

面向遥感真实性检验的地面观测数据质量控制方法与体系

李曲曲^{1,2}, 闻建光^{1,2}, 肖青^{1,2}, 吴小丹³, 游冬琴¹, 唐勇¹,
廉婷¹, 朴森¹, 赵娜¹

1. 中国科学院空天信息创新研究院 遥感卫星应用国家工程研究中心/遥感与数字地球全国重点实验室, 北京 100101;

2. 中国科学院大学, 北京 100049;

3. 西南交通大学 地球科学与环境工程学院, 成都 611756

摘要: 遥感真实性检验的核心是获取像元尺度真值, 地面观测数据是遥感真实性检验中像元尺度参考真值的主要来源。然而, 地面观测数据在采集、传输、处理等过程中受到多种因素的影响, 导致其具有一定的误差, 从而影响像元真值的准确性和遥感真实性检验的可靠性。为此, 本研究面向遥感真实性检验像元真值获取的需求, 系统梳理了地面观测数据的主要误差来源和存在的主要问题, 从规范性、完整性、时效性、准确性、一致性和时空代表性等方面, 总结了当前地面观测数据质量控制的方法与体系, 并以地表反照率像元真值获取的地面观测数据为例, 探讨了其质量控制与评价方法的应用效果。本研究可为推动面向遥感真实性检验的地面观测数据标准化治理、提升遥感像元尺度真值质量提供技术支持。

关键词: 地面观测数据, 遥感产品, 真实性检验, 误差, 质量控制, 地表反照率

中图分类号: P2

引用格式: 李曲曲, 闻建光, 肖青, 吴小丹, 游冬琴, 唐勇, 廉婷, 朴森, 赵娜. 2026. 面向遥感真实性检验的地面观测数据质量控制方法与体系. 遥感学报, 30(3): 662-679

Li Q Q, Wen J G, Xiao Q, Wu X D, You D Q, Tang Y, Lian T, Piao S and Zhao N. 2026. Quality control methods and framework for *in situ* observation data in remote sensing validation. National Remote Sensing Bulletin, 30(3): 662-679 [DOI:10.11834/jrs.20265086]

1 引言

遥感地面观测数据作为模型算法的输入以及真实性检验中参考真值的主要来源, 是科学研究和应用的基础 (Doldirina, 2015; Balsamo 等, 2018; Xu 等, 2024)。随着遥感和地面观测技术的迅速发展, 遥感产品已逐渐成为资源、环境、生态、水文等领域的重要信息源。中国高分辨率对地观测系统通过搭载多谱段、多模式的观测载荷, 构建了覆盖可见光、红外与微波等多类型的对地观测能力, 能够持续提供遥感共性产品, 为多领域应用提供了强有力的数据支撑 (柳钦火 等, 2023)。然而, 遥感数据经由辐射定标、几何校

正、算法反演等环节均可能引入误差, 导致遥感产品本身存在一定的不确定性 (闻建光 等, 2023b), 真实性检验是评价其质量的重要手段 (晋锐 等, 2017)。用于检验的参考真值主要来源于地面观测数据 (吴小丹 等, 2019; 张圆 等, 2020), 准确可靠的地面观测数据是遥感产品真实性检验的重要参考 (张仁华 等, 2010)。由于观测仪器本身存在的误差、外部环境因素的干扰、各地面台站工作人员对观测规范的理解以及观测方法的差异, 导致观测数据的质量良莠不齐, 地面观测数据的质量控制和评价对于客观评价定量遥感产品的精度具有非常重要的意义 (李新, 2013; Chave 等, 2019)。

收稿日期: 2025-03-12; 预印本: 2025-12-09

基金项目: 国家重点研发计划(编号:2023YFB3905805)

第一作者简介: 李曲曲, 研究方向为遥感产品真实性检验。E-mail: liququ21@mails.ucas.ac.cn

通信作者简介: 闻建光, 研究方向为定量遥感。E-mail: wenjg@aircas.ac.cn

世界上很多公开发布的数据集，例如全球通量观测网络提供的 FLUXNET 2015 数据集 (Pastorello 等, 2020)、全球历史气候数据集 (Peterson 和 Vose, 1997)、寒区陆面过程试验数据集 CLPX (the Cold Land Processes Field Experiment) (Parsons 等, 2004)、国际土壤水分网络数据集 ISMN (the International Soil Moisture Network) (Dorigo 等, 2021) 等都进行了严格的质量控制。

中国在气象、生态、遥感、水文等领域建立了大量地面观测站 (闻建光 等, 2023a)，不同观测站点的站网建设由不同的科研机构 and 政府部门负责，缺乏统一的质量控制规范，造成了数据之间缺乏一致的质量基准，影响了数据的跨区域、跨平台联合使用。因此，地面观测数据的质量控制成为提升数据质量、促进科学研究和数据共享的重要课题。

质量控制 QC (Quality Control) 是指依据观测要素的时空变化特征和各观测要素之间的相互联系的规律进行分析，标注不合理的可疑数据 (刘小宁和任芝花, 2005)。质量控制作为数据治理过程中的核心环节，旨在确保数据的规范性、完整性、时效性、准确性、一致性和代表性 (国家市场监督管理总局和中国国家标准化管理委员会, 2018)。传统的质量控制方法通常包括基于物理模型的边界范围检查、一致性检查等 (Guo 和 Liu, 2015; 任芝花 等, 2015; 林兴稳, 2018; Bian 等, 2019)。这些方法通过对数据进行严格的校验与筛查，能够有效识别异常值。近年来，随着计算机科学特别是机器学习和人工智能技术的飞速发展，许多学者致力于将自动化技术与人工智能方法相结合，提出了许多创新性的质量控制框架，基于深度学习的异常检测方法已经被广泛应用于农业、气象和环境监测领域 (Dandriofosse 等, 2024)，特别是在时间序列数据的质量控制方面，如 Bandaru 等 (2024) 通过开发一种双向长短期记忆 (LSTM) 模型，能够有效识别并剔除土壤水分时间序列数据中的异常值。一些研究依赖于将测量值与气象模型 (Båserud 等, 2020) 或机器学习算法 (Kim 等, 2020; Sha 等, 2021) 的预测进行比较，可以有效发现并修正观测数据中的异常。Toreti 等 (2022) 开发了一个近乎实时的质量控制系统 QuackMe (Quality Checks of Meteorological Data)，获取高质量的每日气象信息，用于作物产量预报

系统和欧洲干旱监测。这些研究表明，质量控制方法的改进不仅能够提高数据处理效率，还能增强数据的准确性和高效利用，为后续的气象预测、农业决策等应用提供更加可靠的基础数据支持。

尽管这些质量控制方法已经在不同领域得到了应用，但大多仅关注观测数据本身的完整性与一致性等，现有的质量控制方法尚未形成统一的体系，行业内方法繁多但缺乏系统性，这使得不同研究成果的可比性和通用性较低。特别是在遥感真实性检验中，质量控制除了保证数据的有效性，还要关注其在时空尺度上是否具有代表性，已有研究表明，真实性检验结果的误差与地面观测的代表性误差之间存在显著相关性，尤其在代表性误差较大时，验证误差主要来源于地面观测在空间尺度上的代表性不足 (Wu 等, 2019)。因此，如何总结和规范现有的质量控制方法，增加在像元尺度上对数据时空代表性的约束，并提出一种更加系统化、标准化的面向真实性检验的地面观测数据质量控制流程，以提升地面观测数据用于遥感真实性检验的有效性，成为实际应用中的一项亟待解决的难题。

为此，本研究依据 Web of Science 核心合集和中国科学引文数据库，以地面观测数据、质量控制和真实性检验为关键词，搜索 2000 年—2025 年发表的 320 篇文章，首先梳理遥感地面观测数据的主要误差来源，分析目前面临的主要问题；然后根据地面观测数据的特点，总结各类质量控制与评价的方法；最后，本研究以地表反照率为例，探讨了质量控制方法的应用效果和可行性，以期地面观测数据的质量控制提供理论支持和实践指导。

2 地面观测数据主要误差来源

面向真实性检验的地面观测数据的总体误差来源主要包括观测误差和代表性误差 (李新, 2013)。

$$\varepsilon = \varepsilon_o + \varepsilon_r \quad (1)$$

式中， ε 表示地面观测数据的总体误差， ε_o 表示观测误差， ε_r 表示代表性误差。其中，观测误差包括器测误差和测量误差，这是观测过程中固有的误差类型，主要来源于仪器的有限精度、测量过程中的操作方法误差等。代表性误差是由于在有限空间分辨率下，采样点的信息不能有效代表面

元信息而产生。

2.1 器测误差

器测误差是指测量过程中由于仪器的有限精度所引起的误差，主要受到观测仪器磨损老化、零点不准、仪器校准因子使用错误、传感器污垢堆积、传感器物理定位有问题、设备能源供给、数据传输介质受损、网络延迟、外界自然环境变化等因素的影响（刘丰和郭建文，2013；Pastorello等，2014）。器测误差常常具有明显的规律性和累积性，对测量结果影响较大，可以通过使用前对仪器进行精准标定、定期检修和校准，使用中合理操作并进行检查校正，观测后对数据加入修正系数进行标定等措施以减少长期运行导致的偏差积累。此外，真实性检验场网可以记录和分析仪器运行状态及环境元数据，便于追溯异常信息来源。采用冗余观测和多传感器交叉验证，以提升异常值识别能力；人工巡检和现场记录也在发现设备故障或人为干扰中发挥着不可替代的作用，以上这些措施都可以减小器测误差的影响。

2.2 测量误差

测量误差是指由于测量方法不完善或所依据的理论不严密引起的方法误差，以及测量过程中操作不当所引起的误差，与仪器本身无关，通常可以分为系统误差和随机误差两类。系统误差由测量方法、理论模型等系统性因素引起，具有规律性和可重复性，可通过优化测量方法、改进理论模型或引入统计校正等方法来加以控制。随机误差则主要源于测量过程中的偶然因素，导致数据存在杂乱、异常、不精确等特征，其误差大小和符号往往没有一定的规律，难以直接估计，但在足够多的测量总体上一般服从高斯分布，具有一定的抵偿性（常国宾等，2024），因此可以通过多次测量取平均值的办法来降低其影响，此外，还可以通过规范操作流程、加强观测人员培训等手段，提高观测数据的准确性与可靠性。

2.3 代表性误差

代表性误差主要是指在用采样数据向其所代表的空间总体进行推断过程中，由于代表性不足而产生的推断误差，代表性误差与空间异质性密切相关（李新，2013）。其产生的原因主要包括：采样点的空间代表性不足，采样点数量有限无法

覆盖区域内的整体空间结构等等。代表性误差同样无法被完全消除，但可以通过优化空间采样进行有效控制（Wen等，2022；Wu等，2021）。样点优化布设的核心在于保证其具有最佳的全局代表性。目前优化采样的方法主要包括3大类：经典采样法（如简单随机抽样、分层抽样等）（Cochran，1977）、空间采样法（如空间系统抽样、空间三明治抽样等）（Cressie，2015；Wang等，2012）和基于模型的采样法（基于地统计克里格的抽样、基于累积代表度的抽样等）（Gruijter等，2006；Wu等，2016）。确定采样点后，在样方内进行多点重复采样，获取地表空间异质性的同时，减小偶然误差的影响。

真实性检验中还需要考虑时间代表性误差，地面观测与遥感观测存在时间差异时，地面数据能否有效反映遥感过境时刻的地表状态，即为时间代表性的评价重点。时间代表性的关键在于地物状态的时间稳定性：例如叶面积指数在数天内变化缓慢，即使观测时刻不同，其时间代表性仍然较好；而地表温度受辐射影响显著，则需要严格的时间匹配（如控制在半小时内）才能保证代表性。因此，时间代表性应结合地物变量的动态特征进行评估，而同步观测是提高时间代表性的一种重要手段。

3 面临的主要问题

由于观测数据质量问题的不可预见性，面向真实性检验的数据质量控制仍面临诸多挑战（Pastorello等，2017），主要体现在如下两个方面。

3.1 数据多源异构,自动化处理复杂

遥感地面观测数据多源异构、类型多样，以半结构化数据和非结构化数据为主（王桥，2021），人工观测的数据原始文档以手写记录和拍摄的图片为主，整理后的数据以Word、Excel格式为主，自动观测数据包括txt、xlsx、xml、dat文件、图片以及日志文档等，不同来源和格式的数据往往缺乏统一的标准，导致数据在结构、精度、时效性等方面存在差异，进一步增加了数据整合的复杂性，为后续的数据组织和质量控制带来困难。

3.2 缺乏面向真实性检验的质量控制体系

遥感地面观测数据作为真实性检验参考真值

的主要来源，其自身的质量直接决定了验证结果的可靠性。然而，目前多数质量控制方案仅关注数据内部的一致性和物理合理性，缺乏完整的面向真实性检验像元尺度的质量控制体系。在真实性检验中，除了关注数据本身的质量，还关注地面数据在遥感像元尺度的时空代表性。因此，构建一套覆盖数据预处理、质量判别、代表性分析的完整质量控制流程，已成为提升遥感真实性检验可靠性的重要方向。

4 质量控制方法体系

遥感地面观测数据既包括无线传感器网络、通量塔等连续观测的数据，又包含人工多点采集的数据（李新 等，2023）。对于地面数据的质量可以用质量元素进行描述，质量元素是指产品满足用户要求和使用目的的基本特性（曾衍伟和龚健雅，2004）。根据地面观测数据的特点，可以将一级质量元素分为规范性、完整性、时效性、准确性、一致性和时空代表性（图1），二级质量元素是对一级质量元素的具体描述，即二级质量元素是最基本的数据质量信息的载体（刘丰和郭建文，2013）。

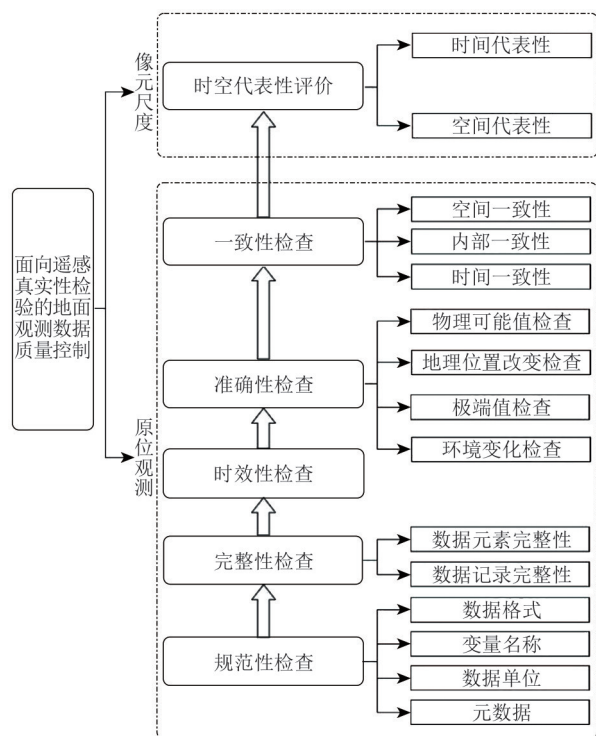


图1 面向遥感真实性检验的地面观测数据质量控制体系
Fig. 1 Quality control framework for *in situ* observation data in remote sensing validation

4.1 规范性检查

规范性检查主要是检查获取的观测数据的格式是否符合地面观测相关的国家标准规范或行业标准规范的要求，从元数据是否规范、同一类数据单位是否统一、变量名称是否规范以及同类型数据存储格式是否符合规范等方面进行检查，对于不规范的数据进行整理，处理成标准统一、格式规范的数据，方便进一步应用。

4.2 完整性检查

4.2.1 数据元素完整性

数据元素完整性检查的内容包括文件命名、观测变量、数据观测周期等的完整性（马明国，2010）。首先是对文件完整性进行检查，检查观测数据的字段信息是否完整，记录的观测数据是否缺失，如果数据有缺失，要联系相应的观测人员核实数据缺失的原因，如果是记录时遗漏数据，要及时进行补充，如果数据出现冗余，要对重复数据进行剔除。其次，检查每个变量是否按照实际的观测周期提交数据，如辐射数据观测周期为10 min，则提交的数据中每隔10 min就应该有新的观测记录，如果数据观测频次不够，则需检查仪器是否出现故障。最后，核实提供的数据是否是按照约定的相应变量对应的数据。

4.2.2 数据记录完整性

对数据的完整率、重复性等情况进行质量复查，在数据处理中，存在不同类型的空值，包括传输空值、计算空值和仪器故障空值等（数据记录如“N/A”，“NAN”，“-999”或空）。传输空值源自数据传输过程中由于信号丢失或格式转换失败所导致的数据缺失；计算空值指由其他属性数据计算而来的数据由于计算失败而产生的缺失值；仪器故障空值则是由于观测设备异常或中断，导致连续时间段内数据无法采集与传输，引发长时段的数据缺失，这些空值会对后续分析带来系统性偏差，需在质量控制中加以识别与修正。自动观测的数据可以根据仪器自带的质量标识或空值属性设置值，对数据进行过滤。对于人工观测数据的空值，需要核实是否在纸质数据整理时少录数据。

4.3 时效性检查

数据的时效性检查是指对数据的更新频率、最新性和有效性等方面进行评估的过程，是确保数据质量的一个重要方面，是及时准确的数据支持决策的前提（李默涵和李建中，2016）。数据时效性可以分为基于时间戳的时效性和基于规则的时效性（Gong等，2023）。基于时间戳的时效性是指通过检查数据的时间戳，判断数据是否在期望时间内更新，对于自动观测网络传输数据的时效性，刘丰和郭建文（2013）通过式（2）进行时效性判断，

$$rt = 1 - e^{-(t_0 - t)f} \quad (2)$$

式中， t_0 表示数据采集时间， t 表示数据接收时间，而 f 则代表数据采集频率。当 rt 趋近于0时，表示数据的时效性越强；而当 rt 趋近于1时，则意味着数据的时效性较差。

基于规则的时效性，是指根据业务需求或应用场景，判断数据是否及时更新。对于人工观测的地面数据，可以从数据的汇总频率、入库更新及时性等方面进行时效性评估，例如监测植被生长状态的叶面积指数，人工观测需要能够反映目

标时间段的真实状态，通常按周记录数据，保证日期范围内记录的数量或频率分布满足业务需求（Gong等，2023）。可以通过规定观测周期数据缺失率 mr 进行判断：

$$mr = 1 - \frac{m}{M} \quad (3)$$

式中， M 表示规定日期范围内计划观测数据量， m 表示数据实际采集量。 mr 越趋近于0，表明数据的时效性越强；而当 mr 趋近于1时，则数据时效性较差。

4.4 准确性检查

4.4.1 物理可能值检查

物理可能值检查用于确保观测数据在合理的物理范围内，通过预先设定一组合理的物理范围或阈值，来过滤或标记超出合理范围的异常数据。如地表反照率数据，其物理可能值范围限定为（0，1），基于植被指数计算的干旱指数，取值范围（-1，1）。如果观测值不在该范围内，则该数据被舍弃或标记为异常数据，且不能参与其后相关参数的计算。表1列出了常见的几种遥感地面观测参数的物理可能值范围。

表1 常见的几种遥感地面观测参数的物理可能值范围

Table 1 Physically possible value ranges for common remote sensing *in situ* observation parameters

参数	物理可能值	单位	参数	物理可能值	单位
地表反射率	(0, 1)		植被指数	(-1, 1)	
归一化离水辐射率	(0, 1)		叶面积指数	(0, 10)	m ² /m ²
地表发射率	(0, 1)		植被覆盖度	(0, 1)	
气溶胶光学厚度	(0, 10)		光合有效辐射吸收比例	(0, 1)	
地表反照率	(0, 1)		森林地上生物量	(0, 1000)	t/hm ²
光合有效辐射	(0, 1000)	W/m ²	蒸散发	(0, 10)	mm/d
地表温度	(150, 350)	K	土壤水分	(0, 1)	m ³ /m ³

4.4.2 极端值检查

极端值检查主要是根据先验知识建立地面观测数据范围的“软边界”，如果观测数据超过由先验知识建立的边界范围，则可以认为该数据是可疑数据。极端值检查的关键在于合理设定最大极端值和最小极端值（赵虹，2015），如果最大极端值设定过高或者最小极端值设定过低，可能无法有效检测和排除异常数据。相反，如果最大极端值设定过低或者最小极端值设定过高，则可能错误地排除正常的地面观测数据，使得质量控制的

结果并不合理。

在传统的极端值检查中，通常采用数据的平均值加减4倍标准差来确定其最大和最小极端值（任芝花等，2006）。拉依达准则（又称为 3σ 准则）（Wang等，2018；Wang等，2021）进一步缩小了极端值可能的范围，假设有一系列样本数据 $X = (x_1, x_2, \dots, x_i, \dots, x_m)$ ，其中 X 的均值为 μ ，标准差为 σ ，则如果满足下式（4），即当某个数据与平均值的偏差 >3 倍标准差（ 3σ ）时，则认为该数据可疑。

$$|x_i - \mu| \geq 3\sigma \quad (4)$$

这种方法的假设前提是数据量足够多，即 m 足够大，且数据满足正态分布，当数据量较少时，可能会删除正常的的数据，这种可能性会随着数据量的增加而降低。另外，当某个异常点或者某些异常点特别大时，会导致判定异常点的极端值范围 3σ 也特别大，此时，只有明显异常的数据才能被剔除，稍微异常的数据可能都会落在极端值范围内，难以被有效检测，从而使得采用拉依达准则方法的效率比较低。

Guo 和 Liu (2015) 改进了 3σ 准则的算法，通过扩展时间序列数据并对其进行多项式拟合以及计算标准偏差 σ ，比较每个数据值与拟合函数的残差绝对值是否 $>3\sigma$ ，来判断数据是否异常。改进后的 3σ 准则在计算中使用数据序列的拟合值可以准确的模拟数据的趋势，通过残差识别异常数据，可以提高可疑数据的检测精度。

4.4.3 环境变化检查

环境变化检查主要是针对地面数据中所有与大气相关的辐射参数的观测过程中，云的变化对观测结果的影响，有效识别云的变化，给出有云的标志，并对其影响进行订正。

判断天气是否有云最直接的方法是目视观测，实验人员可以通过直接观察天空的云量来实时记录天气状况，这种方法在短时间内是可行的，但对于需要长时间连续观测的野外站点而言，难以实现（吴小丹等，2019）。随着现代人工智能技术的发展，利用深度学习的方法，将全天空成像系统自动拍摄的云图片，进行晴空和有云分类，并确定云量，进而对应当天数据是否可用（Kazantzidis 等，2012；Kalkan 等，2022）。Li 等（2011）通过混合阈值算法综合了固定和自适应阈值方法的优点，将彩色云图归一化为蓝/红通道比图像，根据其标准偏差分别由不同的阈值算法处理，提高了云检测的准确率。

当云信息没有获取或不可用时，可以利用辐射数据进行判断。由于在有云的天气条件下，下行短波辐射值低于晴天的下行短波辐射值，因此可以利用这一特征来标记或剔除有云天气的数据（吴小丹等，2019）；通过辐射传输模拟计算得到理论上的晴空下行辐射，当地表实际入射太阳辐射与其比值 >0.8 时，可以认为这种情况是晴空

（Papachristopoulou 等，2024）。

4.4.4 地理位置改变检查

由于城市建设和发展、自然灾害、气候变化、土地利用类型改变以及其他人为活动的影响，导致原站点观测环境显著变化，不能满足观测要求，影响到数据的准确性，使得数据不具有代表性，导致站点搬迁到一个新的位置以确保观测数据的准确性。站点迁移前后观测环境发生了明显的变化，对数据的均一性影响非常显著。常用的检验由于地理位置改变导致观测资料非均一性的方法有基于惩罚最大 F 检验（PMFT）（Wang，2008）、惩罚最大 t 检验（PMT）（Wang 等，2007）和标准正态均一性检验（SNHT）方法（Pandžić 等，2020；韩海涛等，2021）。

SNHT 法使用参考序列能够有效识别出断点，但检验时的置信度容易受到断点位置的影响，当断点位于时间序列开始和结束两端时，发生错误检测的概率较高，即置信度较低（Toreti 等，2011）。PMFT 和 PMT 两种方法分别通过对比变化点前后的方差和均值来检测变化点，通过引入惩罚项有效减少了误报，但在数据噪声较大的情况下，方差或均值的变化可能会掩盖掉真实的变化点。PMT 方法需要有一个时间序列作为参考序列，但不可避免地会受到参考序列中非均一要素的影响，可能判断出虚假的断点，而 PMFT 方法适用于不带参考站的均一性检验（Wan 等，2017），避免了参考序列的不合理和非均一性因素的影响，但忽略了台站本身的趋势特征（肖晶晶等，2018）。

4.5 一致性检查

4.5.1 内部一致性检查

内部一致性检查，又可以称之为数据的比较检查，旨在识别同一时刻或同一时间段内观测的多个相关变量之间是否遵循既定的物理规律（Dandriofosse 等，2024）。例如：日最低气压应低于日平均气压，日平均气压又应低于日最高气压；同样，地面日最低气温应低于地面日平均气温，而地面日平均气温应低于地面日最高气温。上述关系体现了变量之间的物理一致性，若观测数据不满足这些基本的物理逻辑关系，则反映出观测数据可能存在异常或测量误差，因此内部一致性检查常作为数据质量控制流程中的关键环节，用

于快速筛查和标记可疑数据。

4.5.2 时间一致性检查

时间一致性检查是指在某一连续时间段内的观测数据应该保持一个稳定平缓的变化,而不是有突然的变化存在,即观测的要素随着时间的变化应该是稳定的(赵虹, 2015)。时间一致性检查的目的是通过检查观测要素随时间的变化来判断该要素在一定时间范围内的变化是否符合特定规律。如果在观测时间内,数据超出了允许的最大变化阈值,本研究则认为该数据为可疑数据。

利用观测站的资料进行时间内插预估,比较预估值与实测值进行时变检查,有效过滤长时间序列中的随机噪声(Steinacker, 2023),还可以通过子序列异常检测方法,将时间序列数据通过滑动窗口选取多个固定长度的子序列,分别提取子序列的均值、标准差、偏度、峰度等统计特征,通过设定动态阈值判断子序列是否异常。该方法的优点是可以使每个相对的滑动窗口内保留原有的时间序列属性(杨寒雨等, 2023)。

4.5.3 空间一致性检查

空间一致性检查主要是检查观测数据在一定的空间范围内的变化是否符合其空间规律。根据地理学第一定律,很多地面观测变量在地理空间上存在着相关性,如果地面观测站点空间距离较近,他们的观测结果往往比较相近(卢宾宾等, 2025)。

与下垫面无关的参数(温度、湿度等),可以通过计算估计值与观测值之差来判断,常用的估计值计算方法有反距离加权插值(Lu和Wong, 2008; 赵虹, 2015),多项式内插(Guidoum, 2025),最优插值(Miatselskaya等, 2022),空间回归(Cerlini等, 2020)等。值得注意的是,在方法选择上需要考虑观测网的密度、地表的时空变异特征等,如反距离加权插值的方法计算高效,适用于实时业务系统,而空间回归方法通过引入地理协变量(如高程、坡向)可解释性更优。

4.6 时空代表性评价

4.6.1 时间代表性评价

时间代表性定义为在给定时空域内的一组地面测量,在多大尺度上反映了特定应用尺度上的

实际情况(Nappo等, 1982)。时间代表性主要是评价观测参数的测量周期是否足以捕捉气候和生物条件的自然变化,然而,时间代表性的合理性不能仅依赖于观测时间的长短,需要综合考虑特定参数的时间序列特征、数据的完整性和时间相关性以及整体时间覆盖情况等(Chu等, 2017; Baca-López等, 2021)。对于某些观测参数,如土壤热通量、蒸散发等,观测数据需要至少覆盖一个完整的生长季,才能保证验证的时间代表性。

近年来,已有研究提出了多种时间代表性的评价方法,包括趋势分析(Chang等, 2021)、周期分析(Assfalg等, 2009)、样本均值和标准差比较(Anttila等, 2012)、蒙特卡洛模拟(Sharp等, 2015)、滑动窗口分析(杨寒雨等, 2023)以及相关性分析等(Chu等, 2017)。这些方法各有侧重,能够从不同角度评估数据的时间代表性。例如,趋势分析通过分析长期变化趋势来判断某一时段的代表性,判断站点观测期内是否能够充分捕捉到驱动变量的平均状况和年际变异性(Chu等, 2017; Xu等, 2024);周期分析则着重于捕捉数据中的周期性特征;样本均值和标准差比较方法则从统计角度分析数据的代表性(Anttila等, 2012);此外,蒙特卡洛模拟通过大量随机抽样计算来评估时间代表性。

4.6.2 空间代表性评价

空间代表性是指在一个特定的时空维度上测量的样本数据,能够在多大程度上反映在这个特定范围下的真实情况(徐保东等, 2015)。由于观测塔的设计从来没有最大限度地扩大空间覆盖范围,因此人们通常考虑从观测位置向外扩展到特定的范围问题,即空间代表性(Schimel等, 2015)。在使用地面观测数据对遥感产品进行建模、反演、验证等应用时,应对其进行空间代表性评价,确保数据具有较好的空间代表性。观测数据是否具有空间代表性主要影响因素是空间异质性导致的点面特征不匹配。

半方差函数模型是常用的地表空间异质性衡量方法,说明站点观测在像元尺度的空间代表性(Curran, 1988; Román等, 2009; Wang等, 2014; Wu等, 2022; Li等, 2025; 吴小丹等, 2025),只要空间数据满足平稳性假设,可用半方差函数描述和量化其空间变异性和自相关性,当研究区

内空间异质性较小时，说明该区域内地物分布较均匀，站点代表性较好（马晋，2021）。

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Z(x_i) - Z(x_i + h))^2 \quad (5)$$

式中， $\gamma(h)$ 表示半方差函数， $Z(x_i)$ 表示位置 x_i 处的样本值， h 表示任意2个样本之间的空间距离， $N(h)$ 表示距离为 h 时的点对数目。

为了消除目标变量本身数值大小的影响，Xu等（2016）定义了描述相对异质性的指标基台系数（CS）：

$$CS(s) = \frac{\sqrt{c_0(s) + c(s)}}{\bar{z}(x(s))} \times 100 \quad (6)$$

式中， s 代表像元空间分辨率的大小， $\bar{z}(x(s))$ 表示目标变量在统计区域内的平均值， $c_0(s)$ 和 $c(s)$ 分别是块金值和偏基台值。CS越小，表明地表的空间异质性越弱，站点的空间代表性越好（徐保东，2018）。

点面特征比较方法通过比较站点观测值和代表区域范围内数据的统计指标（如均值）之间的关系，如果小于一定的阈值，则认为站点观测能够代表相应的区域特征（Wu等，2022）。针对植被参数的特点，Xu等（2016）设计了能够定量衡量点面一致性的指标：主要植被类型比例（DVTP）相对绝对误差（RAE），用于对地面观测 LAI 的空间代表性进行评价，上述指标的表达式分别为

$$DVTP(s) = \frac{A(s)}{\sum_{i=1}^{N(s)} n_i(s)} \times 100 \quad (7)$$

$$RAE(s) = \frac{|z(x_s) - \bar{z}(x(s))|}{\bar{z}(x(s))} \times 100 \quad (8)$$

式中， s 代表像元大小， $A(s)$ 代表站点观测的植被类型在低分像元内的面积， $n_i(s)$ 代表第 i 种植被类型的面积， $N(s)$ 代表像元内总的植被类型个数， $z(x_s)$ 代表站点位置对应在植被参数趋势面的像元值， $\bar{z}(x(s))$ 代表像元尺度内植被参数的均值。当DVTP越高、RAE越小时，说明点面一致性越好，空间代表性越好。

在对地面数据的代表性进行评价时，往往时间和空间两个维度是密不可分的，时空代表性反映了对整个研究区域及其时间演变过程的代表性程度。常用的方法包括欧氏距离（Su等，2022；He等，2015；Wu等，2021）、余弦相似度（Wu等，2016）等，可以定量分析数据在空间和时间

维度上的相关性和差异性。

5 应用实例

5.1 研究区及数据

地表反照率是调节地表辐射收支和地气相互作用的关键参数之一（Dickinson，1995），被广泛应用于各种气候预测模型、地表能量平衡和全球气候变化的研究中（Han等，2024）。本研究以地表反照率数据为例，介绍其质量控制与评价。

研究区为中国科学院怀来遥感综合试验站的林地观测区（40.3478°N，115.7768°E），该观测区内地物为人工林，地势平坦，主要有常绿针叶植被油松、阔叶落叶植被海棠和观赏性植被枫树，林地内布设有7个辐射数据地面观测站点，如图2所示。站点的具体地理位置如表2所示。

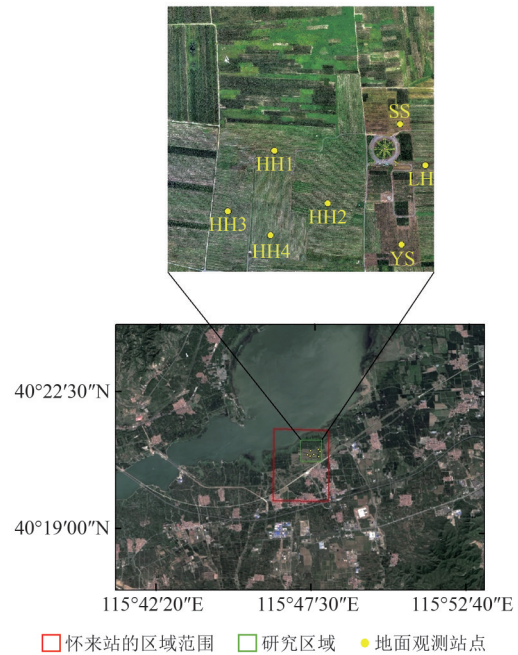


图2 怀来试验站 Landsat 8 真彩色(波段4-3-2)合成图像
Fig. 2 True-color composite image(bands 4-3-2)of Huailai Station from Landsat 8

所用地面数据的观测时间范围为2020年5月15日—2021年11月25日，数据类型为上、下行短波辐射数据，该数据由CMP3辐射表自动测得，时间分辨率为10 min，辐射表经过了严格的标定，由观测仪器本身带来的不确定性可以忽略不计。辐射表的架设高度为5 m，观测足迹直径范围为30 m。将测得的上行短波辐射与下行短波辐射作比即可得到地表反照率。本研究中选择太阳天顶

角最小的当地正午时刻前后半小时的数据来代表当日地表反照率的观测值。

表2 怀来站7个地面观测站点的地理位置

Table 2 Geographical locations of the 7 *in situ* observation sites at Huailai station

站点名	经度/(°E)	纬度/(°N)
HH1	115.7903	40.35093
HH2	115.7922	40.34953
HH3	115.7887	40.34929
HH4	115.7902	40.34866
LH	115.7956	40.35059
SS	115.7947	40.35169
YS	115.7948	40.34846

本案例开展地面观测数据的质量控制面向的是500 m尺度的正午时刻地表反照率产品的真实性检验,如MODIS反照率产品MCD43A3(Schaaf等,2002)等,因此在时间尺度上地面数据对应选择正午时刻的观测数据,由于地面站点观测尺度和500 m分辨率产品的像元尺度差异大,借助高分地表反照率产品(游冬琴等,2023)作为中间参考,表征粗分辨率尺度像元内的空间异质性。

5.2 数据质量控制

5.2.1 规范性检查

对地表反照率数据的格式、内容、变量名称进行规范性检查,观测资料符合规范。

5.2.2 完整性检查

对数据的完整率、重复性、观测周期等情况进行质量复查,经检查发现,在红色海棠3(HH3)站点出现数据间断情况,2020年7月3日—2020年7月8日数据缺失,经过向数据观测人员核实,确定该时间段内HH3站点供电出现问题,导致观测数据缺失。其余站点的观测数据符合文件完整性、周期完整性以及变量完整性的要求。

5.2.3 时效性检查

由于获取的数据是由站点提供的历史数据,不是实时传输的观测数据,因此略过时效性检查。

5.2.4 准确性检查

地表反照率的物理可能值范围限定为(0,

1),一旦超过了这个范围,将其视为异常数据。在极端值检查中,结合观测要素的实际情况,计算数据的平均值 ± 2.5 倍标准差作为最大/最小极端值(Peterson和Vose,1997)。图3展示了其中一个站点HH1的极端值检查结果。经过上述检查发现,7个站点的数据中,共计17 d的数据未通过极端值检查。

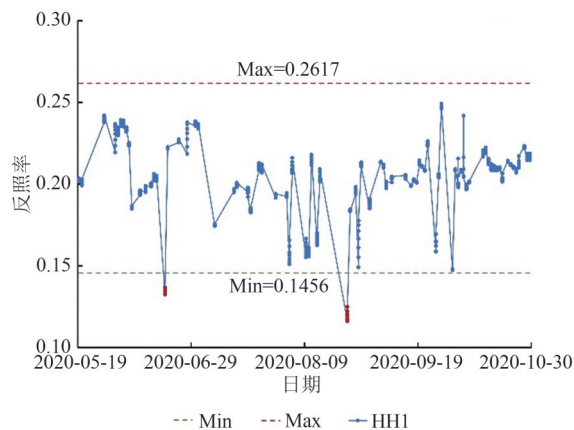


图3 红色海棠1(HH1)站点反照率极端值检查结果(红色表示数据异常)

Fig. 3 Outlier inspection results of albedo at HongHaitang 1 (HH1) site

5.2.5 一致性检查

在本研究中,使用的实验数据是上行短波辐射和下行短波辐射数据,按照内部一致性检查的要求,应满足上行短波辐射小于下行短波辐射这一物理条件,而上行短波辐射与下行短波辐射的比值即为地表反照率,在物理可能值检查过程中,通过界定地表反照率的物理范围为0—1,已经筛选出满足上行短波辐射 $\leq$ 下行短波辐射这一条件的数据,因此,不再做重复检查,数据通过内部一致性检查。

时间一致性检查所采用的判别方法是中位值平均滤波法,取地表反照率在当日正午时刻前后各半小时内的相对偏差不超过5%,以此来判断数据在当日正午时刻前后各半小时内是否稳定。通过检查发现,在2020年5月—2020年11月,除SS站点外,其他6个地面观测站点的数据中均存在异常值,此次检查中共计29个时次的数据没有通过质量检查,数据异常率为0.715%。结合历史气象资料数据分析,出现异常值的原因是由于观测时

刻出现云遮挡,标记该异常时次的的数据,对通过时间一致性检查的每日观测数据取平均值,作为当日地表反照率的地面实测值。

选取与待检验站点下垫面一样或者相似的两个及以上站点的观测数据,计算它们的平均值,将平均值与待检验站点的观测数据值进行对比,如果偏差超过了给定的阈值(本研究设为15%),则认为该数据为可疑数据。经过一致性检查后的数据如图4所示。

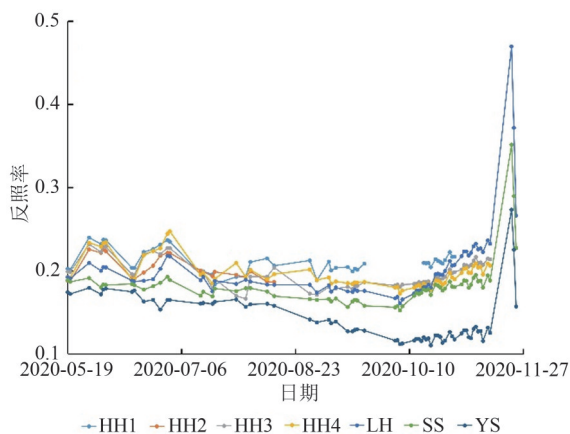


图4 怀来站地表反照率参考数据集

Fig. 4 Reference dataset of surface albedo at Huailai station

5.2.6 代表性评价

为了评价地面站点观测是否能够代表16 m分辨率的高分反照率,本研究中利用更高分辨率的数据(高分二号多光谱数据4 m分辨率)在地面观测站点周围的近红外波段DN值的空间分布特征来检验站点所在高分16 m像元内部是否均一,用该波段的原因主要是考虑到近红外波段受大气影响较小。除2021-03-08是8 m的空间分辨率外,其余时间的均为4 m空间分辨率。2021-03-08对应一个高分16 m像元应该为2×2个像元,考虑到可能存在几何偏差,统计3×3个像元值;其余时间对应高分16 m像元应为4×4个像元,考虑几何偏差,统计5×5个高分像元。

图5展示了红色海棠3(HH3)站点所在高分16 m像元在4个典型地表状态时期的DN值分布。所选日期涵盖了2020-06-02(林木生长初期,海棠花期)、2020-08-10(生长繁盛期,完全植被覆盖)、2020-11-22(凋落期,林地与裸土混合)以及2021-03-08(冬末裸土为主)。结果显示,在不同地表状态下,该站点所在高分16 m像元内部的DN值分布仍较稳定,具有良好的空间均一性。

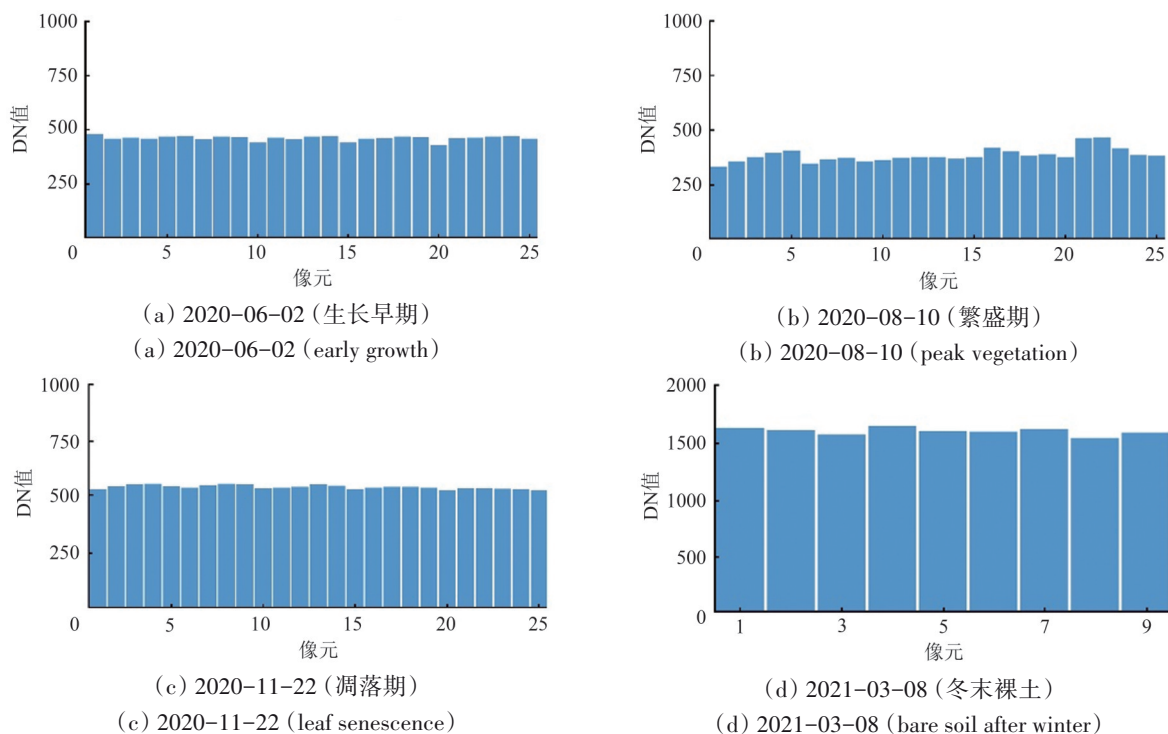


图5 HH3站点所在高分16 m像元内DN值在4个典型时期的分布情况

Fig. 5 Distribution of DN values within the 16 m pixels of Gaofen pixels containing the HH3 site during four representative periods

空间代表性是借助高分 16 m 反照率数据进行计算的, 挑选 3 景典型时期的高分一号反照率数据影像作为代表, 以各个站点所在高分像元为中心, 分别向外扩充 33×33 个像元, 对应地表 500 m×500 m 的空间范围, 利用半变异函数模型计算半方差, 进而评价地面站点观测的空间代表性。图 6 展示了红色海棠 3 (HH3) 站点半变异函数模拟的结果, 表 3 给出了 7 个站点的空间代表性评价的相关定量指标。结合图 6 和表 3 中可见: 研究区域内不同站点的空间异质性情况不同, 且随着时间不断变化, 在植被生长初期 (2020-06-04), 基台值和基台系数较大, 空间代表性相对较弱, 在植被生长旺季 (2020-07-15), 地表全部为林地覆盖, 各站点的基台值和基台系数都比较小, 表明站点的

空间异质性较弱, 空间代表性较好; 在植被生长后期 (2020-11-02), 树叶凋零, 林地和裸土混合, 空间代表性相对植被生长旺季减弱。此外, 7 个地面观测站点中, HH1、HH3、LH 和 SS 明显符合以上生长规律; 而对于 HH2、HH4 以及 YS, 则在植被生长旺季表现为空间异质性较强于其他时期, 结合研究区实际情况, 主要有 2 个原因: 一是两个海棠站点的花期较晚, 在 2020-07-15 这一时段长势不均匀; 二是油松所在区域种植面积较小, 相比之下, 空间异质性在生长旺季表现更明显。整体来看, 在植被生长季内, 各个站点的基台值较小, 地表空间异质性较弱, 地表相对比较均一, 空间代表性较好。

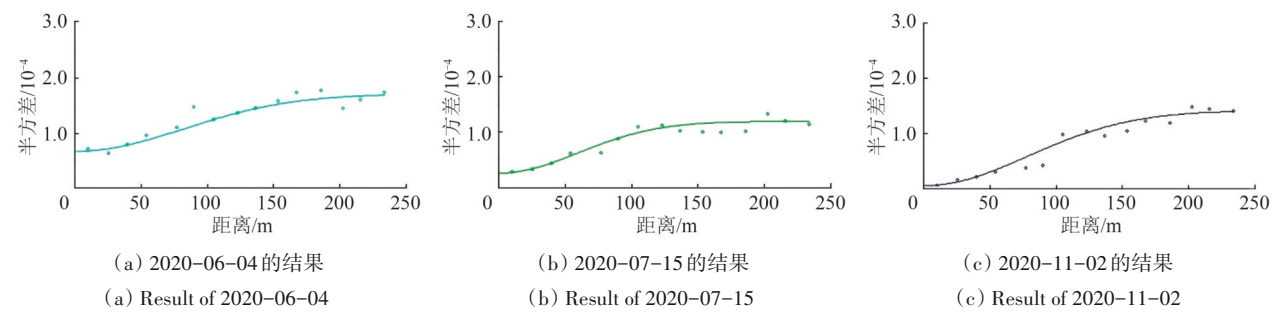


图 6 红色海棠 3(HH3) 站点半变异函数模拟结果图
Fig. 6 Simulation results of semivariogram at HongHaitang 3(HH3) site

从表 3 可以看出, 半变异函数模拟的结果, 基台值和基台系数大部分情况下是相一致的, 表现为基台值越大, 基台系数也越大。但是, 在绿色海棠 (LH) 站点不满足以上规律, 2020-07-15 的

基台值小于其他时间的, 而该时期对应的基台系数却大于 2020-11-02 时期的基台系数, 出现这种情况的原因是因为在进行半变异函数拟合的过程中, 基台值的计算会受到地表反照率值本身的影响。

表 3 研究区实验期间空间代表性定量指标

Table 3 Quantitative metrics of spatial representativeness in the study area during the experimental period							
站点	日期	块金值	偏基台值	基台值	变程	平均值	基台系数
HH1	2020-06-04	7.21E-05	4.41E-05	1.16E-04	49.584	0.16779	6.4228
	2020-07-15	2.17E-05	7.50E-05	9.67E-05	67.527	0.15923	6.1754
	2020-11-02	6.31E-06	2.55E-04	2.61E-04	119.260	0.16710	9.6710
HH2	2020-06-04	7.79E-05	8.96E-05	1.68E-04	89.835	0.16323	7.9296
	2020-07-15	1.79E-05	2.09E-04	2.27E-04	93.721	0.15894	9.4781
	2020-11-02	7.11E-06	9.73E-05	1.04E-04	93.393	0.15982	6.3949
HH3	2020-06-04	6.50E-05	1.03E-04	1.68E-04	116.250	0.16180	8.0201
	2020-07-15	2.57E-05	9.34E-05	1.19E-04	84.771	0.15967	6.8342
	2020-11-02	5.95E-06	1.36E-04	1.41E-04	111.060	0.16382	7.2612
HH4	2020-06-04	5.24E-05	1.20E-04	1.73E-04	73.851	0.16152	8.1331
	2020-07-15	2.98E-05	1.78E-04	2.08E-04	97.062	0.16128	8.9376
	2020-11-02	4.84E-06	9.02E-05	9.51E-05	104.150	0.16026	6.0841

							续表
站点	日期	块金值	偏基台值	基台值	变程	平均值	基台系数
LH	2020-06-04	5.43E-05	8.25E-05	1.37E-04	36.326	0.16296	7.1788
	2020-07-15	1.04E-05	9.50E-05	1.05E-04	60.699	0.15322	6.7001
	2020-11-02	1.53E-05	9.59E-05	1.11E-04	73.446	0.15800	6.6743
SS	2020-06-04	3.78E-05	9.46E-05	1.32E-04	31.620	0.16617	6.9235
	2020-07-15	8.52E-06	7.18E-05	8.03E-05	61.117	0.15326	5.8481
	2020-11-02	1.39E-05	1.59E-04	1.73E-04	90.989	0.15763	8.3530
YS	2020-06-04	5.85E-05	7.92E-05	1.38E-04	50.537	0.15951	7.3581
	2020-07-15	1.11E-05	2.10E-04	2.21E-04	80.844	0.15650	9.4953
	2020-11-02	1.50E-05	8.46E-05	9.96E-05	108.570	0.15206	6.5641

响，而计算基台系数，则消除了地表反照率本身数值大小的影响，从而保证不同站点的观测都具有可比性。

6 结 论

地面观测数据的质量控制是确保数据可靠性与科学性的关键环节，随着观测技术的发展，观测手段不断丰富，数据量大幅增加，这对质量控制提出了更高的要求。本研究通过对地面观测数据质量控制的方法、技术及其应用实例的探讨，总结了当前的研究进展。质量控制的核心在于确保数据的规范性、完整性、时效性、准确性、一致性和代表性。传统的质量控制方法主要依赖于人工检查和基本的统计分析。随着质量控制技术的发展，能够在海量数据中快速识别出异常和不一致的数据点，从而降低人工干预的需求。

(1) 面向真实性检验的地面观测数据的质量控制，是在传统地面观测数据质量控制基础上的一次深化与拓展。除了常规的规范性、一致性等检查外，主要体现在数据的时空代表性以及从站点尺度向像元尺度推算过程中引入的不确定度，这一过程不仅涉及对观测数据本身的处理，还要综合考虑站点的空间分布、变量在特定区域内的分布特征、数据的时间跨度等多方面的因素，以确保数据质量控制的全面性和科学性，为真实性检验提供可靠的像元尺度参考真值奠定基础。

(2) 地面观测数据质量控制的标准规范需要进一步加强。虽然当前已经形成了面向真实性检验的地面观测数据的质量控制技术体系，但各行业对地面观测数据的质量要求差异，如何定义符合应用需求的地面观测数据质量，需要根据需求

制定数据质量标准和控制方法。同时，推动各观测网络采用统一的数据格式规范与元数据标准，增强数据的互操作性与可集成性。

(3) 推动质量控制方法体系在真实性检验系统平台的集成与落地，是实现大规模、高效率数据管理与应用的关键方向。在进一步完善数据质量控制标准体系的基础上，借助现有的数据管理系统和高性能计算框架，根据不同的地表参量和观测条件，发展具备可配置性与自适应能力的质量控制模块，实现地面观测数据的自动校验与分析，解决真实性检验对地表参量从点到面元的观测需求，降低人工干预成本，提升数据处理的时效性与可靠性。

(4) 本研究构建的面向真实性检验的质量控制指标体系，重点在于识别并剔除不满足真实性检验要求的观测数据，以保证验证数据的可靠性与代表性。需要指出的是，本研究聚焦的是质量控制而非质量评价，尚未涉及各类指标结果的融合与加权评分。未来工作将基于本体系进一步发展数据质量评价方法，建立统一的加权评分机制，以实现遥感真实性检验地面数据整体质量的定量评价与分级。

参考文献(References)

Anttila S, Ketola M, Vakkilainen K and Kairesalo T. 2012. Assessing temporal representativeness of water quality monitoring data. *Journal of Environmental Monitoring*, 14(2): 589-595 [DOI: 10.1039/C2EM10768F]

Assfalg J, Bernecker T, Kriegel H P, Kröger P and Renz M. 2009. Periodic pattern analysis in time series databases//14th International Conference on Database Systems for Advanced Applications. Brisbane: Springer: 354-368 [DOI: 10.1007/978-3-642-00887-0_32]

- Baca-López K, Fresno C, Espinal-Enríquez J, Martínez-García M, Camacho-López M A, Flores-Merino M V and Hernández-Lemus E. 2021. Spatio-temporal representativeness of air quality monitoring stations in Mexico City: implications for public health. *Frontiers in Public Health*, 8: 536174 [DOI: 10.3389/fpubh.2020.536174]
- Balsamo G, Agusti-Panareda A, Albergel C, Arduini G, Beljaars A, Bidlot J, Blyth E, Bousserez N, Boussetta S, Brown A, Buizza R, Buontempo C, Chevallier F, Choulga M, Cloke H, Cronin M F, Dahoui M, De Rosnay P, Dirmeyer P A, Drusch M, Dutra E, Ek M B, Gentile P, Hewitt H, Keeley S P E, Kerr Y, Kumar S, Lupu C, Mahfouf J F, McNorton J, Mecklenburg S, Mogensen K, Muñoz-Sabater J, Orth R, Rabier F, Reichle R, Ruston B, Pappenberger F, Sandu I, Seneviratne S I, Tietsche S, Trigo I F, Uijlenhoet R, Wedi N, Woolway R I and Zeng X B. 2018. Satellite and in situ observations for advancing global earth surface modelling: a review. *Remote Sensing*, 10(12): 2038 [DOI: 10.3390/rs10122038]
- Bandaru L, Irigireddy B C, Pvnir K and Davis B. 2024. DeepQC: a deep learning system for automatic quality control of in-situ soil moisture sensor time series data. *Smart Agricultural Technology*, 8: 100514 [DOI: 10.1016/j.atech.2024.100514]
- Båserud L, Lussana C, Nipen T N, Seierstad I A, Oram L and Aspelien T. 2020. TITAN automatic spatial quality control of meteorological in-situ observations. *Advances in Science and Research*, 17: 153-163 [DOI: 10.5194/asr-17-153-2020]
- Bian Z Q, Chong W, Ding L and Yang W. 2019. Analysis and research on quality control method of global radiation observation data. *The Journal of Engineering*, 2019(23): 8975-8979 [DOI: 10.1049/joe.2018.9161]
- Cerlini P B, Silvestri L and Saraceni M. 2020. Quality control and gap-filling methods applied to hourly temperature observations over central Italy. *Meteorological Applications*, 27(3): e1913 [DOI: 10.1002/met.1913]
- Chang G B, Zhang S B and Liu Z P. 2024. Understanding adjustment from an information theoretic perspective. *Geomatics and Information Science of Wuhan University*, 49(2): 313-323 (常国宾, 张书毕, 刘志平. 2024. 从信息论角度理解间接平差. *武汉大学学报(信息科学版)*, 49(2): 313-323) [DOI: 10.13203/j.whugis20210383]
- Chang K L, Schultz M G, Lan X, McClure-Begley A, Petropavlovskikh I, Xu X B and Ziemke J R. 2021. Trend detection of atmospheric time series: incorporating appropriate uncertainty estimates and handling extreme events. *Elementa: Science of the Anthropocene*, 9(1): 00035 [DOI: 10.1525/elementa.2021.00035]
- Chave J, Davies S J, Phillips O L, Lewis S L, Sist P, Schepaschenko D, Armston J, Baker T R, Coomes D, Disney M, Duncanson L, Hérault B, Labrière N, Meyer V, Réjou-Méchain M, Scipal K and Saatchi S. 2019. Ground data are essential for biomass remote sensing missions. *Surveys in Geophysics*, 40(4): 863-880 [DOI: 10.1007/s10712-019-09528-w]
- Chu H S, Baldocchi D D, John R, Wolf S and Reichstein M. 2017. Fluxes all of the time? A primer on the temporal representativeness of FLUXNET. *Journal of Geophysical Research: Biogeosciences*, 122(2): 289-307 [DOI: 10.1002/2016jg003576]
- Cochran W G. 1977. *Sampling Techniques*. 3rd ed. New York: John Wiley and Sons
- Cressie N A C. 2015. *Statistics for Spatial Data*. Hoboken: John Wiley and Sons
- Curran P J. 1988. The semivariogram in remote sensing: an introduction. *Remote Sensing of Environment*, 24(3): 493-507 [DOI: 10.1016/0034-4257(88)90021-1]
- Dandirfossé S, Jago A, Huart J P, Michaud V, Planchon V and Rosillon D. 2024. Automatic quality control of weather data for timely decisions in agriculture. *Smart Agricultural Technology*, 8: 100445 [DOI: 10.1016/j.atech.2024.100445]
- Dickinson R E. 1995. Land processes in climate models. *Remote Sensing of Environment*, 51(1): 27-38 [DOI: 10.1016/0034-4257(94)00062-R]
- Doldirina C. 2015. Open data and earth observations: the case of opening up access to and use of earth observation data through the global earth observation system of systems. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 6(1): 73-85
- Dorigo W, Himmelbauer I, Aberer D, Schremmer L, Petrakovic I, Zappa L, Preimesberger W, Xaver A, Annor F, Ardö J, Baldocchi D, Bitelli M, Blöschl G, Bogaen H, Brocca L, Calvet J C, Camarero J J, Capello G, Choi M, Cosh M C, van de Giesen N, Hajdu I, Ikonen J, Jensen K H, Kanniah K D, de Kat I, Kirchengast G, Kumar Rai P, Kyrouac J, Larson K, Liu S X, Loew A, Moghaddam M, Martínez Fernández J, Mattar Bader C, Morbidelli R, Musial J P, Osenga E, Palecki M A, Pellarin T, Petropoulos G P, Pfeil I, Powers J, Robock A, Rüdiger C, Rummel U, Strobel M, Su Z B, Sullivan R, Tagesson T, Varlagin A, Vreugdenhil M, Walker J, Wen J, Wenger F, Wigneron J P, Woods M, Yang K, Zeng Y J, Zhang X, Zreda M, Dietrich S, Gruber A, van Oevelen P, Wagner W, Scipal K, Drusch M and Sabia R. 2021. The International Soil Moisture Network: serving Earth system science for over a decade. *Hydrology and Earth System Sciences*, 25(11): 5749-5804 [DOI: 10.5194/hess-25-5749-2021]
- Gong Y D, Liu G Z, Xue Y Z, Li R and Meng L Z. 2023. A survey on dataset quality in machine learning. *Information and Software Technology*, 162: 107268 [DOI: 10.1016/j.infsof.2023.107268]
- Grujter J J, Bierkens M F P, Brus D J and Knotters M. 2006. *Sampling for Natural Resource Monitoring*. Berlin, Heidelberg: Springer [DOI: 10.1007/3-540-33161-1]
- Guidoum A. 2025. Statistical modeling and mapping of rainfall in the endorheic basins of Northern Algeria: a comparison of spatial interpolation methods. *Acta Geophysica*, 73(2): 1679-1699 [DOI: 10.1007/s11600-024-01392-6]
- Guo J W and Liu F. 2015. Automatic data quality control of observations in wireless sensor network. *IEEE Geoscience and Remote Sensing Letters*, 12(4): 716-720 [DOI: 10.1109/LGRS.2014.2359685]
- Han H T, Wang Y P and Zhang J. 2021. Research progress on the effect of station relocation on climate data homogeneity in North-

- west China. *Plateau Meteorology*, 40(2): 448-454 (韩海涛, 王雅萍, 张瑾. 2021. 中国西北地区台站迁移对气候资料均一性影响的研究进展. *高原气象*, 40(2): 448-454) [DOI: 10.7522/j.issn.1000-0534.2020.00032]
- Han Y, Wen J G, Xiao Q, You D Q, Meng L, Wu S B, Hao D L, Tang Y, Chen X, Liu Q H and Zhao C C. 2024. Asymmetry in the diurnal variation of land surface albedo and its impacts on daily mean albedo estimation. *Journal of Geophysical Research: Atmospheres*, 129(14): e2023JD039728 [DOI: 10.1029/2023JD039728]
- He H L, Zhang L Y, Gao Y Z, Ren X L, Zhang L, Yu G R and Wang S Q. 2015. Regional representativeness assessment and improvement of eddy flux observations in China. *Science of the Total Environment*, 502: 688-698 [DOI: 10.1016/j.scitotenv.2014.09.073]
- Jin R, Li X, Ma M G, Ge Y, Liu S M, Xiao Q, Wen J G, Zhao K, Xin X P, Ran Y H, Liu Q H and Zhang R H. 2017. Key methods and experiment verification for the validation of quantitative remote sensing products. *Advances in Earth Science*, 32(6): 630-642 (晋锐, 李新, 马明国, 葛咏, 刘绍民, 肖青, 闻建光, 赵凯, 辛晓平, 冉有华, 柳钦火, 张仁华. 2017. 陆地定量遥感产品的真实性检验关键技术与试验验证. *地球科学进展*, 32(6): 630-642) [DOI: 10.11867/j.issn.1001-8166.2017.06.0630]
- Kalkan M, Bostancı G E, Güzel M S, Kalkan B, Özşarı Ş, Soysal Ö and Köse G. 2022. Cloudy/clear weather classification using deep learning techniques with cloud images. *Computers and Electrical Engineering*, 102: 108271 [DOI: 10.1016/j.compeleceng.2022.108271]
- Kazantzidis A, Tzoumanikas P, Bais A F, Fotopoulos S and Economou G. 2012. Cloud detection and classification with the use of whole-sky ground-based images. *Atmospheric Research*, 113: 80-88 [DOI: 10.1016/j.atmosres.2012.05.005]
- Kim H J, Park S M, Choi B J, Moon S H and Kim Y H. 2020. Spatio-temporal approaches for quality control and error correction of atmospheric data through machine learning. *Computational Intelligence and Neuroscience*, 2020: 7980434 [DOI: 10.1155/2020/7980434]
- Li M H and Li J Z. 2016. Data currency determination: key theories and technologies. *Intelligent Computer and Applications*, 6(6): 72-75 (李默涵, 李建中. 2016. 数据时效性判定: 关键理论和技术. *智能计算机与应用*, 6(6): 72-75) [DOI: 10.3969/j.issn.2095-2163.2016.06.020]
- Li Q Q, Wen J G, Wu X D, Xiao Q, You D Q and Tang R Q. 2025. The optimal deployment of ground samples: whether spatial heterogeneity is dominated by randomness or structure factors? *IEEE Geoscience and Remote Sensing Letters*, 22: 2500505 [DOI: 10.1109/LGRS.2024.3514356]
- Li Q Y, Lu W T and Yang J. 2011. A hybrid thresholding algorithm for cloud detection on ground-based color images. *Journal of Atmospheric and Oceanic Technology*, 28(10): 1286-1296 [DOI: 10.1175/JTECH-D-11-00009.1]
- Li X. 2014. Characterization, controlling, and reduction of uncertainties in the modeling and observation of land-surface systems. *Science China Earth Sciences*, 57(1): 80-87 (李新. 2013. 陆地表面系统模拟和观测的不确定性及其控制. *中国科学: 地球科学*, 43(11): 1735-1742) [DOI: 10.1007/s11430-013-4728-9]
- Li X, Liu S M, Liu Q H, Xiao Q, Che T, Ma M G, Jin R, Ran Y H, Wen J G, Xu Z W and Li Z Y. 2023. Heihe remote sensing experiments: retrospect and prospect. *National Remote Sensing Bulletin*, 27(2): 224-248 (李新, 刘绍民, 柳钦火, 肖青, 车涛, 马明国, 晋锐, 冉有华, 闻建光, 徐自为, 李增元. 2023. 黑河遥感试验: 回顾与展望. *遥感学报*, 27(2): 224-248) [DOI: 10.11834/jrs.20235013]
- Lin X W. 2018. *Land Surface Albedo Remote Sensing Products Validation over Rugged Terrain*. Beijing: University of Chinese Academy of Sciences (Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences) (林兴稳. 2018. 山区地表反照率遥感产品真实性检验方法研究. 北京: 中国科学院大学(中国科学院遥感与数字地球研究所))
- Liu F and Guo J W. 2013. Study on quality control approach for Heihe wireless sensor network observation data. *Remote Sensing Technology and Application*, 28(2): 252-257 (刘丰, 郭建文. 2013. 面向黑河无线传感网络观测数据的质量控制方法研究. *遥感技术与应用*, 28(2): 252-257) [DOI: 10.11873/j.issn.1004-0323.2013.2.252]
- Liu Q H, Wen J G, Zhou X, Zhao J, Li Z Y, Li X, Ma M G, Wang W Z, Liao X H, Liu S M, Fan W J, Xiao Q, Zhong B, Li J, Xin X Z, Li L, Jia L, Gao Z H, Jin J D, Liang S, Xin J, Liao C J and Wu Y R. 2023. Technique system of remote sensing product generation and validation of GF common products. *National Remote Sensing Bulletin*, 27(3): 544-562 (柳钦火, 闻建光, 周翔, 赵坚, 李增元, 李新, 马明国, 王维真, 廖小罕, 刘绍民, 范闻捷, 肖青, 仲波, 李静, 辛晓洲, 李丽, 贾立, 高志海, 金家栋, 梁师, 邢进, 廖楚江, 吴一戎. 2023. 高分遥感共性产品生成和真实性检验技术体系. *遥感学报*, 27(3): 544-562) [DOI: 10.11834/jrs.20232531]
- Liu X N and Ren Z H. 2005. Progress in quality control of surface meteorological data. *Meteorological Science and Technology*, 33(3): 199-203 (刘小宁, 任芝花. 2005. 地面气象资料质量控制方法研究概述. *气象科技*, 33(3): 199-203) [DOI: 10.19517/j.1671-6345.2005.03.002]
- Lu B B, Ge Y, Qin K and Dong G P. 2025. Theoretical and technical frameworks for geographically weighed models. *National Remote Sensing Bulletin*, 29(3): 596-609 (卢宾宾, 葛咏, 秦昆, 董冠鹏. 2025. 地理加权建模理论与技术框架. *遥感学报*, 29(3): 596-609) [DOI: 10.11834/jrs.20244064]
- Lu G Y and Wong D W. 2008. An adaptive inverse-distance weighting spatial interpolation technique. *Computers and Geosciences*, 34(9): 1044-1055 [DOI: 10.1016/j.cageo.2007.07.010]
- Ma J. 2021. *Retrieval and Validation of Long-Term Land Surface Temperature from Satellite Remote Sensing*. Chengdu: University of Electronic Science and Technology of China (马晋. 2021. 长时序地表温度遥感反演与真实性检验方法研究. 成都: 电子科技大学) [DOI: 10.27005/d.cnki.gdzku.2021.005062]
- Ma M G. 2010. Quality control and evaluation for the ground observation data of the watershed allied telemetry experimental research. *Remote Sensing Technology and Application*, 25(6): 766-771 (马明国. 2010. 黑河综合遥感联合试验地面观测数据质量控制与评价. *遥感技术与应用*, 25(6): 766-771) [DOI: 10.11873/j.issn.

- 1004-0323.2010.6.766]
- Miatselskaya N S, Bril A I, Chaikovskiy A P, Yukhymchuk Y Y, Milinevski G P and Simon A A. 2022. Optimal interpolation of aeronet radiometric network observations for the evaluation of the aerosol optical thickness distribution in the Eastern European region. *Journal of Applied Spectroscopy*, 89(2): 296-302 [DOI: 10.1007/s10812-022-01357-x]
- Nappo C J, Caneill J Y, Furman R W, Gifford F A, Kaimal J C, Kramer M L, Lockhart T J, Pendergast M M, Pielke R A, Randerson D, Shreffler J H and Wyngaard J C. 1982. The workshop on the representativeness of meteorological observations, June 1981, Boulder, Colo. *Bulletin of the American Meteorological Society*, 63(7): 761-764
- Pandžić K, Kobold M, Oskoruš D, Biondić B, Biondić R, Bonacci O, Likso T and Curić O. 2020. Standard normal homogeneity test as a tool to detect change points in climate-related river discharge variation: case study of the Kupa River Basin. *Hydrological Sciences Journal*, 65(2): 227-241 [DOI: 10.1080/02626667.2019.1686507]
- Papachristopoulou K, Fountoulakis I, Bais A F, Psiloglou B E, Papadimitriou N, Raptis I P, Kazantzidis A, Kontoes C, Hatzaki M and Kazadzis S. 2024. Effects of clouds and aerosols on downwelling surface solar irradiance nowcasting and short-term forecasting. *Atmospheric Measurement Techniques*, 17(7): 1851-1877 [DOI: 10.5194/amt-17-1851-2024]
- Parsons M A, Brodzik M J and Rutter N J. 2004. Data management for the Cold Land Processes Experiment: improving hydrological science. *Hydrological Processes*, 18(18): 3637-3653 [DOI: 10.1002/hyp.5801]
- Pastorello G, Agarwal D, Papale D, Samak T, Trotta C, Ribeca A, Poindexter C, Faybishenko B, Gunter D, Hollowgrass R and Canfora E. 2014. Observational data patterns for time series data quality assessment//2014 IEEE 10th International Conference on e-Science. Sao Paulo: IEEE: 271-278 [DOI: 10.1109/eScience.2014.45]
- Pastorello G, Gunter D, Chu H S, Christianson D, Trotta C, Canfora E, Faybishenko B, Cheah Y W, Beekwilder N, Chan S, Dengel S, Keenan T, O'Brien F, Elbashandy A, Poindexter C, Humphrey M, Papale D and Agarwal D. 2017. Hunting data rogues at scale: data quality control for observational data in research infrastructures//2017 IEEE 13th International Conference on e-Science (e-Science). Auckland: IEEE: 446-447 [DOI: 10.1109/eScience.2017.64]
- Pastorello G, Trotta C, Canfora E, Chu H S, Christianson D, Cheah Y W, Poindexter C, Chen J Q, Elbashandy A, Humphrey M, Isaac P, Polidori D, Reichstein M, Ribeca A, van Ingen C, Vuichard N, Zhang L M, Amiro B, Ammann C, Arain M A, Ardö J, Arkebauer T, Arndt S K, Arriga N, Aubinet M, Aurela M, Baldocchi D, Barr A, Beamesderfer E, Marchesini L B, Bergeron O, Beringer J, Bernhofer C, Berveiller D, Billesbach D, Black T A, Blanken P D, Bohrer G, Boike J, Bolstad P V, Bonal D, Bonnefond J M, Bowling D R, Bracho R, Brodeur J, Brümmer C, Buchmann N, Burban B, Burns S P, Buysse P, Cale P, Cavagna M, Cellier P, Chen S P, Chini I, Christensen T R, Cleverly J, Collalti A, Consalvo C, Cook B D, Cook D, Coursolle C, Cremonese E, Curtis P S, D'Andrea E, da Rocha H, Dai X Q, Davis K J, De Cinti B, de Grandcourt A, De Ligne A, De Oliveira R C, Delpierre N, Desai A R, Di Bella C M, di Tommasi P, Dolman H, Domingo F, Dong G, Dore S, Duce P, Dufrêne E, Dunn A, Dušek J, Eamus D, Eichelmann U, ElKhidir H A M, Eugster W, Ewenz C M, Ewers B, Famulari D, Fares S, Feigenwinter I, Feitz A, Fensholt R, Filippa G, Fischer M, Frank J, Galvagno M, Gharun M, Gianelle D, Gielen B, Gioli B, Gitelson A, Goded I, Goeckede M, Goldstein A H, Gough C M, Goulden M L, Graf A, Griebel A, Gruening C, Grünwald T, Hammerle A, Han S J, Han X G, Hansen B U, Hanson C, Hatakka J, He Y T, Hehn M, Heinesch B, Hinko-Najera N, Hörtnagl L, Hutley L, Ibrom A, Ikawa H, Jackowicz-Korczynski M, Janouš D, Jans W, Jassal R, Jiang S C, Kato T, Khomik M, Klatt J, Knohl A, Knox S, Kobayashi H, Koerber G, Kolle O, Kosugi Y, Kotani A, Kowalski A, Kruijt B, Kurbatova J, Kutsch W L, Kwon H, Launiainen S, Laurila T, Law B, Leuning R, Li Y N, Liddell M, Limousin J M, Lion M, Liska A J, Lohila A, López-Ballesteros A, López-Blanco E, Loubet B, Loustau D, Lucas-Moffat A, Lüers J, Ma S Y, Macfarlane C, Magliulo V, Maier R, Mammarella I, Manca G, Marcolla B, Margolis H A, Marras S, Massman W, Mastepanov M, Matamala R, Matthes J H, Mazzenga F, McCaughey H, McHugh I, McMillan A M S, Merbold L, Meyer W, Meyers T, Miller S D, Minerbi S, Moderow U, Monson R K, Montagnani L, Moore C E, Moors E, Moreaux V, Moureaux C, Munger J W, Nakai T, Neiryneck J, Nesic Z, Nicolini G, Noormets A, Northwood M, Noretto M, Nouvellon Y, Novick K, Oechel W, Olesen J E, Ourcival J M, Papuga S A, Parmentier F J, Paul-Limoges E, Pavelka M, Peichl M, Pendall E, Phillips R P, Pilegaard K, Pirk N, Posse G, Powell T, Prasse H, Prober S M, Rambal S, Rannik Ü, Raz-Yaseef N, Rebmann C, Reed D, de Dios V R, Restrepo-Coupe N, Reverter B R, Roland M, Sabbatini S, Sachs T, Saleska S R, Sánchez-Cañete E P, Sanchez-Mejia Z M, Schmid H P, Schmidt M, Schneider K, Schrader F, Schroder I, Scott R L, Sedláč P, Serrano-Ortiz P, Shao C L, Shi P L, Shironya I, Siebicke L, Šigut L, Silberstein R, Sirca C, Spano D, Steinbrecher R, Stevens R M, Sturtevant C, Suyker A, Tagesson T, Takanashi S, Tang Y H, Tapper N, Thom J, Tomassucci M, Tuovinen J P, Urbanski S, Valentini R, van der Molen M, van Gorsel E, van Huissteden K, Varlagin A, Verfaillie J, Vesala T, Vincke C, Vitale D, Vygodskaya N, Walker J P, Walter-Shea E, Wang H M, Weber R, Westermann S, Wille C, Wofsy S, Wohlfahrt G, Wolf S, Woodgate W, Li Y L, Zampieri R, Zhang J H, Zhou G Y, Zona D, Agarwal D, Biraud S, Torn M and Papale D. 2020. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, 7(1): 225 [DOI: 10.1038/s41597-020-0534-3]
- Peterson T C and Vose R S. 1997. An overview of the Global Historical Climatology Network temperature database. *Bulletin of the American Meteorological Society*, 78(12): 2837-2850 [DOI: 10.1175/1520-0477(1997)078<2837:aootgh>2.0.co;2]
- Ren Z H, Xu S, Sun H N and Zhang Q. 2006. Quality check and analysis on the historic global surface synoptic observations. *Journal of*

- Applied Meteorological Science, 17(4): 412-420 (任芝花, 许松, 孙化南, 张强). 2006. 全球地面天气报历史资料质量检查与分析. 应用气象学报, 17(4): 412-420 [DOI: 10.3969/j.issn.1001-7313.2006.04.004]
- Ren Z H, Zhang Z F, Sun C, Liu Y M, Li J, Ju X H, Zhao Y F, Li Z P, Zhang W, Li H K, Zeng X J, Ren X W, Liu Y and Wang H J. 2015. Development of three-step quality control system of real-time observation data from AWS in China. Meteorological Monthly, 41(10): 1268-1277 (任芝花, 张志富, 孙超, 刘一鸣, 李俊, 鞠晓慧, 赵煜飞, 李志鹏, 张玮, 李洪康, 曾行吉, 任晓炜, 刘莹, 王海军). 2015. 全国自动气象站实时观测资料三级质量控制系统的研制. 气象, 41(10): 1268-1277 [DOI: 10.7519/j.issn.1000-0526.2015.10.010]
- Román M O, Schaaf C B, Woodcock C E, Strahler A H, Yang X Y, Braswell R H, Curtis P S, Davis K J, Dragoni D, Goulden M L, Gu L H, Hollinger D Y, Kolb T E, Meyers T P, Munger J W, Privette J L, Richardson A D, Wilson T B and Wofsy S C. 2009. The MODIS (Collection V005) BRDF/albedo product: assessment of spatial representativeness over forested landscapes. Remote Sensing of Environment, 113(11): 2476-2498 [DOI: 10.1016/j.rse.2009.07.009]
- Schaaf C B, Gao F, Strahler A H, Lucht W, Li X W, Tsang T, Strugnell N C, Zhang X Y, Jin Y F, Muller J P, Lewis P, Barnsley M, Hobson P, Disney M, Roberts G, Dunderdale M, Doll C, d'Entremont R P, Hu B X, Liang S L, Privette J L and Roy D. 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sensing of Environment, 83(1/2): 135-148 [DOI: 10.1016/S0034-4257(02)00091-3]
- Schimel D, Pavlick R, Fisher J B, Asner G P, Saatchi S, Townsend P, Miller C, Frankenberg C, Hibbard K and Cox P. 2015. Observing terrestrial ecosystems and the carbon cycle from space. Global Change Biology, 21(5): 1762-1776 [DOI: 10.1111/gcb.12822]
- Sha Y K, Gagne II D J, West G and Stull R. 2021. Deep-learning-based precipitation observation quality control. Journal of Atmospheric and Oceanic Technology, 38(5): 1075-1091 [DOI: 10.1175/JTECH-D-20-0081.1]
- Sharp R A, Mudford O C and Elliffe D. 2015. A data-based method for selecting parameters of momentary time sampling to provide representative data. European Journal of Behavior Analysis, 16(2): 279-294 [DOI: 10.1080/15021149.2015.1083294]
- State Administration for Market Regulation and Standardization Administration of the People's Republic of China. 2018. GB/T 36344-2018 Information technology—Evaluation indicators for data quality. Beijing: Standards Press of China (国家市场监督管理总局, 中国国家标准化管理委员会). 2018. GB/T 36344-2018 信息技术 数据质量评价指标. 北京: 中国标准出版社)
- Steinacker R. 2023. Mean value splines and their use for climatological time series. International Journal of Climatology, 43(9): 4326-4336 [DOI: 10.1002/joc.8089]
- Su L, Gao C C, Ren X L, Zhang F Y, Cao S S, Zhang S Q, Chen T D, Liu M Q, Ni B C and Liu M. 2022. Understanding the spatial representativeness of air quality monitoring network and its application to PM_{2.5} in the mainland China. Geoscience Frontiers, 13(3): 101370 [DOI: 10.1016/j.gsf.2022.101370]
- Toreti A, Bratu M, Müller E, van der Grijn G, Oliveira I, Zampieri M and Ceglar A. 2022. Advancing near-real-time quality controls of meteorological observations. Bulletin of the American Meteorological Society, 103(4): E1078-E1087 [DOI: 10.1175/BAMS-D-21-0171.1]
- Toreti A, Kuglitsch F G, Xoplaki E, Della-Marta P M, Aguilar E, Prohom M and Luterbacher J. 2011. A note on the use of the standard normal homogeneity test to detect inhomogeneities in climatic time series. International Journal of Climatology, 31(4): 630-632 [DOI: 10.1002/joc.2088]
- Wan G N, Yang M X, Liu Z C, Wang X J and Liang X W. 2017. The precipitation variations in the Qinghai-Xizang (Tibetan) Plateau during 1961-2015. Atmosphere, 8(5): 80 [DOI: 10.20944/preprints201701.0128.v1]
- Wang C, Caja J and Gómez E. 2018. Comparison of methods for outlier identification in surface characterization. Measurement, 117: 312-325 [DOI: 10.1016/j.measurement.2017.12.015]
- Wang J F, Stein A, Gao B B and Ge Y. 2012. A review of spatial sampling. Spatial Statistics, 2: 1-14 [DOI: 10.1016/j.spasta.2012.08.001]
- Wang Q. 2021. Progress of environmental remote sensing monitoring technology in China and some related frontier issues. National Remote Sensing Bulletin, 25(1): 25-36 (王桥). 2021. 中国环境遥感监测技术进展及若干前沿问题. 遥感学报, 25(1): 25-36 [DOI: 10.11834/jrs.20210572]
- Wang S K, Zhang Z Y, Wang P G and Tian Y R. 2021. Failure warning of gearbox for wind turbine based on 3 σ -median criterion and NSET. Energy Reports, 7: 1182-1197 [DOI: 10.1016/j.egyr.2021.09.146]
- Wang X L. 2008. Penalized maximal F test for detecting undocumented mean shift without trend change. Journal of Atmospheric and Oceanic Technology, 25(3): 368-384 [DOI: 10.1175/2007JTECH-982.1]
- Wang X L, Wen Q H and Wu Y H. 2007. Penalized maximal t test for detecting undocumented mean change in climate data series. Journal of Applied Meteorology and Climatology, 46(6): 916-931 [DOI: 10.1175/jam2504.1]
- Wang Z S, Schaaf C B, Strahler A H, Chopping M J, Román M O, Shuai Y M, Woodcock C E, Hollinger D Y and Fitzjarrald D R. 2014. Evaluation of MODIS albedo product (MCD43A) over grassland, agriculture and forest surface types during dormant and snow-covered periods. Remote Sensing of Environment, 140: 60-77 [DOI: 10.1016/j.rse.2013.08.025]
- Wen J G, Liu Q H, Li Z Y, Li X, Liu S M, Xiao Q, Gao Z H, Ma M G, Che T, Liu L Y, Fang H L, Yan G J, Ge Y, Chen E X, Zhang Y, Ma L L, Wu X D and Chen X. 2023a. A review of the development of remote sensing field experiments and product validation in China. National Remote Sensing Bulletin, 27(3): 573-583 (闻建光, 柳钦火, 李增元, 李新, 刘绍民, 肖青, 高志海, 马明国, 车涛, 刘良云, 方红亮, 阎广建, 葛咏, 陈尔学, 张勇, 马灵玲, 吴小丹, 陈曦). 2023a. 中国遥感实验与真实性检验的发展思考. 遥感学报,

- 27(3): 573-583) [DOI: 10.11834/jrs.20232673]
- Wen J G, Wu X D, Wang J P, Tang R Q, Ma D J, Zeng Q C, Gong B C and Xiao Q. 2022. Characterizing the effect of spatial heterogeneity and the deployment of sampled plots on the uncertainty of ground “truth” on a coarse grid scale: case study for Near-Infrared (NIR) surface reflectance. *Journal of Geophysical Research: Atmospheres*, 127(11): e2022JD036779 [DOI: 10.1029/2022JD036779]
- Wen J G, Xiao Q, Zhong S Y, Tang Y, Chen X, Wei Q F, Wu X D, Lin X W, Ouyang X Y, You D Q and Liu Q H. 2023b. Technology system for product validation and algorithm test of GF common products and an application example. *National Remote Sensing Bulletin*, 27(3): 780-789 (闻建光, 肖青, 钟守熠, 唐勇, 陈曦, 魏秋方, 吴小丹, 林兴稳, 欧阳晓莹, 游冬琴, 柳钦火. 2023b. 高分遥感共性产品算法测评与真实性检验技术体系及应用实例. 遥感学报, 27(3): 780-789) [DOI: 10.11834/jrs.20221716]
- Wu X D, Wen J G, Xiao Q, Lin X W, You D Q, Yin G F and Liu Q H. 2025. Theory and method of ground observation on satellite pixel scales: taking the numerical land surface parameters as an example. *National Remote Sensing Bulletin*, 29(5): 1227-1244 (吴小丹, 闻建光, 肖青, 林兴稳, 游冬琴, 尹高飞, 柳钦火. 2025. 面向公里级遥感像元的地面观测理论与方法——以数值型陆表参数为例. 遥感学报, 29(5): 1227-1244) [DOI: 10.11834/jrs.20254296]
- Wu X D, Wen J G, Xiao Q, Liu Q, Peng J J, Dou B C, Li X H, You D Q, Tang Y and Liu Q H. 2016. Coarse scale in situ albedo observations over heterogeneous snow-free land surfaces and validation strategy: a case of MODIS albedo products preliminary validation over northern China. *Remote Sensing of Environment*, 184: 25-39 [DOI: 10.1016/j.rse.2016.06.013]
- Wu X D, Wen J G, Xiao Q, You D Q, Gong B C, Wang J P, Ma D J and Lin X W. 2022. Spatial heterogeneity of albedo at subpixel satellite scales and its effect in validation: airborne remote sensing results from HiWATER. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 4407114 [DOI: 10.1109/tgrs.2021.3124026]
- Wu X D, Wen J G, Xiao Q, You D Q, Lin X W, Wu S B and Zhong S Y. 2019. Impacts and contributors of representativeness errors of *in situ* albedo measurements for the validation of remote sensing products. *IEEE Transactions on Geoscience and Remote Sensing*, 57(12): 9740-9755 [DOI: 10.1109/tgrs.2019.2928954]
- Wu X D, Wen J G, Xiao Q, You D Q, Wang J P, Ma D J and Lin X W. 2021. A multiscale nested sampling method for representative albedo observations at various pixel scales. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 8193-8207 [DOI: 10.1109/jstars.2021.3105562]
- Wu X D, Xiao Q, Wen J G and You D Q. 2019. Advances and challenges in the validation of remote sensing albedo products. *Journal of Remote Sensing (in Chinese)*, 23(1): 11-23 (吴小丹, 肖青, 闻建光, 游冬琴. 2019. 异质地表反照率遥感产品真实性检验研究现状及挑战. 遥感学报, 23(1): 11-23) [DOI: 10.11834/jrs.20198057]
- Xiao J J, Li Z Q, Guo F F, Yao Y P, Ma H, Wang K and Wen Q P. 2018. Construction and analysis of annual precipitation series from 1901 to 2017 in Zhejiang province. *Climate Change Research*, 14(6): 553-561 (肖晶晶, 李正泉, 郭芬芬, 姚益平, 马浩, 王阔, 温泉沛. 2018. 浙江省1901—2017年降水序列构建及变化特征分析. 气候变化研究进展, 14(6): 553-561) [DOI: 10.12006/j.issn.1673-1719.2018.019]
- Xu B D. 2018. Leaf Area Index Retrieval and Product Validation over Heterogeneous Land Surfaces. Beijing: University of Chinese Academy of Sciences (Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences) (徐保东. 2018. 非均质地表叶面积指数反演及产品真实性检验. 北京: 中国科学院大学(中国科学院遥感与数字地球研究所))
- Xu B D, Li J, Liu Q H, Huete A R, Yu Q, Zeng Y L, Yin G F, Zhao J and Yang L. 2016. Evaluating spatial representativeness of station observations for remotely sensed leaf area index products. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(7): 3267-3282 [DOI: 10.1109/JSTARS.2016.2560878]
- Xu B D, Li J, Liu Q H, Xin X Z, Zeng Y L and Yin G F. 2015. Review of methods for evaluating representativeness of ground station observations. *Journal of Remote Sensing (in Chinese)*, 19(5): 703-718 (徐保东, 李静, 柳钦火, 辛晓洲, 曾也鲁, 尹高飞. 2015. 地面站点观测数据代表性评价方法研究进展. 遥感学报, 19(5): 703-718) [DOI: 10.11834/jrs.20154178]
- Xu C, Wang W, Hu Y J and Liu Y W. 2024. Evaluation of ERA5, ERA5-Land, GLDAS-2.1, and GLEAM potential evapotranspiration data over mainland China. *Journal of Hydrology: Regional Studies*, 51: 101651 [DOI: 10.1016/j.ejrh.2023.101651]
- Yang H Y, Zhao X Y and Wang L. 2023. Review of data normalization methods. *Computer Engineering and Applications*, 59(3): 13-22 (杨寒雨, 赵晓永, 王磊. 2023. 数据归一化方法综述. 计算机工程与应用, 59(3): 13-22) [DOI: 10.3778/j.issn.1002-8331.2207-0179]
- You D Q, Wen J G, Tang Y, Liu Q, Zhong S Y, Han Y, Gong B C, Zhong B, Wu S L and Liu Q H. 2023. The GF land surface albedo product based on the high-spatial-and-temporal-resolution BRDF priori-knowledge and its preliminary validation. *National Remote Sensing Bulletin*, 27(3): 738-747 (游冬琴, 闻建光, 唐勇, 刘强, 钟守熠, 韩源, 宫宝昌, 仲波, 吴善龙, 柳钦火. 2023. 基于高时空分辨率BRDF先验知识的高分卫星地表反照率产品及其初步验证. 遥感学报, 27(3): 738-747) [DOI: 10.11834/jrs.20231717]
- Zeng Y W and Gong J Y. 2004. Implementing technique of spatial data quality control and evaluation. *Geomatics and Information Science of Wuhan University*, 29(8): 686-690 (曾衍伟, 龚健雅. 2004. 空间数据质量控制与评价方法及实现技术. 武汉大学学报(信息科学版), 29(8): 686-690) [DOI: 10.3969/j.issn.1671-8860.2004.08.007]
- Zhang R H, Tian J, Li Z L, Su H B, Chen S H and Tang X Z. 2010. Principles and methods for the validation of quantitative remote sensing products. *Science China Earth Sciences*, 53(5): 741-751 (张仁华, 田静, 李召良, 苏红波, 陈少辉. 2010. 定量遥感产品真实性检验的基础与方法. 中国科学: 地球科学, 40(2): 211-222) [DOI: 10.1007/s11430-010-0021-3]
- Zhang Y, Jia Z Z, Liu S M, Xu Z W, Xu T R, Yao Y J, Ma Y F, Song L

S, Li X, Hu X, Wang Z Y, Guo Z X and Zhou J. 2020. Advances in validation of remotely sensed land surface evapotranspiration. *Journal of Remote Sensing* (in Chinese), 24(8): 975-999 (张圆, 贾贞贞, 刘绍民, 徐自为, 徐同仁, 姚云军, 马燕飞, 宋立生, 李相, 胡骁, 王泽宇, 郭枝虾, 周纪. 2020. 遥感估算地表蒸散发真实性检验研究进展. *遥感学报*, 24(8): 975-999) [DOI: 10.11834/jrs.

20209099]

Zhao H. 2015. *Research on Quality Control Methods for Ground Observation Data and Its Application in Numerical Models*. Nanjing: Nanjing University of Information Science and Technology (赵虹. 2015. 地面观测资料的质量控制方法研究及其在数值模式中的应用. 南京: 南京信息工程大学)

Quality control methods and framework for *in situ* observation data in remote sensing validation

LI Ququ^{1,2}, WEN Jianguang^{1,2}, XIAO Qing^{1,2}, WU Xiaodan³, YOU Dongqin¹, TANG Yong¹, LIAN Ting¹, PIAO Sen¹, ZHAO Na¹

1. National Engineering Research Center for Satellite Remote Sensing Applications/State Key Laboratory of Remote Sensing and Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China

Abstract: *In situ* observation data serve as the primary source of pixel-scale reference truth for the validation of remote sensing products and play a crucial role in remote sensing modeling, parameter inversion, and products quality assessment. Reliable validation depends on the availability of *in situ* observations that are not only accurate but also representative of the spatial and temporal characteristics of coarse scale remote sensing pixels. Over recent decades, numerous *in situ* observation networks and datasets have been established worldwide. Despite these advances, the quality of *in situ* observation data remains a major limiting factor in remote sensing validation. Various sources of uncertainty introduced during data acquisition, transmission, and processing may degrade data reliability and propagate errors into truth value and products validation results. Therefore, Quality Control (QC) is essential to ensure the accuracy, reliability, and applicability of *in situ* observation data. The objective of this study is to systematically review the major sources of error affecting *in situ* observation data and to summarize existing quality control and evaluation methods for *in situ* observation data in remote sensing validation.

This study systematically reviews the major sources of error and existing problems in *in situ* observation data for remote sensing validation. On this basis, existing quality control and evaluation methods are reorganized and summarized from six key dimensions: data standardization, completeness, timeliness, accuracy, consistency, and spatiotemporal representativeness, forming a structured quality control framework oriented toward remote sensing validation. To demonstrate the practical application of the reviewed methods, land surface albedo is selected as a representative variable, and typical validation studies are examined to illustrate how quality control and evaluation procedures are applied to mitigate instrumental, measurement, and representative errors.

The study indicates that quality control of *in situ* observation data for remote sensing validation currently faces two major challenges. First, the multisource and heterogeneous characteristics of *in situ* observation datasets, including differences in instruments, observation protocols, data formats, and sampling strategies, complicate unified processing and automated quality assessment. Second, a unified and systematic quality control framework specifically designed for remote sensing validation is still lacking, resulting in fragmented methods and inconsistent evaluation criteria across different datasets and applications. In addition, representative errors caused by scale mismatches between *in situ* observations and satellite pixels remain a major source of uncertainty. By synthesizing existing studies, this paper provides a comprehensive overview of quality control and evaluation methods for *in situ* observation data, clarifying their roles in mitigating instrumental errors, measurement errors, and representative errors, as well as their applicability and limitations in pixel-scale validation.

For remote sensing validation, quality evaluation should place particular emphasis on the spatiotemporal representativeness of *in situ* observations at the pixel scale, taking into account *in situ* sites distribution, surface heterogeneity, variable characteristics, and temporal coverage. By systematically summarizing quality control methods and framework for *in situ* observation data, this review provides a reference for standardized data processing and quality evaluation, thereby supporting more reliable and scientifically robust pixel-scale validation of remote sensing products.

Key words: *in-situ* observation data, remote sensing products, validation, error, quality control, land surface albedo

Supported by National Key Research and Development Program of China (No. 2023YFB3905805)