

模拟退火算法在皮尔逊III型分布参数估计中的应用

摘 要：为研究模拟退火算法在水文频率曲线适线中的适用性，以某一站点 1949 年-1995 年的年降水量资料为例，采用模拟退火算法按照纵向离差绝对值和最小准则(ABS)对皮尔逊III型分布进行参数优化，并与矩法的计算结果作对比。结果表明：模拟退火算法是一种可行的水文频率曲线参数估计方法，其估计精度优于矩法。

关 键 词：模拟退火算法；皮尔逊III型曲线；参数优化

水文频率计算是基于水文变量的样本资料探求总体的统计规律，对未来的水文情势作出预估，进而为确定工程规模、防洪标准、以及供水保证率等重要资料。根据《水利水电工程设计洪水计算规范》(SL44-2006)，皮尔逊III型曲线是我国水文频率分析计算建议使用的线型，其分布参数估计的传统方法主要有矩法、双权函数法、线性矩法和目估适线法等^[1]。矩法计算简单，但精度偏低；双权函数法虽然计算精度有所提高，但未能在单权函数法基础上进一步降阶；目估适线法则掺有较多的主观因素。近年来，随着计算机技术和人工智能优化算法的不断发展，优化算法被成功运用到水文计算中。其中，模拟退火算法(Simulated Annealing, SA)对于一般无约束优化问题具有易实现和渐进收敛的特点。本文首先利用矩法求得参数估计值，再根据纵向离差绝对值和最小准则(ABS)，利用 Matlab 软件编程研究了模拟退火算法在皮尔逊III型分布参数估计中的应用。

1. 模拟退火算法

1.1 基本原理

模拟退火算法是一种以概率 1 收敛于全局最优解的随机搜索算法，由 Kirkpatrick 等^[2]在 1983 年提出。算法的思想来源于固体退火原理：将固体加温至温度充分高，再让其徐徐冷却，加热时，固体内部粒子随温升变为无序状，内能增大，而冷却时粒子渐趋有序，在每个温度都达到平衡态，最后在常温时达到基态，内能减为最小。基于对该过程的模拟，算法在求解过程中的一个解 x 及其目标函数 $f(x)$ 分别与固体的一个微观状态 x 及其能量 E_x 相对应，算法进程递减的控制参数 T 担当固体退火过程中温度的角色，则对于 T 的每一取值，算法采用 Metropolis 准则，持续进行“产生新解-判断-接受或舍弃”的迭代过程而达到该

温度下的平衡点。算法的求解过程如下：

(1) 给定初始退火温度 T_k (令 $k=0$)、终止温度 T_{final} ，产生随机初始解 x_0 。

(2) 在温度 T_k 下重复执行如下操作，直至达到温度 T_k 的平衡状态。

①在解 x 的领域中产生新的可行解 x' 。

②计算 x' 的目标函数 $f(x')$ 和 x 的目标函数 $f(x)$ 的差值 Δf 。

③依照概率 $\min\{1, \exp(-\Delta f / T_k)\} > \text{random}[0,1]$ 接收 x' ，其中 $\text{random}[0,1]$ 是 $[0,1]$ 区间内的随机数。

(3) 退火操作：按退火策略进行降温，具体有对数下降、快速降温、直线下降和指数退温等^[3]。本文采用指数退温，即 $T_{k+1} = CT_k, k \leftarrow k+1$ ，其中 $C \in (0,1)$ ，一般取 0.95。若满足收敛判据，则退火过程结束；否则，转至步骤(2)。

其中退火温度控制着求解过程向最优值的优化方向进行，同时它又以概率 $\exp(-\Delta f / T_k)$ 来接收劣质解，因此算法可以跳出局部极值点，只要初始温度足够高，退火过程足够慢，算法就能收敛到全局最优解。

1.2 Metropolis 准则

Metropolis 准则赋予模拟退火算法一种概率突跳特性，是该算法在计算过程中能摆脱局部最优解的关键，其含义是：给定当前粒子一个初始状态，即确定一个初始解 x_0 ，并计算其目标函数值 $f(x_0)$ 作为解的当前状态。然后对当前解随机扰动生成一个新解 $f(x_{new})$ 。如果 $f(x_{new}) \leq f(x_0)$ ，接受当前解作为新解，即令 $x_0 \leftarrow x_{new}$ ，否则按照式 (1) 判断是否接受新解。

$$(1) \quad \begin{cases} \text{if } \left\{ \exp \frac{-[f(x_{new}) - f(x_0)]}{k_B T} \right\} > \text{rand}(0,1) & \text{接受} \\ \text{else} & \text{不接受} \end{cases}$$

其中， k 为 Boltzmann 常数， T 为当前计算温度。

1.3 算法流程图

模拟退火算法流程图如下图所示。

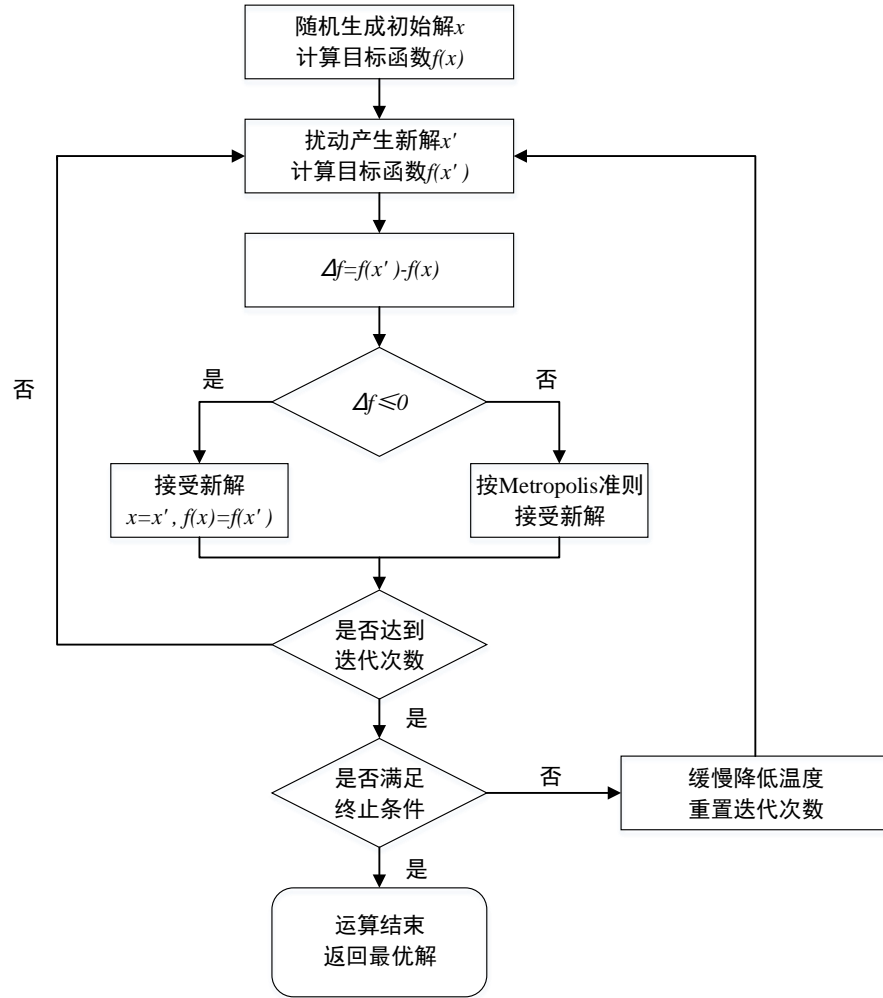


图 1 模拟退火算法流程图

2. 水文频率曲线适线准则

适线法按不同的适线准则分为三种，即纵向离差平方和最小准则(OLS)、纵向离差绝对值和最小准则(ABS)、纵向相对离差平方和最小准则(WLS)，其中以离差平方和最小准则(OLS)最为常用，但理论研究成果表明，纵向离差绝对值和最小准则(ABS)精度最好^[4]。本文以 ABS 准则为目标函数。皮尔逊Ⅲ型曲线 ABS 准则适线目标函数为：

$$S_A(\hat{\theta}) = \min \left\{ \sum_{i=1}^n \left| x_i - \bar{x}(1 + C_v \varphi) \right| \right\} \quad (2)$$

式中： $S_A(\hat{\theta})$ 为 ABS 准则的目标函数； $\hat{\theta}$ 为待求参数集合， $\hat{\theta} = \{\bar{x}, C_v, C_s\}$ ； \bar{x} 为系列均值； C_v 为系列变差系数； C_s 为系列偏态系数； n 为系列长度； x_i 为系列值； φ 为离均系数，为 P_i 和 C_s 的函数，可查皮尔逊Ⅲ型分布系数表或进行数值计

算得到； P_i 为水文频率。

3. 实例应用

以某一站点 1949 年-1995 年的年降水量资料为例，选用皮尔逊III型分布作为其频率曲线线型，以矩法的估计值为初值，采用模拟退火算法分别按照 ABS 准则进行分布参数优化计算，并与初值作对比。

3.1 矩法预估初值

矩法对皮尔逊III型分布的三个统计参数 \bar{x}, C_v, C_s 的无偏估值公式如下：

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3)$$

$$C_v = \sqrt{\frac{\sum_{i=1}^n (K_i - 1)^2}{n-1}} \quad (4)$$

$$C_s = \frac{n^2}{(n-1)(n-2)} \frac{\sum_{i=1}^n (K_i - 1)^3}{nC_v^3} \approx \frac{\sum_{i=1}^n (K_i - 1)^3}{(n-3)C_v^3} \quad (5)$$

其中， $K_i = \frac{x_i}{\bar{x}}$ ，称模比系数。

计算见表 1，最终求得 $\bar{x}=395.32$ ， $C_v=0.358$ ， $C_s=0.518$ ， $S_A(\hat{\theta})=593.04$ 。

表 1 某站实测年径流量计算表

年份	年径流量 ($\times 10^8 \text{m}^3$)	序号	按大到小 排列	模比系数 K_i	K_i-1	$(K_i-1)^2$	$(K_i-1)^3$	$P = \frac{m}{n+1}$ (%)	Φ_i	纵向离差绝对 值
1949	680.6	1	750.9	1.899	0.899	0.809	0.727	2.083%	2.2980	30.3518
1950	468.4	2	680.6	1.721	0.721	0.520	0.375	4.167%	1.9053	15.6309
1951	489.2	3	650.9	1.646	0.646	0.418	0.270	6.250%	1.6607	20.5572
1952	450.6	4	620.8	1.570	0.570	0.325	0.185	8.333%	1.4694	17.5238
1953	436.8	5	586.2	1.483	0.483	0.233	0.112	10.417%	1.2952	7.5739
1954	586.2	6	585.2	1.480	0.480	0.231	0.111	12.500%	1.1894	21.5578
1955	567.9	7	567.9	1.436	0.436	0.190	0.083	14.583%	1.0835	19.2417
1956	473.9	8	549.1	1.389	0.389	0.151	0.059	16.667%	0.9776	15.4256
1957	357.8	9	539	1.363	0.363	0.132	0.048	18.750%	0.8717	20.3095
1958	650.9	10	534	1.351	0.351	0.123	0.043	20.833%	0.7765	28.7815
1959	391	11	489.2	1.237	0.237	0.056	0.013	22.917%	0.6974	4.8145
1960	201.2	12	485	1.227	0.227	0.051	0.012	25.000%	0.6182	2.1895
1961	452.4	13	474	1.199	0.199	0.040	0.008	27.083%	0.5508	0.7306
1962	750.9	14	473.9	1.199	0.199	0.039	0.008	29.167%	0.4834	10.1717
1963	585.2	15	468.4	1.185	0.185	0.034	0.006	31.250%	0.4204	13.5866
1964	304.5	16	452.4	1.144	0.144	0.021	0.003	33.333%	0.3603	6.0840
1965	370.5	17	450.6	1.140	0.140	0.020	0.003	35.417%	0.3003	12.7814
1966	351	18	436.8	1.105	0.105	0.011	0.001	37.500%	0.2403	7.4787
1967	294.8	19	427	1.080	0.080	0.006	0.001	39.583%	0.1802	6.1761
1968	360.9	20	425	1.075	0.075	0.006	0.000	41.667%	0.1262	11.8149
1969	276	21	391	0.989	-0.011	0.000	0.000	43.750%	0.0738	14.7610
1970	549.1	22	372	0.941	-0.059	0.003	0.000	45.833%	0.0213	26.3368
1971	534	23	370.5	0.937	-0.063	0.004	0.000	47.917%	-0.0311	20.4127
1972	349	24	365	0.923	-0.077	0.006	0.000	50.000%	-0.0836	18.4885
1973	350	25	360.9	0.913	-0.087	0.008	-0.001	52.083%	-0.1353	15.2705
1974	372	26	357.8	0.905	-0.095	0.009	-0.001	54.167%	-0.1870	11.0525
1975	292	27	357	0.903	-0.097	0.009	-0.001	56.250%	-0.2387	4.5345
1976	485	28	351	0.888	-0.112	0.013	-0.001	58.333%	-0.2904	3.2166
1977	427	29	350	0.885	-0.115	0.013	-0.002	60.417%	-0.3422	3.1121
1978	620.8	30	349	0.883	-0.117	0.014	-0.002	62.500%	-0.3943	9.4831
1979	539	31	306	0.774	-0.226	0.051	-0.012	64.583%	-0.4464	26.1458
1980	474	32	305	0.771	-0.229	0.052	-0.012	66.667%	-0.4985	19.7747
1981	292	33	304.5	0.770	-0.230	0.053	-0.012	68.750%	-0.5506	12.9037
1982	228	34	294.8	0.746	-0.254	0.065	-0.016	70.833%	-0.6035	15.1146
1983	357	35	292	0.739	-0.261	0.068	-0.018	72.917%	-0.6576	10.2487
1984	425	36	292	0.739	-0.261	0.068	-0.018	75.000%	-0.7118	2.5828
1985	365	37	277.3	0.701	-0.299	0.089	-0.027	77.083%	-0.7694	9.1334
1986	241	38	276	0.698	-0.302	0.091	-0.028	79.167%	-0.8270	2.2839
1987	267	39	267	0.675	-0.325	0.105	-0.034	81.250%	-0.8925	2.0093
1988	305	40	241	0.610	-0.390	0.152	-0.060	83.333%	-0.9633	17.9847
1989	306	41	238.9	0.604	-0.396	0.157	-0.062	85.417%	-1.0364	9.7487
1990	238.9	42	228	0.577	-0.423	0.179	-0.076	87.500%	-1.1182	9.0672
1991	277.3	43	217.9	0.551	-0.449	0.201	-0.090	89.583%	-1.2000	7.5858
1992	170.8	44	208.5	0.527	-0.473	0.223	-0.106	91.667%	-1.3052	2.1021
1993	217.9	45	201.2	0.509	-0.491	0.241	-0.118	93.750%	-1.4162	6.3071
1994	208.5	46	187.9	0.475	-0.525	0.275	-0.144	95.833%	-1.5529	12.3511
1995	187.9	47	170.8	0.432	-0.568	0.323	-0.183	97.917%	-1.7860	28.2464
									总和	593.0402

3.2 参数优化

以矩法预估值为初始值，利用 Matlab 软件中优化工具箱的模拟退火算法对皮尔逊III型分布参数进行优化。针对迭代中 φ 值的确定问题，考虑从标准化降水指数（SPI）与 Z 指数进行计算。

（1）标准化降水指数（SPI）是表征某时段降水量出现的概率多少的指标之一，在计算出某时段内降水量的 Γ 分布概率后，再进行正态标准化处理，最终用标准化降水累积频率分布来划分干旱等级^[5]。

在对 Γ 分布概率进行正态标准化处理后，得到标准化正态分布函数：

$$P(x < x_0) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-SPI^2/2} dx \quad (6)$$

对式(6)进行近似求解可得：

$$SPI = \begin{cases} -\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right), t = \sqrt{\ln\left(\frac{1}{P^2}\right)}, P < 0.5 \\ t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}, t = \sqrt{\ln\left(\frac{1}{(1-P)^2}\right)}, 0.5 \leq P < 1 \end{cases} \quad (7)$$

上式中， $c_0=2.515517$ ， $c_1=0.802853$ ， $c_2=0.010328$ ， $d_1=1.432788$ ， $d_2=0.189269$ ， $d_3=0.001308$ 。而在工程水文学中，概率往往以大于某一确定值的形式出现，所以在本次优化计算中，式(7)的计算式应在 F 条件不变的情况下相互调换。

（2）Z 指数是目前在我国使用较为广泛的气象干旱指标，它假定降水量服从皮尔逊III型分布，通过对降水量进行频率分析来确定干旱的程度^[6]。同样对降水量进行正态化处理，转换为以 Z 为变量的标准正态分布，可得：

$$Z = \frac{6}{C_s} \left(\frac{C_s}{2} \varphi + 1 \right)^{\frac{1}{3}} - \frac{6}{C_s} + \frac{C_s}{6} \quad (8)$$

综合标准化降水指数（SPI）与 Z 指数相关内容，二者在理论上应相等，故可以从两者关系，反求 φ 值。

利用 Matlab 软件编程进行计算，主程序代码如下：

```
编辑器 - C:\Users\XuLei\Desktop\最优化结课论文\ABS准则下的模拟退火算法\P3SAABS.m
P3ThoFre.m  ABS.m  P3SAABS.m  +
1  function [ex,cv,cs,s1]=P3SAABS(ex1, cv1, cs1, ep, ey)
2  x=[ex1; cv1; cs1];
3  rand('twister', 71); % These two commands are only included to
4  randn('state', 59); % make the results reproducible.
5  options=saoptimset('TemperatureFcn',@temperatureexp,'Tolfun',1e-8,'InitialTemperature',100,'Display','off');
6  [xe,fval] = simulannealbnd(@(x)ABS(ey, P3ThoFre(x(1), x(2), x(3),ep)), x);
7  ex=xe(1);
8  cv=xe(2);
9  cs=xe(3);
10 s1=fval;
11 fprintf('ex is%9.4f\n',ex);
12 fprintf('cv is%8.4f\n',cv);
13 fprintf('cs is%8.4f\n',cs);
14 fprintf('S is%10.4f\n',s1);
15
```

图 2 主程序 P3SAABS 代码

相关函数代码如下：

```
编辑器 - C:\Users\XuLei\Desktop\最优化结课论文\ABS准则下的模拟退火算法\P3ThoFre.m
P3ThoFre.m  ABS.m  P3SAABS.m  +
1  function y = P3ThoFre(varargin)
2  ex=varargin{1};cv=varargin{2};cs=varargin{3};p=varargin{4};
3  c0=2.515517;c1=0.802853;c2=0.010328;
4  d1=1.432788;d2=0.189269;d3=0.001308;
5  m=length(p);
6  for i=1:m
7      if p(i)<0.5
8          w=sqrt(-2*log(p(i)));
9          u(i)=(w-(c0+c1*w+c2*w^2)/(1+d1*w+d2*w^2+d3*w^3));
10         else
11             w=sqrt(-2*log(1-p(i)));
12             u(i)=-(w-(c0+c1*w+c2*w^2)/(1+d1*w+d2*w^2+d3*w^3));
13         end
14         kt(i)=u(i)+cs/2*((u(i)^2-1)/3)+cs^2*(u(i)^3/108-u(i)/18)+cs^3*(1/216-u(i)^3/216)+cs^4*u(i)/1296-cs^5/23328;
15     end
16     y=ex+kt*ex*cv;
17 end
```

图 3 P3ThoFre 函数代码

```
编辑器 - C:\Users\XuLei\Desktop\最优化结课论文\ABS准则下的模拟退火算法\ABS.m
P3ThoFre.m  ABS.m  P3SAABS.m  +
1  function [sa] = ABS(ey, y)
2  % Purpose:
3  %
4  sa = sum(abs(ey - y));
```

图 4 ABS 函数代码

输入参数值，得到以下运行结果：

```
命令窗口
>> P3SAABS(395.32, 0.318, 0.518, [0.020833333 0.041666667 0.0625 0.083333333 0.104166667 0.125 0.145833333 0.166666667 0.1875 0.21], [750.9 680.6 650.9 620.8 586.2 585.2 567.9 549.1 539 534 489.2 485 474 473.9 468.4 452.4 450.6 436.8 420.8])
Optimization terminated: change in best function value less than options.FunctionTolerance.
ex is 393.6598
cv is 0.3892
cs is 0.8227
fx S is 448.8307
```

图 5 参数输入及运行结果

将优化结果同矩法预估值比较，见下表。

表 2 参数估计结果比对

适线准则	参数估计方法	\bar{x}	C_V	C_S	目标函数值
ABS	矩法	395.32	0.318	0.528	593.04
	SA	393.66	0.389	0.823	448.83

由计算结果可以看出，利用模拟退火算法优化的纵向离差绝对值和小于矩法所得值，误差更小。此外，模拟退火算法求得的 $C_S \approx 2.12C_V$ ，更符合一般年径流问题特性^[4]。综合来看，模拟退火算法是一种可行的水文频率曲线参数估计方法。

参考文献

- [1] 水利部. SL44-2006, 水利水电工程设计洪水计算规范[S]. 北京: 中国水利水电出版社, 2006.
- [2] Kirkpatrick S., Gelatt C. D. JR, Vecchi M. P. Optimization by Simulated Annealing[J]. Science, 1983, 220(4598): 671-680.
- [3] 宋松柏, 康艳. 3 种智能优化算法在设计洪水频率曲线适线法中的应用[J]. 西北农林科技大学学报(自然科学版), 2008, 36(2): 205-210.
- [4] 王文川. 工程水文学[M]. 北京: 中国水利水电出版社, 2013.
- [5] 刘蕾, 张鑫, 徐静. SPI 指数和 Z 指数在青海省东部农业区应用分析[J]. 中国农村水利水电, 2015, (04): 97-100.
- [6] 王帅兵, 李常斌, 杨林山, 等. 基于标准化降水指数与 Z 指数的洮河流域干旱趋势分析[J]. 干旱区研究, 2015, 32(3): 565-572.