

黄河流域九省区 1-km 分辨率月度相对湿度数据集 (2000–2015)

蔡红艳^{1*}, 蒋啸^{1,2}, 杨小唤^{1,2}

1. 中国科学院地理科学与资源研究所, 资源与环境信息系统国家重点实验室, 北京 100101;

2. 中国科学院大学, 北京 100101

摘要: 相对湿度与大气污染物的扩散、迁移和转化密切相关, 是区域生态环境质量评估的重要指标。基于黄河流域所在的九省区及周边 940 个气象站点 2000–2015 年逐月相对湿度观测数据, 在 ANUSPLIN 软件平台上以高程作为独立协变量, 采用三次样条函数进行插值, 生成黄河流域九省区相对湿度时空分布数据集 (2000–2015)。数据时间分辨率为月, 空间分辨率为 1 km, 投影方式为阿尔伯斯等面积投影, 坐标系为 WGS-84。数据集存储为.tif 格式。

关键词: 黄河流域; 相对湿度; 2000–2015

DOI: <https://doi.org/10.3974/geodp.2021.02.09>

CSTR: <https://cstr.escience.org.cn/CSTR:20146.14.2021.02.09>

数据可用性声明:

本文关联实体数据集已在《全球变化数据仓储电子杂志 (中英文)》出版, 可获取:

<https://doi.org/10.3974/geodb.2021.03.12.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2021.03.12.V1>.

1 前言

相对湿度由大气中实际水汽含量和空气温度共同决定, 是衡量地区干湿程度的重要指标^[1–3]。同时, 区域干湿状况直接关系到污染物在大气中的扩散、迁移和转化^[4–6], 与人体健康和生活舒适度密切相关, 因此, 相对湿度也是区域环境质量与城市宜居等评估的重要指标。

黄河流域横跨我国东中西三大区域, 覆盖 9 省区, 是我国生态保护及经济发展的重点区域, 也是生态保护与经济发展矛盾较为突出的区域。黄河流域是我国典型的气候敏感区之一, 大部分处于干旱、半干旱地区, 流域水资源分布不均, 生态环境脆弱。近 20 年来, 由于不均衡的水资源开发, 导致湖泊面积萎缩, 洪涝灾害频发, 干旱灾害

收稿日期: 2021-02-09; 修订日期: 2021-04-20; 出版日期: 2021-06-25

基金项目: 中华人民共和国科学技术部 (2017YFC0503803)

*通讯作者: 蔡红艳 Y-8555-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, caihy@igsrr.ac.cn

数据引用方式: [1] 蔡红艳, 蒋啸, 杨小唤. 黄河流域九省区 1-km 分辨率月度相对湿度数据集 (2000–2015) [J]. 全球变化数据学报, 2021, 5(2): 175–180. <https://doi.org/10.3974/geodp.2021.02.09>. <https://cstr.escience.org.cn/CSTR:20146.14.2021.02.09>.

[2] 蔡红艳, 蒋啸, 杨小唤. 黄河流域九省区 1-km 分辨率月度相对湿度数据集 (2000–2015) [J/DB/OL]. 全球变化数据仓储电子杂志, 2021. <https://doi.org/10.3974/geodb.2021.03.12.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2021.03.12.V1>.

加剧^[7,8]，高工业化水平造成空气质量恶化，大气污染严重^[9]。2019 年 9 月，习近平总书记已明确提出将黄河流域的生态保护与高质量发展列为重大国家战略^[10]。

在此背景下，黄河流域九省区的相对湿度时空分布数据有助于深入了解该区域的干湿变化规律，可为区域生态环境质量评估与治理提供数据支撑，对于贯彻落实黄河流域生态保护与高质量发展战略具有重要意义。

2 数据集元数据简介

《黄河流域九省区 1-km 分辨率月度相对湿度数据集（2000–2015）》^[11]的名称、作者、地理区域、数据年代、时间分辨率、空间分辨率、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

表 1 《黄河流域九省区 1-km 分辨率月度相对湿度数据集（2000–2015）》元数据简表

条 目	描 述
数据集名称	黄河流域九省区 1-km 分辨率月度相对湿度数据集（2000–2015）
数据集短名	RHU_9PYRB
作者信息	蔡红艳 Y-8555-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, caihy@igsnr.ac.cn 蒋啸 AAE-1541-2021, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室; 中国科学院大学, jiangx.20b@igsnr.ac.cn 杨小唤 AAC-8887-2021, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室; 中国科学院大学, yangxh@lreis.ac.cn
地理区域	黄河流域 数据年代 2000–2015 年 时间分辨率 逐月 九省区
数据格式	.tif 数据量 压缩后 1.45 GB 空间分辨率 1 km
数据集组成	数据集包含 2 个文件夹： 文件夹“Studyarea”为研究区边界, 数据格式为.shp。“9PYRB_RHU”包含 2000-2015 年逐月的相对湿度时空分布数据, 数据格式为.tif。数据集共包含 775 个数据文件, 相对湿度时空分布文件的文件名包含其时相信息, 如“9PYRB_RHU200001.tif”为 2000 年 1 月的黄河流域九省区的相对湿度数据, 空间分辨率为 1 km, 单位为% 中华人民共和国科学技术部（2017YFC0503803）
基金项目	
数据计算环境	ANUSPLIN、ArcGIS, 中国科学院地理科学与资源研究所内平台
出版与共享服务平台	全球变化科学研究数据出版系统 http://www.geodoi.ac.cn
地址	北京市朝阳区大屯路甲 11 号 100101, 中国科学院地理科学与资源研究所
数据共享政策	全球变化科学研究数据出版系统的“数据”包括元数据（中英文）、通过《全球变化数据仓储电子杂志（中英文）》发表的实体数据集和通过《全球变化数据学报（中英文）》发表的数据论文。其共享政策如下：（1）“数据”以最便利的方式通过互联网系统免费向全社会开放, 用户免费浏览、免费下载；（2）最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源；（3）增值服务用户或以任何形式散发和传播（包括通过计算机服务器）“数据”的用户需要与《全球变化数据学报（中英文）》编辑部签署书面协议, 获得许可；（4）摘取“数据”中的部分记录创作新数据的作者需要遵循 10%引用原则, 即从本数据集中摘取的数据记录少于新数据集总记录量的 10%, 同时需要对摘取的数据记录标注数据来源 ^[12]
数据和论文检索系统	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS/ISC, GEOSS

3 数据研发方法

3.1 算法原理

ANUSPLIN 软件是基于普通薄盘和局部薄盘样条函数对多变量进行内插的工具，适用于时间序列气象数据的插值。局部薄盘光滑样条是对薄盘光滑样条原型的扩展，除普通的样条自变量外，允许引入协变量子模型。输出统计数据与 N 个观测值 z_i 的局部薄盘样条模型相关，其公式为：

$$z_i = f(x_i) + b^T y_i + e_i (i = 1, 2, \dots, N)$$
 (1)

式中， x_i 是独立变量的 d 维向量； f 是 x_i 的未知光滑函数； y_i 为独立协变量的 p 维向量； b 为 y_i 系数的 p 维未知向量； e_i 是随机误差^[13]。

函数 f 和系数 b 通过式(2)，即由最小二乘法来确定：

$$\sum_i^n \left[\frac{z_i - f(x_i) - b^T y_i}{w_i} \right]^2 + \rho J_m(f)$$
 (2)

式中， $J_m(f)$ 是函数 $f(x_i)$ 的粗糙度测算函数，定义为 f 的 m 阶偏导， ρ 为光滑参数，在数据保真与曲面粗糙度之间起平衡作用， w_i 为已知的相对误差。

在 ANUSPLIN 插值过程中，最佳模型主要依据广义交叉验证（Generalised Cross Validation, GCV）、广义极大似然（Generalised Maximum Likelihood, GML）、均方误差（Mean Square Error, MSE）及信号自由度（Signal）等统计参数来确定。当 GCV 较小，且 Signal 小于站点的一半，则认为拟合过程已找到最优光滑参数，所选模型适用于插值，曲面拟合效果较好。

3.2 技术路线

技术路线如图 1 所示。数据集制作的主要过程包括：（1）相对湿度站点数据预处理。首先对来源于国家气象科学数据中心的 2000–2015 年逐月的气象站点相对湿度原始数据进行质量检查。总体来说，黄河流域九省区观测数据质量较好，数据较为完备。主要进行了数据异常值检测与处理，针对数据缺测或无观测任务的气象站点进行了剔除。将处理后的数据存储为 ASCII 格式，为插值做准备；

（2）插值方案优选。设计了 4 组插值方案如表 2 所示，并以 2015 年的相对湿度数据进行模型测试。根据 GCV 值，发现当自变量为经度、纬度、高程，独立协变量为高程，

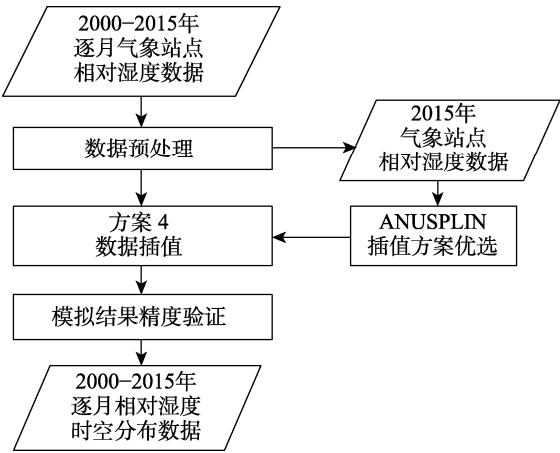


图 1 数据集研发技术路线图

样条次数为 3 次时插值方案最优；（3）数据插值。以优选方案对 2000–2015 年逐月气象站点相对湿度进行空间插值；（4）模拟结果的精度验证。主要采用偏差和相对误差指标，通过站点观测值和模拟值的比较，进行结果验证；（5）形成 2000–2015 年逐月相对湿度时空分布数据集。

表 2 ANUSPLIN 插值模型方案及 GCV

编号	自变量	独立协变量	样条次数	GCV
1	经度、纬度	—	2	13.5
2	经度、纬度、高程	高程	2	13.4
3	经度、纬度	高程	3	12.2
4	经度、纬度、高程	高程	3	12.1

4 数据结果与验证

4.1 数据集组成

数据集包括两个文件夹。文件夹“Studyarea”为研究区边界，数据格式为.shp。“9PYRB_RHU”包含 2000-2015 年逐月的相对湿度时空分布数据，数据格式为.tif。数据集共包含 775 个数据文件，相对湿度时空分布文件的文件名包含其时相信息，如“9PYRB_RHU200001.tif”为 2000 年 1 月的黄河流域九省区的相对湿度数据。

4.2 黄河流域九省区相对湿度空间分布

以 2015 年 3、6、9、12 月为例，展示了黄河流域九省区相对湿度的空间分布特征，如图 2 所示。鉴于篇幅，其他月份的相对湿度空间分布图未展示。黄河流域九省区不同月份的相对湿度差异明显，夏秋两季（6、9 月）的相对湿度较高、春冬（3、12 月）两季相对湿度较低，且冬季略高于春季。相对湿度不同月份的变化与区域年内气候条件密切相关，与降水呈正相关，与气温呈负相关^[14]。在空间分布上，四个月的相对湿度均表现为东南高、西北低与流域内降雨空间分布一致^[15]。

4.3 数据结果验证

依据黄河流域九省区及周边气象站 2015 年 1–12 月的相对湿度实测数据与插值结果，选用偏差 MBE 及相对误差 Er 进行验证，其计算公式如下：

$$MBE = \frac{1}{N} \sum_{i=1}^N (P_i - O_i) \tag{3}$$

$$Er = \frac{P_i - O_i}{O_i} \times 100\% \tag{4}$$

式中， P_i 为插值结果， O_i 为气象站点观测数据。 $|MBE|$ 及 $|Er|$ 越接近 0 则说明插值越接近观测值，插值结果越可靠。

如图 3 所示，以验证年份为例， $|MBE|$ 值介于 0.00–0.10， $|Er|$ 介于 3.6%–5.6%，说明插值数据精度较高，模拟值与观测值较为吻合。

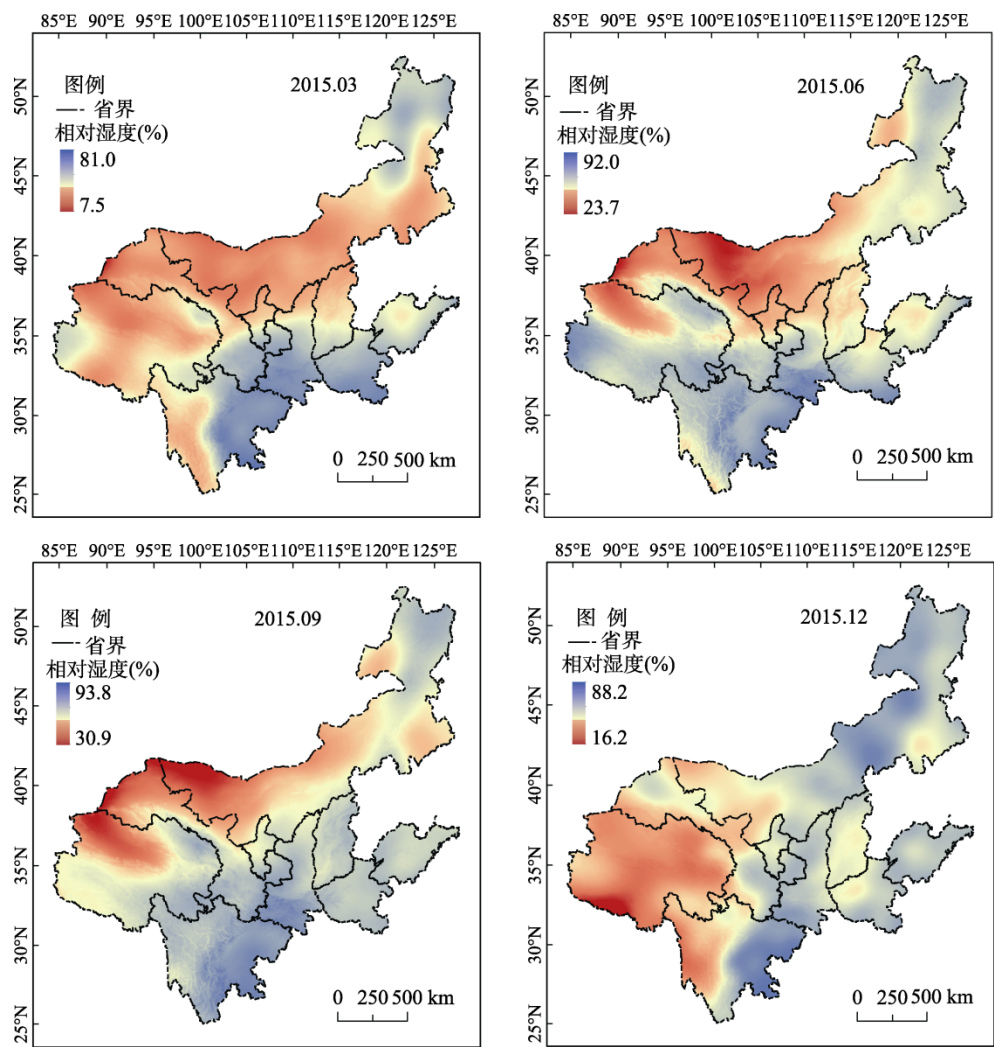


图 2 2015 年不同月份黄河流域九省区相对湿度空间分布图

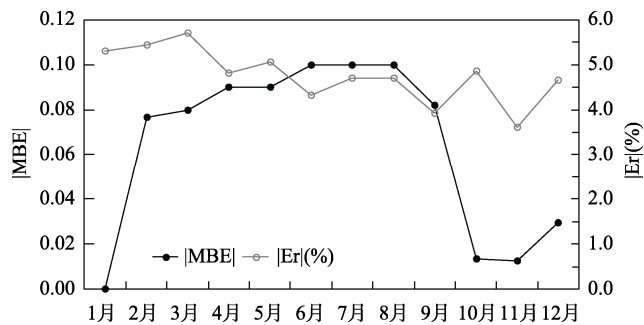


图 3 2015 年黄河流域九省区相对湿度插值精度验证图

5 讨论和总结

本数据集以黄河流域九省区及周边的站点观测相对湿度数据为插值数据，选用三次样条函数并以高程作为独立协变量进行空间插值，通过精度验证，插值结果可靠。

数据集提供的 2000–2015 年月尺度的 1-km 空间分辨率相对湿度数据,有助于揭示黄河流域九省区相对湿度变化的时空特征,可为该区域生态环境质量评估与治理提供可靠的数据支撑。

插值结果显示黄河流域九省区夏秋季的相对湿度较高,而冬春两季较低,这与区域降水的时间变化格局较为一致。在空间分布上,整个区域的相对湿度呈现出东南高、西北低的空间分布特征。在数据插值过程中,本文测试了 4 种插值方案,发现对于相对湿度而言,以高程、经度、纬度作为自变量,高程作为独立协变量,插值方案最优。但也应该注意,最优方案可能会随气象指标或区域不同存在差异。因此,在利用 ANUSPLIN 平台进行气象指标插值时,需要进行最优插值方案的确定。本数据集不足之处为,当以高程作为独立协变量时,插值范围受限于高程范围,因山东省北部沿海地区无高程数据,本数据集在相应区域存在缺失,但缺失数据不足三十五万分之一,对全域数据的使用影响不大。

作者分工: 蔡红艳对数据集生产进行总体设计及数据分析与论文撰写工作;蒋啸负责数据论文的撰写、相对湿度数据的整理与处理;杨小唤负责数据论文的完善。

利益冲突声明: 本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献

- [1] 刘明春,杨晓玲,殷玉春等.武威市相对湿度气候特征及预报[J].干旱区研究,2012,29(4): 654–659.
- [2] Du, J., Cooper, F., Fueglistaler, S. Statistical analysis of global variations of atmospheric relative humidity as observed by AIRS [J]. *Journal of Geophysical Research Atmosphere*, 2012, 117: D12315.
- [3] 刘玉英,李宇凡,谢今范等.东北高空湿度变化特征及其与地面气温和降水的关系[J].地理科学,2016,36(4): 628–636.
- [4] Eck, T. F., Holben, B. N., Kim, J., *et al.* Influence of cloud, fog, and high relative humidity during pollution transport events in South Korea: aerosol properties and PM_{2.5} variability [J]. *Atmospheric Environment*, 2020, 232: 1–16.
- [5] 张卓,王维和,王后茂等.相对湿度对吸水性气溶胶指数的影响[J].遥感学报,2019,23(6): 1177–1185.
- [6] 索南看卓,任国玉,贾文茜等.武汉城市相对湿度气候学特征与长期变化趋势[J].气候与环境研究,2018,23(6): 715–724.
- [7] 黄建平,张国龙,于海鹏等.黄河流域近 40 年气候变化的时空特征[J].水利学报,2020,51(9): 1048–1058.
- [8] 黎云云.气候和土地利用变化下流域干旱评估——传播-驱动-预测研究[D].西安:西安理工大学,2018.
- [9] Wu, D. Hazy weather research in China in the last decade: a review [J]. *Acta Scientiae Circumstantiae*, 2012, 32(2): 257–269.
- [10] 习近平.在黄河流域生态保护和高质量发展座谈会上的讲话[J].求是,2019,(20): 1–5.
- [11] 蔡红艳,蒋啸,杨小唤.黄河流域九省区 1-km 分辨率月度相对湿度数据集(2000–2015) [J/DB/OL]. 全球变化数据仓储电子杂志, 2021. <https://doi.org/10.3974/geodb.2021.03.12.V1>. <https://cstr.science.org.cn/CSTR:20146.11.2021.03.12.V1>.
- [12] 全球变化科学研究数据出版系统.全球变化科学研究数据共享政策[OL]. <https://doi.org/10.3974/dp.policy.2014.05> (2017 年更新).
- [13] 刘志红, Li, L. T., McVicar, T. R., 等.专用气象数据空间插值软件 ANUSPLIN 及其应用[J].气象,2008,34(2): 92–100.
- [14] 刘荣璐,李宝富,廉丽姝.1960–2015 年西北干旱区相对湿度时空变化与气候要素的定量关系[J].水土保持研究,2020,27(6): 233–239, 246.
- [15] 邵晓梅,严荣昌,魏红兵.基于 Kriging 插值的黄河流域降水时空分布格局[J].中国农业气象,2006,27(2): 65–69.