

土壤水、盐、入渗变异特性及其相互关系的空间序列分析

李鸿杰 杜 历
(宁夏水利科学研究所, 750021)

摘 要

用序列分析的理论方法对土壤水、盐、入渗的变异特性及其相互关系进行分析,结果是令人满意的。最大熵谱分辨率高,揭露出优势周期;互谱分析揭露出序列之间的相互关系,多维AR(p)模型拟合效果甚佳。

关键词 序列分析,最大熵谱,自相关,相依性,独立性

为了较充分地研究土壤入渗变异特性,并研究土壤水分、盐分对土壤入渗的影响,作者做79组土壤入渗观测试验,同时在对对应点采样做土壤含水量及土壤含盐量测定分析。本文采用空间序列分析的理论和方法,对土壤水分、土壤盐分、土壤入渗的空间变异特性及土壤水分、土壤盐分对土壤入渗的影响进行分析研究,结果比较满意。作者认为,该试验研究对于盐渍土壤的灌溉排水具有重要意义。

一、试验及结果

试验在宁夏回族自治区平罗县东风试验场进行。试验地在7.95km²试区中央,地块

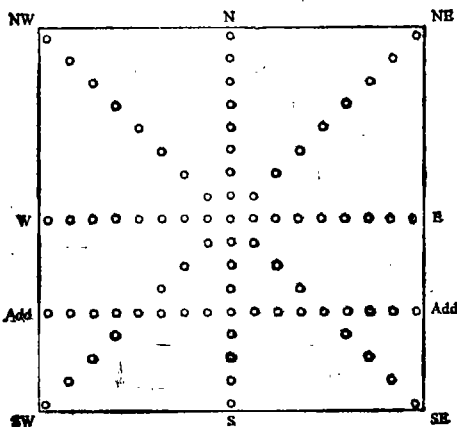


图1 试验布置图

Fig. 1 Arrangement of the experiment

50m × 50m。土壤质地为轻壤质土(该类土壤占试区耕地面积的72%)。地茬为春小麦套种玉米收获后10天左右(冬灌前,未耕翻),作物生长及地表看不出差别。土壤入渗试验采用双环入渗仪法。双环入渗仪用铁皮制作,内环直径0.80m,外环直径1.13m,环深0.30m,试验插深0.10m。试验布置成5条直线(纵、横点间距3m,斜线试点间距 $3\sqrt{2}$ m),共79组(见图1)。在试点正中采集土壤水分及土壤盐分样品,采集土柱长度5cm(0—5)。试验昼夜不停,连续进行,白天覆盖塑膜以减少蒸发。为使试验更接近生产实际,双环储水0.10m,开始试验,

表 1
Table 1

Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Mean
E-W S.	0.459	0.502	0.621	0.552	0.469	0.607	0.652	0.510	0.822	0.306	0.617	0.464	0.621	0.714	0.880	0.486	0.472	0.573
E-W A.	0.032	0.047	0.048	0.047	0.048	0.178	0.109	0.069	0.064	0.012	0.031	0.042	0.019	0.036	0.043	0.034	0.032	0.052
E-W o.	0.830	1.008	1.155	1.069	0.979	1.961	1.623	1.193	1.513	0.396	0.995	0.912	0.877	1.153	1.413	0.878	0.837	1.105
E-W e.	0.137	0.131	0.108	0.153	0.143	0.099	0.082	0.108	0.069	0.086	0.082	0.082	0.118	0.115	0.082	0.099	0.073	0.103
E-W o.	13.250	12.550	12.250	11.860	11.760	12.400	16.990	13.260	12.430	12.350	13.700	12.240	12.370	13.570	15.690	15.830	15.570	13.415
N-S s.	0.350	0.524	0.512	0.440	0.261	0.462	0.321	0.443	0.822	0.811	0.298	0.713	0.713	0.510	0.801	0.683	0.696	0.550
N-S A.	0.053	0.045	0.048	0.022	0.021	0.029	0.029	0.020	0.064	0.079	0.013	0.015	0.018	0.013	0.016	0.019	0.032	0.031
N-S a.	0.899	1.018	1.033	0.691	0.431	0.791	0.605	0.671	1.513	1.613	0.386	0.948	0.992	0.671	1.072	0.948	1.103	0.905
N-S e.	0.111	0.102	0.143	0.089	0.092	0.086	0.082	0.092	0.069	0.095	0.121	0.111	0.134	0.156	0.105	0.111	0.115	0.106
N-S o.	14.150	13.840	12.780	11.530	13.300	11.300	12.200	13.940	12.430	9.380	9.800	10.540	10.870	8.460	11.020	9.440	10.650	11.507
NE-SW S.	0.525	0.722	0.545	0.334	0.430	0.827	0.424	0.567	0.822	0.485	0.638	1.010	0.595	0.608	0.907	0.886	1.093	0.671
NE-SW A.	0.060	0.034	0.016	0.012	0.013	0.020	0.011	0.014	0.064	0.021	0.018	0.063	0.014	0.017	0.034	0.023	0.096	0.031
NE-SW a.	1.141	1.148	0.754	0.428	0.575	2.309	0.553	0.761	1.513	0.739	0.886	1.723	0.768	0.812	1.302	1.254	2.039	1.100
NE-SW e.	0.108	0.131	0.118	0.208	0.208	0.131	0.259	0.134	0.069	0.131	0.115	0.095	0.102	0.128	0.082	0.102	0.102	0.130
NE-SW o.	14.250	11.300	14.670	12.180	12.480	15.400	9.330	14.490	12.430	9.460	9.520	13.990	11.860	9.190	12.590	15.410	12.490	12.414
SE-NW S.	0.275	0.855	0.427	0.350	0.499	0.230	0.841	0.857	0.822	0.394	0.563	0.512	0.368	0.642	0.615	0.487	0.704	0.556
SE-NW A.	0.026	0.047	0.016	0.014	0.020	0.013	0.034	0.035	0.064	0.016	0.019	0.034	0.015	0.030	0.018	0.026	0.080	0.029
SE-NW a.	0.523	1.397	0.593	0.475	0.733	0.327	1.293	1.272	1.513	0.556	0.815	0.908	0.506	1.019	0.860	0.793	1.491	0.886
SE-NW e.	0.108	0.121	0.134	0.220	0.118	0.140	0.137	0.108	0.069	0.086	0.140	0.102	0.175	0.230	0.092	0.108	0.086	0.127
SE-NW o.	8.650	0.840	25.440	5.450	9.650	0.420	9.930	9.470	12.430	15.000	11.640	13.020	13.060	13.260	13.340	14.660	15.380	12.096
add S.	0.345	0.311	0.448	0.962	0.499	0.453	0.340	0.654	0.713	0.718	0.638	1.646	0.595	0.619	0.657	0.395	0.352	0.608
add A.	0.024	0.014	0.022	0.052	0.020	0.018	0.021	0.028	0.018	0.023	0.029	0.011	0.014	0.015	0.040	0.009	0.068	0.025
add a.	0.584	0.418	0.697	1.547	0.733	0.657	0.543	0.976	0.992	0.992	0.998	2.145	0.768	0.835	1.129	0.427	1.020	0.909
add e.	0.131	0.208	0.150	0.153	0.118	0.181	0.273	0.118	0.134	0.159	0.121	0.108	0.102	0.118	0.131	0.105	0.082	0.140
add o.	8.970	10.990	8.520	7.990	9.650	9.270	9.150	6.930	10.870	6.120	10.620	12.790	11.860	12.890	13.040	11.470	11.950	10.181

记时观测,直至结束。

土壤入渗试验资料根据 Philip 半解析公式整理:

$$I = S t^{\frac{1}{2}} + A t \quad (1)$$

式中 I 为土壤累积入渗量 (cm); S 为土壤吸渗率 ($\text{cm}/\text{min}^{\frac{1}{2}}$); A 为土壤稳定入渗率 (cm/min); t 为累计入渗时间 (min)。

令 $f(x) = [(S_i t_{ij}^{\frac{1}{2}} + A_i t_{ij}) - I_{ij}]$, 用 (2) 式进行迭代计算, 即可求得 S_i, A_i [1]。

$$\begin{cases} x_{k+1} = x_k + p_k(\mu) \\ (A_k^T A_k + \mu I) p_k = -A_k^T f(x_k) \end{cases} \quad (2)$$

经整理的试验结果见表 1。

二、试验结果的空序分析

本文将土壤水、盐、入渗经整理的试验结果看作空间随机序列的一次实现, 进行序列分析。

所谓序列分析, 就是用数学模型近似地描述动态数据, 并通过对模型的分析, 更本质地认识数据的内在结构和复杂性, 进而达到由表及里, 控制规律和实现预测的目的。

由于 ARMA(p, q) 模型形式简单, 便于分析数据的结构和内在性质, 并且和具有有理谱密度的平稳零均值序列存在着相互对应的关系, 是进行最佳预报和控制(最小方差意义下)的较常用的模型。又 AR(p) 模型是极大熵谱准则下的最优线性模型形式, 而且其参数可以通过 Burg 递推得到; 最大熵谱估计具有分辨率高, 适用于短序列的优点(能准确地挑选短序列的优势周期) [2]。本文主要用 AR(p) 模型的最大熵谱估计研究土壤水分、盐分、土壤入渗空间变异特性, 还将用互谱(交叉谱)及多维序列分析的理论方法, 研究土壤水分和土壤盐分对土壤入渗的影响。

(一) 检验及预处理

一般来说, 试验资料大都是非平稳序列, 可分解为:

$$Z = D + S + \varepsilon \quad (3)$$

上式中 D 为趋势函数; S 为平稳序列; ε 为白噪声。

建模之前, 应先判定序列是否具备平稳性, 还要检验序列的正态性。为判定所建模型是否合理, 还应检验模型噪声(残量)序列的独立性。

1. 平稳性(独立性)检验 对于试验数据序列 $x_i, i = 1, 2, \dots, N$ 可以给出它的均值, 方差和自相关函数的无偏或渐近无偏估计:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (4)$$

$$\sigma^2 = s^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (5)$$

$$r(\tau) = \frac{1}{N} \sum_{i=1}^{N-\tau} \left(\frac{x_i - \bar{x}}{s} \right) \left(\frac{x_{i+\tau} - \bar{x}}{s} \right) \quad \tau = 0, 1, \dots, m \quad (6)$$

由已知序列可以计算出自相关函数的一串值 r_1, r_2, \dots, r_m 。

检验 (1)

假设给定的序列是平稳的, 它的自相关函数的所有值都近似地等于 0。如果序列前 m 个自相关系数 $r_k, k = 1, 2, \dots, m$ 。有

$$|r_k| \leq \frac{1.96}{\sqrt{N}} \quad (7)$$

则 95% 的置信度认为 r_k 与零没有显著差异, 从而判定该序列是平稳的。

检验 (2)

由于 k 个独立 $N(0, 1)$ 随机变量的平方和服从自由度为 k 的 χ^2 分布, 对于预先给定的信度 α , 由 χ^2 分布表可查出 $\chi_{k\alpha}^2$, 其中:

$$P(\chi_k^2 \geq \chi_{k\alpha}^2) = \alpha$$

现假设 $\sqrt{N}r_k(N, \alpha), k = 1, 2, \dots, k$ 是 k 个独立 $N(0, 1)$ 分布随机变量, 则:

$$Q_k = N \sum_{k=1}^k r_k^2(N, \alpha) \quad (8)$$

为自由度等于 k 的中心 χ^2 分布, 那么假设成立时, 应当有 $Q_k \leq \chi_{k\alpha}^2$ 。若 $Q_k > \chi_{k\alpha}^2$, 就否定 $\sqrt{N}r_k(N, \alpha), k = 1, 2, \dots, k$ 是独立 $N(0, 1)$ 分布随机变量的假设, 也即否定了平稳假设。此即 Box-Pierce χ^2 检验。

如果给定序列经检验否定了平稳假设, 证明有趋势项, 或曰存在“漂移”^[9]。趋势项通常可以取 $t(h)$ 的一次项、二次项等多项式或对数、指数函数等; 亦可通过差分处理使序列化为平稳序列。本次采用二次多项式拟合和差分处理。

处理后仍须检验, 直至通过。

2. 正态性检验及序列正态化处理 本文采用偏态系数 (c_i) 法, 对试验数据 ($x_i, i = 1, 2, \dots, N$) 进行正态性检验。计算公式

$$c_i = M_3/\delta^3 \quad (9)$$

式中, $M_3 = \sum_{i=1}^N (x_i - \bar{x})^3/N$

判别式:

$$|c_i| < 2\sqrt{6/N} \quad (10)$$

上式成立, 接受序列正态假设。对于拒绝正态假设的序列, 采用对数变换法进行处理^[5]。

设非正态序列 $x_i, i = 1, 2, \dots, N$ 进行如下对数变换:

$$x_i = \log(x_i + a) \quad (11)$$

式中, a 为位移因子, 由下式确定其初值:

$$\log(x_{\max} + a) + \log(x_{\min} + a) = 2\log(x_m + a)$$

式中, x_m 为序列的平均值。变换中 x_i 的顺序不变。

调整 a , 使 x_i 通过正态性检验。

(二) 参数序列谱分析及结果

1. 理论与方法 设 $\{x_i\}, i = 1, 2, \dots, n$ 为一个数学期望为 0 方差为 1 的平稳序

列,它的统计特性完全由其协方差函数列 $\{r_k\}$, $k = 0, \pm 1, \dots, \pm m$ 或谱密度函数 $g(\lambda)$ 表示出来,且 $\{r_k\}$ 和 $g(\lambda)$ 相互对应。

如果用谱熵:

$$H(g) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \log g(\lambda) d(\lambda) \quad (12)$$

来表示 $g(\lambda)$ 所含信息量大小:

$$g(\lambda) = \sum_{k=-\infty}^{\infty} r_k e^{-i2\pi k\lambda} \quad \text{应满足约束} \quad (13)$$

$$r_k = \int_{-\frac{1}{2}}^{\frac{1}{2}} g(\lambda) e^{i2\pi k\lambda} d\lambda \quad (14)$$

求得 $g(\lambda)$ 使 (2.10) 达到最大。此即最大熵准则。

偶联 (13)、(14), 可解得^[2]:

$$g(\lambda) = \frac{\delta^2}{\left| 1 - \sum_{k=1}^m a_k e^{-i2\pi k\lambda} \right|^2} \quad (15)$$

该序列的 AR(m) 模型表示为:

$$x_i = \sum_{k=1}^m a_k x_{i-k} + \varepsilon_i \quad \begin{matrix} (i = 1, 2, \dots, n) \\ (k = 1, 2, \dots, m) \end{matrix} \quad (16)$$

由 (15) 得到如下关系式:

$$\begin{bmatrix} r_0 & \dots & r_m \\ \vdots & \ddots & \vdots \\ r_m & \dots & r_0 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} r_1 \\ \vdots \\ r_m \end{bmatrix} \quad (17)$$

由此解出的 a_1, \dots, a_m 就是 AR(m) 模型的参数。

实际计算时, a_k , $k = 1, 2, \dots, m$ 一般用 Burg 递推求得。(本文略)

本文以 FPE 准则为基础,对照谱图,以能揭露主要谱峰又不至使谱线分裂为依据定阶:

$$\sqrt{N} \leq m \leq N / \ln N \quad (18)$$

2. 结果及分析 在最大熵准则下,用 AR(p) 模型对试验参数序列进行模拟,依据 (18) 定阶, AR(m) 模型的阶数 $m = 4-6$ 。AR(m) 模型的拟合值与试验数据的残差序列分别通过了平稳性和正态性检验,模型被接受。E-W 到模拟的具体结果见图 2。

由谱图可看出,最大熵谱的谱图都不是单峰的,多为双峰,亦有三峰;谱峰陡峻尖锐,高低差别显著。最大峰值对应的周期分别是: $>30m$ 的长周期, $20-30m$ 的中长周期和 $6-10m$ 的短周期。其中 S 的周期是, $>30m$ 、 $25-30m$ 及 $10m$; A 的周期是, $28-30m$ 、 $20-22m$ 、 $8-12m$ 及 $6m$; θ 的周期是, $28-30m$ 、 $22m$ 、 $10-15.5m$ 及 $6-10m$, e 的周期是, $>30m$ 、 $12.7m$ 及 $7.5m$ 。

序列分析揭露出看似均质的土壤各点参数处处存在差异。并且这些差异只是在大于 $30m$ 的点距上才有各自的独立性,才是完全随机的,或者说,在轻壤质土分带的大范围

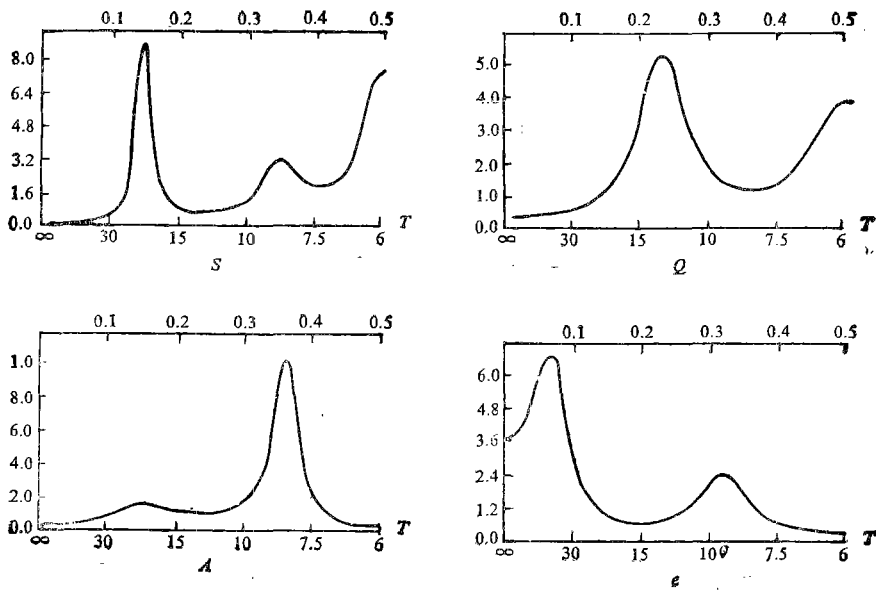


图 2 E-W 列最大熵谱图
Fig. 2 E-W line maximum entropy spectrogram

内,才可粗略地将土壤看成是均质的;但是在 30m 以内的点距上这些差异存在着相依性,且可用数个主要周期迭加的线性模型来描述,也就是说土壤存在着空构结构的相依性。这些小范围内的相依性不是地质构造引起的,主要是由人们的生产活动,如种植、灌溉(包括浑水放淤)、耕作引起的。究竟何种作用为主,怎样起作用,有待进一步研究。

(三) 互谱分析及结果

为了研究土壤水分、土壤盐分对土壤入渗的影响及土壤水分同土壤盐分的关系,我们引入互谱概念,并对参数序列进行互谱分析。这样分析的结果要比个别点的参数相互比较更为全面可靠。

1. 理论与方法 考虑两个序列 x_i 和 y_i , 它们的互谱(或称交叉谱)定义为:

$$S_{xy}(f) = P_{xy}(f) - iQ_{xy}(f) \quad (19)$$

一般称实部 $P_{xy}(f)$ 为协谱,虚部 $Q_{xy}(f)$ 为正交谱。用 $W(f)$ 表示 $\omega(t)$ 的福氏变换,则:

$$P_{xy}(f) = R_e[X(f)Y^{-1}(f)] \quad (20)$$

$$Q_{xy}(f) = I_m[X(f)Y^{-1}(f)] \quad (21)$$

构造两个新序列 a_i 和 b_i :

$$a_i = x_i + iy_i$$

$$b_i = x_i + y_i$$

分别作福氏变换得到:

$$A(f) = X(f) + iY(f)$$

$$B(f) = X(f) + Y(f)$$

序列 a_i 和 b_i 的自谱分别为:

$$P_{aa}(f) = \lim_{T \rightarrow \infty} \frac{1}{T} [P_{xx}(f) + P_{yy}(f) + 2I_m(xy^{-1})] \quad (22)$$

$$P_{bb}(f) = \lim_{T \rightarrow \infty} \frac{1}{T} [P_{xx}(f) + P_{yy}(f) + 2R_s(xy^{-1})] \quad (23)$$

将(20)代入(23),(21)代入(22)即得序列 x_i 和 y_i 的互谱表达式^[3,4]:

$$P_{xy}(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \{ [P_{bb}(f) - P_{xx}(f) - P_{yy}(f)] / 2 \} \quad (24)$$

$$Q_{xy}(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \{ [P_{aa}(f) - P_{xx}(f) - P_{yy}(f)] / 2 \} \quad (25)$$

(24)、(25)右边的量均可用最大熵法分别对序列 x_i, y_i, a_i 和 b_i 计算得到。

人们感兴趣的凝聚谱、位相谱及滞后长度谱分别由(26)、(27)和(28)得到^[3,4]:

$$Z_{xy} = \frac{P_{xy}^2 + Q_{xy}^2}{P_{xx} \cdot P_{yy}} \quad (26)$$

$$Q_{xy}(f) = \arctg \left(\frac{Q_{xy}}{P_{xy}} \right) \quad (27)$$

$$L(T) = \frac{Q_{xy}(f)T}{2\pi} \quad (28)$$

2. 结果及分析 用以上介绍的方法对试验参数进行互谱分析。因为,我们主要通过样本估计的凝聚谱研究两个序列之间的关系,因此,需要对凝聚谱的显著性进行检验。

$$F = \frac{(\gamma - 1)Z}{1 - Z} \quad (29)$$

$$\gamma = \left(2n - \frac{m}{2} \right) / m$$

式中, n 为样本大小, m 为阶数。

若 $F > F_\alpha$ 则显著:

α 为信度, F_α 通过查表可得。

具体本文: $n = 17, m = 5, \gamma = 6.3$

查 F_α 表, $F_{0.05} = 2.81, F_{0.01} = 4.34$

由(29)得

$Z > 0.36$ 接受显著假设(信度 0.05)

$Z > 0.45$ 接受显著假设(信度 0.01)

E-W 列的分析及检验结果见图 3。

由互谱谱图,我们得出的结论是:(1) $S-\theta$ 负相关,因设土壤均质 $S = 1-\theta$; $A-\theta$ 正相关, θ 大,下渗快; $e-\theta$ 负相关,水去盐存, $S-e, A-e$ 负相关,土壤盐分影响土壤的渗透性,(2)除 E-W、N-S 方向的 $A-\theta$ 在 6, 7.5m 短周期有较好关系外,其他在 30m 或大于 30m 长周期上有密切关系;(3) $S-e$ 关系较差; NE-SW、SE-NW 中, $A-Q$ 关系不好。(反映了土壤水分取样和人渗试验的时间差)。

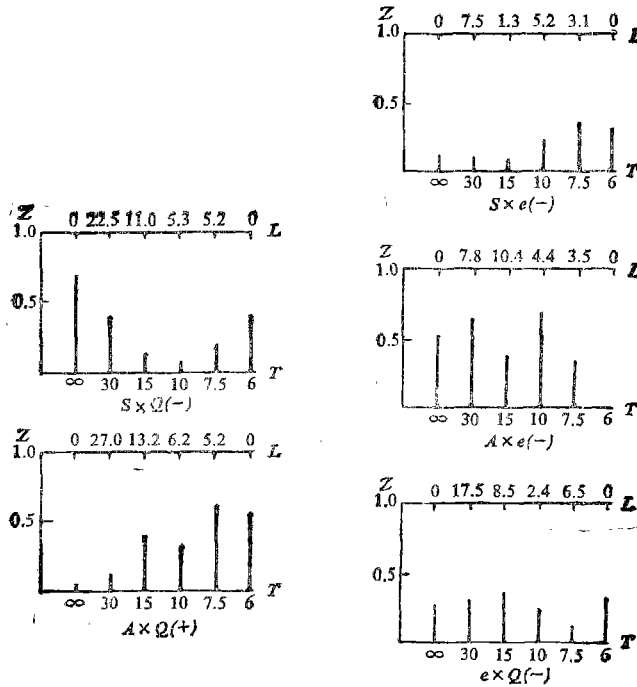


图3 E-W 列交叉谱谱图
Fig. 3 E-W line cross spectrum spectrogram

(四) 多维序列 AR 模型及分析结果

引入多维序列 AR 模型对试验参数进行分析,是为了从序列上研究土壤水分、土壤盐分及土壤入渗(多输入)与土壤入渗(单输出)的回归关系。

设 $Z_i = (x_i^{(1)}, \dots, x_i^{(l)}, y_i, \dots, y_i^{(r)})^T$, 其中 $y_i^{(1)}, \dots, y_i^{(r)}$ 为 r 个输入变量, $x_i^{(1)}, \dots, x_i^{(l)}$ 为 l 个输出变量的 k 维 ($k = l + r$) 随机序列, $\{Z_i\}$ 满足如下形式的自回归模型:

$$Z_i = \sum_{j=1}^P A_{pi} Z_{i-j} + U_j \tag{30}$$

(30) 可表示为:

$$x_i = \sum_{j=1}^P A_{pi}^{(1,1)} x_{i-j} + \sum_{j=1}^P A_{pi}^{(1,2)} y_{i-j} + U_j \tag{31}$$

特别地,如果输出量 x_i 是标量,则系统是单输出多输入的情形,则(31)可表为:

$$x_i = \sum_{j=1}^P a_{pi}^{(1,1)} x_{i-j} + \sum_{i=1}^r \sum_{j=1}^P a_{pi}^{(1,i+1)} y_{i-j} + U_j \tag{32}$$

这是一类带有多输入的一维自回归模型。本文仅用(32)。

E-W 列多维序列分析的结果见图4。

由图4的图形读图仅能看到1条曲线,这是由于误差平方和 $< 1 \times 10^{-5}$, 不足以在如此小的图形显示出来,看到的是模拟值与实测值完全重合。

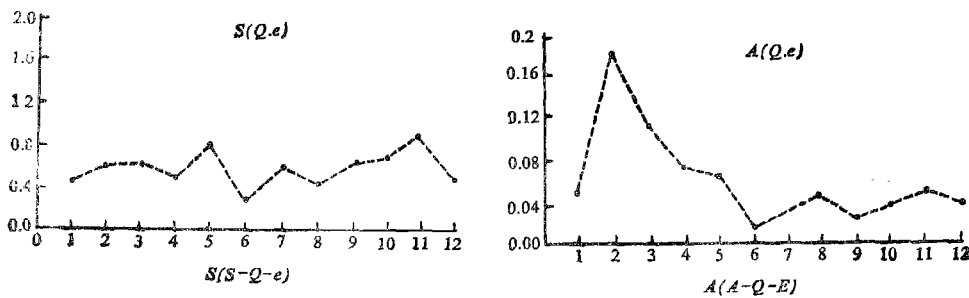


图 4 E-W 列多维谱分析模拟结果

Fig. 4 E-W line multichannel spectral analysis of analog result

这种分析的好处是在已经进行过多次测验的工作面上,对测验困难或花费大的项目可以减少测点,由其他项目的完整资料通过模型对减去测点的项目进行预测。如对缺少测点的费时劳力的 S 或 A 进行缺测点的预测插补。

参 考 文 献

1. 李鸿杰、杜历, 1990: 土壤入渗空间变异特性初步研究。水利学报, 第 8 期, 35—42 页。
2. 项静恬、杜金观等, 1986: 动态数据处理—时间序列分析。气象出版社。
3. 黄嘉佑、李黄, 1984: 气象中的谱分析。气象出版社。
4. 曹鸿兴, 1983: 局地天气预报的数据分析方法。气象出版社。
5. 方再根, 1988: 计算机模拟和蒙特卡洛方法。北京工业学院出版社。
6. Davidoff B. Lewis J. W. and Selim H. M., 1986: Temperature variability along a transect. *Soil Science*, 142(2):114—123.

SPATIAL SERIAL ANALYSIS OF VARIABILITIES OF SOIL MOISTURE SALINITY AND INFILTRATION AND THEIR INTER RELATIONS

Li Hongjie and Du Li

(Ningxia Institute of Water Conservancy, 750021)

Summary

The variabilities of soil moisture, Salinity and infiltration and their inter relation were analysed by using the theoretical method of serial analysis in this paper.

Satisfactor results were obtained by the method. The maximum entropy spectral method ometry estimation used had a high resolution ratio and could expose the superior epe-riod. The inter relations among the series were exposed by the mutual spectral analysis. The results obtained by a multidimension AR(p) model could fit well with the mtaswied results.

Key words Serial analysis, Maximum entropy, Autocorrelation dependency, Independence