专题 黄土高原生态环境

中国典型流域地表特征与时间序列水文气象数据集（2003–2020 年）

马慧[1,2]，张凌[1*]，冯克庭[3,4]

1. 中国科学院西北生态环境资源研究院，冰冻圈科学与冻土工程全国重点实验室，兰州 730000
2. 中国科学院大学，北京 100000
3. 中国科学院地球环境研究所，黄土科学全国重点实验室，西安 710061
4. 中国科学院地球环境研究所，黄土科学数据中心，西安 710061

摘要：流域特征与水文气象数据集对于深化对水文水资源演变规律的认识、推动水文模型开发与性能评估等具有重要作用。本研究构建了一个覆盖中国约四分之一国土面积、涵盖 21 个典型流域的地表特征与时间序列水文气象数据集。利用自动出口重定位（Automatic Outlet Relocation, AOR）算法，并结合高分辨率流向数据，实现对流域边界的自动化划分，划分结果与水利部统计数据高度一致（R^2 = 0.99）。在此基础上，通过空间叠加与掩膜提取，获得了包括地形、土地利用、土壤、植被和水文气象在内的 25 个流域特征变量。这些特征变量大多来源于专门针对中国区域研发的高质量、高分辨率产品，融合了地面观测与遥感等多源数据，其精度优于同类全球产品。本数据集具有广泛的应用价值，可为大尺度流域的水文建模、水文影响评估及水资源管理等研究提供重要的数据支撑。

关键词：中国；数据集；流域；径流；水文气象；水资源

数据库（集）基本信息简介

数据库（集）名称	中国典型流域地表特征与时间序列水文气象数据集（2003–2020 年）
数据通信作者	张凌（zhanglingky@lzb.ac.cn）
数据作者	马慧、张凌、冯克庭
数据时间范围	2003–2020年
地理区域	地理范围为21°52′50″–45°33′33″N，79°37′50″–129°24′16″E
数据量	883.99 MB
数据格式	*.shp, *.xlsx，*tif
数据服务系统网址	https://doi.org/10.57760/sciencedb.28506
基金项目	国家自然科学基金（42271286）；中国科学院青年创新促进会（2023454）；甘肃省重点研究计划（23ZDKA0004）。
数据库（集）组成	本数据集由三个部分组成：（1）流域边界矢量图；（2）流域特征与时间序列水文气象数据集，包括径流、降水、气温、土地利用与覆被、地形、土壤、植被等要素；（3）流域特征变量栅格数据。

www.sciengine.com/CSD
ISSN 2096-2223
CN 11-6035/N

文献 CSTR：
32001.14.11-6035.csd.2025.0144.zh

文献 DOI：
10.11922/11-6035.csd.2025.0144.zh

数据 DOI：
10.57760/sciencedb.28506

文献分类：地球科学

收稿日期：2025-07-30
录用日期：2025-10-28
发表日期：2025-12-16

* 论文通信作者
张凌：zhanglingky@lzb.ac.cn

引　言

流域特征与水文气象数据集是开展水文水资源研究的基础，对于深入理解水文过程的动态演变及其长期变化趋势具有重要意义。通过整合时间序列水文气象数据（如河流径流、降水、气温、风速和潜在蒸散发）与流域地表特征信息（如地形、地质、土壤和土地覆盖），构建完整的流域综合数据集，不仅为大尺度水文水资源研究提供了新的视角，也有助于破解水资源管理中的关键科学问题，支撑区域水资源演变的定量分析及气候变化影响评估[1-3]。

在国际上，多个国家和地区基于大样本研究的流域特征与气象数据框架（CAMELS）构建了流域水文气象特征数据集，如澳大利亚（CAMELS-AUS）[4]、巴西（CAMELS-BR）[5]、瑞士（CAMELS-CH）[6]、德国（CAMELS-DE）[7]、丹麦（CAMELS-DK）[8]、法国（CAMELS-FR）[9]、英国（CAMELS-GB）[10]和印度（CAMELS-IND）[11]等。CAMELS（Catchment Attributes and Meteorology for Large-Sample Studies）数据集旨在提供覆盖大量流域的水文气象时间序列数据及其景观属性，包括气候指数、水文特征、土地覆盖、土壤和地质等要素[1,12-13]。其核心目标在于通过统一、开放的流域属性、气象数据和径流观测资料，为大样本流域水文学（Large-Sample Hydrology, LSH）研究提供数据支持，并确保广泛的空间代表性[10]。这些数据集不仅提供了高质量的水文气象时间序列数据，还系统整合了地形、土壤、植被等地表特征，为深入理解流域过程与区域异质性提供了重要基础。然而，在中国区域基于 CAMELS 框架的流域水文气象数据集仍相对较少[14]，尤其缺乏关键的径流观测数据。径流作为流域水文模拟与评估的核心变量，其缺失不仅削弱了气象因子与水文响应之间的耦合分析能力，也在很大程度上制约了流域水文建模、水资源评估与管理等方面的深入研究。因此，构建覆盖中国区域、融合径流观测的高质量 CAMELS 框架数据集，对于提升我国在流域水文过程模拟、水资源调控与气候变化影响评估等方面的研究水平，具有重要的科学价值和现实意义。

本研究构建了一个覆盖中国九大水系的 21 个典型流域的地表特征与时间序列水文气象数据集。利用自动出口重定位（AOR）算法工具，结合高分辨率流向数据，对 21 个流域的边界进行了自动划分，并利用水利部统计数据对划分结果进行验证。在此基础上，分别提取各流域包括地形、土地利用、土壤与植被，以及水文气象（包括径流、降水、气温、潜在蒸散发）等共计 25 个特征变量。本数据集可为大尺度流域水文建模、水文响应分析及水资源评估与管理等提供坚实的数据基础，具有重要的价值与应用潜力。

1　数据采集和处理方法

1.1　研究区概况

本研究所选的 21 个典型流域广泛分布于黄河、长江、松花江、珠江等中国主要河流水系，覆盖总面积约 240.8 万平方公里，占全国国土面积约四分之一，如图 1 所示。这些流域在气候类型、地形地貌及水文过程等方面差异显著，能够较好地反映中国主要的水文水资源特征，从而保证样本的典型性与代表性。

1.2　数据收集与处理

本研究数据涉及水文气象、土壤与植被、土地利用及地形四大类别，共包含 25 个变量。

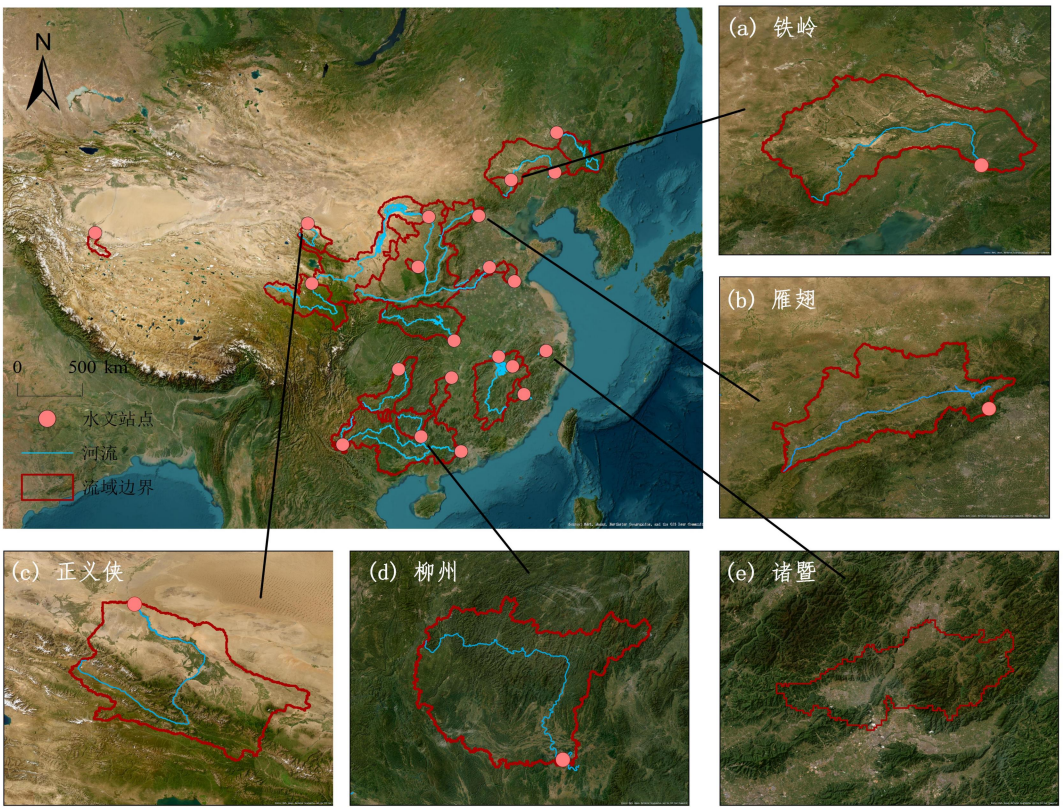


图 1　典型流域的地形特征与空间分布

Figure 1　Topographic features and spatial distribution of the typical catchments

1.2.1　水文气象数据

水文气象数据包括径流、降水量、最高气温、最低气温、平均气温及潜在蒸散发。其中，流域的径流数据主要来源于历年《中国河流泥沙公报》（http://www.mwr.gov.cn/sj/tjgb/zghlnsgb/），时间跨度为 2003–2020 年。气象数据包括降水、气温和潜在蒸散发（PET），来源于多源数据融合的中国高分辨多要素气象驱动产品（ChinaMet）[15]。该产品通过将气象站观测数据、卫星遥感数据和再分析产品融合，并结合气候学辅助插值方法与机器学习技术构建而成，空间分辨率为 1 公里，时间分辨率为日尺度。

1.2.2　土壤与植被数据

土壤数据主要包括土壤水力参数数据和土壤质地数据。土壤水力参数（包括土壤剖面深度和孔隙度）数据来源于中国土壤水力参数数据集（China Dataset of Soil Hydraulic Parameters, CSHP，http://globalchange.bnu.edu.cn/）[16]。该数据集基于中国范围内 8,979 个土壤剖面和 1:100 万中国土壤图开发，涵盖土壤剖面深度、孔隙度等关键水力学属性参数。CSHP 数据为时间不变特征，其空间分辨率与土壤图相一致（约 1 公里）。土壤质地数据来源于全球统一土壤数据库（Harmonized World Soil Database, HWSD），该数据集由联合国粮农组织（FAO）与国际应用系统分析研究所（IIASA）联合开发（https://iiasa.ac.at/models-tools-data/hwsd），提供全球范围内的土壤信息[17]。数据的空间分辨率为 30 角秒（约 1 公里），涵盖表层（0–30 厘米）和亚表层（30–100 厘米）两个土壤深度层次。土壤质地包括黏土（Clay）、粉土（Silt）和沙土（Sand）含量的空间分布。

植被动态变化采用归一化植被指数（NDVI）进行量化。植被指数数据（0.05°，日尺度）来源于先进的甚高分辨率辐射计（AVHRR）地表反射率产品[18]，由美国国家海洋和大气管理局环境信息中心（National Oceanic and Atmospheric Administration，NOAA）提供（https://www.earthdata.nasa.gov/sensors/avhrr）。AVHRR NDVI 数据自 1981 年起提供长时间序列，空间分辨率为 0.05°（约 5 公里），时间分辨率为日尺度。

1.2.3 土地利用数据

土地利用数据源于中国土地利用覆盖数据集（China Land Cover Dataset, CLCD, https://doi.org/10.5281/zenodo.5816591）[19]，该数据集基于 Landsat 遥感影像联合研制，覆盖 1990 年至 2021 年间的年度土地利用类型信息，空间分辨率为 30 米。灌溉数据集来源于中国灌溉耕地分布图（CIrrMap250）[20]，该数据集由 Zhang 等人基于多源遥感数据与统计调查数据，结合机器学习方法生成，时间范围为 2000 年至 2020 年，空间分辨率为 250 米（https://doi.org/10.6084/m9.figshare.24814293.v2）。

1.2.4 地形数据

地形数据包括高程和坡度两部分。高程数据来源于 ASTER 全球数字高程模型（ASTER Global Digital Elevation Model, ASTER GDEM），其原始空间分辨率为 90 米，从以下网址获取：https://asterweb.jpl.nasa.gov/gdem.asp。坡度数据基于处理后的高程数据，通过 ArcGIS 软件计算获得。

1.3 数据生产

本研究数据生产的流程图如图 2 所示，首先利用自动出口重定位（Automatic Outlet Relocation, AOR）算法工具，并结合 1 公里分辨率的流向数据及流域出口位置信息，对 21 个流域的边界进行自动划分。然后，基于流域边界，对栅格数据（包括气象、土壤和植被、土地利用与地形等）进行空间叠加与掩膜提取，统计各流域属性的平均值。最后，将流域尺度属性数据与《中国河流泥沙公报》发布的径流数据进行匹配，生成中国典型流域地表特征与时间序列水文气象数据集（2003–2020 年）。该数据集以 Excel 表格和栅格数据两种方式存储：前者汇总了流域尺度的地表特征和时间序列水文气象数据，后者汇总了流域不同属性的空间分布信息，空间分辨率为 1 公里，采用 WGS84 坐标系，时间范围覆盖 2003–2020 年。

1.3.1 流域划分

本研究通过自动出口重定位（Automatic Outlet Relocation, AOR）算法工具对 21 个流域的边界进行划分[21]。与传统的 ArcGIS 划分方法相比，AOR 算法在方法原理及大规模应用中具有显著优势。传统方法需要研究者反复试错以确定累积流量阈值来生成数字河网，并对流域出口进行手动定位，不仅效率低且主观性强。AOR 算法利用累积流量沿主河道的梯度信息，能够自动将出口点移动至最优河网位置：首先通过稳定格点进行主河道重定位，然后沿主河道和潜在支流追踪累积流量变化，找到靠近观测出口且满足流量增长条件的支流格点，并最终选取距离出口最近的格点作为最终出口。在美国本土 1,398 个流域的实验中，AOR 算法的划分准确率较 ArcGIS 重定位方法提升约 1.16 倍，计算效率提升约 10 倍。

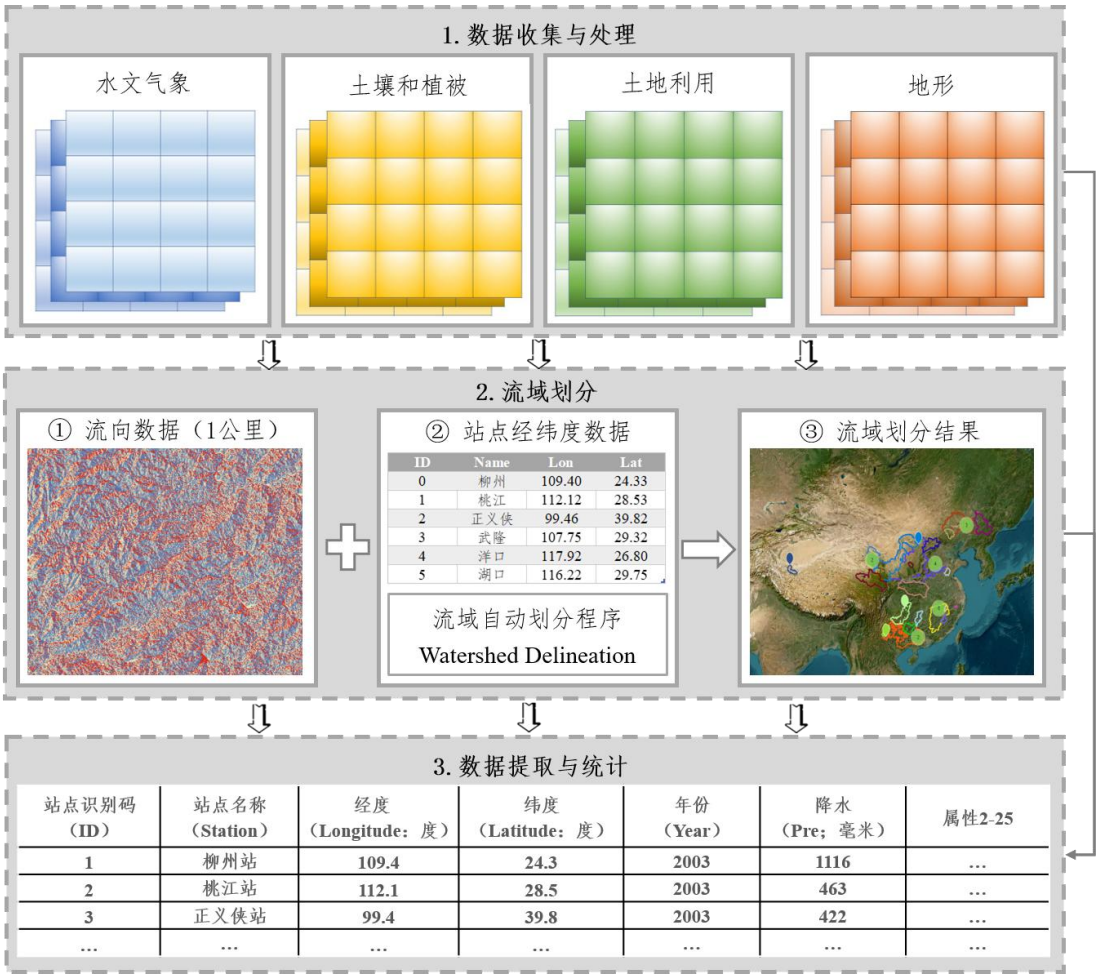


图 2　本研究数据生产的流程图

Figure 2　Flowchart of the data production in this study

利用 AOR 算法工具，并结合全球高分辨率（1 公里）流向数据[22]，对 21 个流域的边界进行自动划分。如图 3a 所示，划分结果与水利部统计的流域面积高度一致，决定系数（R^2）达 0.99，表明流域划分结果精度较高。图 3b 展示了 21 个流域的面积分布直方图，多数流域（67%）面积小于 10 万平方公里，流域最大面积达 80 万平方公里。

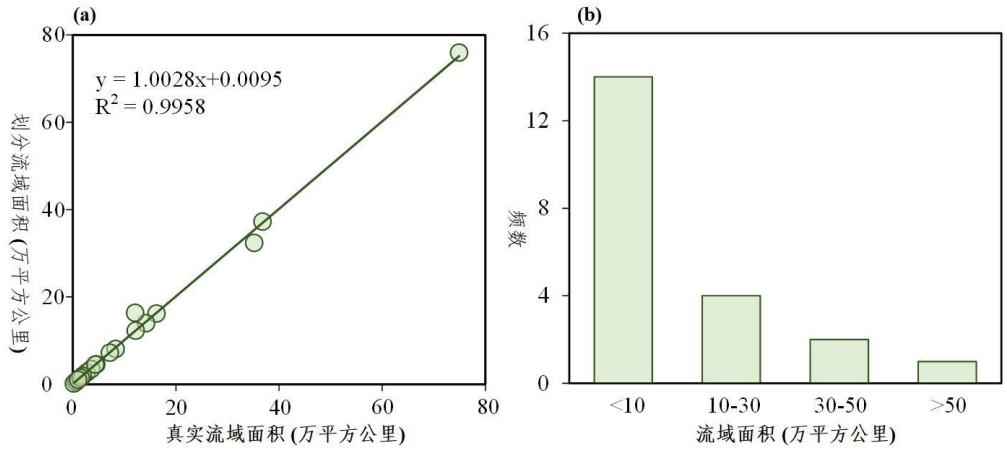


图 3　流域面积划分结果精度（a）及流域面积直方图（b）

Figure 3　Accuracy of catchment area delineation (a) and catchment area histogram (b)

1.3.2 流域变量特征的提取和统计

本研究基于流域边界共计算了 25 个特征变量，如表 1 所示。各变量的计算方法如下：（1）通过将径流量（单位：10^8 m^3/年）除以流域面积获得径流深（单位：mm/年）；（2）气象、土壤、植被和地形特征通过对流域范围内相应变量空间平均获得；（3）土地利用特征则采用加权平均的方法进行计算，见公式（1）。

$$Var（T）=\frac{\sum_{i=1}^{n} T_i A_i}{\sum_{i=1}^{n} A_i} \tag{1}$$

其中，$Var（T）$表示土地利用类型 T 的面积占流域总面积的比例；T_i 为流域中第 i 个网格单元土地利用类型所占比例；A_i 为网格单元面积；n 为该流域内所有网格单元的数量。

表 1　流域地表特征与水文气象变量描述

Table 1　Description of land surface characteristics and hydro-meteorological variables of catchments

数据类型	变量名	描述
水文气象	径流深（SflowD）	流域年平均径流深（毫米）
	径流量（SflowV）	流域年径流量（亿方）
	降水（Pre）	流域年平均降水量（毫米）
	潜在蒸散发（PET）	流域年平均潜在蒸散发量（毫米）
	最高气温（Tempmax）	流域年平均最高气温（℃）
	最低气温（Tempmin）	流域年平均最低气温（℃）
	气温（Tempmean）	流域年平均气温（℃）
土壤与植被	土壤剖面深度（PDEP）	流域平均土壤剖面深度（米）
	孔隙率（POR）	流域平均土壤孔隙率（%）
	归一化植被指数（NDVI）	流域年平均归一化植被指数（-）
	沙土（Sand）	流域平均沙土含量（%）
	粉土（Silt）	流域平均粉土含量（%）
	黏土（Clay）	流域平均黏土含量（%）
土地利用	荒地（Barren）	流域荒地面积占比（-）
	耕地（Cropland）	流域耕地面积占比（-）
	森林（Forest）	流域森林面积占比（-）
	草地（Grassland）	流域草地面积占比（-）
	不透水（Impervious）	流域不透水地表面积占比（-）
	灌木（Shrub）	流域灌木面积占比（-）
	积雪（Snowice）	流域积雪面积占比（-）
	水体（Water）	流域水体面积占比（-）
	湿地（Waterland）	流域湿地面积占比（-）
	灌溉耕地（CIrrMap）	流域灌溉耕地面积占比（-）
地形	海拔（Elevation）	流域平均海拔（米）
	坡度（Slope）	流域平均坡度（°）

2 数据样本描述

本数据集包含站点识别码（ID，用于识别矢量文件中各个流域在空间上的分布）、站点名称（Station）、经纬度（Longitude 和 Latitude）、年份（Year）、面积（Area），如表 2 所示，以及如表 1 所示的 25 个流域特征与水文气象变量，共计 31 个字段。图 4 展示了其中 4 个流域（柳州站、皇庄站、桃江站和高要站）的降水-径流变化关系。

表 2　数据集的字段含义描述表

Table 2　Field descriptions for the dataset

站点识别码（ID）	站点名称（Station）	经度（Longitude；度）	纬度（Latitude；度）	年份（Year）	面积（Area;平方公里）	属性 1-25
1	柳州站	109.4	24.33333	2003	45386	…
1	柳州站	109.4	24.33333	2004	45386	…
1	柳州站	109.4	24.33333	2005	45386	…
1	柳州站	109.4	24.33333	2006	45386	…
1	柳州站	109.4	24.33333	2007	45386	…
1	柳州站	109.4	24.33333	…	45386	…

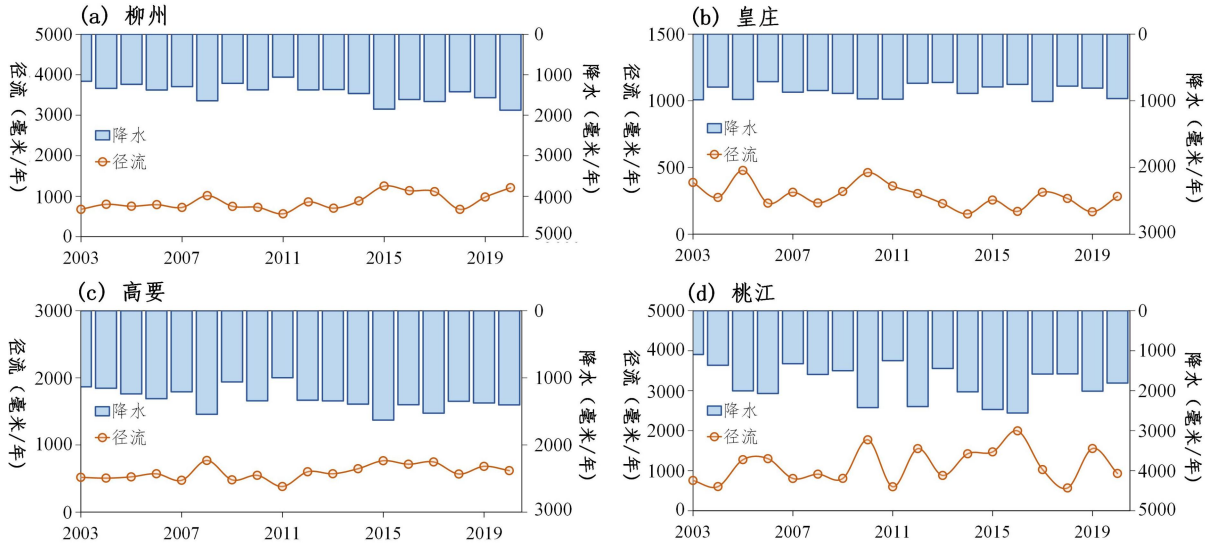


图 4　四个流域降水-径流变化关系

Figure 4　Precipitation–runoff relationships for four catchments

3 数据质量控制和评估

本数据集整合了《中国河流泥沙公报》[23]、多源数据融合的中国高分辨率多要素气象驱动产品[15]、土壤水力参数数据[16]、先进的甚高分辨率辐射计（AVHRR）植被指数产品[18]、中国灌溉耕地分布图[20]，以及 ASTER 全球数字高程模型等多个数据集。大部分特征变量来源于专门针对中国区域研发的高质量、高分辨率产品，融合了地面观测与遥感等多源数据，其精度优于同类全球产品。国外同

类产品通常面向全球开发，在中国区域使用的观测数据有限，难以准确反映中国复杂多样的流域特征[4-11]。与国内现有同类数据相比，本数据集不仅整合了流域气象、土地利用、地形等属性数据，还纳入了关键的径流观测数据。因此，本数据集在数据精度及内容的完整性和综合性方面均具有明显优势，可更好地支持大尺度水文模型模拟和水资源评估等应用[14]。

本研究使用的流域特征与水文气象数据的质量评估详见表 3。

表 3　中国典型流域地表特征与水文气象数据质量信息表

Table 3　Data quality information for land surface characteristics and hydrometeorological variables of the typical catchments in China

数据类型	数据名称	参数数据质量
水文气象	径流量数据	该数据采用我国主要流域（如黄河、长江）水利委员会，以及各省、市等水利部门的水文控制站实测径流量数据[23]。
	气象数据	该数据来源于多源融合的中国高分辨率多要素气象驱动产品 ChinaMet。水资源公报数据独立验证表明：该产品精度较高（R^2=0.92，RMSE=177 mm/年），不确定性小于其他同类产品[24]。
土壤与植被	土壤水力参数数据	该数据集是基于 8,979 个土壤剖面数据和《中国土壤图》以及土壤转移函数（PTFs）研制，是目前适用于中国陆面建模的“最优”土壤属性数据集[16]。该数据集的精度受限于源图及剖面观测的不足，未来可通过增加剖面样点的空间覆盖并优化传递函数模型来提升可靠性。
	植被指数数据	该数据来源于美国国家海洋和大气管理局（NOAA）发布的甚高分辨率辐射计（AVHRR）地表反射率产品。通过比较同一天观测的 AVHRR 和 MODIS 正下方校正后的地表反射率，计算得到 NDVI 的均方根差（RMSD）为 0.074，表明两者在大尺度植被监测中具有较好的一致性[18]。
	土壤质地数据	该数据来源于全球统一土壤数据库（Harmonized World Soil Database, HWSD）。该数据库是由国际应用系统分析研究所（IIASA）与联合国粮食及农业组织（FAO）联合建立的高分辨率土壤数据库，空间分辨率为 30 角秒[17]。
土地利用	土地利用覆盖数据	基于 5,463 个人工目视解译样本，CLCD 的整体分类精度达到 79.31 %。第三方样本验证表明，CLCD 在总体精度上优于其他已有产品，包括 MCD12Q1、ESACCI_LC、FROM_GLC 和 GlobeLand30[19]。数据受 Landsat5 时空覆盖不均衡和数据缺口的限制，未来可结合 Landsat MSS、Sentinel-2 与 Sentinel-1 等多源数据以提升产品分辨率和精度。
	灌溉耕地数据	基于 2 万多个参考点验证表明：中国时间序列灌溉耕地分布图（CIrrMap250）的总体精度为 0.79–0.88，优于其他已有产品[20]。该数据集受数据（尤其是统计数据）不确定性的影响，分辨率较粗，未来可引入更高分辨率遥感数据，并考虑作物类型的影响，提升产品精度。
地形	地形数据	该数据来源于日本经济产业省（METI）与美国国家航空航天局（NASA）联合发布的先进星载热发射与反射辐射计（ASTER）全球数字高程模型（GDEM），空间分辨率为 90 米[25]。

本研究构建的数据集目前包含 21 个典型流域，其覆盖面积仅占全国国土面积的四分之一，尤其

在西北干旱区和西部地区的流域数据十分稀缺。同时，数据集中的径流数据为年尺度，虽然适用于流域径流趋势分析及区域水量平衡研究，但难以对干旱和洪水等极端水文事件以及径流的季节变化进行分析。为此，未来将进一步从《水资源公报》和《中华人民共和国水文水利年鉴》等公开资料中收集流域径流观测数据，扩大样本覆盖范围，提升流域数据的空间代表性和时间分辨率。此外，未来将进一步整合河流泥沙、水质等相关数据，丰富数据库的内容，为流域水资源管理和生态环境评估提供更加全面的数据支撑。

4 数据使用方法和建议

本数据集包含流域边界及流域地表特征与时间序列水文气象数据（2003–2020）。所有流域边界合并为一个矢量文件（文件名：CN_21basins.shp），可在 ArcMap 软件中直接加载查看流域空间分布。流域地表特征与水文气象数据以 Excel 文件形式提供，所有流域数据汇总在同一个工作表中（文件名：CN_21basins_attributes_hydrometerology.xlsx）。此外，数据集还包含 GeoTIFF 格式的流域特征变量与水文气象空间分布数据。已在中国科学数据银行和黄土科学数据中心对数据使用方法（示例代码）进行了详细说明。

本数据集可应用的范围主要包括以下三个方面：（1）结合土地利用变化、水库和灌溉等信息，开展流域径流变化的归因分析，揭示径流变化的驱动机制[26]；（2）结合地下水储量变化数据，基于水量平衡原理，可利用该数据库反演流域实际蒸散发量，用于验证遥感蒸散发产品精度[27-29]；（3）数据库也可为大尺度陆面或水文模型的参数标定与性能评估提供参考[30]。

在使用过程中需注意：（1）本数据集在西北干旱区的流域较少，更适用于湿润和半湿润地区的分析和研究；（2）流域地表特征与水文气象汇总数据（Excel 格式）与变量空间分布数据（GeoTIFF 格式）可通过站点识别码（ID）及站点名称进行匹配。例如，ID 为 1、站点名称为“柳州”的流域降水空间分布数据（2003 年）对应路径为：“CN_basins\流域特征变量与水文气象空间分布数据\1 气象数据\1 降水数据\tif 年数据\1_柳州_station\1_柳州_station_prec_2003.tif”；（3）流域变量的空间分布数据中不在流域边界范围内的值为 nan，因此在用 Python 或者 MATLAB 进行分析或统计时需要对其进行剔除。

作者分工职责

马慧（1999—），女，甘肃省天水市人，硕士研究生，研究方向为水资源。主要承担工作：数据处理与论文初稿撰写。

张凌（1988—），男，四川省巴中市人，博士，副研究员、项目研究员，研究方向为灌溉水资源可持续利用。主要承担工作：论文构思、论文指导与修改。

冯克庭（1980—），男，宁夏回族自治区中卫市人，硕士，正高级工程师，研究方向为遥感、地学大数据应用、科学数据共享。主要承担工作：论文指导、文章校验。

参考文献

[1] ADDOR N, DO H X, ALVAREZ-GARRETON C, et al. Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges[J]. Hydrological Sciences Journal, 2020, 65(5): 712–725.

DOI:10.1080/02626667.2019.1683182.

[2] GUPTA H V, PERRIN C, BLÖSCHL G, et al. Large-sample hydrology: a need to balance depth with breadth[J]. Hydrology and Earth System Sciences, 2014, 18(2): 463 – 477. DOI:10.5194/hess-18-463-2014.

[3] KRATZERT F, KLOTZ D, SHALEV G, et al. Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets[J]. Hydrology and Earth System Sciences, 2019, 23(12): 5089 – 5110. DOI:10.5194/hess-23-5089-2019.

[4] FOWLER K J A, ACHARYA S C, ADDOR N, et al. CAMELS-AUS: hydrometeorological time series and landscape attributes for 222 catchments in Australia[J]. Earth System Science Data, 2021, 13(8): 3847 – 3867. DOI:10.5194/essd-13-3847-2021.

[5] CHAGAS V B P, CHAFFE P L B, ADDOR N, et al. CAMELS-BR: hydrometeorological time series and landscape attributes for 897 catchments in Brazil[J]. Earth System Science Data, 2020, 12(3): 2075 – 2096. DOI:10.5194/essd-12-2075-2020.

[6] HÖGE M, KAUZLARIC M, SIBER R, et al. CAMELS-CH: hydro-meteorological time series and landscape attributes for 331 catchments in hydrologic Switzerland[J]. Earth System Science Data, 2023, 15(12): 5755 – 5784. DOI:10.5194/essd-15-5755-2023.

[7] LORITZ R, DOLICH A, ACUÑA ESPINOZA E, et al. CAMELS-DE: hydro-meteorological time series and attributes for 1582 catchments in Germany[J]. Earth System Science Data, 2024, 16(12): 5625 – 5642. DOI:10.5194/essd-16-5625-2024.

[8] LIU J, KOCH J, STISEN S, et al. CAMELS-DK: hydrometeorological time series and landscape attributes for 3330 Danish catchments with streamflow observations from 304 gauged stations[J]. Earth System Science Data, 2025, 17(4): 1551 – 1572. DOI:10.5194/essd-17-1551-2025.

[9] DELAIGUE O, GUIMARÃES G M, BRIGODE P, et al. CAMELS-FR dataset: a large-sample hydroclimatic dataset for France to explore hydrological diversity and support model benchmarking[J]. Earth System Science Data, 2025, 17(4): 1461 – 1479. DOI:10.5194/essd-17-1461-2025.

[10] COXON G, ADDOR N, BLOOMFIELD J P, et al. CAMELS-GB: hydrometeorological time series and landscape attributes for 671 catchments in Great Britain[J]. Earth System Science Data, 2020, 12(4): 2459 – 2483. DOI:10.5194/essd-12-2459-2020.

[11] MANGUKIYA N K, KUMAR K B, DEY P, et al. CAMELS-IND: hydrometeorological time series and catchment attributes for 228 catchments in Peninsular India[J]. Earth System Science Data, 2025, 17(2): 461 – 491. DOI:10.5194/essd-17-461-2025.

[12] ADDOR N, NEWMAN A J, MIZUKAMI N, et al. The CAMELS data set: catchment attributes and meteorology for large-sample studies[J]. Hydrology and Earth System Sciences, 2017, 21(10): 5293 – 5313. DOI:10.5194/hess-21-5293-2017.

[13] NEWMAN A J, CLARK M P, SAMPSON K, et al. Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance[J]. Hydrology and Earth System Sciences, 2015, 19(1): 209 – 223. DOI:10.5194/hess-19-209-2015.

[14] HAO Z, JIN J, XIA R L, et al. CCAM: China catchment attributes and meteorology dataset[J]. Earth System Science Data, 2021, 13(12): 5591 – 5616. DOI:10.5194/essd-13-5591-2021.

[15] 张凌, 胡英屹, 赵彦博, 等. 多源数据融合的中国高分辨多要素气象驱动产品（ChinaMet）[DS/OL]. 国家冰川冻土沙漠科学数据中心 (http://www.ncdc.ac.cn), 2025. https://www.doi.org/10.12072/ncdc.nieer.db6722.2025. [ZHANG L, HU Y Y, ZHAO Y B, et al. ChinaMet: A High-resolution Multi-element Meteorological Driving Dataset for China via Multi-source Data Fusion[DS/OL]. National Cryosphere Desert Data Center, 2025. https://www.doi.org/10.12072/ncdc.nieer.db6722.2025.. DOI:10.12072/ncdc.nieer.db6722.2025.]

[16] DAI Y J, SHANGGUAN W, DUAN Q Y, et al. Development of a China dataset of soil hydraulic parameters using pedotransfer functions for land surface modeling[J]. Journal of Hydrometeorology, 2013, 14(3): 869–887.

[17] FAO, IIASA. Harmonized World Soil Database version 2.0[EB/OL]. (2023-09-18) [2025-12-08]. https://doi.org/10.4060/cc3823en.

[18] VERMOTE E,. NOAA CDR Program. NOAA Climate Data Record (CDR) of AVHRR Normalized Difference Vegetation Index (NDVI), Version 5[EB/OL]. (2019 - 07 - 11) [2025-12-08]. https://doi.org/10.7289/V5ZG6QH9.

[19] YANG J, HUANG X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019[J]. Earth System Science Data, 2021, 13(8): 3907–3925. DOI:10.5194/essd-13-3907-2021.

[20] ZHANG L, XIE Y H, ZHU X F, et al. CIrrMap250: annual maps of China’s irrigated cropland from 2000 to 2020 developed through multisource data integration[J]. Earth System Science Data, 2024, 16(11): 5207–5226. DOI:10.5194/essd-16-5207-2024.

[21] XIE J X, LIU X M, BAI P, et al. Rapid watershed delineation using an automatic outlet relocation algorithm[J]. Water Resources Research, 2022, 58(3): e2021WR031129. DOI:10.1029/2021WR031129.

[22] EILANDER D, VAN VERSEVELD W, YAMAZAKI D, et al. A hydrography upscaling method for scale-invariant parametrization of distributed hydrological models[J]. Hydrology and Earth System Sciences, 2021, 25(9): 5287–5313. DOI:10.5194/hess-25-5287-2021.

[23] 中华人民共和国水利部. 中国河流泥沙公报[R/OL]. 北京: 中国水利水电出版社 (http://www.mwr.gov.cn/sj/tjgb/zghlnsgb), 2020. [MINISTRY OF WATER RESOURCES, PEOPLE'S REPUBLIC OF CHINA. The River Sediment Bulletin of China[R/OL]. Beijing: China Water & Power Press, 2020. http://www.mwr.gov.cn/sj/tjgb/zghlnsgb.]

[24] HU Y Y, ZHANG L. Added value of merging techniques in precipitation estimates relative to gauge-interpolation algorithms of varying complexity[J]. Journal of Hydrology, 2024, 645: 132214. DOI:10.1016/j.jhydrol.2024.132214.

[25] MINISTRY OF ECONOMY, TRADE, AND INDUSTRY (METI) OF JAPAN, UNITED STATES NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA). ASTER Global Digital Elevation Map Announcement[EB/OL]. (2019 - 08 - 05) [2025 - 12 - 08]. https://asterweb.jpl.nasa.gov/gdem.asp.

[26] TANG X P, FU G B, ZHANG S L, et al. Attribution of climate change and human activities to streamflow variations with a posterior distribution of hydrological simulations[J]. Hydrology and Earth System Sciences, 2022, 26(20): 5315–5339. DOI:10.5194/hess-26-5315-2022.

[27] PALMROTH S, KATUL G G, HUI D F, et al. Estimation of long-term basin scale evapotranspiration

from streamflow time series[J]. Water Resources Research, 2010, 46(10): 2009WR008838. DOI:10.1029/2009WR008838.

[28] XIONG J H, ABHISHEK, XU L, et al. ET-WB: water-balance-based estimations of terrestrial evaporation over global land and major global basins[J]. Earth System Science Data, 2023, 15(10): 4571–4597. DOI:10.5194/essd-15-4571-2023.

[29] DILE Y T, AYANA E K, WORQLUL A W, et al. Evaluating satellite-based evapotranspiration estimates for hydrological applications in data-scarce regions: a case in Ethiopia[J]. Science of the Total Environment, 2020, 743: 140702. DOI:10.1016/j.scitotenv.2020.140702.

[30] BECK H E, PAN M, LIN P R, et al. Global fully distributed parameter regionalization based on observed streamflow from 4, 229 headwater catchments[J]. Journal of Geophysical Research: Atmospheres, 2020, 125(17): e2019JD031485. DOI:10.1029/2019JD031485.

论文引用格式

马慧，张凌，冯克庭. 中国典型流域地表特征与时间序列水文气象数据集（2003–2020 年）[J/OL]. 中国科学数据, 2025, 10(4). (2025-12-16). DOI: 10.11922/11-6035.csd.2025.0144.zh.

数据引用格式

马慧，张凌，冯克庭. 中国典型流域地表特征与时间序列水文气象数据集（2003–2020 年）[DS/OL]. V3. Science Data Bank, 2025. (2025-09-29). DOI: 10.57760/sciencedb.28506.

A dataset of land surface characteristics and time-series hydrometeorological data for typical catchments in China (2003–2020)

MA Hui[1,2], ZHANG Ling[1*], FENG Keting[3,4]

1. State Key Laboratory of Cryospheric Science and Frozen Soil Engineering, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, P.R. China

2. University of Chinese Academy of Sciences, Beijing 100000, P.R. China

3. State Key Laboratory of Loess Science, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an 710061, P.R. China

4. Loess Science Data Center, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an 710061, P.R. China

*Email: zhanglingky@lzb.ac.cn

Abstract: Catchment characteristics and hydro-meteorological datasets play a critical role in deepening our understanding of hydrological processes and water resource dynamics, as well as in supporting the development and performance evaluation of hydrological models. In this study, we present a

comprehensive dataset that integrates land surface attributes and time series hydro-meteorological variables across 21 representative catchments, covering approximately one-quarter of China's terrestrial area. Catchment boundaries were delineated automatically using the Automatic Outlet Relocation (AOR) algorithm based on high-resolution flow direction data. The delineated catchment areas agree well with official statistics from the Ministry of Water Resources (R^2 = 0.99). Using these catchment boundaries, we derived 25 catchment attributes, including topography, land use, soil properties, vegetation, and hydro-meteorological indicators, through spatial overlay and mask extraction. Most of these attributes were derived from high-quality, high-resolution products specifically developed for China, integrating multi-source data such as ground observations and remote sensing, and offering higher accuracy than comparable global products. This dataset has broad application potential and can greatly support large-scale hydrological modeling, hydrological impact assessment, and water resource management.

Keywords: China; dataset; catchments; runoff; hydrometeorology; water resources

Dataset Profile

Title	A dataset of land surface characteristics and time-series hydrometeorological data for typical catchments in China (2003–2020)
Data corresponding author	ZHANG Ling (zhanglingky@lzb.ac.cn)
Data authors	MA Hui, ZHANG Ling, FENG Keting
Time range	2003–2020
Geographical scope	21°52′50″ – 45°33′33″N, 79°37′50″ – 129°24′16″E
Data volume	883.99 MB
Data format	*.shp, *.xlsx,*tif
Data service system	https://doi.org/10.57760/sciencedb.28506
Sources of funding	The National Natural Science Foundation of China (42271286); the Youth Innovation Promotion Association of Chinese Academy of Sciences (2023454); the Key Research Program of Gansu Province (Grant No. 23ZDKA0004).
Dataset composition	This dataset consists of three components: (1) a vector map of catchment boundaries; (2) a dataset of catchment characteristics and time series hydro-meteorological variables, including runoff, precipitation, temperature, land use and land cover, topography, soil, and vegetation; and (3) gridded data of catchment features.