



文献 CSTR:
32001.14.11-6035.csd.2023.0096.zh



文献 DOI:
10.11922/11-6035.csd.2023.0096.zh
数据 DOI:
10.57760/sciencedb.08027

文献分类: 地球科学

收稿日期: 2023-04-23
开放同评: 2023-05-08
录用日期: 2023-08-15
发表日期: 2023-12-25

* 论文通信作者
熊俊峰: jfxiong@niglas.ac.cn

2001–2020 年中国东北区域土壤水蚀数据集

吴瀚逸^{1,2,3,4}, 熊俊峰^{1,4*}, 侯渲^{4,5}, 林晨⁴, 许金朵^{4,5}, 马荣华^{4,5}

1. 可持续发展大数据国际研究中心, 北京 100094
2. 北京师范大学遥感科学国家重点实验室, 北京 100875
3. 北京师范大学环境演变与自然灾害教育部重点实验室, 北京 100875
4. 中国科学院南京地理与湖泊研究所, 南京 210008
5. 国家科技资源共享服务平台, 国家地球系统科学数据中心湖泊-流域分中心, 南京 210008

摘要: 东北黑土区是我国重要的粮食产区和生态保障区, 近年来由于气候变化和过度垦殖, 水土流失加剧, 严重威胁东北地区粮食安全和生态环境。高时空分辨率土壤水蚀模拟对实现东北地区农业持续发展和土地退化监测等可持续发展目标有着重要意义。本数据集基于 Google Earth Engine (GEE) 云平台, 融合多源遥感数据, 解构 RUSLE 模型, 对模型各因子进行算法组合优选。结合水利部公开的东北区域主要水文监测站的年输沙量和输沙模数数据与 RUSLE 模型估算的土壤侵蚀模数对比验证, 以时序相关性、均方根误差以及平均绝对误差三个精度作为评价指标选取最优因子算法组合, 得到东北区域 250 m 土壤水蚀模数估算结果。本数据集可以较好地表征 2001–2020 年东北区域土壤水蚀模数空间分布及时间序列变化, 为东北地区的水土流失防治以及土壤侵蚀评估提供有效参考。

关键词: 土壤水蚀模数; 模型优选; 东北区域; RUSLE

数据库(集)基本信息简介

数据库(集)名称	2001–2020 年中国东北区域土壤水蚀数据集
数据作者	吴瀚逸, 熊俊峰, 侯渲, 林晨, 许金朵, 马荣华
数据通信作者	熊俊峰 (jfxiong@niglas.ac.cn)
数据时间范围	2001–2020 年
地理区域	东北地区: 黑龙江省、吉林省、辽宁省、河北省东北部(承德、秦皇岛市)、内蒙古东部(锡林郭勒盟、赤峰、通辽、兴安盟、呼伦贝尔市)
空间分辨率	250 m
数据量	2.71 GB
数据格式	*.tif
数据服务系统网址	https://doi.org/10.57760/sciencedb.08027
基金项目	可持续发展大数据国际研究中心主任青年基金(CBAS2022DF014); 国家自然科学基金青年科学基金项目(42201400); 江苏省自然科学基金资助项目(BK20221058); 全球变化背景下中国区域湖泊响应数据库(CAS-WX2021SF-0306); 中国科学院南京地理与湖泊研究所科学数据中心(CAS-WX2022SDC-SJ05)。
数据库(集)组成	本数据集包含 20 个*.tif 文件。文件以对应年份命名。

引言

土壤侵蚀是土地退化的重要原因，引发的一系列生态环境问题已严重威胁到区域的粮食生产与水土保持^[1-3]。东北黑土区是我国重要的粮食产区 and 生态保障区。近年来，由于气候变化和过度垦殖，导致了严重的水土流失，威胁到东北地区粮食安全和水土资源安全^[4-6]。因此对于土壤侵蚀进行准确的监测、评估以及侵蚀量计算，有助于水土保持措施的配置与优化。

由于精准的实地测量和高精度土壤水蚀数据获取方式在大区域尺度内难以实施且耗费的人力物力较大，因此利用基于遥感和 GIS 开发的土壤侵蚀模型对于土壤侵蚀进行定量模拟计算，是当前土壤侵蚀计算以及水土流失动态监测的重要途径之一。在大尺度区域的土壤水蚀研究中最为常用的模型有通用土壤侵蚀模型 (Universal Soil Loss Equation, USLE) 和其修订模型 RUSLE (Revised Universal Soil Loss Equation)，以及经改进后适用于中国土壤侵蚀特征的 CSLE (Chinese Soil Loss Equation) 模型。USLE/RUSLE 系列侵蚀模型将复杂的土壤侵蚀过程解构为土壤特性、地形、植被覆盖、土地管理以及降雨和径流等多个影响因子进行计算，可以较好地对土壤侵蚀进行定量描述，故本研究采用 RUSLE 模型进行水蚀量计算。如表 1 所示。

表 1 部分土壤水蚀数据产品

Table 1 Partial data products of soil water erosion

序号	文献	研究区	模型	空间分辨率	年份
1	ZHANG 等 ^[7]	三北工程区	RUSLE	1 km	2000–2020
2	王莉等 ^[8]	雅鲁藏布江、年楚河、拉萨河 (一江两河) 流域	RUSLE	1 km	2001–2015
3	章文波 ^[9]	亚欧重点区域 20 国	CSLE	300 m	2015
4	章文波 ^[10]	青藏高原	CSLE	300 m	1992/2005/2015
5	杨勤科 ^[11]	巴基斯坦 15 个小流域	CSLE	2.5 m	2019

以上土壤水蚀数据产品在大尺度土壤水蚀估算研究中通常利用单一土壤水蚀影响因子算法，对于具有不同地理分异规律的地区，缺少区域性优化，时空分辨率不够精细；并且，对于数据精度缺少直接验证。同时考虑到东北地区与全国其他区域存在地理环境差异，基于其他地理条件开发的模型算法在东北地区的适用性有待验证，因此本数据集将当前研究中常用的模型针对东北区域做出优选与结果的直接验证。

针对当前土壤水蚀模型应用过程中区域性优化和实测数据验证缺失的短板，本研究基于多源遥感数据，解构 RUSLE 模型，对模型每个因子选取多种经典算法进行排列组合，结合水文站数据直接验证筛选最优算法组合，模拟东北地区 2001–2020 年土壤水蚀模数时空分布，建立东北地区土壤水蚀时空数据集，对实现东北地区农业持续发展和土地退化监测等可持续发展目标有着重要意义。

1 数据采集和处理方法

1.1 数据生产流程

本文拟采用技术路线如下首先基于多源遥感数据，利用 RUSLE 模型，对模型每个因子选取多种经典算法进行排列组合计算，结合水文站数据直接验证筛选最优算法组合，得到东北地区不同三级

流域分区的土壤水蚀模数估算结果。本文的技术路线如图 1 所示，包含以下处理步骤：

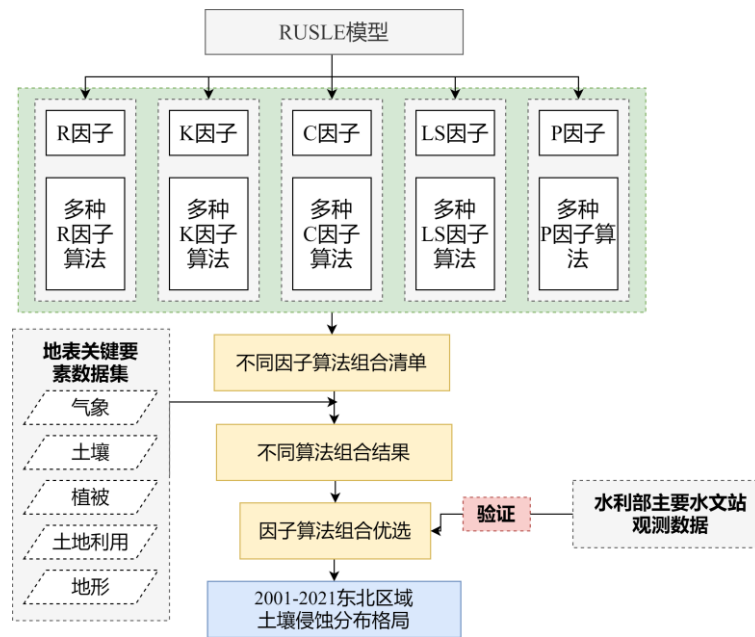


图 1 技术路线图

Figure 1 Technical flow chart

1.2 模型及算法

RUSLE 模型是美国农业部基于通用土壤侵蚀模型 USLE 修正的一种适用区域较广的土壤水蚀模数估算模型^[12-13]。该模型综合考虑了降雨、土壤理化性质、坡度坡长、地表植被覆盖以及水土保持措施共 5 个因子。其基本公式如下：

$$A = R \times K \times LS \times C \times P \quad (1)$$

其中，A 代表土壤侵蚀模数，单位为 $\text{t ha}^{-1} \text{y}^{-1}$ ；R 是降水侵蚀因子，单位为 $\text{MJ mm ha}^{-1} \text{h}^{-1} \text{y}^{-1}$ ，代表降水对于土壤颗粒的冲刷侵蚀作用，作为土壤水蚀的主要动力因子；K 是土壤可蚀性因子，代表土壤本身理化性质在土壤侵蚀过程中的影响作用；LS 代表地形因子，包括坡度（S）和坡长（L）两个因子；C 表示地表植被和作物覆盖因子，表示在植被覆盖与种植耕作对于土壤侵蚀的影响，该因子与植被冠层、地表植被覆盖以及其他地表植被状况有关；P 代表水土保持措施因子，代表着梯田、沟渠、围栏等水土保持措施对土壤水蚀的影响，值域为 0–1 之间，越接近 1 表示水土保持措施越少。

1.3 数据来源

本研究使用的降水数据是来自美国航空航天局（National Aeronautics and Space Administration, NASA）和日本宇宙航空研究开发机构（Japan Aerospace Exploration Agency, JAXA）合作开展全球卫星降水计划（Global Precipitation Measurement mission, GPM），其利用多传感器多卫星多算法结合卫星网络和雨量计反演得到更高精度的降水数据，其能够提供全球范围基于微波和红外的半小时的降水数据产品^[14]。GPM 的数据产品主要分为：Level-1、Level-2 和 Level-3。本研究使用的是 Level-3 的数据是在 Level-2 的基础上对固定时间和空间尺度进行插值，数据完整性和一致性较高，包括逐半小时降水以及逐月降水数据。土壤信息数据是国际应用系统分析研究所（International Institute for Applied Systems Analysis, IIASA）及联合国粮农组织（Food and Agriculture Organization of the United

Nations, FAO) 倡导下建立的世界土壤数据库 (Harmonized World Soil Database version 1.2, HWSD V1.2) [15]。其中, 中国地区数据源为 1995 年全国第二次土壤普查由南京土壤所提供的 1:1 000 000 土壤数据, 数据格式为 grid, 投影为 WGS84, 采用的土壤分类系统主要为 FAO-90。DEM 高程数据是由美国宇航局 NASA 制作由美国地质联邦调查局 (United States Geological Survey, USGS) 免费发布的高程数据, 包括北纬 60° 到南纬 56° 之间陆地表面较高质量的海拔数据 [16]。数据来源详见表 2。

表 2 数据来源表

Table 2 Data source table

序号	数据名称	空间分辨率	时间分辨率	数据网址
1	Global Precipitation Measurement v6 (GPM)	0.1°	逐半小时/逐月	https://disc.gsfc.nasa.gov/datasets?keywords=GPM&page=1
2	Harmonized World Soil Database v1.2 (HWSD)	1 km	1995 年全国第二次土壤普查	https://iiasa.ac.at/models-tools-data/hwsd
3	Terra Vegetation Indices MOD13A2 V6.1	1 km	逐旬	https://ladsweb.modaps.eosdis.nasa.gov/search/order/1/MOD13A2--6
4	SRTM Digital Elevation Data v4 (SRTM DEM)	90 m	2008 年	https://www.usgs.gov/products/data/data-releases
5	中国科学院地理科学与资源研究所中国多时期土地利用/土地覆盖遥感监测数据集	30 m	逐 5 年	https://www.resdc.cn

1.4 数据处理

1.4.1 R 因子

首先对逐半小时的 GPM 降水数据进行日合成得到逐日的降水数据, 根据章文波等 [17] 提出的 R 因子计算算法, 由逐日降水数据计算得到逐年的 R 因子时空分布, 记为 R1; 其次, 根据周伏建和黄炎和 [18] 提出的降水因子计算模型, 由 GPM 逐月降水数据进行计算得到逐年降水侵蚀因子, 记为 R2。具体计算公式参照表 3。

1.4.2 K 因子

基于 HWSD 数据选取了土壤不同粒径颗粒组成含量以及土壤有机碳含量, 通过目前土壤水蚀研究中常用的 EPIC [19]、Shrazis [20] 以及 Torri.D [21] 模型计算表示土壤的可蚀性, 得到土壤可蚀性因子空间分布, 依次记为 K1, K2 和 K3。具体计算公式参照表 3。

1.4.3 C 因子

选取 MODIS (Moderate-resolution Imaging Spectroradiometer) 提供的 MOD13A2 植被指数数据的归一化差值植被指数 (Normalized Difference Vegetation Index, NDVI) 数据, 首先利用蔡崇法和丁树文 [22] 提出的分段的非线性转换模型计算得到 C 因子空间分布, 记为 C1; 其次利用 Van Leeuwen and Sammons [23] 提出的带参数的指数非线性模型计算得到 C 因子空间分布 (其中经验性参数取值为 $\alpha = 2.5$; $\beta = 1$), 记为 C2。具体计算公式参照表 3。

1.4.4 LS 因子

坡度坡长因子 LS 代表着地形条件对于土壤侵蚀的影响。本研究基于 SRTM（Shuttle Radar Topography Mission）90 米分辨率的数字高程模型数据，利用符素华等^[24]的坡度坡长法计算出坡度坡长因子，记为 LS1。具体计算公式参照表 3。

1.4.5 P 因子

水土保持措施因子代表着地表对于土壤水蚀的保护措施对于土壤水蚀的影响。本研究基于文献调研统计以及中国 5 年一期的土地利用数据，学者们在东北区域常用的水土保持措施因子的取值进行赋值计算。详细取值如表 4，参考文献见数据集中的补充材料。

表 3 R、K、C 和 LS 因子

Table 3 R, K, C and LS factors

算法因子	编号	文献	计算方法	参数说明
R	R1	章文波等, 2002 ^[17]	$R = \alpha \sum_{j=1}^k (D_j)^\beta; \quad \alpha = 21.586 \times \beta^{-7.1891}; \quad \beta = 0.8363 + \frac{18.114}{P_{d12}} + \frac{24.455}{P_{y12}};$	R 是降雨侵蚀力因子, α , β 是模型参数, 根据降雨特征进行估算, D_j 是降水侵蚀力计算时段内的第 j 天的降雨量, k 是时段内的具体天数, P_{d12} 是时段内日降水量大于 12mm 的日平均降雨量, P_{y12} 是日均降雨量大于 12 mm 的年均降雨量, r_i 是日降雨量。
	R2	周伏建和黄炎和, 1995 ^[18]	$R = 17.02 \times \sum_{i=1}^{12} (-1.5527 + 0.1792r_i)$	
C	C1	蔡崇法和丁树文, 2000 ^[22]	$C = \begin{cases} 1, & FVC = 0 \\ 0.6508 - 0.3436 \times \lg FVC, & 0 < FVC \leq 78.3\%; \\ 0, & FVC > 78.3\% \end{cases}$ $FVC = NDVI_S$ $= (NDVI - NDVI_{low}) / (NDVI_{high} - NDVI_{low})$	C 是植被覆盖度因子, FVC 是植被覆盖度, NDVI 是归一化差值植被指数, $NDVI_{low}$ 是纯裸土像元的 NDVI 值, $NDVI_{high}$ 是纯植被像元的 NDVI 值, α , β 是模型参数, 决定了 NDVI 曲线的形状。
	C2	Van Leeuwen & Sammons, 2004 ^[23]	$C = \exp \left[-\alpha \times \frac{NDVI}{(\beta - NDVI)} \right]$	

算法因子	编号	文献	计算方法	参数说明
K	K1	Williams, 1990 ^[19]	$K = \left\{ 0.2 + 0.3 \times \exp \left[-0.0256 \times Sd \times \left(1 - \frac{Sl}{100} \right) \right] \right\} \times \left[\frac{Sl}{Cl + Sl} \right]^{0.3} \times \left\{ 1.0 - 0.25 \times \frac{Coc}{Coc + \exp(3.72 - 2.95 \times Coc)} \right\} \times \left\{ 1.0 - \frac{0.7 \times (1 - Sd/100)}{1 - Sd/100 + \exp[-5.51 + 22.9 \times (1 - Sd/100)]} \right\}$	K是土壤可蚀性因子, Sd是砂粒的重量百分比, Sl是粉粒的重量百分比, Cl是黏粒的重量百分比, Coc是土壤有机碳含量, D _g 是土壤几何平均粒径。
	K2	Shirazi & Boersma, 1984 ^[20]	$K = 7.954 \times \left\{ 0.0017 + 0.0494 \times \exp \left[-\frac{1}{2} \times \left(\frac{\log(D_g) + 1.675}{0.6986} \right)^2 \right] \right\}$	
	K3	Torri 等, 1997 ^[21]	$K = 0.0293 \times (0.65 - D_g + 0.24D_g^2) \times \exp \left\{ -0.0021 \times \frac{Coc}{Cl} - 0.00037 \times \left(\frac{Coc}{Cl} \right)^2 - 4.02 + Cl^2 \right\}$	
LS	LS1	符素华等, 2015 ^[24]	$L = (\lambda/22.13)^m; \quad m = \begin{cases} 0.2 & \theta < 0.5^\circ \\ 0.3 & 0.5^\circ \leq \theta < 1.5^\circ \\ 0.4 & 1.5^\circ \leq \theta < 3^\circ \\ 0.5 & \theta \geq 3^\circ \end{cases}$ $S = \begin{cases} 10.8 \sin(\theta) + 0.03 & \theta < 5.14^\circ \\ 16.8 \sin(\theta) + 0.5 & 5.14^\circ \leq \theta \end{cases}$	L是坡长因子, S是坡度因子, λ是水平投影坡长, m是可变的坡长指数, θ是坡度。

表 4 P 因子赋值表

Table 4 P factor value table

土地利用类型	P 值
林地	1
草地	1
旱地	0.352
水田	0.1
不透水面	0
水体	0
裸地	1

2 数据样本描述

2.1 数据信息及结构

本数据集包含 2001–2020 年东北区域土壤水蚀模数数据。数据集空间分辨率为 250 m，时间分辨率为 1 年，数值单位为 $(t\ ha^{-1}\ y^{-1})$ 。数据集保存为“2001–2020 年东北区域土壤水蚀模数.zip”压缩文件，压缩文件大小为 2.71 GB。其中文件以对应年份进行命名，如“NE_2001”代表 2001 年东北区域土壤水蚀模数，共 20 个文件，背景值为-1。

2.2 数据样本

以下给出 2001–2020 年，5 年一期的土壤水蚀结果时空分布图（图 2）。总体上东北地区的土壤侵蚀热点为西、北、南的大、小兴安岭与长白山地区，这三个地区的海拔都较周围高，土壤侵蚀敏感性也较强。其中以长白山南侧土壤水蚀模数较大，主要是因为该地区处在黄海暖流北上所产生的水汽通道上，南侧为迎风坡，因而长白山地区的降水多。整个流域的降水量非常丰沛，导致该地区计算出的土壤水蚀模数偏大。山区的侵蚀状况主要受到植被变化、冻融作用等多营力的影响，随时间变化呈现出波动的趋势；平原地区的侵蚀状况随时间变化呈现出轻度和微度的侵蚀区域缓慢增大的趋势，强烈的侵蚀呈现减小趋势。

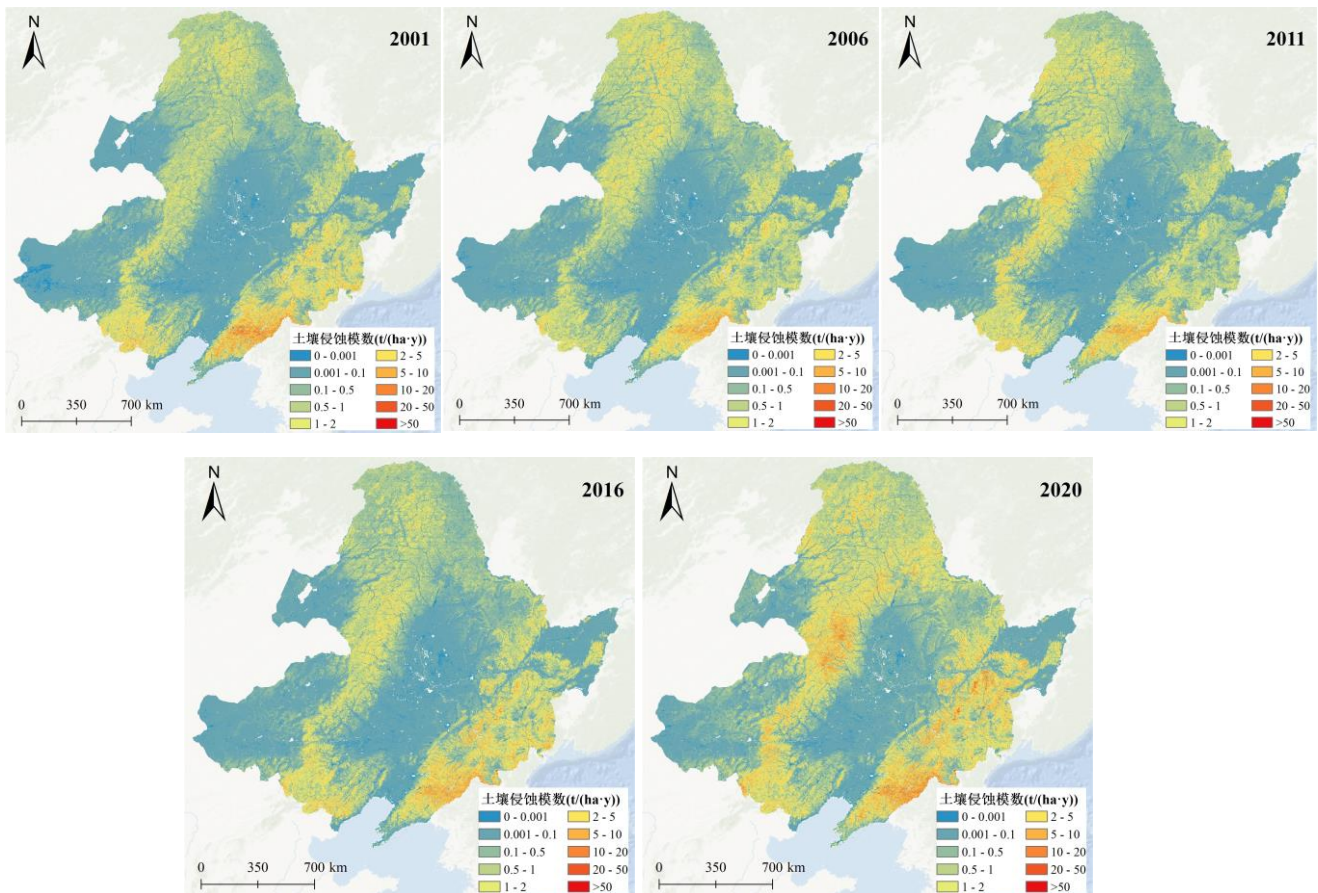


图 2 2001–2020 年东北区域土壤水蚀估算分布图

Figure 2 Distribution map of estimated soil water erosion in Northeast China from 2001 to 2020

3 数据质量控制和评估

3.1 质量控制

本研究将水利部发布的 2001–2020 年《河流泥沙公报》^[25]中东北区域主要水文监测站测量得到的年输沙量以及输沙模数与模型估算的土壤侵蚀模数进行直接验证，并基于验证结果选取最优的因子算法组合生成最终侵蚀模数估算结果。步骤如下：1) 首先根据水文站点分布位置与流域河段位置的相对关系以及站点对应的控制流域面积对初始流域评价单元进行合并或者迁移，建立一个站点对一个流域单元或一个站点对多个流域单元的流域评价单元。2) 其次对每个评价单元建立以站点测量得到的年输沙量为横坐标，利用土壤水蚀估算模型计算的流域总土壤水蚀量为纵坐标建立散点图，并结合时间序列变化趋势以及公开资料记载的重大河流泥沙事件剔除异常值。3) 最后将水文观测站点测量得到的年输沙量与对应流域评价单元内计算出的土壤水蚀模数的总和对比并计算二者在控制流域面积尺度上的评估指标，比较不同组合方式在三级流域尺度上的验证指标，选取精度最优的组合模型生成最终结果。

3.2 质量评估

在本研究中将实测水文站点数据与对应流域估算的土壤水蚀模数进行统计比较，引入相关系数（Correlation coefficient, Cov）、均方根误差（Root Mean Square Error, RMSE）、平均绝对误差（Mean Absolute Error, MAE）等统计指标进行精度评估。相关计算公式如下：

相关系数（Cov）：

$$\text{Cov} = \frac{\sum_{i=1}^n (x_i^{obv} - \overline{x_i^{obv}})(x_i^{est} - \overline{x_i^{est}})}{\sqrt{\sum_{i=1}^n (x_i^{obv} - \overline{x_i^{obv}})^2} \sqrt{\sum_{i=1}^n (x_i^{est} - \overline{x_i^{est}})^2}} \quad (2)$$

均方根误差（RMSE）：

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i^{obv} - x_i^{est})^2} \quad (3)$$

平均绝对误差（MAE）：

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i^{obv} - x_i^{est}| \quad (4)$$

其中， x_i^{obv} 为站点观测值， x_i^{est} 为对应流域估算的土壤水蚀模数值， n 为站点观测值数量， $\overline{x_i^{obv}}$ 为站点观测值平均值， $\overline{x_i^{est}}$ 为水蚀模数估算值的平均值。

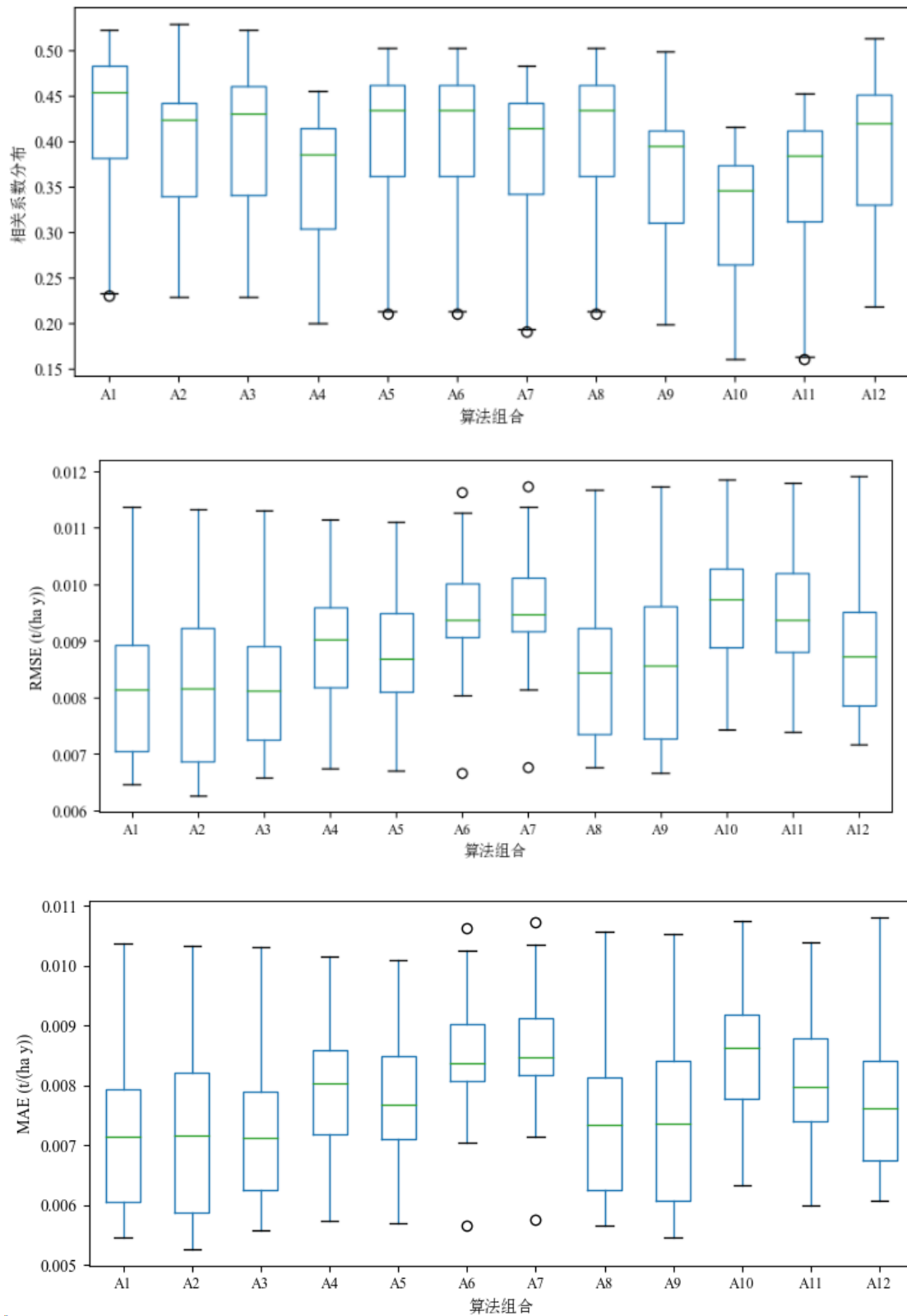


图3 相关系数, RMSE 与 MAE 统计箱线图 (其中 A1–A12 为 12 种算法组合方式, 如表 5 所示)

Figure 3 Statistical boxplot analysis of correlation coefficient, RMSE and MAE (A1–A12 is a combination of 12 algorithms, as shown in table 5)

各因子组合排列方案如表 5 所示, 每种因子组合方案的相关系数, RMSE 与 MAE 三个评估指标统计如图 3 所示, 最终得到 12 种土壤水蚀估算结果及验证结果。由于因子计算方法的适用条件以

及地理要素之间的差异，根据 R、K、C 因子的不同算法计算出的水蚀结果主要会根据降雨、土壤理化性质以及地表植被覆盖状况的不同，而产生不同的结果。由图 3 可知最佳的组合因子组合方案为 A1 组合，其中使用的为 R1-K1-C1-LS1-P 的因子组合，其流域估算的土壤水蚀模数值与站点观测值相关性分布主要为 0.4–0.5 之间，相较于其他组合方案整体稍高，RMSE 分布位于 0.007–0.009 之间相较于其他组合方案整体偏低。A6, A7, A10 组合方式相关性指标明显偏低且 RMSE 与 MAE 偏高，说明在东北区域以上三种因子组合的估算模型较不适用。A2–A5、A8、A9、A11、A12 相关性分布于 0.3–0.45，且分布较为离散，RMSE 与 MAE 相较于 A1 组合偏高，整体分布于 0.007–0.01 之间。

表 5 模型各因子组合表

Table 5 The table of combination for various model factors

组合编号	R	K	C	LS	P
A1	R1	K1	C1	LS1	P
A2	R2	K1	C1	LS1	P
A3	R1	K1	C2	LS1	P
A4	R2	K1	C2	LS1	P
A5	R1	K2	C1	LS1	P
A6	R2	K2	C1	LS1	P
A7	R1	K2	C2	LS1	P
A8	R2	K2	C2	LS1	P
A9	R1	K3	C1	LS1	P
A10	R2	K3	C1	LS1	P
A11	R1	K3	C2	LS1	P
A12	R2	K3	C2	LS1	P

4 数据价值

水土流失对于我国东北地区的生态环境以及粮食种植影响较为严重。同时受到气候、土壤、植被、土地利用等综合地理环境因素的影响，东北区域黑土流失严重，耕层变浅，肥力下降，土壤盐碱化加剧。水土过程模拟是治理土地退化的关键基础环节。本研究提供的 2001–2020 年长时间序列的高精度、高时空分辨率的土壤水蚀模数估算结果，对东北地区水土保持以及生态环境监测具有重要意义。

本数据集是公开发布的东北区域 2001–2020 年土壤水蚀模数估算数据，具有较高的时空分辨率和实际精度，可为东北区域的土壤侵蚀评估以及水土保持决策提供基础数据，帮助决策者和生产者制定合理的政策和风险管理策略。同时，本数据集也具有较高的科学价值与研究意义，基于模型优选思想为空间异质性较高区域的土壤侵蚀估算提供了新的思路与角度。在后续的研究中，我们将尝试使用精度更高的源数据来计算各个影响因子并使用月尺度监测数据进行土壤侵蚀的验证。

数据作者分工职责

吴瀚逸（2000—），男，福建南安人，硕士研究生，研究方向为农业遥感、环境遥感。主要承担工作：

研究方案设计、算法构建、产品生产、精度验证和论文撰写。

熊俊峰（1993—），男，江西九江人，助理研究员，研究方向为环境遥感。主要承担工作：研究思路与方案设计、论文撰写。

侯渲（1998—），男，甘肃天水人，工程师，研究方向数据库建设、数据共享。主要承担工作：数据收集整理、数据产品评估。

林晨（1984—），男，江苏南京人，研究员，研究方向自然资源遥感监测。主要承担工作：研究思路设计。

许金朵（1982—），女，江苏徐州人，高级工程师，研究方向数据库建设、数据共享。主要承担工作：研究框架设计，数据提供。

马荣华（1972—），男，山东临沂人，研究员，研究方向湖泊环境遥感。主要承担工作：研究方案设计。

参考文献

- [1] 王志杰, 简金世, 焦菊英, 等. 基于 RUSLE 的松花江流域不同侵蚀类型区泥沙输移比估算[J]. 水土保持研究, 2013, 20(5): 50–56. [WANG Z J, JIAN J S, JIAO J Y, et al. Estimation of sediment delivery ratio in different soil erosion regions in the Songhua River Basin based on RUSLE[J]. Research of Soil and Water Conservation, 2013, 20(5): 50–56.]
- [2] 王硕, 方海燕, 和继军. 植被覆盖和降雨因子变化及对东北黑土区土壤侵蚀的影响[J]. 水土保持通报, 2021, 41(2): 66–75. DOI: 10.16115/j.cnki.issn.1005-7129.2020.06.006. [WANG S, FANG H Y, HE J J. Effects of variation in vegetation cover and rainfall on soil erosion in black soil region, northeastern China[J]. Bulletin of Soil and Water Conservation, 2021, 41(2): 66–75. DOI: 10.16115/j.cnki.issn.1005-7129.2020.06.006.]
- [3] 刘盼, 任春颖, 王岩松, 等. 基于 RUSLE 模型的梅河口市土壤侵蚀动态分析[J]. 水土保持通报, 2019, 39(1): 172–179, 325. DOI: 10.16843/j.sswc.2017.06.006. [LIU P, REN C Y, WANG Y S, et al. Dynamic analysis of soil erosion in Meihekou city based on RUSLE model[J]. Bulletin of Soil and Water Conservation, 2019, 39(1): 172–179, 325. DOI: 10.16843/j.sswc.2017.06.006.]
- [4] 李佳蕾, 孙然好, 熊木齐, 等. 基于 RUSLE 模型的中国土壤水蚀时空规律研究[J]. 生态学报, 2020, 40(10): 3473–3485. DOI: 10.5846/stxb201903290610. [LI J L, SUN R H, XIONG M Q, et al. Estimation of soil erosion based on the RUSLE model in China[J]. Acta Ecologica Sinica, 2020, 40(10): 3473–3485. DOI: 10.5846/stxb201903290610.]
- [5] 鄂立思, 孙丽, 李苗, 等. 基于 GIS 和 RUSLE 乌裕尔河流域土壤侵蚀定量评价[J]. 哈尔滨师范大学自然科学学报, 2015, 31(2): 154–158. DOI: 10.3969/j.issn.1000-5617.2015.02.042. [E L S, SUN L, LI M, et al. Quantitative assessment of soil erosion in wuyuer river basin based on GIS and RUSLE[J]. Natural Science Journal of Harbin Normal University, 2015, 31(2): 154–158. DOI: 10.3969/j.issn.1000-5617.2015.02.042.]
- [6] 张光辉, 杨扬, 刘瑛娜, 等. 东北黑土区土壤侵蚀研究进展与展望[J]. 水土保持学报, 2022, 36(2): 1–12. DOI: 10.13870/j.cnki.stbcxb.2022.02.001. [ZHANG G H, YANG Y, LIU Y N, et al. Advances and prospects of soil erosion research in the black soil region of northeast China[J]. Journal

- of Soil and Water Conservation, 2022, 36(2)1-12. DOI: 10.13870/j.cnki.stbcxb.2022.02.001.]
- [7] ZHANG N J, XIAO Y D, HOU R X, et al. Dataset of Ecosystem Soil Conservation Capacity Assessment in Sanbei Project Area (2000-2020)[DB/OL]. Science Data Bank, 2021. <https://cstr.cn/31253.11.sciencedb.01287>. CSTR:31253.11.sciencedb.01287.
- [8] 王莉, 张凡. “一江两河”地区土壤侵蚀及其影响因子空间分布(2001-2015)[DB/OL]. 国家青藏高原科学数据中, 2022. <https://doi.org/10.11888/Terre.tpd.272397>. <https://cstr.cn/18406.11.Terre.tpd.272397>. [WANG L, ZHANG F. Soil erosion and its influencing factors in the Mid-Yarlung Tsangpo River Region (2001-2015) [DB/OL]. National Tibetan Plateau/Third Pole Environment Data Center, 2022. <https://doi.org/10.11888/Terre.tpd.272397>. <https://cstr.cn/18406.11.Terre.tpd.272397>.]
- [9] 章文波. 重点区域 20 国 300 米分辨率土壤侵蚀强度数据集(2015)[DB/OL]. 国家青藏高原科学数据中心, 2021. DOI:10.11888/Soil.tpd.271737. [Zhang W. Dataset of Soil Erosion Intensity with 300 m resolution in 20 countries in key regions (2015) [DB/OL]. National Tibetan Plateau/Third Pole Environment Data Center, 2021. DOI:10.11888/Soil.tpd.271737.]
- [10] 章文波. 青藏高原 300 米分辨率土壤侵蚀强度数据集(1992、2005、2015)[DB/OL]. 国家青藏高原科学数据中心, 2019. DOI:10.11888/Disas.tpd.270224. [Zhang W. Dataset of soil erosion intensity with 300m resolution in Tibetan Plateau (1992, 2005, 2015) [DB/OL]. National Tibetan Plateau/Third Pole Environment Data Center, 2019. DOI:10.11888/Disas.tpd.270224.]
- [11] 杨勤科. 巴基斯坦 15 个小流域 2.5 米分辨率土壤水蚀模数数据集(2019)[DB/OL]. 国家青藏高原科学数据中心, 2020, DOI:10.11888/Soil.tpd.270429. [YANG Q K. Dataset of soil water erosion modulus with 2.5 m resolution in 15 watersheds of Pakistan (2019)[DB/OL]. National Tibetan Plateau/Third Pole Environment Data Center, 2020, DOI:10.11888/Soil.tpd.270429.]
- [12] RENARD K, FOSTER G R, WEESIES G, et al. Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE)[J]. Agricultural Handbook, 1997.
- [13] MCCOOL D K, BROWN L C, FOSTER G R, et al. Revised slope steepness factor for the universal soil loss equation[J]. Transactions of the ASAE, 1987, 30(5): 1387–1396. DOI: 10.13031/2013.30576.
- [14] Huffman G J, Stocker E F, Bolvin D T, et al. GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06, Greenbelt, MD[DB]. Goddard Earth Sciences Data and Information Services Center (GES DISC), 2019. DOI:10.5067/GPM/IMERG/3B-HH/06.
- [15] Fischer G, Nachtergaele F, Prieler S, et al. Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008)[DB]. IIASA, Laxenburg, Austria and FAO, Rome, Italy, 2008, <https://previous.iiasa.ac.at>.
- [16] JARVIS A, H I REUTER, A NELSON, et al. Hole-filled SRTM for the globe Version 4[DB], CGIAR-CSI SRTM 90m, 2008, <https://srtm.csi.cgiar.org>.
- [17] 章文波, 谢云, 刘宝元. 利用日雨量计算降雨侵蚀力的方法研究[J]. 地理科学, 2002, 22(6): 705-711. DOI: 10.3969/j.issn.1000-0690.2002.06.012. [ZHANG W B, XIE Y, LIU B Y. Rainfall erosivity estimation using daily rainfall amounts[J]. Scientia Geographica Sinica, 2002, 22(6): 705-711. DOI: 10.3969/j.issn.1000-0690.2002.06.012.]
- [18] 周伏建, 黄炎和. 福建省降雨侵蚀力指标 R 值[J]. 水土保持学报, 1995, 9(1): 13–18. DOI: 10.3969/j.issn.1005-3409.2004.01.017. [ZHOU F J, HUANG Y H. R value of rainfall erosivity index in

- Fujian Province[J]. Journal of Soil and Water Conservation, 1995, 9(1): 13–18. DOI: 10.3969/j.issn.1005-3409.2004.01.017.]
- [19] EASTERLING W E, ROSENBERG N J, MCKENNEY M S, et al. Preparing the erosion productivity impact calculator (EPIC) model to simulate crop response to climate change and the direct effects of CO₂[J]. Agricultural and Forest Meteorology, 1992, 59(1/2): 17–34. DOI: 10.1016/0168-1923(92)90084-H.
- [20] SHIRAZI M A, BOERSMA L. A unifying quantitative analysis of soil texture[J]. Soil Science Society of America Journal, 1984, 48(1): 142–147. DOI: 10.2136/sssaj1984.03615995004800010026x.
- [21] TORRI D, POESEN J, BORSELLI L. Predictability and uncertainty of the soil erodibility factor using a global dataset[J]. CATENA, 1997, 31(1/2): 1–22. DOI: 10.1016/S0341-8162(97)00036-2.
- [22] 蔡崇法, 丁树文, 史志华, 等. 应用 USLE 模型与地理信息系统 IDRISI 预测小流域土壤侵蚀量的研究[J]. 水土保持学报, 2000, 14(2): 19-24. DOI: 10.3321/j.issn: 1009-2242.2000.02.005. [CAI C F, DING S W, SHI Z H, et al. Study of applying USLE and geographical information system IDRISI to predict soil erosion in small watershed[J]. Journal of Soil and Water Conservation, 2000, 14(2): 19-24. DOI: 10.3321/j.issn: 1009-2242.2000.02.005.]
- [23] LEEUWEN, W J D V, SAMMONS G. Vegetation dynamics and erosion modeling using remotely sensed data (MODIS) and GIS [C]. Tenth biennial USDA forest service remote sensing applications conference, 2004, 5-9.
- [24] 符素华, 刘宝元, 周贵云, 等. 坡长坡度因子计算工具[J]. 中国水土保持科学, 2015, 13(5):105-110. DOI: 10.3969/j.issn.1672-3007.2015.05.016. [FU S H, LIU B Y, ZHOU G Y, et al. Calculation tool of topographic factors[J]. Science of Soil and Water Conservation, 2015, 13(5): 105-110. DOI: 10.3969/j.issn.1672-3007.2015.05.016.]
- [25] 中华人民共和国水利部. 中国河流泥沙公报[M]. 北京: 中国水利水电出版社, 2020. [Ministry of Water Resources of the People's Republic of China. China river sediment bulletin [M]. Beijing: China Water Resources and Hydropower Press, 2020.]

论文引用格式

吴瀚逸, 熊俊峰, 侯渲, 等. 2001 – 2020 年中国东北区域土壤水蚀数据集[J/OL]. 中国科学数据, 2023, 8(4). (2023-06-19). DOI: 10.11922/11-6035.csd.2023.0096.zh.

数据引用格式

吴瀚逸, 熊俊峰, 侯渲, 等. 2001 – 2020 年中国东北区域土壤水蚀数据集[DS/OL]. V2. Science Data Bank, 2023. (2023-05-08). DOI: 10.57760/sciencedb.08027.

A dataset of soil water erosion of Northeast China from 2001 to 2020

WU Hanyi^{1,2,3,4}, XIONG Junfeng^{1,4*}, HOU Xuan^{4,5}, LIN Chen⁴, XU Jinduo^{4,5}, MA Ronghua^{4,5}

1. International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, P.R.China

2. State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing 100875, P.R.China

3. Key Laboratory of Environmental Change and Natural Disaster, Ministry of Education, Beijing Normal University, Beijing 100875, P.R.China

4. Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, P.R.China

5. Lake-Watershed Science Data Center, National Earth System Science Data Center, National Science & Technology Infrastructure of China, Nanjing 210008, P.R.China

*Email: jfxiong@niglas.ac.cn

Abstract: The black soil area in Northeast China plays an important role in food production and ecological security in China. However, over-cultivation practices have led to severe soil erosion, which seriously threatens food security and ecological environment in Northeast China. Accurate simulations of water erosion with high spatio-temporal resolution are important for advancing sustainable development goals, such as promoting sustainable agriculture and monitoring land degradation in Northeast China. This dataset is based on Google Earth Engine (GEE) cloud platform, integrating multi-source remote sensing data, deconstructing RUSLE model, and optimizing algorithm combinations for each factor of the model. We compared and verified the annual sand transport volume and sand transport modulus data from the main hydrological monitoring stations in Northeast China with the soil erosion modulus estimated by the RUSLE model. Then, we selected the optimal factor algorithm combination based on three accuracy metrics (time series correlation, root mean square error and mean absolute error), and obtained the estimation results of soil water erosion modulus with the resolution of 250 m. This dataset can better depict the spatial distribution and temporal changes of soil erosion modulus in Northeast China from 2001 to 2020. It can serve as an effective reference for soil erosion control and assessments in Northeast China.

Keywords: soil water erosion; model preference; northeast China; RUSLE

Dataset Profile

Title	A dataset of soil water erosion of Northeast China from 2001 to 2020
Data corresponding author	XIONG Junfeng (jfxiong@niglas.ac.cn)
Data authors	WU Hanyi, XIONG Junfeng, HOU Xuan, LIN Chen, XU Jinduo, MA Ronghua
Time range	2001–2020
Geographical scope	Northeast China: Heilongjiang Province, Jinlin Province, Liaoning Province, Northeast Hebei Province (Chende City, Qinhuangdao City), Eastern Inner Mongolia Autonomous Region (Xilingol League, Chifeng City, Tongliao City, Xing'an League,

	Hulunbuir City)
Spatial resolution	250 m
Data volume	2.71 GB
Data format	*.tif
Data service system	< https://doi.org/10.57760/sciencedb.08027 >
Sources of funding	The Director Fund of the International Research Center of Big Data for Sustainable Development Goals (CBAS2022DF014); the National Natural Science Foundation of China (Grant No. 42201400); the Natural Science Foundation of Jiangsu Province for Youths (Grant No. BK20221158); Network Security and Informatization Plan of Chinese Academy of Sciences (CAS-WX2021SF-0306); Scientific Data Center of Nanjing Institute of Geography & Limnology Chinese Academy of Sciences (CAS-WX2022SDC-SJ05).
Dataset composition	The dataset is composed of 20 files in TIF format, each named in accordance with the corresponding year, respectively.