中国地面气候资料日值数据集(V3.0)数据提取

China daily ground climate data extraction

中国地面气候资料日值数据集包含了1951年1月以来的气压、气温、降水量、蒸发量、相对湿度、风向风速、日照时数和0cm地温要素的逐日数据。该数据集是以TXT文本格式存储，使用时还需要进一步提取和处理，本文便是利用Python程序进行数据的快速提取。

中国地面气候资料日值数据集（V3.0）是遥感、GIS、水文、气候变化等研究领域中的常用数据源。该数据集包含了中国699个基准、基本气象站自1951年1月以来的气压、气温、降水量、蒸发量、相对湿度、风向风速、日照时数和0cm地温要素的逐日数据。

该数据集是以TXT文本格式存储，每个文件中存储了逐月的所有气象站的逐日数据，数据的存储结构简单，但使用起来还需要进一步提取和处理，例如，若提取某一地区某一时间段的降水数据，则需要从该时间段内的每个文件中找出该地区包括的气象站，然后提取数据，格式化存储。这便是本文的写作目的，利用Python程序进行数据的快速提取。

1. 数据集说明

数据集中文名称：中国地面气候资料日值数据集(V3.0)

数据集代码：SURF\_CLI\_CHN\_MUL\_DAY

数据集版本：V3.0

数据集建立时间：20120804

数据TXT文件名称说明：

数据文件命名由数据集代码（SURF\_CLI\_CHN\_MUL\_DAY）、要素代码（XXX)、项目代码（XXXXX）、年份标识（YYYY）和月份标识（MM）组成。其中，SURF表示地面气象资料，CLI表示地面气候资料，CHN表示中国，MUL表示多要素，DAY表示日值数据。

文件命名为：

1. 本站气压：SURF\_CLI\_CHN\_MUL\_DAY-PRS-10004-YYYYMM.TXT

2. 气温：SURF\_CLI\_CHN\_MUL\_DAY-TEM-12001-YYYYMM.TXT

3. 相对湿度：SURF\_CLI\_CHN\_MUL\_DAY-RHU-13003-YYYYMM.TXT

4. 降水：SURF\_CLI\_CHN\_MUL\_DAY-PRE-13011-YYYYMM.TXT

5. 蒸发：SURF\_CLI\_CHN\_MUL\_DAY-EVP-13240-YYYYMM.TXT

6. 风向风速：SURF\_CLI\_CHN\_MUL\_DAY-WIN-11002-YYYYMM.TXT

7. 日照：SURF\_CLI\_CHN\_MUL\_DAY-SSD-14032-YYYYMM.TXT

8. 0cm地温：SURF\_CLI\_CHN\_MUL\_DAY-GST-12030-0cm-YYYYMM.TXT

示例：SURF\_CLI\_CHN\_MUL\_DAY-TEM-12001-201812.txt

特征值说明：

台站海拔高度 +100000，当台站海拔高度为估测值时，在估测数据基础上加100000

各要素项 32766，数据缺测或无观测任务

气压日极值 +20000，气压极值取自定时值，在原值上加20000

日最小相对湿度 +300，最小相对湿度取自定时值，在原值上加300

风速 +1000，当风速超过仪器测量上限时，在上限数据基础上加1000

风向 1-17，用数字表示风向方位，17表示静风

+100，当表示风向为八风向时，在原值上加100

90X，风向出现X个时，风向数据用个数X表示

95X，风向至少出现X个时，风向数据用个数X表示

降水量 32700，表示降水"微量"

32XXX，XXX为纯雾露霜

31XXX，XXX为雨和雪的总量

30XXX，XXX为雪量(仅包括雨夹雪，雪暴）

蒸发量 32700，表示蒸发器结冰

+1000，蒸发器中注入的水全部蒸发，在注入的水量数据基础上加1000

0cm地温 +10000，实际温度（零上）超仪器上限刻度，在上限数据基础上加10000

-10000，实际温度（零下）超仪器下限刻度，在下限数据基础上减10000

2. 数据提取

了解了数据集的存储格式，就可以根据自己的需求提取相应的数据。方便使用和处理的气象数据存储格式通常以数据类型为字段（列）、以日期或时间为记录（行）。

示例

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| DATE | TEM | PRE | RHU | PRS |
| 2018/01/01 | xx | xx | xx | xx |
| 2018/01/02 | xx | xx | xx | xx |
| 2018/01/03 | xx | xx | xx | xx |
|  |  |  |  |  |

数据提取Python代码

xxx

本示例代码对特征值处理较为简单，可根据特征值说明处理数据特征值，以便满足具体应用需要。