Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Mapping crop residue cover in global black soil regions based on a new multi-band spectral index and an integrated thresholding approach

Yifeng Cui
Institute of Geographic Sciences and Natural Resources Research
Nanshan You
younS@igsnrr.ac.cn

Institute of Geographic Sciences and Natural Resources Research
Ronggao Liu
Institute of Geographic Sciences and Natural Resources Research
Zhichao Li
Institute of Geographic Sciences and Natural Resources Research
Chao Zhang
National University of Singapore
Mohammad Reza Alizadeh
Massachusetts Institute of Technology
Na Chen
Massachusetts Institute of Technology
César Terrer
Massachusetts Institute of Technology
Jinwei Dong
dongjw@igsnrr.ac.cn

Institute of Geographic Sciences and Natural Resources Research

Method Article

Keywords:

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11627777/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Mapping crop residue cover in global black soil regions based on a new multi-band spectral index and an integrated thresholding approach

Yifeng Cui[1,2,5†], Nanshan You[1†], Jinwei Dong[1,5*], Zhichao Li[1,5], Chao Zhang[3], Mohammad Reza Alizadeh[2], Na Chen[4], César Terrer[2], Ronggao Liu[1*]

[1] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China.

[2] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

[3] Department of Civil and Environmental Engineering, National University of Singapore, Singapore, 117576, Singapore.

[4] Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

[5] University of Chinese Academy of Sciences, Beijing, 100049, China.

† These authors contributed equally to this work.

* Corresponding authors: dongjw@igsnrr.ac.cn (J.D.); liurg@igsnrr.ac.cn (R.L.)

Abstract: Crop residue cover is a defining feature of conservation tillage, supporting soil health, carbon sequestration, and agricultural productivity, particularly across the world's black soil regions. However, large-scale monitoring remains challenging due to the spectral heterogeneity of post-harvest fields and the scarcity of ground-truth samples. Here, we propose a new multi-band crop residue cover spectral index (MCRCSI), and an accuracy-weighted composite thresholding (AWCT) for robust crop residue mapping on single-date multispectral imagery. MCRCSI integrates the cellulose- and lignin-related SWIR contrast of crop residues with their joint Red and near-infrared reflectance response, enhancing separability from soil and photosynthetic vegetation. AWCT integrates thresholds derived from Otsu's method, k-means clustering, and Gaussian mixture models using accuracy-based weights, enabling adaptive threshold determination across different cropping systems. Across eight representative sites in the United States, Ukraine, Northeast China, and Argentina, we evaluated the MCRCSI-AWCT framework using 15,591 independent validation samples. It achieved a mean overall accuracy (OA) of 93.6% and a mean F1 score of 94.0%, with site-level OA ranging from 81.0% to 98.0% (SD = 5.6%). MCRCSI was sensitive to changes in crop residue fraction, with distinct responses at low (<20%) and high (>80%) cover levels, indicating potential for fraction retrieval. Its strong separation of unharvested crops from post-harvest crop residue also suggests potential for harvest event detection. Together, these findings provide a scalable framework for mapping crop residue cover and monitoring field management practices over time.

Keywords: crop residue mapping; multi-band crop residue cover spectral index; accuracy-weighted composite thresholding; Sentinel-2; global black soil regions.

1. Introduction

Crop residue (straw) is the remains of crop biomass left on the field after harvest. Annually, global crop residue production reaches several billion tons (Smerald et al., 2023), representing a massive biomass resource. The strategic management of these residues, whether through field cover, removal from land, or in-situ burning, carries profound socio-environmental implications (Zhao et al., 2024). Specifically, crop residue management is central to Conservation Agriculture (CA) (Smith et al., 2019) and Climate-Smart Agriculture (CSA) (Simelton et al., 2021) frameworks, aligning agricultural productivity with global climate and soil health goals.

Crop residue cover represents a vital agricultural mulching practice whose importance is most pronounced in the world's black soil regions (Fontana et al., 2022). As critical global food baskets, these highly productive regions are subject to intensive land-use regimes that have led to severe pedological degradation. This decline manifests as accelerated soil erosion (Kassam et al., 2019) and the depletion of organic matter (Demattê et al., 2025), both of which are primarily attributed to persistent tillage and residue removal. The 1930s Dust Bowl in the U.S. Midwest remains a stark historical warning of how unsustainable management can trigger catastrophic soil loss (Lambert et al., 2020). Conversely, residue cover serves to restore soil organic matter (Meng et al., 2025), enhance nutrient cycling, and mitigate erosion (Feng et al., 2025). Furthermore, it optimizes water retention (Shojaeezadeh et al., 2024), reduces tillage labor, and decreases air pollution by eliminating open burning (Yang et al., 2024). In addition, residue cover alters surface energy fluxes and water exchange processes (Jia et al., 2025), which contributes significantly to local and regional climate regulation. Consequently, accurate and consistent monitoring of residue cover at large scales is essential for assessing agricultural management and understanding its multi-dimensional impacts on soil, climate, and food production.

However, monitoring crop residue covers across regions and years remains challenging due to the heterogeneity of cropland surfaces, seasonal dynamics, and the spectral similarity among residues, soils, and senescent vegetation. To date, various observation platforms have

been employed to detect crop residue cover, including ground and laboratory measurements (Daughtry et al., 2004), near-surface platforms such as unmanned aerial vehicles (UAVs) (Zhou et al., 2020), and satellite-based remote sensing (Hively et al., 2019). Among these, satellite provides the most practical approach due to its broad spatial coverage, temporal continuity, and cost-effectiveness. The Landsat series (Azzari et al., 2019; Dong et al., 2024) and Sentinel-2 missions (Deshpande et al., 2024; Zhou et al., 2025) jointly serve as the core optical data for residue studies by combining long-term continuity with fine spatial resolution (10–30 m).

Despite the rapid development of global land-cover and cropland mapping products, existing global-scale datasets remain insufficient for monitoring crop residue cover and post-harvest agricultural management. Products such as the Copernicus Global Land Service land-cover layers (Buchhorn et al., 2020), ESA WorldCover (Zanaga et al., 2021), and GLAD (Khan et al., 2026) global cropland datasets have provided valuable information on broad land-cover categories, cropland extent, and cropland dynamics at regional to global scales. However, these products are primarily designed to answer where cropland is located, rather than how cropland is managed after harvest (Cui et al., 2024). Within the cropland class, residue-covered fields, bare tilled soils, senescent vegetation, and recently disturbed surfaces are generally not explicitly separated. As a result, these datasets can serve as useful cropland masks or ancillary inputs, but they cannot directly quantify crop residue cover, distinguish residue-covered from non-residue-covered cropland, or support annual assessments of conservation tillage practices. This limitation is particularly critical in global black soil regions, where differences in post-harvest residue management strongly influence soil erosion risk, soil organic matter dynamics, surface energy balance, and land–atmosphere interactions. Therefore, a residue-specific, temporally consistent, and transferable mapping framework is needed to complement existing global land-cover and cropland products by moving from cropland extent mapping toward within-cropland management characterization.

Existing crop residue studies generally follow two methodological paradigms: (1) soft classification: area estimation of crop residue coverage and (2) hard classification of crop

residue cover, which assigns each pixel to a discrete class (e.g., residue-covered vs. non-residue-covered). The first paradigm focuses on quantitative area estimation of crop residue coverage, establishing continuous relationships between satellite-derived features and field measurement (Dong et al., 2023; Gao et al., 2022). Empirical regression models, such as multiple linear or partial least squares regression have been widely used with spectral indices to retrieve crop residue coverage (Daughtry et al., 2006; Hively et al., 2018). To mitigate mixed-pixel effects, spectral mixture analysis (SMA) further decomposes each pixel into fractional abundances of soil, vegetation, and residue components (Yue et al., 2019; Pacheco and Mcnairn, 2010; Feng et al., 2025). Despite their widespread use, these quantitative methods often lack robustness and scalability. While effective at local scales, their performance depends heavily on labor-intensive ground-truth samples and site-specific calibration, restricting transferability across regions with diverse agro-ecological conditions.

The second paradigm emphasizes hard classification of crop residue cover into discrete classes. A widely adopted practice is to apply a fixed empirical threshold, commonly 30 percent of crop residue coverage, to identify residue-covered fields (Dong et al., 2024; Zhang et al., 2025). Recent advances have incorporated machine learning – such as Random Forest (Azzari et al., 2019) and deep learning models including Convolutional Neural Networks (CNN) (Tao et al., 2022; Chen et al., 2024b) and U-Net (Zhou et al., 2025) to improve crop residue classification and large-scale mapping. While these data-driven models effectively capture complex spectral-spatial features, their practical utility is frequently constrained by the substantial demand for high-quality training samples and the significant computational overhead required for model tuning. Furthermore, the inherent black-box nature of these models often results in poor transferability across different sensors and geographical regions.

Spectral index and threshold-based approaches offer an interpretable and computationally efficient framework for agricultural monitoring by integrating multi-band reflectance information without relying on extensive training data. They have been successfully applied to crop type mapping such as rice (Xu et al., 2023) and soybean (Chen et al., 2023) during the growing season. However, crop residue cover emerges as a distinct landscape element

only after harvest, representing a fundamentally different spectral problem. At this stage, the landscape becomes a heterogeneous mix of crop residues, bare soil, and senescent vegetation, all of which are further complicated by fluctuating moisture. Although several tillage-related indices, including the Normalized Difference Tillage Index (NDTI) (Van Deventer et al., 1997) and the Bare Soil Index (BSI) (Diek et al., 2017), have been proposed to characterize soil disturbance, they are not explicitly designed to capture the spectral signal of residue cover itself and therefore exhibit limited robustness under complex post-harvest conditions. This limitation is structural rather than parametric and cannot be fully resolved through threshold tuning alone. Furthermore, converting continuous index values into binary residue maps remains challenging due to large variations in index distributions across regions and years, highlighting the need for a robust composite thresholding framework that integrates complementary decision paradigms rather than relying on a single thresholding method.

To address these limitations, this study develops a Multi-band Crop Residue Cover Spectral Index (MCRCSI) for large-scale residue mapping across global black soil regions using Sentinel-2 imagery. Eight representative test regions in the major black soil regions worldwide including China, the United States, Ukraine, and Argentina were selected for evaluation. Residue, soil, and vegetation samples were obtained from field surveys and expanded through rule-based visual interpretation. Specifically, this study has three objectives: (1) to design a new residue index of MCRCSI based on the spectral separability of crop residue from soil, vegetation, and other land cover types; (2) to develop an Accuracy-Weighted Composite Thresholding (AWCT) approach to separate crop residue cover from other land cover types; and (3) to evaluate the generalization of MCRCSI–AWCT framework across geographic regions and sensor platforms. These objectives correspond to the subsequent analyses of index performance, threshold stability, and cross-region and cross-sensor generalization. In addition, the applicability of MCRCSI to multi-temporal observations was explored to demonstrate its potential for intra-annual monitoring of crop residue dynamics. Overall, this study provides a scalable and transferable framework for large-scale crop residue cover mapping and dynamic monitoring.

2. Study area and Data

2.1 Study area

Regions with extensive black soil coverage are collectively termed black soil regions (BSR) (Tong et al., 2024). In this study, the extent of BSR was masked using soil type of Mollisol defined by the USDA-NRCS Soil Taxonomy. To ensure global representativeness, we selected eight 1° × 1° test grids covering four continental BSRs: North America, Asia, Europe, and South America, with two grids selected from each region (S1–S8 in **Fig. 1**). In all test areas, cropland accounts for over 60% of the total land area, confirming that each grid adequately represents the dominant agricultural landscape within its respective BSRs. Detailed characteristics of the study areas are provided in **Table 1**.

S1 and S2 are situated in the U.S. Midwest Corn Belt, where fertile soils support corn (April–October) and soybean (May–October) rotations (Wang et al., 2020) and conservation agriculture was widely adopted following the 1930s Dust Bowl (Lambert et al., 2020). S3 and S4 in the Ukraine feature humus-rich soils historically contributing ~10% of global grain trade (Kravchenko et al., 2018), though the region now faces degradation risks from intensive use and conflict. S5 and S6 span the Northeast China BSR (Songnen, Sanjiang, and Liaohe Plains), yielding ~25% of national grain and undergoing significant crop structure adjustments since the 21st century (Di et al., 2026). S7 and S8 in the Argentine Pampas have seen >120 years of cultivation under a subtropical–temperate climate, dominated by a soybean–maize double-cropping system growing from October to May.

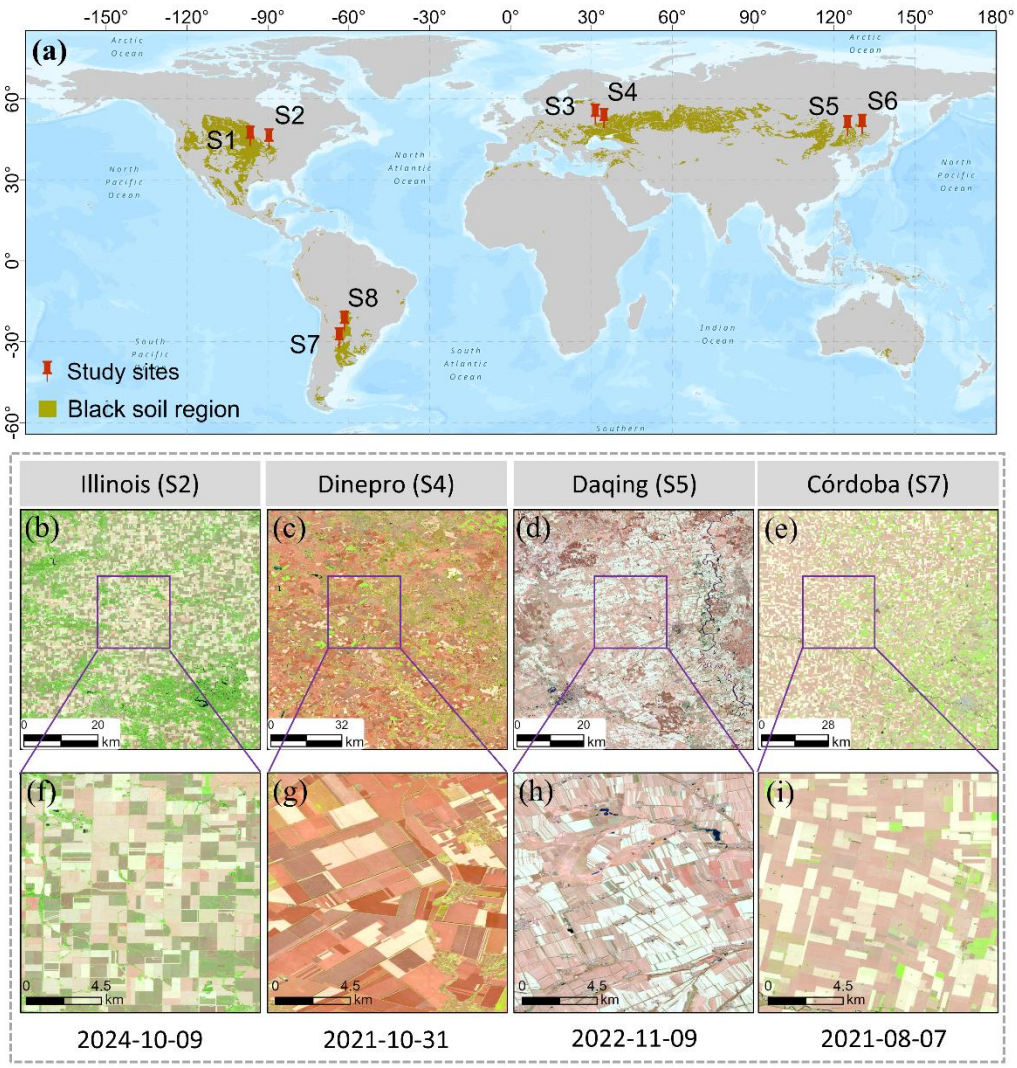


Fig. 1. Geographic distribution and overview of the study sites used for algorithm testing. (a) Location of study area S1–S8 across global black soil region. Panel (b)–(i) display Sentinel-2 false-color composites (SWIR2–NIR–Red) for four representative sites (S2, S4, S5, and S7), with the corresponding acquisition date indicated below each image.

175 **Table 1** Overview of study sites and images used in this study

Site	Location	Crops (pattern)	Sample source	Date	MGRS Tile
S1	Nebraska, United States	Maize, soybean (single)	Visual interpretation	20241013	14TQM, 15TTG, 14TPN, 14TPM, 14TQN
S2	Illinois, United States		Field survey	20241009	16TBL, 15TYF 16TBM, 15TYG 16TCL, 16TCM
S3	Poltava Ukraine	Maize, wheat, sunflower, and rapeseed (single)	Visual interpretation	20211031	36UVV, 36UWV 36UVA, 36UWA
S4	Dinepro, Ukraine				36UXU, 36UXV 36UWV, 36UWU
S5	Daqing, China	Maize, soybean	Field survey	20221109	51TXL, 51TXM
S6	Jiamusi, China	Maize, soybean, and rice (single)		20231024	52TFS, 52TER 52TES, 52TFR
S7	Cordoba, Argentina	Maize, soybean, and cotton (double)	Visual interpretation	20210807	20HMJ, 20HMK 20HLJ, 20HLK
S8	Santiago, Argentina			20210807	20JNR, 20JPR 20JNS, 20JPS

176 *2.2 Remote sensing dataset*

177 We used Sentinel-2 images for algorithm development, and used the Landsat and MODIS
178 images for cross-sensor consistency assessment (**Table S1)**. Sentinel-2 satellites provide 13
179 spectral bands from the visible to shortwave-infrared regions at 10–60 m spatial resolutions
180 with a revisit frequency of ~5 days, greatly enhancing its capability to monitor non-growing-
181 season cropland surface conditions. In this study, the Harmonized Sentinel-2 MSI Level-2A
182 (Surface Reflectance) product was processed in Google Earth Engine (GEE). To ensure
183 consistent spatial management, the official Sentinel-2 Military Grid Reference System
184 (MGRS) (Bauer-Marschallinger and Falkner, 2023) grid shapefile was used to define image
185 tiles, and the acquisition dates and corresponding MGRS codes are summarized in **Table 1**.

For each study site, all Sentinel-2 tiles listed were acquired on the same date and subsequently mosaicked into a single scene-level image.

To examine the generalizability of the algorithm across satellite sensors, we used the Landsat 8/9 Level-2 Collection 2 Tier 1 products and MOD09GA.061 Terra Surface Reflectance product from GEE and harmonized their reflectance with Sentinel-2 MSI using the cross-sensor conversion coefficients from Zhang et al. (2018). Cross-sensor image pairs were selected within a ±1-day acquisition window and resampled to the same spatial resolution. Obvious snow cover and abrupt change pixels were removed for cross-sensor consistency analysis.

2.3 Ground reference data

Ground reference data were obtained from field surveys and visually interpreted samples using Sentinel-2 imagery and crop type maps (Table S2). Field campaigns were conducted in the Northeast China and U.S. Midwest BSRs in 2023 and 2024, where residue-covered plots were directly surveyed. These field samples served as references for visual interpretation of Sentinel-2 composites, allowing reference data expansion to the Ukraine and Argentina BSRs. The combined dataset provided comprehensive ground truth for index development, threshold calibration, and independent accuracy validation. For subsequent analyses, the reference samples were split into calibration set and independent validation set at an approximately 30:70 ratio. The calibration set was used to evaluate the candidate thresholding methods and calibrate the final thresholds in AWCT, whereas the independent validation set was only used in the final accuracy assessment and was not involved in threshold selection or weight estimation.

2.3.1 Samples from windshield surveys

A total of 38 sampling grid cells (0.5° × 0.5°) were selected across the northeast China BSR. Two field campaigns were conducted in northeast China BSR during March to April and October 2023. Survey routes were primarily established along inter-province roads across the Northeast China Plain. For each sampling plot, field personnel aimed to reach the center

of agricultural fields for image acquisition, while supplementary photographs were collected from nearby accessible locations when direct access was restricted. Geographic coordinates were recorded using smartphones and geotagged and timestamped photos were captured from multiple directions **Fig. 2**a–b. These field photographs were then cross-checked with Sentinel-2 imagery acquired on the same or adjacent dates to ensure accurate spatial correspondence and reliable visual interpretation. In the southern BSR of Northeast China (e.g., Liaoning and Jilin provinces), where fields are generally smaller and crop residue cover is spatially fragmented, a subset of samples was found to be near residue-nonresidue boundaries, requiring manual correction during subsequent data processing. All field samples were subsequently used for spectral analysis and validation of the index. Field campaign in the United States BSR followed the same protocol and was conducted in October 2024 across states of Illinois, Iowa, Minnesota, and Wisconsin (Cui et al., 2026).

2.3.2 Sample augmentation based on visual interpretation of Sentinel-2 composites

Direct field sampling in Ukraine and Argentina was not feasible due to logistical and financial constraints. Furthermore, conducting large-scale ground surveys across all BSRs would have required prohibitive resources. Conventional high-resolution sources, such as Google Earth imagery, also fail to capture the brief off-season window when crop residues are present. These limitations necessitated the use of false-color composites (FCC) (Patra et al., 2006; Li et al., 2024), which facilitate the differentiation of agricultural features through distinct spectral and color contrasts. In this study, we employed the SWIR2–NIR–Red band combination to generate composites that enhance the visual characteristics of residue-covered fields. These bands were selected because SWIR bands are highly sensitive to the lignin and cellulose components of crop residues (Lamb et al., 2022; Yue et al., 2020), whereas the Red band responds strongly to vegetation biomass and leaf area index (Heiskanen, 2006). Compared to traditional True-Color Composites (TCC), this unique combination provides a more intuitive visual perception and makes residue features significantly easier to identify (**Fig.S1**).

In the FCC images, residue-covered fields typically exhibited light or pale yellowish tones, which contrasted clearly with the dark reddish-brown hues of black soil (bare soil) and the vibrant green tones of vegetated surfaces (**Fig. 2**e–f). These characteristic chromatic patterns provided the basis for establishing a unified interpretation protocol. Field-confirmed samples collected in both Northeast China and the U.S. Midwest were first used as reference examples to calibrate the visual interpretation criteria and were subsequently used to guide sample augmentation in regions without local field observations.

To minimize subjectivity in visual interpretation, all candidate samples were independently interpreted using the same Sentinel-2 FCC image and reference criteria. Available crop maps were additionally used to cross-check crop type and agricultural context and to reduce the inclusion of non-agricultural or otherwise inappropriate samples. To further standardize the interpretation and screen chromatically anomalous samples, the color characteristics of the retained candidates were evaluated in the CIE chromaticity space (Ohno, 2000). The surface reflectance values derived from three-band FCC images were transformed into the CIE xy chromatic color space using the equations (1) –(5):

$$x = \frac{X}{X + Y + Z} \tag{1}$$

$$y = \frac{Y}{X + Y + Z} \tag{2}$$

$$X = 2.7689R_{swir2} + 1.7517R_{nir} + 1.1302R_{red} \tag{3}$$

$$Y = 1.0000R_{swir2} + 4.5907R_{nir} + 0.0601R_{red} \tag{4}$$

$$Z = 0.0000R_{swir2} + 0.0565R_{nir} + 5.5943R_{red} \tag{5}$$

The CIE (x)–(y) representation emphasizes chromatic information while reducing the influence of overall brightness, allowing samples with similar spectral-color characteristics to form relatively coherent distributions in chromaticity space (Fig. 2k). Field-confirmed residue samples from Northeast China and the U.S. Midwest were used to characterize the

reference CIE (x)–(y) domain of crop residues. Candidate residue samples derived from visual interpretation were projected into the same color space and compared with this reference distribution (**Fig. 2**l). Samples showing strong consistency with the reference residue color domain were retained, whereas chromatically anomalous samples potentially affected by vegetation, bare soil, shadows, mixed surface conditions, or other spectral contamination were excluded. Samples located close to field boundaries or containing evident mixed-pixel signals were also removed. The CIE-based analysis was therefore used as an auxiliary quality-control procedure rather than as an independent classifier.

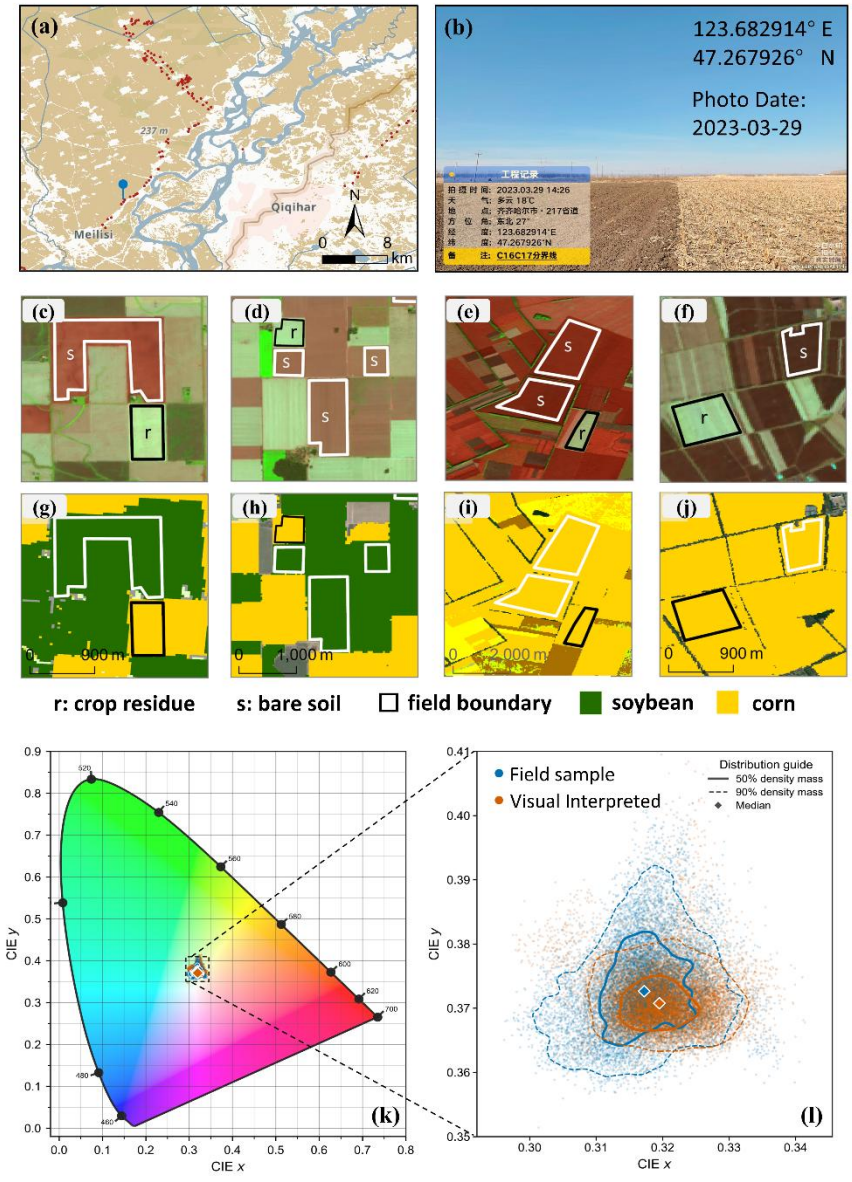


Fig. 2. Collection and cross-regional consistency assessment of crop-residue reference samples. (a) Field photographs and residue sampling conducted in Meilis District, Qiqihar City, Northeast China, on 29 March 2023. Ground observations were collected from

sampling areas located near the centers of fields to represent surface residue-cover conditions of the dominant crop types. (b) Example geotagged field photograph documenting surface crop residue during the field campaign. (c–f) Representative Sentinel-2 images the U.S. Midwest, Argentina, Ukraine, and Northeast China, respectively, used to delineate crop-residue-covered fields (r) and bare-soil fields (s), with field boundaries outlined. (g–j) Corresponding crop type maps for the U.S. Midwest, Argentina, Ukraine, and Northeast China, respectively, with soybean shown in green and maize shown in yellow. (k) Distribution of the reference samples in the full CIE *x*–*y* chromaticity space; the dashed box denotes the enlarged region shown in (l). (l) Zoomed chromaticity distributions of field-sampled reference data from Northeast China (CHN) and the United States Midwest (USA) and visually interpreted samples from Argentina (ARG) and Ukraine (UKR). Solid and dashed contours enclose the 50% and 90% density mass, respectively, and diamonds indicate the median chromaticity coordinates of each sample source.

Using the field-validated samples from Northeast China and the United States as references, the same interpretation protocol was extended to Ukraine and Argentina, resulting in a final dataset of 11,188 residue cover samples and 10,996 non-residue cover samples (**Table S3**). Of these 22,184 reference samples, 6,593 were allocated to the calibration set and 15,591 to the independent validation set.

2.4 Ancillary datasets

2.4.1 Crop type and crop calendar

Four crop type maps were employed as reference datasets to assist in interpreting crop residue composition patterns within each black soil region (BSR) (Table S2). For Northeast China, we used our previously developed and continuously updated crop type map covering 2019–2023 (You et al., 2021). For the United States, the USDA/NASS Cropland Data Layer (Boryan et al., 2011) was adopted. For Ukraine, the 2020 crop type map published by Chen et al. (2024a) was used, and for Argentina, we employed the Argentina National Map of Crops spanning 2018–2024 (De Abelleyra et al., 2024). USDA FAS/IPAD maize and soybean calendars (https://www.fas.usda.gov/regions) provided ancillary guidance for defining regional non-growing seasons (**Fig.S2**). Harvests generally end by October in the Northern Hemisphere BSRs and occur during March–May in Argentina, informing region-specific off-season image selection and index development.

2.4.2 Residue fraction and relative water content data

Two complementary spectral datasets were used to evaluate the effects of the residue fraction and water content on the performance of the proposed spectral index.

To evaluate the sensitivity of MCRCSI to changes in residue fraction, we used residue fraction data from the USGS agricultural field-condition spectral dataset (Hively et al., 2021). In this dataset, residue fraction represents the fractional ground cover of non-photosynthetic vegetation (NPV), which predominantly comprises crop residue and senescent plant material. At each sampling location, a high-resolution digital photograph was acquired concurrently with the field reflectance spectrum and cropped to match the field of view of the spectroradiometer. A systematic grid of either 132 or 156 points was superimposed on each photograph, and each point was visually classified as green vegetation (GV), NPV, or soil. Residue fraction was calculated as the number of points classified as NPV divided by the total number of classified points and was subsequently expressed as a percentage. To minimize interference from green vegetation and isolate the spectral response of residue–soil mixtures, only samples with no green vegetation cover (GV = 0) were retained, resulting in 432 samples. The photograph-derived residue fraction was used as the independent reference variable for evaluating the responses of MCRCSI and the other residue indices to changes in residue fraction.

To quantify the sensitivity of MCRCSI to residue moisture, we used the relative water content (RWC) associated with digitized maize- and soybean-residue spectra reported by Quemada and Daughtry (2016). Five RWC levels were available for maize (1%, 40%, 68%, 81%, and 100%) and soybean (23%, 50%, 70%, 80%, and 100%). Each crop–RWC combination was represented by a single spectrum covering 410–2400 nm at 5-nm intervals. These spectra were convolved with the Sentinel-2 relative spectral response functions to derive band-equivalent reflectance, from which MCRCSI and the other eight residue indices were calculated.

2.4.3 Other residue-related spectral indices used in the study

For comparative evaluation, eight tillage- and vegetation-related spectral indices that can be directly and consistently calculated from Sentinel-2 MSI surface reflectance were selected

331 from previous studies as summarized in **Table 2**. These include the BSI, NDTI, Crop Residue
332 Covered Spectral Index (CRCSI), Dead Fuel Index (DFI), Normalized Difference Vegetation
333 Index (NDVI), Normalized Difference Senescent Vegetation Index (NDSVI), Normalized
334 Difference Residue Index (NDRI) and Broadband spectral Angle Index (BAI). For BAI, we
335 used the red–NIR–SWIR2 combination, denoted BAI (346), following the numbering
336 scheme of Yue et al. (2020), in which the six selected broadbands (blue, green, red, NIR,
337 SWIR1, and SWIR2) are numbered 1–6, respectively. This notation is distinct from the
338 Sentinel-2 band numbering used in all index formulas, as defined in **Table S1**.

339 **Table 2** Spectral indices used in this study for comparison. The spectral bands are defined
340 according to the Sentinel-2 band numbering scheme.

Index	Formula	Source
BSI	((B12+B4)-(B8+B2))/((B12+B4)+(B8+B2))	(Diek et al., 2017)
CRCSI	100×B4×(B11-B12)/(1-mean(B4+B11))	(Zhang et al., 2025)
DFI	100×(1-B12/B11)×(B4/B8A)	(Cao et al., 2010)
NDTI	(B11-B12)/(B11+B12)	(Van Deventer et al., 1997)
NDVI	(B8-B4)/(B8+B4)	(Rouse Jr et al., 1974)
NDSVI	(B11-B4)/(B11+B4)	(Qi et al., 2002)
NDRI	(B4-B12)/(B4+B12)	(Gelder et al., 2009)
BAI (346)	$BAI(346) = 180° - \tan^{-1}\left(\frac{B8 - B4}{0.06728}\right) + \tan^{-1}\left(\frac{B12 - B8}{0.54784}\right)$	(Yue et al., 2020; Wu et al., 2026)

341 **3. Methods**

342 The residue cover mapping workflow (**Fig. 3**) consisted of three stages. (1) Study area and
343 sampling: eight test sites across four global black soil regions (BSRs) were selected to

represent diverse agroecosystems, and reference samples were established through field surveys, visual interpretation, and CIE-based filtering, followed by division into independent calibration and validation sets. (2) Index development: crop maps and crop calendars were combined with Sentinel-2 imagery to define the off-season observation period. Spectral analysis, band selection, formula derivation, sensitivity analysis, and comparison with candidate indices were then used to develop and evaluate MCRCSI. (3) Index-based residue cover mapping: MCRCSI layers were generated from single-scene Sentinel-2 imagery, and AWCT was applied to integrate candidate thresholds derived using K-means, Otsu, and GMM according to their calibration-set F1 scores, yielding an accuracy-weighted threshold for final residue-cover mapping. The resulting map was subsequently assessed using the independent validation set.

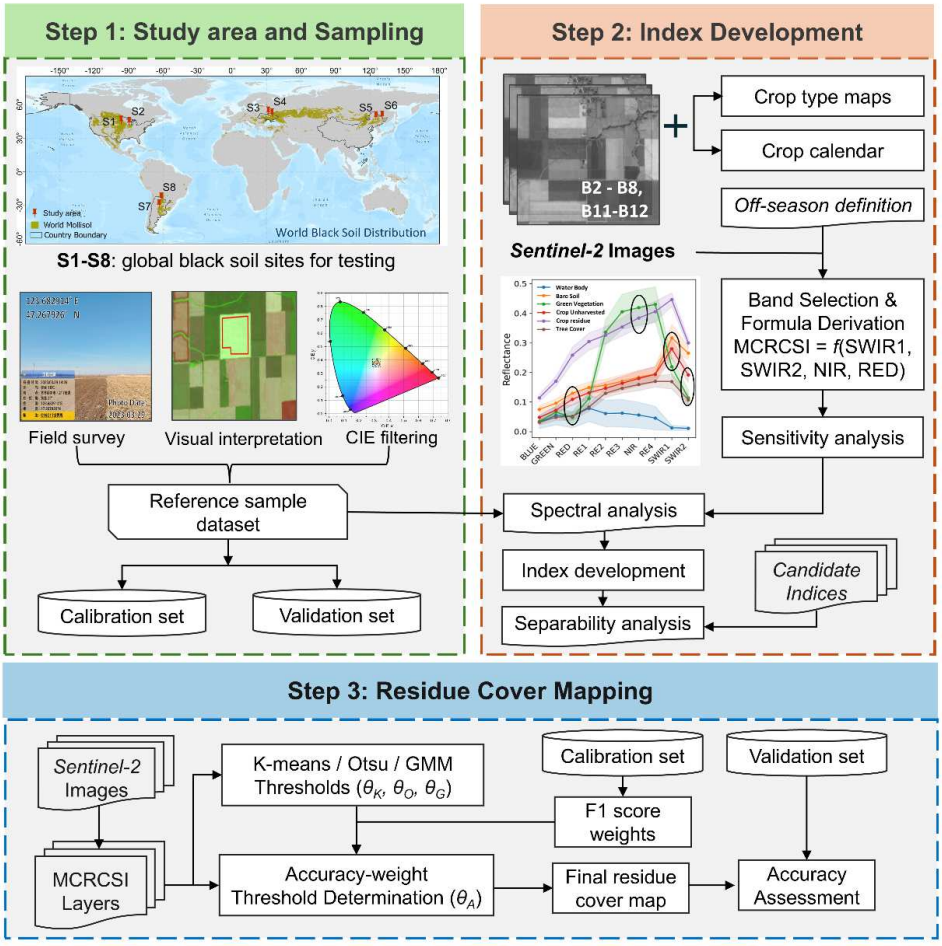


Fig. 3. The overall methodological framework for crop residue cover mapping.

3.1 Spectral analysis

358 3.1.1 Typical land cover types in non-growing seasons

359 Crop residues become exposed as mature crops are harvested, marking the start of the off-
360 season, or non-growing season, which extends until subsequent sowing. Off-season surfaces
361 are typically heterogeneous mixtures of residues: bare soil, unharvested crops, green and
362 woody vegetation, snow, and water may coexist within individual fields (**Fig.S3**). Their
363 proportions and spatial arrangements vary across regions, fields, and observation dates,
364 generating complex spectral signatures that underpin our remote sensing–based classification
365 system. Based on dominant land-use types across the four representative BSRs and their
366 spectral separability, we defined six major off-season cropland surface classes: (1) Residue
367 cover (RC), (2) Bare soil (BS), (3) Unharvested crops (CU), (4) Tree cover (TC), (5) Green
368 vegetation (GV) and (6) Water body (WB). Detailed definitions of all classes are provided in
369 **Table 3**.

370 **Table 3** Description of typical land cover types in croplands during non-growing season

Types	Abbreviation	Description
Residue Cover	RC	Cropland densely covered by one or multiple layers of crop residue, where the underlying background soil is virtually invisible in the vertical view, representing a pure pixel of residue.
Bare Soil	BS	Soil in an exposed state without any vegetation or crop residue cover, which includes unplanted areas or farmland recently disturbed by tillage or plowing.
Crop Unharvested	CU	Cropland currently hosting crops that are at or nearing maturity and are ready for harvest but not yet been subjected to agricultural harvesting activities.
Tree Cover	TC	Shrubs, trees, and other similar vegetation located within the spatial boundary of the cropland,

		possessing a discernible green vegetation canopy at the pixel scale, such as agricultural shelterbelts.
Green Vegetation	GV	Crops that have been planted and are currently in the growth stage, with active photosynthesis and developed green leaf canopy that provides shield over the underlying soil.
Water Body	WB	Water that temporarily obscures cropland surfaces.

3.1.2 Spectral characteristics of crop residue

Based on samples of each land-cover type and single-date Sentinel-2 imagery, we analyzed the spectral reflectance of post-harvest fields landscapes (**Fig. 4**). Water bodies exhibit the lowest reflectance (< 0.05) across all wavelengths, indicating strong absorption. Bare soil shows a gradual reflectance increase with wavelength, peaking near SWIR1 and slightly declining toward SWIR2, consistent with typical soil spectra. Vegetation types (green vegetation and tree canopy) display strong chlorophyll absorption in the red band, a sharp rise in the red-edge bands, and high NIR reflectance; the latter declines more steeply in SWIR due to higher water absorption. Unharvested crops (e.g., mature maize) exhibit intermediate spectral behavior, with reduced NIR reflectance relative to actively growing vegetation, reflecting lower leaf area and moisture content. Residue-covered surfaces (bright residue) show distinct and stable spectral signatures. In the visible bands, their reflectance lies between that of bare soil and vegetation, reflecting partial scattering by cellulose and lignin. In the NIR, residue reflectance (~0.3–0.4) is higher than soil but lower than vegetation, forming a consistent and discriminative peak. In the SWIR bands, residue shows the greatest contrast with bare soil around SWIR1, attributable to low water content and enhanced structural scattering. This confirms SWIR1 as the key diagnostic band for residue identification (**Fig.S4**).

Analysis of the Sentinel-2 red-edge bands (RE1–RE4) confirmed their ability to enhance residue–soil separability. However, since these bands are absent in Landsat and MODIS, their

use is limited for cross-sensor and long-term applications. To ensure broader applicability, this study focuses on the consistent and physically interpretable spectral behavior of the VIS–NIR–SWIR1 bands.

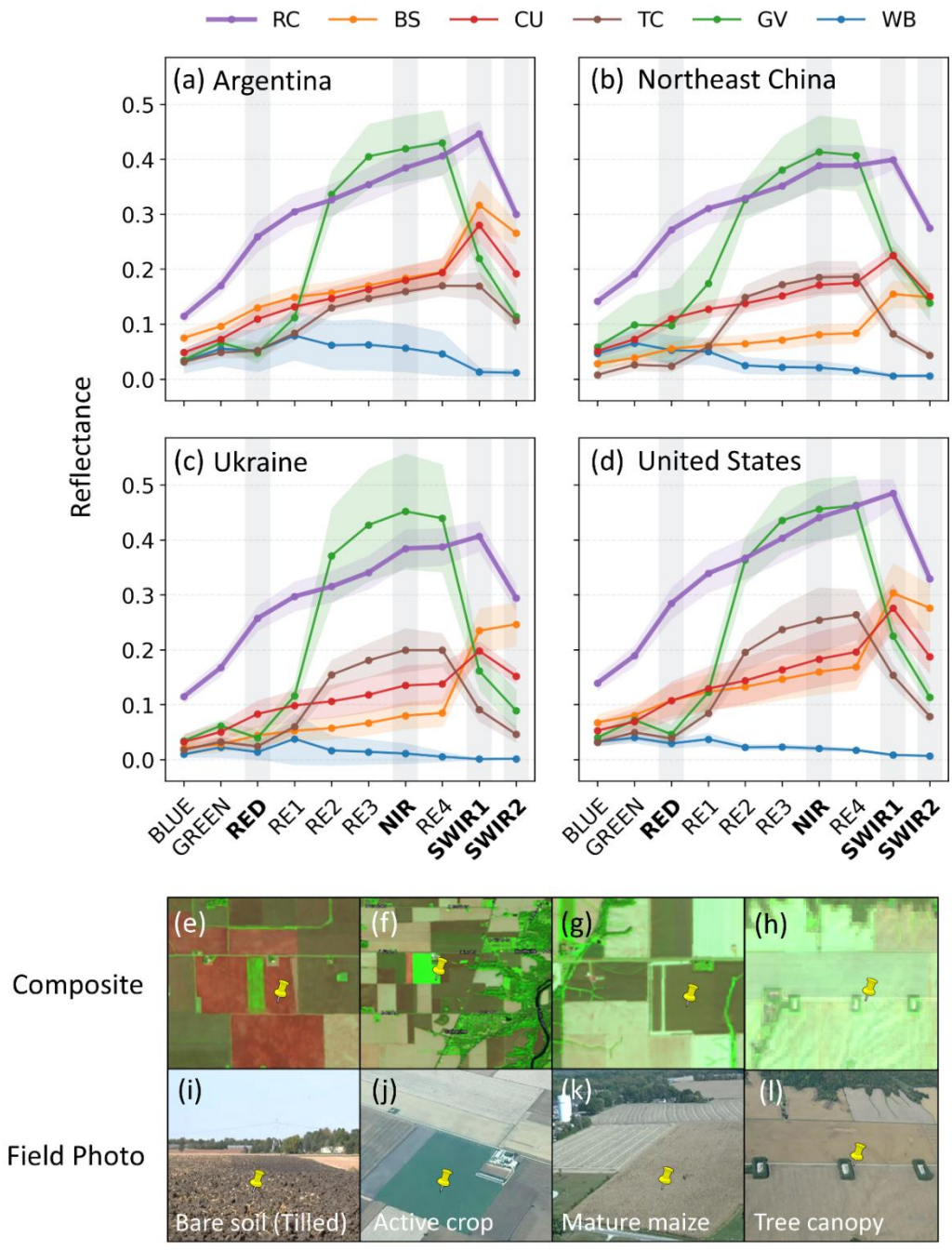


Fig. 4. Spectral reflectance and visual characteristics of land cover types in black soil regions. (a–d) Surface reflectance curves across Sentinel-2 MSI bands (BLUE to SWIR2) for Argentina (7 Aug 2021), Northeast China (9 Nov 2022 and 24 Oct 2023), Ukraine (31 Oct 2021), and the United States (9 and 13 Oct 2024). Land cover classes include residue cover (RC), bare soil (BS), crop unharvested (CU), tree cover (TC), green vegetation (GV), and water bodies (WB). (e–l) Side-by-side comparison of Sentinel-2 composites and

corresponding field/aerial photographs for representative classes: tilled bare soil (e, i), active crop (f, j), mature maize (g, k), and tree canopy (h, l).

We used the M-index (Kaufman and Remer, 1994), defined as the difference in class means divided by the sum of their standard deviations, to assess spectral separability between crop residue and other land-cover types across spectral bands (**Table S4**). For each region, the coefficient of variation (CV) was calculated across the five M-indices comparing RC with BS, CU, GV, TC, and WB, then averaged across the four regions. Differences in MCRCSI among RC, BS, CU, and TC were assessed using Welch's one-way ANOVA, Holm-adjusted pairwise Welch tests, and Hedges' g with 95% confidence intervals.

3.2 Multi-band Crop Residue Cover Spectral Index (MCRCSI)

MCRCSI was developed based on three characteristic spectral properties of crop residues. First, cellulose- and lignin-related absorption in the SWIR domain produces stronger attenuation toward SWIR2, resulting in a pronounced reflectance contrast between SWIR1 and SWIR2 for residue cover (**Fig. 5**a). Second, residue-covered surfaces generally exhibit relatively high reflectance in both the red and NIR bands, allowing the product to emphasize their characteristic broadband brightness while reducing interference from surfaces that are dark in either band (**Fig. 5**b–c). Third, the relatively high SWIR1 reflectance of crop residues provides the basis for the ratio denominator: as SWIR1 reflectance increases, the denominator decreases and the residue-related spectral response is amplified (**Fig. 5**d). Accordingly, MCRCSI integrates the SWIR1–SWIR2 absorption contrast, Red–NIR brightness, and SWIR1-based ratio enhancement to improve the spectral separability of crop residues among other land cover types.

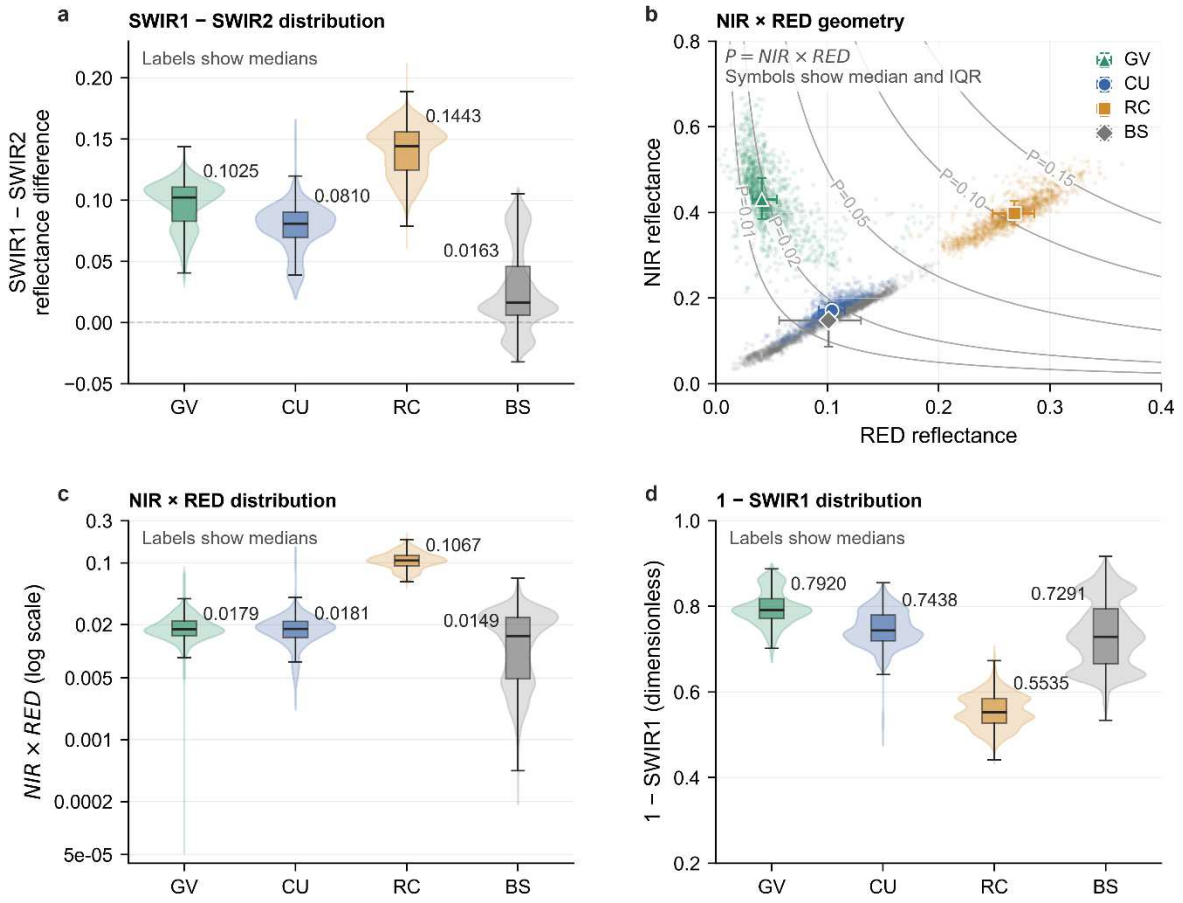


Fig. 5. Spectral components of MCRCSI across four land-cover classes. (a) Distributions of the SWIR contrast term, SWIR1 − SWIR2. (b) Red–NIR reflectance space with contours of equal product values, $P = \mathrm{NIR} \times \mathrm{RED}$. Coloured symbols indicate class medians, and horizontal and vertical error bars represent interquartile ranges (IQRs). (c) Distributions of the NIR × RED product on a logarithmic scale. (d) Distributions of the denominator, 1 − SWIR1. Green, blue, orange and grey represent green vegetation (GV), crop unharvested (CU), residue cover (RC) and bare soil (BS), respectively.

Based on these spectral characteristics, MCRCSI was formulated using reflectance from four spectral bands: Red, NIR, SWIR1, and SWIR2. Specifically, the Red band highlights the contrast between residue and green vegetation in terms of chlorophyll absorption, while the NIR band reflects the intermediate reflectance of residue between vegetation and bare soil. The difference between SWIR1 and SWIR2 captures the characteristic that residue typically exhibits higher reflectance in SWIR1 than in SWIR2, thereby enhancing discrimination from bare soil. Residue demonstrates a sharp SWIR2 absorption driven by cellulose–lignin chemistry, yielding a steep SWIR1–SWIR2 gradient. The numerator combines these complementary spectral features through a multiplicative formulation:

$$N = 1000|\rho_{\mathrm{NIR}} \times \rho_{\mathrm{Red}} \times (\rho_{\mathrm{SWIR1}} - \rho_{\mathrm{SWIR2}})| \quad (6)$$

where ρNIR, ρRed, ρSWIR1 and ρSWIR2 denote the corresponding surface reflectance. The product of the NIR and red reflectance weights the SWIR1–SWIR2 contrast according to their joint reflectance response: the contrast is enhanced when both reflectance is relatively high and suppressed when either is low (**Fig. 5**b-c). The absolute-value operator was applied to express the magnitude of the SWIR1–SWIR2 spectral gradient, thereby restricting MCRCSI to a non-negative range and maintaining a consistent index direction (**Fig.S5**d). The factor of 1000 merely rescales the resulting values to a more readily interpretable numerical range; it neither changes the relative spectral response nor improves numerical stability.

The original denominator, 1-ρSWIR1, increases the contribution of surfaces with high SWIR1 reflectance. However, it may become very small when ρSWIR1 approaches one, potentially causing numerical inflation. Although this condition was not observed for valid crop-residue, soil, vegetation, water or snow samples (**Fig.S5**a-b), it occurred in a small number of residual cloud-contaminated pixels (**Fig.S5**c). We therefore introduced a hierarchical anti-saturation procedure consisting of cloud masking, denominator regularization and output clipping. Pixels identified as cloud or cloud shadow by the sensor quality mask, or with SWIR1 reflectance outside the valid surface-reflectance interval of 0-1, were treated as invalid before index calculation. For the remaining pixels, the stabilized MCRCSI was calculated as

$$\mathrm{MCRCSI}_{\mathrm{safe}} = \begin{cases} \mathrm{NoData} & \mathrm{if} Q_{\mathrm{invalid}} = 1 \\ \min\left[C, \dfrac{1000|\rho_{\mathrm{NIR}} \times \rho_{\mathrm{Red}} \times (\rho_{\mathrm{SWIR1}} - \rho_{\mathrm{SWIR2}})|}{\max(|1-\rho_{\mathrm{SWIR1}}|, \varepsilon)}\right] & \mathrm{if} Q_{\mathrm{invalid}} = 0 \end{cases} \quad (7)$$

Where

$$Q_{\mathrm{invalid}} = Q_{\mathrm{cloud}} \vee Q_{\mathrm{cloud\ shadow}} \vee \left[\rho_{\mathrm{SWIR1}} \notin [0,1]\right] \quad (8)$$

The denominator protection value was set to ε=0.10, which constrains the maximum denominator-related amplification to 1/ε=10. This threshold does not modify the currently

462 observed non-cloud samples because their minimum absolute denominator was 0.4332. In
463 contrast, 57 of the 359 cloud observations had |1-ρSWIR1|≤0.10, confirming that the
464 protection primarily targets residual cloud contamination rather than valid crop-residue
465 surfaces.

466 The upper bound was set to C=100. This value exceeds the maximum MCRCSI observed
467 among all non-cloud land-cover classes (81.27), while preventing residual cloud pixels from
468 producing values several orders of magnitude larger. A saturation-quality flag was
469 additionally defined as

$$Q_{\text{sat}} = \mathbb{I}\left[|1-\rho_{\text{SWIR1}}|\leq\varepsilon \vee \text{MCRCSI}_{\text{protected}}\geq C\right] \quad (9)$$

470 where $\mathbb{I}[\cdot]$ is the indicator function and MCRCSI protected denotes the value obtained after
471 denominator regularization but before final clipping. Pixels for which Qsat=1 were retained
472 as explicitly flagged observations rather than interpreted as high-confidence crop-residue
473 pixels. The values ε=0.10 and C=100 are operational safeguards derived from the observed
474 spectral distributions and should not be interpreted as universal physical constants.

475 *3.3 Accuracy-Weighted Composite Thresholding (AWCT)*

476 MCRCSI distributions can vary across regions and years due to differences in crop
477 composition, residue conditions, soil background, surface moisture, and other scene-specific
478 factors. Consequently, a thresholding method that performs well for one MCRCSI
479 distribution may not remain optimal under another. In addition, commonly used unsupervised
480 thresholding algorithms identify class boundaries according to different statistical criteria and
481 assumptions (Chen et al., 2024c), and may therefore respond differently to bimodal, weakly
482 separated, skewed, or multimodal MCRCSI distributions. Rather than prescribing a single
483 thresholding algorithm a priori, we therefore evaluated three representative and
484 complementary methods: Otsu's thresholding method (Otsu), K-means clustering (K-means),
485 and Gaussian Mixture Model (GMM), to assess their threshold estimates and classification
486 performance under varying scene conditions. Based on their complementary strengths, an

Accuracy-Weighted Composite Thresholding (AWCT) approach was further developed to integrate the three candidate thresholds and improve the consistency of residue mapping across diverse regions and years.

Otsu's method (Tellman et al., 2021) maximizes between-class variance in the index histogram, providing a purely statistical, data-driven threshold particularly suited to clearly bimodal distributions. K-means clustering (Wang et al., 2019) iteratively partitions pixels into two clusters by minimizing within-cluster variance based on feature-space distances and defines the threshold as the midpoint between the cluster centers, performing well under weak bimodality. The GMM (Bishop and Nasrabadi, 2006) represents the MCRCSI distribution as a weighted combination of two Gaussian components corresponding to residue-covered and residue-free surfaces, with the threshold determined by the intersection of their weighted probability density functions. The two-component assumption was validated on representative MCRCSI datasets through AIC and BIC comparisons of models with one to five components (**Fig.S6**).

We assessed the performance of each candidate thresholding method using a dedicated calibration set, and the corresponding F1 scores were used as weights for threshold integration. Specifically, Otsu, K-means, and GMM were first applied to derive three candidate thresholds, which were then evaluated against the calibration samples to obtain their respective F1 scores. The final composite threshold was derived through accuracy-weighted averaging:

$$T_{\mathrm{AWCT}} = \frac{\sum_i (\mathrm{F1}_i \cdot T_i)}{\sum_i \mathrm{F1}_i} \tag{10}$$

where T_i and $\mathrm{F1}_i$ represent the threshold and F1 score of the i-th method, respectively. Importantly, the calibration samples used for AWCT weighting were kept separate from the independent validation set used for the final accuracy assessment. Thus, the validation samples did not participate in threshold selection or weight estimation, ensuring that the reported classification accuracy represented an independent evaluation of the resulting AWCT classification. Through this procedure, the three candidate thresholds derived from

Otsu, K-means, and GMM were integrated according to their calibration-set F1 scores to obtain the final AWCT threshold. The AWCT threshold was then applied to the MCRCSI layers derived from Sentinel-2 multi-spectral images to generate final residue maps.

3.4 Accuracy assessment

We evaluate the accuracy of the algorithm by statistical metrics, including Overall Accuracy (OA), User's Accuracy (UA), Producer's Accuracy (PA), F1-score, and the Matthews Correlation Coefficient (MCC). Overall Accuracy (OA) represents the proportion of correctly classified samples relative to the total number of samples, reflecting the overall agreement between classification results and reference data. User's Accuracy (UA) indicates the reliability of a positive prediction for a given class, while Producer's Accuracy (PA) represents the proportion of reference samples of a given class that were correctly identified. The F1-score, defined as the harmonic mean of UA and PA, provides a balanced assessment of classification performance, particularly when class distributions are uneven, and was therefore used as a key metric for determining the optimal threshold in this study. In contrast, the Matthews Correlation Coefficient (MCC) takes into account all four elements of the confusion matrix, including true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), and remains robust even under conditions of class imbalance, making it one of the most comprehensive single metrics for binary classification tasks.

4. Results

4.1 Separability of MCRCSI

MCRCSI clearly distinguishes crop residue from bare soil and vegetation across the four BSRs (**Fig. 6** and **Figs. S7–S9**). It preserves visible differences between residue-covered fields and surrounding non-residue surfaces while suppressing signals from bare soil and green vegetation. Compared with CRCSI, MCRCSI exhibits smoother tonal transitions and less pronounced bright saturation within residue-covered fields. Its distinction between residue and unharvested crops is also clearer than that provided by DFI and BAI (346),

particularly in the USA and ARG regions. MCRCSI further provides more consistent visual discrimination of residue than BSI, NDTI, NDRI and NDSVI, and more clearly distinguishes residue from bare soil than NDVI.

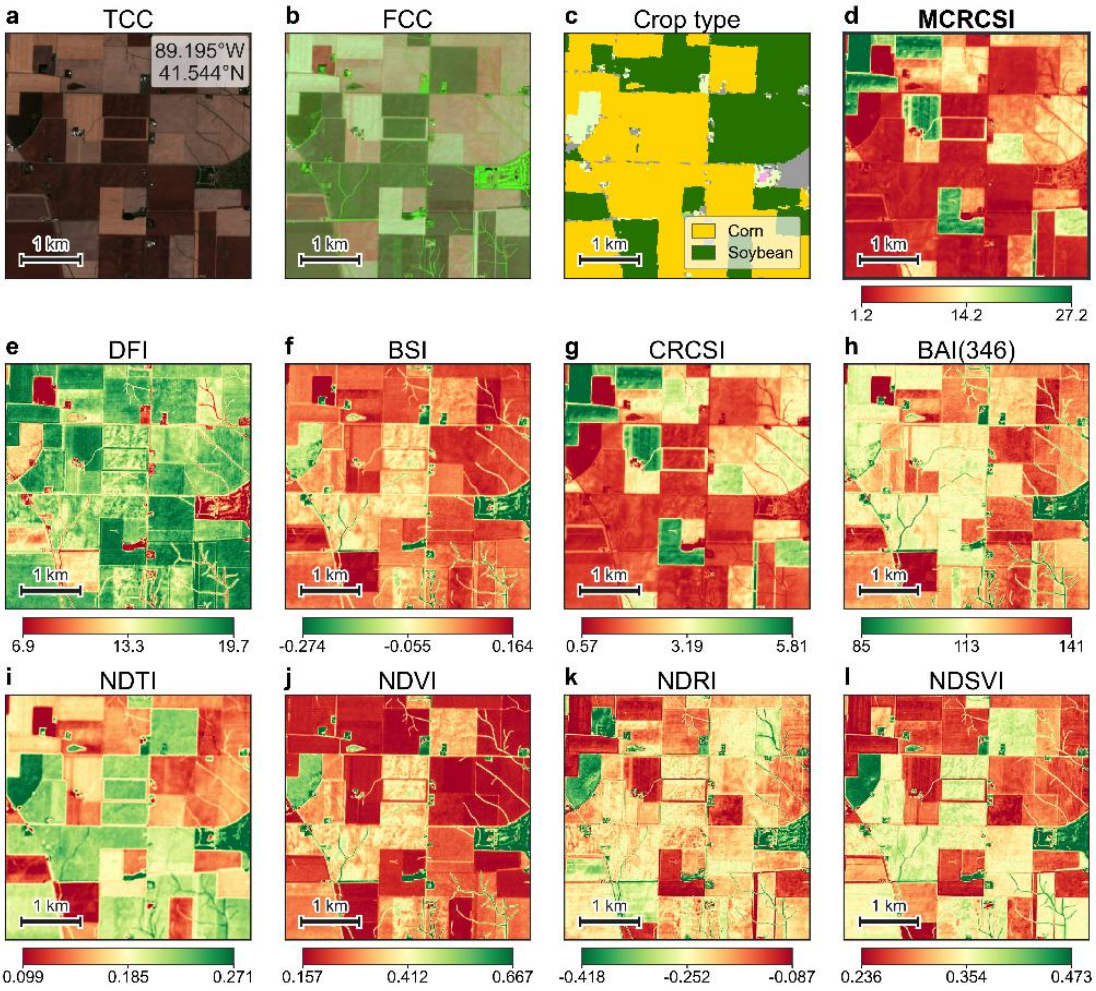


Fig. 6. Visual comparison of tillage indices in study area S2 (Illinois, USA). (a) True-color composite (TCC), (b) false-color composite (FCC), (c) crop-type map showing corn and soybean, and spatial distributions of (d) MCRCSI, (e) DFI, (f) BSI, (g) CRCSI, (h) BAI (346), (i) NDTI, (j) NDVI, (k) NDRI, and (l) NDSVI. For consistent visual interpretation, green denotes relatively high crop residue cover and red denotes relatively low crop residue cover in all index maps. Accordingly, for the positive-response indices (DFI, CRCSI, MCRCSI, NDTI, NDVI, and NDSVI), larger index values are shown in green, whereas for the negative-response indices (BSI, BAI (346), and NDRI), smaller index values are shown in green.

MCRCSI exhibits strong inter-class separability across all four regions, with M-index values ranging from 1.93 to 3.34 for the RC–BS, RC–CU, RC–GV, and RC–TC comparisons (**Table S5**). In the MCRCSI boxplots, the boxes representing RC do not overlap with those representing non-residue classes in any of the four regions (Fig. 7a). Welch's one-way

ANOVA revealed significant differences in MCRCSI among BS, RC, CU, and TC within each region (all $P < 0.001$). Subsequent two-sided Welch's *t*-tests showed significantly higher mean MCRCSI values for RC than for BS, CU, and TC in all four regions (all Holm-adjusted $P < 0.001$; Fig. 7k). Within-class variation remained evident, particularly in the wider RC interquartile range in the USA; nevertheless, the RC boxes did not overlap with those of BS, CU, or TC within any region (Fig. 7a). MCRCSI also showed clearer RC–CU, RC–TC, and RC–BS separation than DFI, BAI (346), and NDVI, respectively (Fig. 7a, c, d, i; **Table S5**). Across the five RC-related contrasts in **Table S5**, MCRCSI achieved a mean M-index of 2.94 and the lowest mean within-region CV among the evaluated indices (14.6%; (Fig. 7j)). Although CRCSI yielded a higher mean M-index (3.80), its mean within-region CV was also higher (22.7%), indicating that MCRCSI provided more uniform separation strength among class contrasts within regions. Together, these results show that MCRCSI maintained clear RC discrimination across the four study regions despite within-class variability, while exhibiting relatively balanced separability among the evaluated class contrasts.

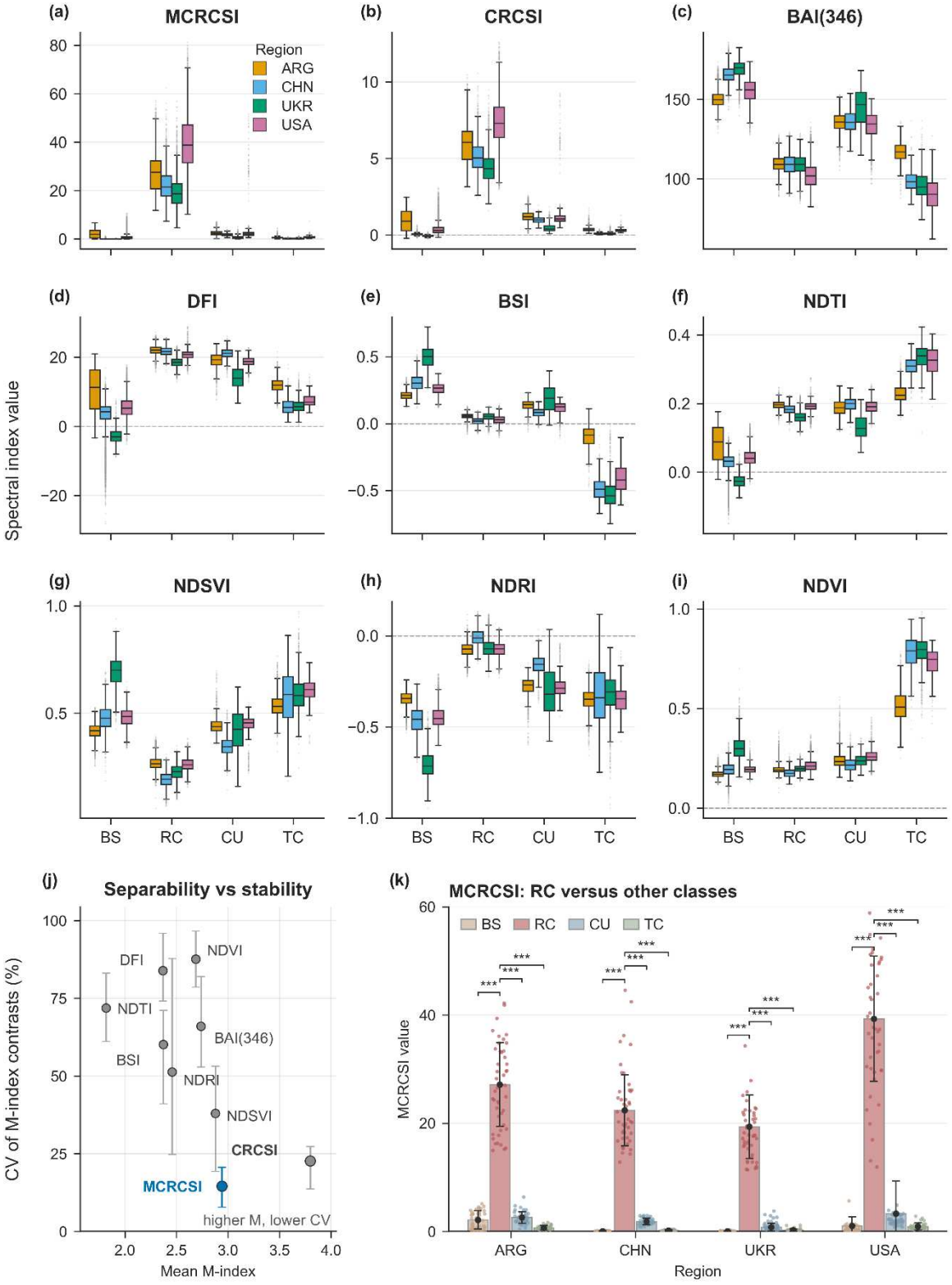
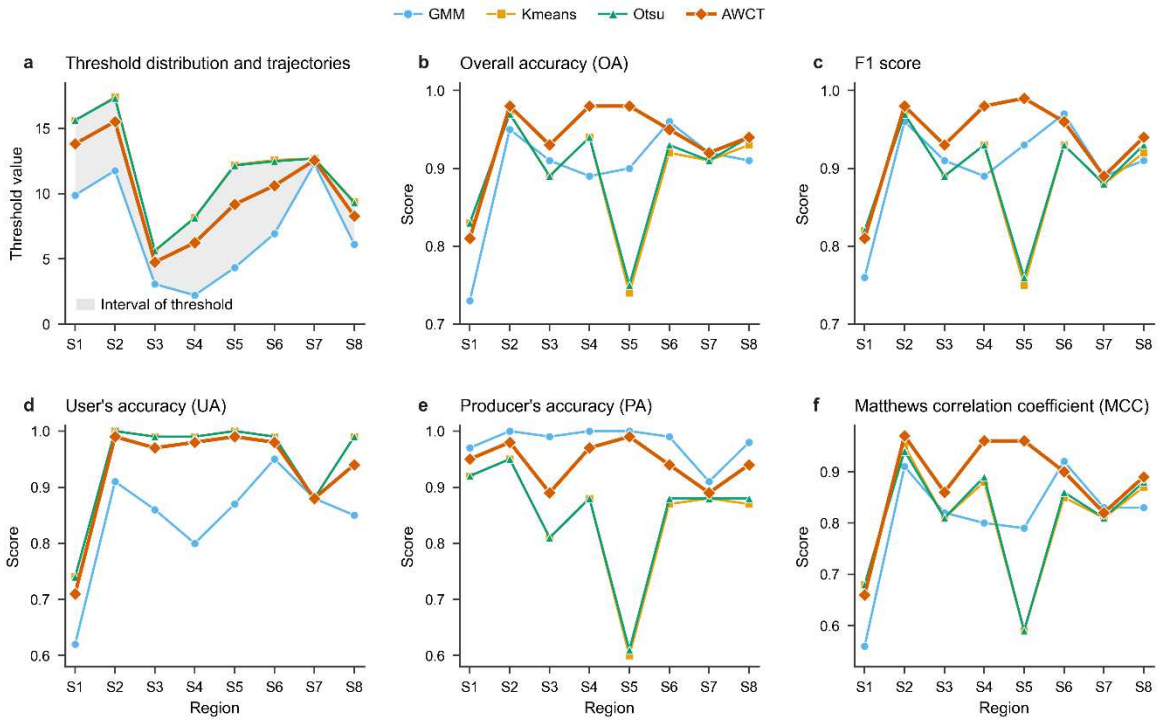


Fig. 7. Cross-region spectral responses, separability, and stability of crop-residue indices. (a–i) Distributions of MCRCSI, CRCSI, BAI (346), DFI, BSI, NDTI, NDSVI, NDRI, and NDVI for four surface classes—bare soil (BS), residue cover (RC), unharvested crops (CU), and tree cover (TC)—across Argentina (ARG), Northeast China (CHN), Ukraine (UKR), and the U.S. Midwest (USA). Boxplots indicate the median and interquartile range (IQR), with whiskers extending to 1.5 × IQR. (j) Separability–consistency comparison. Each point represents an index. The x-axis shows mean M-index across four regions and five

contrasts (RC–BS, RC–CU, RC–GV, RC–TC, and RC–WB; GV, green vegetation; WB, water bodies; **Table S5**). The y-axis shows mean regional CV ($SD/mean \times 100\%$), calculated across the five contrasts within each region; vertical bars span the minimum–maximum regional CVs. Higher M and lower CV indicate stronger and more uniform separability, respectively; narrower bars indicate less regional variation. (k) Bars represent mean MCRCSI values, and error bars indicate ±1 standard deviation. All available observations were used to calculate summary statistics and perform statistical tests. Brackets indicate comparisons between RC and each of the other three classes using two-sided Welch's *t*-tests, with Holm correction applied to the three comparisons within each region (***, $P < 0.001$).

4.2 Performance of MCRCSI-AWCT for residue mapping

The three individual thresholding methods yielded systematically different thresholds across the eight test regions (**Fig. 8**a). GMM consistently produced the lowest thresholds, whereas K-means and Otsu yielded nearly identical, higher estimates. AWCT thresholds fell between these estimates and were consistently closer to K-means and Otsu than to GMM. The largest separation among the individual methods occurred in S5, where thresholds ranged from 4.32 to 12.21 and AWCT yielded 9.18. Conversely, the methods converged in S7, with thresholds ranging from 12.27 to 12.69 and an AWCT threshold of 12.55. Across regions, AWCT thresholds were highest in S2 (15.52), followed by S1 (13.81), and lowest in S3 (4.75).



598

599 **Fig. 8. Metric evaluation and threshold comparison of different methods across eight**
600 **study sites (S1–S8).** (a) Thresholds derived from GMM, K-means, Otsu, and AWCT
601 methods, (b)–(f) Accuracy metrics including (b) Overall Accuracy (OA), (c) F1-score, (d)
602 User's Accuracy (UA), (e) Producer's Accuracy (PA), and (f) Matthews correlation
603 coefficient (MCC).

604 The resulting residue maps revealed distinct field-scale patterns across the eight test regions
605 (**Fig. 9**). Residue-covered areas generally followed regular parcel boundaries in the U.S.
606 Midwest (S1–S2) and Argentina (S7–S8), whereas the Ukrainian sites (S3–S4) contained
607 fewer but larger residue-covered fields. Northeast China (S5–S6) exhibited finer, more
608 fragmented patterns, with more extensive residue cover in S6 than in S5. Across regions, the
609 mapped residue patches broadly aligned with the field structures visible in the false-colour
610 composite (FCC) imagery.

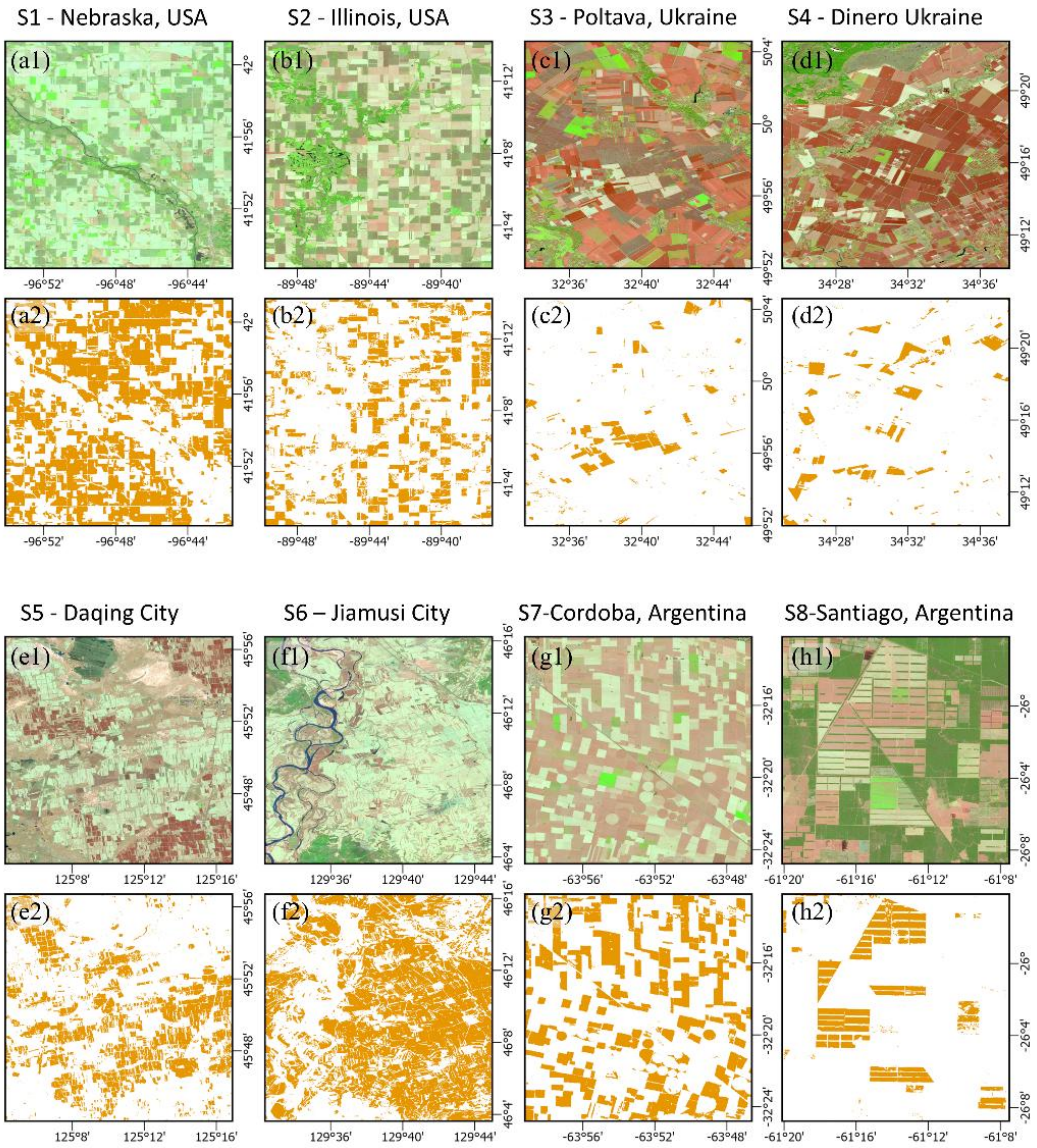


Fig. 9. Sentinel-2 based crop residue mapping in S1–S8. (a1–h1) show false-color images, (a2–h2) present the mapping results.

MCRCSI-AWCT achieved an overall OA of 0.936, F1 score of 0.935, and MCC of 0.878, exceeding the three individual methods by 4.0–4.5 percentage points in OA, 0.033–0.049 in F1, and 0.070–0.073 in MCC (**Fig. 8**b–f). It also achieved a balanced PA (0.944) and UA (0.930). Regionally, performance was highest in S2, S4, and S5, with OA reaching 0.98 and F1 ranging from 0.98 to 0.99, whereas S1 had the lowest OA and F1, both at 0.81. Relative to the best individual method in each region, the largest OA gains occurred in S5 and S4, at 8 and 4 percentage points, respectively. The improvement was particularly pronounced over K-means and Otsu in S5, where OA increased by 23–24 percentage points. However, AWCT

622 slightly underperformed the best individual method in S1 and S6 and matched its OA in S7
623 and S8, indicating that the benefits varied across regions.

624 *4.3 Comparison with machine learning classifiers*

625 To further evaluate the spatial performance of the MCRCSI–AWCT framework, we
626 compared its mapping results with those from three commonly used machine learning (ML)
627 classifiers: Random Forest (RF), XGBoost, and LightGBM. All classifiers were trained using
628 the same labeled samples and four Sentinel-2 surface-reflectance bands (red, near-infrared,
629 SWIR1, and SWIR2).

630 Across the four regions, the proposed method ("This study" in **Table 4**) achieved the highest
631 mean balanced accuracy (83.41%), precision (92.89%), F1 score (80.95%), and MCC (0.664).
632 Compared with RF, which was the strongest ML classifier in terms of mean balanced
633 accuracy, these values represented improvements of 7.58, 18.69, and 3.09 percentage points
634 in balanced accuracy, precision, and F1 score, respectively, together with an MCC increase
635 of 0.127. In contrast, the proposed method produced a lower mean recall (72.67%) than RF
636 (86.93%), XGBoost (88.98%), and LightGBM (88.05%). The regional results showed that
637 this advantage was consistent in three of the four study regions. The proposed method
638 achieved the highest balanced accuracy in the U.S. Midwest (84.81%), Northeast China
639 (83.30%), and Argentina (86.92%), exceeding the best-performing ML classifier in these
640 regions by 6.97, 10.43, and 15.94 percentage points, respectively. It also produced the highest
641 F1 score and MCC in these three regions. Ukraine was the main exception, where RF
642 achieved a balanced accuracy of 83.60%.

643 **Table 4** Regional classification performance of RF, XGBoost, LightGBM, and the proposed
644 MCRCSI–AWCT method. Mean values are unweighted averages across the four study
645 regions. Bold values indicate the best result within each region and metric. Residue cover was
646 treated as a positive class.

Sites	Method	Balanced Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	MCC
U.S Midwest	RF	77.84	73.84	96.89	83.81	0.615
	XGBoost	57.30	58.49	**98.33**	73.35	0.263

Sites	Method	Balanced Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	MCC
Ukraine	LightGBM	76.23	72.15	97.72	83.01	0.595
	MCRCSI-AWCT	**84.81**	**92.92**	76.61	**83.98**	**0.697**
	RF	**83.60**	**100.00**	**67.19**	**80.38**	**0.615**
	XGBoost	81.80	99.94	63.69	77.80	0.585
	LightGBM	80.51	99.94	61.12	75.85	0.563
	MCRCSI-AWCT	78.60	**100.00**	57.19	72.77	0.533
Northeast China	RF	70.90	62.89	96.85	76.26	0.486
	XGBoost	72.01	63.87	96.69	76.93	0.503
	LightGBM	72.87	64.55	**97.08**	77.54	0.520
	MCRCSI-AWCT	**83.30**	**88.33**	76.31	**81.88**	**0.674**
Argentina	RF	70.98	60.09	86.79	71.01	0.432
	XGBoost	68.61	55.76	**97.21**	70.87	0.434
	LightGBM	69.39	56.57	96.29	71.27	0.441
	MCRCSI-AWCT	**86.92**	**90.31**	80.57	**85.16**	**0.751**
Mean	RF	75.83	74.20	86.93	77.86	0.537
	XGBoost	69.93	69.52	**88.98**	74.74	0.446
	LightGBM	74.75	73.30	88.05	76.92	0.530
	MCRCSI-AWCT	**83.41**	**92.89**	72.67	**80.95**	**0.664**

The spatial comparison further revealed clear differences in within-field mapping patterns (**Fig. 10**). Across the four landscapes, pixels receiving a majority vote for residue (vote counts of 2 or 3) formed more scattered and discontinuous patches, whereas the proposed method generally delineated residue cover as more contiguous field-scale units. The displayed areas also contained numerous apparent misclassifications by the ML classifiers, including isolated residue assignments and within-field omissions that disrupted otherwise homogeneous field patterns. These errors were particularly evident in the elongated parcels of central U.S. Midwest and the heterogeneous croplands of Northeast China (**Fig. 10**a3–a4 and c3–c4). These qualitative spatial patterns were consistent with the higher precision and MCC of the

proposed method, although its lower recall indicated that some residue-covered areas remained undetected.

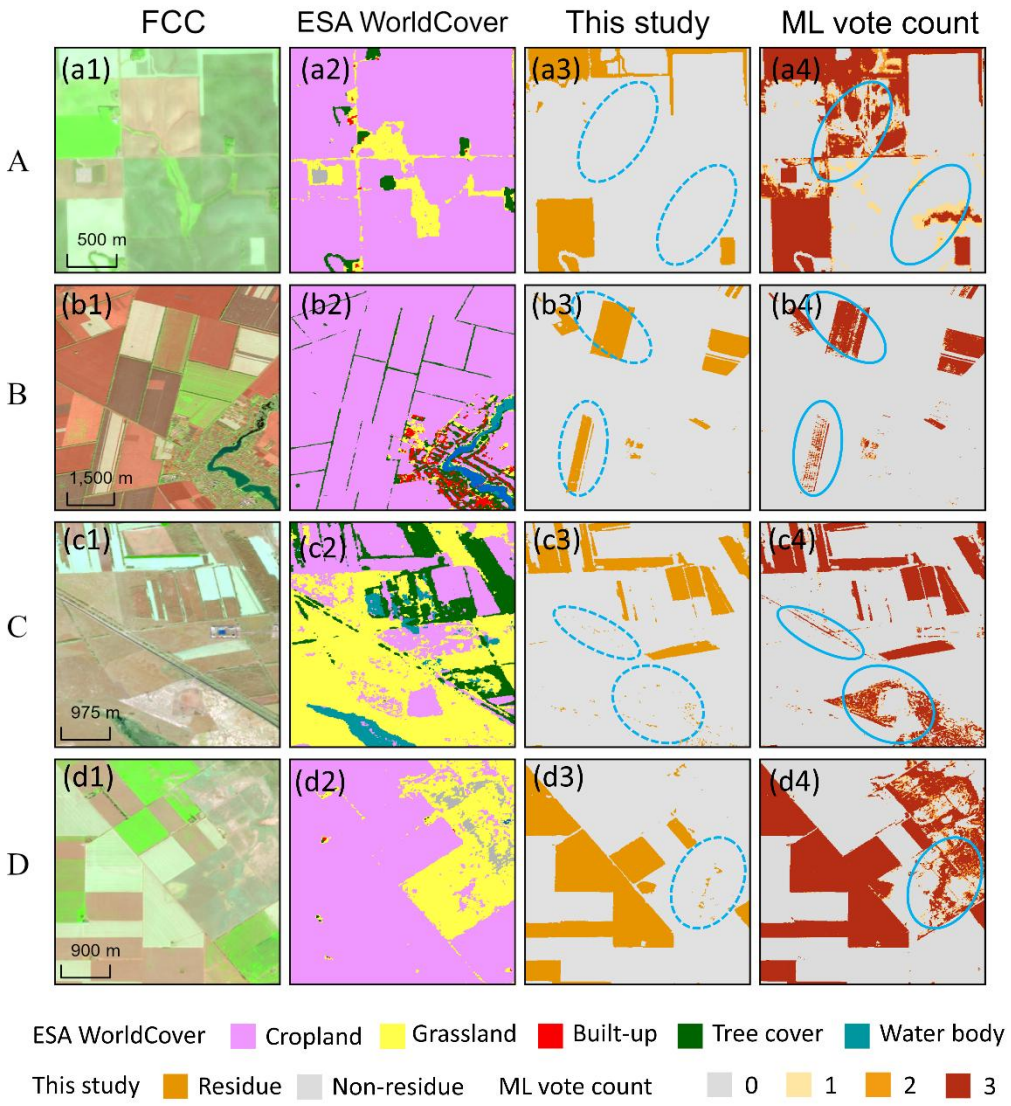


Fig. 10. Spatial comparison of residue-cover mapping by the proposed method and tree-based machine-learning classifiers. Four representative agricultural landscapes are shown: A, the U.S. Midwest (42°18′24″ N, 96°49′ W); B, central Ukraine (49.5329° N, 32.1916° E); C, Northeast China (46°32′7″ N, 124°43′3″ E); and D, the Argentine Pampas (32.1585° S, 63.5340° W). Each row presents a false-color composite (FCC), an ESA WorldCover land-cover map, the binary residue/non-residue map produced by MCRCSI-AWCT (this study), and a machine-learning (ML) vote-count map. The ML vote count indicates how many of the three classifiers (RF, XGBoost, and LightGBM) assigned each pixel to the residue class, ranging from 0 to 3. Vote counts of 2 or 3 correspond to the majority-vote residue classification. Blue ellipses highlight representative areas where the ML results show more fragmented or scattered residue-cover patterns than the proposed-method results.

5. Discussion

5.1 Advantages of the proposed method

Our proposed MCRCSI–AWCT framework offers four advantages over existing studies. First, MCRCSI enhanced the discrimination between crop residue (RC) and unharvested mature crops (CU), which is a particularly challenging class pair in residue cover mapping because senescent standing crops can exhibit spectral characteristics similar to those of post-harvest residues. This limitation was evident among the comparison indices: For RC–CU separation, MCRCSI achieved a mean M statistic of 2.60 across the four regions, higher than NDTI (0.33), NDRI (1.83), NDSVI (2.13), DFI (0.68), NDVI (0.77), BSI (1.12), and BAI (346) (1.99), and slightly lower than CRCSI (2.93) (**Table S5**). This improvement arises from applying a joint NIR × Red gate to the residue-sensitive SWIR contrast (SWIR1-SWIR2). During crop senescence, chlorophyll-related changes affect visible and red-edge/NIR reflectance (Xu et al., 2019), while SWIR reflectance also varies with crop phenology (Zhong et al., 2016). Among the indices evaluated here, NDTI uses only SWIR1 and SWIR2, NDRI combines Red and SWIR2, and BAI (346) uses Red, NIR, and SWIR2, while the multiband DFI and BSI integrate these spectral domains primarily through ratios or additive contrasts. In contrast, the NIR × Red product in MCRCSI acts as an AND-like constraint, retaining a strong response only when both bands exhibit residue-like reflectance and thereby preferentially suppressing CU surfaces that resemble residue in only one spectral domain. Accordingly, the ratio of the median response of RC to that of CU reached 13.89 for the joint NIR × Red gate, compared with 6.08 for the Red-only gate and 5.51 for the NIR-only gate (**Fig.S10**a). Although the RC–CU M statistic of MCRCSI was slightly lower than that of CRCSI, MCRCSI produced a substantially greater relative response contrast: its median RC-to-CU response ratio was 13.89, 2.42 times the CRCSI ratio of 5.75 (**Fig.S10**b). This improved discrimination between pre- and post-harvest surfaces also provides a spectral basis for harvest monitoring (**Fig.S11**) (Tang et al., 2026), enabling the proposed method to infer harvest status from single-date imagery and to track crop-to-residue transitions using multi-temporal observations.

Second, MCRCSI exhibited a pronounced sensitivity to variations in residue cover fraction. Based on the USGS field spectral dataset from Hively et al. (2021), we found MCRCSI produced the largest robust standardized median difference between low-cover (≤20%) and high-cover (≥80%) samples among the evaluated indices, reaching 2.45 compared with 2.29 for CRCSI, 1.74 for NDSVI, 1.66 for DFI, and 1.59 for NDTI (**Fig. 11**). The strong response of MCRCSI across the residue-cover gradient, particularly its notable contrast between sparse and dense residue fractions, provides a favorable spectral basis for extending the index from binary residue detection to quantitative fractional cover retrieval when calibrated against field-measured cover fractions.

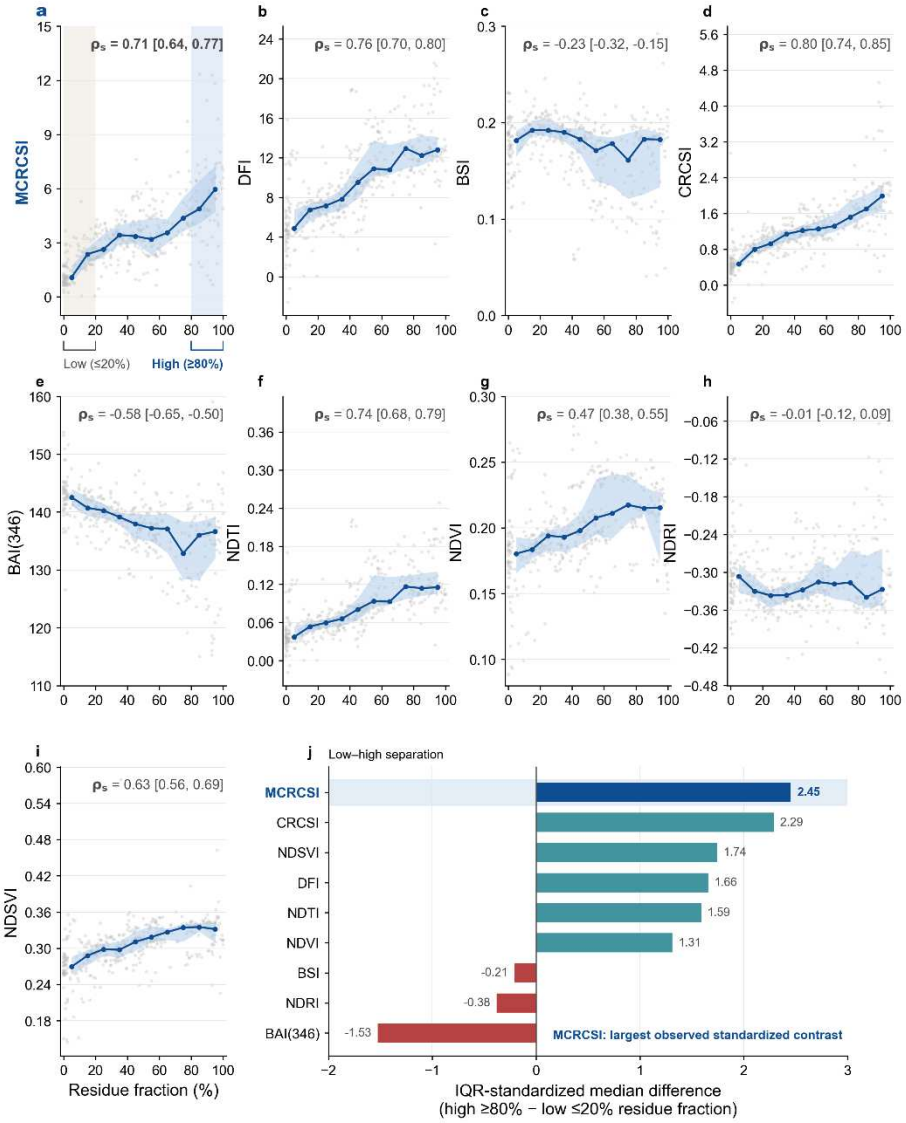


Fig. 11. Spectral-index response to crop residue fraction change. (a–i) Relationships between residue fraction and nine Sentinel-2 spectral indices; lines and ribbons show bin

710 medians and interquartile ranges, respectively. Values of ρ_s are Spearman's rank correlation
711 coefficients with 95% confidence intervals. (j) IQR-standardized median differences
712 between high (≥80%) and low (≤20%) residue fractions, calculated as $(\text{median}_{high} -$
713 $\text{median}_{low})/(\text{IQR}_{all}/1.349)$, where IQR_{all} is calculated across all samples. A value of 0
714 indicates no difference between group medians, whereas +1 or −1 indicates that the high-
715 fraction median is one robust standard deviation above or below the low-fraction median,
716 respectively. Larger absolute values indicate stronger separation. MCRCSI shows the largest
717 positive separation.

718 Our proposed framework shows high adaptability to geographically distinct regions.
719 MCRCSI-AWCT shows high performance across global black soil regions in Argentina,
720 China, Ukraine, and the USA. Only use of a small set of calibration samples, MCRCSI-
721 AWCT can reach the F1 score higher than 0.81 for residue mapping in eight test regions.
722 Previous studies have often focused on local regions or limited spatial extent, such as county-
723 level mapping (Zhang et al., 2025) or field-scale experiments (Zhou et al., 2025). In contrast,
724 the cross-continental evaluation conducted here demonstrates that MCRCSI can adapt to
725 diverse cropping systems, within global black soil regions. The additional experiments
726 beyond black soil regions show the potential of MCRCSI-AWCT adapting to other soil types,
727 such as sandy soils, and brown soils (**Fig.S12**). This high adaptability can be mainly
728 attributed to our two-step semi-supervised thresholding approach. AWCT proved to be more
729 reliable and stable than the stand-alone unsupervised clustering approaches. As the residue-
730 cover class proportion changed, AWCT adaptively adjusted the weights assigned to
731 individual clustering methods while maintaining a relatively stable threshold range of 12.67–
732 14.19 (**Fig.S13**a-b). It also achieved the highest median MCC under varying levels of class
733 imbalance samples (**Fig.S13**c-d). These findings suggest that AWCT can accommodate shifts
734 in class prevalence more effectively than Otsu, K-means, or GMM alone. Meanwhile, AWCT
735 only needs a small set of calibration samples to synthesize the final threshold. Compared with
736 traditional machine learning models, AWCT needs smaller number of samples but can
737 maintain a comparable accuracy (83.41% vs. 69.93–75.83%, **Table 4**).

738 In addition, the proposed MCRCSI supports intra-annual residue cover dynamic monitoring.
739 In Northeast China, MCRCSI captured the pronounced decline in residue cover during the
740 pre-sowing period, with the mapped residue fraction decreasing from 23.2% on 4 March to

741 0.4% on 13 April 2023, consistent with progressive residue removal or incorporation during
742 spring field preparation (**Fig. 12**a–d). In contrast, the U.S. Midwest exhibited the opposite
743 seasonal trajectory during autumn harvest, where residue cover increased from 0.6% on 9
744 September to 26.7% on 24 October 2024 and remained high at 25.1% on 29 October (**Fig.**
745 **12**e–h). This single-scene detection capability is important since residue cover can change
746 rapidly over short periods, enabling timely monitoring of post-harvest management without
747 relying on temporal compositing. In practical applications, however, cloud and snow
748 contamination should be rigorously masked to avoid spurious temporal changes.

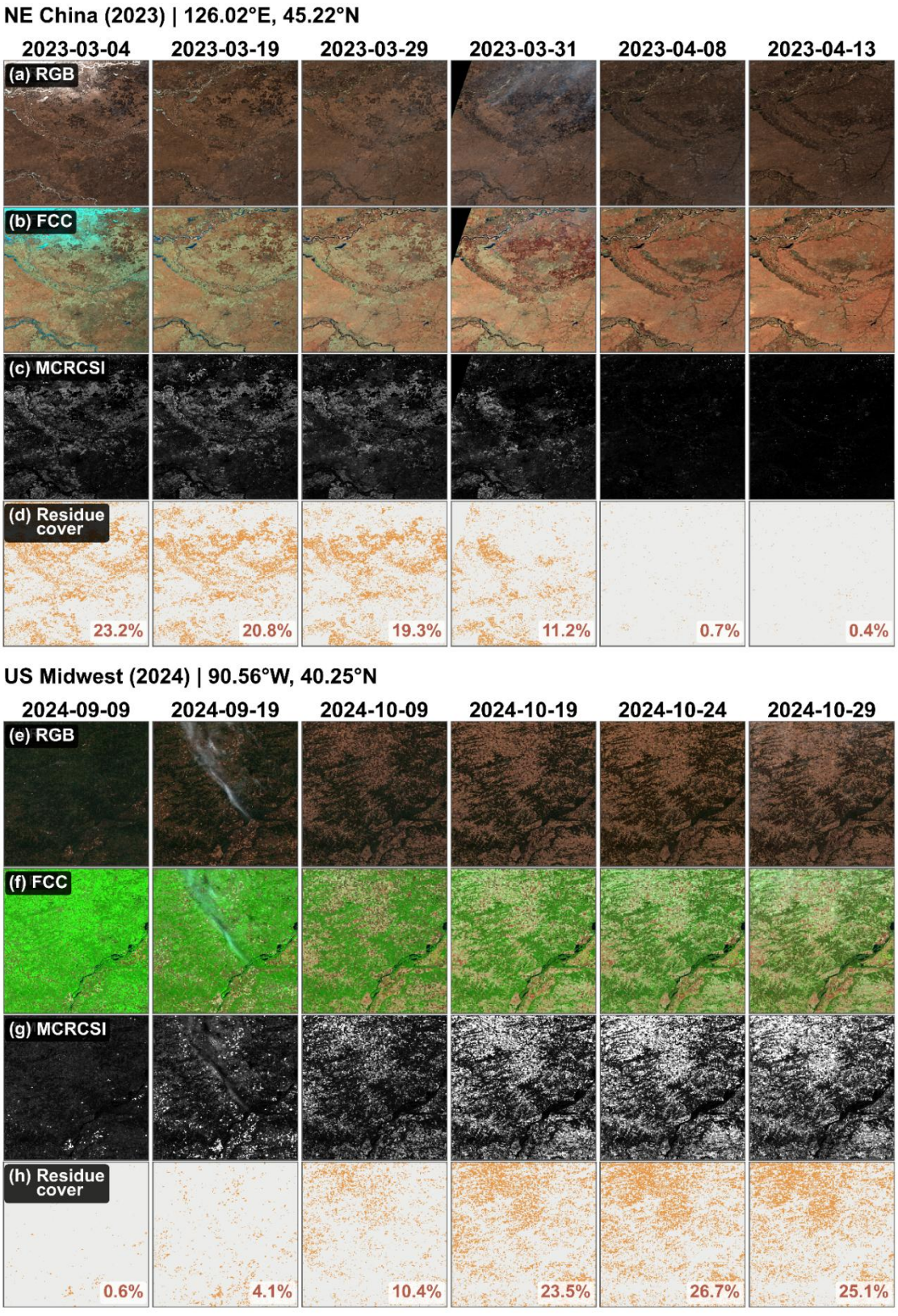


Fig. 12. Contrasting temporal trajectories of MCRCSI-derived residue-like cover in NE China and the US Midwest. (a–d) RGB, false-color composite (FCC), MCRCSI and threshold-derived residue-cover maps for six selected dates in NE China in 2023. (e–h) Corresponding observations for the US Midwest in 2024. FCC images use the SWIR2–NIR–Red combination; RGB and FCC images are displayed using a fixed 0–4500 channel stretch,

and MCRCSI images use a fixed grayscale stretch of 0–20. Residue-like pixels (orange) were defined as pixels with MCRCSI values above the fixed regional threshold, whereas pixels at or below the threshold are shown in light grey. Percentages in (d) and (h) represent the proportion of residue-like pixels for each date.

5.2 Explanation of the spatial-temporal variability of MCRCSI threshold

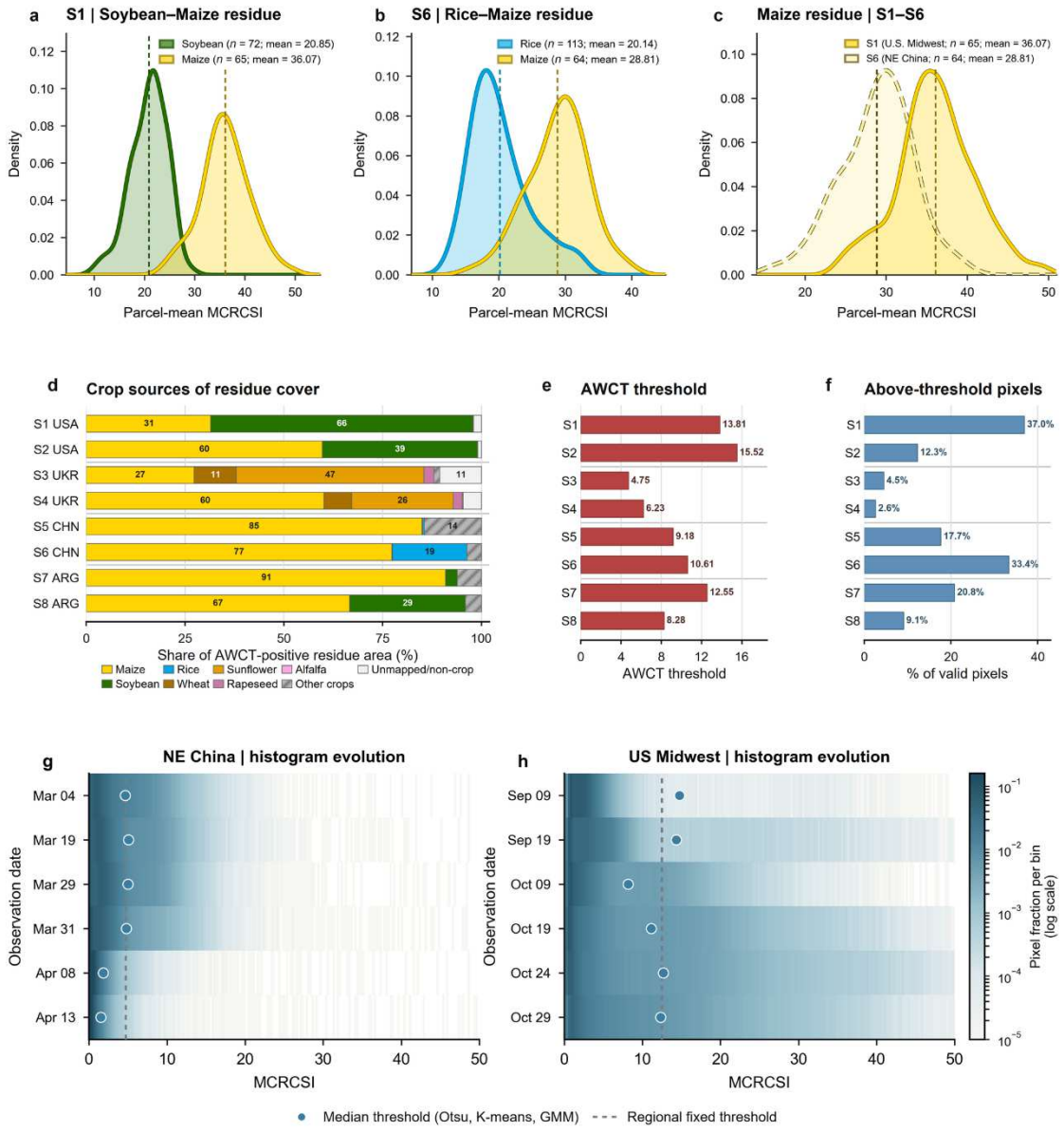
MCRCSI thresholds for distinguishing residue from non-residue varied among the eight test regions, ranging from 4.75 in S3 (Ukraine) to 15.52 in S2 (U.S. Midwest) (**Table S6**). This variability reflects spatial and temporal shifts in scene-level MCRCSI distributions driven by crop-specific residue responses, regional crop composition, and post-harvest surface dynamics. These differences highlight the need for adaptive rather than globally fixed thresholds across heterogeneous agricultural systems.

Spatial variability in MCRCSI was primarily associated with the crop types contributing to residue cover and their relative abundance within each region. Residues from different crops yielded different MCRCSI values, likely because variation in their biochemical composition and physical structure produced distinct spectral responses. Within the same scene, maize and soybean residues showed systematically different MCRCSI distributions (**Fig. 13**a). Crop composition then determines the relative contribution of these crop-specific components to the scene-level distribution. Conceptually, the regional MCRCSI distribution can therefore be regarded as a mixture of crop-specific distributions weighted by their relative abundance. Landscapes dominated by crops with relatively high residue MCRCSI values consequently tend to exhibit a stronger high-value component than landscapes dominated by crops with lower residue responses (**Fig. 13**d–f), thereby shifting the threshold required to separate residue from background. This mechanism is particularly relevant to soybean-dominated systems: because soybean residue generally exhibits lower MCRCSI values than maize residue, applying a threshold derived from maize-dominated landscapes could increase omission of soybean residue.

Crop composition, however, does not fully explain regional threshold differences. Even residues from the same crop exhibited substantial variation among regions. For example, maize residues in S1 and S6 occupied distinct MCRCSI ranges (**Fig. 13**c), indicating that

784 crop type does not uniquely determine the absolute MCRCSI distribution. Such within-crop
785 variability likely reflects differences in residue amount and condition, fragmentation, soil
786 background, moisture, weathering, and management context.

787 Temporal variability arises mainly from changes in residue abundance and surface condition
788 during the post-harvest period. As harvesting progresses, increasing exposure of crop
789 residues strengthens the residue-related high-MCRCSI component of the scene distribution.
790 Subsequent residue removal, incorporation, and tillage can progressively reduce observable
791 residue cover and weaken or contract this high-value component (**Fig. 13**g–h). Consequently,
792 observations acquired at different stages of the harvest–management sequence may exhibit
793 substantially different MCRCSI distributions even within the same region.



794

795 **Fig. 13. Spatial and temporal variability of MCRCSI distributions.** (a–b) MCRCSI
796 distributions of major crop residues in S1 (soybean and maize) and S6 (rice and maize), with
797 dashed lines indicating crop-specific means. (c) Comparison of maize-residue MCRCSI
798 distributions between S1 and S6. (d–f) Crop composition, AWCT thresholds, and
799 proportions of valid pixels exceeding the thresholds across S1–S8. (g–h) Temporal MCRCSI
800 distributions in Northeast China and the U.S. Midwest; circles indicate median thresholds
801 from Otsu, K-means, and GMM, and dashed lines indicate regional fixed thresholds.
802 Background colour indicates the fraction of valid pixels within each MCRCSI bin on a
803 logarithmic scale, with darker colours representing higher pixel fractions.

804 *5.3 Influence of residue moisture on MCRCSI and residue cover mapping*

Moisture content can alter the MCRCSI values and further introduce error for residue mapping. We systematically assess the effects of residue moisture on MCRCSI using relative water content (RWC) dataset Quemada and Daughtry (2016). We found that increasing residue moisture systematically reduced MCRCSI even when residue cover remained unchanged. MCRCSI decreases by 63.7% for maize residue and 47.1% for soybean residue when RWC increases to 100% (**Fig. 14**a–b). The standardized response of MCRCSI (0.281 decrease per 10% RWC increase) fell within the overall range of the evaluated indices, and its endpoint change was smaller than those of DFI, NDRI, BSI, and NDTI. However, the moisture sensitivity of MCRCSI was higher than BAI (346), which was specifically designed to reduce residue-water effects (Yue et al., 2020).

This effect was also revealed by Sentinel-2 time-series observations. During a snow-free rainfall event in town of Shuangcheng, Northeast China (45.38° N, 126.31° E), 7.25 mm of precipitation increased surface soil moisture by 0.0316 m^3 m^{-3} between 25 and 27 October 2022 (**Fig.S14**a-c). Pixels with pre-event MCRCSI <10 changed little (+0.083), whereas those with MCRCSI of 10 – 20 and ⩾20 decreased by 1.606 and 3.668 units, respectively (**Fig. 14**c). Because higher MCRCSI values indicate a greater likelihood of residue cover, rainfall-induced wetting preferentially shifted likely residue pixels toward lower MCRCSI values. The negative relationship between changes in NDMI and MCRCSI further supports this moisture effect (**Fig. 14**d). Consequently, imagery acquired shortly after rainfall may underestimate residue-covered area, particularly where classification relies on thresholds calibrated under drier conditions.

Snow cover, which is a common seasonal surface disturbance in high-latitude black soil regions, produced an even stronger interference for residue mapping. For potential residue pixels, median MCRCSI decreased by 79.6%, from 14.18 on 9 November to 2.89 on 21 November, and further to 1.45 on 26 November as snow developed (**Fig. 14**e–f). As snow progressively covered post-harvest residues, the proportion of exposed residue within the satellite-observed surface decreased while the snow-covered fraction increased, thereby weakening the residue spectral signal. Moreover, subsequent snowmelt can further wet

exposed residues, prolonging the suppression of MCRCSI even after snow cover begins to recede. MCRCSI was also negatively correlated with NDMI (r = −0.559) and NDSI (r = −0.582) (**Fig.S14**d), indicating systematic downward shifts under wetter and snow-affected conditions. These results imply that moisture, freezing, and snow can introduce substantial temporal inconsistency in residue-cover maps even without actual changes in residue abundance. Therefore, relatively dry and snow-free observations should be prioritized, with explicit cloud/snow masking and moisture screening; for multi-temporal mapping, temporal compositing (Cui et al., 2026) or moisture-condition-specific threshold calibration may further reduce omission bias and improve temporal comparability.

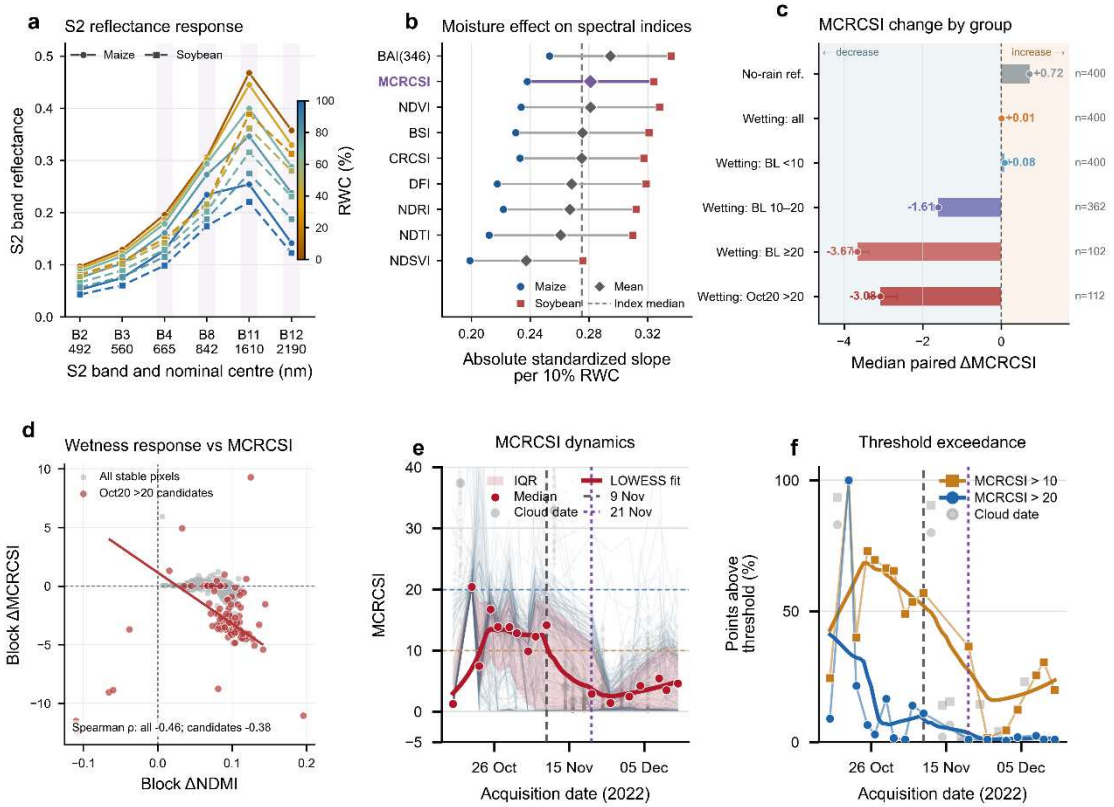


Fig. 14. Moisture and snow effects on MCRCSI from residue spectra and Sentinel-2 time-series observations. (a) Sentinel-2 band reflectance of maize (solid lines, circles) and soybean (dashed lines, squares) residues across five relative water content (RWC) levels. Colours denote RWC; shading marks the bands used in MCRCSI. (b) Absolute standardized response slopes per 10% increase in RWC for nine spectral indices. Slopes are standardized by each crop–index trajectory's across-level standard deviation. Connected circles and squares represent maize and soybean; diamonds show their means, and the dashed line marks the median across indices. (c) Paired MCRCSI changes for all stable pixels during the no-rain reference and wetting intervals. Wetting-event subgroups comprise three BL strata (BL

852 <10, 10 ≤ BL <20 and BL ≥20) and Oct20 >20 candidates. Bars and points show medians
853 across 1-km blocks, each summarized by its median paired pixel change; error bars indicate
854 95% bootstrap confidence intervals, and n denotes block counts. (d) Block-median ΔNDMI–
855 ΔMCRCSI relationships during wetting, where NDMI is the normalized difference moisture
856 index. Grey and red points represent all stable pixels and candidates, respectively; the red
857 line denotes the candidate Theil–Sen fit, with Spearman's ρ reported. (e) MCRCSI dynamics
858 for candidate residue-cover pixels in Northeast China, showing medians, interquartile ranges
859 and locally weighted scatterplot smoothing (LOWESS) fits. Horizontal dashed lines mark
860 thresholds of 10 and 20. (f) Proportions exceeding MCRCSI thresholds of 10 (orange squares)
861 and 20 (blue circles), with LOWESS fits. In (e,f), grey symbols denote cloud-affected
862 observations excluded from fitting; vertical lines mark the last snow-free observation (9
863 November) and first observed snow (21 November). Ref. = 15–20 Oct; Wetting = 25–27 Oct;
864 BL = pre-rain MCRCSI on 25 Oct. Oct20 >20 denotes candidates preselected with
865 MCRCSI >20 on 20 October. All dates are in 2022.

866 *5.4 Consistency of MCRCSI across sensors*

867 We further examine the consistency of MCRCSI across S2, Landsat, and MODIS to assess
868 the transferability across sensors. Images were acquired on the same or adjacent days, with
869 reflectance resampled to a common resolution before MCRCSI calculation and paired values
870 randomly sampled within the cropland mask. The strong Sentinel-2–Landsat relationships
871 across the four BSRs (R^2 = 0.92–0.94; **Fig. 15**a–d) indicate that MCRCSI preserves largely
872 consistent spatial variability between the two sensors, supporting their potential integration
873 for longer-term monitoring. Nevertheless, systematic departures from the 1:1 relationship
874 suggest that sensor-specific differences remain and that cross-sensor calibration may be
875 required before constructing a harmonized time series (Li and Roy, 2017).

876 In contrast, agreement between Sentinel-2 and MODIS was substantially weaker and more
877 regionally variable (R^2 = 0.19–0.67; **Fig. 15**e–h). A plausible contributor is sub-pixel spectral
878 mixing within the 500 m MODIS footprint: even after cropland masking, individual pixels
879 may integrate fields with different crop or residue conditions together with non-cropland
880 surfaces, and resampling alone cannot make these measurements spectrally equivalent to
881 finer-resolution observations. The spatial comparisons further corroborated this
882 interpretation (**Fig.S15**). Landsat 8/9 preserved much of the field- and patch-scale variability
883 evident in Sentinel-2 imagery, whereas MODIS represented the same landscapes as
884 substantially smoother and more spatially aggregated patterns, with field boundaries, narrow

strips, and localized MCRCSI extremes largely unresolved. These findings support integrating Sentinel-2 and Landsat for longer-term MCRCSI monitoring after cross-sensor calibration, whereas MODIS is better suited to regional-scale applications than direct field-scale comparison.

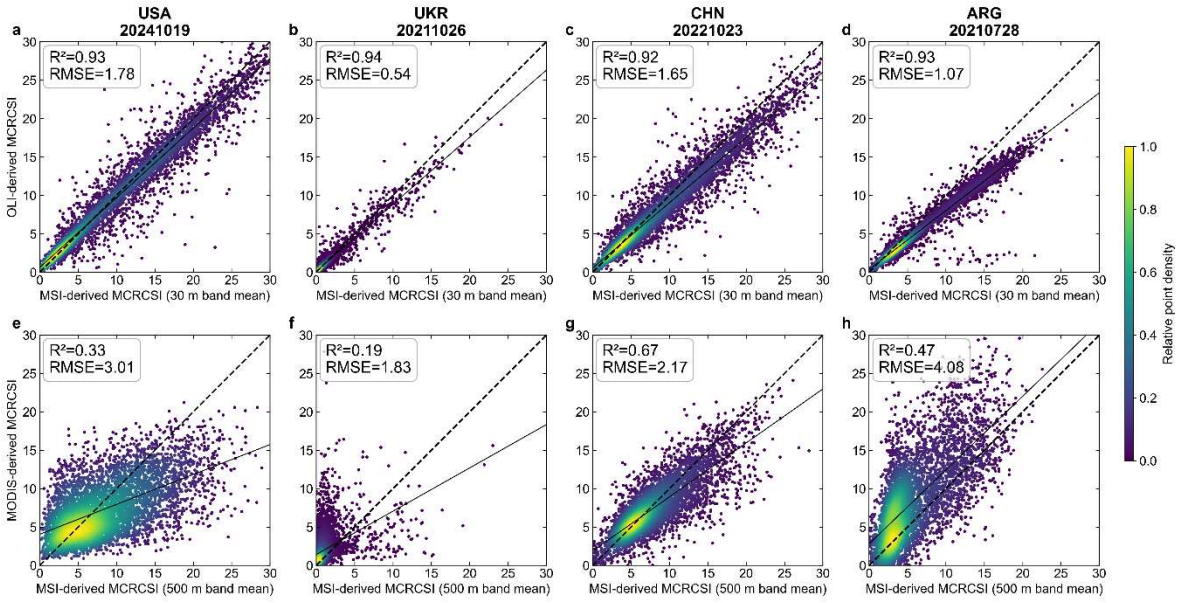


Fig. 15. Cross-sensor consistency of MCRCSI after spatial-resolution matching. (a–d) Comparison of Landsat OLI-derived MCRCSI with Sentinel-2 MSI-derived MCRCSI after aggregating MSI reflectance to 30 m in the USA, Ukraine, China, and Argentina, respectively; (e–h) corresponding comparison of MODIS-derived MCRCSI with MSI-derived MCRCSI after aggregating MSI reflectance to 500 m. Colors indicate point density. Solid lines denote linear regressions and dashed lines the 1:1 relationship; $R^2$and RMSE are reported in each panel. Only pixels with ≥50% cropland were included.

5.5 Uncertainties and future development

Despite the improved spectral separability and threshold robustness demonstrated in this study, several uncertainties should be considered when applying the proposed framework. Residual cloud contamination, particularly thin or incompletely masked clouds, can artificially increase MCRCSI values and cause false residue detections, making rigorous cloud and snow masking essential, especially in cold or frequently cloudy agricultural regions (Li et al., 2022). At low residue cover, spectral mixing with soil also limits detectability: apparent detection increased from 71% at 10–15% cover to 87% at 15–20%, first exceeding the 80% criterion, while cross-validation at 15% and 20% thresholds suggested an

906 operational lower bound of approximately 15% under the tested conditions (**Fig.S16**a–c).
907 Crop-dependent spectral variability can increase omission errors when a single threshold is
908 applied. Because soybean residues generally exhibit lower MCRCSI values than maize and
909 rice residues, they are more likely to fall below the scene-level threshold. In S1, the global
910 AWCT threshold produced a soybean-residue omission rate of 10.28%, compared with 1.89%
911 using a crop-specific threshold **(Fig.S16**d–e**).** Temporal window selection further affects
912 comparability because tillage, burning, residue redistribution, and snow can rapidly modify
913 post-harvest surfaces (Zhou et al., 2024); therefore, observation periods should be aligned
914 with study objectives and regional crop phenology, particularly for large-area or annual
915 mapping. Sensor configuration remains an important constraint because MCRCSI relies on
916 SWIR information, limiting direct application to high-resolution platforms without suitable
917 SWIR bands, such as Jilin-1 and PlanetScope, although WorldView-3 provides an exception
918 (**Table S7**).

919 Future development should focus on improving robustness across residue-cover levels, crop
920 types, seasons, and sensors. Multi-temporal observations and fractional-cover retrieval may
921 improve characterization of sparse residues where binary classification becomes less reliable,
922 while crop-specific calibration and standardized monitoring windows could further reduce
923 regional and seasonal inconsistencies. Future sensors with improved SWIR spectral
924 resolution may better resolve lignocellulose absorption features and strengthen residue–soil
925 discrimination (Lamb et al., 2025). In particular, the planned Landsat Next (Landsat 10)
926 mission will provide finer spectral sampling in the SWIR domain (Roy et al., 2026), including
927 multiple narrower SWIR-2 bands, offering new opportunities for crop-residue
928 characterization and potentially extending residue monitoring toward consistent global
929 applications. In parallel, integrating complementary observations from Landsat, MODIS, and
930 SAR could improve temporal continuity while balancing spatial resolution and revisit
931 frequency, ultimately supporting more consistent residue monitoring across regions and
932 seasons.

6. Conclusions

Accurate and consistent monitoring of crop residue cover is essential for assessing conservation tillage practices and their biogeophysical impact on the Earth's climate system. However, large-scale mapping remains constrained by limited reference data, spectral confusion among crop residues, soil, and vegetation, and substantial spatial heterogeneity. To address these challenges, this study developed the Multi-band Crop Residue Cover Spectral Index (MCRCSI) to enhance residue signals while reducing background interference. MCRCSI combines cellulose- and lignin-absorption information in SWIR2 with complementary spectral information from the red, NIR, and SWIR1 bands. We further developed an Accuracy-Weighted Composite Thresholding (AWCT) that integrates Otsu, k-means, and GMM for adaptive residue-cover classification. Across eight representative sites in four major black soil regions, the MCRCSI–AWCT achieved a mean OA of 93.6% and a mean F1-score of 94.0% based on 15,591 independent validation samples. Agreement between Sentinel-2 and Landsat 8/9 (R^2 = 0.92-0.94) further supported the cross-sensor consistency of MCRCSI. Beyond binary residue mapping, the high sensitivity of MCRCSI to changes in residue cover may support crop residue fraction estimation and harvest event detection. Together, the framework provides a scalable basis for monitoring crop residue cover across regions and satellite sensors, with potential for near-real-time mapping. It may therefore support conservation tillage assessment, long-term monitoring of field management, and research on sustainable agriculture and climate mitigation.

CRediT authorship contribution statement

Yifeng Cui: Conceptualization, writing – review & editing. Nanshan You: Supervision, writing – review & editing. Jinwei Dong: Supervision, Conceptualization, writing – review & editing. Zhichao Li: writing – review & editing. Chao Zhang: writing – review & editing. Mohammad Reza Alizadeh: Writing – review & editing. Na Chen: Writing – review & editing. Ronggao Liu: Supervision, Conceptualization. César Terrer: Supervision.

Acknowledgments:

This study was supported by the National Key Research and Development Program of China (2023YFD1500200), the National Natural Science Foundation of China (42271375, 42461144212, 42525108, and 72221002).

Conflicts of Interest:

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The satellite datasets, including Sentinel-2, Landsat series and MODIS images are available in the data archive of Google Earth Engine (https://developers. google.com/earth-engine/datasets). The ground reference data used in this study is available upon reasonable request.

971 **References**

972 Azzari, G., Grassini, P., Edreira, J. I. R., Conley, S., Mourtzinis, S., and Lobell, D. B.: Satellite
973 mapping of tillage practices in the North Central US region from 2005 to 2016, Remote Sens. Environ.,
974 221, 417-429, 10.1016/j.rse.2018.11.010, 2019.
975 Bauer-Marschallinger, B. and Falkner, K.: Wasting petabytes: A survey of the Sentinel-2 UTM tiling
976 grid and its spatial overhead, ISPRS J. Photogramm. Remote Sens., 202, 682-690, 2023.
977 Bishop, C. M. and Nasrabadi, N. M.: Pattern recognition and machine learning, 4, Springer2006.
978 Boryan, C., Yang, Z., Mueller, R., and Craig, M.: Monitoring US agriculture: the US Department of
979 Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program, Geocarto Int.,
980 26, 341-358, 10.1080/10106049.2011.562309, 2011.
981 Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., and Smets, B.: Copernicus
982 global land cover layers—collection 2, Remote Sens., 12, 1044, 2020.
983 Cao, X., Chen, J., Matsushita, B., and Imura, H.: Developing a MODIS-based index to discriminate
984 dead fuel from photosynthetic vegetation and soil background in the Asian steppe area, Int. J. Remote
985 Sens., 31, 1589-1604, 2010.
986 Chen, B., Tu, Y., An, J., Wu, S., Lin, C., and Gong, P.: Quantification of losses in agriculture
987 production in eastern Ukraine due to the Russia-Ukraine war, Commun. Earth Environ., 5, 336,
988 10.1038/s43247-024-01488-3, 2024a.
989 Chen, H., Li, H., Liu, Z., Zhang, C., Zhang, S., and Atkinson, P. M.: A novel Greenness and Water
990 Content Composite Index (GWCCI) for soybean mapping from single remotely sensed multispectral
991 images, Remote Sens. Environ., 295, 113679, 10.1016/j.rse.2023.113679, 2023.
992 Chen, W., Shang, K., Wang, Y., Qi, W., Ding, S., and Zhang, X.: A mixed convolution and distance
993 covariance matrix network for fine classification of corn straw cover types with fused hyperspectral
994 and multispectral data, Int. J. Appl. Earth Obs. Geoinformation, 134, 104213, 2024b.
995 Chen, Z., Zhang, H., Zhang, M., Wu, Y., and Liu, Y.: Mangrove mapping in China using Gaussian
996 mixture model with a novel mangrove index (SSMI) derived from optical and SAR imagery, ISPRS
997 J. Photogramm. Remote Sens., 218, 466-486, 2024c.

998 Cui, Y., Dong, J., Zhang, C., You, N., Ren, S., Zhu, P., Duan, Q., Liu, R., and Terrer, C.: Mapping
999 Annual Crop Residue Cover Across Global Mollisol Regions at 10-m Resolution (2019–2024), Earth
1000 Syst. Sci. Data Discuss., 2026, 1-44, 2026.
1001 Cui, Y., Liu, R., Li, Z., Zhang, C., Song, X.-P., Yang, J., Yu, L., Chen, M., and Dong, J.: Decoding
1002 the inconsistency of six cropland maps in China, The Crop Journal, 12, 281-294, 2024.
1003 Daughtry, C., Hunt Jr, E., and McMurtrey Iii, J.: Assessing crop residue cover using shortwave
1004 infrared reflectance, Remote Sens. Environ., 90, 126-134, 2004.
1005 Daughtry, C. S., Doraiswamy, P., Hunt Jr, E., Stern, A., McMurtrey Iii, J., and Prueger, J.: Remote
1006 sensing of crop residue cover and soil tillage intensity, Soil Tillage Res., 91, 101-108, 2006.
1007 de Abelleyra, D., Iturralde Elortegui, M. d. R., Zelaya, K., Portillo, J., Melilli, M., Franzoni, A.,
1008 Volante, J., Goytia, Y., Kurtz, D., Ybarra, D., Barrios, R., Benedetti, P., Morales, C., Berriolo, J.,
1009 Lanceta Pereyra, M., Scavone, A., Murray, F., Santillan, J., Continelli, M., Gerlero, G., Salas, D.,
1010 Reinaldi, J., Lopez Juane, P., Gomez, D., Krapovickas, S., Sapino, V., Regonat, A., Cracogna, M.,
1011 Espíndola, C., Valiente, S., Parodi, M., Colombo, F., Scarel, J., Ayala, J., Martins, L., Basanta, M.,
1012 Rausch, A., Almada, G., Boero, L., Calcha, J., Chiavassa, A., Calandroni, M., Murillo, M., Pascale,
1013 B., Borrachi, S., Erreguerena, J., Besteiro, I., Oyesqui, L., Lazaeta, M., Pezzola, A., Winschel, C.,
1014 Rodriguez Pérez, J., German, A., and Viatale, J.: Argentina National Map of Crops 2023/2024,
1015 Zenodo [dataset], 10.5281/zenodo.13984185, 2024.
1016 Demattê, J. A., Poppiel, R. R., Novais, J. J. M., Rosin, N. A., Minasny, B., Savin, I. Y., Grunwald,
1017 S., Chen, S., Hong, Y., and Huang, J.: Frontiers in earth observation for global soil properties
1018 assessment linked to environmental and socio-economic factors, The Innovation, 6, 100985, 2025.
1019 Deshpande, M. V., Pillai, D., Krishna, V. V., and Jain, M.: Detecting and quantifying zero tillage
1020 technology adoption in Indian smallholder systems using Sentinel-2 multi-spectral imagery, Int. J.
1021 Appl. Earth Obs. Geoinf., 128, 103779, 2024.
1022 Di, Y., Dong, J., You, N., Li, Z., Moreno-Martínez, Á., Izquierdo-Verdiguier, E., Sun, J., and Fu, P.:
1023 30 m-resolution annual crop type maps in Northeast China from 2001 to 2022, Sci. Data, 2026.
1024 Diek, S., Fornallaz, F., Schaepman, M. E., and De Jong, R.: Barest pixel composite for agricultural
1025 areas using landsat time series, Remote Sens., 9, 1245, 2017.

1026 Dong, Y., Xuan, F., Li, Z., Su, W., Guo, H., Huang, X., Li, X., and Huang, J.: Modeling the corn
1027 residue coverage after harvesting and before sowing in northeast china by random forest and soil
1028 texture zoning, Remote Sens., 15, 2179, 2023.
1029 Dong, Y., Xuan, F., Huang, X., Li, Z., Su, W., Huang, J., Li, X., Tao, W., Liu, H., and Chen, J.: A
1030 30-m annual corn residue coverage dataset from 2013 to 2021 in Northeast China, Sci. Data, 11, 216,
1031 10.1038/s41597-024-02998-7, 2024.
1032 Feng, L., Wang, Y., Fensholt, R., Tong, X., Tagesson, T., Zhang, X., Ardö, J., Zhou, J., Shao, W.,
1033 and Dou, Y.: Globally increased cropland soil exposure to climate extremes in recent decades, Nat.
1034 Commun., 16, 4354, 2025.
1035 Fontana, A., Labaz, B., Monger, C., Morras, H., Wang, L., Pfeiffer, M., Taboada, M. A., and Valle,
1036 S.: Status and challenges of black soils, FAO2022.
1037 Gao, L., Zhang, C., Yun, W., Ji, W., Ma, J., Wang, H., Li, C., and Zhu, D.: Mapping crop residue
1038 cover using Adjust Normalized Difference Residue Index based on Sentinel-2 MSI data, Soil Tillage
1039 Res., 220, 105374, 2022.
1040 Gelder, B., Kaleita, A., and Cruse, R.: Estimating mean field residue cover on midwestern soils using
1041 satellite imagery, Agron. J., 101, 635-643, 2009.
1042 Heiskanen, J.: Estimating aboveground tree biomass and leaf area index in a mountain birch forest
1043 using ASTER satellite data, Int. J. Remote Sens., 27, 1135-1158, 2006.
1044 Hively, W. D., Lamb, B. T., Daughtry, C. S., Shermeyer, J., McCarty, G. W., and Quemada, M.:
1045 Mapping crop residue and tillage intensity using WorldView-3 satellite shortwave infrared residue
1046 indices, Remote Sens., 10, 1657, 2018.
1047 Hively, W. D., Shermeyer, J., Lamb, B. T., Daughtry, C. T., Quemada, M., and Keppler, J.: Mapping
1048 crop residue by combining Landsat and WorldView-3 satellite imagery, Remote Sens., 11, 1857,
1049 2019.
1050 Hively, W. D., Lamb, B. T., Daughtry, C. S. T., Serbin, G., Dennison, P. E., Kokaly, R. F., Wu, Z.,
1051 and Masek, J. G.: Evaluation of SWIR Crop Residue Bands for the Landsat Next Mission, Remote
1052 Sens., 13, 3718, 2021.

1053 Jia, A., Wang, D., Peng, J., Ma, Z., and Liang, S.: Dynamic parameterization of global land surface
1054 albedo components: Bare soil, non-photosynthetic vegetation, and photosynthetic vegetation, Remote
1055 Sens. Environ., 329, 114943, 2025.
1056 Kassam, A., Friedrich, T., and Derpsch, R.: Global spread of conservation agriculture, Int. J. Environ.
1057 Stud., 76, 29-51, 10.1080/00207233.2018.1494927, 2019.
1058 Kaufman, Y. J. and Remer, L. A.: Detection of forests using mid-IR reflectance: An application for
1059 aerosol studies, IEEE Trans. Geosci. Remote Sens., 32, 672-683, 1994.
1060 Khan, A., Potapov, P., Turubanova, S., Hansen, M. C., Tyukavina, A., Poulson, A., Hernandez-Serna,
1061 A., Song, X.-P., Harris, N., and Stolle, F.: Global annual cropland dynamics 2015–2024, Remote
1062 Sens. Environ., 340, 115438, 2026.
1063 Kravchenko, Y., Balayev, A., Mazăre, V., Zhang, X., Liu, X., and Herbert, S. J.: Conservation
1064 Practices on Ukrainian Mollisols : A Mini Review, Int. J. Energy Environ., 12, 2018.
1065 Lamb, B. T., Hively, W. D., Jennewein, J., Thieme, A., Soroka, A. M., Santos, L., Jones, D., and
1066 Mirsky, S.: Multiyear crop residue cover mapping using narrow-band vs. broad-band shortwave
1067 infrared satellite imagery, Soil Tillage Res., 251, 106524, 2025.
1068 Lamb, B. T., Dennison, P. E., Hively, W. D., Kokaly, R. F., Serbin, G., Wu, Z., Dabney, P. W., Masek,
1069 J. G., Campbell, M. J., and Daughtry, C. S. T.: Optimizing Landsat Next Shortwave Infrared Bands
1070 for Crop Residue Characterization, Remote Sens., 14, 6128, 2022.
1071 Lambert, A., Hallar, A. G., Garcia, M., Strong, C., Andrews, E., and Hand, J. L.: Dust impacts of
1072 rapid agricultural expansion on the Great Plains, Geophys. Res. Lett., 47, e2020GL090347, 2020.
1073 Li, H., Huang, J., Zhang, C., Ning, X., Zhang, S., and Atkinson, P. M.: An efficient and generalisable
1074 approach for mapping paddy rice fields based on their unique spectra during the transplanting period
1075 leveraging the CIE colour space, Remote Sens. Environ., 313, 114381, 2024.
1076 Li, J. and Roy, D. P.: A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit
1077 Intervals and Implications for Terrestrial Monitoring, Remote. Sens., 9, 902, 2017.
1078 Li, Z., Shen, H., Weng, Q., Zhang, Y., Dou, P., and Zhang, L.: Cloud and cloud shadow detection for
1079 optical satellite imagery: Features, algorithms, validation, and prospects, ISPRS J. Photogramm.
1080 Remote Sens., 188, 89-108, 2022.

Meng, X., Bao, Y., Zhang, X., Luo, C., and Liu, H.: A long-term global Mollisols SOC content prediction framework: Integrating prior knowledge, geographical partitioning, and deep learning models with spatio-temporal validation, Remote Sens. Environ., 318, 114592, 2025.

Ohno, Y.: CIE fundamentals for color measurements, NIP & Digital Fabrication Conference, 540-545,

Pacheco, A. and McNairn, H.: Evaluating multispectral remote sensing and spectral unmixing analysis for crop residue mapping, Remote Sens. Environ., 114, 2219-2228, 2010.

Patra, S., Shekher, M., Solanki, S., Ramachandran, R., and Krishnan, R.: A technique for generating natural colour images from false colour composite images, Int. J. Remote Sens., 27, 2977-2989, 2006.

Qi, J., Marsett, R., Heilman, P., Bieden-bender, S., Moran, S., Goodrich, D., and Weltz, M.: RANGES improves satellite-based information and land cover assessments in southwest United States, Eos, Transactions American Geophysical Union, 83, 601-606, 2002.

Quemada, M. and Daughtry, C. S.: Spectral indices to improve crop residue cover estimation under varying moisture conditions, Remote Sens., 8, 660, 2016.

Rouse Jr, J. W., Haas, R. H., Deering, D., Schell, J., and Harlan, J. C.: Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation, 1974.

Roy, D. P., Wulder, M. A., Gorelick, N., Hansen, M., Healey, S., Hostert, P., Huntington, J., Radeloff, V. C., Scambos, T., and Schaaf, C.: The next Landsat: Mission turning point?, Remote Sens. Environ., 332, 115087, 2026.

Shojaeezadeh, S. A., Al-Wardy, M., Nikoo, M. R., Mooselu, M. G., Alizadeh, M. R., Adamowski, J. F., Moradkhani, H., Alamdari, N., and Gandomi, A. H.: Soil erosion in the United States: Present and future (2020–2050), Catena, 242, 108074, 2024.

Simelton, E., Carew-Reid, J., Coulier, M., Damen, B., Howell, J., Pottinger-Glass, C., Tran, H. V., and Van Der Meiren, M.: NBS framework for agricultural landscapes, Front. Environ. Sci., 9, 678367, 2021.

Smerald, A., Rahimi, J., and Scheer, C.: A global dataset for the production and usage of cereal residues in the period 1997-2021, Sci. Data, 10, 685, 10.1038/s41597-023-02587-0, 2023.

Smith, P., Calvin, K. V., Nkem, J. N., Campbell, D., Cherubini, F., Grassi, G., Korotkov, V. N., Le Hoang, A., Lwasa, S., McElwee, P., Nkonya, E., Saigusa, N., Soussana, J.-F., Taboada, M. A.,

Manning, F. C., Nampanzira, D. K., Arias-Navarro, C., Vizzarri, M., House, J. I., Roe, S., Cowie, A. L., Rounsevell, M. D. A., and Arneth, A.: Which practices co-deliver food security, climate change mitigation and adaptation, and combat land degradation and desertification?, Global Change Biol., 26, 1532 - 1575, 2019.

Tao, W., Dong, Y., Su, W., Li, J., Xuan, F., Huang, J., Yang, J., Li, X., Zeng, Y., and Li, B.: Mapping the Corn Residue-Covered Types Using Multi-Scale Feature Fusion and Supervised Learning Method by Chinese GF-2 PMS Image, Front. Plant Sci., 13, 901042, 2022.

Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C., Brakenridge, G. R., Erickson, T. A., and Slayback, D. A.: Satellite imaging reveals increased proportion of population exposed to floods, Nature, 596, 80-86, 2021.

Tong, Y., Angelini, M. E., Yigini, Y., and Luotto, I.: Global black soil distribution, Front. Agric. Sci. Eng., 11, 271-281, 10.15302/J-Fase-2024567, 2024.

Van Deventer, A., Ward, A., Gowda, P., and Lyon, J.: Using thematic mapper data to identify contrasting soil plains and tillage practices, Photogramm. Eng. Remote Sens., 63, 87-93, 1997.

Wang, S., Azzari, G., and Lobell, D. B.: Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques, Remote Sens. Environ., 222, 303-317, 2019.

Wang, S., Di Tommaso, S., Deines, J. M., and Lobell, D. B.: Mapping twenty years of corn and soybean across the US Midwest using the Landsat archive, Sci. Data, 7, 307, 10.1038/s41597-020-00646-4, 2020.

Wu, X., Zhou, Q., Guan, K., Wang, S., Hipple, J., Peng, B., Chen, Z., and Qin, R.: A framework to detect tillage practices from space: A demonstration in the US Midwest, Remote Sens. Environ., 337, 115323, 10.1016/j.rse.2026.115323, 2026.

Xu, M., Liu, R., Chen, J. M., Liu, Y., Shang, R., Ju, W., Wu, C., and Huang, W.: Retrieving leaf chlorophyll content using a matrix-based vegetation index combination approach, Remote Sens. Environ., 224, 60-73, 2019.

Xu, S., Zhu, X., Chen, J., Zhu, X., Duan, M., Qiu, B., Wan, L., Tan, X., Xu, Y. N., and Cao, R.: A robust index to extract paddy fields in cloudy regions from SAR time series, Remote Sens. Environ., 285, 113374, 10.1016/j.rse.2022.113374, 2023.

Yang, Y., Tilman, D., Jin, Z., Smith, P., Barrett, C. B., Zhu, Y.-G., Burney, J., D'Odorico, P., Fantke, P., Fargione, J., Finlay, J. C., Rulli, M. C., Sloat, L., Jan van Groenigen, K., West, P. C., Ziska, L., Michalak, A. M., Team, t. C.-A., Lobell, D. B., Clark, M., Colquhoun, J., Garg, T., Garrett, K. A., Geels, C., Hernandez, R. R., Herrero, M., Hutchison, W. D., Jain, M., Jungers, J. M., Liu, B., Mueller, N. D., Ortiz-Bobea, A., Schewe, J., Song, J., Verheyen, J., Vitousek, P., Wada, Y., Xia, L., Zhang, X., and Zhuang, M.: Climate change exacerbates the environmental impacts of agriculture, Science, 385, eadn3747, doi:10.1126/science.adn3747, 2024.

You, N., Dong, J., Huang, J., Du, G., Zhang, G., He, Y., Yang, T., Di, Y., and Xiao, X.: The 10-m crop type maps in Northeast China during 2017-2019, Sci. Data, 8, 41, 10.1038/s41597-021-00827-9, 2021.

Yue, J., Tian, Q., Dong, X., and Xu, N.: Using broadband crop residue angle index to estimate the fractional cover of vegetation, crop residue, and bare soil in cropland systems, Remote Sens. Environ., 237, 111538, 2020.

Yue, J., Tian, Q., Tang, S., Xu, K., and Zhou, C.: A dynamic soil endmember spectrum selection approach for soil and crop residue linear spectral unmixing analysis, Int. J. Appl. Earth Obs. Geoinf., 78, 306-317, 2019.

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N.-E., Ramoino, F., and Arino, O.: ESA WorldCover 10 m 2020 v100, 10.5281/zenodo.5571936, 2021.

Zhang, H. K., Roy, D. P., Yan, L., Li, Z., Huang, H., Vermote, E., Skakun, S., and Roger, J.-C.: Characterization of Sentinel-2A and Landsat-8 top of atmosphere, surface, and nadir BRDF adjusted reflectance and NDVI differences, Remote Sens. Environ., 215, 482-494, 10.1016/j.rse.2018.04.031, 2018.

Zhang, W., Li, W., Wang, C., Yu, Q., Tang, H., and Wu, W.: A novel index for mapping crop residue covered cropland using remote sensing data, Comput. Electron. Agric., 231, 109995, 10.1016/j.compag.2025.109995, 2025.

Zhao, X., Li, R.-C., Liu, W.-X., Liu, W.-S., Xue, Y.-H., Sun, R.-H., Wei, Y.-X., Chen, Z., Lal, R., Dang, Y. P., Xu, Z.-Y., and Zhang, H.-L.: Estimation of crop residue production and its contribution

1167 to carbon neutrality in China, Resources Conservation and Recycling, 203, 107450,
1168 10.1016/j.resconrec.2024.107450, 2024.
1169 Zhong, L., Hu, L., Yu, L., Gong, P., and Biging, G. S.: Automated mapping of soybean and corn
1170 using phenology, ISPRS J. Photogramm. Remote Sens., 119, 151-164, 10/gd8dqd, 2016.
1171 Zhou, D., Li, M., Li, Y., Qi, J., Liu, K., Cong, X., and Tian, X.: Detection of ground straw coverage
1172 under conservation tillage based on deep learning, Comput. Electron. Agric., 172, 105369, 2020.
1173 Zhou, Q., Guan, K., Wang, S., Hipple, J., and Chen, Z.: From satellite-based phenological metrics to
1174 crop planting dates: Deriving field-level planting dates for corn and soybean in the US Midwest,
1175 ISPRS J. Photogramm. Remote Sens., 216, 259-273, 2024.
1176 Zhou, Y., Ferdinand, M. S., van Wesemael, J., Dvorakova, K., Baret, P. V., Van Oost, K., and van
1177 Wesemael, B.: A framework for mapping conservation agricultural fields using optical and radar time
1178 series imagery, Remote Sens. Environ., 328, 114858, 10.1016/j.rse.2025.114858, 2025.

Supplementary Materials

Mapping crop residue cover in global black soil regions based on a new multi-band spectral index and an integrated thresholding approach

Yifeng Cui[1,2,5†], Nanshan You[1†], Jinwei Dong[1,5*], Zhichao Li[1,5], Chao Zhang[3], Mohammad Reza Alizadeh[2], Na Chen[4], César Terrer[2], Ronggao Liu[1*]

[1] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China.

[2] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

[3] Department of Civil and Environmental Engineering, National University of Singapore, Singapore, 117576, Singapore.

[4] Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

[5] University of Chinese Academy of Sciences, Beijing, 100049, China.

† These authors contributed equally to this work.

* Corresponding authors: dongjw@igsnrr.ac.cn (J.D.); liurg@igsnrr.ac.cn (R.L.)

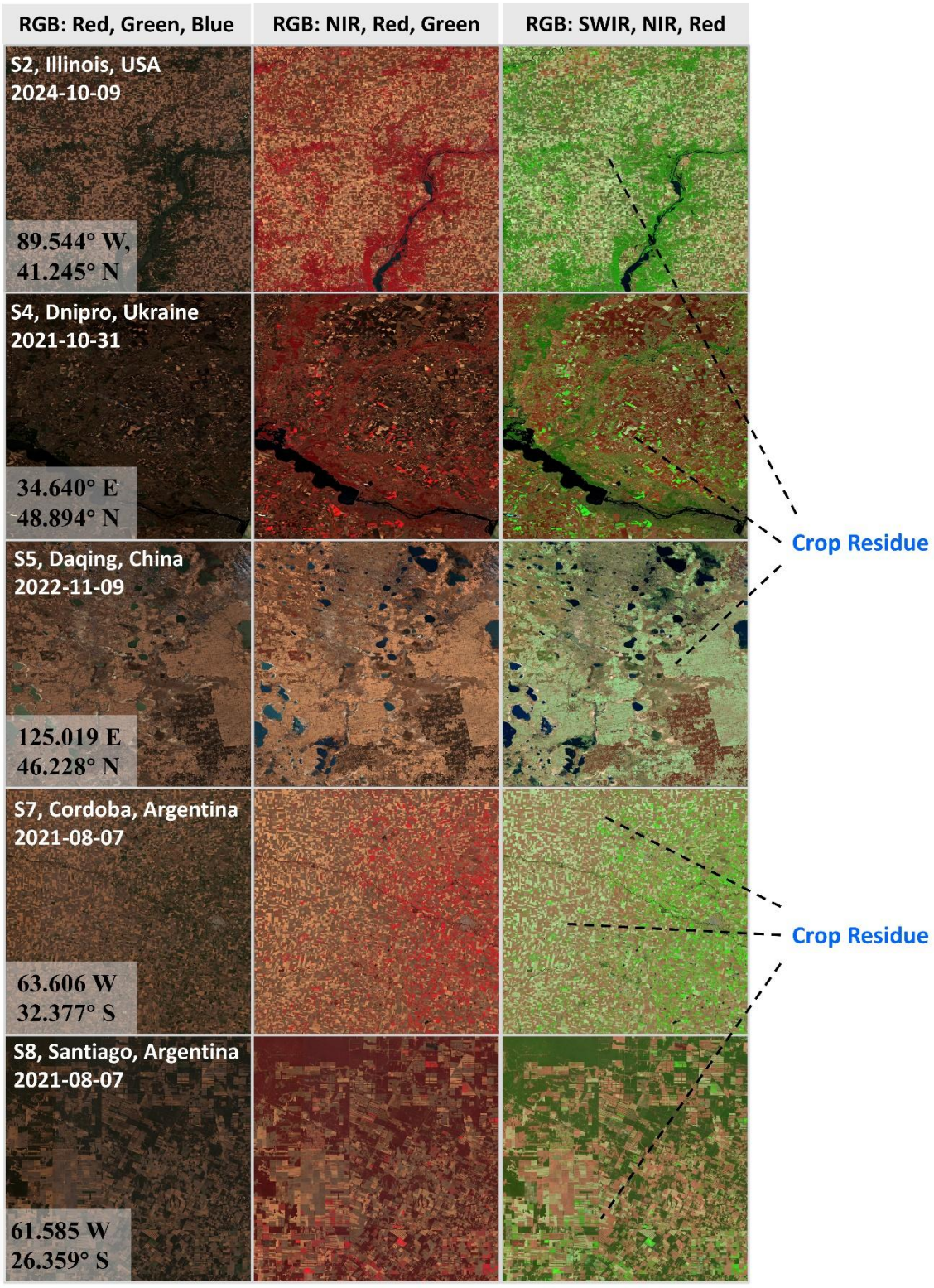
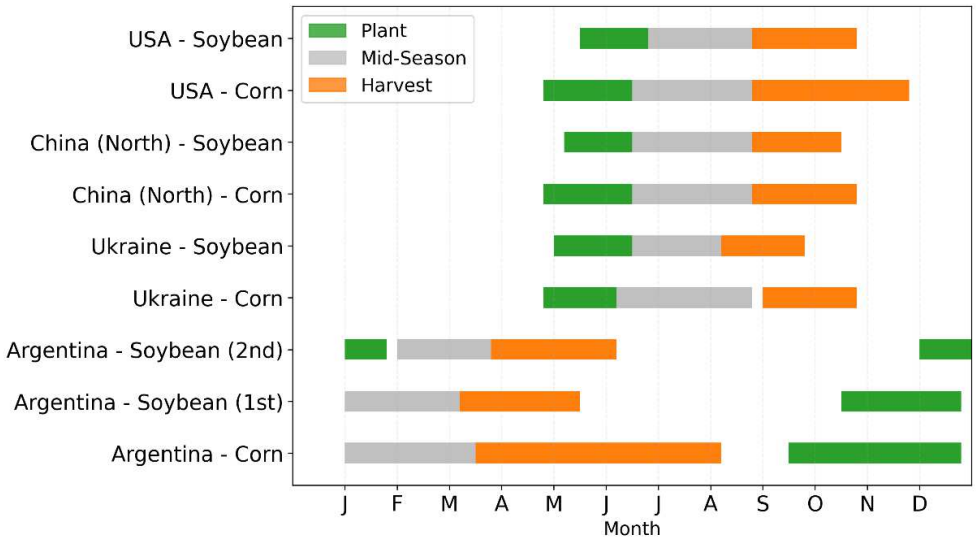


Fig.S1 Interpretation of residue cover samples on Sentinel-2 false-color composite imagery.



1201

1202 **Fig.S2** Crop calendar of major crops across global black soil region

1203

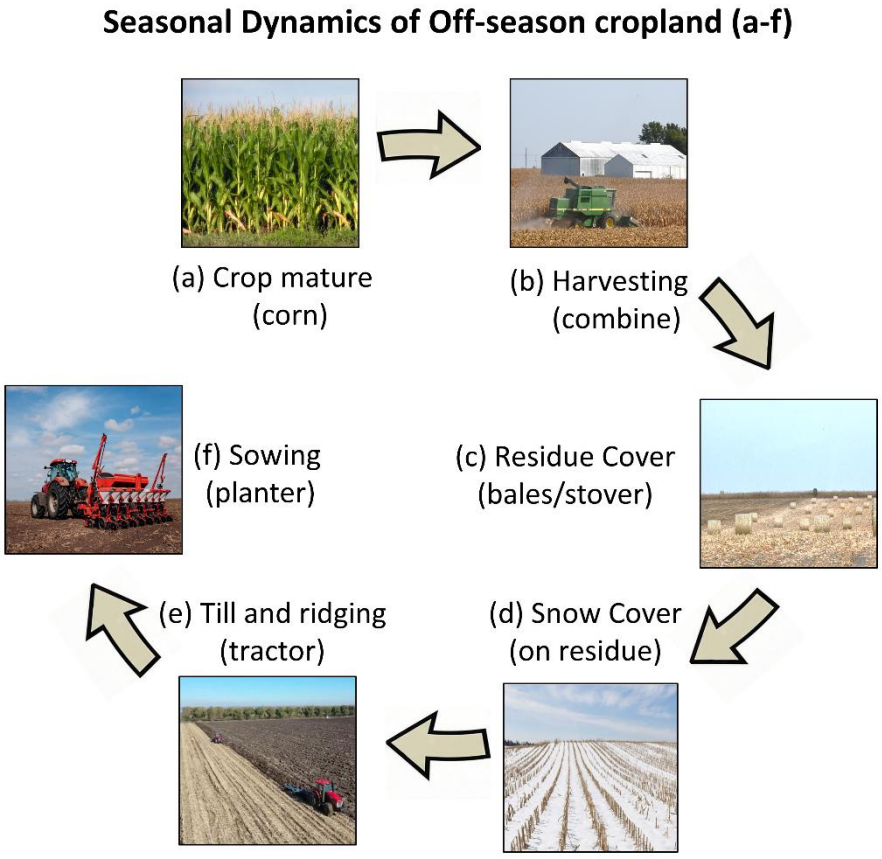


Fig.S3 Seasonal transitions of off-season cropland. The panel (a–f) illustrates the temporal transition of agricultural surfaces, defining the off-season (non-growing season) as the period spanning from crop maturity (a) and harvesting (b), through residue cover (c) and snow accumulation (d), to tillage (e) and the sowing (f) of the subsequent crop.

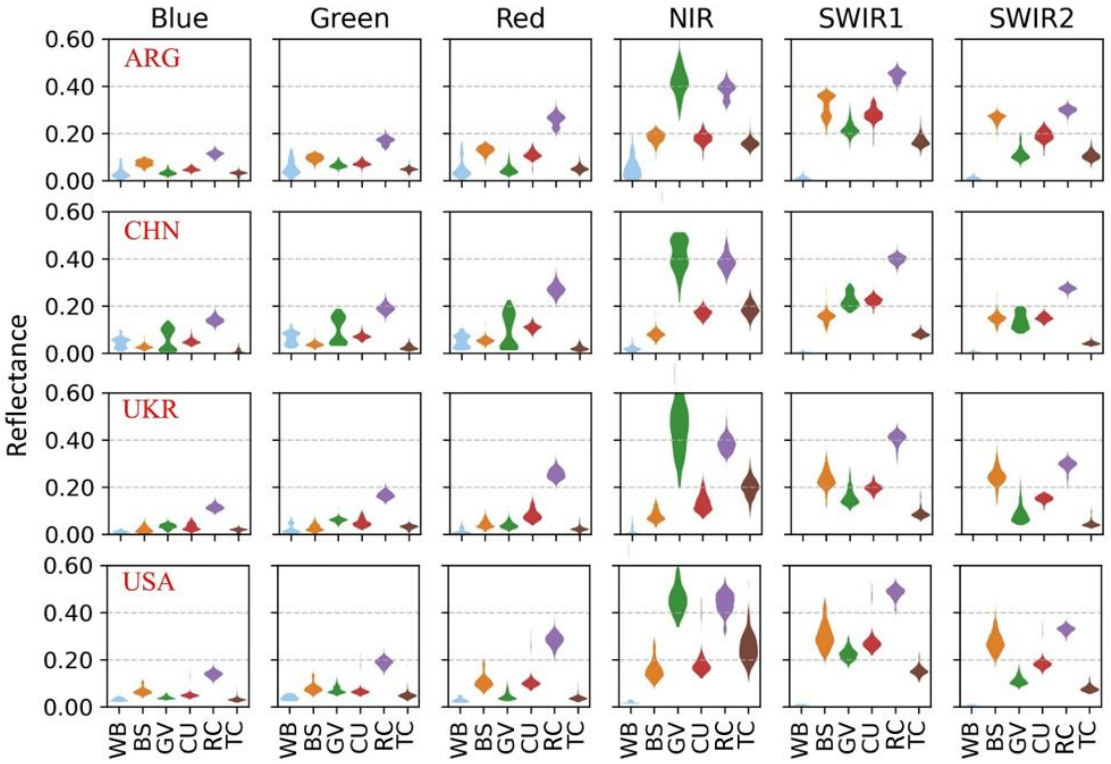


Fig.S4 Violin plots of band reflectance for land cover types in black soil regions. Abbreviations: WB – Water Body; BS – Bare Soil; GV – Green Vegetation; CU – Crop Unharvested (mature but unharvested maize); RC– Residue Cover; TC – Tree Cover.

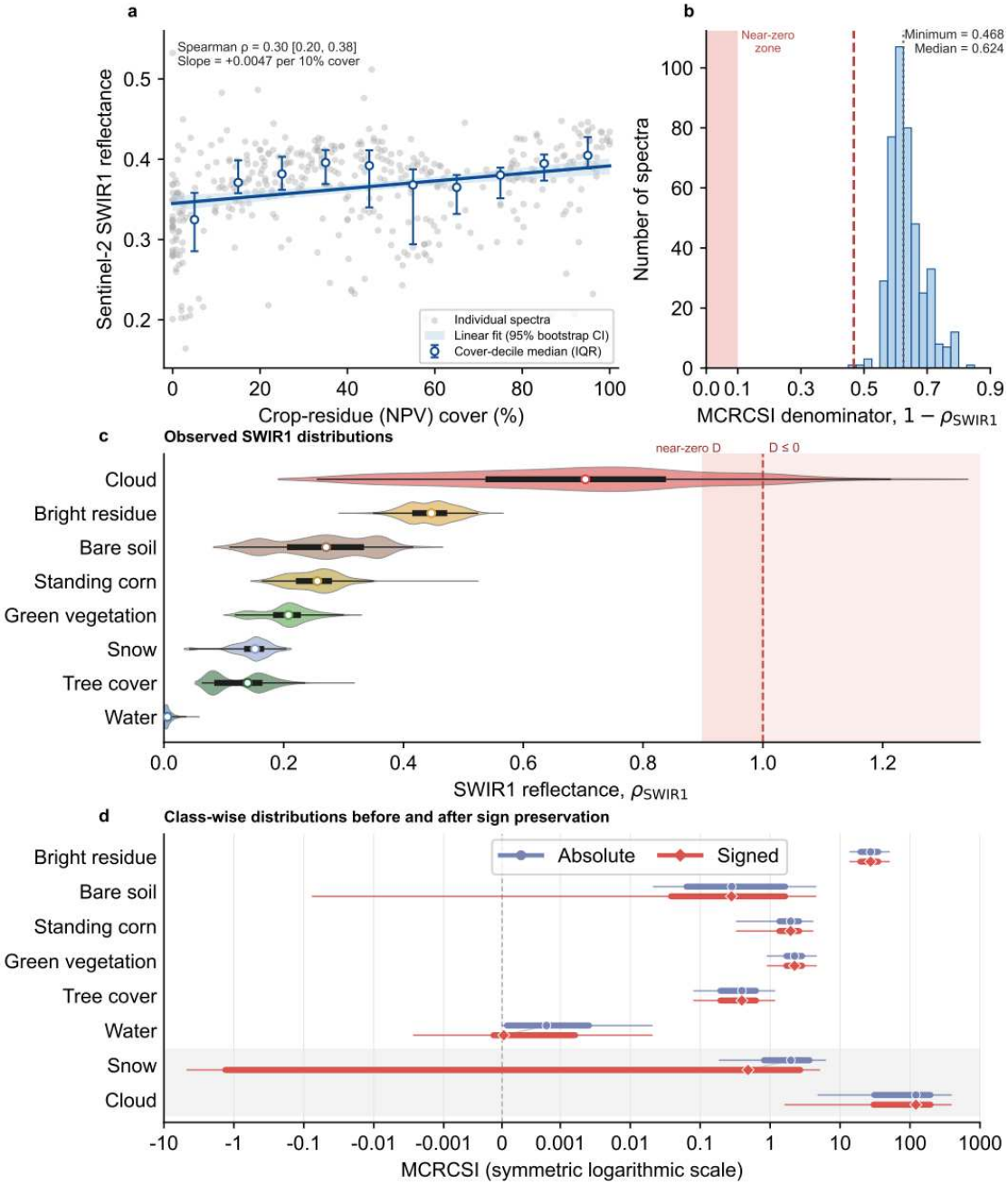


Fig.S5 Numerical stability and sign sensitivity of MCRCSI across residue-cover conditions and major land-cover classes. (a) Relationship between SWIR1 reflectance (ρ_{SWIR1}) and crop-residue (NPV) cover for 432 spectra. Gray points represent individual spectra, the blue line denotes the linear fit with the 95% bootstrap confidence interval, and open circles with error bars indicate cover-decile medians and interquartile ranges (IQRs). (b) Distribution of the MCRCSI denominator, $(1-\rho_{\mathrm{SWIR1}})$, for the same residue spectra. The shaded region denotes the near-zero denominator zone, showing that all observed values remained well above this range. (c) Observed SWIR1 reflectance distributions across major land-cover classes. The shaded region indicates near-zero denominator conditions, and the dashed line marks $D = 0$ ($\rho_{\mathrm{SWIR1}} = 1$); potentially unstable values occurred primarily in residual cloud observations. (d) Class-wise distributions of absolute and sign-preserving MCRCSI, shown on a symmetric logarithmic scale. Comparison of the two formulations

demonstrates that sign preservation mainly affects extreme or anomalous observations, particularly snow and cloud, whereas the distributions of major non-cloud surface classes remain broadly consistent.

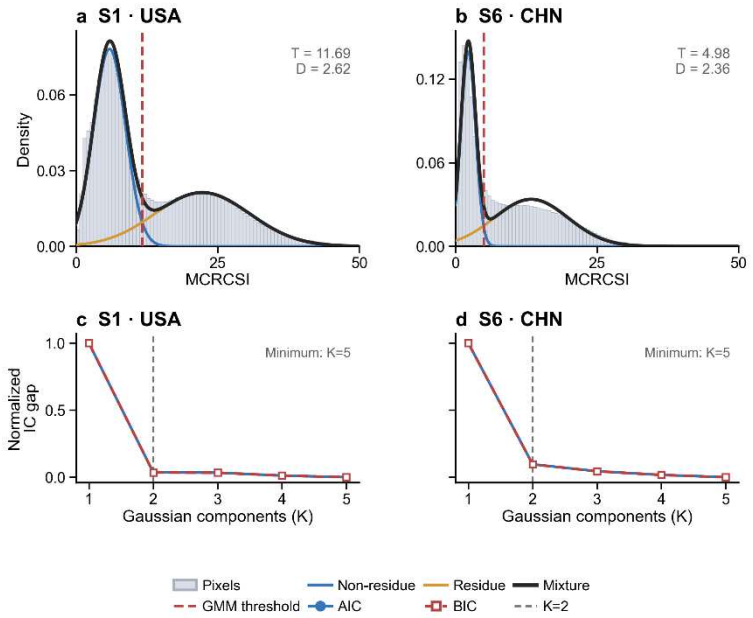
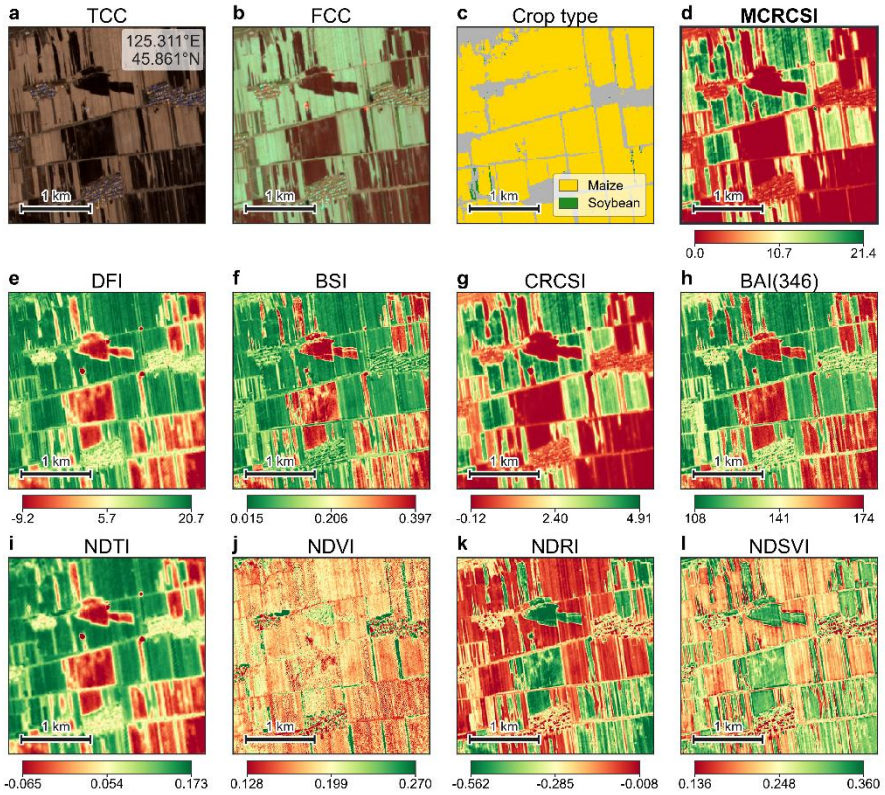


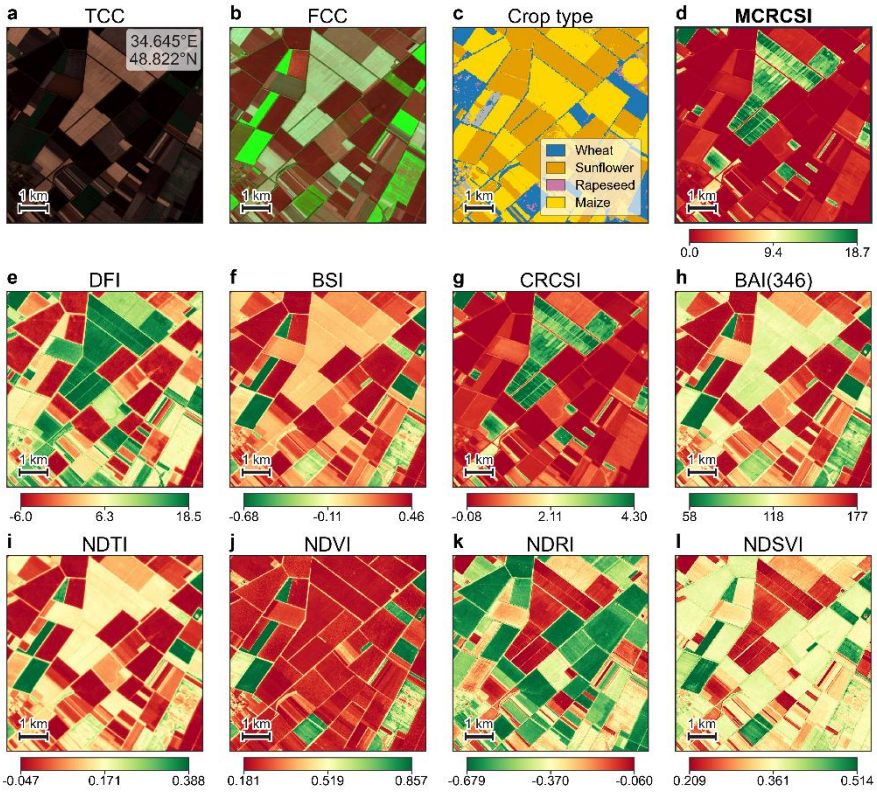
Fig.S6 Evaluation of the two-component GMM assumption for MCRCSI thresholding. (a–b) Representative MCRCSI distributions and two-component GMM fits for S1 (USA) and S6 (China), with dashed lines indicating the derived thresholds. (c–d) Normalized AIC and BIC across 1–5 Gaussian components, showing a sharp improvement from K=1 to K=2 and only marginal changes with additional components.



1239

1240 **Fig.S7 Visual comparison of tillage indices in study area S5 (Daqing, China).**

1241



1242

1243 **Fig.S8 Visual comparison of tillage indices in study area S3 (Poltava, Ukraine).**

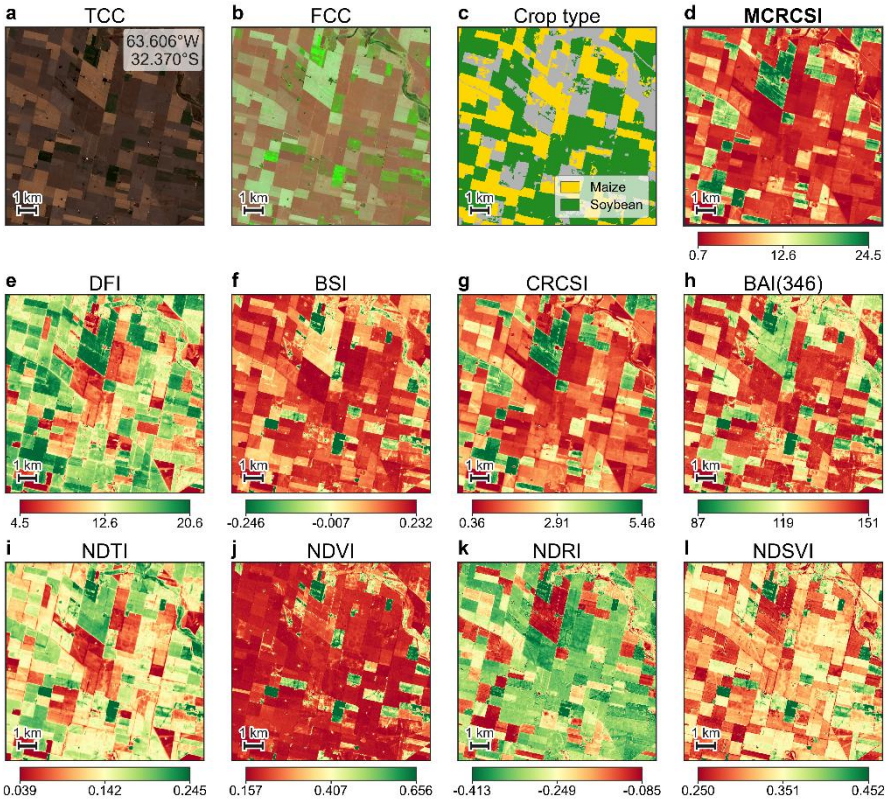


Fig.S9 Visual comparison of tillage indices in study area S7 (Cordoba, Argentina).

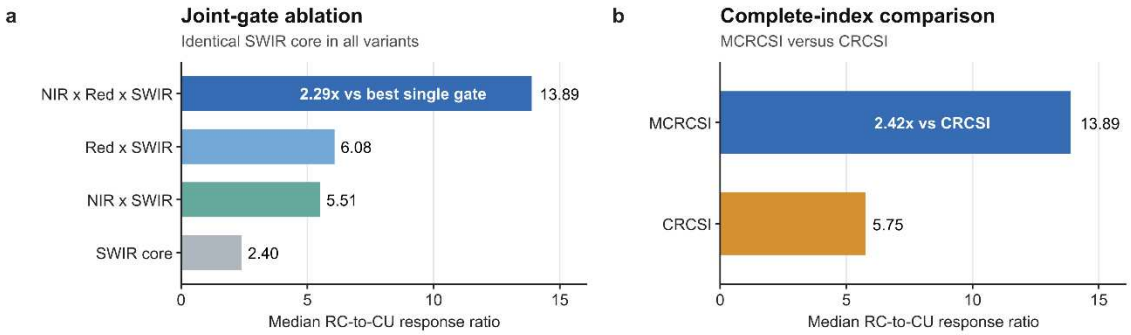


Fig.S10 Joint NIR × Red gating enhances the residue-to-unharvested-crop response margin. (a) Fixed-SWIR-core comparison experiment quantifying the contributions of NIR × Red gating. The median Residue Cover (RC)-to-Crop Unharvested (CU) response ratio increased from 2.40 for the SWIR core alone to 5.51 and 6.08 with the NIR and Red single-band gates, respectively. Applying the joint NIR × Red gate increased the ratio to 13.89, 2.29-fold higher than the better-performing single-band gate (Red × SWIR). (b) Complete-index comparison showing that MCRCSI achieved a median RC-to-CU response ratio of 13.89, compared with 5.75 for CRCSI, corresponding to a 2.42-fold increase in the typical response margin. For each index or ablation variant, the RC-to-CU response ratio was calculated as the median index value of RC samples divided by the median index value of CU samples; larger ratios indicate greater relative separation between the two classes.

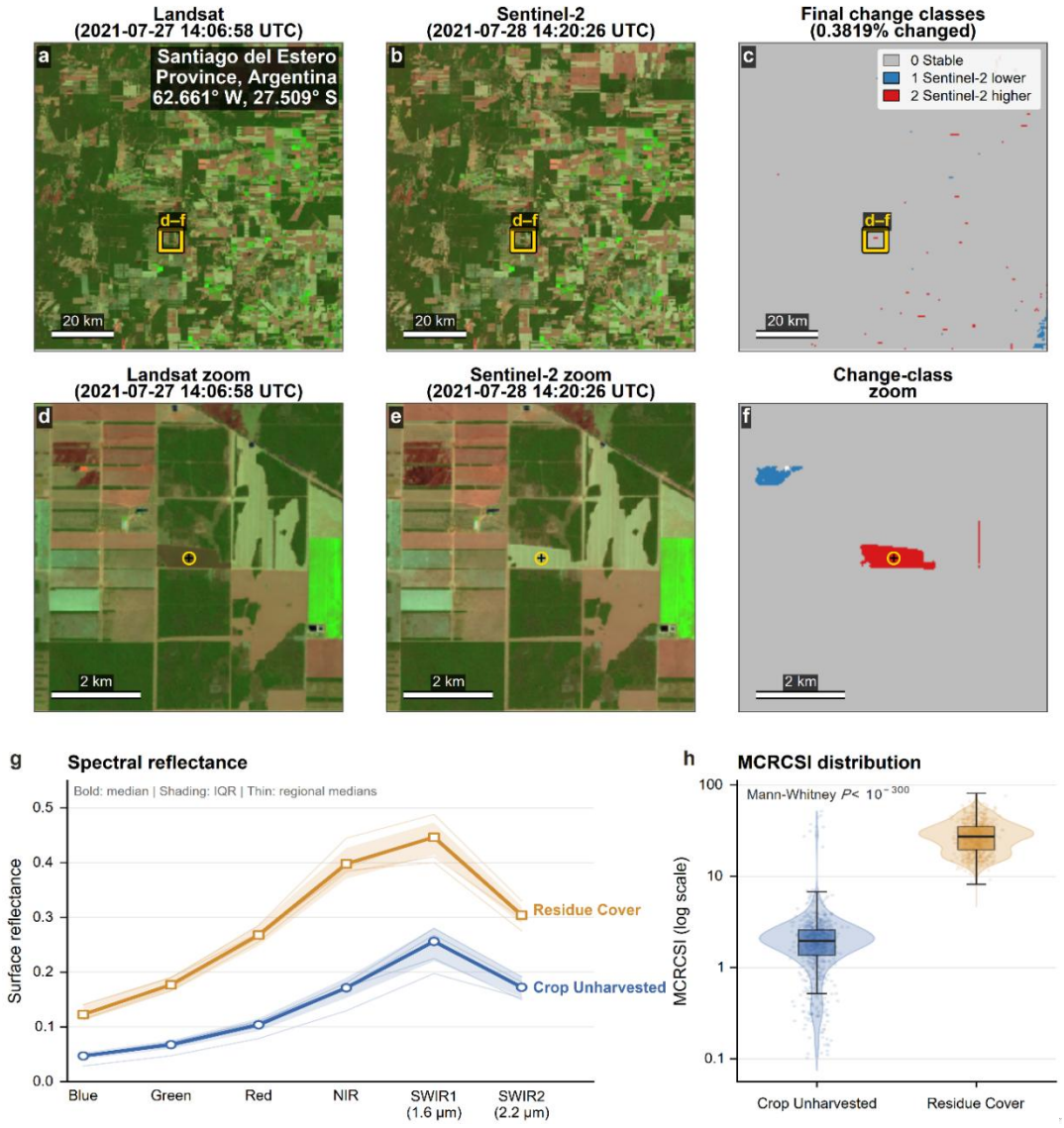


Fig.S11 Detection of short-term land-surface changes from near-contemporaneous Landsat and Sentinel-2 observations in the Argentina test region and spectral separability between crop unharvested (CU) and residue cover (RC). (a-b) Landsat image acquired on 27 July 2021 and Sentinel-2 image acquired on 28 July 2021, respectively, centered near 62.661° W, 27.509° S. (c) MCRCSI-based change classes between the two observations, with gray indicating stable pixels, blue lower MCRCSI in Sentinel-2, and red higher MCRCSI in Sentinel-2. (d-e) Enlarged views of the areas outlined by the yellow boxes in (a) and (b), and (f) the corresponding change-class map, highlighting representative burned and harvested fields. Burning resulted in decreased MCRCSI, whereas harvesting produced increased MCRCSI associated with the emergence of surface crop residues. (g) Sentinel-2 reflectance spectra of CU and RC, with bold lines indicating overall medians, shading the interquartile ranges, and thin lines regional medians. (h) MCRCSI distributions on a logarithmic scale, showing marked separation between CU and RC (Mann–Whitney U test, $P < 10^{-300}$).

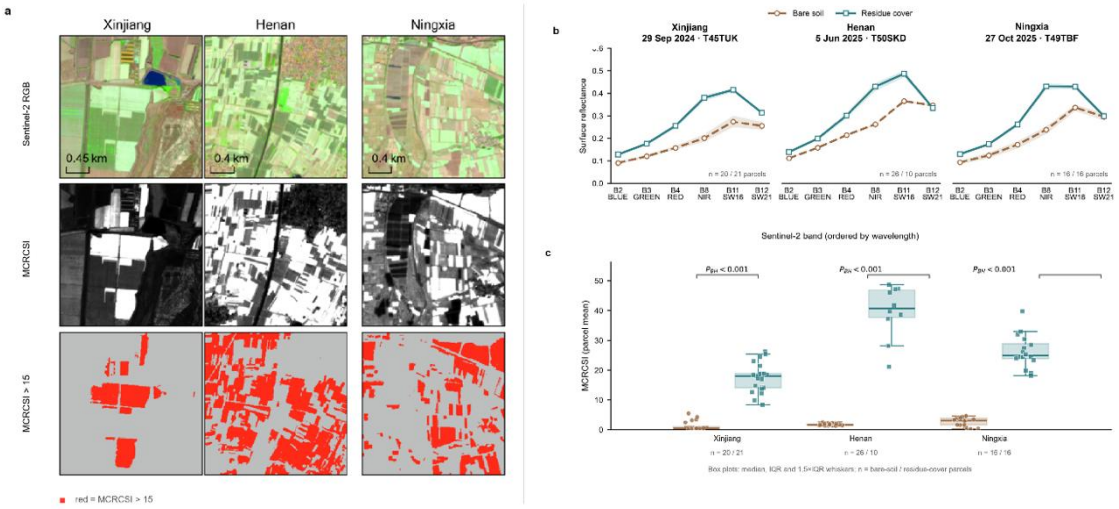


Fig.S12 Transferability of MCRCSI across three non-black soil agricultural regions. (a) Sentinel-2 imagery, MCRCSI maps, and residue-cover delineation in Xinjiang, Henan, and Ningxia. (b) Mean Sentinel-2 spectral profiles of bare soil and residue-covered parcels. (c) Comparison of parcel-level MCRCSI values between bare soil and residue cover, showing significant separation in all three regions ($P < 0.001$).

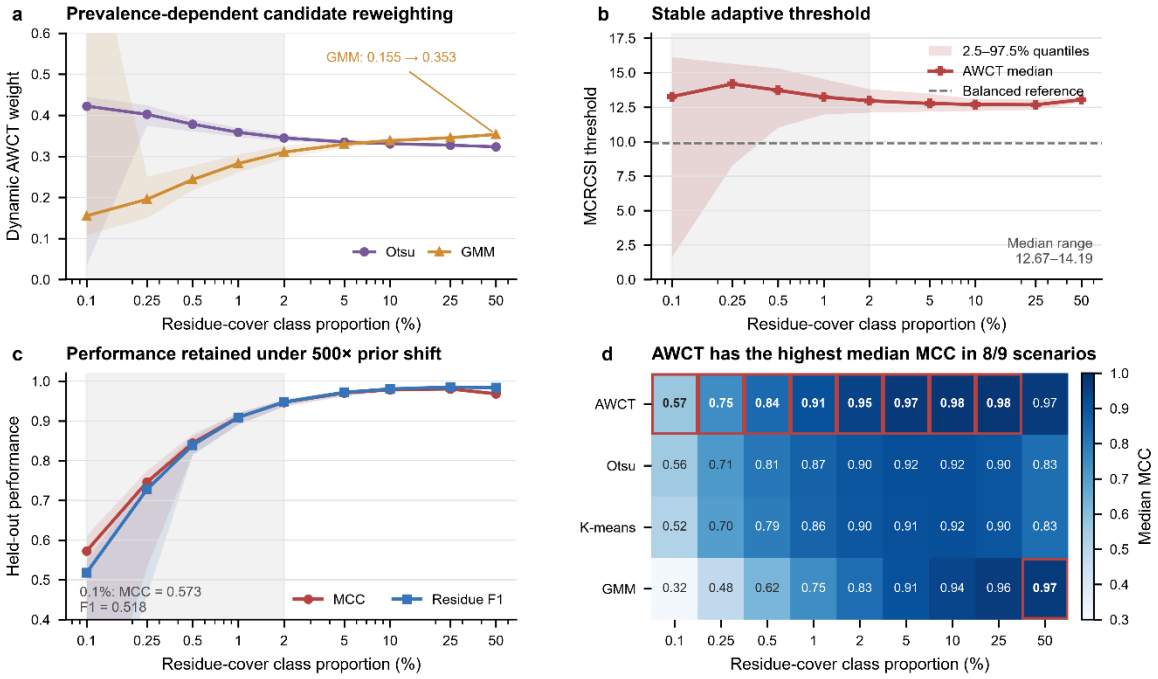


Fig.S13 Adaptive behavior of AWCT under extreme class imbalance. (a) Prevalence-dependent reweighting of Otsu and GMM candidate thresholds; (b) resulting AWCT threshold stability across residue fractions of 0.1–50%; (c) holdout MCC and residue F1 performance; and (d) median MCC comparison among AWCT and individual thresholding methods. Each scenario used 5,000 samples and 100 Monte Carlo repetitions; shaded bands indicate the 2.5–97.5% empirical quantiles.

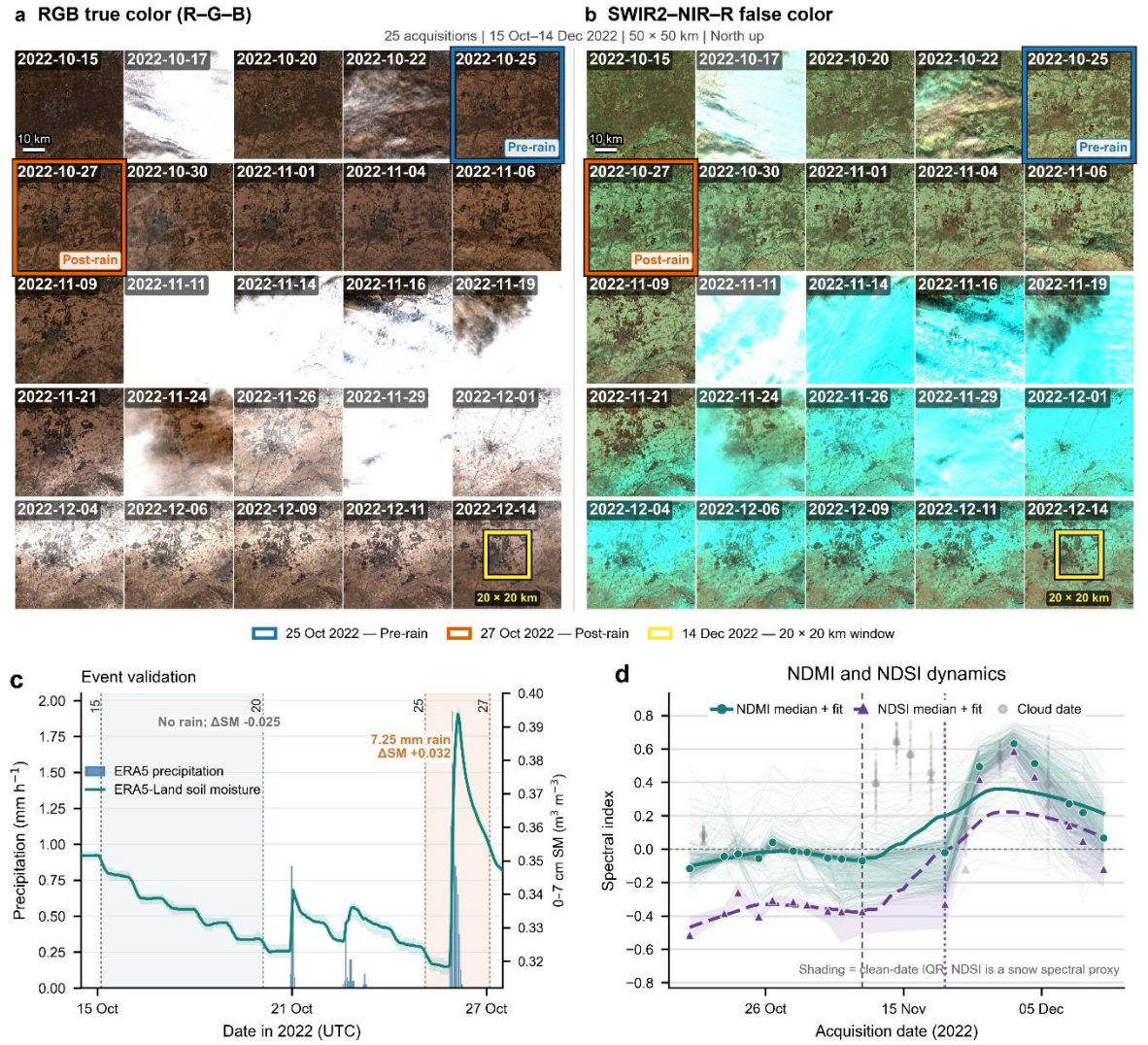


Fig.S14 Multi-temporal satellite observations and meteorological evidence of rainfall- and snow-related changes in post-harvest cropland landscapes. (a–b) True-color (red–green–blue) and false-color (SWIR2–NIR–red) composites of 25 Sentinel-2 acquisitions from 15 October to 14 December 2022. Each image covers the same 50 × 50 km area, with north at the top. Blue and orange outlines highlight the pre-rainfall (25 October) and post-rainfall (27 October) acquisitions, respectively; yellow outlines delineate a 20 × 20 km window in the 14 December images. (c) Hourly ERA5 precipitation (bars) and ERA5-Land volumetric soil moisture in the upper 0–7 cm (line). Shaded intervals indicate a rain-free drying period during 15–20 October, with a soil-moisture change (ΔSM) of −0.025 $m^3\ m^{-3}$, and a rainfall-related wetting period during 25–27 October, with 7.25 mm of cumulative rainfall and ΔSM of +0.032 $m^3\ m^{-3}$. (d) Temporal dynamics of the Normalized Difference Moisture Index (NDMI) and Normalized Difference Snow Index (NDSI). Circles and triangles represent acquisition-date medians of NDMI and NDSI, respectively; solid and dashed curves show their fitted trends. Shaded bands indicate the interquartile ranges for cloud-free acquisition dates, and gray symbols denote cloud-affected observations. The gray and purple vertical lines mark 9 and 21 November, respectively. NDSI is used as a spectral proxy for snow cover.

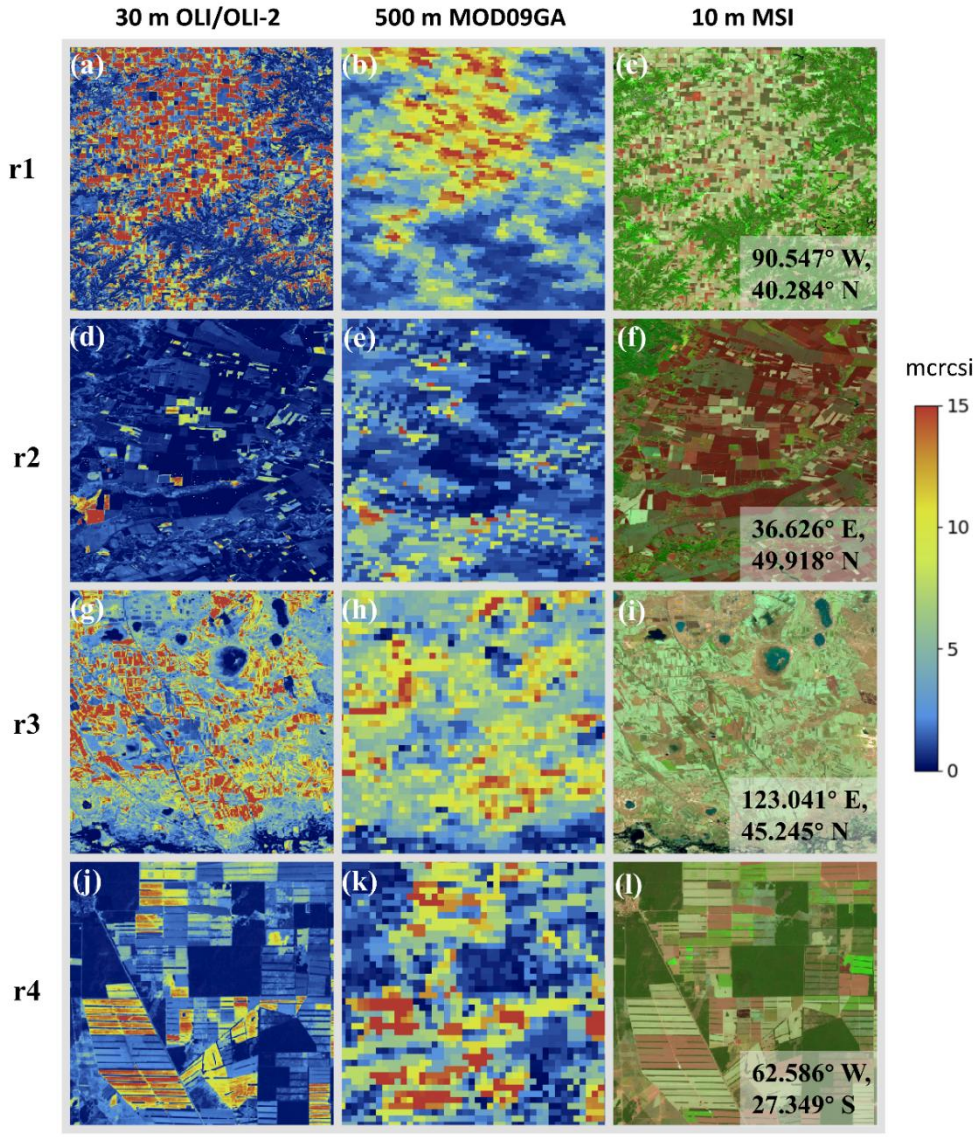


Fig.S15 Cross-sensor comparison of same-date MCRCSI. Comparison of MCRCSI derived from Landsat, MODIS, and Sentinel-2 sensors across four representative regions: r1, Illinois, USA; r2, Ukraine border; r3, Songnen Plain, Northeast China; and r4, Grand Chaco, Argentina. Coordinates indicate the central location of each study site.

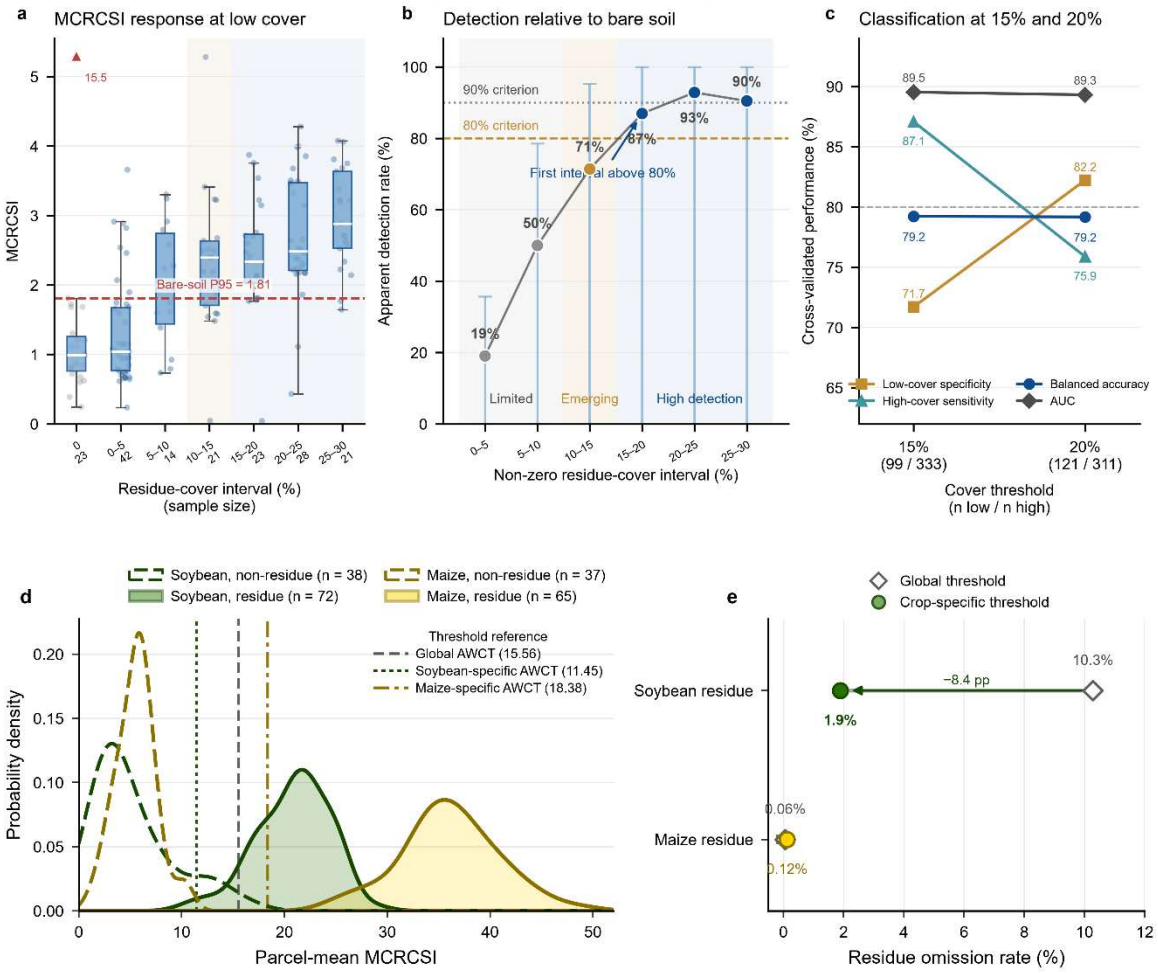


Fig.S16 MCRCSI sensitivity to low crop-residue cover and crop-specific differences in residue detectability. (a) MCRCSI distributions across residue-cover intervals, with the bare-soil 95th percentile shown as a reference; (b) apparent detection rate relative to bare soil, with bootstrap confidence intervals; and (c) cross-validated classification performance using 15% and 20% residue-cover thresholds. Specificity represents the proportion of low-cover samples (< threshold) correctly classified as low cover, sensitivity represents the proportion of high-cover samples (≥ threshold) correctly classified as high cover, balanced accuracy is the mean of sensitivity and specificity, and ROC–AUC represents the area under the receiver-operating-characteristic curve. Data in (a–c) were derived from the USGS field spectral library after convolution to Sentinel-2 bands. (d) Parcel-mean MCRCSI distributions for soybean and maize fields with and without residue cover; vertical lines indicate the global AWCT threshold (15.56) and crop-specific thresholds for soybean (11.45) and maize (18.38). Dashed curves represent non-residue parcels (soybean, n = 38; maize, n = 37), and filled curves represent residue-covered parcels (soybean, n = 72; maize, n = 65). (e) Residue omission rates obtained using global and crop-specific thresholds, showing a marked reduction in soybean omission from 10.3% to 1.9%, while maize omission remained very low under both thresholds.

1336

1337 **Table S1** Spectral band information of satellite images used in this study

Satellite	Spectral Bands #	Central wavelength (nm)	Band width (nm)	Spatial Resolution (m)
Sentinel-2 A/B/C MSI	B2-Blue	490	459 - 525	10
	B3-Green	560	542 - 578	10
	B4-Red	665	649 - 680	10
	B5-RedEdge1	705	697 - 711	20
	B6-RedEdge2	740	733 - 748	20
	B7-RedEdge3	785	772 - 792	20
	B8-NIR	842	779 - 885	10
	B8A-Narrow NIR	865	854 - 875	20
	B11-SWIR1	1610	1568 - 1659	20
	B12-SWIR2	2190	2115-2289	20
Landsat 8/9 OLI/ OLI-2	B2-Blue	480	450 – 510	30
	B3-Green	560	530 – 590	30
	B4-Red	655	640 – 670	30
	B5-NIR	865	850 – 880	30
	B6-SWIR1	1610	1570 – 1650	30
	B7-SWIR2	2200	2110 – 2290	30
MODIS Terra MOD09GA.061	b03-Blue	469	459–479	500
	b04-Green	555	545–565	500
	b01-Red	645	620–670	500
	b02-NIR	860	841–876	500
	b06-SWIR1	1640	1628–1652	500
	b07-SWIR2	2130	2105–2155	500

1338

1339 **Table S2** Crop maps used in this study

Crop type maps	Period	Spatial Resolution(m)	Source
Northeast China Crop Maps	2019–2023, yearly	10	You et al. (2021)
USDA-NASS Crop Data Layer	2019–2024, yearly	30	Boryan et al. (2011)
Ukraine Crop Maps	2020	10	Chen et al. (2024a)
Argentina Map of National Crops	2018–2024, yearly	30	De Abelleyra et al. (2024)

1340

Table S3 Number of independent Residue and Non-residue samples in the study sites

Site	Residue samples	Residue (%)	Non-residue samples	Non-residue (%)	Total
S1	1,000	41.9%	1,387	58.1%	2,387
S2	1,179	50.7%	1,148	49.3%	2,327
S3	1,180	45.7%	1,400	54.3%	2,580
S4	1,277	52.5%	1,156	47.5%	2,433
S5	2,881	63.2%	1,678	36.8%	4,559
S6	1,628	59.9%	1,092	40.1%	2,720
S7	1,093	35.6%	1,975	64.4%	3,068
S8	950	45.0%	1,160	55.0%	2,110
All	11,188	50.4%	10,996	49.6%	22,184

1344 **Table S4** M-index between residue cover (RC) and other land-cover types including bare soil
1345 (BS), unharvested maize (CU), tree canopy (TC), green vegetation (GV), and water body
1346 (WB) across different regions.

Region	Type	Blue	Green	Red	NIR	SWIR1	SWIR2
Northeast China	RC vs WB	2.54	2.78	4.19	7.32	17.78	16.87
	RC vs GV	1.39	1.28	1.83	0.24	3.54	3.02
	RC vs BS	4.39	5.16	5.51	5.48	5.66	3.66
	RC vs CU	3.16	3.79	3.95	3.78	4.90	4.75
	RC vs TC	5.17	5.53	6.72	3.01	10.45	11.39
	RC vs Snow	3.08	2.77	2.31	1.85	5.18	3.23
	RC vs Cloud	1.93	1.97	1.91	1.75	1.23	1.51
Argentina	RC vs BS	1.55	2.38	2.90	3.50	1.86	0.99
	RC vs TC	3.80	4.56	5.07	4.17	5.73	5.73
	RC vs WB	2.38	2.36	3.15	4.25	13.67	13.50
	RC vs GV	3.46	3.61	4.54	0.37	4.31	4.45
	RC vs CU	3.07	3.41	3.25	3.48	3.06	2.77
U.S. Midwest	RC vs BS	2.44	2.80	3.10	3.22	2.27	0.91
	RC vs TC	4.61	4.24	6.16	1.77	7.23	8.52
	RC vs GV	4.69	3.48	5.23	0.15	5.09	6.10
	RC vs WB	5.09	5.26	6.84	8.24	17.00	17.73
	RC vs CU	2.69	2.82	2.81	2.67	3.01	3.02
Ukraine	RC vs BS	3.55	4.53	5.33	5.15	2.55	0.75
	RC vs CU	3.04	3.54	3.67	3.50	4.54	3.78
	RC vs GV	3.37	3.93	5.81	0.48	3.91	3.33
	RC vs TC	4.55	5.37	6.88	2.49	6.38	6.40
	RC vs WB	5.54	4.15	6.69	7.00	14.08	11.75

1347

1348

1349 **Table S5** M-index of residue-covered croplands vs. others across black soil regions

BSR	Type	NDTI	NDRI	NDSVI	DFI	CRCSI	**MCRCSI**	NDVI	BSI	BAI
Argentina	RC-BS	1.84	3.39	2.26	1.58	2.68	**2.69**	0.59	2.98	3.86
	RC-CU	0.15	2.26	2.62	0.82	2.97	**2.80**	0.76	1.23	1.99
	RC-GV	2.28	2.55	3.32	4.61	3.65	**2.87**	4.90	3.37	2.65
	RC-TC	0.93	2.63	3.11	3.11	4.13	**3.33**	3.22	1.56	0.58
	RC-WB	1.63	1.68	2.12	0.32	4.89	**3.57**	0.63	0.53	2.12
China	RC-BS	2.69	3.98	3.20	2.44	4.48	**3.34**	0.22	3.36	4.68
	RC-CU	0.49	1.62	1.97	0.13	3.27	**2.85**	0.74	1.08	2.07
	RC-GV	0.82	0.83	0.88	1.82	2.32	**1.93**	1.65	1.68	1.55
	RC-TC	3.09	1.48	2.29	4.03	4.52	**3.28**	5.42	4.33	0.69
	RC-WB	1.15	3.41	4.51	0.15	4.85	**3.39**	3.70	0.90	4.82
Ukraine	RC-BS	5.23	5.54	4.85	6.15	4.44	**3.11**	1.44	4.17	4.70
	RC-CU	0.64	1.36	1.52	0.99	3.23	**2.74**	0.84	1.11	2.14
	RC-GV	1.32	1.39	2.67	4.97	3.74	**2.63**	6.14	2.76	1.95
	RC-TC	3.22	1.41	2.76	3.26	4.12	**3.01**	5.53	3.54	0.75
	RC-WB	1.08	1.09	2.02	0.20	4.59	**3.14**	1.44	0.17	4.45
United States	RC-BS	3.48	3.55	2.82	3.21	3.25	**2.80**	0.29	3.17	3.39
	RC-CU	0.05	2.06	2.41	0.78	2.25	**2.02**	0.74	1.07	1.75
	RC-GV	2.94	2.52	3.45	4.60	3.76	**2.85**	5.92	4.36	2.50
	RC-TC	2.31	2.17	3.37	3.82	4.11	**3.11**	5.07	3.39	0.69
	RC-WB	0.99	4.25	5.43	0.33	4.68	**3.33**	4.52	2.69	7.41

1350

Table S6 Accuracy assessment of different testing regions based on different threshold models. The highest value in each metric is highlighted in bold.

Region	Model	Thresholds	OA	PA	UA	F1	MCC
S1	GMM	9.88	0.73	0.97	0.62	0.76	0.56
	Kmeans	15.59	0.83	0.92	0.74	**0.82**	0.68
	Otsu	15.63	**0.83**	0.92	0.74	**0.82**	0.68
	AWCT	13.81	0.81	0.95	0.71	0.81	0.66
S2	GMM	11.76	0.95	1.0	0.91	0.96	0.91
	Kmeans	17.40	0.97	0.95	1.0	0.97	0.95
	Otsu	17.33	0.97	0.95	1.0	0.97	0.94
	AWCT	15.52	**0.98**	0.98	0.99	**0.98**	0.97
S3	GMM	3.06	0.91	0.99	0.86	0.91	0.82
	Kmeans	5.62	0.89	0.81	0.99	0.89	0.81
	Otsu	5.63	0.89	0.81	0.99	0.89	0.81
	AWCT	4.75	**0.93**	0.89	0.97	**0.93**	0.86
S4	GMM	2.21	0.89	1.0	0.8	0.89	0.8
	Kmeans	8.17	0.94	0.88	0.99	0.93	0.88
	Otsu	8.14	0.94	0.88	0.99	0.93	0.89
	AWCT	6.23	**0.98**	0.97	0.98	**0.98**	0.96
S5	GMM	4.32	0.90	1.0	0.87	0.93	0.79
	Kmeans	12.21	0.74	0.6	1.0	0.75	0.59
	Otsu	12.16	0.75	0.61	1.0	0.76	0.59
	AWCT	9.18	**0.98**	0.99	0.99	**0.99**	0.96
S6	GMM	6.93	**0.96**	0.99	0.95	**0.97**	0.92
	Kmeans	12.57	0.92	0.87	0.99	0.93	0.85
	Otsu	12.47	0.93	0.88	0.99	0.93	0.86
	AWCT	10.61	0.95	0.94	0.98	0.96	0.9
S7	GMM	12.27	**0.92**	0.91	0.88	**0.89**	0.83
	Kmeans	12.69	0.91	0.88	0.88	0.88	0.81
	Otsu	12.69	0.91	0.88	0.88	0.88	0.81
	AWCT	12.55	**0.92**	0.89	0.88	**0.89**	0.82
S8	GMM	6.12	0.91	0.98	0.85	0.91	0.83
	Kmeans	9.38	0.93	0.87	0.99	0.92	0.87
	Otsu	9.31	0.94	0.88	0.99	0.93	0.88
	AWCT	8.28	**0.94**	0.94	0.94	**0.94**	0.89
All	GMM	7.07	0.896	0.980	0.842	0.902	0.808
	Kmeans	11.70	0.891	0.848	0.948	0.886	0.805
	Otsu	11.67	0.895	0.851	0.948	0.889	0.808
	AWCT	10.12	**0.936**	**0.944**	0.930	**0.935**	**0.878**

1355 **Table S7** Band parameters of major commercial Earth observation satellites in the 2020s.

Satellite	Source	Resolution (Multi-spectral)	Revisit period	Band parameters	Equipped with SWIR bands
Jilin-1	China	2 m	35–37 times per day	Pan, Blue, Green, Red, NIR	N
Gaojing-1	China	2 m	≤1 day	Pan, Blue, Green, Red, NIR	N
PlanetScope	Planet Labs, United States	~3 m	~1 day	Coastal, Blue, Green-I, Green, Yellow, Red-Edge, NIR	N
Pléiades	Airbus, France	2 m	~1day	Pan, Blue, Green, Red, NIR	N
WorldView-3	Maxar, United States	MS 1.24 m; SWIR 3.7 m	<1 day	Pan, 8×VNIR (Coastal, Blue, Green, Yellow, Red, Red-Edge, NIR1, NIR2), 8×SWIR, CAVIS	Y

1356