



文献 DOI:
10.11922/csdata.170.2016.0112
数据 DOI:
10.11922/sciencedb.319

文献分类: 地球科学

收稿日期: 2016-11-14
开放同评: 2016-12-13
发表日期: 2017-03-31

2000 ~ 2012 年全国气温和降水 1 km 网格空间插值数据集

王军邦^{1*}, 王居午², 叶辉¹, 刘亚¹, 何洪林¹

摘要: 随着高分辨率卫星遥感技术的快速发展及其在生态学中的广泛应用, 与时空分辨率相匹配的气象空间数据, 成为生态系统过程模拟及变化驱动力分析的重要数据, 也是全球气候变化背景下对空间异质性较高的陆地生态系统进行长期时空动态变化及机制研究的关键资料。为了得到与中分辨率成像光谱仪 (MODIS) 陆表产品时空分辨率相一致的气象空间数据, 进而为估算中国陆地生态系统植被净初级生产力等提供气象输入数据, 本研究采用 2000 ~ 2012 年的全球日气候历史数据网络 (Daily Global Historical Climatology Network-Daily, GHCN-D) 和中国气象局国家气象信息中心制作的“中国地面气候资料日值数据集”, 选取日降水和日均温两个要素, 对其进行读取、合并、检查、统计、空间插值批处理代码的生成等操作, 最后由 ANUSPLINE 软件插值生成 1 km 空间网格栅格数据。本栅格数据的公开共享和下载服务, 可为我国陆地生态系统的时空变化研究提供数据支持。

关键词: 气温; 降水; 插值; 1 km 网格; 中国

数据库 (集) 基本信息简介

数据库 (集) 中文名称	2000 ~ 2012 年全国气温和降水 1 km 网格空间插值数据集		
数据库 (集) 英文名称	An interpolated temperature and precipitation dataset at 1-km grid resolution in China (2000 – 2012)		
联系人	王军邦 (jbwang@igsnr.ac.cn)		
数据作者	王军邦、王居午、叶辉、刘亚、何洪林		
数据时间范围	2000 ~ 2012 年	地理区域	中国大陆地区
空间分辨率	1 km 网格	数据量	5.1 GB
数据格式	ArcGIS 32 位浮点型 (.flt)		
数据服务系统网址	http://www.cern.org.cn/data/initDRsearch?classcode=DPAPER , http://www.sciencedb.cn/dataSet/handle/319		
基金项目	中国科学院科技服务网络计划 (STS 计划, KFJ-SW-ST-167) 国家科技攻关计划 (2016YFC0500203) 国家自然科学基金项目 (31270520)		
数据库 (集) 组成	数据集由两部分组成, 数据量为 5.1 GB; 其一为 2000 ~ 2012 年的降水栅格数据集; 其二为 2000 ~ 2012 年的气温栅格数据集。文件命名分别为 PRCP、TEM。其中: 1: PRCP 是 2000 ~ 2012 年的降水栅格数据集, 数据量为 2.58 GB; 2: TEM 是 2000 ~ 2012 年的气温栅格数据集, 数据量为 2.51 GB。		

引言

气象观测数据是开展陆地生态系统长期变化研究的基础资料。随着遥感技术的发展, 生态系统相关研究已经由传统的样地单点或斑块尺度, 发展为应用空间分辨率 30 m 到 10 km 及以上的空间数据的研究^[1-3]。开展长期动态和空间格局变

1. 中国科学院地理科学与资源研究所, 生态系统网络观测与模拟重点实验室, 北京 100101;
2. 东北农业大学, 资源与环境学院, 哈尔滨 150030
* 联系人 (Email: jbwang@igsnr.ac.cn)

化研究, 需要同时在时间和空间尺度上与之匹配的气象数据。现有的空间气象数据主要是基于全球的气候研究, 缺少与高时空分辨率的遥感数据相匹配的气象空间数据, 对区域气候变化的影响分析带来困难, 因此高时空分辨率的气象数据十分重要^[4-6]。

本研究以 2000 ~ 2012 年全球日气候历史数据网络 (Daily Global Historical Climatology Network-Daily, GHCN-D) 和中国气象局国家气象信息中心制作的“中国地面气候资料日值数据集”为基础, 选取日降水和日均温两个要素, 通过中国科学院地理科学与资源研究所生态系统网络观测与模拟重点实验室开发的《基于 Matlab 的国际主要气象站点数据空间插值预处理系统》(软件登记号: 2014SR142569, 软件分类号: 30200-7500) 和目前国际上应用较多的澳大利亚国立大学开发的商业软件 ANUSPLIN 气象数据空间插值方法, 生成空间网格为 1 km 的气温和降水栅格数据。本数据与中分辨率成像光谱仪 (MODIS) 陆表产品时空分辨率相一致, 能够为估算中国陆地生态系统植被净初级生产力等提供基础气象数据, 为我国陆地生态系统的时空变化研究提供数据支持。

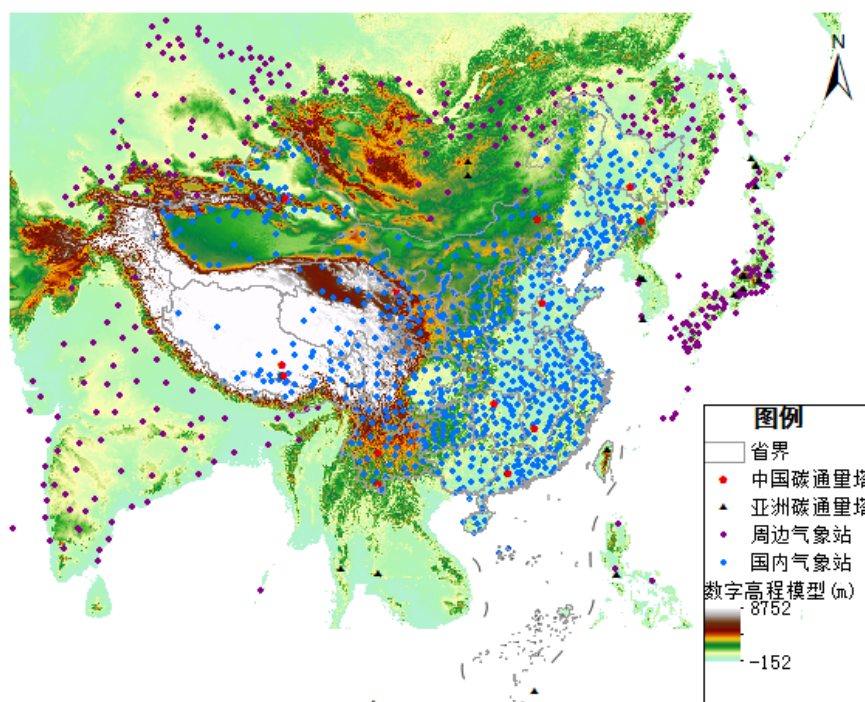


图 1 研究区数字高程模型及气象台站分布图

1 数据采集和方法

本文所用气象站点数据包括来自 GHCN-D 和中国气象局国家气象信息中心制作的“中国地面气候资料日值数据集”。其中前者有站点数 345 个, 后者共有 753 个地面气象观测站 (图 1), 包含日降水和日均温两个要素。通过《基于 Matlab 的国际主要气象站点数据空间插值预处理系统》将 GHCN-D 和中国气象局的数据进行标准化, 数字高程模型 (Digital elevation models, DEM) 使用中国国际科学数据服务平台的“中国 90 米分辨率高程数据”, 再用 ArcGIS 软件将 DEM 数据转换成 ANUSPLIN 软件能识别的 ASCII 格式的数据 (图 2)。

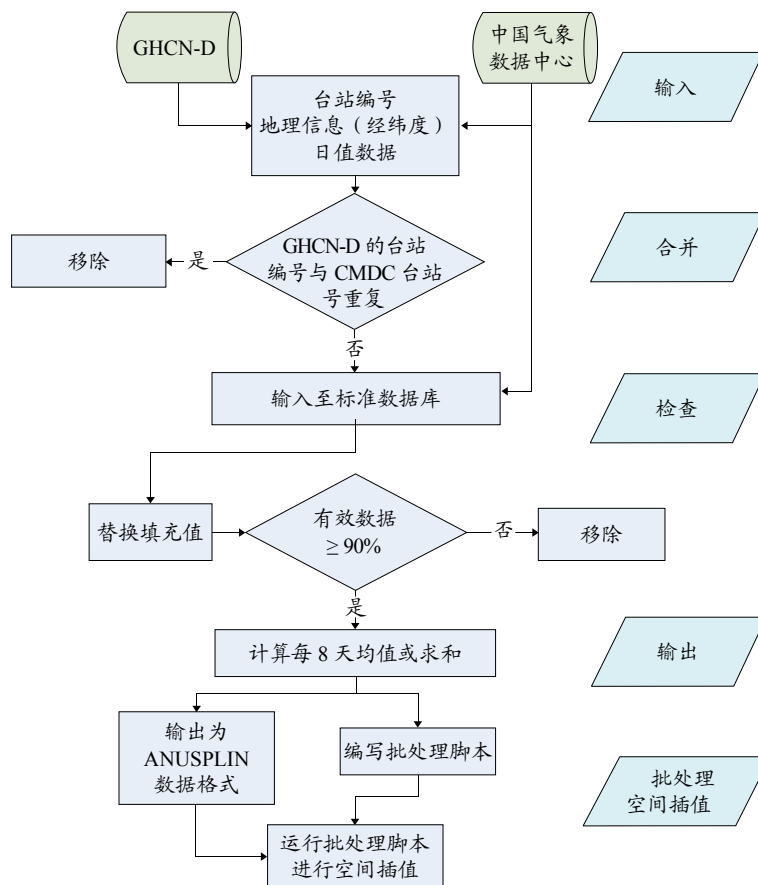


图 2 数据生成流程图

1.1 ANUSPLIN 原理

气象要素空间栅格化的核心是空间内插技术。ANUSPLIN 软件是采用薄盘光滑样条法进行转换分析和对多变量数据进行插值的工具，现已广泛应用于澳大利亚和欧美各国^[7-10]。这个软件通过提供综合的统计分析、数据诊断和空间分布标准误差来完成整个进程，同时还支持多数据输入和多拟合曲面输出。该软件由 8 个程序模块组成，由于本研究所使用资料的台站数少于 2000 个，因而只需执行 SPLINA（薄盘光滑样条函数）和 LAPGRD（进行局部薄盘光滑样条表面估值，并计算贝叶斯标准误差）等主要程序模块。ANUSPLIN 软件的核心算法是局部薄盘光滑样条法^[11]，其理论统计模型为：

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

其中， z_i 是位于空间 i 点的因变量， $f(x_i)$ 是关于 x_i 的未知光滑函数， x_i 为 d 维样条独立变量， b^T 是关于 y_i 的 p 维系数， y_i 为 p 维独立协变量， e_i 是期望值为 0 且方差为 $w_i \sigma^2$ 的自变量随机误差， w_i 为已知的相对误差方差， σ^2 为通常未知的所有数据点上的误差方差常数， N 是观测值的个数。

当式（1）缺少第 2 项，即没有协变量（ $p=0$ ）时，模型可以简化为普通的薄盘光滑样条模型；而当式（1）缺少第 1 项时，模型变为简单的多元线性回归模型，实际使用中 ANUSPLIN 软件不允许出现这种状况。函数 f 和系数 b 由最小二乘估计来确定：

$$\sum_{i=1}^N \left[\frac{z_i - f(x_i) - b^T y_i}{w_i} \right]^2 + \rho J_m(f) \quad (2)$$

其中, $J_m(f)$ 为函数 f 的粗糙度测度函数, 即函数 f 的 m 阶偏导; ρ 为正值光滑参数, 当 $\rho \rightarrow 0$ 时, 函数 f 接近于精确内插式; 当 $\rho \rightarrow +\infty$ 时, 函数 f 接近于一个最小二乘多项式, 取决于粗糙度 m 。光滑参数的值通常取决于拟合曲面的最小预测误差, 由广义交叉验证 (Generalized cross validation, GCV) 来确定。而 GCV 的计算是逐个移去数据点, 利用在同样的 ρ 下的剩余数据点来估算该点的残差。这在 ANUSPLIN 的输出日志文件中有记录。

1.2 数据处理流程

本数据集的产生主要包括 5 个部分, 分别是数据的读取、合并、数据检查、统计、批处理代码的生成, 最后是 ANUSPLIN 的插值。主要输入数据包括 GHCN-D 台站信息、中国国家气象台站信息、DEM 数据、GHCN-D 数据、中国地面气候资料日值数据。读取不同数据源的日气象要素数据后, 合并为统一的标准格式。数据的检查是提取各个要素真实值, 将观测数据的样本数小于 90% 的站点作为无效站点剔除, 然后运用反距离权重插值法 (IDW) 对有效站点的缺值进行插补, 从而保证插值过程和插值结果的可信度, 弥补了气象台站分布稀疏不均的不足。本数据集基于日尺度的降水和气温, 统计为每 8 天均值 (其中降水为每 8 天累加), 最后合成为年尺度的降水和气温。主要数据来源包括:

(1) DEM 数据采集

DEM 数据是 ASCII 格式的数据, 本研究将中科院国际科学数据服务平台发布的、根据美国航空航天局的 SRTM GRID 数据处理得到的成品数据, 转化成研究要求的数据, 并根据研究范围的大小, 对数据进行重新裁剪。

(2) 中国地面气候资料日值数据

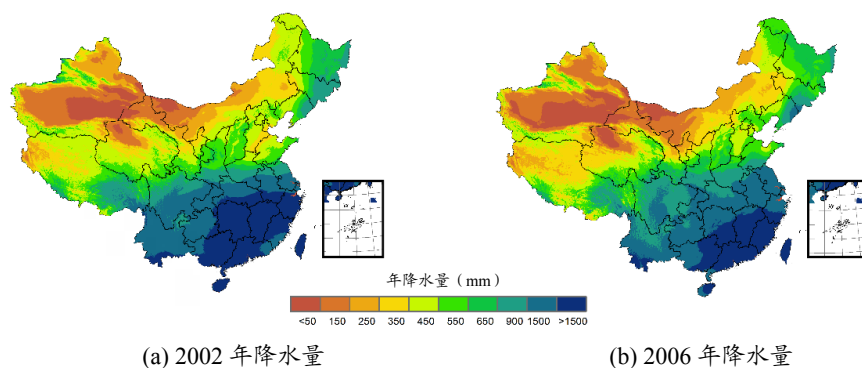
来源于各省、市、自治区气候资料处理部门逐月上报的《地面气象记录月报表》的信息化资料。

(3) GHCN-D 数据

GHCN-D 数据通过访问 <ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily> 下载获取。

2 数据样本

2000 ~ 2012 年全国 1 km 空间分辨率气温和降水栅格数据集的命名遵循如下规则: PRCP_YYYY.flr 和 TEM_YYYY.flr 通过文件名, 能够获取关于此数据文件的相应信息。例如 PRCP_2006.flr、TEM_2006.flr 中, PRCP 表示全国 1 km 空间分辨率降水栅格数据产品; TEM 表示全国 1 km 空间分辨率气温栅格数据产品; YYYY 表示年份。气温与降水数据样本的展示如图 3, 其中 a 和 b 分别是 2002 年和 2006 年的年降水量, c 和 d 分别是 2002 年和 2006 年的年平均气温。



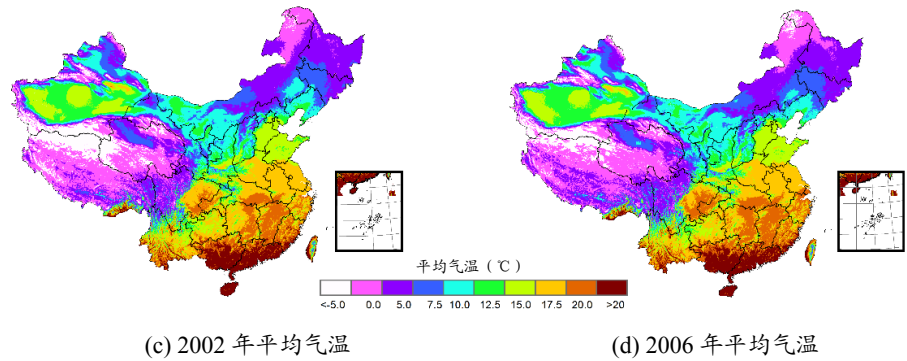


图 3 基于气象台站观测数据空间插值为 1km 网格的降水与气温空间分布示例

3 数据质量控制和评估

为了验证本数据集的精度，我们采用了 7 个亚洲通量观测台站 2000 年以来的气象要素 (<http://asiaflux.net>) 作为验证数据 (图 4，其中 a 和 c 是空间插值数据和碳通量塔观测值间的散点图及其线性拟合结果，b 和 d 是它们的频数分布图)。这弥补了传统交叉验证的缺陷，保证了数据的可信度。根据插值结果和验证数据对比，插值结果和通量站点显著线性相关，其中降水 $R^2=0.67$, $p < 0.01$ ，均方根误差 $RMSE=18.50$ mm。降水的累积分布图表明，在各降水梯度上，插值的降水和站点数据的分布一致，说明插值结果能较好地反映出真实的降水水平。温度 $R^2=0.94$ ， $p < 0.01$ ，插值结果能解释真实结果的 94%。

图 5 为中国地面气温和降水 $0.5^\circ \times 0.5^\circ$ 格点数据集 (V 2.0) 与本研究所有年份的 6 月 22 日每 8 天的数据均值结果对比。除降水在东部沿海地区有明显的差异以外，降水和气温数值在区域分布上整体一致。

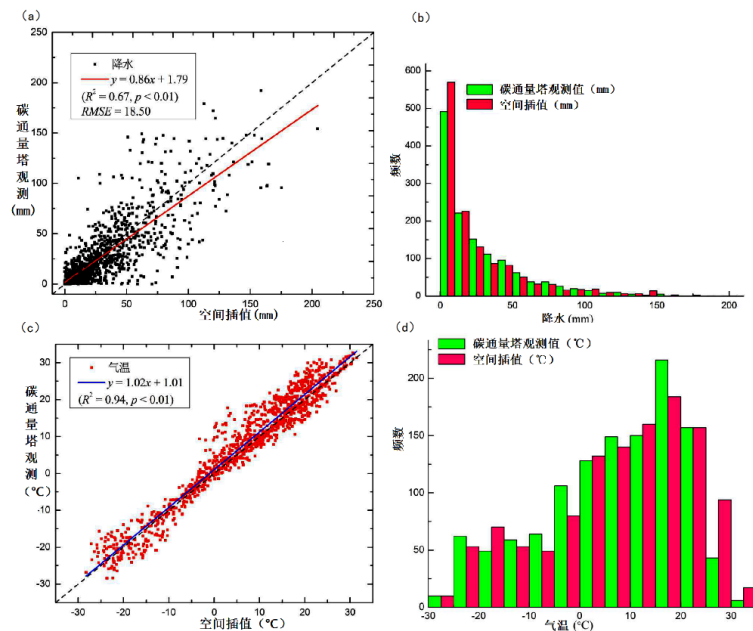


图 4 基于气象台站观测空间插值数据与碳通量塔上测定的降水和气温对比验证的结果

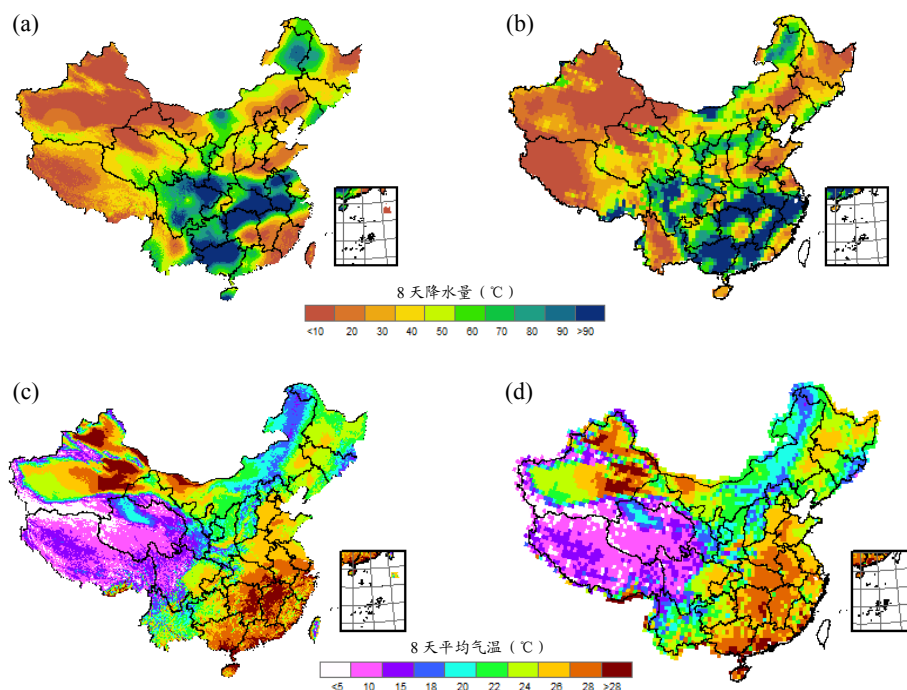


图 5 本研究插值的 1 km 网格数据 (a, c) 与中国气象局插值的 $0.5^{\circ} \times 0.5^{\circ}$ 网格降水和气温数据集 (b, d, 缺少台湾数据) 对比

4 数据使用方法和建议

2000 ~ 2012 年全国气温和降水 1 km 网格空间插值数据集通过 CERN 综合中心数据资源服务网站 (<http://www.cnern.org.cn>) 或 Science Data Bank (<http://www.sciencedb.cn/dataSet/handle/319>) 提供下载服务。用户登录系统后, 在首页点击“数据论文数据”图标或在“数据资源”栏目选择“数据论文数据”, 进入相应页面后可以下载完整数据。本数据集是在 ArcGIS 环境下制作生成的浮点型二进制栅格数据, 即 ArcGIS 的浮点型格式, 这种格式的数据除以 .flt 为后缀的 32 位浮点型数据文件外, 还包括同文件名的分别以 .hdr 和 .prj 为后缀的 ArcGIS 格式的数据头文件和投影文件, 可在 ArcGIS 环境直接调用或以 Matlab 等软件编程批量读取和分析。

致谢

本数据原始台站观测数据由中国气象局气象数据共享网提供, 中国生态系统研究网络和中国科学院计算机网络信息中心大数据技术与应用发展部提供了数据存放平台和技术支撑, 在此表示衷心感谢。

作者分工职责

王军邦 (1974—), 男, 青海湟中人, 博士, 副研究员, 研究方向为生态系统生态学。主要承担工作: 总体方案设计。

王居午 (1991—), 男, 河北邯郸人, 硕士研究生, 主要研究方向为土地利用与规划。主要承担工作: 基础数据整理收集及论文撰写工作。

叶辉 (1985—), 男, 江西九江人, 博士研究生, 研究方向为碳循环模型和生态信息学。主要承担工作: 基础数据处理、1 km 分辨率的降水和气温空间插值。

刘亚 (1990—), 男, 江西赣州人, 硕士研究生, 主要研究方向为土地生态与 3S 综合应用。主要承担工作: 基础数据整理收集。

何洪林 (1971—), 男, 湖南冷水江人, 博士, 研究员, 主要研究方向生态信息学。主要承担工作: 总体方案设计。

参考文献

- [1] Chapin III F, Matson P, Mooney H. Principles of terrestrial ecosystem ecology[M]. New York: Springer Science & Business Media, 2011: 369 – 397.
- [2] Bellard C, Bertelsmeier C, Leadley P, et al. Impacts of climate change on the future of biodiversity[J]. Ecology Letters, 2012, 15(4): 365 – 377.
- [3] Frey S, Lee J, Melillo J, et al. The temperature response of soil microbial efficiency and its feedback to climate[J]. Nature Climate Change, 2013, 3(4): 395 – 398.
- [4] Schuur E, McGuire A, Schädel C, et al. Climate change and the permafrost carbon feedback[J]. Nature, 2015, 520(7546): 171 – 179.
- [5] Hulme M. A 1951 – 80 global land precipitation climatology for the evaluation of general circulation models[J]. Climate Dynamics, 1992, 7(2): 57 – 72.
- [6] Fan Y, Dool H. A global monthly land surface air temperature analysis for 1948 – present[J]. Journal of Geophysical Research: Atmospheres, 2008, 113(D1).
- [7] Hutchinson M. The application of thin plate smoothing splines to continent-wide data assimilation[J]. BMRC research report, 1991 (27): 104 – 113.
- [8] Hutchinson M. Interpolation of rainfall data with thin plate smoothing splines. Part I: Two dimensional smoothing of data with short range correlation[J]. Journal of Geographic Information and Decision Analysis, 1998, 2(2): 139 – 151.
- [9] Hutchinson M. Interpolation of rainfall data with thin plate smoothing splines. Part II: Analysis of topographic dependence[J]. Journal of Geographic Information and Decision Analysis, 1998, 2(2): 152 – 167.
- [10] Hutchinson M. ANUSPLIN Version 4.2 User Guide[M]. Canberra: Center for Resource and Environmental Studies, the Australian National University, 2001.
- [11] 尚宗波, 高琼, 杨莫安. 利用中国气候信息系统研究年降水量空间分布规律 [J]. 生态学报, 2001, 21(5): 689 – 694.

引用数据

- (1) 王军邦, 王居午, 叶辉, 等. 2000 ~ 2012 年全国气温和降水 1 km 网格空间插值数据集 [DB/OL]. Science Data Bank. DOI: 10.11922/sciencedb.319.

An interpolated temperature and precipitation dataset at 1-km grid resolution in China (2000 – 2012)

Wang Junbang, Wang Juwu, Ye Hui, Liu Ya, He Honglin

ABSTRACT With the rapid development of satellite remote sensing techniques with high spatial resolution and their wide applications in the field of ecology, high-resolution meteorological spatial data have become popular in ecosystem process simulation systems, especially on spatiotemporal dynamics of terrestrial ecosystems with high spatial heterogeneity in the background of global climate change. Spatial meteorological data with high resolution consistent with the land surface data from MODIS is essential to estimating the net primary productivity of vegetation in terrestrial ecosystems of China. The temperature and precipitation datasets with a grid resolution of 1 km were produced based on the observed data from the National Meteorological Information Center (NMIC) of China Meteorological Administration and the Daily Global Historical Climatology Network-Daily (GHCN-D) during 2000–2012. The stations were selected in a region larger than the research area and data quality was evaluated. The eight-day average and total values were calculated for temperature and precipitation, respectively. Finally, the ANUSPLIN software was used to interpolate the observed meteorological data to a grid with high resolution (1-km). Those two annual data products can be used to further study the spatiotemporal changes in terrestrial ecosystems of China.

KEYWORDS temperature; precipitation; interpolation; 1 km grid; China

引文格式：王军邦, 王居午, 叶辉, 等. 2000 ~ 2012 年全国气温和降水 1 km 网格空间插值数据集 [J/OL]. 中国科学数据, 2017, 2(1). DOI: 10.11922/csdata.170.2016.0112.