

秦榕,玛依努尔·阿布拉,井立红,等.新疆1993—2012年辐射观测资料质量分析[J].沙漠与绿洲气象,2014,8(2):54-58.
doi:10.3969/j.issn.1002-0799.2014.02.007

新疆1993—2012年辐射观测资料质量分析

秦榕¹,玛依努尔·阿布拉¹,井立红²,杨霏³,陶淘¹,肖静⁴

(1.新疆气象信息中心,新疆 乌鲁木齐 830002;2.塔城地区气象局,新疆 塔城 834700;
3.乌鲁木齐市气象局,新疆 乌鲁木齐 830002;4.莎车气象局,新疆 喀什 844700)

摘要:利用新疆11个辐射站最近20 a(1993—2012年)共2 640个辐射信息化数据文件,从原始数据表-33及月辐射数据(R)文件中遴选出疑误数据共计7 025条,其中更改错误数据323个,对疑误、错误数据进行了质量检查及评估,并应用数理统计方法分析了其时空分布特征以及总辐射、散射辐射、直接辐射、反射辐射、净全辐射出现疑误数据的概率,同时对出现疑误、错误数据的原因进行了分析,结果表明:新疆最近20 a辐射疑误数据主要发生并集中在一、二级辐射观测站,其中数据文件部分出现疑误较多,记录格式部分出现错误相对较少;疑误数据出现的主要原因是由于辐射参数的设置只能考虑正常天气状况,而异常天气情况比较复杂,很难完全涵盖。此外直接辐射表误差和自动跟踪架误差也是出现疑误数据较多的原因。整体而言,自2003年辐射数据采集质控实现自动化后,经质量控制软件检查的R文件出现疑误、错误数据的概率显著减少,资料可靠,可为科研工作提供有利的技术支持。

关键词:辐射资料;质量评估;新疆

中图分类号:P416

文献标识码:B

文章编号:1002-0799(2014)02-0054-05

为保证地面气象辐射观测资料最大可能的合理性,在资料形成时有其规定的要素允许值范围,超出此允许值范围的资料被定义为错误资料。近年来有学者对每日每时上传的地面气象辐射实时资料和每月上旬传输的非实时资料即月辐射数据(R)文件的质量控制方法做过研究分析^[1-3],并得出一些对实际工作有一定指导意义的结论,但有关辐射资料质量控制的分析尚未见公开报道。文章利用相关软件,探讨在质量控制辐射数据(R)文件时,需要对经质量控制后系统自动生成的疑误信息进行人工判断,并结合有关技术规定和问题解答给予适当处理,以确保辐射观测数据的真实、准确、可靠,进而为业务科研工作提供科学的、准确的数据。

1993—2012年辐射数据信息化文件在新疆已投入科技服务与业务使用,共有11个辐射观测站的观测数据实现了信息化,其中国家一级站2个:包括

乌鲁木齐、喀什;二级站5个:包括伊犁、哈密、阿勒泰、塔城、和田;三级站4个:包括焉耆、若羌、吐鲁番、阿克苏。由于历史资料时间跨度大,新旧辐射数据文件格式不统一,而期间《OSSMO2004地面测报软件》又做过多次修改升级,对质量控制过程中出现的疑误数据增加了很多新的补充规定,但历史数据文件未重新经过系统的质量控制,这或多或少都给辐射数据的质量带来一些影响。因此,结合不断更新的补充规定对辐射信息化文件进行重新评估和分析,不仅优化了辐射数据的质量,且可为其他气象数据的信息化处理提供参考和借鉴,同时可以使得气象资料数据集的用户了解疑误数据的种类和原因,掌握新疆辐射数据整体质量状况,从而为客观的、定量的、安全的使用辐射数据提供技术支持。

1 资料与方法

利用最近20 a(1993—2012年)新疆11个辐射站逐月信息化辐射数据R文件(对部分老辐射数据R文件使用《OSSMO2004地面气象测报软件》进行了转换),采用国家局下发的《OSSMO2004地面气象

收稿日期:2013-04-20;修回日期:2013-06-20

基金项目:公益性气象行业专项(GYHY201206013)资助。

作者简介:秦榕(1961-),女,高级工程师,主要从事地面辐射数据质量控制及分析。E-mail:qinrong1961@sina.cn

测报业务软件》和《基础气象资料专项工作——历史辐射气象资料质量控制软件(1.0.0)》分别对新疆 11 个站最近 20 a 共 2 640 个辐射数据文件进行质量检查,检查内容主要包括:总辐射、散射辐射、直接辐射、反射辐射、净全辐射等五项要素,应用数理统计方法对疑误、错误数据个数的空间分布、时间分布、各要素分布等进行评估分析。

方法:

$$\text{平均单站月疑误数据个数} = \frac{\text{总疑误数据个数}}{\text{总站月数}}$$

$$\text{疑误数据率} = \frac{\text{总疑误数据个数}}{\text{总数据个数}}$$

2 辐射疑误(错误)数据的时空分布特征

2.1 空间分布

经计算,新疆 11 个辐射站最近 20 a 辐射数据的疑误数据共计 7 025 条(包括更改的错误数据 323 个),平均单站月疑误数据个数为 2.7 个/(站·月),疑误数据率为 0.0%。经分析发现,新疆大部分辐射观测站均出现过疑误数据,其中一级站由于观测要素多,因此出现的疑误也最多,达 4 573 个,占疑误数据总数的 65%;二级站次之(2 423 个),占疑误数据总数的 34%;三级站共计 29 个,占疑误数据总数的 0.4%。

2.2 时间分布

2.2.1 年际变化

图 1a 为新疆 11 个辐射站疑误数据个数的年际变化图,从图中可以看出,疑误数据出现较多的年份有 1993、1994、1999、2007、2012 年,平均在 400 个以上,占疑误数据总数的 31%;其次为 1997、2000、2004、2006、2011 年,平均为 350 个,占疑误数据总数的 26%;2005 年疑误数据最少,共 213 个,占疑误数据总数的 2%。

图 1b 为新疆 11 个辐射站错误数据个数的年际变化图,从图中可以看出,错误数据出现最多的年份是 2005 年,经对比审核发现,其原因是由于软件在处理要素极值出现时间上应该按...(即空白)处理,在做 R 文件维护时,却将空白记录默认为“////”而引起。错误数据出现多的年份是 1996—1997 年,且主要发生在乌鲁木齐站,原因在于乌鲁木齐国家一级辐射站使用了超检仪器,导致异常数据偏多。

统计分析表明,自 2003 年辐射数据采集质控陆续实现自动化以后,经过质量控制软件检查的 R 文件出现错误的概率显著减少,即使有个别数据出现

疑异,也基本是由于记录格式、数据出现野值或软件使用初期其自身的不完善造成的。

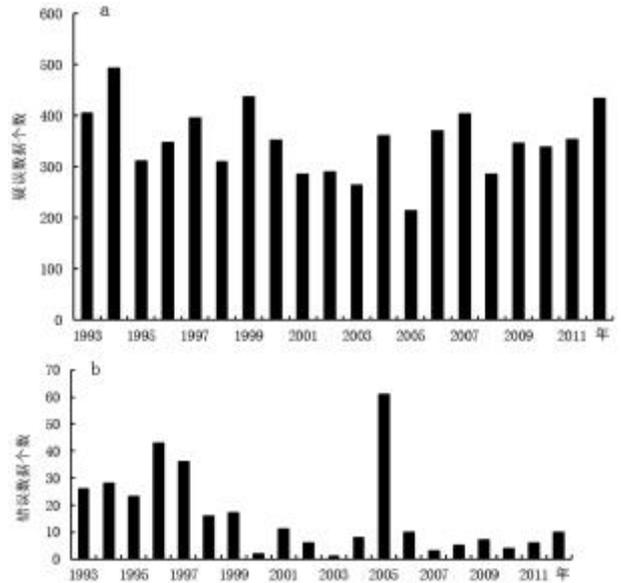


图 1 新疆辐射站疑误数据个数(a)和错误数据个数(b)年际变化

2.2.2 月、季变化

图 2 为新疆 11 个辐射站疑误、错误数据个数的月际变化直方图,从图中可以看出,各月均有疑误、错误数据出现,但其月际分布不均匀,其中 1 月、12 月疑误、错误数据个数最多,究其原因主要是冬季新疆天空状况、下垫面状况复杂多变所致,6 月、8 月疑误数据个数相对较少,6 月、9 月错误数据最少。

四季中冬季疑误(错误)数据个数最多,达 2 773 (131)个,春季、秋季次之,分别为 1 404(71)个、1 449 (61)个,夏季最少 1 162(57)个。原因在于夏秋季节新疆气候干旱少雨,天空晴朗稳定,大气透明度极

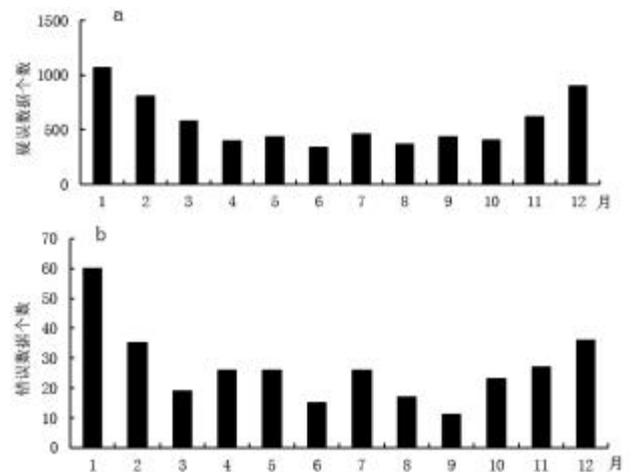


图 2 新疆辐射站疑误数据个数(a)和数据个数(b)月变化

佳,使得疑误(错误)数据相对较少。

3 辐射五要素疑误(错误)数据及原因分析

3.1 五要素疑误(错误)数据个数分布

对新疆 11 个辐射站各要素进行质量控制后提出疑误信息 7 025 条,利用检查结果绘制了各要素疑误数据个数图(图 3a),从图中可以看出,总辐射、净全辐射出现疑误数据最多,其次为反射辐射和直接辐射,最少为散射辐射,其中总辐射出现疑误数据 2 185 条,占疑误数据总数的 32%,净全辐射出现疑误数据 2 173 条,占疑误数据总数的 32%;反射辐射出现疑误数据 1 421 条,占疑误数据总数的 21%,直接辐射出现疑误数据 1 063 条,占疑误数据总数的 16%,散射辐射出现疑误数据 11 条,占疑误数据总数的 0.0%。

图 3b 为五要素出现错误数据个数统计,从图中可以看出,共修改错误数据 323 个,其中,净全辐射出现错误数据最多,为 62 个,占错误总数的 19%;反射比错误数据 58 个,占错误总数的 18%;反射辐射错误数据 40 个,占错误总数的 12%,总辐射错误数据 30 个,占错误总数的 9%;直接辐射错误数据 34 个,占错误总数的 11%;水平面直接辐射错误数据 31 个,占错误总数的 10%;散射辐射错误数据 6 个,占错误总数的 2%;大气混浊度错误数据 3 个,占错误总数的 1%;封面封底错误数据 59 个,占错误总数的 18%。

3.2 辐射疑误(错误)数据信息统计

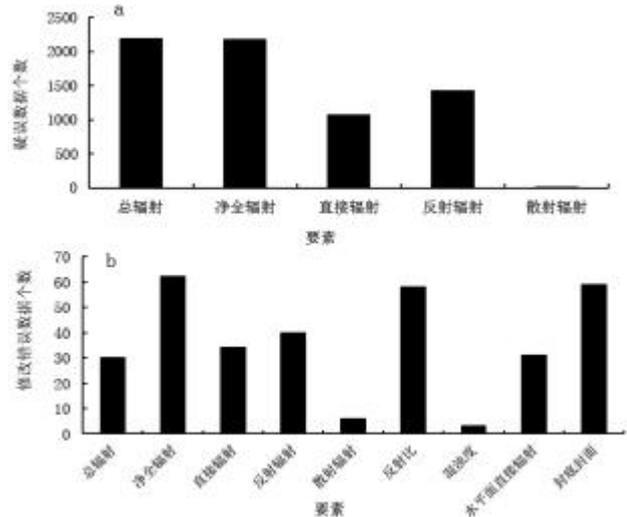


图 3 新疆辐射站各要素疑误(a)和错误(b)数据个数

表 1 为新疆辐射站各类疑误(错误)信息个数统计,从要素分布中可以看出各站点各要素均有疑误数据出现,黑体部分为定性的错误数据。

3.3 辐射疑误(错误)数据原因分析

3.3.1 系统原因分析

3.3.1.1 仪器原因

直接辐射仪器多用全自动跟踪架进行观测,由于跟踪架长时间在露天条件下运转,很难做到完全准确跟踪太阳,此外观测数据是否可靠还与辐射观测员是否按照规程操作,是否认真对准太阳有关^[9]。目前直接辐射资料数据总体偏小,其误差大小是随

表 1 新疆辐射站各类疑误信息个数统计

疑误类型(出现要素)	出现站点	疑误个数
超出气候极限值(反射辐射、净全辐射、直接辐射)	阿勒泰、塔城、伊宁、乌鲁木齐、喀什、和田、哈密	1 800
总射日总量超过可能总射日总量的 120%	阿勒泰、塔城、焉耆、伊宁若羌、哈密	160
当日无日照,而总辐射与散射日总量之差>10.0 MJ/m ²	喀什、乌鲁木齐	3
要素时总量连续 4 个(7 个)时次≥4.0 MJ/m ² 、1.5 MJ/m ² 或 2.5 MJ/m ² (或数据相同)	塔城、哈密、乌鲁木齐、若羌、焉耆	66
要素辐射辐照度不为 0 或 //,出现时间不为...或 ///	若羌、和田、喀什、乌鲁木齐、伊宁、塔城、阿勒泰、哈密	569
要素日总量不等于时总量合计	和田、阿勒泰、焉耆、哈密、喀什、乌鲁木齐、若羌	25
反射比不等于反射辐射日总量比总辐射日总量或辐照度一和混浊度 XXX 不——对应	乌鲁木齐、喀什	177
总辐射日最大辐照度<反射(散射、净射)辐射日最大辐照度	乌鲁木齐、喀什塔城、阿勒泰、伊宁、哈密	187
要素连续 2 个时次或以上时次缺测,日总量应缺测	喀什、焉耆、若羌	4
总辐射时总量>1.0 时,净全辐射>总辐射的时总量	阿勒泰、塔城、伊宁、乌鲁木齐、喀什、和田、哈密	349
总射日总量不等于散射日总量与水平直接辐射曝辐量的和	乌鲁木齐、喀什	34
总射日(时)总量小于反射(散射、净射)日(时)总量	乌鲁木齐、喀什、伊宁、塔城	1 347
净全辐射时总量全天均≥0.0 MJ/m ² 或净全辐射时总量全天均≤0.0 MJ/m ²	阿勒泰、塔城、乌鲁木齐、伊宁、和田、喀什、哈密	1 316
水平面直接辐射总量超过直接辐射日总量	乌鲁木齐、喀什	930

机的。总辐射观测的误差约为5%左右,直接辐射表观测误差约为2%。散射辐射准确值虽然必须乘以大于1的遮光环订正系数才能获取,但它还受季节、纬度和各地云量多少的影响。所以对大面积出现总辐射日(时)总量小于反射辐射日(时)总量、总辐射日(时)总量小于散射辐射日(时)总量、总辐射日(时)总量小于净全辐射日(时)总量的记录,国家气象局定义 $\leq 1.0 \text{ MJ/m}^2$ 属正常记录,对超出此范围的记录需认真分析查找原因,做出正确判断。

由于全波段辐射观测仪器制造难度大,加上国际上尚无全辐射测量的世界标准等原因,导致目前我国测量的净全辐射误差约为10%~20%左右^[4],因此会出现总辐射时总量 $>1.0 \text{ MJ/m}^2$ 时,净全辐射大于总辐射时总量的疑误记录。

3.3.1.2 台站参数设置

设置辐射审核参数时,为及时发现采集出现的异常数据,而将参数界限值设置为多年人工观测平均值,就会有超出气候极限值的现象发生。

3.3.1.3 软件原因

日最大辐射辐照度为0,出现时间不为“...”、总辐射日最大辐照度 $<$ 反射辐射日最大辐照度、总辐射日最大辐照度 $<$ 散射辐射日最大辐照度,均系受天空有云对记录的影响,同时由于软件对特殊情况处理判断无法过细造成。

3.3.1.4 天空及下垫面状况

降水、仪器加盖、天空有云、多云、阴、太阳周围有微量云、飞行物(禽鸟飞过)、重霜、冬季积雪较深台站等特殊情况出现,需要人机结合分析疑误信息,确有疑误应对记录进行处理。

3.3.1.5 数据采集异常

仪器采集数据时某些要素偶尔会出现一些误差,如:漂零点漂移、明显偏离真值的粗大误差、随机误差以及由于小尺度天气系统扰动而引起的微气象误差,这些现象在此次质控时也出现数次,如经分析能够确定偏离真值,应予以删除。

3.3.2 净全辐射疑误数据原因分析

净全辐射时总量全天均 $\leq 0.0 \text{ MJ/m}^2$ 、净全辐射时总量全天均 $\geq 0.0 \text{ MJ/m}^2$,这两种疑误信息发生的原因都是因为降水出现,仪器加盖或多云天气导致;此外新疆北疆地区出现重霜、有微量云、飞行物扰乱(禽鸟飞过)或积雪深厚的台站容易出现此疑误,此时要根据具体情况分析处理。

3.3.3 直接辐射疑误数据原因分析

该日无日照,而直接辐射日总量 $>10.0 \text{ MJ/m}^2$,

多是由于直接辐射表在有云等情况下跟踪误差引起;水平面直接辐射总量超过直接辐射日总量是由于太阳周围有云的情况下会出现此类疑误信息。

3.3.4 计算错误

反射比不等于反射辐射日总量比总辐射日总量、要素日总量不等于时总量合计、总射日总量不等于散射日总量与水平直接辐射曝辐量的和,这些疑误数据信息多在20世纪90年代手工计算错误出现,其他时间均为计算进位影响。

4 结论

(1)新疆11个辐射站最近20 a辐射数据的疑误数据共计7 025条,其中修改错误数据323个,平均单站月疑误数据个数为2.7个/站·月,疑误数据率为0.0%;其中一级站出现疑误最多,二级站次之,三级站最少。

(2)疑误数据主要原因是由于辐射参数是按正常天气状况下设置,而异常天气情况比较复杂,很难完全涵盖。其次在湿度大、阴天、风小的夜间及风雨雪天气、天空有云、降水、下垫面出现重霜、冬季积雪等均会出现疑误记录,此外直接辐射表误差、全自动跟踪架误差、净表误差和散表误差也是出现疑误数据的原因。

(3)自2003年辐射数据采集质控实现自动化后,经质量控制软件检查的R文件出现错误的概率显著减少;整体来说11个站最近20 a信息化资料可靠,疑误数据在很正常的范围内,经质量检查后的辐射数据是真实、可靠的,可以为气象业务和社会各界提供数据服务。

参考文献:

- [1] 王新华,罗四维,刘小宁,等.国家级地面自动站A文件质量控制方法及软件开发[J].气象,2006,32(3):107-112.
- [2] 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[J].气象学报,2005,63(4):526-533.
- [3] 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[J].气象,2007,33(1):19-24.
- [4] 秦榕,井立红,何亚平,等.净全辐射异常值得审核方法[J].气象,2011,37(5):645-648.
- [5] 胡玉峰.自动气象站原理与测量方法[M].北京:气象出版社,2004:6.
- [6] 中国气象局监测网络司.地面气象观测规范技术问题解答(第1号)[M].北京,2005.
- [7] 中国气象局监测网络司编写组.地面气象观测规范[M].北京:气象出版社,2003.

- [8] 秦榕,何亚平,侯玲红,等.新疆辐射观测资料质量分析[J].沙漠与绿洲气象,2010,4(增刊):196-198.
- [9] 玛依努尔,秦榕,黎莹辉.新疆辐射数据人机审核方法[J].沙漠与绿洲气象,2011,5(增刊):99.

Analysis of Radiation Observational Data Quality from 1993 to 2012 in Xinjiang

QIN Rong¹, Mayinuer·ABULA¹, JIN Lihong², YANG Xian³, TAO Tao¹, XIAO Jing⁴

(1.Xinjiang Meteorological Information Center, Urumqi Xinjiang 830002, China; 2.Tacheng Meteorological Bureau, Tacheng, 834700, China; 3.Urumqi Meteorological Administration, Urumqi 830002, China; 4.Shachen Meteorological Bureau, Kashi 844700, China)

Abstract 2 640 data files of radiation informatization over last 20 years (1993-2012) from 11 radiation stations in Xinjiang were examined. With original data sheet-33 and monthly radiation data (R) files, 7 025 doubted data were selected, including 323 erroneous data to be corrected. Such doubted and erroneous data were treated with quality inspection and evaluation, and their probabilities with respect to spatial-temporal distribution characteristics, total radiation, scattered radiation, direct radiation, reflected radiation and net total radiation were analyzed with mathematical statistical method; and the reasons were also analyzed. The results indicated that the doubted data occurs and concentrates mainly at the primary and secondary radiation observation stations, mostly in the form of data files and record format. The major reason was that only normal weather conditions were considered for the setting of radiation parameters, but abnormal weather conditions are too complicated to be covered fully. Besides, direct radiation meter error and auto-tracking mount error also account for many doubted data. On the whole, the probability of doubted data and erroneous data declined significantly in R files examined with quality control software since radiation data acquisition quality control was automated in 2003. The data from R files are reliable to provide technical support to scientific research.

Key words radiation data; quality evaluation; Xinjiang