

53-58

中国地质统计学(空间信息统计学)发展的回顾与前景

侯景儒

(北京科技大学地质系·100083)

p628.2



作者学术简历 侯景儒, 61 岁, 现任北京科技大学地质研究所教授, 中国金属学会冶金地质学会副理事长, 中国地质学会数学地质委员会副主任委员, 中国地质教育协会副秘书长。曾任北京科技大学地质系副主任。长期致力于地质统计学(空间信息统计学)及数学地质的研究、应用和教学工作, 先后负责由国家自然科学基金资助的科研项目和冶金部地质总局下达的科研项目等, 并多项获奖。多年来, 撰写出版了《矿床统计预测与地质统计学的理论与应用》(冶金工业出版社, 1993 年), 《地质统计学的理论与方法》(地质出版社, 1990 年), 《地质统计学及其在矿产储量计算中的应用》(1982 年, 冶金工业出版社), 《非参数及多元地质统计学的理论分析及应用》(冶金工业出版社, 1994 年) 等专著及教材 10 部; 并公开发表学术论文近 70 篇。

摘要 讨论了地质统计学(空间信息统计学)发展的若干问题, 其内容包括 4 部分: 1. 地质统计学的理论及方法; 2. 关于软件系统的现状; 3. 关于实际应用; 4. 地质统计学发展前景的若干看法。

关键词 地质统计学, 空间信息统计学, 克立格法, 数学地质, 变异函数

在中国乃至世界范围内, 地质科学已经由找矿为主拓宽到为解决当代社会发展面临的许多重大问题服务。社会对地质科学的要求越来越高, 地质工作者面临的任务日趋复杂和艰巨。为了满足人类社会对地质学的各种要求, 地质科学也正经历着重大的变革, 其变革的标志之一就是地质科学逐步向量化、精确化及自动化发展, 数学地质(Mathematical Geology)在某种意义上来说就是地质科学向前发展的必然结果, 因为, 数学地质是定量研究各种地质事件发生、发展的内在规律及空间形式的科学。

当前, 已经有许多数学地质的理论及方法用来研究和解决各种地质及其他问题, 地质统计学(Geostatistics)就是其中最具实用

意义的一种。地质统计学是以区域化变量理论(Theory of regionalized Variable)为基础, 以变异函数(Variogram)为基本工具来研究那些展布于空间并呈现出一定的随机性和结构性的自然现象(包括地质现象)的科学。显然, 凡是要研究某些变量(或特征)的空间分布特性并对其进行最优估计, 或要模拟所研究对象的离散性、波动性或其他性质时都可应用地质统计学的理论与方法。根据地质统计学理论及方法的内涵以及发展现状, 我认为把地质统计学理解或称之为空间信息统计学(Spatial-information Statistics)更为合适。

我国自 1978 年开始对地质统计学的理论、方法及应用的研究至今已近 20 年了, 我国的地统计学研究经历了曲折的历程之后,

在地质统计学的理论、方法及应用方面均达到了成熟的阶段。这里要特别指出的是,《地质与勘探》在地质统计学的推广、普及和应用等方面起到了十分重要的作用。

本文仅就地质统计学在中国的发展现状及前景进行初步探讨。

1 地质统计学的理论及方法研究取得了显著成绩,达到了较高水平

地质统计学的研究对象是区域化变量,而区域化变量具有随机性及结构性两重性质,对某区域化变量在研究区的变异性研究是通过计算其变异函数并进行结构分析得到的。地质统计学就是在结构分析的基础上采用各种地质统计学方法来研究各种实际问题,所采用的理论方法从总的趋势看分为两大类:一类是采用最优无偏估计方法进行估计,建立研究对象的估值模型;另一类则是采用各种模拟方法进行模拟,建立研究对象的模拟模型。我国在这两个方面均进行了卓有成效的研究。

1.1 最优无偏估计方法研究

地质统计学中应用的估计方法主要是各种克立格法(Kriging)。所谓克立格法就是一种求最优、线性或非线性、无偏内插估计值的方法,最基本和常用的方法是普通克立格法(Ordinary-kriging)。设任一待估域 V (或点)的某一区域化变量的真值 Z_i 的估计值为 Z_i^* , Z_i^* 可以通过该待估域 V (或点)影响范围内 n 个有效信息值 $Z(x_\alpha)$ 的线性组合得到:

$$Z_i^* = \sum_{\alpha=1}^n \lambda_\alpha Z(x_\alpha) \quad (1)$$

地质统计学方法与其他方法不同之处在于求 Z_i^* 时的权系数 λ_α 能保证 Z_i^* 为无偏线性估计值。而不同的克立格法的区别则在于根据区域化变量的变异特征、信息值的几何位置特征以及对估计精度要求的不同,为求得(1)式中 λ 必须求解不同的克立格方程组。我国

在这方面做了大量的研究工作并成功的应用于生产实际。例如,当信息值服从正态分布及平稳假设时,求解 λ 的普通克立格方程组是:

$$\begin{cases} \sum_{\alpha=1}^n \lambda_\alpha C(x_i, x_\alpha) - \mu = C(x_i, V) \\ \sum_{\alpha=1}^n \lambda_\alpha = 1 \quad (\beta=1, 2, \dots, n) \end{cases} \quad (2)$$

当信息值非平稳时,求解 λ 的泛克立格方程组:

$$\begin{cases} \sum_{\beta=1}^n \lambda_\beta C(x_i, x_\beta) - \sum_{\alpha=1}^n \lambda_\alpha f_\alpha(x_i) = \bar{C}(x_i, V) \\ \sum_{\alpha=1}^n \lambda_\alpha f_\alpha(x_i) = f_\alpha(V) \quad (\alpha=1, 2, \dots, n) \end{cases} \quad (3)$$

当信息值服从对数正态分布时,求解 λ 的是如下对数正态克立格方程组:

$$\begin{cases} \sum_{\alpha=1}^n \lambda_\alpha C(x_i, x_\alpha) - \mu = C(x_i, V) \\ \sum_{\alpha=1}^n \lambda_\alpha = 1 \end{cases} \quad (4)$$

当信息值不服从正态分布,其分布曲线呈现出长尾巴时,为了消除特异值(outlier)对估计值的过高影响,求(1)式中 λ_α 时,可求得如下指示克立格方程组:

$$\begin{cases} \sum_{\alpha=1}^n \lambda_\alpha C(x_i, x_\alpha; Z) - \mu = \bar{C}(x_i, V; Z) \\ \sum_{\alpha=1}^n \lambda_\alpha(Z) = 1 \quad (\alpha=1, 2, \dots, n) \end{cases} \quad (5)$$

为了提高估计精度,我们应用多个信息(多元信息)来求(1)式中的 λ_α 时,我们可以采用多元地质统计学,这时求 λ_α 的方程组则为更为复杂的协同克立格(cokriging)方程组:

$$\begin{cases} \sum_{k=1}^K \sum_{\alpha=1}^{n_k} \lambda_{k\alpha} \bar{C}_{k\alpha}(x_i, x_\alpha) - \mu_k = \bar{C}_{k\alpha}(V, x_i) \\ \sum_{\alpha=1}^{n_k} \lambda_{k\alpha} = 1 \quad (\alpha=1, 2, \dots, n_k) \\ \sum_{\alpha=1}^{n_k} \lambda_{k\alpha} = 0 \quad (k=1, 2, \dots, K) \end{cases} \quad (6)$$

在水文、石油、农林、环保等科学研究中,所研究的变量不仅具有空间特征,而且具有时间特征,这时,我们把研究的变量看成是时间—空间域的随机函数,从而给出时—空域中多元信息的地质统计学理论及方法。求(1)式中

的 λ_t 时,必须求得如下方程组:

$$\begin{cases} \sum_{t=1}^T \sum_{\alpha_i=1}^{n_t} \lambda_{\alpha_i} \bar{C}_{t,\alpha_i}(x_{t,\alpha_i}, x_{\alpha_i}) - \mu_t = \bar{C}_{t,\alpha_i}(V_{t,\alpha_i}, x_{\alpha_i}) \\ \quad (\alpha_i = 1, 2, \dots, n_t) \\ \sum_{\alpha_i=1}^{n_t} \lambda_{\alpha_i} = 1 \quad (t = 1, 2, \dots, T) \\ \sum_{\alpha_i=1}^{n_t} \lambda_{\alpha_i} = 0 \quad (t \neq t_0) \end{cases} \quad (7)$$

(7)式适用于空间域信息量大于时间域信息量,当时间域信息量大于空间信息量时,应对(7)式进行修正。

此外,为了适应我国矿产勘查工作中一些具体要求,我们还应用地质统计学的某些基本理论,提出一些可行的方法。例如,我们提出了在矿产储量计算中矿石储量分级的定量的地质统计学标志 R_t :

$$R_t = \bar{\sigma}_k^2 / D^2(Z) \quad (8)$$

我们提出了应用地质统计学方法选择最优勘探网度(网形)的方法。为了合理地确定储量计算中的特高品位,提高储量计算的精度,我们推荐采用基于地质统计学基本理论的邻域法来识别并处理特高品位(或称特异值),按照 Hawkins 1980 年提出的识别特异值的统计量: $I = n(G - M)^2 / (n + 1)\sigma^2$ (9) 识别特高品位是合理的。

在勘探阶段,我们采用地质统计学方法或其他方法计算的储量为原地资源(in situ reserves),而在进行采矿设计时,还要进行可回采储量(recoverable reserves)的估计,我们曾经详细的研究了可回采储量的总体估计方法,即采用线性地质统计学的有关方法把原地资源修正为可回采储量的总体估计量。可回采储量的局部估计对采矿设计及矿山地质等均十分重要,其估计方法属非线性平稳地质统计学范畴,其主要方法有二元高斯分布法和多元高斯分布法。我国有个别采矿设计部门只对二元高斯分布法进行了研究,其研究深度及应用的广度仍远不能满足需要,目

前尚处于探索研究阶段。

1.2 条件模拟(conditional simulation)

上述的各种克立格法均有各自的优点,它们在一定条件下能给出最优无偏估计值,但是,它们缺点是由(10)式(修匀关系式)表现出的修匀效应:

$$D^2(v/G) \approx D_k^2(v/G) + \bar{\sigma}_k^2 \quad (10)$$

从而无法给出某变量的真实值的离散程度,为此,地质统计学家提出了条件模拟这一方法。

众所周知,地质过程的数据模拟是基于所研究过程的主要地质条件及特征进行各种模拟,构造出诸如确定性静态模型及动态模型,和随机静态模型及动态模型,当地质过程既有确定性因素又有随机性因素时,则可以构造出混合性模型。在随机性模型中最常用的是蒙特-卡洛法。

地质统计学中的数据模型比诸如蒙特-卡洛法更精确更有用。因为地质统计学中的随机模拟具有以下特征:(1)它要求伪随机数服从一定的概率分布,具有一定的数学期望及方差;(2)它要求保持一定的空间自相关性,即保持与实际数据有相同的协方差函数或变异函数,这是因为区域化变量既有随机性又有空间结构性;(3)它要求在各观测点处的模拟值等于该点处的实测值,满足前两条者称为“非条件模拟”。上述三条均满足者称为“条件模拟”。

地质统计学中的上述数据模拟比传统的随机模拟更有特点,表现在:(1)它保持变量的空间自相关性(协方差函数及变异函数)不变,从而更适于模拟区域化变量;(2)使观测点上的模拟值等于实测值,因而,观测点越多,模拟就越接近实际;(3)它有可能实现三维空间模拟,因为,已经提出的“转向带法”可以把一维模拟扩展到三维模拟,节省机时及内存;(4)通过条件模拟可以求出与实际值一样的协方差函数或变异函数,从而可重现实际值的离差方差。

条件模拟已经发展成为地质统计学的重要组成部分,它的基本理论方法在煤田地质、石油地质及其他领域得到广泛的应用。

2 地质统计学计算程序现状分析

由于地质统计学理论及各类克立格法比较复杂,地质统计学方法应用一般要借助于电子计算机,所以,地质统计学软件的研制开发基本上与地质统计学方法及应用是同步的。目前,具有较好应用价值的地质统计学软件是集数据库,地质统计学的各类方法计算和CAD图形处理技术于一体的多功能交互式大型组合软件。具有代表性的软件系统有以下几种。

2.1 ISATIS 地质统计学程序系统

该软件由法国巴黎高等矿院地质统计学研究中心(centre de Geostatistiques, EN-SMP)研制的,软件的运行环境为工作站 UNIX 操作系统,并由以下几部分构成:(1)数据库及数据管理(主要使用的是 GTX 数据库);(2)数据变异性分析及各类统计(包括各种变异函数及协方差函数的计算及拟合);(3)各类地质统计学方法计算,这是该软件系统的最关键部分,它几乎包含了当今大多数克立格方法;(4)非条件模拟及条件模拟方法;(5)其他快速插值的方法;(6)图形显示。

2.2 HERESIM-3D 地质统计学软件

这是法国巴黎高等矿院地质统计学中心与法国石油局共同研制的软件系统。该软件包含的内容与上述 ISATIS 软件有相似之处,但它对油气田地质及开采方面的问题考虑更周到。该软件包含以下几方面:(1)应用岩石学理论细致的研究地层及岩性,处理沉积顺序、旋回,空间分布及沉积环境等问题;(2)油气藏的随机模拟;(3)表述油气藏的各种物理参数(如孔隙度等)的研究;(4)描述油藏。

2.3 加拿大 GEOSTAT 国际地质统计学软件系统

该系统开发公司推出软件系统以普通克立格法为主,也有其他克立格法。而加拿大矿产和能源中心开发的 CAMET 地质统计学程序系统在地质勘查、储量计算及环境污染的研究方面更具特色。加拿大 IGC 公司(International Geosystems corporation)研制的 GLS 系统(Geology System)是一个标准地质编录数据库系统。在 GLS 基础上研制的 RCS 系统(Reserve Calculated System)应用了普通克立格法、泛克立格法等。

此外,美国的 Newmont(丹佛)公司、澳大利亚的 KRJA Systems 公司研制开发的软件系统中均利用了地质统计学的有关理论及方法。由美国斯坦福大学应用地球科学系 C. V. Deutsch 和 A. G. Journel 共同编写的 GSLIB(Geostatistical Software Library and User's Guide)一书(附软盘)在普及和推广地质统计学理论、方法及软件研制方面发挥了一定作用。

中国地质统计学家、数学地质学家和有关工程技术人员在研究地质统计学的理论及方法的同时,为了将其应用于生产实际,他们也研制出各种软件系统,具有代表性的有:

(1)CGES(Chinese Geology and Exploration System)中文地质勘探系统软件。这是冶金部黄金指挥部计算遥感中心在加拿大 IGC 公司 GLS 和 RCS 软件系统基础上全面汉化改造并扩充功能而完成的。该系统由中文地勘编录子系统(CGLS)、地质数据库子系统(GDBS)、储量计算子系统(CRCS)和中文地勘成图子系统(CGGS)4大部分组成,利用该系统可完成自标准化地勘编录,建库、克立格法储量计算以及各类地勘报告附图的生成和绘制。为提高水平,目前已开始对该软件系统的 Windows 新版软件的研制。

(2)KPX 2.0 固体矿产勘查评价自动化系统。这是地矿部信息研究院在 KPX 1.0(1991)版本上完成的。KPX 1.0 版本是以 PC 机为硬件环境,实现了计算机辅助勘查评

价的某些基本功能,KPX 2.0 版本除增强了成图系统之外,还增加了一些地质统计学方法(主要是普通克里格法)。该软件包括 4 个子系统,即地质编录及数据管理子系统,传统储量计算方法子系统,地质统计学储量计算子系统和图形处理子系统。

(3)地质统计学方法研究程序集,这是北京科技大学地质系专门为地质统计学理论、方法研究以及供地勘单位使用的软件,该程序集包括:①2D—GEOST 程序系统,该系统可以完成二维条件下的普通克里格法、对数正态克里格法、泛克里格法计算;②2D—INKRG 程序系统,该系统主要完成指示克里格法计算;③2D—MALGEOSTA 程序系统,这是一套专门从事二维条件下多元地质统计学计算的程序集,主要包括协同克里格法和因子克里格法,以及时间—空间域的多元信息的地质统计学计算;④3D—COKR 程序系统,该系统主要完成三维条件下的协同克里格计算。该程序集包括的地质统计学方法较多,但侧重于方法研究。

我国其他单位在地质统计学程序研制方面也做了大量有益的工作,如有北京有色设计院、长沙有色设计院、核工业部地质局、中国地质大学、西安石油学院、有色总公司北京地质所、南昌有色设计院和中国矿业大学等。

3 地质统计学方法的实际应用

按照地质统计学的定义及其研究内容,地质统计学的应用范围很广,而且根据研究对象的不同特点采用不同的地质统计学方法,从而提高了精度。地质统计学应用的领域有以下几方面:(1)地质勘查与矿业工程,包括地质储量和可回采储量计算;品位优化及经济分析;勘探方法研究;采矿设计中有关参数的优化;矿山地质及管理。(2)石油及天然气勘探开发,包括油藏描述中的有关问题;油气田地质特征(储油构造特征及油气内部物理性质)研究;油气藏储量计算等。(3)水文

地质研究,如含水层标志的空间变异性;地下水化学成分;地下水运动规律与成分变化的动态模拟等。(4)工程地质研究,如土壤力学性质在时间、空间上的变化以及有关参数的最优估计;工程地质研究中某些工程可靠性及风险性评价等。(5)环境科学研究,如污染事件的有效预报;污染物质在时间—空间域中的分布等。此外,地质统计学还成功的应用于渔业资源研究、海洋地球物理、农林科学、图象处理、制图技术、公共卫生、财政分析及材料科学等方面。

目前,地质统计学在我国主要是应用于矿产资源的评估及储量计算,物化探数据处理及地质勘探等领域。随着我国经济体制的深化改革,地质统计学的应用将会更为广泛和有效应用。为了推进地质统计学在矿产储量计算中的实际应用,原全国矿产储量委员会先后数次向全国发文。1995 年还专发了关于“运用地质统计学方法提交地质勘探报告的编写提纲和审查提纲”的试行意见,从而促进了地质统计学的实际应用。经国家储委审批的驾劳金矿、刁泉银矿、玉龙铜矿等矿床的储量报告或以地质统计学计算为主,或附以应用地质统计学计算结果。

4 几点看法

地质数据处理(包括矿产储量计算)的精度及数据空间分布特征的研究是提高地质研究水平的重要内容,也是地质工作(或广泛的矿业领域)实现自动化的重要标志,地质统计学在这方面是大有作为的,为了提高地质统计学在矿业(乃至国民经济)中发挥更大作用,取得更大的经济效益,我国的地质统计学研究应着重以下几方面:

(1)加强地质统计学理论及方法的深入研究。当前除不断提高平稳线性地质统计学研究水平外,重点应对非参数地质统计学、多元地质统计学、条件模拟及非线性地质统计学的研究,提高我国地质统计学的理论水平。

(2)为了适应当前经济发展的需要,充分发挥地质统计学在找矿勘探与矿山开采之间、勘探工作与经济评价之间、国内与国外地质工作之间的桥梁及纽带作用,提高矿业工作的水平,必须大力宣传与推广地质统计学,而不断提高各种地质统计学方法在各个领域的应用水平是关键。也就是说,应用地质统计学方法研究有关问题要见成效,有效益。

(3)加强各种地质统计学方法计算程序系统的研制及普及,提高各类程序系统的水平是关键。

(4)开拓地质统计学在我国的应用领域,使地质统计学在国民经济建设中发挥它应该发挥的作用。时刻跟踪地质统计学国际先进水平,为我国的国民经济建设全方位服务。

(5)积极培养高水平的地质统计学人才。

参考文献

- 1 侯景儒,郭光裕.矿产统计预测与地质统计学的理论与应用.北京:冶金工业出版社,1993.
- 2 侯景儒.地质统计学在我国的应用及其发展.地质与勘探,1991,27(2).
- 3 赵鹏大,孟宪国等.从第26届国际地质大会看数学地质的现状与发展.中国数学地质,1994(5).
- 4 刘承祚.数学地质在1992—1993年期间的主要进展.中国数学地质,1994(5).
- 5 侯景儒,黄竟先等.非参数及多元地质统计学的理论分析及其应用.北京:冶金工业出版社,1994.
- 6 侯景儒.实用地质统计学(理论·方法·应用).北京:北京科技大学教材,1992.
- 7 王世祚,余先川.数学地质发展的趋势与展望.中国数学地质,1995(6).
- 8 侯景儒,郎宝平等.空间域及时一空域多元信息的地质统计学研究及应用.中国数学地质,1995(6).
- 9 国家自然科学基金委员会.地质科学.北京:科学出版社,1991.
- 10 侯景儒.地质统计学发展现状及对若干问题的讨论.黄金地质,1996,2(1).
- 11 C. Deutsch, A. G. Journel: GSLIB. Geostatistical software Library and User's Guide. Oxford University Press, 1992.
- 12 J. Rivoirard. Introduction to Disjunctive Kriging and Non-linear Geostatistics. Clarendon Press, Oxford, 1994.
- 13 E. H. Isaaks, R. M. Srivastava. Applied Geostatistics. Oxford University Press 1989.
- 14 Hans Wackernagel. Multivariate Geostatistics, an introduction with application. Springer, 1995.
- 15 Centre de Géostatistique. ISATISL Table of Contents. Fontainebleau, France, 1994.
- 16 Centre de Géostatistique. Introducing HERESIM 3D 3.0 user Guide, Fontainebleau, France, 1995.

REVIEW AND PROSPECT OF GEOSTATISTICS (SPATIAL—INFORMATION STATISTICS) DEVELOPMENT IN CHINA

Hou Jingru

Abstract The paper discussed some problems about geostatistical (Spatial-information statistics) development which involve four parts: 1. Present situation of theory and methods of geostatistical (patial-information statistical) studing; 2. About software of geostatistics; 3. Application of geostatistics; 4. Some veiw of geostatistical (patial-information statistical) development in future.

Key wards geostatistics, Spatial-information statistics, kriging, mathematical geology, variogram.