

**中国国家级地面气象站基本气象要素  
日值数据集（V3.0）  
评估报告**

国家气象信息中心

2012 年 8 月

# 目 录

1 数据集概要说明.....	1
2 数据集制作过程与数据处理规定 .....	2
2.1 数据集制作过程 .....	2
2.2 数据处理规定 .....	4
2.2.1 日平均值统计方法 .....	4
2.2.2 特殊数据表征规定 .....	5
2.3 质量控制 .....	6
2.3.1 地面基础气象资料建设——基础数据质量检测与控制 .....	6
2.3.2 数据集质量控制.....	8
3 数据集质量评估.....	9
3.1 数据质量评估指标.....	9
3.2 数据集质量评估结果 .....	10
3.2.1 1951~2010 年数据质量评估结果.....	11
3.2.2 2011 年 1 月~2012 年 5 月数据质量评估结果 .....	15
3.2.3 2012 年 6~7 月数据质量评估结果 .....	16
4 较“中国高密度台站地面基本气象要素日值数据集（V2.0）”的改进.....	17
5 小结 .....	21

# 中国国家级地面气象站基本气象要素日值数据集（V3.0）

## 评估报告

至今，国家气象信息中心已制作并提供服务的我国地面气象站基本气象要素日值数据产品有两个版本。其一为“中国地面气候资料日值数据集”，该数据集包含中国基本、基准气象站 1951 年以来的地面基本气象要素逐日观测数据；其二为 2008 年 4 月发布的“中国高密度台站地面基本气象要素日值数据集（V2.0）”，该数据集包含了中国基本、基准气象站、一般气象站和农垦、森工、生产建设兵团等行业站在内的 2466 个站点 1951~2007 年地面基本气象要素逐日观测数据，其中中国基本、基准气象站、一般气象站共计 2425 个。

本次制作发布的“中国国家级地面气象站基本气象要素日值数据集（V3.0）”包含了中国基本、基准气象站、一般气象站在内的主要 2474 个站点 1951 年 1 月以来地面基本气象要素逐日观测数据以及相应时期的台站信息（台站经纬度、海拔高度）。该数据集相对于上述已发布的两类数据产品，各要素观测数据本身无论从数据完整性还是质量上均有明显提高，并且在数据文件中随要素数据给出了各个时期对应的台站信息。尤其 2474 个站 1951~2010 年的数据质量提高非常明显，该段数据均来源于 2011-2012 年中国气象局开展的“地面基础气象资料建设”项目的工作成果，并且在数据集制作过程中经过了较严格的质量控制。在地面基础气象资料建设以及数据集制作过程中，对发现的 2474 个站 1951~2010 年可疑和错误数据普遍给与了人工核查与更正，并对因数字化遗漏造成的类似“数据缺测”现象进行了补录。

## 1 数据集概要说明

“中国国家级地面气象站基本气象要素日值数据集（V3.0）”包含了中国基本、基准气象站、一般气象站在内的主要 2474 个站点 1951 年 1 月以来地面基本气象要素逐日观测数据。

### （1）数据来源

数据集中 1951~2010 年数据，基于地面基础气象资料建设项目归档的

“1951~2010 年中国 2474 个国家级地面站数据更正后的月报数据文件（A0/A1/A）基础资料集”研制。2011 年 1 月~2012 年 5 月数据基于各省上报到国家气象信息中心的地面月报数据文件（A 文件）研制。2012 年 6~7 月数据基于国家气象信息中心实时库数据研制。实时库中该部分数据来自实时上传的地面自动站逐小时数据文件（Z 文件）及日值数据文件。

## （2）数据集要素

本次制作的数据集涉及的观测要素项如下，其它要素项数据将后续制作给出：

- ① 日平均本站气压、日最高本站气压、日最低本站气压；
- ② 日平均气温、日最高气温、日最低气温；
- ③ 日平均相对湿度、日最小相对湿度；
- ④ 日平均风速、日最大风速、最大风速的风向、日极大风速、极大风速的风向；
- ⑤ 日平均 0cm 地温、日最高 0cm 地温、日最低 0cm 地温；
- ⑥ 20-8 时降水量、8-20 时降水量、20-20 时降水量；
- ⑦ 日小型蒸发量、日大型蒸发量；
- ⑧ 日累计日照时数。

资料时间范围：1951 年 1 月~最新。

资料空间范围：以中国 31 个省（区、市）内 2474 个国家级地面观测站为主。

# 2 数据集制作过程与数据处理规定

## 2.1 数据集制作过程

### （1）确定数据集内容和文件组织方式

首先根据气象要素分类与编码（QX/T 133-2011），确定数据集代码名称，数据文件命名方式，文件存放目录结构，各要素的排列方式和数据类型、单位（具体见该数据集格式说明文件：SURF\_CLI\_CHN\_MUL\_DAY\_FORMAT.doc）。规定了数据质量控制码（详见 2.3.2）。在对数据进行分析后，设置了特征值（详见 2.2.2）用以表示数据集中的特殊数据。

## **(2) 数据提取与统计**

基于 2012 年基础气象资料建设工作组归档的“1951~2010 年中国 2474 个国家级地面站数据更正后的月报数据文件（A0/A1/A）基础资料集”研制得到 1951~2010 年 2474 站本站气压、气温、相对湿度、风向风速、0cm 地温逐日 4 次定时观测数据与日极值数据、以及累计降水量、日蒸发量和累计日照时数、逐月台站信息。基于省级逐月上报到国家气象信息中心的地面月报数据文件（A 文件）研制得到 2011 年 1 月~2012 年 5 月上述要素项观测数据及台站信息。基于国家气象信息中心实时库数据（Z 文件），研制得到 2012 年 6~7 月相应要素项数据及台站信息。

按照 2.2.1 节规定的统计方法由各要素逐日定时观测数据统计得到日平均值，包括日平均本站气压、日平均气温、日平均相对湿度、日平均风速、日平均 0cm 地温。

## **(3) 数据质量控制**

中国国家级地面气象站基本气象要素日值数据集（V3.0），主要基于基础气象资料建设工作组归档的“1951~2010 年中国 2474 个国家级地面站数据更正后的月报数据文件（A0/A1/A）基础资料集”制作而成。在 2011 年 3 月~2012 年 6 月中国气象局开展的地面基础气象资料建设工作中，对 1951~2010 年 2474 个国家站地面月报数据文件中的观测数据进行了反复质量检测与控制，期间纠正了大量的错误数据，并对数字化遗漏数据进行了补录，使得数据质量得到明显提升。另外在数据集制作过程中，对数据集中 1951~2010 年的要素数据进行了质量控制，对发现的可疑和错误数据普遍给与了人工核查与更正，并最终对所有要素数据标注质量控制码。

各省上报至国家气象信息中心的 2011 年 1 月~2012 年 5 月地面月报数据文件，在日常资料处理业务中经过了严格的“台站-省级-国家级”三级质量控制。数据集中 2011 年 1 月~2012 年 5 月所有日数据由该期间的月报数据文件提取及统计得到。同时相应观测数据质量控制码取自月报数据文件中的省级质控码，日平均统计数据的质量控制码为月报文件中相应日定时观测数据的最大质控码。

实时上传的地面自动站观测数据经过了台站质量控制，国家气象信息中心将实时上传的地面自动站逐小时数据文件和日值观测数据文件中的观测数据及台

站质量控制码入到实时库中用于数据服务。本数据集中 2012 年 6~7 月相应要素观测数据及质量控制码取自实时库,日平均统计数据的质量控制码为实时库中相应日定时观测数据的最大质控码。

#### **(4) 生成数据集**

按照数据格式说明文件“SURF\_CLI\_CHN\_MUL\_DAY\_FORMAT.doc”规定的数据集内容、文件组织方式以及各文件中数据格式编制程序,生成数据集各要素数据文件实体。

数据集中全部数据都放在实体文件目录下,一个月一个要素项目所有站日观测数据及相应的数据质量控制码为一个文件,文件命名为“SURF\_CLI\_CHN\_MUL\_DAY-XXX-XXXXX-YYYYMM.TXT”,其中“XXX”为要素代码,“XXXXX”为项目代码,“YYYY”为年份,“MM”为月份。

#### **(5) 统计各要素的实有率和质量情况**

以实有率统计各要素数据的完整性,根据数据质量控制结果统计各要素数据的质量状况。

#### **(6) 撰写相关文档**

撰写“数据集质量评估报告”、“数据集说明文档”、“元数据文档”、“台站信息列表”。

## **2.2 数据处理规定**

### **2.2.1 日平均值统计方法**

由上文中地面气象月报数据文件或实时库中提取得到的各要素逐日 4 次定时(02:00、08:00、14:00、20:00)观测数据统计日平均值,统计规定如下:

① 日平均气温、相对湿度、0cm 地温均为 4 次定时观测值的平均值;

② 一般情况下,日平均本站气压、日平均风速均由 4 次定时观测值平均得到,但当三次人工观测站在未配自记仪器的情况下,由 08:00、14:00、20:00 定时观测值平均得到;

③ 按上述①②规定统计日平均值时,当参加统计的某定时值缺测,则相应要素日平均值缺测处理。

## 2.2.2 特殊数据表征规定

数据集中使用的一些具有特定含义的数据，来表示数据的特殊情况，称为特征值。其规定说明见表 1。

表 1 特征值说明

数据项	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

### （1）台站海拔高度为估测值的数据表征

早期，许多台站海拔高度值非仪器实测，属估测，估测值精度可能较低。为了区分实测值与估测值，规定当台站海拔高度为估测值时，在估测数据基础上加 100000。

### （2）无观测任务项及缺测数据表征

由地面月报文件研制得到的数据集中，当台站有值班观测任务，但某要素项按规定不观测或有观测值但无需数字化，或虽应观测但因仪器故障等原因数据缺测时，数据集中均用特征值“32766”表示。

### （3）日极值取之定时值的表征

当日最高（最低）气压或日最小相对湿度取之定时观测值时，气压极值在原值基础上加 20000，湿度极值在原值基础上加 300。

#### （4）风向风速数据表征

如表 1 所示，当风速超过仪器测量上限时，在上限数据基础上加 1000 表示；风向用 1-17 编码表示，编码与风向的对应关系见表 2。历史上人工测量风向普遍如表 2 所示为 16 方位，但早期部分站风向测量精度为 8 方位。为了区分上述两种测量，规定风向为 8 方位测量时，在风向编码值的基础上加 100。

当日最大风速（日极大风速）出现多次时，各次对应的风向可能不同，此时风向用个数表示。当风向明确出现 X 个时，风向数据用 90X 表示；当仅确定风向至少出现 X 个时，风向数据用 95X 表示。

表 2 风向方位的编码

风向	编码	风向	编码	风向	编码	风向	编码	风向	编码	风向	编码
N	1	ENE	4	SE	7	SSW	10	W	13	NNW	16
NNE	2	E	5	SSE	8	SW	11	WNW	14	C(静风)	17
NE	3	ESE	6	S	9	WSW	12	NW	15		

#### （5）降水量表征

当降水量为微量时，降水数据用 32700 表示；当测量对象为纯雾露霜时，降水数据用 32XXX（XXX 表示测量数据，下同）表示；当测量对象为雨和雪的汇合物时，降水数据用 31XXX 表示；当测量对象为雨夹雪或雪暴水当量时，降水数据用 30XXX 表示。

#### （6）蒸发量表征

当蒸发器结冰不进行当日观测时，蒸发数据用 32700 表示；当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。

#### （7）0cm 地温超仪器可观测范围时的表征

当实际地表温度（零上）超仪器上限刻度时，在上限刻度基础上加 10000 表示；当实际地表温度（零下）超仪器下限刻度时，在下限刻度基础上减 10000 表示。

### 2.3 质量控制

#### 2.3.1 地面基础气象资料建设——基础数据质量检测与控制



中国国家级地面气象站基本气象要素日值数据集（V3.0），主要基于基础气象资料建设工作组归档的“1951~2010 年中国 2474 个国家级地面站数据更正后的月报数据文件（A0/A1/A）基础资料集”制作而成。2011-2012 年在地面基础气象资料建设过程中，专门研制了《地面基础气象资料质量检测与控制工作流程》、《地面数字化基础气象数据质量检测方案》、《地面基础气象资料质量控制方案》及《地面基础气象数据质量核查与更正规范》，以对 1951~2010 年国家级地面台站观测数据进行质量检测与控制、数据核查与更正。通过该项建设工作基本解决了 1951~2010 年 2474 站 150 多万份月报数据文件中台站参数、要素方式位、要素观测值与实际观测不符等现象。台站参数、要素方式位、要素观测值的修正比例分别各占其被检总量的 1.10%、1.97%、2.72%，新增数据文件 11199 个站月。发现并纠正了如下 12 项具体问题：

表 3 地面基础气象资料建设工作中，各要素项正点数据、日极值、日累计值数据更正情况

要素	项目	数据修正量 (个)	数据修正率 (%) (数据修正量/被检数据总量)
本站气压	正点气压	2,460,057	0.90
	日最高气压	206,344	0.74
	日最低气压	216,024	0.77
气温	正点气温	2,567,373	0.89
	日最高气温	389,193	0.85
	日最低气温	391,044	0.85
相对湿度	正点相对湿度	2,655,702	0.94
	日最小相对湿度	243,676	0.83
风	正点风速	2,577,829	0.94
	正点风速的风向	3,158,007	1.15
	日最大风速	256,049	1.02
	日最大风速的风向	268,864	1.07
	日极大风速	74,234	1.01
	日极大风速的风向	89,187	1.21
0cm 地温	正点 0cm 地温	2,520,357	1.01
	日最高 0cm 地温	315,159	0.75
	日最低 0cm 地温	322,244	0.76
降水	20-08、08-20、20-20 时降水量	1,114,732	0.81
蒸发	日小型蒸发量	397,737	0.97
	日大型蒸发量、各时大型蒸发量	105,819	1.07
日照	日照总时数、各时日照时数	2,942,753	1.98

- ① 地面月报数据文件命名错误及文件中数据格式错误现象；
- ② 观测数据超出要素允许值范围的错误现象；

- ③ 要素数据之间或同要素日值与日极值矛盾现象；
- ④ 多份数据文件中同站同期数据不同现象（观测数据表现为非唯一性）；
- ⑤ 台站参数（经纬度、海拔高度、气压计高度等）错误现象；
- ⑥ 因数字化遗漏造成的要素整月或部分日无数据现象（疑似数据缺测）；
- ⑦ 要素方式位与实际观测数据不一致现象；
- ⑧ 日风向风速表征错误现象；
- ⑨ 数据文件中要素资料被其它资料错误替代现象；
- ⑩ 大、小蒸发器观测数据录入顺序错误现象；
- ⑪ 因传感器接反或数据记录错误导致的各层地温数据间异常现象；
- ⑫ 风向传感器疑似雷击造成的风向异常现象等。

地面月报数据文件中上述数据问题若不解决，均会影响由此制作的数据集质量，造成提取的数据错误或不完整。中国国家级地面气象站基本气象要素日值数据集（V3.0）中所包含的本站气压、气温、相对湿度、风、0cm地温、降水量、蒸发量、日照数据，通过地面基础气象资料建设工作，日照要素数据更正比例稍高，约为2%，其它要素数据更正量在1%左右，各要素具体更正情况见表3。

图1以为57401站1961-1963年5月小型蒸发量数据修正情况为例，展现了修正前后5月份月蒸发量逐年时间变化曲线。由图可见，修正前后蒸发量气候变化趋势截然相反。

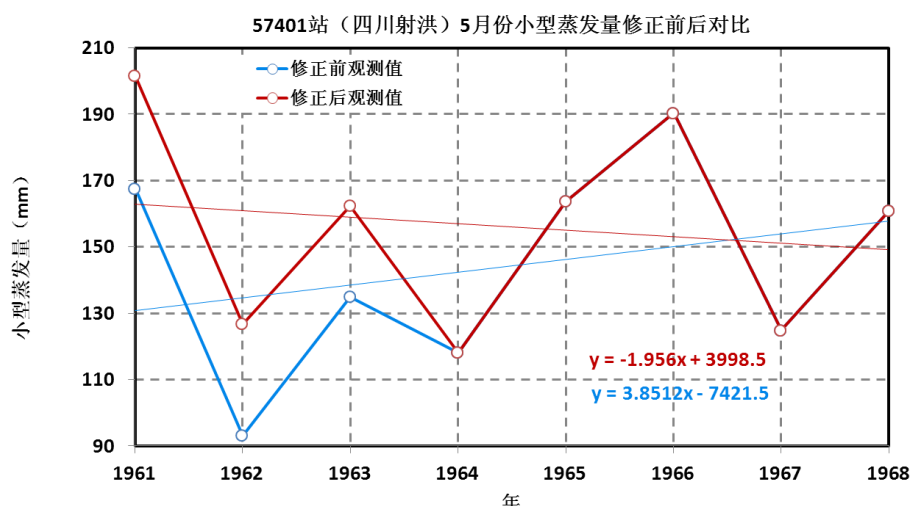


图1 57401站1961~1963年5月小型蒸发量修正前后对比

### 2.3.2 数据集质量控制

在本次数据集制作过程中，对所有要素观测数据均给出了质量信息——质量控制码。质量控制码的规定及含义见表 4。

表 4 数据质量控制码

质量控制码	含义
0	数据正确
1	数据可疑
2	数据错误
8	数据缺测或无观测任务
9	数据未进行质量控制

数据集中 2011 年 1 月~2012 年 5 月所有要素项目极值数据、累计量数据的质量控制码为同期地面月报数据文件中的省级质控码，由地面气象月报数据文件三级质量控制业务系统中的省级质量控制业务给出；日平均值数据的质量控制码取自月报文件中相应定时数据的最大质控码。2012 年 6~7 月日极值数据及累计量数据的质量控制码为实时上传的自动站小时数据文件、日数据文件中的台站质控结果，由台站质控软件自动标注给出；日平均值数据的质量控制码取自实时库中相应定时数据的最大质控码。1951~2010 年所有要素项目数据质量控制码为数据集产品制作过程中，进行如下质量控制后的结果：

- ① 气候界限值或允许值检查；
- ② 台站极值检查；
- ③ 定时值、日平均值与日极值间内部一致性检查；
- ④ 时间一致性检查；
- ⑤ 空间一致性检查；
- ⑥ 人工核查与更正。

在对 1951~2010 年数据集产品制作过程中，对上述①~⑤自动检测出的疑误数据普遍进行了人工核查，对核查后明确为错误的数据进行了纠正，明确为正确的数据其质量控制码改为“0”，无法明确质量信息的数据保留自动检测结果。

## 3 数据集质量评估

### 3.1 数据质量评估指标

评估报告中用实有率评估数据集各要素项的完整性，用正确率评估数据集各

要素项的质量状况。统计量的计算方法如下：

$$\text{实有率} = \frac{\sum_{i=1}^N \text{实有观测数据量}_i}{\sum_{i=1}^N \text{数据总量}_i} \times 100\% \quad (1)$$

$$\text{正确率} = \frac{\sum_{i=1}^N \text{正确数据量}_i}{\sum_{i=1}^N \text{实有观测数据量}_i} \times 100\% \quad (2)$$

(1)、(2) 式中，N 为总台站数；i 表示第 i 个站，i=1,2,3,...N。各统计参数的意义如下：

**实有观测数据量<sub>i</sub>**：数据集中第 i 个站某要素项非“32766”的数据量；

**正确数据量<sub>i</sub>**：数据集中第 i 个站某要素项数据质量控制码为“0”的数据量；

**数据总量<sub>i</sub>**：数据集中第 i 个站某要素项**缺测数据量<sub>i</sub>**与**实有观测数据量<sub>i</sub>**之和。

**缺测数据量<sub>i</sub>**：指按规定应观测但因仪器故障等原因造成的数据集中第 i 个站某要素项无观测数据或无统计数据（数据缺测）的总量，在数据集中缺测数据用“32766”表示。

由上文可知，特征值“32766”在数据集中不仅仅表示数据缺测，有时还表示无观测任务。数据集中**无观测任务**定义如下：当台站有值班观测任务，但某要素项按规定不观测，或虽然进行了观测但因观测仪器精度较低，按规定观测数据无需数字化。

由于月报文件中缺测数据和无观测任务数据表达方式完全相同，由地面月报文件研制得到的日值数据集中，整月缺测数据和无观测任务数据无法明确区分，因此无观测任务数据也同样用特征值“32766”表示。在数据集评估过程中，无观测任务数据不计入**缺测数据量<sub>i</sub>**。规定下列情况均表示无观测任务：

- ① 某要素项从开始有观测数据计算，之前均表示该要素项无观测任务；
- ② 某要素项连续12个月及以上无观测数据，视该要素相应月份无观测任务；
- ③ 当三次人工观测站无自记仪器时，02:00 时本站气压、风向风速视为无观测任务；
- ④ 由于大、小型蒸发非全年观测，规定若某观测项连续 4 个月及以上无数据，视为无观测任务。

### 3.2 数据集质量评估结果

由于数据集中 1951 年 1 月~2010 年 12 月、2011 年 1 月~2012 年 5 月、2012 年 6~7 月三段数据的来源和质量控制渠道不同，下文分别对各段数据进行评估。

#### 3.2.1 1951~2010 年数据质量评估结果

##### (1) 观测站点数变化及站点分布

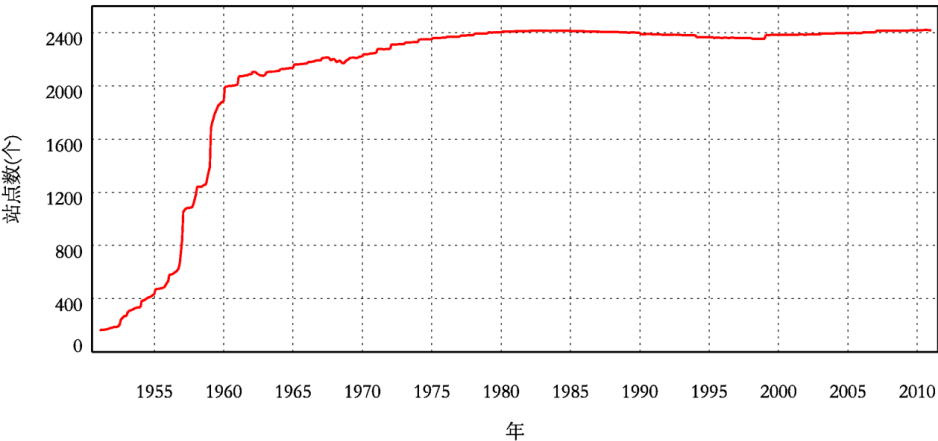


图 2 1951~2010 年国家级站点数逐月变化曲线

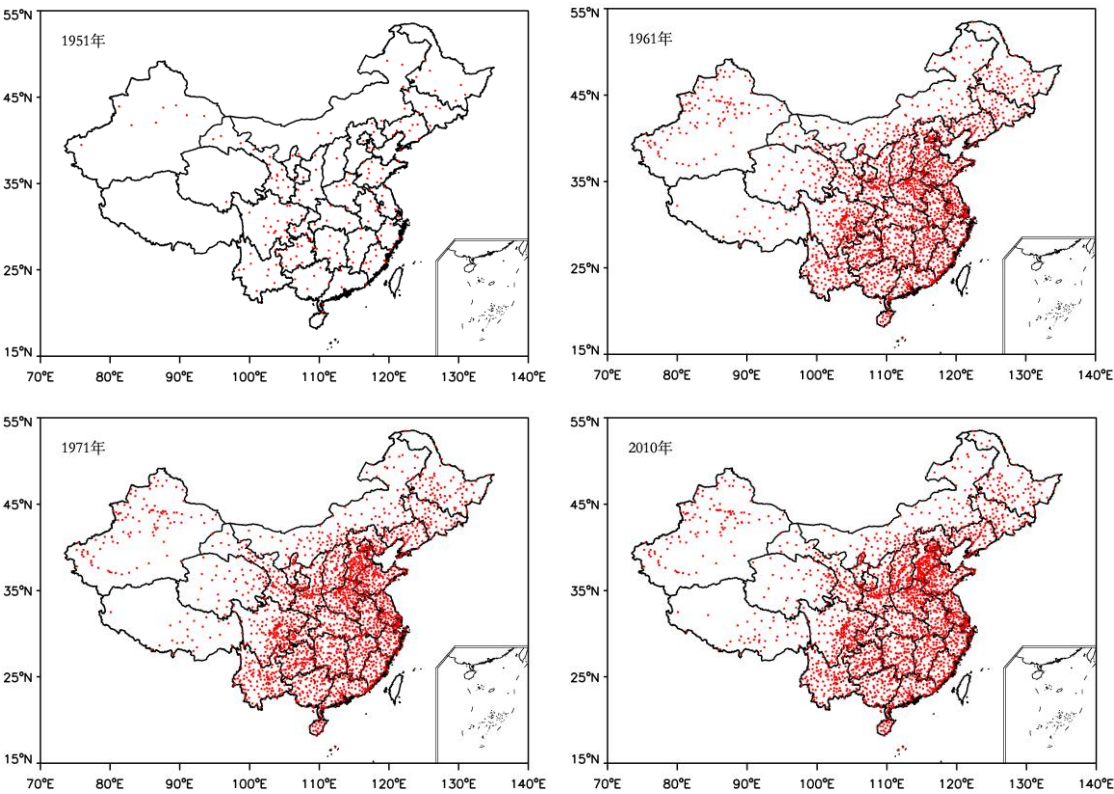


图 3 国家级地面气象站分布

根据地面月报数据文件统计了历年逐月国家级地面台站量,台站量随时间变化曲线见图 2。图 3 分别展示了 1951、1961、1971、2010 年观测站点空间分布。1951 年有观测台站 182 个,我国青藏高原地区站点数极少,一些边境地区无观测站。1955 年之前,国家级台站数不足 500 个,此后几年站点迅速增加,1961 年增至 2088 站。1971 年进一步增加至 2279 站,1975 年之后站点数相对比较稳定,维持在 2400 站左右,2010 年为 2421 站。

下文图 4 中给出了 1951~2010 年各要素实有测站数逐月变化情况。从图中可以看出,气温、相对湿度、风 3 种要素的实有测站数变化曲线与图 2 一致。本站气压 1974 年前测站明显较气温等要素测站少; 0cm 地温 1980 年前测站明显较气温等要素测站少,尤其 1967~1979 年期间。气压、0cm 地温上述时间段测站相对于气温等要素少,主要是由于按规定不观测或观测仪器测量精度较低按规定无需数字化引起的。1980 年后各要素实有测站数均与图 2 一致,稳定在 2400 个台站左右。

## (2) 数据完整性

表 5 1951~2010 年 2474 站各要素项数据完整性状况

要素	项目	总数据量 (个)	实有率 (%)
本站气压	日平均气压	43,090,396	99.84
	日最高气压	27,925,920	99.95
	日最低气压	27,925,920	99.95
气温	日平均气温	45,827,716	99.96
	日最高气温	45,818,708	99.95
	日最低气温	45,823,292	99.95
相对湿度	日平均相对湿度	45,733,500	99.96
	日最小相对湿度	29,399,316	99.74
风	日平均风速	45,688,780	99.73
	日最大风	50,306,136	98.87
	日极大风	14,704,358	98.17
0cm 地温	日平均地温	42,358,968	99.31
	日最高地温	41,842,256	99.32
	日最低地温	42,374,272	99.32
降水	20-08 降水量	45,831,812	99.85
	08-20 降水量	45,831,812	99.85
	20-20 降水量	45,831,812	99.97
蒸发	小型蒸发量	41,024,488	99.37
	大型蒸发量	3,190,325	98.16
日照	日照时数	45,216,864	99.79

表 5 给出了数据集中共 20 项要素数据完整性状况。20 项数据中，除了风速极值项数据、大型蒸发量数据实有率在 98%~99%之间外，其它要素项数据的实有率均在 99%以上。本站气压、气温、相对湿度、降水、日照要素项数据的实有率均在 99.7%以上，接近 100%。从图 4 各要素日数据缺测率逐年变化可以看出，各要素缺测数据主要集中在 1980 年之前，60 年之前数据缺测率相对较高。由图 2 可见 1980 年之前，尤其 60 年前地面气象站属于建设期，建站初期难免数据缺测。由于该期间测站较少，虽然缺测率相对较高，但对整个数据集的完整性影响不大。

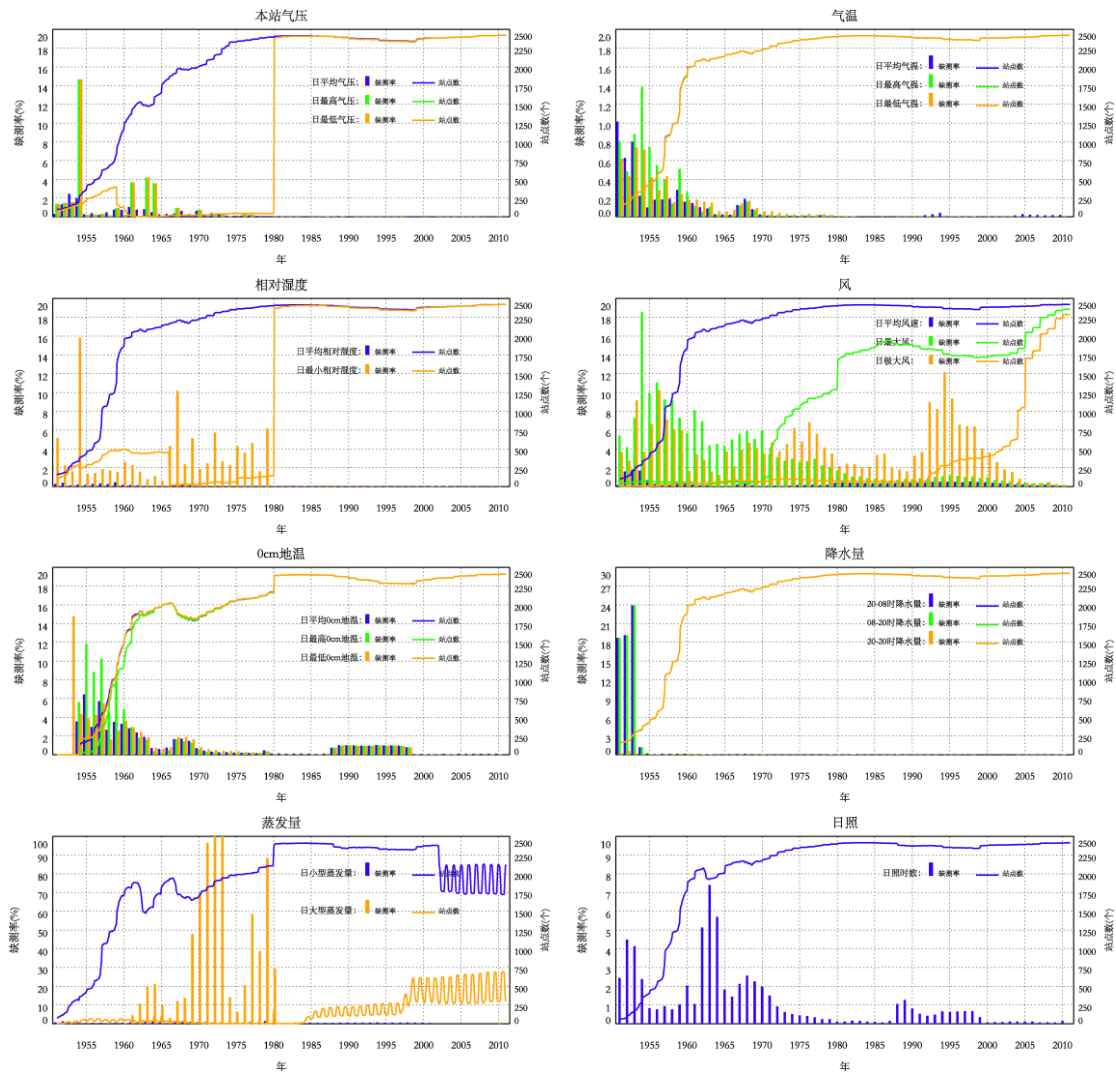


图 4 各要素项数据逐年缺测率和实有观测站点数逐月变化

### (3) 数据质量

由 2.3 节介绍可知，地面基础气象资料建设工作通过数据质量检测与控制、人工核查与更正，对检测出的 2474 个站 1951~2010 年台站信息化观测数据中的

大量错误数据进行了更正，数据质量得到明显提高。另外，在对本数据集 1951~2010 年数据质量控制过程中，同样采用了软件自动质量控制与人工核查更正的方式，即对软件自动检测出的疑误数据普遍进行了人工核查，对核查后明确为错误的的数据进行了纠正，明确为正确的数据其质量控制码改为“0”，无法明确质量信息的数据保留自动检测结果。

表 6 各要素项数据质量状况

要素	项目	正确率 (%)	可疑率 (%)	错误率 (%)
本站气压	日平均气压	99.916	0.084	0.000
	日最高气压	99.979	0.021	0.000
	日最低气压	99.979	0.021	0.000
气温	日平均气温	100.000	0.000	0.000
	日最高气温	99.999	0.001	0.000
	日最低气温	100.000	0.000	0.000
相对湿度	日平均相对湿度	100.000	0.000	0.000
	日最小相对湿度	99.924	0.076	0.000
风	日平均风速	100.000	0.000	0.000
	日最大风	100.000	0.000	0.000
	日极大风	99.999	0.001	0.000
0cm 地温	日平均地温	100.000	0.000	0.000
	日最高地温	100.000	0.000	0.000
	日最低地温	100.000	0.000	0.000
降水	20-08 降水量	100.000	0.000	0.000
	08-20 降水量	100.000	0.000	0.000
	20-20 降水量	100.000	0.000	0.000
蒸发	小型蒸发量	100.000	0.000	0.000
	大型蒸发量	99.999	0.000	0.001
日照	日照时数	100.000	0.000	0.000

通过上述数据处理，本次制作的 1951~2010 年数据集质量较高。从表 6 可见，各要素项数据正确率均超过了 99.9%，接近 100%。除本站气压和最小相对湿度数据可疑率相对略高外，其它要素项数据的可疑率和所有要素数据的错误率均在  $10^{-5}$  以下。

本站气压日值数据可疑主要是由于气压值与台站海拔高度不匹配造成的。通过人工核查，发现气压可疑数据普遍为正确数据，由于台站海拔高度错误造成两者不匹配。台站海拔高度错误，主要是由于早期高度值多为估测值不准确造成的，这也是图 5 中显示的 1965 年之前气压数据可疑率较高的主要原因。由于数据集



中仅对观测要素数据标注质量控制码，台站海拔高度未给出质量信息，为了引起用户注意台站海拔高度问题，当两者不匹配时，虽然气压数据正确，仍保留其可疑质量信息。

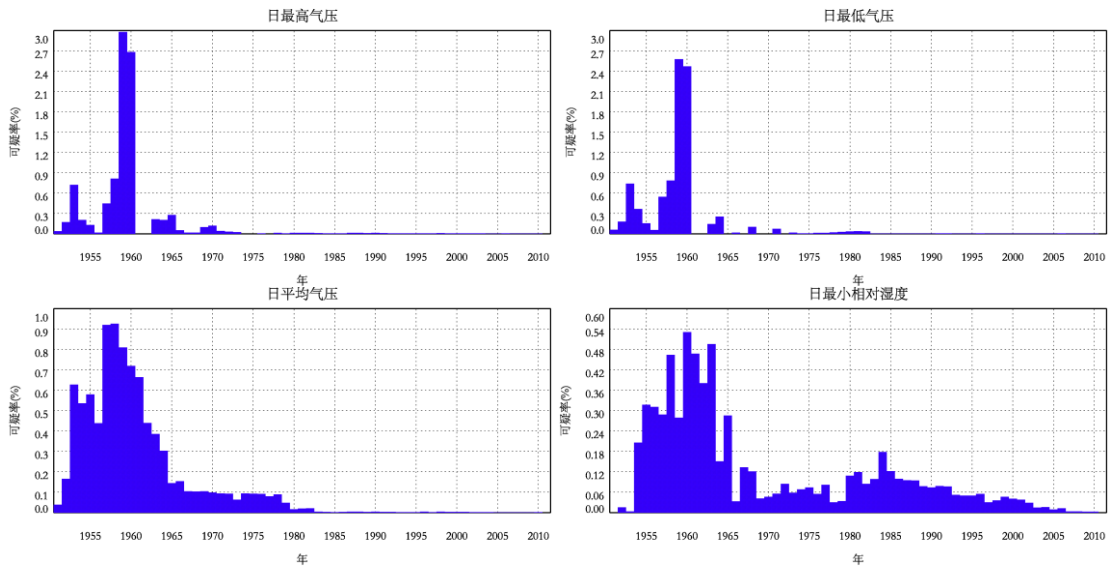


图 5 本站气压、相对湿度日数据可疑率历年变化图

相对湿度可疑数据均为 0%。考虑到大气中含水量非 0，因此在数据集质量控制过程中，规定大气相对湿度不允许等于 0%。通过人工核查，发现 0%的湿度值若考虑仪器允许的测量误差（如 $\pm 5\%$ ），那么湿度可疑数据普遍可视为正确数据。相对湿度数据可疑率逐年变化来看（见图 5），可疑数据主要出现在 1965 年以前。

3.2.2 2011 年 1 月~2012 年 5 月数据质量评估结果

2011 年 1 月~2012 年 5 月数据由省级日常业务进行了质量控制。从表 7 可以看出,2011 年 1 月~2012 年 5 月各要素项数据的完整性较高,普遍超过了 99.9%，参与质量控制的数据中可疑数据所占比例低于 4%，数据质量较好。

表 7 2011 年 1 月~2012 年 5 月各要素项数据完整性和质量状况

要素	观测项目	应观测数据总量 (个)	实有率 (%)	正确率 (%)	可疑率 (%)	未质控比率 (%)
本站气压	日平均气压	1,253,390	99.98	98.418	0.065	1.517
	日最高气压	1,253,390	100.00	98.458	0.019	1.523
	日最低气压	1,253,390	100.00	98.424	0.052	1.524
气温	日平均气温	1,253,390	99.99	98.425	0.056	1.519
	日最高气温	1,253,390	100.00	98.423	0.058	1.519
	日最低气温	1,253,390	100.00	98.333	0.149	1.518

相对湿度	日平均相对湿度	1,253,390	99.98	98.422	0.057	1.521
	日最小相对湿度	1,253,390	100.00	98.460	0.014	1.526
风	日平均风速	1,253,390	99.92	98.439	0.004	1.557
	日最大风	2,506,780	99.92	98.449	0.000	1.551
	日极大风	2,506,476	99.86	98.440	0.001	1.559
0cm 地温	日平均地温	1,247,703	99.94	98.213	0.223	1.564
	日最高地温	1,247,703	99.99	98.123	0.313	1.564
	日最低地温	1,247,703	99.99	98.086	0.358	1.556
降水	20-08 降水量	1,253,182	99.98	95.881	0.007	4.112
	08-20 降水量	1,253,182	99.98	95.836	0.007	4.157
	20-20 降水量	1,253,182	99.98	95.731	0.112	4.157
蒸发	小型蒸发量	1,005,950	99.68	98.323	0.024	1.653
	大型蒸发量	242,558	99.58	98.714	0.027	1.259
日照	日照时数	1,248,737	99.98	0.174	0.000	99.826

### 3.2.3 2012 年 6~7 月数据质量评估结果

2012 年 6~7 月数据及质量控制码取自国家级实时库，因 2012 年 6 月各要素项数据质控码为 9，表示未参与质量控制，因此仅统计 2012 年 6~7 月各要素项数据完整性。从表 8 可以看出，除了蒸发数据实有率在 90%左右外，其它各要素项数据的完整性较高，均超过了 99.6%。

表 8 2012 年 6~7 月各要素项数据完整性状况

要素	观测项目	应观测数据总量 (个)	实有率 (%)
本站气压	日平均气压	147,803	99.78
	日最高气压	147,742	99.84
	日最低气压	147,742	100.00
气温	日平均气温	147,803	99.79
	日最高气温	147,742	99.84
	日最低气温	147,742	100.00
相对湿度	日平均相对湿度	147,742	99.69
	日最小相对湿度	147,742	99.99
风	日平均风速	147,803	99.77
	日最大风	295,484	99.84
	日极大风	295,484	99.84
0cm 地温	日平均地温	146,888	99.65
	日最高地温	147,132	99.84
	日最低地温	147,132	100.00
降水	20-08 降水量	147,864	99.88
	08-20 降水量	147,803	99.89

	20-20 降水量	147,803	99.90
蒸发	小型蒸发量	102,279	90.24
	大型蒸发量	40,500	90.24
日照	日照时数	143,107	100.00

## 4 较“中国高密度台站地面基本气象要素日值数据集(V2.0)”的改进

“中国国家级地面气象站基本气象要素日值数据集(V3.0)”(简称“数据集 2”，下同) 包含 1951~2010 年 2474 个国家级台站本站气压、气温、相对湿度、风向风速、0cm 地温、降水、蒸发、日照日观测与统计数据，并随各要素项数据给出了相应观测月份的台站信息。2008 年 4 月发布的“中国高密度台站地面基本气象要素日值数据集(V2.0)”(简称“数据集 1”，下同) 包含 1951~2007 年 2425 个国家级台站日观测数据，仅一次性给出各站最新的台站信息，且无 0cm 地温要素数据。本评估报告仅对上述 2 个版本数据集中共同包含的 2425 站本站气压、气温、相对湿度、风向风速、蒸发、日照要素项数据进行对比，对比时间段为 1951~2007 年。

表 9 各要素项数据相对修正情况

要素	项目	数据修正量(单位: 个)				修正比例 (%)
		整月新增	整月剔除	其它修正	修正总量	
本站气压	日平均气压	120705	55005	231177	406887	0.95
	日最高气压	39420	33101	226797	299318	1.07
	日最低气压	39420	33101	231424	303945	1.09
气温	日平均气温	146788	105591	217355	469734	1.03
	日最高气温	149224	105392	216748	471364	1.03
	日最低气温	148883	105247	216797	470927	1.03
相对湿度	日平均相对湿度	146213	105388	235735	487336	1.07
	日最小相对湿度	48409	6918291	246797	7213497	24.60
风	日平均风速	392617	105070	251542	749229	1.64
	日最大风	196290	103161	322371	621822	1.25
	日极大风	336123	80757	140055	556935	3.86
降水	20-8 时降水量	146341	152558	717976	1016875	2.22
	8-20 时降水量	146382	152560	351703	650645	1.42
	20-20 时降水量	149119	152015	954819	1255953	2.74
蒸发	小型蒸发量	115091	73379	17921	206391	0.50
	大型蒸发量	8775	20569	24603	53947	1.63
日照	日照时数	182564	52790	30279	265633	0.59

表 9 给出了数据集 2 相对于数据集 1 中 7 种要素 17 类统计项目数据的修正情况，17 类项目共计修正数据量 15,500,438 个。从表中可见，大多数项目的数据修正量在 0.5%-4.0% 之间。但数据集 2 中日最小相对湿度相对于数据集 1“剔除”数据较多，剔除量达到 23.6%。经过核查，有 225,729 个站月最小相对湿度实际无观测数据，而制作数据集 1 时，用 3 次或 4 次定时值中的最小值进行日最小相对湿度数据的替代处理了。数据集 2 从尊重观测事实的角度出发，不再进行上述替代处理，这也是表 9 中日最小相对湿度数据剔除量以及修正比例比其它项目超多的原因。

数据集 2 相对于数据集 1，各要素项均不同程度的新增加了部分站月数据，同时也剔除了不恰当的整月数据。整月新增及剔除的日值数据量见表 9。要素数据整月被剔除，主要原因包含以下几个方面：①该站月资料是由其它站月资料替代生成，无法纠正，相应月报数据文件中的要素项数据被剔除；②某段时期内该站的测站级别为省级辅助站，而非国家站，资料质量整体欠佳，不纳入到国家站资料中；③某要素数据整月错误，无法纠正，相应要素数据剔除；④极少部分因为该站月本来就无相应观测数据，数据集 1 中的数据来源不明。表 9 中其它修正包含下列 3 个方面：①数据集 1 数据错误，数据集 2 更正为正确数据；②数据集 1 中某月部分日数据错误，数据集 2 对其缺测处理；③数据集 1 中某月部分日无数据，数据集 2 进行了补录。

从各要素项修正量空间分布来看（见图 6），数据集 2 相对于数据集 1 中各要素修正数据主要集中在东北、华北、东南沿海、以及西南等地区。从全国分布来看，绝大部分台站（超过总数的 95%）的日降水观测值均有不同程度修正，这与表 9 显示的降水数据修正量较多一致。

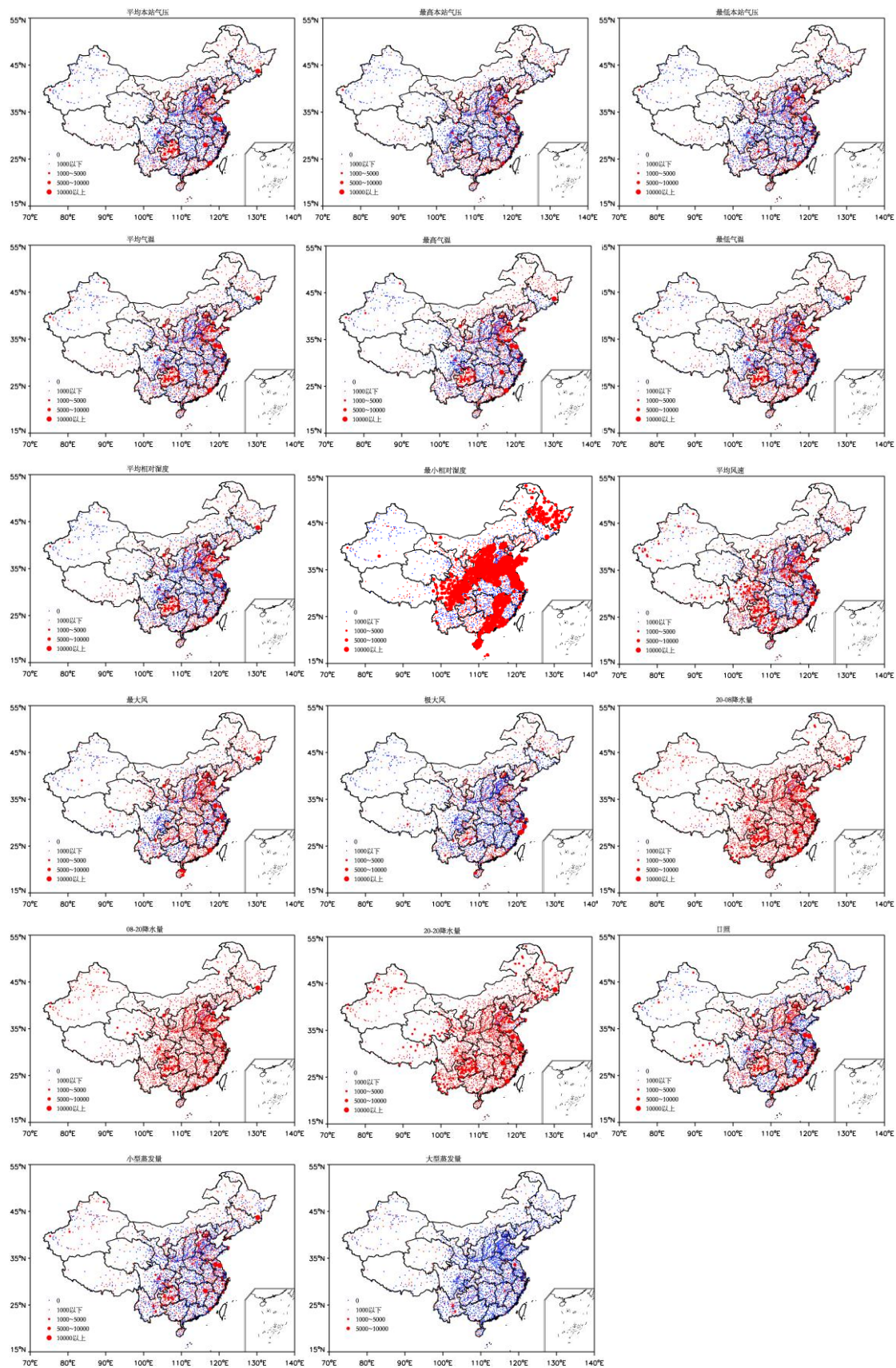
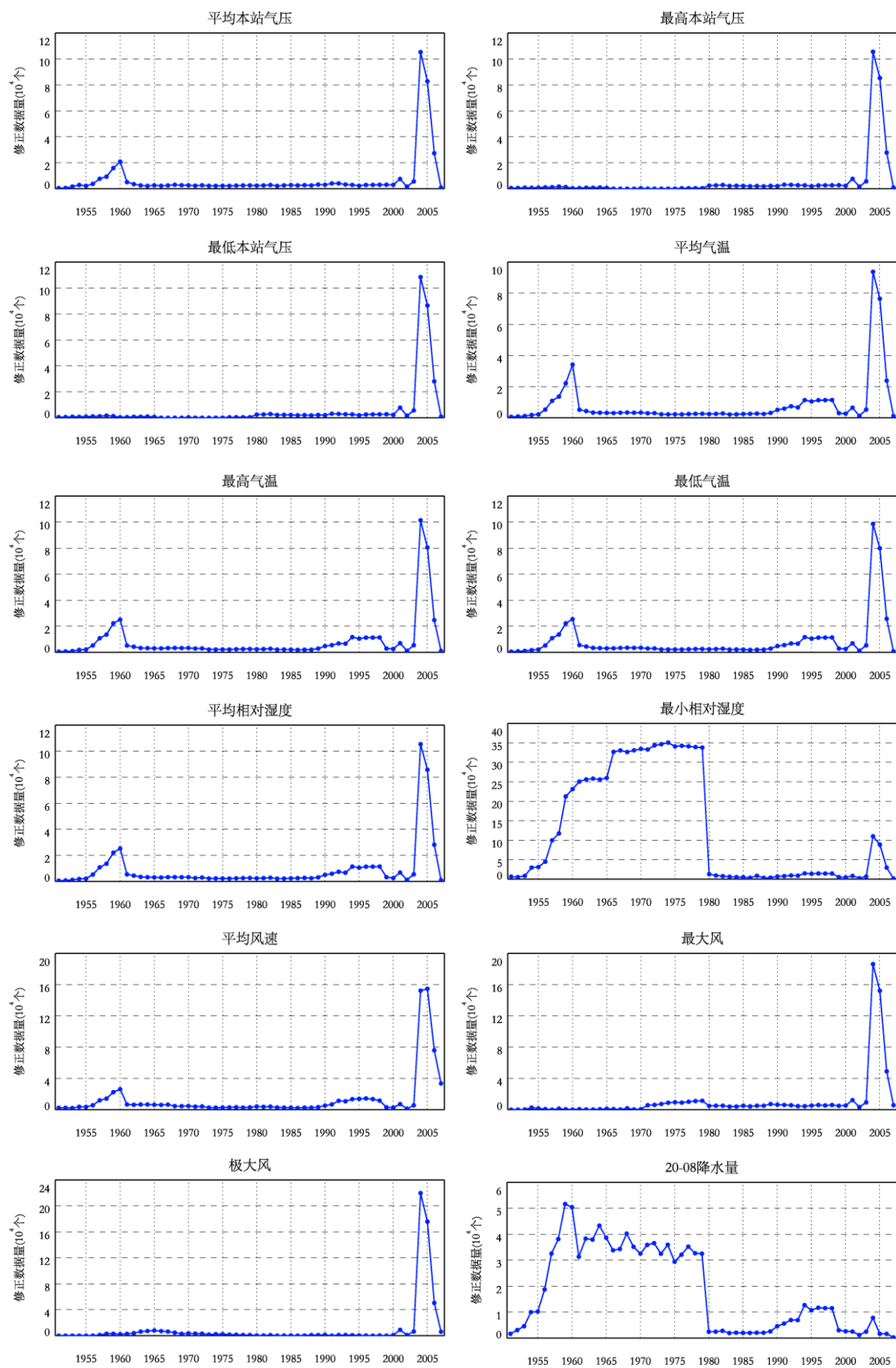


图 6 各要素项数据修正量空间分布





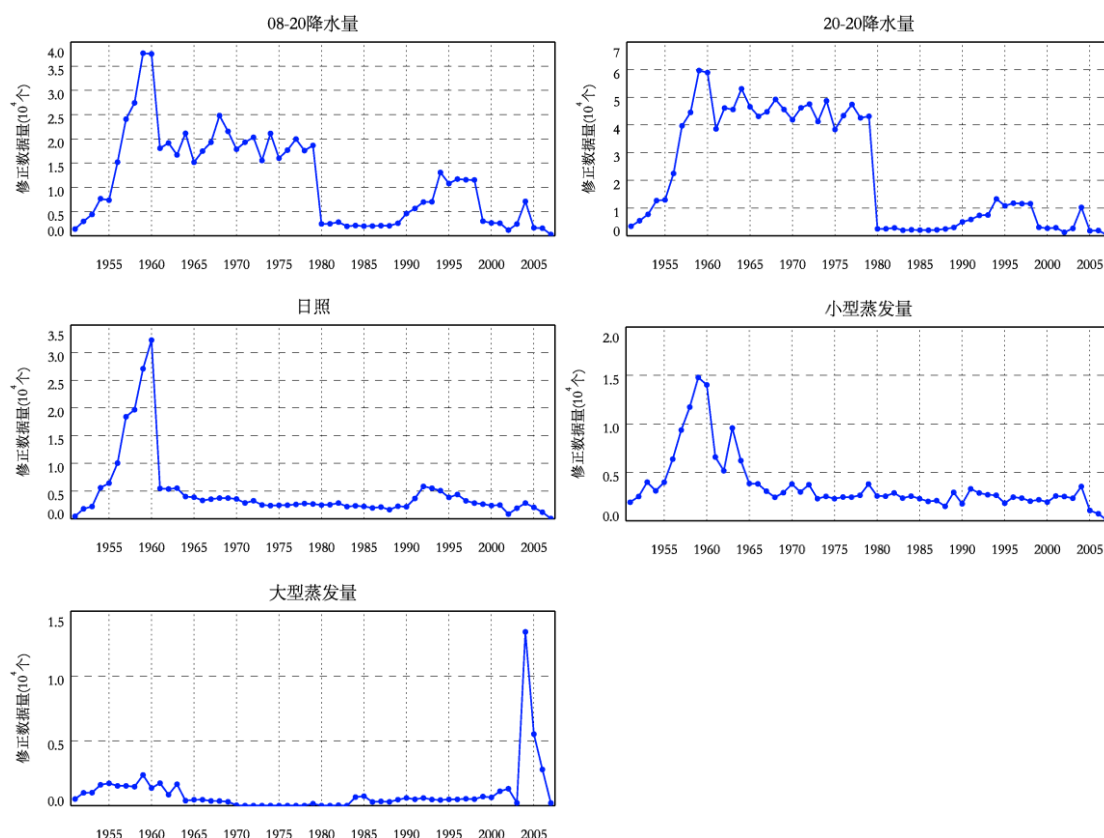


图 7 各要素项数据修正量逐年变化图

从各要素日值修正数据量随时间变化来看（见图 7），多种要素项目 2004~2006 年数据修正量较多。按照人工观测与自动观测平行期上报地面月报文件的有关规定，平行期第 1 年应上报人工观测数据，平行期第 2 年应上报自动观测数据。某些省份的台站在自动观测与人工观测平行期，未正确上报地面月报文件，此次地面基础气象资料建设工作对资料来源进行了重新确定。

降水量要素在 1956~1979 年修正量较大，这一期间 2.0 版数据集中删除了 2278 个站月的日降水量数据，新增 3392 个站月数据，此外其它修正数据量为 1919567 个。小型蒸发量在 1956~1964 年、日照要素在 1956~1960 年修正量较大，主要因为这一期间 2 种要素删除和新增了部分站月数据。

## 5 小结

本次制作发布的“中国国家级地面气象站基本气象要素日值数据集（V3.0）”包含了中国基本、基准气象站、一般气象站在内的主要 2474 个站点 1951 年 1 月以来地面本站气压、气温、相对湿度、风向风速、0cm 地温、降水、蒸发、日照日观测与统计数据以及相应时期的台站信息（台站经纬度、海拔高度）。数据

集中 2474 个站 1951~2010 年的数据在地面基础气象资料建设以及数据集制作过程中均经过了较严格的质量控制,对检测发现的可疑和错误数据普遍给与了人工核查与更正,并对因数字化遗漏造成的类似“数据缺测”现象进行了补录。

**1951~2010 年各要素数据的质量及完整性相对于以往发布的地面同类数据产品明显提高,各要素项数据的实有率普遍在 99%以上,数据的正确率均接近 100%。**

该数据集相对于 2008 年 4 月发布的“中国高密度台站地面基本气象要素日值数据集 (V2.0)”,不仅在数据文件中随日观测数据给出了各个时期对应的台站信息,更重要的是观测数据本身得到明显的修正改善。对比的 1951~2007 年 2425 站 7 种要素 17 类项目共计修正数据量 15,500,438 个,大多数项目的数据修正量在 0.5%-4.0%之间。

## 参考文献

- [1] 中国气象局. 地面气象观测规范. 北京: 气象出版社, 2003.
- [2] 中国气象局. 常规气象资料信息化模式文本汇编. 北京: 气象出版社, 2001.
- [3] 刘小宁, 任芝花. 地面气象资料质量控制方法研究概述. 气象科技, 2005, 33 (3): 199-203.
- [4] 任芝花, 熊安元. 地面自动站观测资料三级质量控制业务系统的研制. 气象, 2007, 33(1): 19-24.
- [5] 任芝花, 刘小宁, 杨文霞. 极端异常气象资料的综合性质量控制与分析. 气象学报, 2005, 63(4): 526-533.
- [6] 任芝花, 熊安元, 邹风玲. 中国地面月气候资料质量控制方法的研究. 应用气象学报, 2007, 18 (4): 516-523.
- [7] 任芝花, 许松, 孙化南等. 全球地面天气报历史资料质量检查与分析. 应用气象学报, 2006, 17 (4): 412-420
- [8] 任芝花, 赵平, 张强等. 适用于全国自动站小时降水资料的质量控制方法. 气象, 2010, 36(7): 123-132.
- [9] 陶士伟, 张跃堂, 陈卫红等. 全球观测资料质量监视评估. 气象, 2006, 32(6): 53-58.
- [10] 陶士伟, 仲跻芹, 徐枝芳等. 2009. 地面自动站资料质量控制方案及应用. 高原气象, 2009, 28(5): 1202-1209.
- [11] 王海军, 杨志彪, 杨代才等. 自动气象站实时资料自动质量控制方法及其应用. 气象, 2007, 33(10): 102-106
- [12] 中国气象局. 中华人民共和国气象行业标准: 地面气象观测资料质量控制, 2010.