



文献 CSTR:

32001.14.11-6035.noda.2022.0003.zh



文献 DOI:

10.11922/11-6035.noda.2022.0003.zh

数据 DOI:

10.11922/sciencedb.j00001.00368

文献分类: 地球科学

收稿日期: 2022-01-21

开放同评: 2022-02-12

录用日期: 2022-06-17

发表日期: 2023-08-11

* 论文通信作者

张豪伟: haoweizhang@whu.edu.cn

2009–2020 年基于 GOSAT 卫星的全球中低纬 二氧化碳柱浓度数据集

许静静¹, 龚威¹, 张劲², 张豪伟^{3*}, 马昕³, 韩舸⁴

1. 武汉大学, 电子信息学院, 武汉 473072

2. 武汉测绘研究院, 武汉 430079

3. 武汉大学, 测绘遥感信息工程国家重点实验室, 武汉 430079

4. 武汉大学, 遥感信息工程学院, 武汉 430079

摘要: 为应对日益加剧的温室效应问题, 全球各国联合签署了《巴黎协定》, 我国也制定了碳达峰、碳中和的计划和政策。二氧化碳作为最主要的温室气体, 是国际关注的重点。因此, 获得高精度、高分辨率的二氧化碳柱浓度时空分布图对于推进“自上而下”评估碳源、碳汇、碳中和的研究至关重要。本研究利用 GOSAT 卫星全球数据, 通过迁移学习理论, 将时间信息作为先验廓线融入空间信息, 对空间预测信息进行调整, 得到高准确度的二氧化碳柱浓度时空预测结果。与中低纬的 TCCON 站点数据对比, 本算法最终得到的月均二氧化碳柱浓度图指标的综合结果 R 为 0.98, RMSE 为 1.38 ppm, 空间分辨率为 0.25°。本数据集由 2009–2020 年月均二氧化碳柱浓度文件组成, 包含 136 个 h5 文件, 可应用于长时间序列的碳源和碳汇计算。

关键词: 二氧化碳柱浓度; GOSAT; 迁移学习; 全球

数据库(集)基本信息简介

数据库(集)名称	2009–2020 年基于 GOSAT 卫星的全球中低纬二氧化碳柱浓度数据集
数据作者	许静静, 龚威, 张劲, 张豪伟, 马昕, 韩舸
数据通信作者	张豪伟 (haoweizhang@whu.edu.cn)
数据时间范围	2009–2020 年
地理区域	地理范围为 60°S – 60°N, 180°W – 180°E
空间分辨率	0.25° × 0.25°
数据量	11.56 GB
数据格式	*.h5
数据服务系统网址	http://doi.org/10.11922/sciencedb.j00001.00368
基金项目	国家自然科学基金 (42171464、41971283、41801261、41827801、41801282); 国家重点研发计划 (2017YFC0212600); 国家对地观测数据中心开放基金 (NODAOP2021005); LIESMARS 特别研究基金。
数据库(集)组成	本数据集包含 2009–2020 年全球中低纬地区 136 个月均二氧化碳柱浓度文件数据 (h5 格式)。文件涵盖了数据点位经纬度、空间和时空插值结果、模型参数、空间分辨率等信息。

引言

近年来,全球气候变化问题得到了越来越广泛的关注,气候变化问题正深刻地影响着人们的生活,全球气候变暖导致海平面上升、影响生态系统稳定性、引发诸如洪涝和干旱等极端恶劣天气^[1-2]。温室效应已经被证实正在发生,并将长期存在,给人类带来不可想象的生存风险。产生温室效应的主要原因是大量排放的二氧化碳气体^[3],其能够减少地球上的长波辐射向外发散,从而使得地球表面温度升高。由于现代化工业对化石能源的大量消耗,二氧化碳的排放速度迅速增加,地球的承载能力已经无法完全满足如此大量的温室气体,为此我国已将“碳达峰”、“碳中和”写入十四五规划。因此,如何有效地检测大气中的二氧化碳柱浓度的时空变化,对于实施减少温室气体排放的措施具有重要指导性意义。

随着遥感技术的发展,各国陆续发射了二氧化碳监测卫星,遥感监测已经成为二氧化碳柱浓度检测的主要方法。目前主流的二氧化碳监测卫星携带的均是被动探测器^[4-6],易受云层、污染等各种因素影响,使得卫星数据在精度和时空覆盖等方面存在一定的缺陷。如何克服上述限制,寻找一种能够获得高精度、高分辨率的二氧化碳柱浓度时空分布图的方法,对于推动“自上而下”评估碳循环、碳源、碳汇研究具有重要意义。

对于卫星观测的二氧化碳柱浓度在时空覆盖上的不足,传统的预测方法是使用空间插值法。这类方法简单快捷,但由于仅考虑了空间特性,没有考虑数据随时间变化的特性^[7-10],其准确性不足。一些学者将时间信息融合进插值方法中,将不同时间段的数据分别进行插值,得到更好的结果^[11-12]。但是这些方法都忽略了二氧化碳柱浓度的时间剖面信息。二氧化碳柱浓度变化在时间尺度上具有连续性,通过这一信息可以有效判断检测的二氧化碳柱浓度是否偏离真实值,从而增加插值的准确性。本研究将二氧化碳柱浓度的时间剖面信息引入插值方法,作为先验信息,对时空插值方法进行改进。形成新的时空插值方法,能够有效去除传统插值方法中异常值对插值结果的影响,具有更高的准确性。在现有的二氧化碳柱浓度数据集中,ZENG 等人^[13]基于 GOSAT-2 卫星反演的二氧化碳柱浓度,相关系数 R 达到 0.92,LI 等人^[14]基于 OCO-2 卫星反演的二氧化碳柱浓度,相关系数 R 达到 0.91,上述工作在数据集覆盖面积和周期上取得了卓越的成就。与以上两个数据集相比,本数据集验证精度分别提升了 6.5%和 7.7%。因此,本数据集取得了更高的精确度。

1 数据采集和处理方法

1.1 数据采集方法

本研究使用了 GOSAT 卫星官方数据产品 (GOSAT Data Archive Service, https://data2.gosat.nies.go.jp/index_en.html),将 2009–2020 年的二氧化碳柱浓度数据作为二氧化碳柱浓度反演的输入数据,使用的具体数据为 GOSAT_FTS_L3_V2.95。GOSAT 是全球首个用于温室气体浓度监测的卫星,发射于 2009 年 1 月,每 3 天就可以收集到全球约 5.6 万个观测点的最新数据,专门用于温室气体浓度的检测。其中傅里叶光谱仪 (TANSO-FTS) 获取的 3 个短波红外波段光谱分别为氧气吸收 A 带、二氧化碳弱吸收带、二氧化碳强吸收带,GOSAT 数据已广泛应用于全球二氧化碳柱浓度自上而下的反演过程。

1.2 数据处理方法和流程

GOSAT 数据提供了较低空间分辨率的二氧化碳柱浓度,为了反演得到每个月的二氧化碳柱浓度图,我们从以下 3 步进行。

(1) 对数据进行预处理,去除其中的异常值。

(2) 使用简单克里金插值方法,对每月的 GOSAT 二氧化碳柱浓度离散数据进行插值,得到全球区域的连续二氧化碳柱浓度数据。

(3) 通过迁移学习理论, 将时间剖面信息与空间插值信息相结合, 使得插值结果在符合空间距离相关性的同时, 更加符合时间剖面连续的特性。

1.2.1 简单克里金插值法进行预测

克里金插值法被称为空间自协方差最佳插值法, 包括普通克里金法、简单克里金法、泛克里金法、协同克里金法等。在本研究中, 我们使用简单克里金插值法对离散的二氧化碳柱浓度数据进行插值。简单克里金法具有无偏特性, 其加权值不受特殊条件约束, 比普通克里金法更加方便使用。

为了获得全球月平均二氧化碳柱浓度分布图, 我们首先对已知的 GOSAT 二氧化碳柱浓度数据进行月平均值计算; 然后, 对于每个月的 GOSAT 二氧化碳柱浓度数据, 利用简单克里金法对数据空白区域进行合理有效的填充。

1.2.2 构建先验廓线曲线库

简单克里金法是利用二氧化碳柱浓度的空间距离信息, 对于每月的二氧化碳柱浓度进行插值时, 仅考虑单月的浓度信息, 为了充分利用时间信息去除空间插值的异常性, 本文构建了时间剖面曲线作为先验廓线, 用于调整插值法得到的二氧化碳柱浓度图。

先验廓线曲线库利用了二氧化碳柱浓度随时间连续变化的特性, 而简单克里金法没有考虑这一时间特性, 其插值结果在相邻的月份之间存在不连续的现象, 甚至出现较大偏差。GOSAT 卫星每 3 天对同一地区过境, 能够提供稳定的月平均二氧化碳柱浓度数据, 利用 GOSAT 卫星数据构建先验廓线曲线库, 可以对异常数据进行校正。

首先, 我们使用 GOSAT_L3 数据, 选取有效的空间点, 有效的空间点的月平均二氧化碳柱浓度有效值应大于 7 个, 以保证在进行时间剖面先验廓线曲线拟合时有足够多的数据, 支撑得到有效的拟合结果。然后利用公式 (1) 对空间某点拟合, 构建这一空间点的连续月平均二氧化碳柱浓度廓线, 用来表征二氧化碳柱浓度随时间的变化情况, 公式如下:

$$F(t) = a + b * t + c * \cos\left(\frac{2\pi t}{12}\right) + d * \sin\left(\frac{2\pi t}{12}\right) + e * \cos\left(\frac{4\pi t}{12}\right) + f * \sin\left(\frac{4\pi t}{12}\right) + \Delta\delta \quad (1)$$

式中 a 为年平均二氧化碳柱浓度; b 为年际间的分量系数; c 、 d 、 e 、 f 为季节分量系数; t 是采样间隔; $\Delta\delta$ 为误差补偿项。

1.2.3 基于迁移学习理论的时间属性整合

利用空间距离进行插值, 仅考虑空间属性对于每月的二氧化碳柱浓度空白区域进行插值。由于插值过程仅考虑单月的浓度信息, 导致插值得到的相邻月份二氧化碳柱浓度存在割裂。同时, 其他空白区域的填充值仅符合空间距离特性, 在时间维度上, 其填充值可能出现偏离现象。为了解决这一问题, 我们将时间信息引入空间插值, 通过迁移成分分析理论, 得到同时符合空间、时间连续特性的二氧化碳柱浓度填充值。

迁移成分分析理论是一种基于特征的迁移学习方法, 它将两个领域的的数据映射到一个高维的再生希尔伯特空间。在这个空间中, 通过最小化源数据和目标数据之间的距离, 使数据的拟合结果能够同时保持两个领域的的数据特性^[15-17]。在本研究中, 我们通过先验廓线的构建, 得到时间域数据, 通过空间插值, 得到空间域数据, 然后利用迁移成分分析理论, 将时间剖面曲线库中相应的空间点位置拟合数据设为源域, 将全球域的空间插值数据设为目标域, 进行迁移学习, 将每个时间剖面数据从源域映射到目标域, 对公式 (1) 中的参数 c 、 d 、 e 、 f 进行再次拟合, 得到参数 a 、 b 的值, 最终拟合结果即为时空插值数据。

2 数据样本描述

本文所采用的数据集为全球中低纬度二氧化碳柱浓度数据集, 共包含 136 个数据, 数据空间分辨率为 0.25° , 每个数据文件大小约为 100 MB。文件命名为 “WHUCOC -YYYYMM-VXX.h5”, 其中 “YYYY” 表示数据的年份, “MM” 表示数据月份, “VXX” 为产品的版本, 时间采用 UTC

时间。每个文件内部包含了 12 个字段，其中“Spatiotemporal XCO₂”为我们输出的数据集字段，其余字段为辅助字段信息，详细的字段介绍如表 1 所示。

表 1 数据字典

Table 1 Data dictionary

序号	字段名称	数据类型	单位	字段介绍
1	Latitude	矩阵	(度, 分)	点位纬度
2	Longitude	矩阵	(度, 分)	点位经度
3	Spatial XCO ₂	矩阵	ppm	空间插值结果
4	Spatiotemporal XCO ₂	矩阵	ppm	时空插值结果
5	ParmentA	矩阵	-	模型参数 a
6	ParmentB	矩阵	-	模型参数 b
7	ParmentC	矩阵	-	模型参数 c
8	ParmentD	矩阵	-	模型参数 d
9	ParmentE	矩阵	-	模型参数 e
10	ParmentF	矩阵	-	模型参数 f
11	Spatial Resolution	浮点型	-	空间分辨率
12	EvaluationLayer	整数	-	不确定性

3 数据质量控制和评估

为了评估时空插值算法的效果，我们使用全球总碳柱观测网络（TCCON）^[18]的观测数据及 GOSAT_L3 产品数据对所获得的月均二氧化碳柱浓度进行准确性验证。TCCON 站的主要设备是地基傅里叶变换光谱仪。该光谱仪可以记录近红外光谱区域的太阳光谱数据，从这些光谱中可以检索出精确的二氧化碳柱平均浓度。考虑到算法预测的二氧化碳柱浓度的空间范围主要在中低纬区域，我们获取 2009–2020 年间全球中低纬地区 TCCON 站点的数据，表 2 列出了所用站点的位置信息，将 TCCON 各个站点观测的月平均二氧化碳柱浓度数据与算法所预测的数据进行匹配，然后使用相关系数 R 和均方根误差 RMSE 两个评价指标分别对算法预测结果进行评估。R 是研究变量之间线性相关程度的量，RMSE 是一种常用的测量数值之间差异的度量。数据集预测值与验证值间应成正比例关系，通过这两个指标可以分别测量两者之间线性相关程度及误差大小。图 1 为所有 TCCON 站点数据与产品数据对比的散点图，其中横轴为 TCCON 数据值，纵轴为相应时间对应位置的产品数据，我们的算法生成的数据与 TCCON 验证数据基本成正比例关系，表明算法具有较高的准确度。图 2 和图 3 分别列出了预测值与各个 TCCON 站点数据的 R、RMSE，结果表明，我们的算法的预测结果在全球中低纬 TCCON 站点评价指标的综合结果 R 为 0.98，RMSE 为 1.38。

在使用 GOSAT_L3 产品对数据集进行评估时，为了评估算法的准确性，我们将模型预测的结果与 2009–2020 年的 GOSAT_L3 数据进行了比较。为了在全局范围内均匀验证，我们为每个 20° 经度区间删除了一列 GOSAT_L3 数据。删除的数据将用作验证的验证集。R 和 RMSE 被用作评估指标来评估验证集和预测数据。对比结果（图 4）显示，2010–2020 年 R 的平均值为 0.96，RMSE 的平均值为 0.53 ppm。由于数据为 2009 年 6 月至 12 月，精度较低，表明该模型不适用于不完整年份。总的来说，我们的模型产品可以填补全球 XCO₂ 的空白区域。与此同时，将我们的算法与同类的算法^[19]进行比较，我们的预测结果的 R 值显著提高，评价指标 R 在所有站点的综合值（0.98）提高了 4.6%，与主流算法相比，我们反演的二氧化碳柱浓度图分辨率 0.25°，高于其他算法的 1° 空间分辨率，我们反演的数据集具有更长的时间跨度，从 2009 年到 2020 年，时间跨度达 12 年。

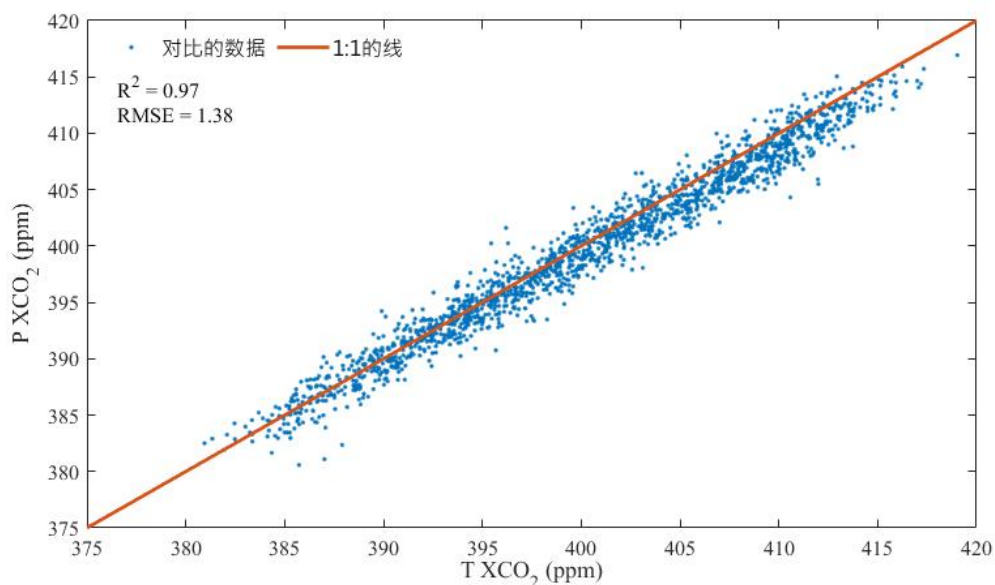


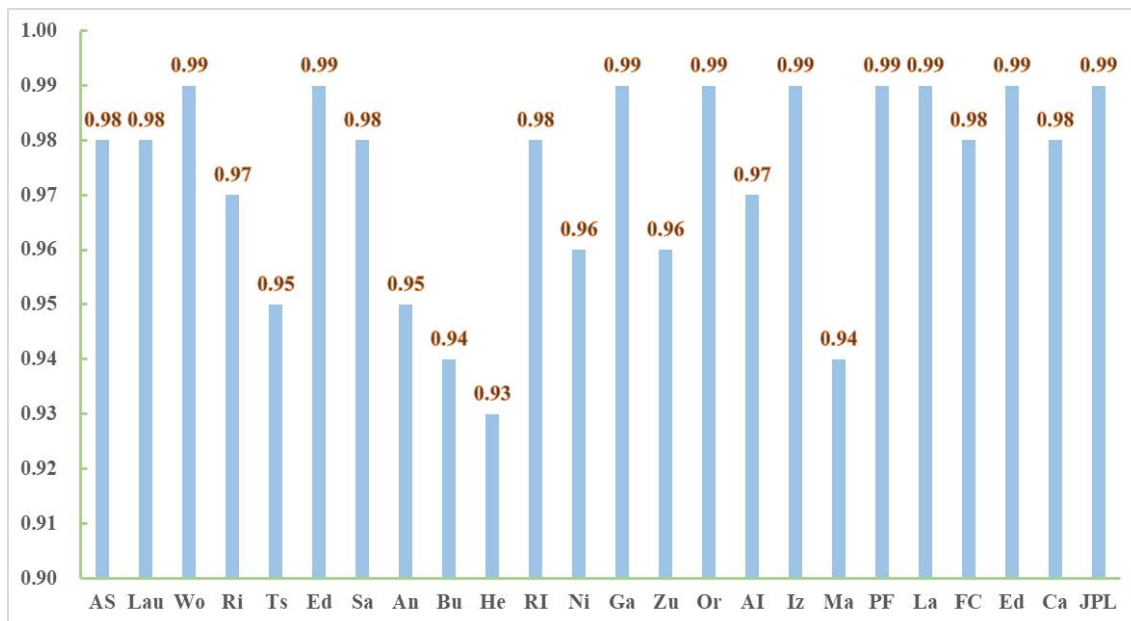
图 1 TCCON 站点数据与产品数据对比

Figure 1 Comparison between TCCON site data and product data

表 2 中低纬插值区域 TCCON 站点位置信息

Table 2 TCCON site location in interpolation regions of middle and low latitudes

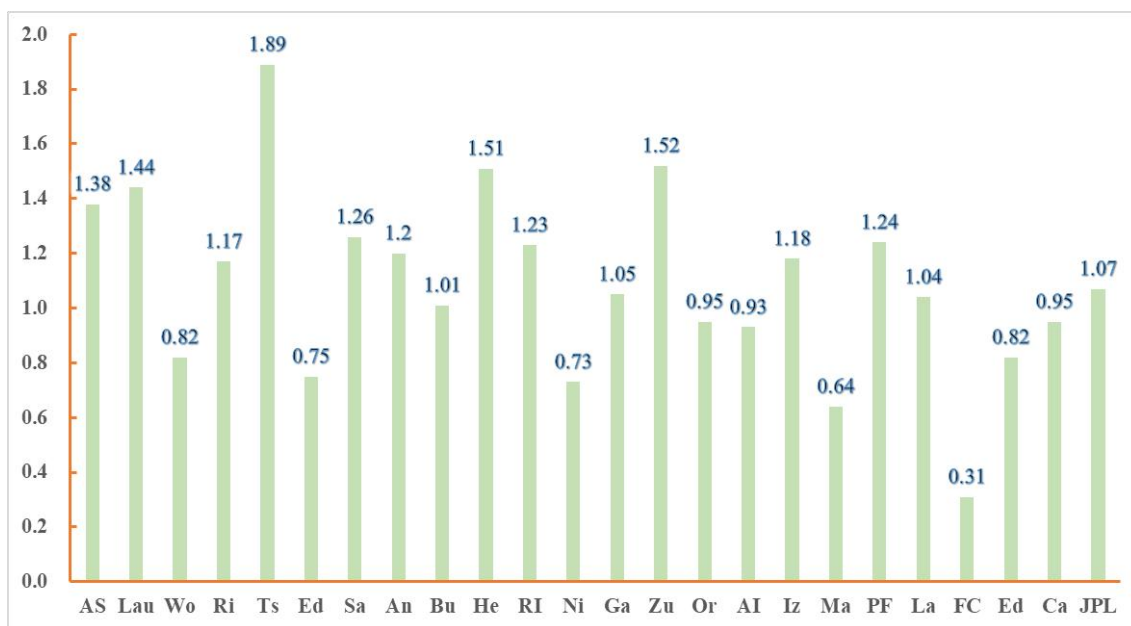
TCCON 站点 (英文缩写)	经度 (度)	纬度 (度)
Jet Propulsion Laboratory (JPL)	-118.18	34.20
Caltech (Ca)	-118.13	34.14
Edwards (Ed)	-117.88	34.96
Four Corners (FC)	-108.48	36.80
Lamont (La)	-97.49	36.60
Park Falls (PF)	-90.27	45.94
Manaus (Ma)	-60.60	-3.21
Izana (Iz)	-16.48	28.30
Ascension Island (AI)	-14.33	-7.92
Orléans (Or)	2.11	47.97
Zugspitze (Zu)	10.98	47.42
Garmisch (Ga)	11.06	47.48
Nicosia (Ni)	33.38	35.14
Réunion Island (RI)	55.49	-20.90
Hefei (He)	117.17	31.90
Burgos (Bu)	120.65	18.53
Anmeyondo (An)	120.65	36.54
Saga (Sa)	130.29	33.24
Edwards (Ed)	130.89	-12.43
Tsukuba (Ts)	140.12	36.05
Rikubetsu (Ri)	143.77	43.46
Wollongong (Wo)	150.88	-34.41
Lauder01&02&03 (Lau)	169.68	-45.04



AS: 所有站点相关系数 R 的均值, 各站点名以英文缩写替代

图 2 各 TCCON 站点观测数据与预测数据 XCO_2 的相关系数 R

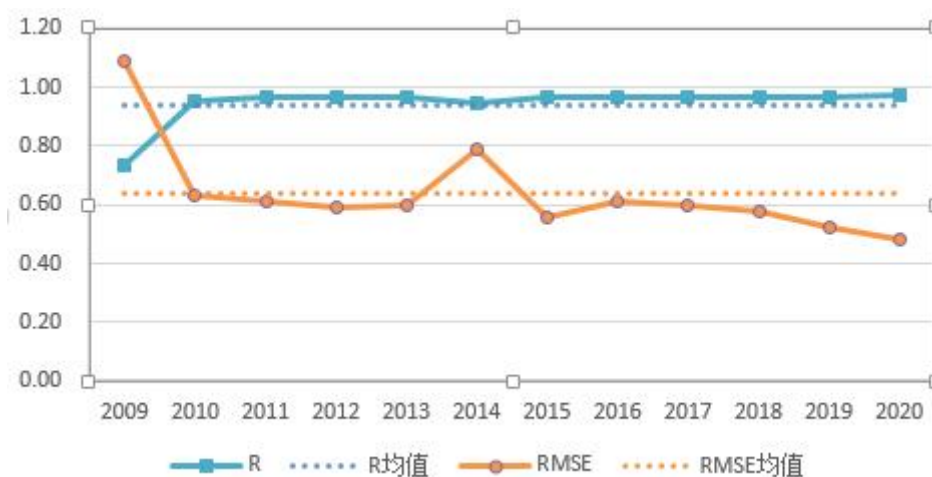
Figure 2 Correlation coefficient R between the observation data at each TCCON site and predicted data of XCO_2



AS: 所有站点相关系数 R 的均值, 各站点名以英文缩写替代

图 3 各 TCCON 站点观测数据与预测数据 XCO_2 的均方根误差 RMSE

Figure 3 Root mean square error (RMSE) between the observation data at each TCCON site and predicted data of XCO_2

图 4 XCO₂ 预测值与 GOSAT_L3 产品对比Figure 4 Comparison between predicted XCO₂ and GOSAT L3 XCO₂

4 数据价值

本数据集是 2009–2020 年全球中低纬度地区二氧化碳柱浓度数据，具有高精度、高分辨率、长时间跨度的特点。本数据集的空间分辨率为 $0.25^{\circ} \times 0.25^{\circ}$ ，与 TCCON 站点数据相比，本数据集的相关系数 R 和均方根误差 RMSE 分别可达到 0.98 和 1.38 ppm。本数据集反映了全球二氧化碳浓度长期动态变化，数据连续性强、实用性高，可应用于全球范围内长时间序列的碳源和碳汇计算。

数据作者分工职责

许静静，女，在读硕士，研究方向为激光遥感技术与二氧化碳气体。主要承担工作：论文撰写，数据质量控制和评估。

龚威，男，教授，研究方向为激光遥感技术与方法。主要承担工作：任务监督。

张劲，男，高级工程师，研究方向为激光遥感技术与方法。主要承担工作：数据制作。

张豪伟，男，硕士，在读博士，研究方向聚焦在大气遥感和二氧化碳气体的研究。主要承担工作：整体思路的设计。

马昕，男，博士，副教授，研究方向为米散射激光雷达及差分吸收探测激光雷达技术。主要承担工作：数据制作。

韩舸，男，副教授，研究方向为新型激光遥感技术研究及其在定量遥感中的应用。主要承担工作：数据制作。

参考文献

- [1] FODEN W B, YOUNG B E, AKAKAYA H R, et al. Climate change vulnerability assessment of species[J]. Wiley Interdisciplinary Reviews: Climate Change, 2019, 10.
- [2] O'NEILL B C, DONE J, GETTELMAN A, et al. The benefits of reduced anthropogenic climate change (BRACE): a synthesis[J]. Climatic Change, 2018, 146(3/4): 287–301. DOI:10.1007/s10584-017-2009-x.
- [3] SHERWOOD S C, WEBB M J, ANNAN J D, et al. An assessment of earth's climate sensitivity using multiple lines of evidence[J]. Reviews of Geophysics (Washington, D C: 1985), 2020, 58(4): e2019RG000678. DOI:10.1029/2019RG000678.
- [4] 刘毅, 王婧, 车轲, 等. 温室气体的卫星遥感—进展与趋势[J]. 遥感学报, 2021, 25(1): 53–64. [LIU Y, WANG J, CHE K, et al. Satellite remote sensing of greenhouse gases: progress and trends[J]. Journal of Remote Sensing, 2021, 25(1): 53–64.]

- [5] LIU Y, WANG J, YAO L, et al. The TanSat mission: preliminary global observations[J]. Science Bulletin, 2018, 63(18): 1200–1207. DOI:10.1016/j.scib.2018.08.004.
- [6] HAN G, XU H, GONG W, et al. Feasibility study on measuring atmospheric CO₂ in urban areas using spaceborne CO₂-IPDA LIDAR[J]. Remote Sensing, 2018, 10(7): 985. DOI:10.3390/rs10070985.
- [7] ZENG Z C, LEI L P, GUO L J, et al. Incorporating temporal variability to improve geostatistical analysis of satellite-observed CO₂ in China[J]. Chinese Science Bulletin, 2013, 58(16): 1948–1954.
- [8] LIU Y, WANG X F, GUO M, et al. Mapping the FTS SWIR L2 product of XCO₂ and XCH₄ data from the GOSAT by the Kriging method - a case study in East Asia[J]. International Journal of Remote Sensing, 2012, 33(10): 3004–3025. DOI:10.1080/01431161.2011.624132.
- [9] TOMOSADA M, KANEFUJI K, MATSUMOTO Y, et al. A prediction method of the global distribution map of CO₂ column abundance retrieved from GOSAT observation derived from ordinary kriging[C]//2009 ICCAS-SICE. Fukuoka. IEEE, : 4869–4873.
- [10] TOMOSADA M, KANEFUJI K, MATSUMOTO Y, et al. Application of the spatial statistics to the retrieved CO₂ column abundances derived from GOSAT data[C]. Proceedings of the 4th WSEAS International Conference on REMOTE SENSING (REMOTE'08), Venice, Italy, 2008.
- [11] TSUTSUMI Y, MORI K, HIRAHARA T, et al. GAW Report No. 184 Technical Report of Global Analysis Method for Major Greenhouse Gases by the World Data Center for Greenhouse Gases[R]. 2009.
- [12] YE Y, DENG J S, HUANG L Y, et al. Modeling and prediction of NPP-VIIRS nighttime light imagery based on spatiotemporal statistical method[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(6): 4934–4946. DOI:10.1109/TGRS.2020.3011695.
- [13] ZENG Z C, LEI L P, HOU S S, et al. A regional gap-filling method based on spatiotemporal variogram model of CO_2 columns[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(6): 3594–3603. DOI:10.1109/tgrs.2013.2273807.
- [14] LI J, JIA K, WEI X Q, et al. High-spatiotemporal resolution mapping of spatiotemporally continuous atmospheric CO₂ concentrations over the global continent[J]. International Journal of Applied Earth Observation and Geoinformation, 2022, 108: 102743. DOI:10.1016/j.jag.2022.102743.
- [15] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345–1359. DOI:10.1109/tkde.2009.191.
- [16] DONG Y N, DU B, ZHANG L P, et al. Dimensionality reduction and classification of hyperspectral images using ensemble discriminative local metric learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(5): 2509–2524. DOI:10.1109/TGRS.2016.2645703.
- [17] DONG Y N, LIANG T Y, ZHANG Y X, et al. Spectral-spatial weighted kernel manifold embedded distribution alignment for remote sensing image classification[J]. IEEE Transactions on Cybernetics, 2021, 51(6): 3185–3197. DOI:10.1109/TCYB.2020.3004263.
- [18] LIU C, WANG W, SUN, Y. TCCON data from Hefei, China, Release GGG2014R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. (2010).
- [19] MA X, ZHANG H W, HAN G, et al. A regional spatiotemporal downscaling method for CO₂ columns[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(10): 8084–8093. DOI:10.1109/TGRS.2021.3052215.

论文引用格式

许静静, 龚威, 张劲, 等. 2009 - 2020 年基于 GOSAT 卫星的全球中低纬二氧化碳柱浓度数据集[J/OL]. 中国科学数据, 2023, 8(3). (2023-06-30). DOI: 10.11922/11-6035.noda.2022.0003.zh.

数据引用格式

许静静, 龚威, 张劲, 等. 2009 – 2020 年基于 GOSAT 卫星的全球中低纬二氧化碳柱浓度数据集 [DS/OL]. Science Data Bank, 2023. (2022-06-30). DOI: 10.11922/sciencedb.j00001.00368.

A dataset of global CO₂ concentration in global middle and low latitudes based on GOSAT satellite data during 2009–2020

XU Jingjing¹, GONG Wei¹, ZHANG Jin², ZHANG Haowei^{3*}, MA Xin³, HAN Ge⁴

1. School of Electronic Information, Wuhan University, Wuhan 473072, P.R. China
2. Wuhan Geomatics Institute, Wuhan 430079, P.R. China
3. State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, P.R. China
4. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, P.R. China

*Email: haoweizhang@whu.edu.cn

Abstract: In response to the increasingly intensifying greenhouse effect, Countries around the world jointly signed the Paris Agreement, and China also made plans and policies of peak carbon dioxide emissions, carbon neutral plans. Carbon dioxide is the focus of international concern as the most important greenhouse gas. So, it's crucial to know how to obtain carbon dioxide concentration of high precision, high resolution temporal and spatial distribution for advancing the top-down assessment of carbon source, carbon sink and carbon neutral research. This paper creatively proposes a method to obtain high-precision and high-resolution temporal and spatial distribution map of global carbon dioxide concentration by using satellite data. We constructed a new prior time curve parameter library for fitting time domain information. In this paper, we used the transfer learning theory to integrate the time information as a prior profile into the spatial information based on the global data of GOSAT satellite. The spatial prediction information was adjusted to obtain more accurate spatio-temporal prediction of carbon dioxide concentration. The spatio-temporal resolution of the product database is 0.25°. Finally, the database has been compared with TCCON data in middle and low latitudes, which shows the correlation coefficient R and RMSE is 0.98 and 1.38 ppm of the monthly average carbon dioxide concentration respectively. The recommended database can be applied to the calculation of carbon sources and carbon sinks on a large scale.

Keywords: CO₂ concentration; GOSAT; transfer learning; global

Dataset Profile

Title	A dataset of global CO ₂ column concentration in middle and low latitudes based on GOSAT satellite data during 2009–2020
Data authors	XU Jingjing, GONG Wei, ZHANG Jin, ZHANG Haowei, MA Xin, HAN Ge
Data corresponding author	ZHANG Haowei (haoweizhang@whu.edu.cn)
Time range	2009–2020
Geographical scope	60°S – 60°N, 180°W – 180°E
Spatial resolution	0.25° × 0.25°
Data volume	11.56 GB
Data format	*.h5
Data service system	< http://doi.org/10.11922/sciencedb.j00001.00368 >
Sources of funding	National Natural Science Foundation of China (Grant No. 42171464, 41971283, 41801261, 41827801, 41801282), National Key Research and Development Program of China (2017YFC0212600), Open Research Fund of National Earth Observation

	Data Center (NODAOP2021005), and LIESMARS Special Research Funding.
Dataset composition	This dataset contains 136 files (h5 format) of monthly mean carbon dioxide concentration from 2009 to 2020 in global middle and low latitudes. The document covers the longitude and latitude of data points, spatial and temporal interpolation results, model parameters, evaluation indexes, spatial resolution and other information.