



武欣, 薛国强, 王研博, 等. 2026. 地球物理多元信息融合数据预处理方法研究进展. 地球物理学进展, 41 (1): 0143-0155, doi: 10. 6038/pg2026H0426.

WU Xin, XUE GuoQiang, WANG YanBo, et al. 2026. Research progress on data preprocessing methods for geophysical multi-information fusion. *Progress in Geophysics* (in Chinese), 41 (1): 0143-0155, doi: 10. 6038/pg2026H0426.

地球物理多元信息融合数据预处理方法研究进展

Research progress on data preprocessing methods for geophysical multi-information fusion

武欣^{1,2,3}, 薛国强^{1,2,3*}, 王研博^{1,2,3}, 崔嵩^{1,2,3}, 石金晶^{1,2,3}

WU Xin^{1,2,3}, XUE GuoQiang^{1,2,3*}, WANG YanBo^{1,2,3}, CUI Song^{1,2,3}, SHI JinJing^{1,2,3}

1. 中国科学院地质与地球物理研究所, 中国科学院矿产资源研究重点实验室, 北京 100029

2. 中国科学院地球科学研究院, 北京 100029

3. 中国科学院大学, 北京 100049

1. Key Laboratory of Mineral Resources, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China

2. Innovation Academy for Earth Science, Chinese Academy of Sciences, Beijing 100029, China

3. College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

摘要 多元信息融合技术是近年来蓬勃发展的一种新兴技术,是人工智能理论与技术在信息分析与加工领域的体现,并在战场态势感知、工业机器人、远程医疗、自动驾驶等领域发挥着重要作用. 随着地球物理探测理论与技术的长足发展,尤其是近年来,随着我国机载/星载对地观测技术的快速进步,产生的海量数据对现有主要基于人工经验的分析模式形成了严峻挑战. 将多元信息融合技术引入地学领域,特别是地球物理探测数据领域,存在日渐迫切的需求. 然而,由于地球物理观测存在方法、场景、装备等多方面差异,所获取的数据在空间展布规范上彼此不同,导致后续信息融合计算难以开展. 为此,须按照统一规范对观测数据进行预处理,以确保数据具有相同的观测密度. 立足于多元信息融合需求,对观测数据进行体系化预处理的理论与技术目前均是欠缺的. 在一些研究中,引用传统的预测技术来开展拟三维(或拟体系化)的数据规范化处理,效果仍有待评估. 为此,本文立足面向地球物理的多元信息融合方法研究,对现有数据规范化预处理方法进行综述,并引入一半航空电磁观测案例,对现有各

Abstract Multimodal information fusion technology is an emerging field that has flourished in recent years, representing the application of artificial intelligence theories and technologies in information analysis and processing. It plays a crucial role in various domains, such as battlefield situational awareness, industrial robotics, remote medical care, and autonomous driving. With the rapid development of geophysical exploration theories and technologies, particularly with the significant advancements in China's airborne and satellite-based Earth observation technologies in recent years, the massive data generated poses a severe challenge to the traditional analysis models, which are primarily based on human experience. There is an increasing demand to introduce multi-information fusion technology into the field of geosciences, especially in the domain of geophysical exploration data. However, due to the differences in methods, scenarios, and equipment in geophysical observations, the data obtained have varying spatial distribution standards, making subsequent information fusion calculations difficult. Therefore, it is necessary to preprocess the observational data according to a unified standard to ensure that the data have consistent observational density. Currently, both the theory and technology for voxel-based preprocessing of observational data, which is crucial for multi-information fusion, are

收稿日期 2025-01-08; 修回日期 2025-06-05.

投稿网址 <http://www.progeophysics.cn>

基金项目 国家自然科学基金项目(42030106, 42074121)资助.

第一作者简介 武欣,男,1982年生,副研究员,研究方向为电磁法高精度探测方法与技术. E-mail: wu_xin18@163.com

* 通讯作者 薛国强,男,1966年生,研究员,研究方向为近源电磁探测理论方法. E-mail: ppxueguoqiang@163.com

类方法的处理性能进行讨论,为进一步发展完善地球物理多元信息融合处理技术体系提供方法依据.

关键词 多元信息融合;数据预测;体素化

中图分类号 P631

文献标识码 A

doi:10. 6038/pg2026H0426

CSTR;32085.14.pg2026H0426

lacking. In some studies, traditional prediction techniques have been used to perform quasi-three-dimensional (or quasi-voxel-based) data standardization, but the effectiveness of these approaches is still under evaluation. Therefore, this paper reviews the existing data standardization preprocessing methods from the perspective of geophysical multi-information fusion methods. It introduces a semi-airborne electromagnetic observation case to discuss the performance of various existing methods, providing a methodological basis for further developing and refining the geophysical multi-information fusion processing technology system.

Keywords Multimodal information fusion; Data prediction method; Voxel-ization

0 引言

经济增长和科技进步的背景下,矿产资源作为社会发展的关键性不可再生资源,其重要性依旧显著(徐德义等, 2023). 面对资源安全形势的日益严峻,实现深部矿床或隐伏矿床的精准定位和探明,以增加矿产资源的储量,已成为资源勘查行业的核心任务(曾庆栋等, 2021). 随着地球科学领域的迅速发展,我们能够更加便捷地获取到大量的多样化地球观测数据,这为进一步的研究提供了坚实的基础(Vance *et al.*, 2024). 然而,当前的主要矛盾在于,尽管找矿难度不断增加且观测数据量持续膨胀,数据分析方法的发展却未能跟上步伐(李芳等, 2016). 主流的矿产资源勘查分析方法在很大程度上依赖于专家的经验 and 直觉. 鉴于此,开发一种能够综合考虑矿床的地球物理、地球化学等多元特征信息的分析方法显得尤为迫切,这将有助于降低对专家经验的依赖,同时提升勘查工作的效率和准确性(底青云等, 2012;吴志春等, 2016). 这实际上是地学领域以及更广泛的大数据领域中的一个普遍且迫切的需求:如何有效地从庞大的数据集中筛选并提取出既可靠又有价值的信息(Meng *et al.*, 2020). 信息融合技术,正是应对这一挑战的关键解决方案,通过整合来自多个传感器的数据以及数据库中的信息,能够提供比单一传感器更精确和具体的推断(Hall and Llinas, 1997; Zheng, 2015). 该技术已被成功应用于战场态势感知、工业机器人、远程综合医学诊断、自动驾驶等多个领域(Bokareva *et al.*, 2006; Muzammal *et al.*, 2020; Velasco-Hernandez *et al.*, 2020).

地球物理观测受限于方法、场景、装备等多方面因素,所获取的数据在空间展布规范上彼此不同. 在

进行信息融合的过程中,不同观测方法获得的数据需要对齐操作,以确保不同数据集对环境的描述没有空间维度上的差异(Drews *et al.*, 2022). 自动驾驶等常规信息融合应用领域,由于研究资源充足,实际上已经有了一套成熟的数据预测、信息对齐方案(Li *et al.*, 2022). 但地球物理的多元信息融合技术尚处发展阶段,目前缺乏专门为此领域设计的先进数据预测工具,这限制了多元信息融合分析的实施. 因此,开发一种高效的数据体素化、空间对齐方法显得尤为关键,它是实现高效、准确多元信息融合分析的前提条件. 预测方法可以利用区域内已知观测点的数据来估计未知点的数据,是一种有效的数据扩充手段,并已在多个研究领域得到广泛应用(Lam, 1983; Xia *et al.*, 2022). 合理使用预测方法能够确保数据具有相同的观测密度,这在一定程度上可以作为数据预测的替代方案. 但是这些方法在数据预测任务中的适用性和有效性仍旧未知,所以需要对其对现有的数据预测方法进行系统的梳理和评估,这将为地学领域的多元信息融合提供坚实的技术支撑和理论基础.

本文立足于地球物理的多元信息融合方法研究,对现有数据规范化预处理的预测方法进行综述,深入探讨了这些方法在数据预测领域的应用潜力和适宜性. 通过具体的案例分析,本文进一步综合评估了它们在实际数据预测任务中的性能和效果. 最后本文就未来数据预测方法的研究方向提供了展望,旨在为地学领域中的多元信息融合研究提供新的视角与思考路径.

1 维诺图方法

维诺图,又称为狄利克雷图或泰森多边形,最早由 Dirichlet 在 1850 年讨论其概念,而 Voronoi 在

1908 年对其进行了数学上的严格定义 (Lejeune Dirichlet, 1850; Voronoi, 1908). 维诺图将平面划分为若干个多边形, 每个多边形内部的任意点距离该多边形的控制点最近. 维诺图的生成依赖于已知点, 并根据这些点的位置确定各个多边形的边界.

维诺图的数学描述可以表示为: 给定平面上的点集 $\{x_1, x_2, \dots, x_n\}$, 维诺图将平面划分为 n 个区域 $V(x_i)$, 其中 $V(x_i)$ 表示所有距离点 x_i 最近的点的集合, 即:

$$V(x_i) = \{x_0 \in \mathbb{R}^2 \mid d(x_0, x_i) < d(x_0, x_j), \forall j \neq i\}, \quad (1)$$

其中 $d(x_0, x_i)$ 表示点 x_0 与 x_i 之间的距离.

基于维诺图的数据预测方法主要有最近邻点方法和自然邻点方法. 最近邻点方法是将研究区域按已知点进行维诺图剖分后, 将待预测点赋值为所在泰森多边形控制点的值 (Goovaerts, 2000). 这种方法简单直接, 能够很好地体现离散观测点的空间控制范围, 适用于弱空间变异性的情况. 其预测结果可以表示为:

$$z^*(x_0) = z(x_i) \quad \text{when} \quad d(x_0 - x_i) < d(x_0 - x_j), \quad (2)$$

其中, $z^*(x_0)$ 为待预测点的值, $z(x_i)$ 为待预测点所在泰森多边形控制点的值.

自然邻点方法使用与待预测点控制的泰森多边形相交的多个泰森多边形进行预测 (Sibson, 1981). 具体而言, 自然邻点方法通过计算与待预测点控制的泰森多边形相交的其他泰森多边形的面积比例来确定权重, 预测公式为:

$$z^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i), \quad (3)$$

其中 λ_i 为第 i 个泰森多边形的权重, $z(x_i)$ 为第 i 个泰森多边形控制点的值, n 为与待预测点相邻的泰森多边形数量, λ_i 的计算基于面积比例:

$$\lambda_i = \frac{A_i}{\sum_{j=1}^n A_j}, \quad (4)$$

其中 A_i 为第 i 个泰森多边形与待预测点控制的泰森多边形相交的面积.

最近邻点方法主要优点在于计算简单, 能够快速获得预测结果 (李飞, 2015). 这种方法通过将每个待预测点的值直接赋予距离其最近的已知数据点的值, 因此计算过程非常高效, 适合于处理大规模数据集的应用场景. 然而, 最近邻点方法也存在显著的缺点, 限制了其在数据预测应用中的适用性. 最明显

的问题在于预测结果呈现阶梯状变化, 即在多边形的边界处预测点的值会发生突变, 而在多边形内部则保持恒定. 这种不连续性使得预测结果显得不自然, 尤其在需要平滑和连续变化的数据预测任务中, 这种方法的局限性尤为突出.

相比之下, 自然邻域方法作为一种更加复杂的预测方法, 其优点在于预测结果较为平滑, 能够较好地反映空间数据的连续变化. 自然邻域方法通过计算待预测点与其周围已知点之间的相邻关系和相交面积来进行加权平均, 从而获得预测结果, 能够捕捉到数据的空间变异性, 使得预测结果更加自然和精确, 适用于空间变异性较强的数据预测. 然而, 这种方法的计算复杂度相对较高, 尤其是在数据点数量较多时, 计算效率会显著降低. 构建维诺图和计算相交面积需要较大的计算资源, 因此在进行大规模数据集的数据预测时, 自然邻域方法可能面临效率问题.

2 反距离权重方法

反距离权重方法 (IDW) 是一种经典的空間预测方法, 最早由 Donald Shepard 在 1960 年代提出 (Shepard, 1968). IDW 的基本思想是利用待预测点周围的已知点进行预测, 通过与待预测点的距离远近来选定预测参考点. 公式如下:

$$z^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i), \quad (5)$$

其中, $z^*(x_0)$ 是待预测点的估计值, $z(x_i)$ 是已知点的值, λ_i 是权重, 与已知点与待预测点的距离相关:

$$\lambda_i = \frac{h_i}{\sum_{i=1}^n h_i}, \quad (6)$$

$$h_i = \frac{1}{d_i^k}, \quad (7)$$

其中, d_i 是待预测点与已知点之间的距离, k 是幂次参数, 常取 2 (张明明等, 2015).

IDW 因其计算简便和易于实现而广受欢迎, 尤其适用于数据点分布密集的场所. 然而, IDW 通常采用圆形搜索域, 这可能忽略了数据的空间方向性变化, 限制了其处理各向异性数据的能力. 如图 1 所示, 为了克服这一局限, 可以对搜索域进行优化, 将其形状从圆形调整为椭圆形, 并根据数据的空间方向性调整椭圆的长轴和短轴, 以实现对各向异性数据的更优适应 (Merwade *et al.*, 2006; Tomeczak, 1998; 颜金彪等, 2021).

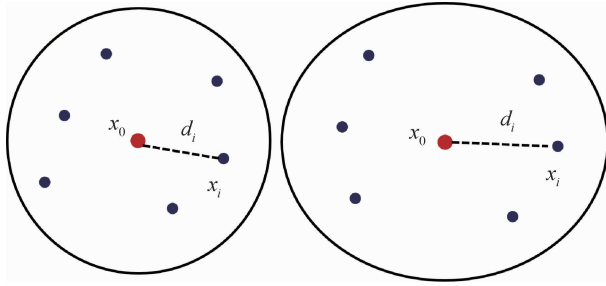


图1 搜索圆域和搜索椭圆域示意图

Fig. 1 Schematic diagrams of searching circular domain and searching elliptic domain

在传统 IDW 方法中,幂次参数是预先设定且固定的,但最佳值可能因数据特性的不同而有所变化. 为了适应这种变化,可以采用数据定制化的自适应调整机制,利用数据的局部统计特性或优化算法来动态确定幂次参数 (Lu and Wong, 2008). 此外,将 IDW 方法与机器学习技术相结合,可以进一步提升权重系数的自动调整能力 (Ma et al., 2019). 通过应用神经网络或回归模型,可以从大量空间数据中学习复杂的分布模式,并提取深层特征,从而优化权重分配,提高预测的准确性和可靠性.

IDW 方法能够高效地处理大规模数据集,可以迅速提供预测结果,满足应用需求. 通过基于距离的权重计算,减少了对参数调节的依赖,简化了模型的配置和应用过程,这种特性显著提升了 IDW 方法的便捷性和实用性,使其在实际应用中广受欢迎. 然而, IDW 方法在数据预测中也面临着一些挑战. 首先, IDW 的权重计算完全依赖于距离,这可能导致极值点和噪声点对预测结果产生过度的影响. 当数据质量不佳或存在异常值时,这种依赖性可能导致预测结果的可靠性降低. 此外, IDW 方法未能充分考虑数据的局部变异性,仅依据距离进行权重分配. 在数据特性变化较大的区域,这种方法可能导致预测误差的增加.

3 克里金方法

克里金方法在地球科学领域扮演着至关重要的角色. 最初在 20 世纪 50 年代, Danie Krige 提出了该方法的原型,并将其应用于矿石储量的估计 (Bárdossy, 1997; Journel and Huijbregts, 1978; Kitanidis, 1997). 随后, George Matheron 引入了地质统计学的概念,为克里金方法提供了坚实的理论基础,并将该方法正式命名为克里金方法 (Matheron,

1963). 假设有某一区域化变量 $z(x)$, 该区域内有 n 个已知点 $\{x_i; i = 1, 2, \dots, n\}$ 及待预测点 x_0 , 待预测点的估计值用表示 $z^*(x_0)$. 克里金方法的核心公式表达如下所示:

$$z^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i), \quad (8)$$

$$\gamma_{ij} = \frac{1}{2} E[(z(x_i) - z(x_j))^2], \quad (9)$$

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} & 1 \\ r_{21} & r_{22} & \cdots & \gamma_{2n} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ -\varphi \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \vdots \\ \gamma_{n0} \\ 1 \end{bmatrix}, \quad (10)$$

其中, γ_{ij} 为半方差函数, φ 表示拉格朗日乘子.

在克里金方法体系中,精确地确定半方差函数是实现有效空间预测的关键环节. 地理学托布勒第一定律揭示了一个基本原则:空间上相互靠近的点在属性上往往表现出较高的相似性 (Oliver and Webster, 1990). 半方差函数正是量化这种空间相似性的重要工具. 利用函数拟合方法,可以构建一个描述半方差与空间距离关系的数学模型,该模型用于计算待预测点与周围已知点间的半方差值,这些值进一步被应用于公式 (10),以计算得出用于预测的权重系数. 这一过程蕴含了克里金方法背后的统计学理论,并体现了其在解析复杂空间数据时的灵活性与适应力. 这一特性使得克里金方法成为地球科学以及其他多个领域中进行空间数据分析的宝贵工具,其应用价值在多个学科领域中得到了广泛认可和应用 (张隽, 2012). 克里金方法经过改进之后还能有效处理具有各向异性特征的空间数据 (姚凌青等, 2009; 王亭, 2013). 一种当前较为先进的方法是将克里金方法和空间变化量分解方法相结合. 首先将变量拆分为两个关键组成部分:基础变量和残差项,这种分解策略旨在捕捉和适应空间数据的趋势变化 (Li et al., 2022; Yang and Jia, 2024). 其中,基础变量部分负责反映区域的空间趋势和各向异性特征,这部分可以通过应用先验知识及数学建模方法进行拟合. 而残差项则表现为各向同性,可以通过标准的克里金方法进行有效估计. 通过这种分解,克里金方法能够更精确地模拟空间数据的内在结构,从而提高预测的准确性和可靠性. 克里金方法也能够与深度学习技术相结合,并展现出了巨大的应用前景. 通过采用神经网络对克里金方法的关键系数进行精细拟合,不仅提升了神经网络模型的可解释

性,还显著提高了预测的准确性(Liu *et al.*, 2022).

尽管克里金方法在空间数据分析中具有显著优势,但其实施过程确实存在一些局限性.该方法涉及复杂的数学和统计运算,如半方差函数的精确拟合和协方差矩阵的构建.在处理大规模数据集时,这些计算可能带来相当高的计算成本,这在一定程度上限制了克里金方法在大数据环境下的广泛应用.此外,克里金方法的效果在很大程度上依赖于半方差函数的恰当选择和参数的准确估计.这些参数对预测的准确性有着直接的影响.如果参数估计不准确,可能会导致预测结果的偏差,从而影响模型的可靠性.因此,进行细致的参数调整、模型验证是确保克里金方法效果的关键步骤,而这些步骤本身也需要大量的计算资源和专业知识.克里金方法还需要大量的高质量数据来支撑其预测的准确性和可靠性.数据的数量不足或质量不高都可能削弱预测结果的有效性,特别是在数据分布不均或包含异常值的情况下,这种影响尤为明显.因此,在应用克里金方法之前,确保数据的充分性和质量是至关重要的前提条件.

4 离散光滑插值方法

离散光滑插值(DSI)方法是地质建模领域中一种广泛采用的预测技术,它在推动地质建模技术发展方面起到了积极作用(Frank *et al.*, 2007; Houling, 1994; Mallet, 1989).该算法的实施基于区域的离散拓扑模型.在算法的初步阶段,需要对构建的拓扑模型及其节点值进行准确定义:设整个拓扑域的节点集合为 Ω ,其中集合中的每个元素 k 表示一个节点.对于任意节点 k ,其邻域 $N(k)$ 定义为所有在有限步数 s_k 内可达的点的集合,需要注意的是, $k \in N(k)$.

在离散光滑插值(DSI)中,待预测点值的确定是通过优化一个特定的准则函数 $R^*(\varphi)$ 来实现的(Mallet, 1992).该函数的目标是寻找一组属性值 $\varphi(k)$,这组值能满足给定的全局粗糙度函数 $R(\varphi)$ 及约束条件 $\rho(\varphi)$ 最小化的,表达式为:

$$R^*(\varphi) = R(\varphi) + \rho(\varphi). \quad (11)$$

$R(\varphi)$ 的定义为公式(12),其中 $\mu(k)$ 为一个给定的正加权函数,在没有特定理由对不同节点进行区分的情况下,建议取值为1.全局粗糙度函数是通过所有节点的局部粗糙度加权求和得到的,这反映了整个预测域的平滑性. $R(\varphi|k)$ 为定义在节点 k 上的局部粗糙度函数,其表达式如公式(13)所示,

$\{v^\alpha(k)\}$ 是给定的权重系数,用于调节节点邻域对该节点粗糙度的贡献.公式(12)、(13)为:

$$R(\varphi) = \sum_k \mu(k) R(\varphi|k), \quad (12)$$

$$R(\varphi|k) = \left| \sum_{\alpha \in N(k)} v^\alpha(k) \varphi(\alpha) \right|^2. \quad (13)$$

约束条件 $\rho(\varphi)$ 的数学表达通过公式(14)给出,其中 ϖ_i^2 是一个预设的正系数. ϖ_i^2 的取值大小直接影响了相应约束在优化过程中的重要性;取值越大,表明该约束条件在满足时的优先级越高.约束条件的累加构成了所有约束的总和 $\rho(\varphi)$,它们共同作用于预测过程,确保解的合理性和准确性.约束条件的选择是多样化的,可以是针对特定节点数值的直接约束,也可以是针对地质体表面或方向的约束.数值约束通常用于确保预测结果满足已知的数据点或地质统计信息,而方向约束则有助于保持地质体的连续性和一致性.通过精心设计的约束条件,DSI方法能够更好地适应地质建模的需求,实现对地质结构的精确模拟.公式(14)为:

$$\rho(\varphi) = \sum_i \varpi_i^2 |A_i \cdot \varphi - b_i|^2. \quad (14)$$

DSI方法的一个核心优势在于其能够灵活地融入多样的约束条件,包括地质边界条件和物理属性限制等(王述红等, 2023).这种灵活性使得DSI方法能够产生更加贴近实际情况的结果,从而增强了数据预测的准确性和可靠性,并被成功应用于地球物理属性的二维和三维预测(窦帆帆等, 2017; Mallet, 1997; 李冰等, 2002; 孟小红等, 2002; Tertois and Koren, 2020; Zhang *et al.*, 2023; Zheng *et al.*, 2019).通过不同的拓扑模型构建技术,DSI能够在特定的拓扑结构下提供更为精细的数学描述,例如,对于矩形拓扑情况的DSI理论已被进一步完善(Irakarama *et al.*, 2021).同样的,DSI方法的预测精度在很大程度上取决于离散拓扑的质量与精度.如果拓扑划分不够精细或存在不合理之处,可能会引起预测结果的失真.虽然细致的拓扑划分能够提升预测精度,但这也相应地增加了对计算和存储资源的需求(刘智勇, 2016).在处理大规模数据时,这种资源需求的增加可能成为实施的瓶颈,限制DSI方法的应用.此外,DSI方法主要适用于离散网格模型,这在一定程度上限制了其在不规则数据集和非网格化数据上的应用能力.在一些复杂的地质环境中,数据可能不适合用网格化的方式表示,这将影响DSI方法的适用性和预测效果.

5 径向基函数方法

径向基函数(RBF)方法是地质建模中一种广泛应用的方法,尤其在隐式地质建模势场法的应用中表现突出(Caumon *et al.*, 2013). 与传统的网格剖分方法不同,该方法不依赖于对计算域的网格化处理,而是依据待预测点及其邻近点的空间关系来确定预测结果(Calcagno *et al.*, 2008; Hillier *et al.*, 2014).

RBF 方法的核心思想在于,首先确定用于预测的函数基和一组特定的参考节点. 接着,将待预测点的值表达为这些参考节点处径向基函数值的线性组合(Buhmann, 2000),表达式为:

$$z^*(x_0) = \sum_{i=1}^n \lambda_i \varphi(x_0 - x_i), \tag{15}$$

$$\underbrace{\begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{bmatrix}}_{\Phi} \underbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}}_{\Lambda} = \underbrace{\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}}_Z, \tag{16}$$

其中 $\varphi(\cdot)$ 表示选取的基函数, $\|\cdot\|$ 表示向量的范数,通常采用欧几里得范数.

RBF 方法的关键是基函数的选择,其中立方体曲面函数和高斯函数是两种较为常用的选项. 立方体曲面函数形式简单,便于计算,而高斯函数则具有强大的拟合能力. 传统的 RBF 方法专注于数值解的获取,并未涉及梯度信息. 然而,在地球科学领域,还期望数值解的梯度能够满足特定的约束条件. 为了满足这一需求,Macêdo 等人提出了埃尔米特径向基函数(HRBF)预测方法,该方法在隐式地质建模中得到了广泛的应用(Guo *et al.*, 2018; Macêdo *et al.*, 2011). HRBF 方法引入了梯度约束,从而能够更好地满足地质应用中对数据连续性和平滑性的要求,如公式(17)所示,其中 H 为黑森矩阵, n_j 为梯度向量. 在 HRBF 方法中,立方体曲面函数因其计算效率和稳定性而成为较为常用的基函数选择(Wang *et al.*, 2024). HRBF 方法的引入,为数据预测提供了一种新的工具,它不仅能够保证数值精度,还能够确保预测模型的梯度满足地质学上的合理性. 这种方法的应用,显著提升了数据预测的质量和可靠性,为复杂地质结构下的数据预测供了强有力的支持. 公式(17)为:

$$z(x) = \sum_{i=1}^n (\alpha_i \varphi(\|x - x_i\|) - \langle \beta_i, \nabla \varphi(\|x - x_i\|) \rangle). \tag{17}$$

RBF 方法在处理大规模数据集时面临显著的计算复杂度问题. 由于每个待预测点都需要综合考虑大量已知点的影响,这导致计算时间和资源需求急剧上升. 这种高计算复杂度可能成为制约算法效率和实用性的关键瓶颈. RBF 方法的性能在很大程度上依赖于基函数的选择. 不同的基函数,例如立方体曲面函数和高斯函数,会显著影响预测结果的精度和稳定性. RBF 方法虽然能够处理复杂的空间关系,但在捕捉数据中的非线性特征方面存在局限. 它通常依赖于基函数的形式,而不是自适应地根据数据的特定特性进行优化. 这种依赖性在处理高度非线性的数据集时可能会限制 RBF 方法的有效性. 此外,RBF 方法在处理含噪声数据时也可能受到噪声的不利影响,这可能导致预测结果的失真. 尽管一些改进方法,如 HRBF 方法,通过引入梯度约束来增强结果的平滑性和连续性,但它们在处理噪声数据方面的效果仍有待提高.

6 深度学习方法

深度学习在数据预测领域的突破性进展已经吸引了业界的广泛关注,并在多个应用场景下取得了令人瞩目的研究成果. 本节将重点介绍完全基于深度学习的预测方法,并评估这些方法在多样化应用场景中的性能表现.

卷积神经网络(CNN)作为深度学习的核心技术,在计算机视觉领域已经发展出多种优秀的网络架构,为预测技术提供了丰富的参考(Goodfellow *et al.*, 2014; He *et al.*, 2016; Wang *et al.*, 2021). Zhu 等人引入了生成对抗网络(GAN)的概念,提出了条件编码-解码生成对抗神经网络,成功实现了数字高程模型的高精度内插(Zhu *et al.*, 2020). 然而,Zhu 等人的工作主要集中于栅格数据,而在许多实际情况中,数据并非栅格化,这限制了传统 CNN 的应用范围. 因此,多层感知机(MLP)作为一种替代的深度学习方法,通过将空间坐标作为网络输入,预测值为输出,为完成预测任务提供了新的思路(Rosenblatt, 1958). Difo 等人 and Hillier 等人分别设计了 MLP 模型来进行布格重力异常和地质建模预测,展示了 MLP 在预测任务中的潜力和灵活性(Difo *et al.*, 2022; Hillier *et al.*, 2023). 尽管 MLP 结构简单,易于实现,但其输入仅为点信息,可能忽略了对提高预测精度至关重要的区域空间信息. 为了解决这一问题,可以借鉴自然语言处理(NLP)领域的技术,其中词元位置信息的处理与空间预测中区域空间信息的

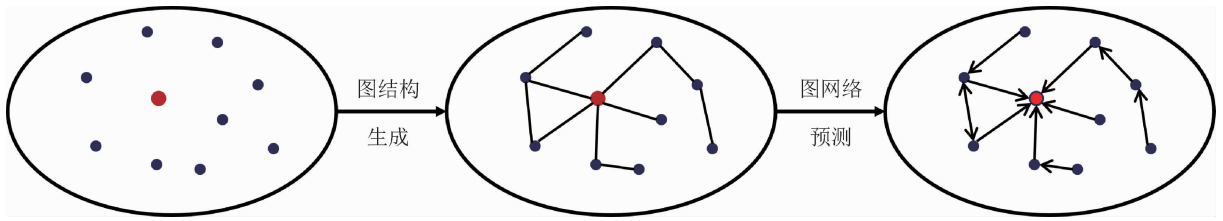


图2 基于图数据结构的数据预测示意图

Fig.2 Schematic of data prediction based on graph data structure

考量有相似之处 (Elman, 1990; Hochreiter and Schmidhuber, 1997). 循环神经网络 (RNN) 和长短期记忆网络 (LSTM) 通过设计能够捕捉序列中远距离依赖的机制, 为处理空间数据提供了灵感. 此外, Transformer 架构的提出及其在 BERT、GPT 等模型中的应用, 为深度学习领域带来了革命性的影响 (Devlin *et al.*, 2019; Radford *et al.*, 2018, 2019; Vaswani *et al.*, 2017). Li 等人受到这些工作的启发, 基于 Transformer 构建了类 BERT 的模型, 开发了一种自监督预测神经网络, 并成功将其应用于降雨量数据的空间预测 (Li *et al.*, 2023b).

尽管如此, 现有方法在处理空间位置信息时仍存在局限性. 图数据结构以其对散乱分布数据的自然表征, 为预测提供了新的视角. 图2构建了待预测点与周围点的图数据结构, 其中箭头表示预测过程中节点间信息的流动 (West, 2001).

通过将属性信息作为节点特征, 并通过边与周围节点交换信息, 可以有效地获取空间数据的全局和局部特征, 最终通过特征映射得到精确的预测结果 (Zhang *et al.*, 2022). 这种方法不仅充分利用了图数据结构的优势, 还结合了深度学习的强大能力, 为空间预测领域带来了新的可能性和研究方向. Appleby 等人提出了一种结合图卷积和图注意力机制的神经网络结构, 用于执行预测任务, 并通过数学推导验证了其与传统克里金方法的一致性 (Appleby *et al.*, 2020). Wu 等人进一步利用图神经网络 (GNN) 设计了归纳图神经网络, 用于时空克里金预测, 展现了 GNN 在处理时空数据方面的潜力 (Wu *et al.*, 2021). 然而, 这些方法通常需要为待预测点提供额外的特征信息. 在实际应用中, 往往只有待预测点的空间位置信息, 而缺乏其他特征数据. 针对这一问题, Li 等人开发了一种不依赖于额外属性信息的空间预测图神经网络模型, 该模型在没有附加特征的情况下, 相较于 Appleby 和 Wu 等人的方法, 展现

出了更优的实验性能 (Li *et al.*, 2023a). 此外, Hillier 等人也将图神经网络技术应用于地质建模中的势场预测问题, 证实了该方法在势场预测中的有效性和适用性 (Hillier *et al.*, 2021). 这些研究表明, 图神经网络因其在捕捉空间关系和处理非规则数据方面的优势, 已成为数据预测领域的一个重要研究方向. 随着图神经网络技术的不断进步, 其在数据预测领域的应用前景将更加广阔.

尽管深度学习技术展现出了巨大的潜力和优势, 但其应用过程中仍存在一些固有的局限性. 首先, 深度学习模型的高效训练往往依赖于大量高质量的标注数据. 在数据预测任务中, 收集足够的标注数据是一项艰巨的任务. 此外, 训练深度学习模型, 特别是复杂的网络结构如卷积神经网络 (CNN)、循环神经网络 (RNN) 和 Transformer, 需要显著的计算资源, 包括高性能图形处理单元 (GPU) 和大量的内存. 这种对高计算资源的依赖性可能阻碍了深度学习模型在资源受限的环境中的部署和应用. 尽管这些模型在处理复杂和高维数据集方面具有显著优势, 但它们对数据中的噪声和异常值也非常敏感. 如果未能妥善处理这些噪声数据, 模型的预测性能可能会受到严重影响, 进而影响预测结果的可靠性. 此外, 深度学习模型的预测能力通常局限于训练数据的分布范围. 对于训练过程中未遇到的新情况或变化, 模型可能无法做出准确预测.

7 实例分析

在构建地下电阻率模型的过程中, 选择恰当的数据预测方法能够显著降低地质勘探中的不确定性, 提升数据质量和模型的准确性, 为构建完整的三维地下结构模型和实现多元信息融合提供必要的数据支撑. 半航空电磁勘探方法, 结合了地面和航空勘探的优势, 利用无人机进行数据采集, 不仅能提高测线的密度, 还能扩大勘探的覆盖范围, 为创建高精度

的地下电阻率模型提供了丰富的数据资源(薛国强等, 2021; Xue *et al.*, 2019). 然而, 尽管该方法在数据采集方面具有显著优势, 其层间分辨率仍然低于层内分辨率. 为了提高层间分辨率, 需要通过有效的数据预测技术来增强层间分辨率.

基于半航空电磁观测数据采集的优势, 本节使用实际案例对一系列预测方法在层间数据预测任务中的性能进行全面的评估. 数据采集工作在广西百色靖西市某地进行, 该区域测线间的地质构造过渡连续, 为对比实验提供了理想的条件, 如图 3 所示(武欣等, 2023). 实验采集并分析了共 19 条半航空瞬变电磁测线的数据, 其中偶数编号的测线被选定为待预测平面. 考虑到工程勘探的实际需求, 测线的布置是规律性的, 并且覆盖范围有限, 这与深度学习数据集构建所需的随机性要求不符, 因此, 未能将基于深度学习的预测方法纳入对比分析. 本实验选取了维诺图方法(自然邻域方法、NN)、反距离权重方法(IDW)、克里金方法(OK)、离散光滑插值方法(DSI)和径向基函数方法(RBF)这些经典方法进行综合对比与分析.



图3 发射源及测线布局图

Fig. 3 Layout of transmitting source and survey lines

为了客观量化不同方法的预测结果差异, 本文采用了三个主要的评价指标: 平均相对误差(MRE)、平均绝对误差(MAE)和均方根误差(RMSE). 这些指标的计算公式如下:

$$MRE = \frac{1}{n} \sum_{i=1}^n \left| \frac{z_i - z_i^*}{z_i} \right|, \tag{18}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |z_i - z_i^*|, \tag{19}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (z_i - z_i^*)^2}{n}}, \tag{20}$$

MRE 的优势在于其能够以相对比例的形式反映预测值与实际值之间的偏差, 这使得它适用于不同数量级的目标变量, 可以提供一个统一的误差度量标准. MAE 通过计算误差的绝对值的平均, 提供了对预测误差的直接度量, 但其对异常值的敏感度较低. RMSE 通过平方误差来增加对较大偏差的惩罚, 从而赋予了较大的误差更大的权重, 这使得 RMSE 对数据中的异常值更为敏感, 能够反映出预测结果中极端偏差的影响.

表1 不同预测方法的误差
Table 1 Errors of different prediction methods

	NN	IDW	OK	DSI	RBF
MRE	0.1289	0.1291	0.1286	0.1300	0.1298
MAE	140.7323	140.9101	140.4600	142.8174	142.1160
RMSE	236.2735	236.0974	235.5265	241.5137	220.0570

通过计算分析, 获得各种方法的预测误差统计数据, 结果如表 1 所示. 结果表明, 克里金方法具有最低的 MRE 和 MAE, 表明该方法在维持整体误差和绝对误差的最小化方面具有显著优势. 此外, 克里金方法的 RMSE 也仅次于最低值, 进一步证实了其预测结果的稳定性和可靠性. 在 MRE 和 MAE 的评价指标中, 维诺图方法表现次优, 这说明该方法在整体误差控制方面具有较好的表现. 然而, 该方法的 RMSE 相对较高, 这说明在预测过程中可能得到较为明显的异常值. 径向基函数方法在 MRE 和 MAE 方面表现一般, 但其 RMSE 值最低, 表明该方法能够提供整体上较为稳定的预测结果. 相比之下, 离散光滑插值方法在所有考虑的误差指标中均表现不佳, 其 MRE、MAE 和 RMSE 均处于最高水平. 这表明离散光滑插值方法在当前应用场景下可能不是最佳选择, 其在属性预测领域的适用性和效果值得进一步的研究和探讨. 反距离权重方法的 MRE、MAE 和 RMS 表现虽然都较为一般, 但该方法在计算资源的需求和计算时间上具有显著优势. 在选择合适的预测方法时, 不仅需要考虑到预测精度, 还需权衡计算效率和资源消耗. 对于需要快速响应的应用场景, 反距离权重方法可能因其较低的计算成本而更为适宜; 而在对预测精度有更高要求的场合, 则可能需要考虑克里金或径向基函数等方法.

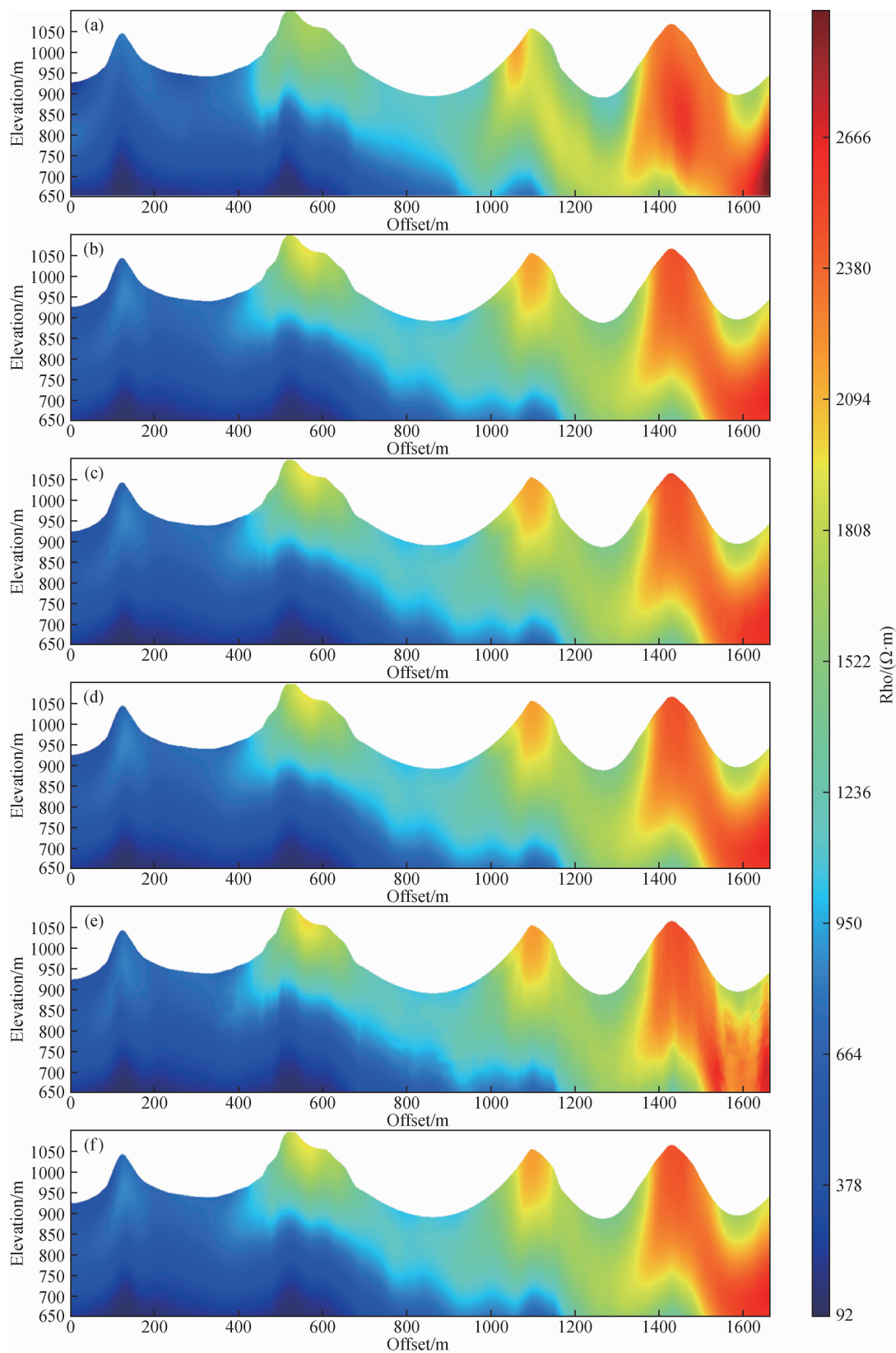


图4 测线14原始数据及数据预测结果对比图
(a)原始数据;(b)维诺图方法预测结果;(c)反距离权重方法预测结果;(d)克里金方法预测结果;
(e)离散光滑插值方法预测结果;(f)径向基函数方法预测结果.

Fig.4 Comparison of the original data and data prediction results for the measured line 14
(a) Original data;(b) Prediction results of the NN method;(c) Prediction results of the IDW method;(d) Prediction results of the OK method;
(e) Prediction results of the DSI method;(f) Prediction results of the RBF method.

为了进一步深入分析各方法在数据预测方面的表现,提供对预测结果分布特征的直观理解,针对测线 14 进行了预测结果的展示,如图 4 所示. 通过图 4 可以发现,除了离散光滑插值方法之外,其他方法在整体上能够与实测数据保持较好的一致性. 尽管这些方法在宏观层面的预测表现具有一定优势,能够提供整体趋势的合理估计,但它们在捕捉局部细节方面的能力仍有局限. 特别是在测线 14 的 1000 ~ 1200 m 处,预测得到的高电阻率区域出现了显著的偏差,原本应位于左侧山脊的电阻率异常区域被错误地预测在了山峰上. 这一现象的根源在于传统预测方法在数据预测过程中的一个关键局限性:它们主要依赖于待预测点附近的局部观测数据,而未能充分整合整个区域的全局信息或地质先验知识. 正是由于这些方法在捕捉全局信息或地质先验知识方面的能力受限,导致其在细节层面的预测精度未能满足精细化预测的要求. 因此,为了提高在不同地质条件下的预测精度,需要探索如何将全局信息或地质先验知识更有效地整合到数据预测过程中.

鉴于当前大型模型技术的快速发展,一个具有现实可能性的创新解决方案是构建专门针对地球科学领域的大型数据预测模型. 该方案的实施可分为以下几个阶段:预训练阶段:首先,利用现有的海量且多样化的地球科学数据资源对模型进行预训练,旨在培养其捕捉数据特征和进行有效预测的能力. 地区尺度微调:其次,在地区尺度上对预训练模型进行第一次针对性的微调,以提高模型捕捉特定区域独特地质特征的能力. 测区尺度微调:进一步地,在测区尺度上,结合现有的数据对模型进行第二次微调,以优化模型对局部地质特征的敏感性和响应能力. 最终预测:经过两次微调后,将待处理的数据输入模型中,以期获得精确的数据预测结果. 这种预训练的大型模型因其规模和复杂性,具备了捕捉和整合数据全局信息的强大能力. 通过两次微调过程,模型不仅保留了对全局信息的敏感性,还融入了地质先验知识,这有望显著提升数据预测的准确性和可靠性,为地球科学领域的研究和应用带来深远的影响.

8 结 语

多元信息融合的实现基础是涉及的数据具有相同的观测密度. 然而,在实际应用中,因为地球物理观测存在方法、场景、装备等多方面差异,所获取的数据在空间展布规范上彼此不同,从而构成了实现

多元信息融合的障碍. 为了克服这一障碍,必须对不同来源和方法得到的数据进行适当的预处理,以确保数据具有相同的观测密度,以支持有效的多元信息融合. 尽管目前尚未有专门针对多元信息融合的数据预测方法,现有的插值计算方法在一定程度上能够满足这一需求. 因此,本文立足于地球物理的多元信息融合,深入分析了现有的预测技术,并探索了它们在实际数值预测任务中的应用潜力,以期多元信息融合领域提供新的视角和解决方案. 案例研究表明,传统预测技术主要依赖于局部区域的信息,未能充分融合全局信息或地质学的先验知识. 这导致这些方法在宏观层面能够提供相对满意的预测结果,但在细节上的表现并不尽如人意. 鉴于当前大模型技术的快速发展,可能的一种解决方案是开发一个专门针对地球科学数据预测的大模型. 对特定的数据预测任务经过合适的微调之后,大模型不仅能保留对全局信息的敏感性,还融入了测区的地质先验知识,从而有望显著提升数据预测的准确性和可靠性.

致 谢 感谢审稿专家提出的修改意见和编辑部的大力支持!

References

- Appleby G, Liu L F, Liu L P. 2020. Kriging convolutional networks. // Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI, 3187-3194, doi: 10.1609/aaai.v34i04.5716.
- Bárdossy A. 1997. Introduction to geostatistics. Institute of Hydraulic Engineering, University of Stuttgart, 28-32.
- Bokareva T, Hu W, Kanhere S, *et al.* 2006. Wireless sensor networks for battlefield surveillance. // Proceedings of the Land Warfare Conference 2006. Brisbane, 1-8.
- Buhmann M D. 2000. Radial basis functions. *Acta Numerica*, 9: 1-38, doi: 10.1017/S0962492900000015.
- Calcagno P, Chiles J P, Courrioux G, *et al.* 2008. Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules. *Physics of the Earth and Planetary Interiors*, 171 (1-4): 147-157, doi: 10.1016/j.pepi.2008.06.013.
- Caumon G, Gray G, Antoine C, *et al.* 2013. Three-dimensional implicit stratigraphic model building from remote sensing data on tetrahedral meshes: theory and application to a regional model of La Popa Basin, NE Mexico. *IEEE Transactions on Geoscience and Remote Sensing*, 51 (3): 1613-1621, doi: 10.1109/TGRS.2012.2207727.
- Devlin J, Chang M W, Lee K, *et al.* 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv: 1810.04805*.
- Di Q Y, Yang C C, Zhu R X. 2012. Key technology development of deep resources exploration and field experimentation. *Bulletin of Chinese Academy of Sciences (in Chinese)*, 27 (3): 389-394, doi: 10.3969/j.issn.1000-3045.2012.03.019.
- Diffio S, Lemotio W, Mezoue Adiang C, *et al.* 2022. Contribution of the artificial neural network (ANN) method to the interpolation of the

- Bouguer gravity anomalies in the region of Lom-Pangar (East-Cameroon). *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, 8(1): 32, doi: 10.1007/s40948-022-00341-1.
- Dou F F, Lin Z Y, Ye Z H. 2017. Study of 3D visualization of discrete data based on GOCAD. *Acta Geologica Sichuan* (in Chinese), 37(1): 159-163.
- Dreus F, Feng D, Faion F, *et al.* 2022. DeepFusion: A robust and modular 3D object detector for lidars, cameras and radars. //2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Kyoto: IEEE, 560-567, doi: 10.1109/IROS47612.2022.9981778.
- Elman J L. 1990. Finding structure in time. *Cognitive Science*, 14(2): 179-211.
- Frank T, Tertois A L, Mallet J L. 2007. 3D-reconstruction of complex geological interfaces from irregularly distributed and noisy point data. *Computers & Geosciences*, 33(7): 932-943, doi: 10.1016/j.cageo.2006.11.014.
- Goodfellow I J, Pouget-Abadie J, Mirza M, *et al.* 2014. Generative adversarial nets. //Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: MIT Press, 2672-2680.
- Goovaerts P. 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, 228(1-2): 113-129, doi: 10.1016/S0022-1694(00)00144-X.
- Guo J T, Wu L X, Zhou W H, *et al.* 2018. Section-constrained local geological interface dynamic updating method based on the HRBF surface. *Journal of Structural Geology*, 107: 64-72, doi: 10.1016/j.jsg.2017.11.017.
- Hall D L, Llinas J. 1997. An introduction to multisensor data fusion. *Proceedings of the IEEE*, 85(1): 6-23, doi: 10.1109/5.554205.
- He K M, Zhang X Y, Ren S Q, *et al.* 2016. Deep residual learning for image recognition. //Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 770-778.
- Hillier M, Wellmann F, Brodaric B, *et al.* 2021. Three-dimensional structural geological modeling using graph neural networks. *Mathematical Geosciences*, 53(8): 1725-1749, doi: 10.1007/s11004-021-09945-x.
- Hillier M, Wellmann F, de Kemp E A, *et al.* 2023. GeoINR 1.0: an implicit neural network approach to three-dimensional geological modelling. *Geoscientific Model Development*, 16(23): 6987-7012, doi: 10.5194/gmd-16-6987-2023.
- Hillier M J, Schetselaar E M, de Kemp E A, *et al.* 2014. Three-dimensional modelling of geological surfaces using generalized interpolation with radial basis functions. *Mathematical Geosciences*, 46(8): 931-953, doi: 10.1007/s11004-014-9540-3.
- Hochreiter S, Schmidhuber J. 1997. Long short-term memory. *Neural Computation*, 9(8): 1735-1780.
- Houlding S W. 1994. 3D Geoscience Modeling: Computer Techniques for Geological Characterization. Berlin, Heidelberg: Springer.
- Irakarama M, Laurent G, Renaudeau J, *et al.* 2021. Finite difference implicit structural modeling of geological structures. *Mathematical Geosciences*, 53(5): 785-808, doi: 10.1007/s11004-020-09887-w.
- Journel A G, Huibregts C J. 1978. *Mining Geostatistics*. London: Academic Press.
- Kitanidis P K. 1997. *Introduction to Geostatistics: Applications in Hydrogeology*. Cambridge: Cambridge University Press.
- Lam N S N. 1983. Spatial interpolation methods: a review. *The American Cartographer*, 10(2): 129-150.
- Lejeune Dirichlet G. 1850. über die Reduction der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1850(40): 209-227.
- Li B, Liu H, Li Y M. 2002. 3-D seismic data discrete smooth interpolation using conjugate gradient. *Chinese Journal of Geophysics* (in Chinese), 45(5): 691-699.
- Li F. 2015. Study on the technology of spatial data interpolation method of mine transient electromagnetic [Master's thesis] (in Chinese). Xuzhou: China University of Mining and Technology.
- Li F, Zheng Q, Wang M, *et al.* 2016. Current status and development trends of multi-source geological information fusion technology in mineral exploration. //Proceedings of the 2016 China Earth Science Joint Academic Annual Conference (in Chinese). Beijing, 545.
- Li J, Shen Y Y, Chen L, *et al.* 2023a. Rainfall spatial interpolation with graph neural networks. //28th International Conference on Database Systems for Advanced Applications. Tianjin: Springer, 175-191, doi: 10.1007/978-3-031-30678-5_14.
- Li J, Shen Y Y, Chen L, *et al.* 2023b. SSIN: Self-supervised learning for rainfall spatial interpolation. *Proceedings of the ACM on Management of Data*, 1(2): 176, doi: 10.1145/3589321.
- Li X, Shi B T, Hou Y N, *et al.* 2022. Homogeneous multi-modal feature fusion and interaction for 3D object detection. //17th European Conference on Computer Vision. Tel Aviv: Springer, 691-707, doi: 10.1007/978-3-031-19839-7_40.
- Li Z Y, Tao H, Zhao D, *et al.* 2022. Three-dimensional empirical Bayesian Kriging for soil PAHs interpolation considering the vertical soil lithology. *Catena*, 212: 106098, doi: 10.1016/j.catena.2022.106098.
- Liu X Q, Su Y J, Hu T Y, *et al.* 2022. Neural network guided interpolation for mapping canopy height of China's forests by integrating GEDI and ICESat-2 data. *Remote Sensing of Environment*, 269: 112844, doi: 10.1016/j.rse.2021.112844.
- Liu Z Y. 2016. Research on surface interpolation algorithm in geological modeling [Master's thesis] (in Chinese). Chengdu: Chengdu University of Technology.
- Lu G Y, Wong D W. 2008. An adaptive inverse-distance weighting spatial interpolation technique. *Computers & Geosciences*, 34(9): 1044-1055, doi: 10.1016/j.cageo.2007.07.010.
- Ma J, Ding Y X, Gan V J L, *et al.* 2019. Spatiotemporal prediction of PM2.5 concentrations at different time granularities using IDW-BLSTM. *IEEE Access*, 7: 107897-107907, doi: 10.1109/ACCESS.2019.2932445.
- Macêdo I, Gois J P, Velho L. 2011. Hermite radial basis functions implecits. *Computer Graphics Forum*, 30(1): 27-42, doi: 10.1111/j.1467-8659.2010.01785.x.
- Mallet J L. 1989. Discrete smooth interpolation. *ACM Transactions on Graphics*, 8(2): 121-144, doi: 10.1145/62054.62057.
- Mallet J L. 1992. Discrete smooth interpolation in geometric modelling. *Computer-Aided Design*, 24(4): 178-191, doi: 10.1016/0010-4485(92)90054-E.
- Mallet J L. 1997. Discrete modeling for natural objects. *Mathematical Geology*, 29(2): 199-219, doi: 10.1007/BF02769628.
- Matheron G. 1963. Principles of geostatistics. *Economic Geology*, 58(8): 1246-1266, doi: 10.2113-3/gsecongeo.58.8.1246.
- Meng T, Jing X Y, Yan Z, *et al.* 2020. A survey on machine learning for data fusion. *Information Fusion*, 57: 115-129, doi: 10.1016/j.inffus.2019.12.001.
- Meng X H, Hou J Q, Liang H Y, *et al.* 2002. The fast realization of discrete smooth interpolation in the interpolation of potential data. *Geophysical and Geochemical Exploration* (in Chinese), 26(4): 302-306, doi: 10.3969/j.issn.1000-8918.2002.04.015.
- Merwade V M, Maidment D R, Goff J A. 2006. Anisotropic considerations while interpolating river channel bathymetry. *Journal of Hydrology*, 331(3-4): 731-741, doi: 10.1016/j.jhydrol.2006.06.018.
- Muzammal M, Talat R, Sodhro A H, *et al.* 2020. A multi-sensor data fusion enabled ensemble approach for medical data from body sensor networks. *Information Fusion*, 53: 155-164, doi: 10.1016/j.inffus.2019.06.021.
- Oliver M A, Webster R. 1990. Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, 4(3): 313-332, doi: 10.1080/02693799008941549.
- Radford A, Narasimhan K, Salimans T, *et al.* 2018. Improving language understanding by generative pre-training.
- Radford A, Wu J, Child R, *et al.* 2019. Language models are unsupervised multitask learners. *OpenAI Blog*, 1(8): 9.
- Rosenblatt F. 1958. The perceptron: a probabilistic model for

- information storage and organization in the brain. *Psychological Review*, 65(6): 386-408, doi: 10.1037/h0042519.
- Shepard D. 1968. A two-dimensional interpolation function for irregularly-spaced data. //Proceedings of the 1968 23rd ACM National Conference. New York: ACM, 517-524, doi: 10.1145/800186.810616.
- Sibson R. 1981. A brief description of natural neighbor interpolation. // Barnett V ed. *Interpreting Multivariate Data*. New York: John Wiley & Sons, 21-36.
- Tertois A, Koren Z. 2020. Solving problems with the discrete smooth interpolation framework, from geomodelling to geophysics and beyond. //First EAGE Digitalization Conference and Exhibition. Vienna: European Association of Geoscientists & Engineers, 1-5, doi: 10.3997/2214-4609.202032041.
- Tomczak M. 1998. Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW)-cross-validation/jackknife approach. *Journal of Geographic Information and Decision Analysis*, 2(2): 18-30.
- Vance T C, Huang T, Butler K A. 2024. Big data in Earth science: Emerging practice and promise. *Science*, 383(6688): eadh9607, doi: 10.1126/science.adh9607.
- Vaswani A, Shazeer N, Parmar N, *et al.* 2017. Attention is all you need. //Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 6000-6010.
- Velasco-Hernandez G, Yeong D J, Barry J, *et al.* 2020. Autonomous driving architectures, perception and data fusion: A review. //2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP). Cluj-Napoca: IEEE, 315-321, doi: 10.1109/ICCP51029.2020.9266268.
- Voronoi G. 1908. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les paralléloèdres primitifs. *Journal für die reine und angewandte Mathematik*, 134: 198-287.
- Wang S H, Li Y M, Yin H, *et al.* 2023. Improvement of GeoSMA-3D program based on multi-constrained DSI algorithm. *Journal of Northeastern University (Natural Science) (in Chinese)*, 44(2): 251-257, doi: 10.12068/j.issn.1005-3026.2023.02.013.
- Wang T. 2013. Research on 3D interpolation approaches based on Kriging with anisotropic spatial structures [Master's thesis] (in Chinese). Nanjing: Nanjing Normal University.
- Wang X L, Guo J T, Fu S H, *et al.* 2024. Towards automatic and rapid 3D geological modelling of urban sedimentary strata from a large amount of borehole data using a parallel solution of implicit equations. *Earth Science Informatics*, 17(1): 421-440, doi: 10.1007/s12145-023-01164-8.
- Wang Z H, Chen J, Hoi S C H. 2021. Deep learning for image super-resolution: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10): 3365-3387, doi: 10.1109/TPAMI.2020.2982166.
- West D B. 2001. *Introduction to Graph Theory*. 2nd ed. Upper Saddle River: Prentice Hall.
- Wu X, Xue G Q, Chen W Y, *et al.* 2023. Research on key technologies of the low-noise system for Semi-airborne electromagnetic method. *Chinese Journal of Geophysics (in Chinese)*, 66(9): 3904-3913, doi: 10.6038/cjg2023Q0891.
- Wu Y K, Zhuang D Y, Labbe A, *et al.* 2021. Inductive graph neural networks for spatiotemporal Kriging. //Proceedings of the 35th AAAI Conference on Artificial Intelligence. Virtually: AAAI, 4478-4485, doi: 10.1609/aaai.v35i5.16575.
- Wu Z C, Guo F S, Lin Z Y, *et al.* 2016. Technology and method of multi-data merging in 3D geological modeling. *Journal of Jilin University (Earth Science Edition) (in Chinese)*, 46(6): 1895-1913, doi: 10.13278/j.cnki.jjuese.201606308.
- Xia H S, Liu Z S, Efremochkina M, *et al.* 2022. Study on city digital twin technologies for sustainable smart city design: A review and bibliometric analysis of geographic information system and building information modeling integration. *Sustainable Cities and Society*, 84: 104009, doi: 10.1016/j.scs.2022.104009.
- Xu D Y, Wang D, Li J H, *et al.* 2023. The security connotation and index system of industry chain and supply chain of strategic mineral resources under the new development pattern. *Journal of Central China Normal University (Natural Sciences) (in Chinese)*, 57(1): 1-12.
- Xue G Q, Chen W, Cheng J L, *et al.* 2019. A review of electrical and electromagnetic methods for coal mine exploration in China. *IEEE Access*, 7: 177332-177341, doi: 10.1109/ACCESS.2019.2951774.
- Xue G Q, Li H, Chen W Y, *et al.* 2021. Progress of transient electromagnetic detection technology for water-bearing bodies in coal mines. *Journal of China Coal Society (in Chinese)*, 46(1): 77-85.
- Yan J B, Wu B, He Q H. 2021. An anisotropic IDW interpolation method with multiple parameters cooperative optimization. *Acta Geodaetica et Cartographica Sinica (in Chinese)*, 50(5): 675-684, doi: 10.11947/j. AGCS.2021.20200148.
- Yang Y, Jia M Y. 2024. 3D spatial interpolation of soil heavy metals by combining Kriging with depth function trend model. *Journal of Hazardous Materials*, 461: 132571, doi: 10.1016/j.jhazmat.2023.132571.
- Yao L Q, Pan M, Cheng Q M, *et al.* 2009. Nested overlap of variograms in 3D Kriging. *Earth Science (Journal of China University of Geosciences) (in Chinese)*, 34(2): 294-298.
- Zeng Q D, Di Q Y, Xue G Q, *et al.* 2021. Modern science and technology in metallogenic and prospecting model studies. *Earth Science Frontiers (in Chinese)*, 28(3): 295-308, doi: 10.13745/j.esf.sf.2021.1.3.
- Zhang J. 2012. Kring-based geophysical data mesh of three-dimensional space [Master's thesis] (in Chinese). Beijing: China University of Geosciences (Beijing).
- Zhang M M, Li X H, Yuan F, *et al.* 2015. Comparative study between three-dimensional strata modeling and surface model interpolation. *Geology of Anhui (in Chinese)*, 25(3): 182-186.
- Zhang Z Q, Wang G W, Carranza E J M, *et al.* 2023. An integrated machine learning framework with uncertainty quantification for three-dimensional lithological modeling from multi-source geophysical data and drilling data. *Engineering Geology*, 324: 107255, doi: 10.1016/j.enggeo.2023.107255.
- Zhang Z W, Cui P, Zhu W W. 2022. Deep learning on graphs: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 34(1): 249-270, doi: 10.1109/TKDE.2020.2981333.
- Zheng S J, Meng X H, Wang J, *et al.* 2019. Potential field data discrete smooth interpolation using conjugate gradient. //International Workshop on Gravity, Electrical & Magnetic Methods and Their Applications. Xi'an, China: Society of Exploration Geophysicists, 85-88, doi: 10.1190/GEM2019-022.1.
- Zheng Y. 2015. Methodologies for cross-domain data fusion: An overview. *IEEE Transactions on Big Data*, 1(1): 16-34, doi: 10.1109/TBDATA.2015.2465959.
- Zhu D, Cheng X M, Zhang F, *et al.* 2020. Spatial interpolation using conditional generative adversarial neural networks. *International Journal of Geographical Information Science*, 34(4): 735-758, doi: 10.1080/13658816.2019.1599122.

附中文参考文献

- 底青云, 杨长春, 朱日祥. 2012. 深部资源探测核心技术研发与应用. *中国科学院院刊*, 27(3): 389-394, doi: 10.3969/j.issn.1000-3045.2012.03.019.
- 窦帆帆, 林子瑜, 叶子华. 2017. 基于GOCAD的离散点云数据三维可视化研究. *四川地质学报*, 37(1): 159-163.
- 李冰, 刘洪, 李幼铭. 2002. 三维地震数据离散光滑插值的共轭梯度法. *地球物理学报*, 45(5): 691-699.
- 李飞. 2015. 矿井瞬变电磁空间数据体插值方法技术研究[硕士论文]. 徐州: 中国矿业大学.

李芳,郑强,王明,等. 2016. 多源地学信息融合技术在矿产勘查中的研究现状及发展趋势. //2016 中国地球科学联合学术年会论文集. 北京, 545.

刘智勇. 2016. 曲面插值算法在三维地质建模中的研究[硕士论文]. 成都: 成都理工大学.

孟小红,侯建全,梁宏英,等. 2002. 离散光滑插值方法在地球物理位场中的快速实现. 物探与化探, 26(4): 302-306, doi: 10.3969/j. issn. 1000-8918. 2002. 04. 015.

王述红,李友明,尹宏,等. 2023. 基于多约束 DSI 算法的 GeoSMA-3D 程序改进. 东北大学学报(自然科学版), 44(2): 251-257, doi: 10.12068/j. issn. 1005-3026. 2023. 02. 013.

王亭. 2013. 顾及各向异性的三维 Kriging 空间插值方法研究[硕士论文]. 南京: 南京师范大学.

武欣,薛国强,陈卫营,等. 2023. 半航空电磁法全流程低噪声关键技术研究. 地球物理学报, 66(9): 3904-3913, doi: 10.6038/cjg2023Q0891.

吴志春,郭福生,林子瑜,等. 2016. 三维地质建模中的多源数据融合技术与方法. 吉林大学学报(地球科学版), 46(6): 1895-1913, doi: 10.13278/j. cnki. jjuese. 201606308.

徐德义,王迪,李军辉,等. 2023. 新发展格局下战略性矿产资源产业链供应链安全内涵及指标体系研究. 华中师范大学学报(自然科学版), 57(1): 1-12.

薛国强,李海,陈卫营,等. 2021. 煤矿含水体瞬变电磁探测技术研究进展. 煤炭学报, 46(1): 77-85.

颜金彪,吴波,何清华. 2021. 顾及各向异性的多参数协同优化 IDW 插值方法. 测绘学报, 50(5): 675-684, doi: 10.11947/j. AGCS. 2021. 20200148.

姚凌青,潘懋,成秋明,等. 2009. 三维 Kriging 方法中的变异函数套合. 地球科学(中国地质大学学报), 34(2): 294-298.

曾庆栋,底青云,薛国强,等. 2021. 成矿模式与找矿模式研究的现代科学技术. 地学前缘, 28(3): 295-308, doi: 10.13745/j. esf. sf. 2021. 1. 3.

张隽. 2012. 基于克里金方法的地球物理三维空间数据网格化[硕士论文]. 北京: 中国地质大学(北京).

张明明,李晓晖,袁峰,等. 2015. 地层三维建模及面模型插值方法对比研究. 安徽地质, 25(3): 182-186.