

■ 海洋水文环境要素 分析方法

左军成 杜 凌 陈美香 徐 青 李 娟 编著

海洋水文环境要素
分析方法



科学出版社

海洋水文环境要素分析方法

左军成 杜 凌 陈美香 徐 青 李 娟 编著

海洋科学浙江省优势专业
海洋科学浙江省一流学科

联合资助

科学出版社

北 京

内 容 简 介

本书分为四章,即资料的预处理和常规分析方法、温盐资料分析与水团分析方法、潮汐潮流分析方法和海流资料分析方法。与以往的同类教材相比,本书在系统介绍海洋水文环境要素分析理论的基础上,介绍了常用的资料预处理方法,对流、潮等海洋要素给出了实际数据分析的步骤,弥补了同类教材中相关内容的缺乏。

本书可作为海洋科学类本科生及相近专业学生的基础课教材,亦可作为相近专业的教学参考用书,对从事相近专业的科技人员有较大参考价值。

图书在版编目(CIP)数据

海洋水文环境要素分析方法 / 左军成等编著. —北京: 科学出版社, 2018. 6

ISBN 978-7-03-057964-5

I. ①海… II. ①左… III. ①海洋水文-要素分析-分析方法
IV. ①P731

中国版本图书馆 CIP 数据核字 (2018) 第 131551 号

责任编辑: 黄 海 / 责任校对: 张凤琴

责任印制: 张克忠 / 封面设计: 许 瑞

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

三河市荣展印务有限公司 印刷

科学出版社发行 各地新华书店经销

*

2018 年 6 月第 一 版 开本: B5 (720×1000)

2018 年 6 月第一次印刷 印张: 16 3/4

字数: 350 000

定价: 69.00 元

(如有印装质量问题, 我社负责调换)

目 录

第一章 资料的预处理和常规分析方法	1
1.1 采样与预处理	1
1.1.1 采样的基本要求	1
1.1.2 预处理	2
1.2 数据表示	13
1.2.1 空间变化	13
1.2.2 时间序列	17
1.2.3 海洋要素特征图	20
1.3 时间序列的常规处理方法	21
1.3.1 功率谱分析	21
1.3.2 小波分析	27
1.3.3 随机动态分析方法	31
第二章 温盐资料分析与水团分析方法	37
2.1 温盐分析的意义	37
2.1.1 温盐分析概况	37
2.1.2 跃层和锋面	40
2.2 定点时间序列分析方法	48
2.2.1 多元统计分析	48
2.2.2 经验模态分解法	56
2.3 温盐场资料的分析方法	65
2.3.1 主成分定义及性质	66
2.3.2 经验正交函数分解法	69
2.4 水团分析方法	82
2.4.1 温-盐图解	82
2.4.2 判别分析	90
第三章 潮汐潮流分析方法	104
3.1 潮汐潮流现象	104
3.1.1 潮汐现象及月中天	104
3.1.2 潮汐的类型	107

3.1.3	潮汐不等现象	109
3.1.4	潮流现象	113
3.2	分潮的导出及分类	115
3.2.1	分潮	115
3.2.2	分潮的分群	127
3.2.3	浅水分潮	127
3.2.4	辐射潮	129
3.3	潮汐调和分析方法	130
3.3.1	潮汐调和分析原理	130
3.3.2	潮汐调和分析的最小二乘法	133
3.3.3	分潮的选取	139
3.3.4	交点因子和交点订正角的计算	143
3.3.5	潮流的调和常数和椭圆要素	151
3.3.6	调和分析步骤	155
3.4	潮汐特征值与工程水位	156
3.4.1	潮汐特征值	157
3.4.2	工程水位	159
3.4.3	陆地高程基准和海图深度基准面	167
第四章	海流资料分析方法	177
4.1	海流的构成	177
4.2	海流的计算	178
4.2.1	Ekman 风漂流的计算	178
4.2.2	惯性流	190
4.2.3	地转流计算	193
4.3	流场动力热力学特征分析	204
4.3.1	流函数	205
4.3.2	涡度	208
4.3.3	热通量和盐通量	209
4.4	海流的观测分析实例	212
4.4.1	单点观测资料的分析	212
4.4.2	海流剖面观测分析	221
4.4.3	海流断面观测分析	223
	参考文献	226
	附表 1 基本天文数据	229

附表 2	主要天文变量周期·····	230
附表 3	引潮力的 Doodson 展开式·····	231
附表 4	分析年观测资料时选取的主要分潮·····	249
附表 5	分析中期观测资料时选取的分潮·····	253
附表 6	计算校核水位时所用龚贝尔 I 型极值分布所用的 λ_{pn} 值·····	258

第一章 资料的预处理和常规分析方法

海洋科学研究对象的特殊性和复杂性决定了海洋数据分析的目的就是最大限度地利用已有的直接观测数据,依据合理有效的分析方法,透过复杂的物理海洋现象,提炼其本质特征和规律,为开发利用海洋提供必要的科学技术支持。本章旨在介绍目前常用的海洋数据处理和表示方法。

1.1 采样与预处理

1.1.1 采样的基本要求

1. 采样间隔

所选仪器的采样间隔必须足够小以能满足提取研究对象的空间和时间尺度上变化的要求。如,每小时进行一次水位资料的采样,就不足以用来研究海浪的性质。

对于一个给定的采样间隔 Δt , 我们能分辨出的最高频率是 Nyquist (或者折叠) 频率 f_N

$$f_N = \frac{1}{2\Delta t} \quad (1.1)$$

譬如,如果采样间隔为 10 小时,那么可分辨的最高频率是 $f_N=0.05\text{cph}$ (cycles per hour, 每小时周)。换言之,式 (1.1) 表示完成一个周期为 $1/f_N$ 的振动至少需要 $2\Delta t$ 的时间 (即获取 3 个采样点,图 1.1)。

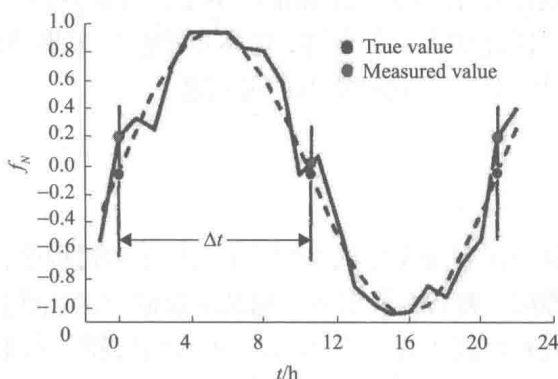


图 1.1 采样图示

上式的一个重要结果是混淆问题。依照该式选取采样间隔而使得 $f < f_N$ 的振动无法分辨出来, $f < f_N$ 频率上(高频部分)的能量将折叠回 $f < f_N$ 的频段上(低频部分)。高频部分的能量并没有消失而是重新分布在我们所研究问题的频段内, 从而使得 $f < f_N$ 内的能谱失真。

Nyquist 频率的概念同样适用于时间和空间。Nyquist 波数是确定采样的基本波长的有效方法。

2. 采样长度

第二个问题是样本要足够长以建立研究过程的重要统计特征。例如, 时间序列观测过程, 要求足够长的采样过程中所研究的现象可以重复几个周期; 空间采样则要求采样空间足够大, 以分辨研究过程中的多个循环。

考虑时间间隔 Δt 的某样本区间, 资料越长则我们就可以越能分辨不同频率的分量。对于 N 个采样点 $T = N\Delta t$, 则时间序列中提取的最低频率是

$$f_0 = \frac{1}{N\Delta t} = \frac{1}{T} \quad (1.2)$$

能分离两个相邻频率 f_1 、 f_2 之差的最小值是

$$\Delta f = |f_2 - f_1| = \frac{1}{N\Delta t} \quad (1.3)$$

理论上, 能够分辨出频率在 $f_0 \leq f \leq f_N$ 之间的所有频率分量。但实际上能分辨出的相邻谱峰的频率是由 Rayleigh 准则确定的, 即两个相邻频率能被分离开的基本条件是其频率差大于式 (1.3) 中的 Δf 。

3. 采样精度

根据前面关于采样间隔和采样长度的讨论, 需要资料足够长且时间间隔又足够短。同时, 采样还要具有足够高的精度, 以使关心的变量的采样精度比背景噪声要大。如果仪器传感器的响应速度和精度不足以分辨观测参量的变化, 那么就很难满足测量要求。因此如果条件允许, 应尽量快速地采集数据, 并可通过平均、平滑和其他处理方法, 以提高观测数据的可信度。

1.1.2 预处理

1. 数据插值

人们在使用各种海洋观测仪器记录海洋水文要素的过程中, 由于技术手段的限制或是仪器故障等原因, 常常会使实际观测到的数据, 在空间上是不均匀的(如沿岸验潮站资料), 而在时间上也常常不均匀[如观测仪器工作不正常导致某段时间数据不可用(gap)或出现奇异值等]。等时间间隔观测到的数据常常会存在奇异值或缺测值, 因此在实际的资料处理过程中就存在如何对时、空非均匀的数据进行处

理的问题。目前,通常的做法是利用其附近(时空域)的观测数据进行适当的插值。

选择什么样的插值方案最合理,要看具体的海洋要素是什么,应该根据要素的特定变化规律来选择能揭示其变化规律的函数进行插值。选择的插补方案应包括:①确定插值的参量(原始资料序列、距平等);②选取插值函数(线性插值、多项式插值、三次样条插值等);③选择合理的插值的判据(精确匹配原始数据、最小二乘估计等);④确定插值后数据。以下介绍几种常用的插值方法。

(1) 线性插值

如果要素的变化满足线性变化规律,则可以采取线性插值。对数据序列 $y(x)$ 的线性插值可写成

$$y(x) = y(a) + \frac{x-a}{b-a}(y(b) - y(a)) \quad (1.4)$$

式中: $x=a$ 和 $x=b$ 是插值所采用的数据的起始时刻(位置)和终止时刻(位置)。

式(1.4)也可作为数据的外推(extrapolation)公式。式(1.4)可以看作是多项式插值的一种特例。

(2) 多项式插值

如果在多个点之间进行插值,就要将简单的线性插值推广为多项式插值。如果某要素 y 与 $i=1,2,\dots,N+1$ 个时间(空间)点存在 N 阶拉格朗日(Lagrange)多项式关系(图1.2)

$$y(x) = a_0 + a_1x + a_2x^2 + \dots + a_Nx^N = \sum_{k=0}^N a_k x^k = y_1 \frac{(x-x_2)(x-x_3)\dots(x-x_{N+1})}{(x_1-x_2)(x_1-x_3)\dots(x_1-x_{N+1})} \quad (1.5)$$

$$+ y_2 \frac{(x-x_1)(x-x_3)\dots(x-x_{N+1})}{(x_2-x_1)(x_2-x_3)\dots(x_2-x_{N+1})} + \dots + y_{N+1} \frac{(x-x_1)(x-x_2)\dots(x-x_N)}{(x_{N+1}-x_2)(x_{N+1}-x_3)\dots(x_{N+1}-x_N)}$$

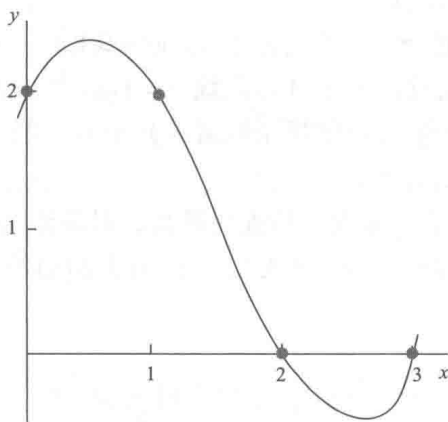


图 1.2 Lagrange 多项式插值

较为常用的三点拉格朗日插值法可以较好地拟合实际曲线,但它的插值曲线

光滑性较差。在资料出现跃层时,拟合曲线也会出现一定程度的摆动,与三次样条插值法相比较,三点拉格朗日插值法的波动幅度稍大一点,不过摆动仅出现在跃层的前拐点处。跃层的强弱、结点步长的大小对其插值的影响均与三次样条插值法相同。由于方法简便,程序短小,目前世界上许多海洋机构仍继续使用这一方法(侍茂崇等,2000)。

(3) 样条插值

样条插值在实际工作中得到广泛应用。样条函数可以更为有效地拟合空间分布非解析的原始资料,并且具有收敛性好、高阶微分近似、截断误差稳定性好的优点。最常用的是三次样条插值函数(图1.3)。

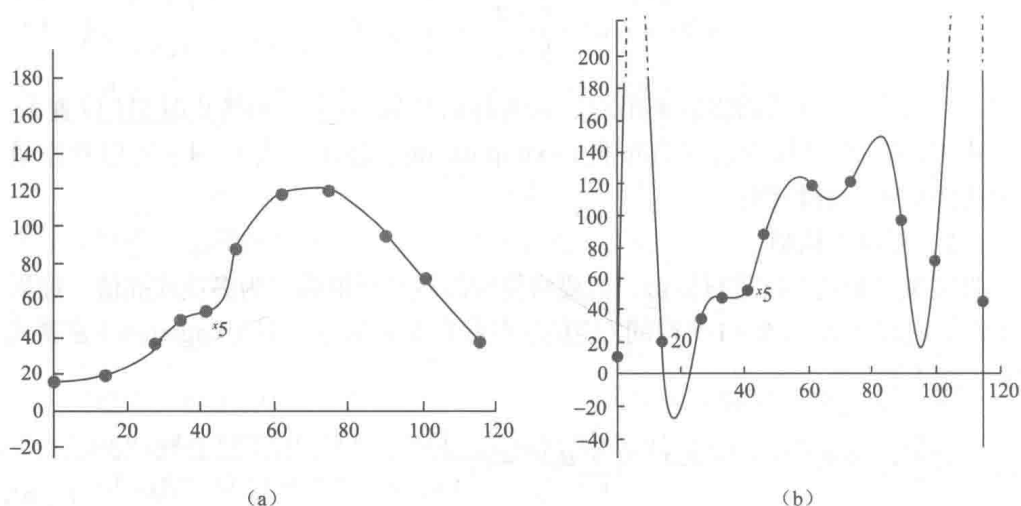


图 1.3 利用 11 点进行三次样条插值 (a) 和多项式插值 (b) 得到的拟合曲线

1) 三次样条插值表达式

设平面上给定 $n+1$ 个点 $(x_i, y_i) (i=0, 1, 2, \dots, n)$ 而且 $a=x_0 < x_1 < x_2 < \dots < x_n=b$, 假设通过这个 $n+1$ 点曲线 $y=f(x)$ 在 (a, b) 区间上有二阶连续导数存在, 在区间 (a, b) 上做 $f(x)$ 的样条插值函数 $S(x)$, 使它满足如下条件:

- i. $S(x_i) = y_i (i=0, 1, 2, \dots, n)$;
- ii. 在区间 (a, b) 上存在一阶及二阶连续导数, 以保证连接处曲线是光滑的;
- iii. 在每一个子区间 $(x_i, x_{i+1}) (i=0, 1, 2, \dots, n-1)$ 上 $S(x)$ 都是三次多项式, 其常用的表达式为

$$S(x) = \left[\frac{3}{h_i^2} (x_{i+1} - x)^2 - \frac{2}{h_i^3} (x_{i+1} - x)^3 \right] y_i + \left[\frac{3}{h_i^2} (x - x_i)^2 - \frac{2}{h_i^3} (x - x_i)^3 \right] y_{i+1} + h_i \left[\frac{1}{h_i^2} (x_{i+1} - x)^2 - \frac{1}{h_i^3} (x_{i+1} - x)^3 \right] m_i - h_i \left[\frac{1}{h_i^2} (x - x_i)^2 - \frac{1}{h_i^3} (x - x_i)^3 \right] m_{i+1} \quad (1.6)$$

式中, $h_i = x_{i+1} - x_i$, $m_i = S'(x_i)$ ($x = x_i$ 处的一阶导数)。

2) 边界条件的确定

根据我们手工绘制拟合曲线的经验, 将所有的点连接成曲线后还要作适当的延长, 而且延长是顺势的, 延长部分近似于直线, 所以在进行插值计算前, 应该在曲线的前端和末端各增加一个点。具体的方法是在原有的样点基础上, 在两端先作线性插值, 插值的位置是在两端延长一个数量级时间单位, 当然, 超出坐标范围的点应该剔除。因此我们可以假设曲线中的 x_0 及 x_n 处的斜率是已知的, 即

$$S'(x_0) = m_0 = \frac{y_1 - y_0}{x_1 - x_0} \quad (1.7)$$

$$S'(x_n) = m_n = \frac{y_n - y_{n-1}}{x_n - x_{n-1}} \quad (1.8)$$

式 (1.7) 和式 (1.8) 就是利用拟合试验数据建立样条插值时的边界条件。边界条件确定得恰当与否, 可以从曲线的拟合效果进行判断。

3) 样条插值的实现

i. 样条函数的建立

建立三次样条插值函数有多种办法, 下面是参照建立 Lagrange 插值公式方法的推导过程建立的, 从表达式 (1.6) 中可以看出, 函数建立的过程实质上是求解 m_i 的过程, 具体可分为以下三个步骤:

第一步: 由 $a_i = \frac{\alpha_i}{2}$, $b_i = \frac{\beta_i}{2}$ 出发, 按下式计算 α_i 、 β_i

$$\begin{cases} \alpha_i = \frac{h_{i-1}}{h_{i-1} + h_i} \\ \beta_i = 3\left[\frac{1 - \alpha_i}{h_{i-1}}(y_i - y_{i-1}) + \frac{\alpha_i}{h_i}(y_{i+1} - y_i)\right] \end{cases}, (i = 1, 2, 3, \dots, n-1) \quad (1.9)$$

式中: α_i 、 β_i 是计算的过程变量, 无实际意义, 以下 a_i 、 b_i 同; y_i 是第 i 个试验数据的纵坐标值; $h_i = x_{i+1} - x_i$, x_i 是第 i 点的横坐标值。

第二步: 利用下式计算 a_i 、 b_i

$$\begin{cases} a_i = \frac{\alpha_i}{2 - (1 - \alpha_i)a_{i-1}} \\ b_i = \frac{\beta_i - (1 - \alpha_i)b_{i-1}}{2 - (1 - \alpha_i)a_{i-1}} \end{cases}, (i = 2, 3, \dots, n) \quad (1.10)$$

第三步：按下式计算 m_i

$$m_i = b_i - a_i m_{i+1} \quad (i = n-1, n-2, \dots, 1) \quad (1.11)$$

将求得的 m_i 代入式 (1.6)，即可得出各个子区间的样条插值函数 $S(x)$ 。

ii. 计算步长的确定

由于三次样条函数是分段函数，即在不同的自变量区间有不同的表达式，所以在实际应用中要注意判断自变量所在的子区间。

使用三次样条插值函数求解曲线上非样点的坐标值时，插值的步长通过试算法确定，即确定一个步长后，看其绘制曲线的效果，如光滑度不够，则逐步减小步长的值，直到曲线光滑度达到要求为止。

2. 空间场资料的客观分析法

空间观测数据的精确估计和反演对海洋学的研究与发展具有重要意义。卫星观测数据作为一种空间数据，具有一般空间数据所具有的特点，蕴含着复杂的非线性动力学机制，在时空分布上具有纷杂多变的时空特征。因而要把这种离散的、不规则的观测资料转换成规则的网格点资料，可以作为数值模拟的初始场或是验证场以及诊断分析的背景场。这种通过某种算法将不规则的原始数据换算到规则的经纬网格点上的方法称之为客观分析方法。较为常用的有 Kriging 插值、反距离加权、Delaunay 三角剖分线性插值、双谐样条 (biharmonic spline) 插值和 Cressman 客观分析等空间内插法。

空间内插对于观测台站稀少或台站分布不合理的地区具有重要的实际意义，并使得利用网格数据进行客观分析成为可能。空间内插法根据研究目的可以有多种分类方法：根据其基本假设可分为几何方法、统计方法、函数方法和物理模型方法等；根据数学本质又可分为局部函数法和逐点内插法。

以下介绍几种常用的空间内插法。

(1) Kriging 插值

Kriging 法被认为是地学统计中最主要的方法之一，是以南非科学家 D. G. Krige 的名字命名的一种空间内插法。从统计意义上讲，Kriging 法是从变量相关性和变异性出发，在有限区域内对空间变量进行无偏、最优估计的一种统计方法；从插值角度上讲，Kriging 法以空间结构分析为基础，充分利用数据空间场的性质，在插值过程中对空间数据求线性最优，可以反映空间场的各向异性。该法的最佳适用条件是空间变量存在着空间相关性。

使用 Kriging 法的主要步骤是：①对空间场进行结构分析。在充分了解空间数据场性质的前提下，建立空间变量的协方差函数，提出变异函数模型（常用的有球面、指数、高斯、阻尼正弦、线性等模型）；②在该模型的基础上进行 Kriging

计算, 求得距离相关量。

令 $V(\hat{X})$ 为点 $\hat{X} = (x, y, t)$ 的观测值, 其中 (x, y) 为空间坐标, t 为时间。假设每一个测量值由真实值和一个随机量 (该样本点的噪声、地球物理和仪器误差) 组成:

$$V(\hat{X}) = \langle U \rangle(\hat{X}) + \varepsilon(\hat{X}) \quad (1.12)$$

为了估计 $\hat{X}_0$ 点上 $\langle U \rangle(\hat{X}_0)$ 的值, 假设 N 个点 $\hat{X}_1, \hat{X}_2, \dots, \hat{X}_N$ 上的观测值 $V(\hat{X}_1), V(\hat{X}_2), \dots, V(\hat{X}_N)$ 组成一个线性公式:

$$\hat{U}(\hat{X}_0) = \sum_{i=1}^N \lambda_i \cdot V(\hat{X}_i) \quad (1.13)$$

其中 $\hat{U}$ 表示为 $\langle U \rangle$ 的无偏估计。

为了估计系数 λ_i , 必须知道以上标量准确的统计描述。假设场 $\langle U \rangle$ 代表一类具有相似统计特性的实体, 其平均值、协方差分别为:

$$E(\langle U \rangle(\hat{X})) = m \quad (1.14)$$

$$E((\langle U \rangle(\hat{X} + H) - m)(\langle U \rangle(\hat{X}) - m)) = C(H) \quad (1.15)$$

其中 $E()$ 和 $C()$ 代表数学期望和协方差。式 (1.14) 表示标量 $\langle U \rangle(\hat{X})$ 的期望独立于点 $\hat{X}$, 式 (1.15) 的含义是变量 $\langle U \rangle(\hat{X})$ 和 $\langle U \rangle(\hat{X} + H)$ 之间的协方差函数不依赖于观测点 $\hat{X}$ 和 $\hat{X} + H$, 但是依赖于它们之间的“距离” H 。同时认为误差 $\varepsilon(x_i, y_i, t_i)$ 与场 $\langle U \rangle$ 不相关, 相互之间也不相关。

$$E(\varepsilon(\hat{X}) \cdot \varepsilon(\hat{Y})) = \delta_{XY} \cdot \sigma^2 \quad (1.16)$$

$$E(\varepsilon(\hat{X}) \cdot \langle U \rangle(\hat{Y})) = 0 \quad (1.17)$$

式中: δ_{XY} 为 Kronecker 公式。

根据此假设, Gauss-Markov 理论认为 $\langle U \rangle(X)$ 是最佳的无偏线性估计, 也就是此估计具有最小的 $E((\langle U \rangle - \hat{U})^2)$ 。

由于式 (1.15) 和 (1.16) 给出的假设不简明直接, 考虑应用标量增长的固有假设, 表示如下:

$$E(\langle U \rangle (\hat{X} + H) - \langle U \rangle (\hat{X})) = 0 \quad (1.18)$$

$$E((\langle U \rangle (\hat{X} + H) - \langle U \rangle (\hat{X}))^2) = G(H) \quad (1.19)$$

式中: X 、 H 固定, H 代表空间和时间间隔, 结构函数 $G(H)$ 对于每一个确定的 H 是恒定的。式 (1.18) 指出不同间隔 H 的 $\langle U \rangle$ 平均独立于空间和时间间隔。式 (1.19) 指出不同间隔 H 的 $\langle U \rangle$ 的方差只是 H 的函数。 $G(H)$ 与协方差函数 C 有关:

$$G(H) = E(((\langle U \rangle (\hat{X} + H) - m) - (\langle U \rangle (\hat{X}) - m))^2) = 2 \cdot (C(0) - C(H)) \quad (1.20)$$

实际中令

$$\Gamma(H) = \frac{1}{2} \cdot G(H) = C(0) - C(H) \quad (1.21)$$

式中: Γ 为 H 的函数, 称为方差图。

线性估计量 $\hat{U}$ 的估计等于系数 λ_i 的估计。由公式

$$L = E((\hat{U} - \langle \hat{U} \rangle)^2) - 2\tau(1 - \sum_{i=1}^N \lambda_i) \quad (1.22)$$

取最小值来确定权重 λ_i 的大小。其中 τ 是 Lagrange 乘数, 用于限制

$$\sum_{i=1}^N \lambda_i = 1 \quad (1.23)$$

这是无偏估计量 (1.18) 的一个结果。通过式 (1.14) 和 (1.18), 式 (1.22) 中的函数 L 可表示为

$$L = \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j C(i, j) - 2C(0, i) + C(0, 0) + 2\tau(\sum_{i=1}^N \lambda_i^2 - 1) + \sum_{j=1}^N \lambda_j^2 \sigma^2 \quad (1.24)$$

$$C(i, j) = E(\langle U \rangle (\hat{X}_i) \cdot \langle U \rangle (\hat{X}_j))$$

其中, 令函数 L 对 $\lambda_1, \lambda_2, \dots, \lambda_N$ 的一阶导数等于 0, 则这些系数 $\lambda_1, \lambda_2, \dots, \lambda_N$ 是 $N+1$ 个线性公式组的解

$$\begin{aligned} \sum_{i=1}^N \lambda_i \cdot C(i, j) - C(j, 0) - \tau + \lambda_j \cdot \sigma^2 &= 0, j=1, \dots, N \\ \sum_{i=1}^N \lambda_i &= 1 \end{aligned} \quad (1.25)$$

当把固有假设考虑进去, 得到下面线性系统

$$\begin{aligned} -\sum_{i=1}^N \lambda_i \cdot \Gamma(j, 0) - \tau + \lambda_j \cdot \sigma^2 &= 0 \\ \sum_{i=1}^N \lambda_i &= 1 \end{aligned} \quad (1.26)$$

上式描述的线性系统称为 Kriging 系统。式 (1.26) 是表征真实值和估计量之间变化的指标

$$E((\hat{U} - <U>)^2) = C(0, 0) + \tau - \sum_{i=1}^N \lambda_i \cdot C(i, 0) \quad (1.27)$$

这些变化值用来确定估计量的质量好坏, 表示估计量与数据的符合程度。Kriging 系统的分辨率表达了式 (1.21) 中方差图 Γ 的最为全面的信息。

为了求解 Kriging 公式, 必须获得尽可能多的方差图 Γ 的信息。根据变量的时空分布, Γ 可表示为

$$\Gamma(x, y) = \varepsilon + a \left(1 - \exp\left(-\frac{(x+cy)}{b}\right) \right) \quad (1.28)$$

式中: a 对应于当变量间不相关时的方差图值; b 是空间滞后系数; c 用来表示变量间的时间相关; ε 对应于估计散射计风矢量时产生的空间噪声。

Kriging 方法的优点在于它基于一些可验证的统计假设, 具有坚实的理论基础。因此, Kriging 方法产生的格点变量的估计量是最佳的, 所有的估计量都依赖于可获得的观测值, 这些估计量的平均误差最小。另外, Kriging 方法提供的方差误差分析的表达式可以表示每一个格点上变量的估计精度, 而且也不会产生回归分析的边界效应。缺点是复杂, 计算量大, 另外变异函数需要根据经验人为选定。

(2) 其他空间插值法

1) 反距离加权法

反距离加权法是最常用的空间插值方法之一, 幂次参数控制着权重系数如何随着离开一个网格点距离的增加而下降。对于较大的方次, 较近的数据点赋以一

个较高的权重份额；对于较小的方次，权重比较均匀地分配给各数据点。可以用下式来表示：

$$Z = \frac{\sum_{i=1}^n \frac{1}{(D_i)^P} Z_i}{\sum_{i=1}^n \frac{1}{(D_i)^P}} \quad (1.29)$$

式中：\$Z\$ 是估计值；\$Z_i\$ 是第 \$i\$ 个样本；\$D_i\$ 是距离；\$P\$ 是幂次，它显著影响内插的结果，其选择标准是使平均绝对误差最小。当采样点与网格点重合时，该网格点被赋予和观测点一致的值，因此这是一个准确插值。它的特征之一是会在网格区域内产生围绕观测点位置的“靶心”。

2) Delaunay 三角化线性内插法

这种方法是通过在数据点之间连线以建立起一个三角形网来进行工作。原始数据点的 Delaunay 三角剖分插值的方法是：所有三角形的边都不能与另外的三角形相交，并且任何一个由这种三角形的 3 个顶点所确定的圆不能包含其他任何一个三角形的顶点。每一个三角形定义了一个覆盖该三角形内网格点的面，三角形的倾斜和标高由定义这个三角形的 3 个原始数据点确定，给定三角形内的全部网格点都要受到该三角形表面的限制。Delaunay 三角化线性内插就是在这种三角剖分的基础上进一步利用线性插值方法来进行插值计算。

3) Cressman 客观分析方法

Cressman 客观分析方法采用的是逐步订正的方法，已被广泛应用于各种诊断分析和数值模拟研究中。逐步订正法最主要的根据是 Cressman 客观分析函数，这种方法是 Cressman 在 1959 年提出的。先给定第一猜测场，然后用实际观测场逐步修正第一猜测场，直到订正后的场逼近观测记录为止。

$$\alpha' = \alpha_0 + \Delta\alpha_{ij} \quad (1.30)$$

其中

$$\Delta\alpha_{ij} = \frac{\sum_{k=1}^K (W_{ijk}^2 \Delta\alpha_k)}{\sum_{k=1}^K W_{ijk}} \quad (1.31)$$

式中：\$\alpha\$ 为任一气象要素；\$\alpha_0\$ 是变量 \$\alpha\$ 在格点 \$(i, j)\$ 上的第一猜测值；\$\alpha'\$ 是变量 \$\alpha\$ 在格点 \$(i, j)\$ 上的订正值；\$\Delta\alpha_k\$ 是观测点 \$k\$ 上的观测值与第一猜测值之差；\$W_{ijk}\$ 是权重因子，在 0.0~1.0 之间变化；\$K\$ 是影响半径 \$R\$ 内的资料站点数。Cressman 客

观分析方法中最重要的是权重函数 W_{ijk} 的确定, 它的一般形式为

$$W_{ijk} = \begin{cases} \frac{R^2 - d_{ijk}^2}{R^2 + d_{ijk}^2}, & (d_{ijk} \leq R) \\ 0, & (d_{ijk} > R) \end{cases} \quad (1.32)$$

其中, 影响半径 R 的选取具有一定的人为因素, 一般取常数。 R 选取的原则是由近及远进行扫描, 常用的几个影响半径是 1、2、4、7 和 10。 d_{ijk} 是格点 (i, j) 到观测点 k 的距离。

Cressman 客观分析方法的缺点主要是统计平滑功能差, 在采样点稀疏、观测资料贫乏的区域进行空间数据内插时, 会出现很多的“空值”斑点(冯锦明等, 2004)。

(3) 客观分析法

从数学的角度出发, 客观分析可以看成是一个插值问题, 但又不能视为完全纯粹的数学插值问题来处理, 原因为:

1) 插值不仅要求数据量足够大, 而且测站分布要较均匀, 一般的插值方法在数据稀少的条件下应用是不可能的;

2) 插值方法对数据的反应很敏感, 个别错误数据会造成较大范围的偏差;

3) 客观分析要以一定的物理依据为出发点。

常用的客观分析方法有如下几种。

1) 改进的多项式法: 利用最小二乘法求得插值多项式的系数, 这种方法要求数据量较多且均匀, 同时计算量也较大;

2) 多元最优插值法: 这是一种用统计方法在均方误差最小意义下的最优线性插值, 适用于三维多因素的客观分析;

3) 逐步订正法: 这是以观测数据间的统计特性为依据的一种方法, 计算量很小, 数据量和分布状态对方法本身没有什么影响, 可以适用于海洋数据的现状。

下面主要介绍一下逐步订正法。逐步订正法要求有一个预备场(或背景场), 在预备场的基础上逐步进行订正, 直到得出满意的结果, 比直接插值效果好。

物理依据: 一个站点的物理量变化值与其附近站点的物理量变化值相关较好, 距离越远, 二者相关越差, 到一定距离处相关为 0。

这提示我们, 如果一个站点上没有观测资料, 它的物理量变化值也应该由附近站点已知的变化值反映出来。由于二者的相关与距离有关, 所以没有观测资料站点上的物理量变化值应由观测的物理量变化值和与距离有关的权重来决定。

设 $A_{i,0}$ 为第 i 站的观测值, A_j 为任一网格点的第 j 次订正值, ΔA_j 为任一网格