

气候业务中全球台站逐日气候资料集的建立

王小玲 任福民 李 威 龚振淞

(国家气候中心 中国气象局气候研究开放实验室, 北京 100081)

提 要: 通过对全球逐日气温和降水观测资料集 (GDCN1.0) 的改进, 建立了适用于开展气候业务的全球台站逐日气候资料集。新建资料集保留了 GDCN1.0 资料集中有较好代表性的资料, 基本保持了原资料的分布, 同时中国范围 (含台湾岛) 的资料内容更加丰富。对气温进行了严格的质量控制, 最大限度地减少了错误值的影响。新建资料集在时间和空间分布上以及质量上都基本满足业务需求。

关键词: 气候业务 全球台站逐日气候资料集 气温 降水

An Operational Global Station Climatological Daily Data Set for Climate

Wang Xiaoling Ren Fumin Li Wei Gong Zhensong

(Laboratory for Climate Studies, CMA, National Climate Center, Beijing 100081)

Abstract: Based on the improvement and supplement of GDCN1.0, a Global Station Climatological Daily Data Set (GSCDDS) is established for climate operation. The data set has the same spacial distribution as GDCN1.0. GSCDDS has more stations in China than GDCN1.0. The data set has good representative both in area and time coverage. Invalid values of temperatures are deleted by a designed quality control procedure.

Key Words: climate operation GSCDDS temperature rainfall

引言

中国气象局早在 1980 年代就开展了对全球气温和降水异常监测业务,但只是月尺度的,目前为止仍然没有开展更短时间尺度的全球气候业务。其中最主要的原因就是缺乏全球逐日历史气候资料集。随着我国减灾防灾工作对及时准确的气候信息需求日益加深,迫切需要开展全球气候滚动监测业务。目前能够通过全球电讯系统(GTS)获取实时的全球逐日气温和降水观测资料,其中有些地区的资料经过严格质量控制^[1];但缺乏相应历史资料。因此,建立一套全球逐日历史气候资料集迫在眉睫。GDCN1.0 是由美国国家气候资料中心(NCDC)在 2002 年整理完成的一套全球逐日历史气候资料集,包含降水量、日最高气温和最低气温三个要素^[2~4],是目前公认的全球最完整的资料集。2004 年初,国家气候中心首次从 NCDC 引入该套资料集。由于该资料集建设之初并非面向业务需求,而是最大限度的保存收集到的资料^[5],使得资料良莠不齐。本文的目的是根据全球气候业务需求,对 GDCN1.0 气候资料集进行改进,建立一套适用于开展气候业务的全球台站逐日气候资料集(Global Station Climatological Daily Data Set, GSCDDS)。

1 GDCN1.0 资料状况及存在的问题

GDCN1.0 资料集包括全球逐日降水量、日最高气温和最低气温 3 个要素,共 32857 个测站,其中降水测站 32834 个(图 1),最高气温测站 14601 个,最低气温测站 14568 个(图 2)。

图 1 和图 2 显示,资料具有较好的空间覆盖率,但也存在空间分布上的不均匀。一

些区域过于密集,而另一些区域较稀疏甚至缺资料,北半球的资料分布较为理想,南半球缺漏地区较多,降水资料在全球范围的分布好于气温。降水资料基本覆盖了北美、欧亚大陆和澳大利亚,其中美国、印度、南美东部和南非比较密集,但南美西部和南部以及除南非以外的非洲大部几乎没有资料。日最高气温与最低气温的空间分布接近,这两个要素在南半球除澳大利亚外,几乎没有资料,而在北半球,美国依然较密集,南亚和非洲北部大部没有资料。

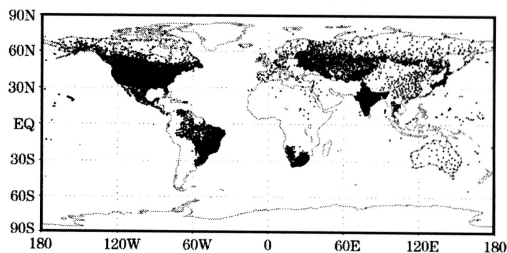


图 1 GDCN1.0 资料集中降水测站分布

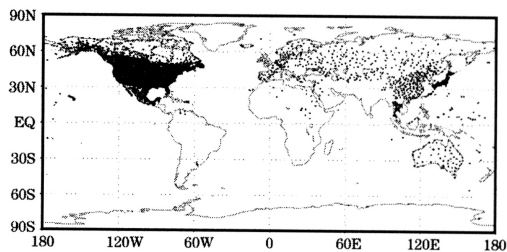


图 2 GDCN1.0 资料集中日最低气温测站分布

资料集在时间上的分布也存在很大差异。美国、澳大利亚、印度、南非南部和南美东部部分地区降水资料记录长度超过 50 年,其中美国的部分测站和澳大利亚东部记录长度超过 100 年,南美大部和南非西部以及加拿大和亚洲中部的部分测站记录长度不足 25 年,其余地区的记录长度在 25~50 年。大多数气温测站的记录长度均在 25 年以上,美国大部和欧亚大陆大部记录长度超

过50年,而北非和加拿大的部分测站记录长度短于25年,有的甚至不足10年。除资料的整个记录长度不一致外,资料的开始和结束年份也存在明显差异。

GDCN1.0资料集中的每一个测站均有相应的国家站号(national station number),但仅有不足十分之一的测站同时具有WMO站号,在很多区域测站没有WMO站号(非WMO测站)。

此外,资料集包含的198个中国测站全部为大陆测站,在我国西部的分布相对稀疏,记录结束时间为1997年12月31日。

2 GSCDDS的建设方案

由于气候变化问题的存在,气候平均值会随时间改变^[6],世界气象组织要求目前业务中应用的气候平均时段为1971~2000年。在GTS上实时交换的全球资料仅限于WMO测站。因此,在业务上应用的气候资料集应满足如下要求:空间覆盖范围尽可能广;所选测站尽可能具有WMO站号;尽可能降低密集分布的测站中短资料带来的负面影响;记录时段尽可能满足气候平均时段。针对GDCN1.0资料集的状况,有必要从空间和时间上以及是否具备世界气象组织台站号方面对资料进行挑选和扩充,并进行资料的质量控制。

2.1 中国资料的扩充

新的中国资料包括两个方面:一是来自国家气象信息中心气候资料室的中国大陆647个测站,资料时段从建站至2001年12月31日,包括日降水量、日平均气温、日最高和最低气温四个要素;另一方面为台湾省18个测站1950~2000年的日降水量资料。新增的中国资料不仅延长了记录长度,加密了空间分布,同时填补了一些区域的空

白,此外还增加了日平均气温这一要素。

2.2 WMO站号测站的选取方案

通过查找相关资料,增补了1000多个测站的世界气象组织站号,此时资料集中的WMO站号测站除南美中西部外,基本代表了资料集的空间分布;同时剔除记录时段短于1年的测站。

2.3 气候平均时段定义

由于不同测站资料的起止年份差异很大,对其气候平均时段的选取不能固定于某一个时段。考虑到世界气象组织对业务工作气候平均时段的要求,本文选取1971年1月1日~2000年12月31日作为气候平均时段。假设资料的记录长度为 M 年,其在1971年1月1日~2000年12月31日间的记录年数为 N ,若 $N \geq 10$,则选择这 N 年作为气候平均时段,若 $N < 10$,则选取 M 作为气候平均时段。

2.4 非WMO站号测站的删选方案

除去被选取的WMO站号测站后,还有2万多个测站(对气温而言有1万多个)为非WMO站号。在地理上,非WMO测站在南美、南非、墨西哥和西太平洋等地弥补有WMO站号测站分布的不足。考虑到增加资料覆盖率的同时应尽可能避免短资料带来的负面影响,确立如下删选方案:

第一步,非WMO站号测站初选:

(1) 计算邻站信息 K

计算某一非WMO站号测站与2.2中已挑选的WMO站号测站(设有 t 个)之间的距离 D_i ,根据临界距离 D_0 (本文取150km),得到该测站的邻站信息

$$K = \sum_{i=1}^t k_i, k_i = \begin{cases} 0, & D_i > D_0 \\ 1, & D_i \leq D_0 \end{cases}$$

(2) 统计该测站的资料记录年数 Y 。

(3) 当 $K=0$ 且 $Y \geq 1$ 年时, 该非 WMO 站号测站通过初选。

第二步, 非 WMO 站号测站优选:

通过初选的非 WMO 站号测站在一些区域仍然很密集, 为尽可能降低密集分布的测站中短资料带来的负面影响, 有必要进一步进行择优筛选:

(1) 在通过初选的站点 (设有 m 个) 范围内统计每一个站的邻站信息 n_j ;

(2) 依据 n_j 从大到小对测站进行排序, $S_j, j=1, \dots, m$;

(3) 依照测站的排序, 逐一对每一个测站 S_j , 按下列步骤从它本身及其 n_j 个邻站一共 n_j+1 个测站中挑选一个最优测站:

① 统计 n_j+1 个测站的气候平均时段指数 $C_x = 3 \times N + (M - N)$, 其中 $x=1, \dots, n_j+1$ 。

② 依据 n_j+1 个测站的剔除次数 l_x (对于第一个测站, n_1+1 个测站的 l_x 皆为 0), 仅对剔除次数 $l_x \leq 1$ 的测站, 从中挑选 C_x 最大的测站为最优测站。

③ 对除最优测站之外的所有测站, 其剔除次数 l_x 加 1。

2.5 气温的质量控制方案

由于观测仪器、资料处理和传输过程可能引起的误差影响, 资料集难免存在一定的错误, 因此, 对资料进行质量控制, 最大限度地减少错误值的可能影响是非常必要的。依据潘晓华等^[7]对气温的质量控制方法, 本文同样也从相同要素空间一致性和相关要素一致性两个方面来剔除最高气温、最低气温和平均气温的错误资料。由于国外的资料只有最高和最低气温两个要素, 只考虑一个相关要素。

3 GSCDDS 的特点

3.1 中国资料更加丰富

新建的资料集包含的中国资料更加丰富, 在要素上增加了日平均气温, 测站数量也大幅增加, 改变了原资料集中中国西部台站稀少的状况, 增加了台湾岛的降水资料, 并将记录长度延长到 2000 年以后。

3.2 空间分布特征

WMO 测站包含 4532 个降水测站和 2185 个气温测站, 在南美中西部的大范围区域没有 WMO 测站。通过对非 WMO 测站的删选, 降水和气温分别增加 572 和 118 个测站。

经过上述两个过程后, 共选取降水测站 5104 个 (图 3), 气温测站 2303 个 (图 4)。无 WMO 站号弥补了南美大部、南非、墨西哥和西太平洋等地降水资料和墨西哥和西太平洋等地气温资料分布的不足。资料在空间上的覆盖范围基本代表了原资料的特征, 且分布均匀。

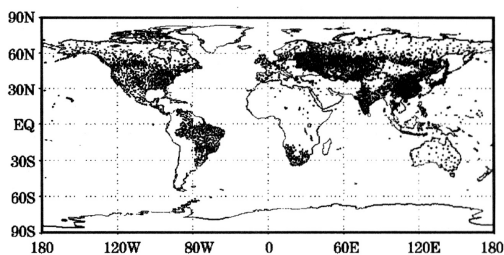


图 3 GSCDDS 降水测站分布

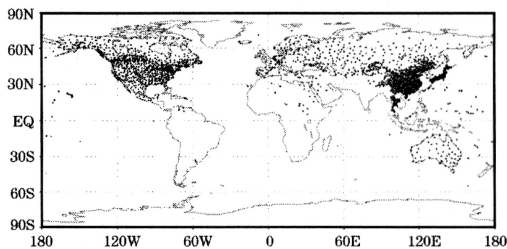


图 4 GSCDDS 气温测站分布

3.3 气候时段分布特征

在所选取的 5104 个降水测站中,其气候平均时段最短记录为 3 年,气候平均时段小于 10 年的站点不足总站数的 3%。在 2303 个气温测站中,气候平均时段最短记录为 1 年,气候平均时段小于 10 年的测站不足总站数的 4%。就降水而言,气候平均时段完全符合世界气象组织标准的占 24.1%,气候平均时段在 1971~2000 年间但其长度超过 20 年短于 30 年的占 48.3%,即气候平均时段在 1971~2000 年间且长度超过 20 年的测站占总站数的 70%以上。对气温而言,气候平均时段完全符合世界气象组织标准的占 52.5%,气候平均时段在 1971~2000 年间且长度超过 20 年的测站达 75%以上。澳大利亚、加拿大和美国大部分测站以及中国绝大多数测站的气候平均时段均为 1971~2000 年。

3.4 气温的质量控制结果

经过质量控制,剔除的所有错误记录占整个资料集记录总数的 0.065%,最大限度地减少了在业务应用中错误资料的影响。

4 结 语

通过以上方案设计,建立了适于开展气候业务的全球逐日气候资料集,新建资料在中国范围(含台湾岛)显著扩充,舍去了大

量非 WMO 测站,保留了时段较长的资料,基本保持了原资料集的空间分布。对气温进行了严格的质量控制,最大限度的减少了错误记录的影响。新建资料集无论在时间和空间分布上,还是在质量上都基本满足业务需要,适合于在气候业务中应用。当然,该套资料集在一些地区仍然存在缺漏,这有待于在未来的工作中进一步完善。

参考文献

- 1 熊安元. 北欧气象观测资料的质量控制. 气象科技, 2003, 31 (5): 314~320.
- 2 Lanzante, John R. Resistant, robust, and nonparametric techniques for the analysis of climate data. Theory and examples, including applications to historical radiosonde station data. *Int. J Climatol.*, 1996, 16: 1197—1226.
- 3 Legates, David R., and Cort J. Willmott. Mean seasonal and spatial variability in gaugencorrected global precipitation. *International Journal of Climatology*, 1990, 10: 111—127.
- 4 Legates, David R. and Cort J. Willmott. Mean seasonal and spatial variability in global surface air temperature. *Theoretical and Applied Climatology*, 1990, 41: 11—21.
- 5 Peterson, Thomas, Harald Daan, Philip Jones. Initial Selection of a GCOS Surface Network. *Bulletin of the American Meteorological Society*. 1997, 78 (10): 2145—2152.
- 6 王绍武. 现代气候学研究进展. 北京: 气象出版社, 2001.
- 7 潘晓华, 翟盘茂. 气温极端值的选取与分析. 气象, 2002, 28 (10): 28—31.