

SAS 软件在气象数据处理中的应用

李够霞

(中国科学院水土保持研究所, 陕西 杨凌 712100)

一、前言

计算机技术的发展促进了社会的信息化进程, 各种处理数据软件的出现使科研人员处理大量数据变得得心应手, 各种形式的软件层出不穷, 选用一种适合自己的软件是提高工作效率的有效途径。

中国生态系统研究网络生态站 (CERN) 的野外气象自动观测系统的仪器所采集的数据不能直接利用, 必须经过提取、转换等过程才可变成有用数据, 但是原始数据量大、要素多。采用 EXCEL 软件进行转换, 不但繁琐, 容易出错, 而且工作量大。作者应用 SAS 软件编写了一套程序, 使这一工作变得非常容易, 同样工作量利用 EXCEL 可能需要十几个工作日, 而使用这套程序几秒钟就可完成。

二、程序功能

自动气象站是每月采集一次数据, 所有气象要素都保存在一个文件里, 即以所有气象要素为表头

的一张二维表, 数据文件类型为文本文件, 扩展名为 .dat, 数据格式如表 1。每年有 12 个 .dat 文件。每张表的时间变量是以 1 小时为时段, 所以数据量比较大, 最终则需转换成每要素一张表, 以每小时为一个变量, 即 24 个时间变量, 如表 2 的形式, 数据文件类型仍为文本文件。扩展名也仍为 .dat,

本套程序是利用 SAS (Statistical Analysis System) 系统编写的, 它可实现从表 1 格式转换成表 2 的格式。表 1 为原始的 .dat 文件, 在这套程序中为气象辐射站所采集的 .dat 文件, 表 2 为某个气象要素以小时为变量的 .dat 文件。

表1 气象数据原始文件

日期	时间	T	T'	Rh	P	Tg	V	W	I	SH
4-21-98	12:05:02	26.3	15.69	32	868.6	31.6	4.6	144	5.5	0.1
4-21-98	13:05:02	26.9	14.44	24	867.5	38.1	7.4	118	17	0.1
4-21-98	14:05:02	27.6	14.86	24	866.6	37.3	5.3	119	0	0.9
:	:									
:	:									

表2 气象数据 24 小时变量

日期	0:05:02	1:05:02	2:0502	3:05:02	4:05:02	5:05:02	6:05:02	.
4-21-98								
4-21-98								
4-21-98								

三、程序流程

要将上述表 1 转换成表 2 的形式, 最终产生全年的每个气象要素的文本文件。使用 SAS 编写程序的数据流程如图 1。

四、源程序

在 c:盘下建立 c:\assas 目录, 每年 1-12 月采集的数据集文件名为 dt1.dat—dt12.dat, 将这 12 个文件拷到该目录下, 执行下列程序, 最终产生的 data1.dat—data9.dat 文件也存在 c:\assas 目录下。

第 1 到第 177 行为源程序。

```

1. /* 将每月产生的.dat 文件分别转换成 SAS 数据集 */
2. data dt1sas;
3. infile 'c:\assas\dt1.dat';
4. input date $10. time $8. v1-v24;
5. run;
6. data dt2sas;
7. infile 'c:\assas\dt2.dat';
8. input date $10. time $8. v1-v24;
9. run;
10. data dt3sas;
11. infile 'c:\assas\dt3.dat';
12. input date $10. time $8. v1-v24;
13. run;
14. data dt4sas;
15. infile 'c:\assas\dt4.dat';

```

```

16. input date : $10. time : $8. v1-v24;
17. run;
18. data dt5sas;
19. infile 'c:\assas\dt5.dat';
20. input date : $10. time : $8. v1-v24;
21. run;
22. data dt6sas;
23. infile 'c:\assas\dt6.dat';
24. input date : $10. time : $8. v1-v24;
25. run;
26. data dt7sas;
27. infile 'c:\assas\dt7.dat';
28. input date : $10. time : $8. v1-v24;
29. run;
30. data dt8sas;
31. infile 'c:\assas\dt8.dat';
32. input date : $10. time : $8. v1-v24;
33. run;
34. data dt9sas;
35. infile 'c:\assas\dt9.dat';
36. input date : $10. time : $8. v1-v24;
37. run;
38. data dt10sas;
39. infile 'c:\assas\dt10.dat';
40. input date : $10. time : $8. v1-v24;
41. run;
42. data dt11sas;
43. infile 'c:\assas\dt11.dat';
44. input date : $10. time : $8. v1-v24;
45. run;
46. data dt12sas;
47. infile 'c:\assas\dt12.dat';
48. input date : $10. time : $8. v1-v24;
49. run;

50. /* 合并 12 个 SAS 数据集 */
51. data dqsas;
52. set dt1sas dt2sas dt3sas dt4sas dt5sas dt6sas dt7sas dt8sas
dt9sas dt10sas dt11sas dt12sas;
53. run;

54. /* 产生以日期及各气象要素为变量的 SAS 数据集 */
55. data file1(keep=date v1)
56. file2 (keep=date v2)
57. file3 (keep=date v3)
58. file4 (keep=date v4)
59. file5 (keep=date v5)
60. file6 (keep=date v13)
61. file7 (keep=date v14)
62. file8 (keep=date v15)
63. file9 (keep=date v16)
64. ;
65. set dqsas;
66. run;

67. /* 产生以小时为变量的 SAS 数据集 */
68. proc transpose data= file1 out=trv1 prefix=t;
69. var v1;
70. by date;
71. run;
72. proc transpose data= file2 out=trv2 prefix=t;
73. var v2;
74. by date;
75. run;
76. proc transpose data= file3 out=trv3 prefix=t;
77. var v3;
78. by date;
79. run;
80. proc transpose data= file4 out=trv4 prefix=t;
81. var v4;
82. by date;
83. run;
84. proc transpose data= file5 out=trv5 prefix=t;
85. var v5;

86. by date;
87. run;
88. proc transpose data= file6 out=trv6 prefix=t;
89. var v13;
90. by date;
91. run;
92. proc transpose data= file7 out=trv7 prefix=t;
93. var v14;
94. by date;
95. run;
96. proc transpose data= file8 out=trv8 prefix=t;
97. var v15;
98. by date;
99. run;
100. proc transpose data= file9 out=trv9 prefix=t;
101. var v16;
102. by date;
103. run;

104. /* 产生只保留日期和时间变量的 SAS 数据集 */
105. data data1(keep= date t1-t24);
106. set trv1;
107. run;
108. data data2(keep= date t1-t24);
109. set trv2;
110. run;
111. data data3(keep= date t1-t24);
112. set trv3;
113. run;
114. data data4(keep= date t1-t24);
115. set trv4;
116. run;
117. data data5(keep= date t1-t24);
118. set trv5;
119. run;
120. data data6(keep= date t1-t24);
121. set trv6;
122. run;
123. data data7(keep= date t1-t24);
124. set trv7;
125. run;
126. data data8(keep= date t1-t24);
127. set trv8;
128. run;
129. data data9(keep= date t1-t24);
130. set trv9;
131. run;

132. /* 将 SAS 数据集转换成文本文件 */
133. data _null_;
134. set data1;
135. file 'c:\assas\data1.dat';
136. put date t1-t24;
137. run;
138. data _null_;
139. set data2;
140. file 'c:\assas\data2.dat';
141. put date t1-t24;
142. run;
143. data _null_;
144. set data3;
145. file 'c:\assas\data3.dat';
146. put date t1-t24;
147. run;
148. data _null_;
149. set data4;
150. file 'c:\assas\data4.dat';
151. put date t1-t24;
152. run;
153. data _null_;
154. set data5;
155. file 'c:\assas\data5.dat';
156. put date t1-t24;
157. run;

```

```

158. data _null_;
159. set data6;
160. file 'c:\assas\data6.dat';
161. put date t1-t24;
162. run;
163. data _null_;
164. set data7;
165. file 'c:\assas\data7.dat';
166. put date t1-t24;
167. run;
168. data _null_;
169. set data8;
170. file 'c:\assas\data8.dat';
171. put date t1-t24;
172. run;
173. data _null_;
174. set data9;
175. file 'c:\assas\data9.dat';
176. put date t1-t24;
177. run;

```

5. 结束语

软件 SAS (Statistical Analysis System) 软件系统是大型集成应用系统, 功能齐全而且强大, 具备完备的数据访问、数据管理、数据分析、报告编制和数据呈现功能。SAS 软件不仅在统计分析、时间序列分析、运筹决策上功能强大, 并可将各种数据以灵活多样的各种报表、图形和三维透视的形式直观地表现出来, 而且在数据的处理上也显示了

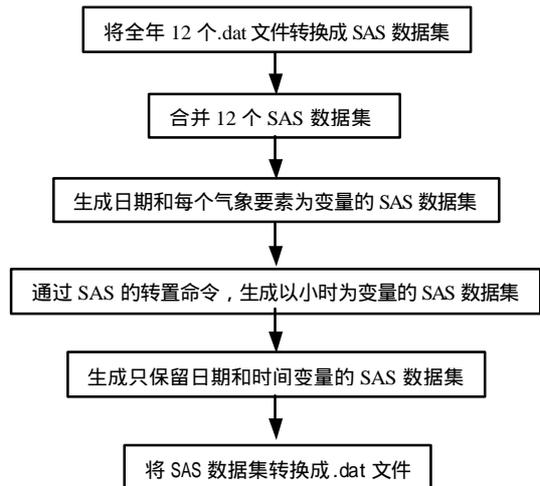


图 1 程序数据流程图

比较强的优势。用 SAS 编程语法简单, 易学易用, 程序处理速度快。不论你从事简单还是复杂的工作, 不论你是 SAS 系统的初学者还是老用户, SAS 系统都能满足你的需要。作者特向你推荐使用 SAS 软件。

(参考文献略)

最新发现与创新

高效信息压缩软件

压缩率比世界著名压缩软件高五至十倍

科技日报报道, 过量信息造成的传输受阻、速度下降, 给生活在网络时代的人们平添了一份新的烦恼。中科院软件所控股是北京九州计算机网络有限公司(下简称九州公司) 推出的最新信息压缩软件 “ 压宝 ”, 为急于摆脱 “ 信息梗阻 ” 苦恼的人们带来了福音。

据悉, 目前世界通用的压缩软件一般分为两种, 即专用压缩软件和普通压缩软件。前者压缩效率比较高, 适用于对音响、影视的大容量信息的压缩, 但使用成本较高, 且仅能由人工进行信息类型判别和相应软件的调用 ; 后者虽使用方便, 成本较低, 但对大容量信息压缩效率低, 多为文本信息的压缩。

九州公司的科研人员独辟蹊径, 应用最新的优化控制理论, 实现了从设计思路到编程技术的全面创新, 可方便地完成对 Windows 环境中 11 类常见计算机文件类型 (如

文本、图像、可执行代码、多类信息复合等) 自动实施高质量、高效率压缩, 平均压缩率比目前世界著名通用压缩软件高 5 倍至 10 倍。该软件及其设计思路均获得美国专利。有关专家指出, 这种以设计思路和编程技术同时获得国际发明专利的, 在世界软件行业都十分少见。

记者在演示现场见到, 经该软件瞬间压缩并还原的图像逼真, 乐曲悦耳, 专家们告诉记者, 此信息是压缩了 20 倍后还原的效果。

(可心)

