据存储、访问、处理以及绘制，并且提供交互式运行界面以及可创建图形化用户界面，还提供与其他程序设计语言的接口(Python、R)。

### 4.4 本章小结

本章详细介绍了空间辐射场三维重构方法的设计与实现过程，主要分成三个部分进行介绍：辐射场数据获取、重构算法设计以及辐射场数据可视化。首先对辐射场数据来源进行了详细介绍，然后介绍辐射场重构算法程序的模块化开发以及每个模块所实现的功能作用，再通过流程图详细说明了辐射场重构算法的实现步骤，最后基于 ROOT 框架介绍辐射场数据可视化的实现方式。

## 第5章 辐射场重构方法评价

### 5.1 辐射场重构方法影响因素研究

本研究分别对简单空间单源辐射场、简单空间多源辐射场、带有屏蔽空间单源辐射场、带有屏蔽空间多源辐射场分别进行 Geant4 模拟，下面分别将源项数量、辐射场空间状况以及测量点数据为因变量，测评本论文提出的辐射场重构方法效果。

Geant4 模拟主要分为初始化和运行：初始化为模拟的参数进行定义，包括 1)定义几何体 2)设置物理过程 3)定义发射粒子；运行为 Geant4 代码执行时，Geant4 内核所执行的具体内容，包括 1)导入几何体结构 2)循环执行 Event(Event 为 Geant4 中粒子发射)。其中，对于不同辐射场的初始化，本研究都设置相同的物理过程，即仅考虑电磁相互作用，电磁相互作用主要包含以下物理过程：光电效应、康普顿效应、电子对效应、韧致辐射、瑞利散射等。对于 Geant4 的几何体定义和粒子源设置，在下面不同影响因素研究中分别进行介绍。

#### 5.1.1 源项数量对辐射场重构效果的影响

为探究源项数量对辐射场重构效果的影响，本论文分别对单源和多源情况分别进行 Geant4 模拟。

对于简单空间单源辐射场，整体辐射场空间设置为  $3m \times 4m \times 5m$  的立方体三维空间，辐射场材料设置为空气，几何体仅设置  $50mm \times 50mm \times 50mm$  的 CsI 材料进行探测，粒子源设置为能量为  $0.662MeV$  的  $^{137}Cs$  点源，点源位置设置在空间几何中心。对于简单空间多源辐射场，几何体也仅设置  $50mm \times 50mm \times 50mm$  的 CsI 材料，粒子源设置将分别定义三个位置不同的点源，分别为能量为  $0.662MeV$  的  $^{137}Cs$  点源，位置设为  $(-1.03m, -1.38m, -2.02m)$ ；能量为  $0.835MeV$  的  $^{54}Mn$  点源，位置设在  $(-1.03m, 1.57m, -1.72m)$  以及能量为  $0.662MeV$  的  $^{137}Cs$  点源，位置设为  $(0.93m, 0.57m, 1.32m)$ 。Geant4 模拟可视化效果如图 5-1 所示。

通过 Geant4 框架分别对单源无屏蔽辐射场和多源无屏蔽辐射场进行模拟仿真，获得空间辐射场数据使用 ROOT 绘制，如图 5-2 所示。

从图 5-2 中可以看出，在无屏蔽的情况下，辐射场剂量分布与源项数量相关。由于 Geant4 没有模拟天然本底，并且  $\gamma$  射线在空气中有一定衰减，因此在各个源项相距较远的情况下， $\gamma$  辐射场剂量率分布源项越多，多源辐射场剂量分布可以看作为多个单源辐

射场剂量分布加和。

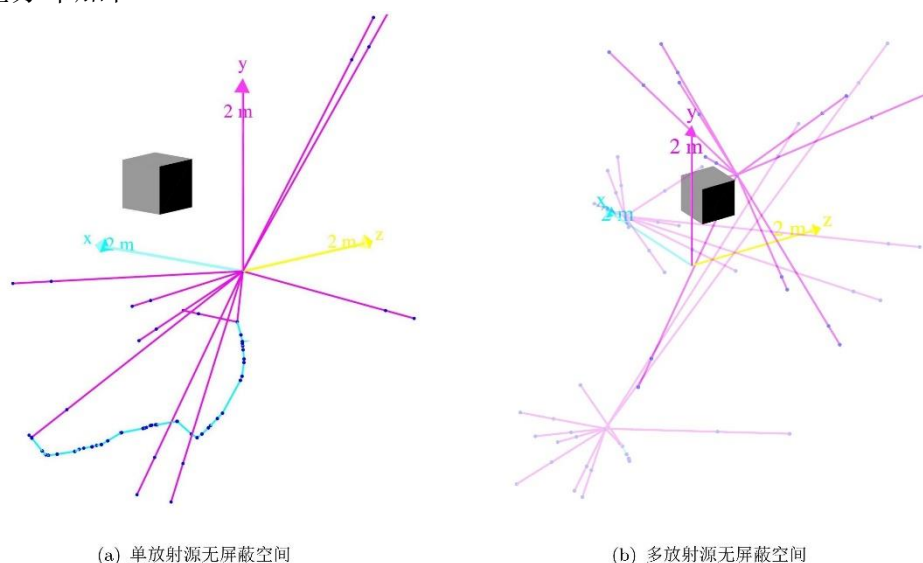


图 5-1 Geant4 模拟单源辐射场和多源辐射场

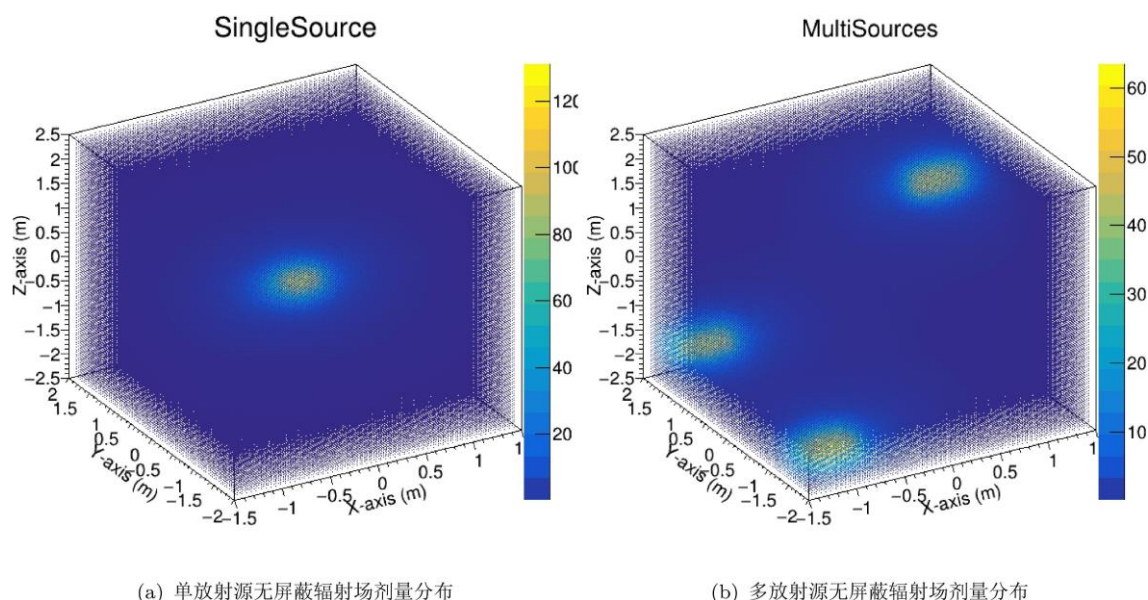


图 5-2 Geant4 模拟单源辐射场和多源辐射场的剂量分布

根据 Geant4 重构出的辐射场，先每隔 50cm 测量一个数据量，共测量  $6 \times 8 \times 10$  个数据点，对辐射场进行插值重构，再根据辐射场重构方法选取若干测量点数据，最终得到重构辐射场数据。将重构出的插值数据与 Geant4 模拟仿真的辐射场剂量率相比，得到不同源项数量与重构片插值的结果如表 5-1 所示。

表 5-1 列举了部分插值点和辐射场整体插值点相对偏差。从整体偏差值数据可以看出，单源辐射场重构整体相对偏差小于多源辐射场重构整体相对偏差。从辐射场插值点相对偏差值可以看出，离测量点越近的重构点，与 Geant4 模拟值的偏差越小；在辐射场空间边缘和源项附近，插值重构相对偏差较大。在空间边缘辐射场偏差较大的原因是

Geant4 模拟辐射场剂量值较低，导致重构相对偏差较大；源项附近插值相对偏差较大的原因为插值算法对梯度较大的领域重构效果不够好。

表 5-1 不同源项数量下重构插值方法对辐射场剂量率的相对偏差

插值坐标(m)	单源辐射场相对偏差(\%)	多源辐射场相对偏差(\%)
(-1.25, -1.75, -2.45)	38.29	61.43
(-1.25, -0.75, -0.65)	2.43	6.78
(-0.85, -1.35, 0.35)	0.22	3.05
(-0.65, -0.75, 0.15)	7.25	0.75
(-0.25, -0.75, -0.45)	1.72	3.01
(-0.25, -0.15, 0.55)	6.98	2.05
(0.35, -1.55, 0.15)	2.76	1.42
(-1.25, 1.05, 1.95)	3.47	0.01
(0.75, 0.85, 1.35)	0.10	31.37
平均相对偏差绝对值	6.15	13.4

### 5.1.2 空间状况对辐射场重构效果的影响

为探究空间状况对辐射场重构效果的影响，本论文分别对无屏蔽空间和带有屏蔽空间情况分别进行 Geant4 模拟。

对于无屏蔽空间辐射场，整体辐射场空间设置为  $3m \times 4m \times 5m$  的立方体三维空间，辐射场材料设定为空气；探测几何体设置为  $50mm \times 50mm \times 50mm$  的 CsI 材料，不设置屏蔽物；粒子源设置为能量为  $0.662MeV$  的  $^{137}Cs$  点源，点源位置设置为辐射场空间几何中心。对于有屏蔽空间辐射场，探测几何同样设置为  $50mm \times 50mm \times 50mm$  的 CsI 材料，设置两个屏蔽物：一个设置为  $2m \times 1m \times 0.02m$  铅板，位置设为  $(0.4m, 0.8m, 1.7m)$ ；另一个屏蔽物设置为  $2.5m \times 1.8m \times 0.3m$  的钢筋混凝土材料，位置设为  $(0.6m, -0.6m, -1.25m)$ ；粒子源设置为能量为  $0.662MeV$  的  $^{137}Cs$  点源，点源位置设置为辐射场空间几何中心。Geant4 模拟可视化效果如图 5-3 所示。

通过 Geant4 框架分别对单点源无屏蔽辐射场和单点源带有屏蔽辐射场进行模拟仿真，获得空间辐射场数据使用 ROOT 绘制，如图 5-4 所示。

从图 5-4 中可以看出，在单点源情况下，辐射场剂量分布与空间状况相关。在屏蔽物周围的领域内，辐射场剂量变化梯度较大，且梯度大小与屏蔽物材料有关，屏蔽材料的线衰减系数越大，辐射场剂量变化梯度越大。

根据 Geant4 重构出的辐射场，先每隔  $50cm$  测量一个数据量，共测量  $6 \times 8 \times 10$  个数据点，对辐射场进行插值重构，再根据辐射场重构方法选取若干测量点数据，最终

得到重构辐射场数据。将重构出的插值数据与 Geant4 模拟仿真的辐射场剂量率相比，得到不同源项数量与重构片插值的结果如表 5-2 所示。

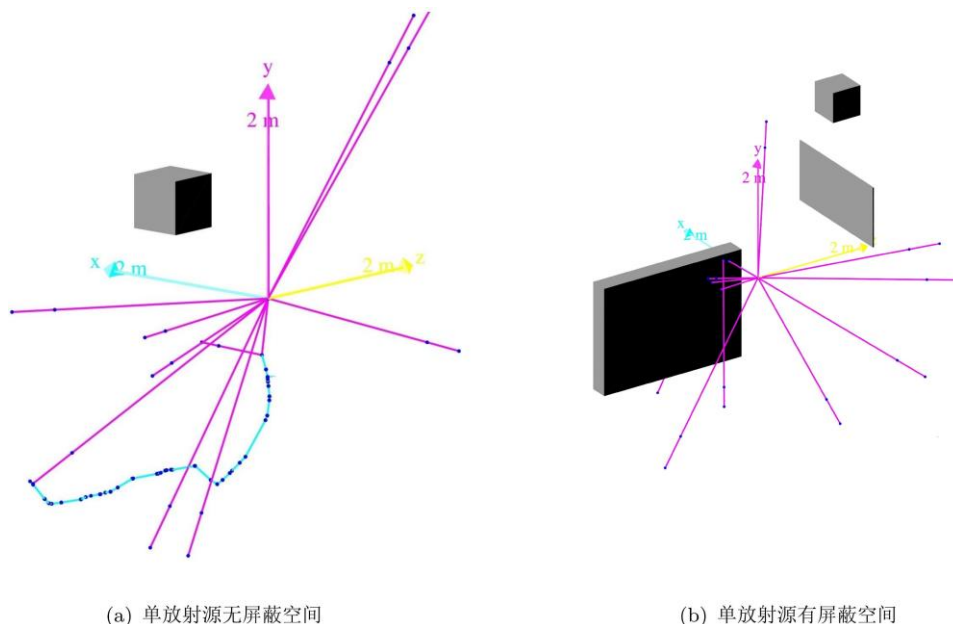


图 5-3 Geant4 模拟无屏蔽空间辐射场和有屏蔽空间辐射场  
SingleSource

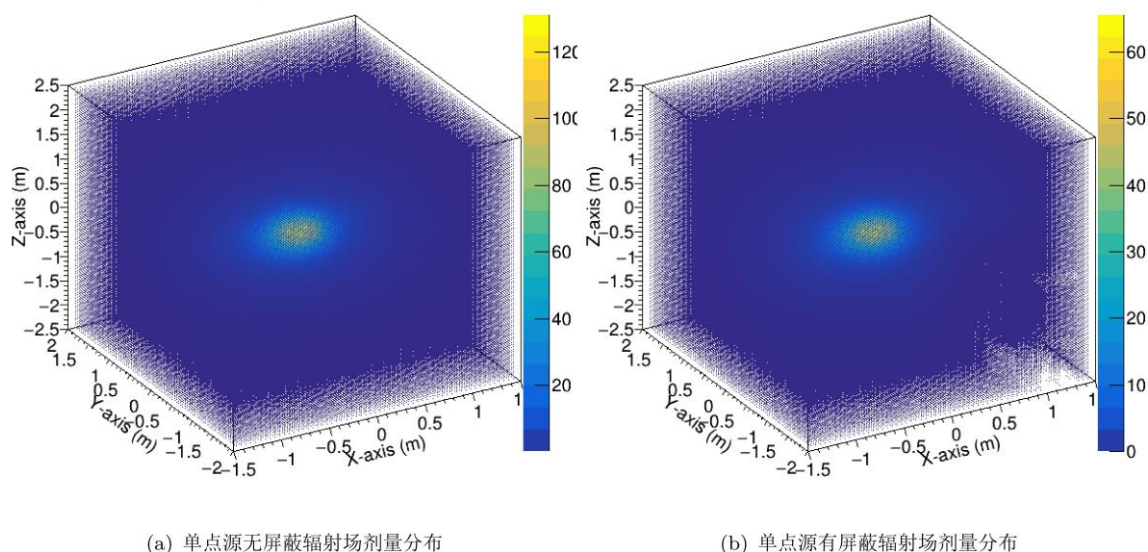


图 5-4 Geant4 模拟单点源无屏蔽辐射场和单点源有屏蔽辐射场的剂量分布

表 5-2 列举了部分插值点和辐射场整体插值点相对偏差。从整体偏差值数据可以看出，无屏蔽辐射场重构整体相对偏差小于带有屏蔽辐射场重构的整体相对偏差。从带有屏蔽的辐射场插值点相对偏差中可以看出，在靠近屏蔽物的地方，插值重构效果不是很好，说明插值重构在变化梯度大的位置重构的效果不够优秀。将空间状况与辐射场相对偏差数据同源项数量与相对偏差数据相比，可以看出空间状况对辐射场重构效果影响更大，或者说，本论文提出的辐射场重构方法对带有屏蔽的辐射场重构效果相对较差。



表 5-2 不同空间状况下重构插值方法对辐射场剂量率的相对偏差

插值坐标(m)	无屏蔽辐射场相对偏差(\%)	有屏蔽辐射场相对偏差(\%)
(-1.25, -1.75, -2.45)	38.29	2.43
(-1.25, -0.75, -0.65)	2.43	3.48
(-0.85, -1.35, 0.35)	0.22	6.03
(-0.65, -0.75, 0.15)	7.25	44.89
(-0.25, -0.75, -0.45)	1.72	59.07
(-0.25, -0.15, 0.55)	6.98	16.57
(0.35, -1.55, 0.15)	2.76	2.19
(-1.25, 1.05, 1.95)	3.47	3.21
(0.75, 0.85, 1.35)	0.10	0.85
平均相对偏差绝对值	6.15	53.46

### 5.1.3 测量点数据对辐射场重构效果影响

为探究测量点数据对辐射场重构效果的影响，本论文分别对单点源无屏蔽空间辐射场、多点源无屏蔽空间辐射场、单点源带屏蔽空间辐射场、多点源带屏蔽空间辐射场四种 Geant4 模拟辐射场进行不同测量点数量、测量点位置对比、分析。

对于以上四种不同类型辐射场，初始测点数量均为  $6 \times 8 \times 10$  个，根据本论文提出的辐射场插值重构方法，对于单点源无屏蔽辐射场最终测点数量为 480 个；多点源无屏蔽辐射场最终测点数量为 489 个；单点源有屏蔽辐射场最终测点数量为 491 个；多点源有屏蔽辐射场最终测量点数量为 483 个。四种辐射场相对偏差值大小及分布如图 5-5。

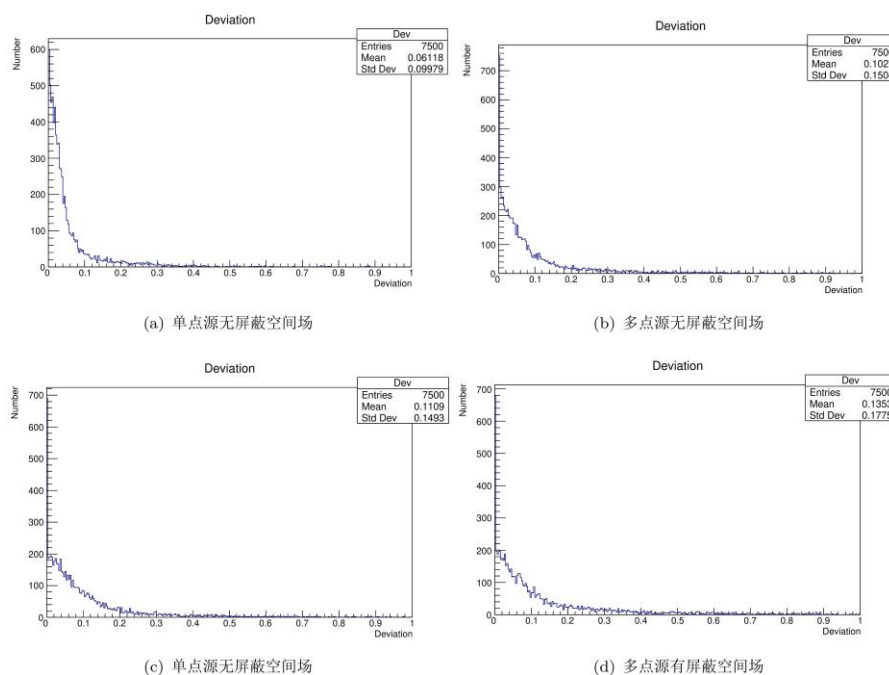


图 5-5 空间辐射场插值重构相对偏差分布

从辐射场重构相对偏差分布中可以看出,相对偏差大小从低到高数量分布大致呈半高斯分布。图 5-5(a)通过高斯拟合后  $\sigma = 0.19$ ; 图 5-5(b)通过高斯拟合后  $\sigma = 0.37$ ; 图 5-5(c)通过高斯拟合后  $\sigma = 0.45$ ; 图 5-5(d)通过高斯拟合后  $\sigma = 0.55$ 。通过对比四种辐射场相对偏差分布半高斯拟合方差  $\sigma$  大小,可以分析出在单点源无屏蔽空间辐射场中,采用本论文提出的辐射场插值重构方法,得出的插值重构数据 68.27% 相对偏差在 5% 以内;在多点源无屏蔽空间辐射场中,得出的插值重构数据 68.27% 相对偏差在 11% 以内;在单点源有屏蔽空间辐射场中,得出的插值重构数据 68.27% 相对偏差在 14% 以内;在多点源有屏蔽空间辐射场中,得出的插值重构数据 68.27% 相对偏差在 18% 以内。

通过改变初始测点数量,分别对以上四种辐射场进行插值重构,得到初始测点数量与相对偏差平均值的关系(初始测量点位置为随机选取)如图 5-6 所示。

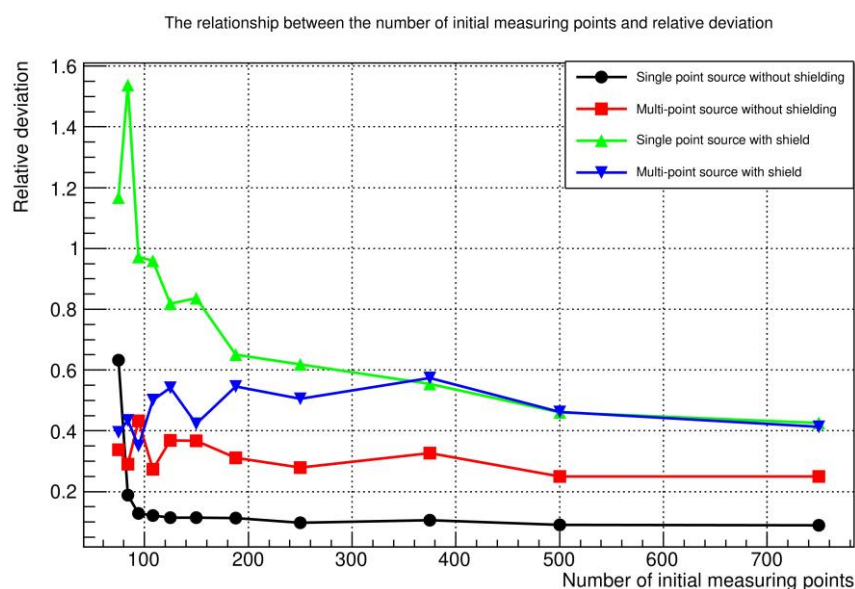


图 5-6 初始测点数量与相对偏差的关系

从图 5-6 中可以看出,初始测量点数据增加时,整体相对偏差呈下降趋势。通过对不同初始测点数据进行插值重构,发现当初始测点数量较少时,通过本论文提出的插值重构算法,在插值过程中添加测点数量越多。对比四种辐射场,可以发现本论文提出的辐射场插值重构方法对无屏蔽空间辐射场重构效果较好,源项数量增加对重构相对偏差有一定影响,但相对偏差大小在一定范围内;而对带有屏蔽辐射场重构效果相对一般,主要原因是对于带有屏蔽空间辐射场,其剂量分布在屏蔽物周围变化较大,插值重构方法对变化梯度较大的数值重构效果一般。

## 5.2 辐射场重构方法比较

本研究分别采用多层 B 样条插值重构方法、克里金插值重构方法以及本论文提出

的插值重构方法对上一节中模拟的四种辐射场进行插值重构。通过对比重构相对偏差以及重构时间来比较三种插值重构方法的在不同辐射场下的重构效果。

### 5.2.1 辐射场重构相对偏差比较

多层 B 样条插值重构算法原理见第二章，是插值算法中常用的一种算法；克里金插值重构算法原理见第三章，是一种多用于地质勘探领域的插值算法；本论文提出的辐射场插值重构算法见第四章，是一种结合多层 B 样条插值算法和克里金插值算法的创新插值重构算法。下面分别对该三种插值重构算法在单点源无屏蔽空间辐射场、多点源无屏蔽空间辐射场、单点源有屏蔽空间辐射场、多点源有屏蔽空间辐射场进行应用，比较其在不同辐射场中的相对偏差大小。

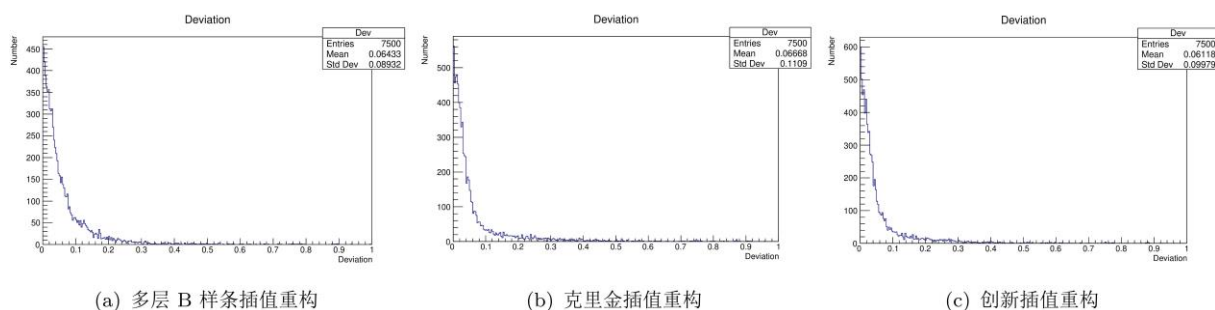


图 5-7 单点源无屏蔽空间辐射场插值重构相对偏差对比

从图 5-7 中可以看出，三种插值算法都能较好的重构出辐射场剂量分布，本论文提出的插值重构方法相比于多层 B 样条插值重构方法和克里金插值重构方法要略好。

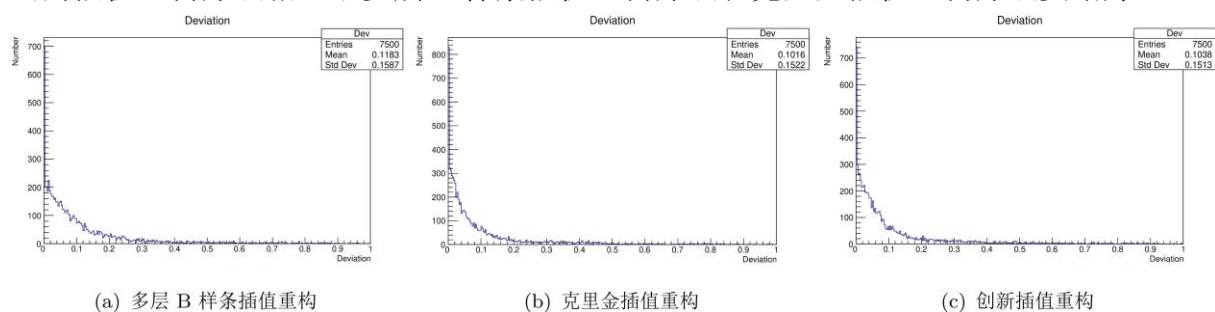


图 5-8 多点源无屏蔽空间辐射场插值重构相对偏差对比

从图 5-8 中可以看出，三种插值算法都能不错的重构出辐射场剂量分布，克里金插值重构方法的相对偏差总体要好于多层 B 样条插值重构方法和本论文提出的插值重构方法，因为在多点源无屏蔽辐射场中，辐射场剂量分布情况通过克里金插值重构方法相比多层 B 样条插值重构方法更加合适，因此重构结果偏差相对较小。

从图 5-9 中可以看出，多层 B 样条插值重构与模拟结构的相对偏差大于克里金插值重构和本论文提出的辐射场插值重构方法相对偏差，说明在带有屏蔽的辐射场下，样条

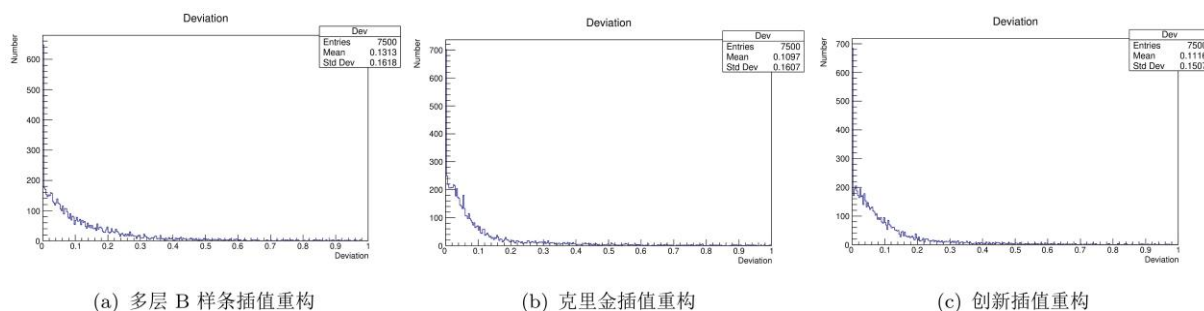


图 5-9 单点源有屏蔽空间辐射场插值重构相对偏差对比

插值重构方法差于克里金插值重构方法。

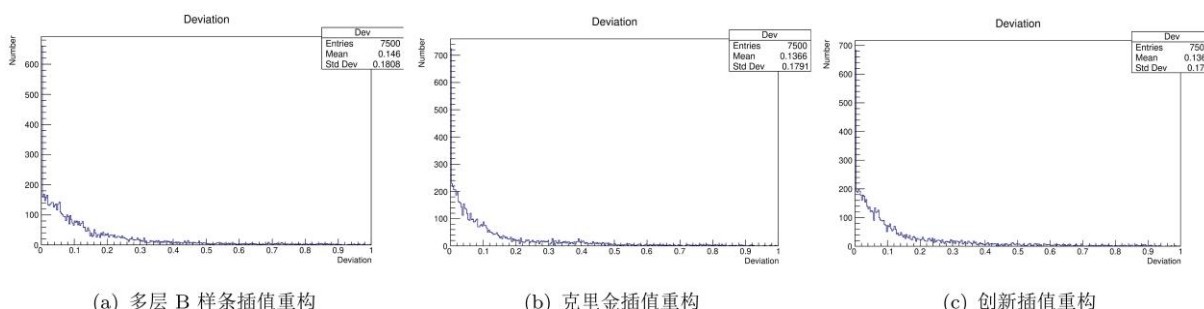


图 5-10 多点源有屏蔽空间辐射场插值重构相对偏差对比

从图 5-10 中可以看出，三种插值算法都能较好的重构出辐射场剂量分布，本论文提出的插值重构算法重构效果与克里金插值重构算法重构效果相似，略好于样条插值重构算法重构的辐射场结果。

通过对比本论文提出的插值重构方法与多层 B 样条插值重构方法和克里金插值重构方法的重构相对偏差，发现本论文提出的插值重构方法在某些插值点要明显好于其他两种方法，但在某些插值点也明显。例如在多点源无屏蔽空间辐射场中，存在多层 B 样条插值重构相对偏差较小的插值点，也存在克里金插值重构相对偏差较小的插值点，亦存在本论文提出的辐射场插值重构相对偏差较小的插值点，如表 5-3 所示。通过比较这些插值方法，发现在辐射场剂量分布较为平缓的区域内，多层 B 样条插值重构出的辐射场与模拟值相对偏差较小；在辐射场剂量分布变化较大的区域，克里金插值重构出的辐射场与模拟值相对偏差较小。

表 5-3 不同重构插值方法下辐射场剂量率插值相对偏差比较

插值坐标(m)	多层 B 样条插值	克里金插值重构	创新插值重构(%)
(-1.45, -0.75, -0.65)	6.81	10.00	8.41
(-1.25, 1.05, 1.95)	1.02	0.86	0.01
(0.75, 0.05, -0.25)	3.98	0.41	2.19

## 5.2.2 辐射场重构时间比较

为探究三种辐射场重构方法在重构时间上的关系，将多层 B 样条插值重构方法、克

里金插值重构方法和本论文提出的辐射场插值重构方法分别在以上四种辐射场进行插值，记录其重构所需的时间，结果如表 5-4 所示。

表 5-4 不同重构插值方法下辐射场重构时间比较

空间辐射场类型	多层 B 样条(s)	克里金插值(s)	本文重构方法(s)
单点源无屏蔽空间	1.20776	0.000233	0.001025
多点源无屏蔽空间	2.02339	0.000188	0.001567
单点源有屏蔽空间	1.9229	0.000169	0.00236
多点源有屏蔽空间	1.93904	0.000223	0.001668

从表 5-4 中可以看出，三种插值重构方法中克里金插值重构所花费的时间最短，本论文提出的插值重构方法时间其次，多层 B 样条插值重构时间最长。在实际辐射场重构应用中，该三种方法重构所花费的时间均可。

### 5.3 本章小结

本章对辐射场插值重构方法进行了详细的验证，并且将本论文提出辐射场插值重构方法与多层 B 样条插值重构方法和克里金插值重构方法进行比较。本章分别探究了源项数量、辐射场空间状况以及测点数量对辐射场插值重构的影响，结果表明源项数量对辐射场插值有一定影响；空间屏蔽物对重构结果的影响比源项数量大；对辐射场插值重构起决定性影响的为初始测点数量。当辐射场测点数量低于一定数量时，插值重构效果都难以达到满意的效果。

通过将本论文提出的插值重构方法与多层 B 样条插值方法和克里金插值重构方法比较，发现在辐射场剂量分布较为平缓的区域，多层 B 样条插值重构方法重构出的辐射场效果较好；在辐射场剂量分布变化梯度较大的区域，克里金插值重构方法重构出的辐射场效果较好。总体来说，本论文提出的插值重构效果在不同区域均能重构出不错的效果。三种插值重构方法在重构时间上均能满足应用要求。

## 总 结

近年来,反演法重构辐射场在辐射场重构中的研究越来越多,辐射场重构是辐射场可视化技术中的关键环节,可应用于核设施退役工程、核辐射安全科普以及矿井辐射剂量监测等等。本论文基于多层 B 样条插值方法和克里金插值方法提出一种适用于三维空间辐射场的插值重构方法,该方法利用有限个离散的辐射场数据能够重构出整个  $\gamma$  辐射场,并提供可视化数据接口,使得重构辐射场能够以图像方式展示。主要研究内容如下:

1、基于多层次 B 样条插值和克里金插值提出一种空间辐射场三维重构方法,能够实现三维区域内散乱数据重构。该重构方法将多层 B 样条插值与克里金插值结果进行结合,能够在某些区域减少重构的相对偏差,从而提高了辐射场重构的精度。一般地,采用本论文提出的辐射场重构方法能够重构出较为满意的辐射场。

2、利用蒙特卡洛应用软件包 Geant4 模拟四种不同类型辐射场,通过对该四种辐射场进行重构,分别探究源项数量、辐射场空间状况以及重构测点数量对辐射场插值重构的影响。结果表明,源项数量对重构相对偏差影响较小;辐射场空间状况对重构相对偏差有一定影响;重构测点数量对辐射场插值起决定性影响。

3、将本论文提出的辐射场重构方法与多层 B 样条插值重构方法、克里金插值重构方法相比较,在单点源无屏蔽空间辐射场中,本论文提出的插值重构方法略好于另外单一插值重构方法;在辐射场剂量分布平缓的区域,多层 B 样条插值重构效果较好;在辐射场剂量分布变化较大的区域,克里金插值重构效果较好。在重构时间方面,克里金插值重构最快,本论文提出的插值重构其次,多层 B 样条插值重构最慢,但三种重构方法重构时间均在工程应用可接受范围内。

上述研究结果表明,本论文提出的空间辐射场三维重构方法具有现实意义。由于时间、精力 and 知识储备的限制,还有很多不够完善的地方,在以后的工作中,可以对以下几个方面做进一步研究:

1、将本文提出的辐射场插值重构方法在实际辐射场中进行应用,分析其在实际应用中的偏差大小,进一步验证其可行性;

2、改进本论文辐射场插值重构中多层 B 样条插值和克里金插值的结合方式,将其比重与辐射场分布梯度关联,提高其插值精度;

3、引入更多的插值重构方法,例如有限元法、三角划分法等等,分析以上方法在哪些辐射场区域分布下精度较高,将其进行组合;

4、将机器学习算法应用于辐射场重构领域中,分析其可行性并与现有重构方法对比。

## 参考文献

- [1] Wang Z, Cai J. Reconstruction of the neutron radiation field on nuclear facilities near the shield using Bayesian inference[J]. Progress in Nuclear Energy, 2020, 118: 103070.
- [2] 曹亚丽, 王尔奇, 王晓峰, 等. 科学传播模式在我国核与辐射安全科普工作中的应用[J]. 核安全, 2014, 1.
- [3] Chen Q, Zhang J. The visualization of 3D radiation dose field in virtual environment[C]//Journal of Physics: Conference Series. IOP Publishing, 2021, 1827(1): 012006.
- [4] Böhm R, Sedlák A, Bulko M, et al. Radon as a Tracer of Lung Changes Induced by Smoking[J]. Risk Analysis, 2020, 40(2): 370-384.
- [5] Harrison J D. Lung cancer risk and effective dose coefficients for radon: UNSCEAR review and ICRP conclusions[J]. Journal of Radiological Protection, 2021: 433-441.
- [6] Dehnert J. Radon exposures of miners at small underground construction sites in old mining: recommendations to improve radiation protection measures by the Saxon Radiation Protection Authority[J]. Health physics, 2020, 118(1): 96.
- [7] Majer M, Roguljić M, Knežević Ž, et al. Dose mapping of the panoramic  $^{60}\text{Co}$  gamma irradiation facility at the Ruđer Bošković Institute—Geant4 simulation and measurements[J]. Applied radiation and Isotopes, 2019, 154: 108824.
- [8] Zhang Z, Song Y, Ma S, et al. A rapid coupling method for calculating the radiation field in decommissioning nuclear power plants[J]. Annals of Nuclear Energy, 2021, 156: 108179.
- [9] Wang Z, Cai J. Inversion of radiation field on nuclear facilities: a method based on net function interpolation[J]. Radiation Physics and Chemistry, 2018, 153: 27-34.
- [10] 李海涛, 邵泽东. 空间插值分析算法综述[J]. 计算机系统应用, 2019, 28(7): 1-8.
- [11] Thevenon J B, Tirel O, Lopez L, et al. Chavir: Virtual reality simulation for interventions in nuclear installations[C]//Proceedings of the 5. International Topical Meeting on Nuclear Plant Instrumentation Controls, and Human Machine Interface Technology. 2006.
- [12] Vermeersch F, Vanbosstraeten C. Software VISIPLAN: a powerful tool for optimisation[C]//2nd European workshop on occupational exposure management at NPPS, Tarragona, Spain. 2000: 5-7.
- [13] He T, Hu L, Long P, et al. Virtual reality-based simulation system for nuclear and radiation safety SuperMC/RVIS[J]. Radiation Protection Dosimetry, 2015: 7-15.



- [14]Grigoryev A M, Tashlykov O L, Popel A A, et al. Determination of radiation field parameters for the problems of routing optimization based on interpolation with radial basis functions[C]//AIP Conference Proceedings. AIP Publishing LLC, 2020, 2313(1): 020007.
- [15]White S R, Wood K T, Martin P G, et al. Radioactive Source Localisation via Projective Linear Reconstruction[J]. Sensors, 2021, 21(3): 807.
- [16]赛雪, 陈颖, 韦孟伏. Multiquadric 散乱数据插值方法在  $\gamma$  辐射场可视化中的应用初探[J]. 核技术, 2016, 39(10): 100501-1-100501-7.
- [17]施法中, 图形学, 查看清, 等. 计算机辅助几何设计与非均匀有理 B 样条 (修订版)[M]. 北京: 高等教育出版社, 2013.
- [18]Franke R. Scattered data interpolation: tests of some methods[J]. Mathematics of computation, 1982, 38(157): 181-200.
- [19]Lee S, Wolberg G, Shin S Y. Scattered data interpolation with multilevel B-splines[J]. IEEE transactions on visualization and computer graphics, 1997, 3(3): 228-244.
- [20]Zhang H, He X. Deep free-form deformation network for object-mask registration[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 4251-4259.
- [21]Nguyen L, Miro J V, Qiu X. Multilevel B-splines-based learning approach for sound source localization[J]. IEEE Sensors Journal, 2019, 19(10): 3871-3881.
- [22]Jiang Y, Zhang N. Applications of scattered interpolation with multilevel B - splines to magnetic anomaly data[J]. Near Surface Geophysics, 2013, 11(5): 579-585.
- [23]Wang H, Amini A A. Cardiac motion tracking with multilevel B-splines and SinMod from tagged MRI[C]//Medical Imaging 2011: Biomedical Applications in Molecular, Structural, and Functional Imaging. International Society for Optics and Photonics, 2011, 7965: 796520.
- [24]Zhang N, Wang Y. Scattered interpolation with multilevel B-splines for seismic reconstructions[C]//72nd EAGE Conference and Exhibition incorporating SPE EUROPEC 2010. European Association of Geoscientists & Engineers, 2010: cp-161-00666.
- [25]Cressie N. The origins of kriging[J]. Mathematical geology, 1990, 22(3): 239-252.
- [26]Lovric M. International Encyclopedia of Statistical Science[M]. Springer, 2011.
- [27]Oliver M A, Webster R. Basic steps in geostatistics: the variogram and kriging[M]. New York, NY: Springer International Publishing, 2015.
- [28]Chiles J P, Delfiner P. Geostatistics: modeling spatial uncertainty[M]. John Wiley & Sons,



2009.

- [29]Zhong B, Liang T, Wang L, et al. Applications of stochastic models and geostatistical analyses to study sources and spatial patterns of soil heavy metals in a metalliferous industrial district of China[J]. Science of the total environment, 2014, 490: 422-434.
- [30]Castrignano A, Boccaccio L, Cohen Y, et al. Spatio-temporal population dynamics and area-wide delineation of *Bactrocera oleae* monitoring zones using multi-variate geostatistics[J]. Precision Agriculture, 2012, 13(4): 421-441.
- [31]Adams C F, Harris B P, Marino II M C, et al. Quantifying sea scallop bed diameter on Georges Bank with geostatistics[J]. Fisheries research, 2010, 106(3): 460-467.
- [32]Matheron G. Les variables régionalisées et leur estimation: une application de la théorie de fonctions aléatoires aux sciences de la nature[M]. Masson et CIE, 1965.
- [33]Gould H, Tobochnik J, Christian W. An introduction to computer simulation methods[M]. New York: Addison-Wesley, 1988.
- [34]张永领, 胡一非, 刘猛, 等. 反应堆退役三维辐射场实时计算及可视化[J]. 辐射防护, 2020, 38(1): 19-25.

## 致 谢

时光荏苒，转眼间四年的本科生涯已经进入尾声。回望四年前，当我刚收到哈尔滨工程大学录取通知书时，那种喜悦之情不可言表，在这四年的校园生活中，欢乐与笑语常伴、至善和至真随行。在此，我特别感激各位老师、同学在本科生涯中对我的帮助与陪伴。

感谢我的指导老师宋玉收老师，宋老师渊博的学识、精益求精的科研态度、一丝不苟的治学态度将使我受益终生。宋老师不仅在论文细节上给予我极大的指导，还从科研思维、研究脉络方面对我进行教导，让我受益良多。

感谢核科学与技术学院所有授我以业的老师，没有这些年知识的积淀，我没有这么大的动力和信心完成这篇论文；感谢我的班主任王贺老师这几年来对我的关心，帮忙与支持。

感谢我的室友们，是你们同我共同维系着寝室那份家的融洽；感谢核学院足球队的小伙伴们，是你们同我在绿茵场上挥汗如雨；感谢定向越野队的同学们，是你们同我清晨在操场上拉练体能。在毕业之际，我祝你们有幸福的明天，美好的未来！

感谢我的家人们，感谢你们对我默默的支持以及无微不至的关怀，在我的成长道路上，是因为有你们的陪伴和关爱，我才有了今天的成绩！

人生的每个阶段都值得被珍惜，美好的岁月中，正是因为有你们我才可以收获和成长。在今后的时间里我不会停下学习的步伐，努力让自己成为更好的人，我相信这是对你们最好的报答！

## 附录 克里金法的多变量扩展

为了解决石油工程、采矿和地质、气象、土壤科学、精确农业、污染控制、公共卫生、渔业、动植物生态、遥感和水文学等方面日益复杂的问题，克里金法已经发展出了许多不同种类的克里金方法。

根据正文中公式 3-1 中确定性函数的不同类型，可以将克里金法分为普通克里金法、泛克里金法、简单克里金法、指示克里金法、协同克里金法、概率克里金法、析取克里金法等。

**普通克里金法** 普通克里金法假设的模型为：

$$V(s) = \mu + \varepsilon(s)$$

其中， $\mu$ 为一个未知常量。

选取普通克里金法，最主要的问题之一就是判断常量平均值的假设是否合理。普通克里金法凭借其显著的灵活性，是现如今应用最为广泛的克里金法。

**泛克里金法** 泛克里金法的假设模型为：

$$V(s) = \mu(s) + \varepsilon(s)$$

其中， $\mu(s)$ 为带有未知参量的函数，例如多项式函数、幂指数函数等。

**简单克里金法** 简单克里金法的假设模型为：

$$V(s) = \mu + \varepsilon(s)$$

其中， $\mu$ 为已知常量，或者没有未知参量的函数。

**指示克里金法** 指示克里金法的假设模型为：

$$I(s) = \mu + \varepsilon(s)$$

其中， $\mu$ 为一个未知参量； $I(s)$ 为一个二进制变量。二进制数据的创建可利用连续数据的阈值实现，或观测数据可以为0或1。指示克里金法可用于离散数据的插值。

**协同克里金法** 协同克里金法的假设模型为：

$$\begin{cases} V_1(s) = \mu_1 + \varepsilon_1(s) \\ V_2(s) = \mu_2 + \varepsilon_2(s) \end{cases}$$

其中， $\mu_1$ 和 $\mu_2$ 为未知常量； $\varepsilon_1(s)$ 和 $\varepsilon_2(s)$ 为两个随机误差各自的自相关函数，它们之间存在互相关。

**概率克里金法** 概率克里金法的假设模型为：

$$\begin{cases} I(s) = I(V(s) > c_t) = \mu_1 + \varepsilon_1(s) \\ V(s) = \mu_2 + \varepsilon_2(s) \end{cases}$$

其中， $\mu_1$ 和 $\mu_2$ 为未知常量， $I(s)$ 为通过使用阈值指示 $I(V(s) > c_t)$ 创建的二进制变量。

**析取克里金法** 析取克里金法的假设模型为：

$$f(V(s)) = \mu_1 + \varepsilon(s)$$

其中， $\mu_1$ 为一个未知常量； $f(V(s))$ 为 $V(s)$ 的一个任意函数。若 $f(V(s)) = I(V(s) > c_t)$ ，则指示克里金就是析取克里金的一种特殊情况。