https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Open Access Earth System Science Data Discussions

Mapping Annual Crop Residue Cover Across Global Mollisol Regions at 10-m Resolution (2019–2024)

Yifeng Cui[1,2,3], Jinwei Dong[1,3], Chao Zhang[4], Nanshan You[1], Shuai Ren[2], Peng Zhu[5], Quan Duan[6], Ronggao Liu[1], César Terrer[2]

[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]University of Chinese Academy of Sciences, Beijing, 100049, China
[4]Department of Civil and Environmental Engineering, National University of Singapore, Singapore, 117576, Singapore
[5]Department of Geography and Institute for Climate and Carbon Neutrality, The University of Hong Kong, Hong Kong SAR, China
[6]Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modelling, Institute for Global Change Studies, Tsinghua University, Beijing, 100084, China

Correspondence to: Jinwei Dong (dongjw@igsnrr.ac.cn)

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Earth System Science Data
Open Access
Discussions

Abstract

Mollisol regions are among the world's major grain-producing areas, yet intensive cultivation and residue removal threaten the long-term sustainability of these fertile soils. Retaining crop residues on the soil surface is an important conservation practice that can reduce erosion, conserve soil moisture, and sustain carbon inputs. However, spatially explicit, multi-year observations remain scarce, limiting consistent assessment of management patterns and their environmental implications. Here we present CrRUC-M (Crop ResidUe Cover across global Mollisol regions), the first annual, 10-m wall-to-wall dataset covering 2019–2024 and derived from Sentinel-2 imagery. We developed a knowledge-guided framework for mapping crop residue cover across croplands. First, annual non-growing-season Sentinel-2 composites were generated to reduce environmental interference. Second, phenological and spectral constraints were applied to screen residue-consistent candidate pixels. Third, final residue cover was mapped by applying the Multi-band Crop Residue Cover Spectral Index (MCRCSI) and region- and year-specific Otsu thresholds to these candidate pixels. The resulting maps were validated using reference samples derived from windshield surveys and visual interpretation, achieving an overall accuracy of 0.83 and F1-score of 0.81. Region-level assessments, comparisons with regional datasets and statistics, and transfer tests using Landsat archives collectively demonstrated that the resulting maps reliably captured fine-scale spatial patterns and interannual variability. Spatiotemporal analyses revealed a pronounced core-periphery gradient in residue-cover frequency, with persistent high-frequency clusters in intensively cultivated areas. Corn was the dominant residue source across Mollisol regions, whereas secondary crop contributions varied with regional cropping systems. CrRUC-M provides the first multi-year crop residue cover benchmark for global Mollisol regions. It can enrich representations of agricultural management practices in Earth system models and provide insights into the effects of crop residue cover on soil carbon dynamics and climate change. The CrRUC-M dataset is publicly available at https://doi.org/10.5281/zenodo.18773303 (Cui et al., 2025).

Keywords: Crop residue cover, Mollisol, tillage, climate-smart agriculture, cropland, Sentinel-2

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

1 Introduction

Crop residues, including stems, leaves, and stalks remaining after harvest, play a critical role in Conservation Agriculture (CA) (Fang et al., 2026). As a protective surface cover, these residues can mitigate soil erosion (Yan et al., 2025), contribute to soil organic matter accumulation (Silva et al., 2026; Liu et al., 2023), and regulate hydrothermal dynamics (Roberts et al., 2005), ultimately bolstering grain yields (Deines et al., 2019) and long-term ecosystem functioning. Such ecosystem services are particularly important in Mollisol regions, which serve as global breadbaskets, but have experienced severe degradation driven by decades of intensive tillage and systemic residue removal (Gibbs and Salmon, 2015; Meng et al., 2024). Under the strategic frameworks of Climate-Smart Agriculture (CSA) and Nature-based Solutions (NbS), maintaining adequate crop residue cover has transcended local agronomy to become a pivotal metric for restoring soil fertility and enhancing agroecosystem resilience (Macfarlane et al., 2024). Hence, a clear understanding of how crop residue cover varies over time and across regions is essential for evaluating conservation tillage adoption and informing soil protection strategies, as well as for understanding its role in land-atmosphere interactions.

Despite its importance, the spatial and temporal patterns of crop residue cover across global Mollisol regions remain poorly quantified, resulting in critical data gaps that hinder cross-scale agroecosystem assessments. These gaps are manifested in several primary dimensions: first, national or subnational statistics on residue cover are often scarce and characterized by significant cross-border inconsistencies (Smerald et al., 2023); second, existing land-use and land-cover datasets typically fail to distinguish specific management practices, such as crop residue cover, from general fallow (Zhang et al., 2021) or cropped classes (Zhou et al., 2022; Hoskovec et al., 2026). Furthermore, while some studies have attempted to synthesize information through meta-analysis to derive residue-related data (Wang et al., 2024), such approaches are inherently limited by the fragmented nature of primary literature, leading to datasets that lack spatiotemporal continuity and struggle to represent dynamic trends at a broader scale. As a result, most prevailing global databases on conservation agriculture report adoption levels only at coarse spatial resolutions or aggregated administrative scales (Porwollik et al., 2019). These products are typically derived from model simulations, gridded statistical upscaling (Zhang et al., 2026), or meta-analytical approaches (Prestele et al., 2018; Kassam et al., 2019; Hoskovec et al., 2026), rather than from spatially explicit, observation-based mapping. While they provide useful macro-level estimates of conservation agriculture adoption, they lack the high-resolution, wall-to-wall information required for field-level modelling and process-based assessments. This limitation is particularly critical for crop residue cover, which represents a core management practice within conservation agriculture yet remains largely absent from globally consistent, observation-based datasets.

These gaps are largely rooted in the inherent complexities of surface monitoring; unlike conventional cropland mapping, crop residue cover exhibits pronounced post-harvest seasonality, strong spatial heterogeneity, and spectral similarity with senescent crops (Liu et al., 2025). These characteristics render it difficult to detect through standard land-surface classification

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

approaches, leaving their total extent and interannual variability uncharacterized. These same complexities also limit satellite-based residue mapping, particularly when observations rely on single-date imagery or conventional land-surface classification approaches. Nevertheless, satellite remote sensing remains the only practical and scalable approach for monitoring crop residue cover across broad Mollisol landscapes, because it provides repeatable, spatially explicit observations over large agricultural regions. When integrated across sensors and acquisition dates, these observations can capture post-harvest temporal signals that help characterize residue dynamics despite spectral confusion and short monitoring windows. However, despite these advantages, existing research remains severely fragmented and geographically constrained, failing to provide a cohesive observational dataset. Most efforts are currently confined to specific “hotspots” of conservation agriculture, such as the U.S. Corn Belt (Azzari et al., 2019; Wu et al., 2026) or Northeast China (Zhang et al., 2025b; Liu et al., 2026), and typically span only isolated years, limiting the ability to characterize long-term conservation tillage trends across diverse Mollisol landscapes. These geographic and temporal gaps are further exacerbated by significant environmental and observational constraints; in high-latitude Mollisol regions, the window for effective monitoring is extremely limited due to persistent cloud cover, seasonal snow, and low solar angles (Bonney and Zhang, 2026). Moreover, the rapid temporal dynamics of post-harvest operations, such as harvest progression (Liu et al., 2025), mean that single-date satellite acquisitions often fail to capture representative crop residue cover levels consistently across regions or years. Consequently, there is still no unified, multi-sensor, harmonized data product that captures residue dynamics consistently across the world’s Mollisol regions (Zhou et al., 2025).

To bridge these knowledge gaps, this study presents CrRUC-M, a wall-to-wall crop residue cover dataset (2019–2024) specifically designed for global Mollisol regions. CrRUC-M is primarily developed using Sentinel-2 MSI imagery, producing annual residue-cover maps at 10 m spatial resolution with enhanced spatial detail and temporal observation density during the non-growing season. In addition, the framework was further extended to Landsat archives to explore cross-sensor transferability and compatibility with long-term Earth observation records. Crop residue cover was mapped using a knowledge-guided framework that integrates prior understanding of crop phenology, land surface processes, and residue spectral characteristics. The dataset provides annual, spatially explicit, wall-to-wall observations of crop residue cover from 2019 to 2024 across global Mollisol regions. As a high-resolution observational record, CrRUC-M supports quantitative assessment of soil conservation practices, analyses of residue-cover impact on soil restoration, and parameterization of biogeophysical and land-surface models. The accompanying Landsat extension additionally provides opportunities for exploratory long-term continuity analyses based on historical satellite archives.

The remainder of this paper is organized as follows. We first describe the methodological framework for generating and validating the CrRUC-M dataset, including non-growing-season image compositing, residue-cover mapping, accuracy assessment, and cross-dataset comparisons. We then characterize the frequency of crop residue cover across global Mollisol regions and to examine its crop-source composition. Finally, we extend the framework to Landsat archives to produce a

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

complementary 30 m residue-cover dataset, enabling cross-sensor consistency evaluation and providing a longer-term basis for future monitoring of residue-cover dynamics.

2 Materials

2.1 Study Area

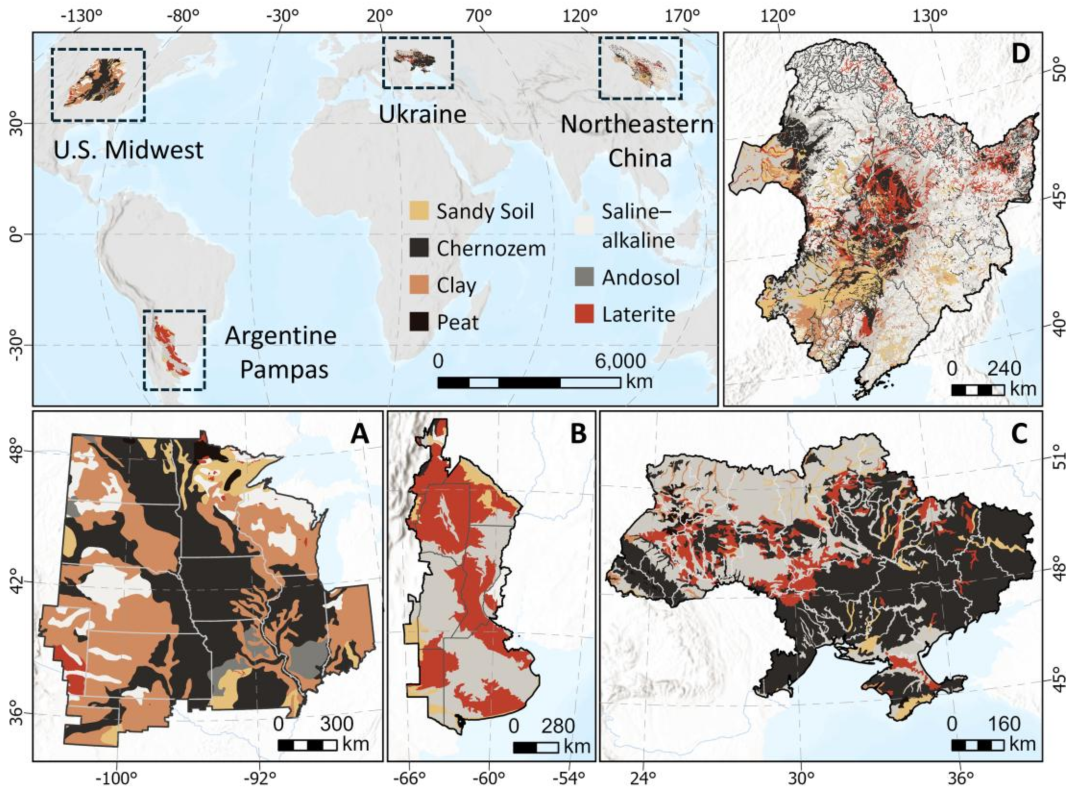
Mollisols are among the most fertile agricultural soils globally and are commonly associated with the broad group of black soils because of their dark, organic-rich surface horizons and high natural productivity. In this study, we use the term global Mollisol regions to refer to the major cropland-dominated black soil regions that broadly correspond to Mollisol or Mollisol-like soil landscapes under different national and international soil classification systems (Liao et al., 2026). These regions are hereafter referred to as Mollisol regions for consistency with the CrRUC-M dataset. Owing to inconsistencies among existing soil classification systems, the spatial extent of black soils or Mollisol-like soils lacks a unified global definition (Liu et al., 2012; Tong et al., 2024). This uncertainty arises from differences in diagnostic criteria, mapping approaches, and national soil survey standards. Therefore, instead of relying on a single global soil map, we delineated the study area using a hybrid approach that integrates administrative boundaries with multi-source datasets (Fig.S1). This design ensures a consistent and operational geospatial extent for generating annual crop residue cover datasets across major global Mollisol regions.

This study targets croplands within global Mollisol agricultural belts across four continents (**Figure 1**). As summarized in **Table 1**, the total study area spans approximately 523.12 × 10^5 km² across the U.S. Midwest, Northeast China, Ukraine, and the Argentine Pampas. The Northern Hemisphere regions include the U.S. Midwest, Northeast China, and Ukraine. The U.S. Midwest represents the largest component of the dataset, covering 209.47 × 10^5 km² across 943 county-level units in 13 states (Fig.S1). Northeast China covers 145.71 × 10^5 km² and is defined by the administrative extents of Heilongjiang, Jilin, and Liaoning Provinces, together with four leagues and cities in eastern Inner Mongolia. The Ukrainian Mollisol region covers 73.82 × 10^5 km², defined by 25 oblasts nationwide. In the Southern Hemisphere, the Argentine Pampas covers 94.12 × 10^5 km² across 11 provinces.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Table 1 Administrative extent of the global Mollisol region in the study.

Mollisol region	Description of administrative units included	Area ($\times 10^4$ km^2)
U.S. Midwest	943 county-level units across 13 states	209.47
Ukrainian	25 oblasts nationwide	73.82
Northeast China	322 counties and 40 prefecture-level cities (or leagues) across Heilongjiang, Jilin, Liaoning, and eastern Inner Mongolia	145.71
Argentine Pampas	11 provinces nationwide	94.12



125 **Figure 1** Study area and soil background of global Mollisol belts. The upper-left panel shows the global extent of the study regions, including U.S. Midwest, Argentine Pampas, Ukraine, and Northeast China. Colored overlays indicate major soil types, including sandy soil, Chernozem (black soil), clay, peat, saline–alkaline soil, Andosol, and laterite, derived from the Harmonized World Soil Database (HWSD). Panels A–C present detailed views of the Mollisol regions in the U.S. Midwest (A), Argentine Pampas (B), and Ukraine (C), respectively, while panel D shows Northeast China.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

2.1 Satellite archives and pre-processing

Sentinel-2 surface reflectance imagery (**Table S1**) from the Harmonized Sentinel-2 MSI Level-2A archive available in Google Earth Engine (GEE; COPERNICUS/S2_SR_HARMONIZED) was used as the primary satellite data source for generating the annual crop residue cover maps in this study. All available observations within the temporal window for each Mollisol region and year were collected (**Table S2**) and pre-processed before annual compositing. For Sentinel-2 imagery, cloud contamination was removed using the Cloud Score+ product (Pasquarella et al., 2023) available in GEE (GOOGLE/CLOUD_SCORE_PLUS/V1/S2_HARMONIZED). Clear-sky pixels were identified based on the cumulative distribution function score (cs_cdf, range 0–1), and pixels with cs_cdf values below 0.9 were masked. Snow and ice pixels were further excluded using the Sentinel-2 Scene Classification Layer (SCL) band. After cloud, haze, snow, and ice masking, the remaining clear-sky observations were used to construct annual non-growing-season composites for residue mapping in Section 3.1.

In addition, Landsat-series imagery was used as an auxiliary data source to evaluate the temporal continuity and cross-platform transferability of the proposed crop residue mapping framework. Specifically, Landsat observations provide a longer historical archive and therefore allow assessment of whether the residue extraction strategy can be extended beyond the Sentinel-2 era (Azzari et al., 2019). Landsat 5 TM, Landsat 7 ETM+, and Landsat 8 OLI surface reflectance imagery from the Collection 2 Tier 1 archive were used for this purpose. Clouds, cloud shadows, and snow were masked using CFMask information derived from the Quality Assessment bands (Zhu and Woodcock, 2012), and only Tier 1 scenes were retained to ensure geometric and radiometric quality. To improve cross-sensor consistency within the Landsat archive, Landsat 5 TM and Landsat 7 ETM+ reflectance values were transformed to Landsat 8 OLI-equivalent reflectance using ordinary least-squares regression coefficients in **Table S3** (Holden and Woodcock, 2016). Spectral alignment was then applied to approximate Sentinel-2 MSI spectral characteristics (Zhang et al., 2018), enabling exploratory comparison between Landsat- and Sentinel-2-derived residue products. It should be noted that the Landsat-based analysis was not used as the primary mapping product in this study, but rather as a supplementary test of temporal extensibility and sensor transferability. Differences in spatial resolution, spectral response functions, and revisit frequency may introduce additional uncertainty, especially in cloud- and snow-prone regions during the non-growing season.

2.2 Ancillary datasets for residue cover mapping

2.2.1 Cropland Mask

The 30 m GLAD Global Cropland Extent dataset (Potapov et al., 2022) was used as the primary cropland mask for the global Mollisol regions. This product provides globally consistent cropland extent information and offers the temporal stability needed for retrospective analyses (Cui et al., 2024a). Although the GLAD cropland product is updated at multi-year intervals,

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

it was considered suitable for this study because cropland patterns in the selected Mollisol regions have remained relatively stable over recent decades (Chen et al., 2025). The GLAD mask therefore provides a consistent baseline for restricting the residue cover analysis to cultivated areas.

2.2.2 Crop Type Maps

Crop type maps were used to characterize regional crop distributions and to support the interpretation of residue sources. We integrated multiple region-specific datasets (**Table S4**). These datasets included the 10 m crop type map for Northeast China (You et al., 2021), the 30 m USDA/NASS Cropland Data Layer for the United States from 2019 to 2024 (Boryan et al., 2011), the 10 m crop type map for Ukraine (Chen et al., 2024), and the 30 m Argentina National Map of Crops from 2018 to 2024 (De Abelleyra et al., 2024).

2.2.3 Crop Calendar

Crop phenology was used to define the temporal window for crop residue detection because crop residues are exposed after harvest and before subsequent planting. We characterized the non-growing season using crop calendar information from the USDA Foreign Agricultural Service, International Production Assessment Division (FAS/IPAD) (Fig.S2). In the Northern Hemisphere, including the U.S. Midwest, Northeast China, and Ukraine, harvest generally occurs from September to October, whereas in the Southern Hemisphere, represented by Argentine Pampas, it primarily takes place from March to May. Although a small proportion of crops in regions such as Northeast China may remain unharvested until the following spring, the vast majority are harvested in autumn (Bao et al., 2025). Based on these phenological constraints, satellite imagery was systematically filtered to capture the critical period between harvest and subsequent planting.

2.3 External reference datasets

To evaluate the consistency of CrRUC-M with established reference products and independent statistical surveys, we used two external datasets that provide complementary information on crop residue cover and tillage practices.

Corn Residue Coverage Data of Northeast China. We used the Corn Residue Coverage (CRC) dataset developed by Dong et al. (2024) as an independent reference for evaluating residue-covered cropland in Northeast China. The CRC dataset provides annual corn residue-cover estimates, ranging from 0 to 100%, from 2013 to 2022, with two observations per year corresponding to the post-harvest and pre-sowing periods. Following the CRC definition, pixels with residue cover greater than 30% were classified as conservation tillage. For comparison with CrRUC-M, CRC-derived conservation tillage areas were aggregated to 0.5 degree grid cells. The resulting gridded estimates were used to evaluate the spatial consistency between CrRUC-M and an existing regional residue-cover dataset in Northeast China.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

USDA-NASS Tillage Survey Report. Crop residue cover estimates from CrRUC-M were compared with agricultural census and survey statistics from the United States Department of Agriculture National Agricultural Statistics Service (USDA-NASS) (Nelson, 2003) at an administrative scale. USDA-NASS provide comprehensive tillage-practice statistics through the U.S. Census of Agriculture and associated surveys. These data report state- and county-level areas under various management systems, including no-till, conservation tillage (excluding no-till), and conventional tillage for the survey years 2017 and 2022. Each indicator is published in terms of total area, number of operations, and average area per operation, enabling a multi-scale assessment of tillage adoption. We used state-level no-till and conservation tillage areas from the survey years of 2022 as independent reference metrics to evaluate CrRUC-M residue-cover estimates across the U.S. Corn Belt.

3 Methodology

We developed a framework to map crop residue cover from satellite imageries (**Figure 2**). The framework is structured into three modules: (1) Compositing of non-growing season (NGS) time-series imageries for residue mapping; (2) a knowledge-guided residue cover mapping, and (3) datasets validation and comparison.

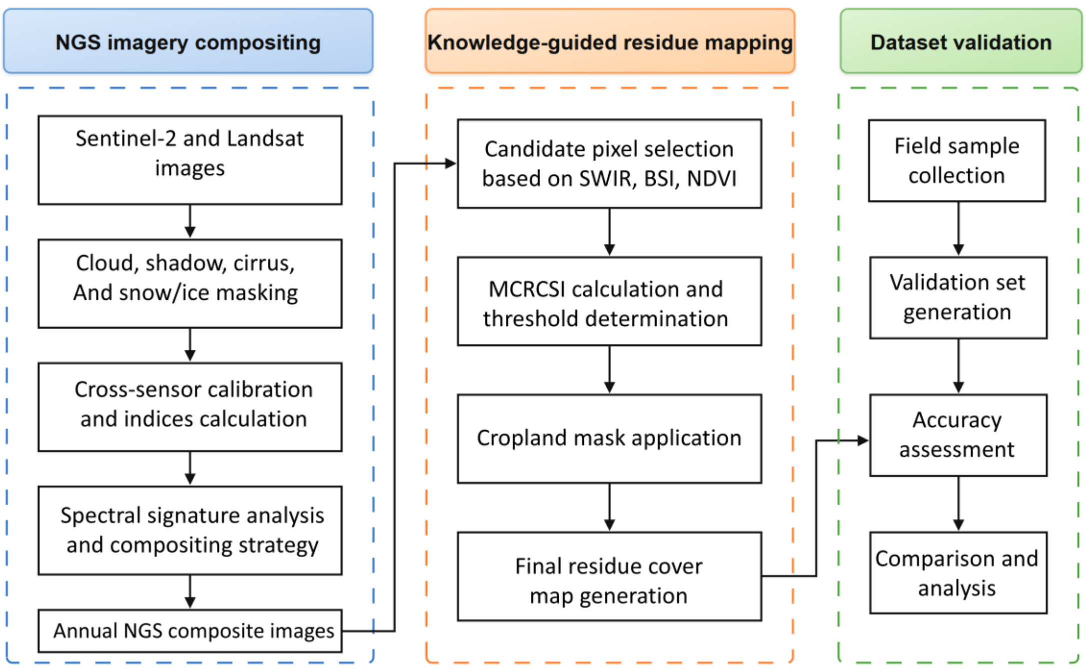


Figure 2 Workflow of the study

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

3.1 Observational compositing during Non-Growing Season (NGS)

3.1.1 Spectral signature analysis during NGS

Generating reliable time-series composites requires an understanding of non-growing season (NGS) spectral dynamics, as inappropriate compositing can introduce significant biases in large-scale analyses(Qiu et al., 2023; Xu et al., 2025). During the non-growing season: the interval between autumn harvest and spring sowing, cropland surfaces in Mollisol regions exhibit high spatiotemporal complexity driven by agricultural operations such as harvesting and tillage, alongside natural shifts like freeze-thaw cycles, snow accumulation, and cloud cover. In high-latitude Northern Hemisphere regions, a stable snow-cover period lasting at least 2.5 months bisects this season into two distinct monitoring phases: the Autumn Residue-Cover Phase and the Spring Exposure Phase. Analysis of a representative Mollisol region in Shuangcheng, China (**Figure 3)**, illustrates that the autumn phase begins in early October as crop residues rapidly accumulate during intensive harvesting, forming a stable, spatially continuous layer by late October that remains detectable until persistent snowfall obscures the surface in December. The subsequent Spring Exposure Phase initiates in early March when rapid snowmelt re-exposes the previous season's residues, providing a second, albeit brief, monitoring opportunity. However, this window is short-lived; by mid-April, intensive spring tillage such as plowing and harrowing quickly incorporates these residues into the soil, leading to darker and more homogeneous bare-soil tones before the next growth cycle begins in early May. These discontinuous and ephemeral windows are frequently obscured by atmospheric interference and moisture fluctuations, emphasizing the critical need for an image compositing algorithm designed to maximize residue signals while minimizing disturbances from clouds, snow, and soil moisture.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Open Access Earth System Science Data Discussions

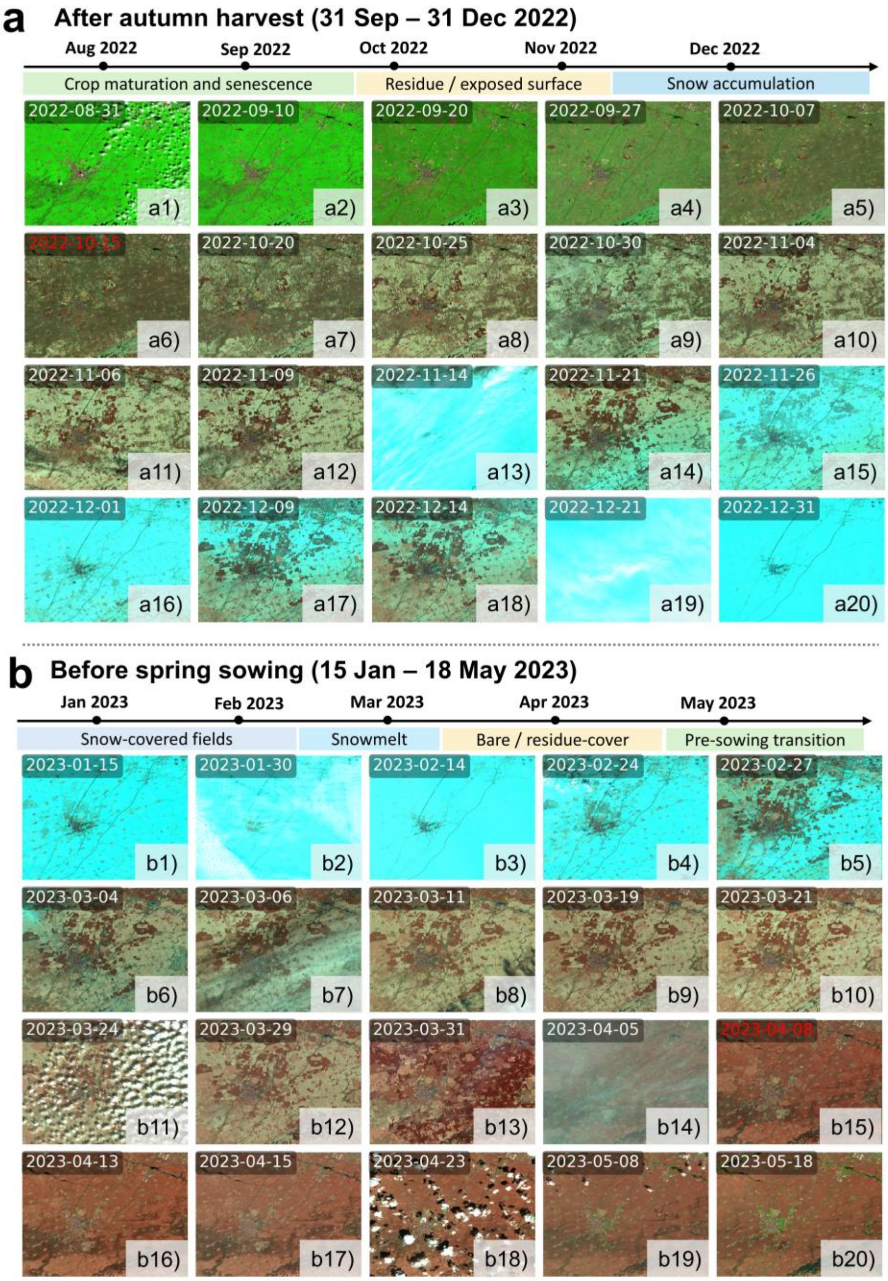


Figure 3 Sentinel-2 false-color composites illustrating post-harvest and pre-sowing surface conditions in the Mollisol region of Northeast China. (a) Time series of Sentinel-2 false-color composites acquired after the 2022 autumn harvest (31 August–31 December 2022), capturing the progressive transition from green vegetation to exposed cropland surfaces, residue cover, and subsequent snow accumulation during early winter. (b) Time series of Sentinel-2 false-color composites acquired before spring sowing in 2023 (15 January–18 May 2023), showing snow-covered fields, snowmelt, and the re-emergence of bare and residue-covered soils prior to planting. Dates of image acquisition are indicated in each panel, and labels (a1–a20, b1–b20) correspond to individual observations in chronological order.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

We visualized the seasonal trajectories during the NGS period to illustrate the spectral transitions associated with crop harvest (**Figure 4**,Fig.S3-5). The compositing strategy is grounded in a clear "peak–valley" inversion of spectral responses triggered by the harvest event. As shown by the time-series analysis in Northeast China (**Figure 4**), the conversion from standing crops to surface residue is characterized by a simultaneous but opposite shift in key indicators: vegetation-related metrics such as NDVI and NIR rapidly decline to their annual minima as photosynthetic activity ceases, whereas residue-sensitive signals, including MCRCSI and SWIR bands, increase to their annual maxima due to enhanced exposure of dry plant material. This abrupt spectral inversion provides a strong physical rationale for image compositing. A maximum-value strategy effectively captures the period of peak residue exposure, while a minimum-value strategy, such as Min-NDVI, isolates time windows with minimal vegetation interference. In this context, harvest represents a critical phenological turning point that generates a synchronized yet contrasting spectral response, enabling reliable discrimination of residue-covered surfaces from both actively growing crops and bare soil.

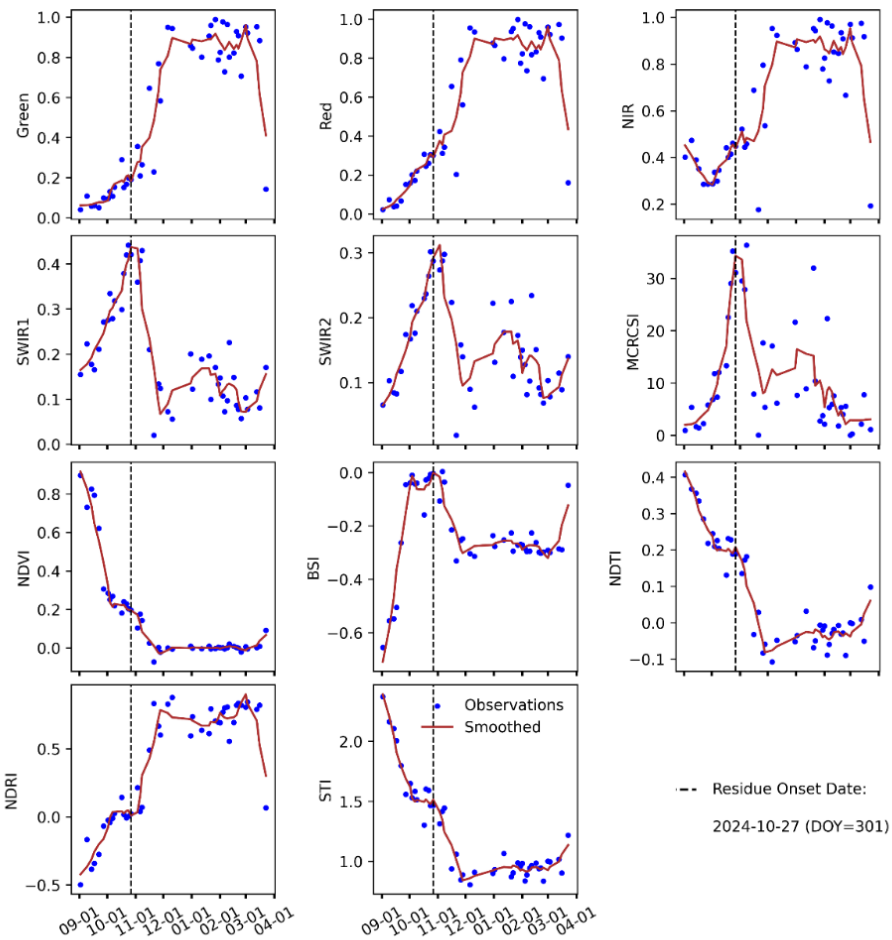


Figure 4 Sentinel-2 time series of residue cover pixel during the 2024 non-growing season in China

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

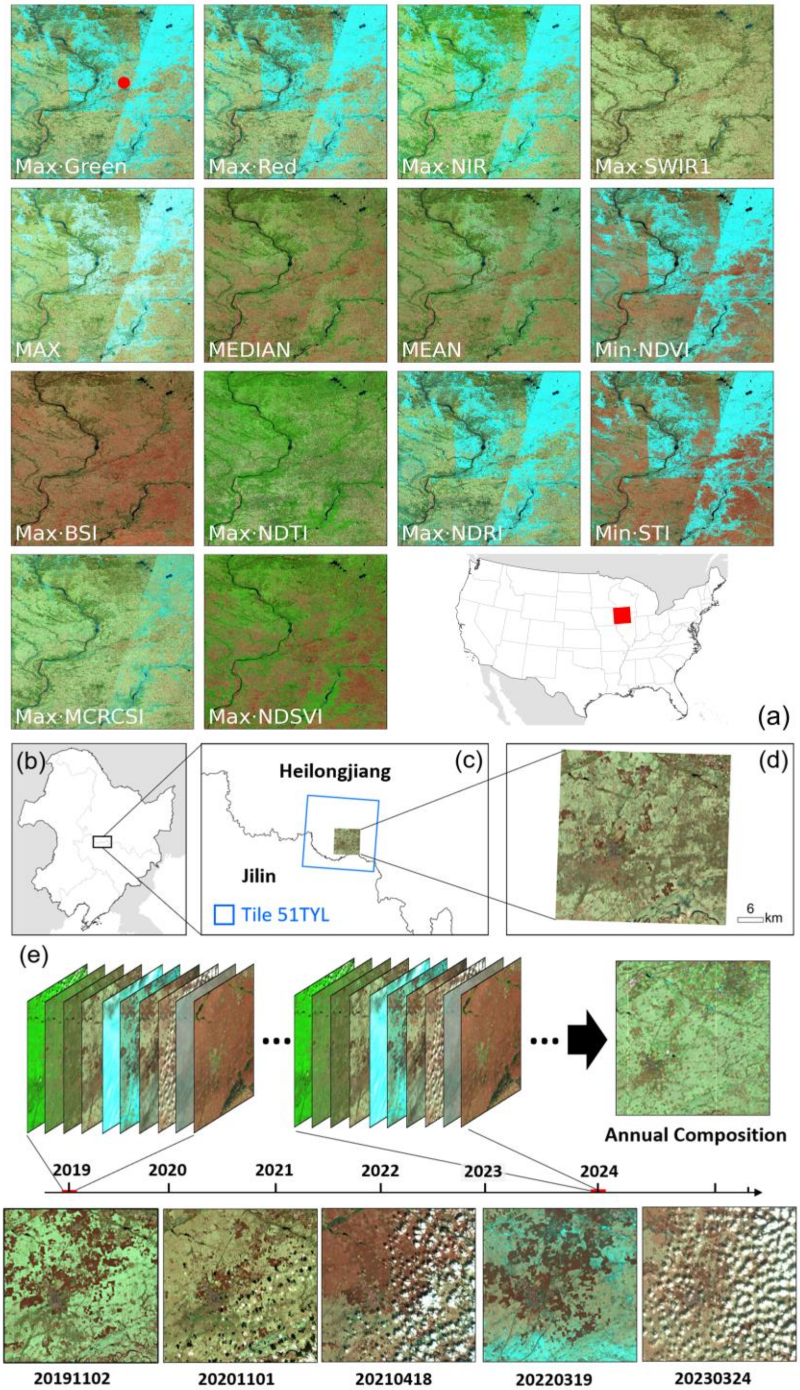
3.1.2 Comparison of compositing methods

Based on the analyse above, we evaluated 14 candidate compositing strategies categorized into statistics-based, band-based, and index-based methods across the global Mollisol regions (**Table 2**). Our comparison revealed a pervasive trade-off between information fidelity and noise suppression: traditional statistical methods like MEAN and MEDIAN effectively smoothed cloud and snow contamination but resulted in significant signal attenuation and homogenization of residue features. Conversely, index-based methods such as Max-MCRCSI excelled at enhancing residue signals but remained highly susceptible to bright artifacts from cloud and snow edges. Among all tested strategies, Max-SWIR1 emerged as the optimal solution (**Figure 5**). By leveraging the physical property that snow reflectance decreases sharply in the shortwave infrared range, the Max-SWIR1 method successfully suppresses snow-related noise while preserving clear, high-contrast surface information. This strategy effectively isolates residue signals while suppressing atmospheric and snow interference.

Table 2 Strategies used for compositing time-series observations

Compositing Methods	Calculations	Reference
Max Green	Observation maximizing Green reflectance	-
Max Red	Observation maximizing Red reflectance	-
Max NIR	Observation maximizing NIR reflectance	-
Max SWIR1	Observation maximizing SWIR1 reflectance	-
MAX	Pixel-wise maximum across observations	Holben (1986)
MEDIAN	Pixel-wise median across observations	Teluguntla et al. (2018)
MEAN	Pixel-wise mean across observations	Griffiths et al. (2019)
Min NDVI	(NIR − Red) / (NIR + Red)	Liu (2016)
Max BSI	((SWIR2 + Red) − (NIR + Blue)) / ((SWIR2 + Red) + (NIR + Blue))	-
Max NDTI	(SWIR1 − SWIR2) / (SWIR1 + SWIR2)	Zheng et al. (2012)
Max NDRI	(Red − SWIR2) / (Red + SWIR2)	-
Min STI	SWIR1 / SWIR2	-
Max MCRCSI	abs(1000 × NIR × Red × (SWIR1 − SWIR2) / (1 − SWIR1))	-
Max NDSVI	(SWIR1−Red)/(SWIR1 + Red)	-

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026



250 **Figure 5** Comparison of time-series compositing methods and illustration of the annual non-growing-season image compositing workflow in representative Mollisol regions**.** (a) Comparison of band-, statistical-, and spectral-index-based compositing methods in the U.S. Midwest Mollisol region. The red point indicates the location used for detailed comparison. (b) Location of the Northeast China study area and selected image. (c) Location of Sentinel-2 tile 51TYL across Heilongjiang and Jilin Provinces and the selected demonstration area. (d)

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Enlarged view of the demonstration area in Shuangcheng District, Harbin, Heilongjiang Province, centered at 126.4065° E and 45.3961° N. (e) Schematic illustration of the annual compositing procedure for 2019–2024, showing the collection of original non-growing-season observations, masking of cloud- and snow-contaminated pixels, and generation of the final annual composite.

3.2 Crop residue cover mapping

3.2.1 Generation of annual NGS composite images

To support high-resolution monitoring of crop residue cover during the non-growing season (NGS), annual composite images were primarily generated using Sentinel-2 imagery for each year within the study period. For each region and year, all available cloud-free Sentinel-2 observations within the predefined NGS temporal window (**Table S2**) were collected and composited according to the optimal compositing strategy identified in Section 3.1. Standard preprocessing procedures, including cloud and shadow masking and quality filtering, were applied prior to compositing. The resulting annual NGS composites preserve the spectral characteristics required for residue detection while minimizing cloud contamination, atmospheric noise, and transient surface conditions. These composites serve as the primary input datasets for the subsequent crop residue cover mapping framework. To further explore the transferability of the framework to historical satellite archives, the same compositing strategy was additionally applied to Landsat imagery (Fig.S6). This complementary Landsat implementation facilitates exploratory long-term continuity analyses while maintaining general methodological consistency with the Sentinel-2–based framework.

3.2.2 Knowledge-guided crop residue cover mapping framework

In theory, all non-abandoned croplands may potentially exhibit crop residue cover after harvest (Zhang et al., 2021). However, the NGS land surface is highly heterogeneous, comprising crop residue, background surfaces (bare soil, water), and multiple spectrally similar interfering surfaces (e.g., evergreen forests, cotton fields, winter wheat and other green vegetation). This complex surface composition, combined with the scarcity of large-scale and temporally consistent training samples, poses significant challenges for supervised classification approaches at continental to global scales. To address this issue, we developed a knowledge-guided residue cover mapping framework. The term *knowledge-guided* refers to the incorporation of prior understanding of (i) agricultural phenology, (ii) the status of land surface during the NGS, and (iii) the spectral properties of crop residues and other land cover types. Specifically, during the NGS, actively growing vegetation is rare in most cropland areas, implying relatively low NDVI values over residue-covered fields; meanwhile, crop residues, which contain cellulose and lignin, typically exhibit enhanced reflectance in the shortwave infrared band, differing from bare soil in combined brightness and spectral index behaviour (**Figure 6**).

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Open Access Earth System Science Data Discussions

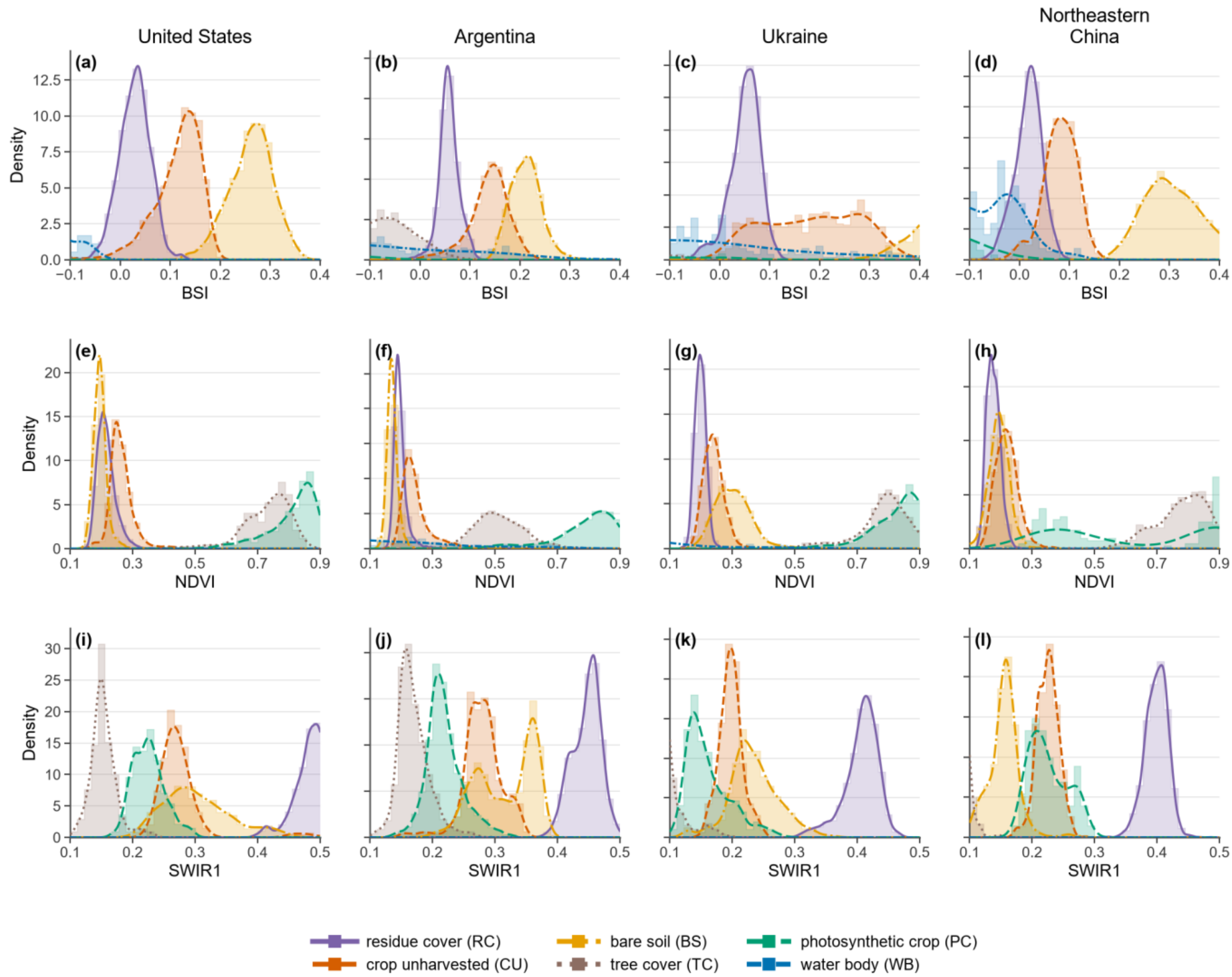


Figure 6 Histograms of BSI, NDVI, and SWIR1 reflectance for representative land-cover types in global Mollisol regions. Columns correspond to regions (U.S. Midwest, Argentine Pampas, Ukraine, and Northeast China), and rows represent spectral variables (BSI, NDVI, and SWIR1). Coloured distributions denote different land-cover types, including residue cover (RC), crop unharvested (CU), bare soil (BS), tree cover (TC), photosynthetic crop (PC), and water body (WB). Panels (a–l) illustrate the regional variability and separability of spectral responses among these classes.

In the first stage, potential residue pixels, referred to as candidate pixels, are broadly identified from the annual NGS composite imagery using a multi-index thresholding strategy grounded in prior knowledge of residue spectral behavior. Candidate pixels are identified by imposing combined constraints of low NDVI to exclude vigorous green vegetation (NDVI $< \tau_{\mathrm{NDVI}}$), low BSI to reduce confusion with exposed soil backgrounds (BSI $< \tau_{\mathrm{BSI}}$), and high SWIR1 reflectance to capture dry, cellulose-rich residue signals (SWIR1 $> \tau_{\mathrm{SWIR1}}$) in Eq. (1). Here, τ_{NDVI} and τ_{BSI} are set to 0.35 and 0.20, respectively, while τ_{SWIR1} is set to 0.35 for Ukraine and U.S. Midwest Mollisol regions and relaxed to 0.30 for Northeast China and Argentine Pampas to

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

account for regional differences in residue brightness and background surface conditions. Pixels satisfying all three conditions are defined as the candidate residue set using Eq. (1):

$$\mathrm{P} \in S_{residue}, \text{ when} \begin{cases} NDVI(P) < \tau_{NDVI} \\ BSI(P) < \tau_{BSI} \\ SWIR1(P) > \tau_{SWIR1} \end{cases} \tag{1}$$

This step does not directly produce the final residue map; instead, it narrows the classification space to pixels exhibiting residue-consistent spectral behaviour, thereby reducing spectral ambiguity and improving separability for threshold segmentation.

In the second stage, final residue-cover pixels were identified from the candidate set using an ensemble thresholding approach based on the Multi-band Crop Residue Cover Spectral Index (MCRCSI) (Cui et al., manuscript under review), which was designed to discriminate residue-covered pixels from background surfaces (Eq. (2)). To derive representative MCRCSI distributions, the study area was divided into 1° × 1° grid cells, and 200 random points were generated within each grid cell and overlaid with the corresponding MCRCSI layers (Fig.S7). The interquartile range (IQR) rule was then applied to remove outliers from the sampled MCRCSI values (You and Dong, 2020). The sampling procedure was applied separately to each Mollisol region r and year t, allowing the resulting distributions to reflect spatiotemporal variations in soil moisture, illumination, residue abundance, and management practices. Otsu's method (Otsu, 1979) was used to determine the primary MCRCSI threshold, $\tau_{\mathrm{Otsu}}^{r,t}$ in Eq. (3), by selecting the partition that maximized the between-class variance of the sampled MCRCSI distribution. Pixels exceeding this threshold were assigned to the residue set according to Eq. (4). Otsu thresholding was selected because it provides a deterministic, non-parametric solution to the one-dimensional binary-separation problem and does not require an explicit assumption regarding the component distributions. K-means clustering ($K = 2$) (Wang et al., 2019) and a two-component Gaussian mixture model (GMM) (You et al., 2023) were additionally applied as sensitivity analyses. K-means represents a distance-based partition, whereas GMM provides a probabilistic threshold under an assumed mixture distribution. The close agreement among the three thresholds (Fig.S8) indicated that threshold selection was relatively insensitive to the specific unsupervised algorithm; therefore, K-means and GMM were used to quantify method-related uncertainty rather than being averaged to define the primary threshold.

$$MCRCSI = \frac{|\rho_{NIR} \times \rho_{RED} \times (\rho_{SWIR1} - \rho_{SWIR2})|}{1 - \rho_{SWIR1}} \times 1000 \tag{2}$$

$$\tau_{\mathrm{Otsu}}^{r,t} = \arg\max_{\tau} \left[\omega_0^{r,t}(\tau)\, \omega_1^{r,t}(\tau) \left(\mu_0^{r,t}(\tau) - \mu_1^{r,t}(\tau)\right)^2\right], \tag{3}$$

$$P \in S_{\mathrm{residue}}^{r,t} \text{ if } \mathrm{MCRCSI}(P) > \tau_{\mathrm{Otsu}}^{r,t} \tag{4}$$

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Finally, the region- and year-specific adaptive threshold was applied to the corresponding candidate pixels identified in the first stage. For each pixel x, the final binary classification label $C_x^{r,t}$ was defined using Eq. (5):

$$C_x^{r,t} = \begin{cases} 1, & x \in S_{\mathrm{cand}}^{r,t} \text{ and } \mathrm{MCRCSI}_x^{r,t} \geq \tau_{\mathrm{Otsu}}^{r,t}, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where $C_x^{r,t} = 1$ denotes crop-residue cover and $C_x^{r,t} = 0$ denotes non-residue cover. Thus, a candidate pixel was classified as residue-covered when its MCRCSI value equalled or exceeded the corresponding region- and year-specific threshold. Candidate pixels below the threshold, together with pixels excluded during the first-stage screening, were assigned to the non-residue class. Based on these binary labels, the final residue-cover set was defined using Eq. (6):

$$S_{\mathrm{residue}}^{r,t} = \{x \mid C_x^{r,t} = 1\}. \tag{6}$$

3.2.3 Cropland mask and map generation

A cropland mask (Section 2.3.1) was applied to restrict crop residue cover mapping to actively cultivated land, thereby excluding non-agricultural surfaces. Following the framework described in Section 3.2.2, annual binary crop residue cover maps were produced for each region and year during 2019–2024, yielding spatially explicit wall-to-wall products. Crop type maps (**Table S4**) were further overlaid with the residue cover maps to assign residue signals to specific crop categories, enabling subsequent analysis of residue source composition.

3.3 Validation sample collection and augmentation

3.3.1 Field survey and ground reference collection

Windshield surveys were conducted to collect ground truth data for crop residue cover mapping in the major Mollisol croplands. Field surveys were conducted in Northeast China during 2022–2023 and in the U.S. Midwest in 2024, covering representative agricultural landscapes dominated by corn and soybean (**Figure 7**a, b). Along the survey routes, georeferenced field observations and photographs were collected to document both crop residue-covered surfaces and other land-surface conditions to assist residue mapping.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

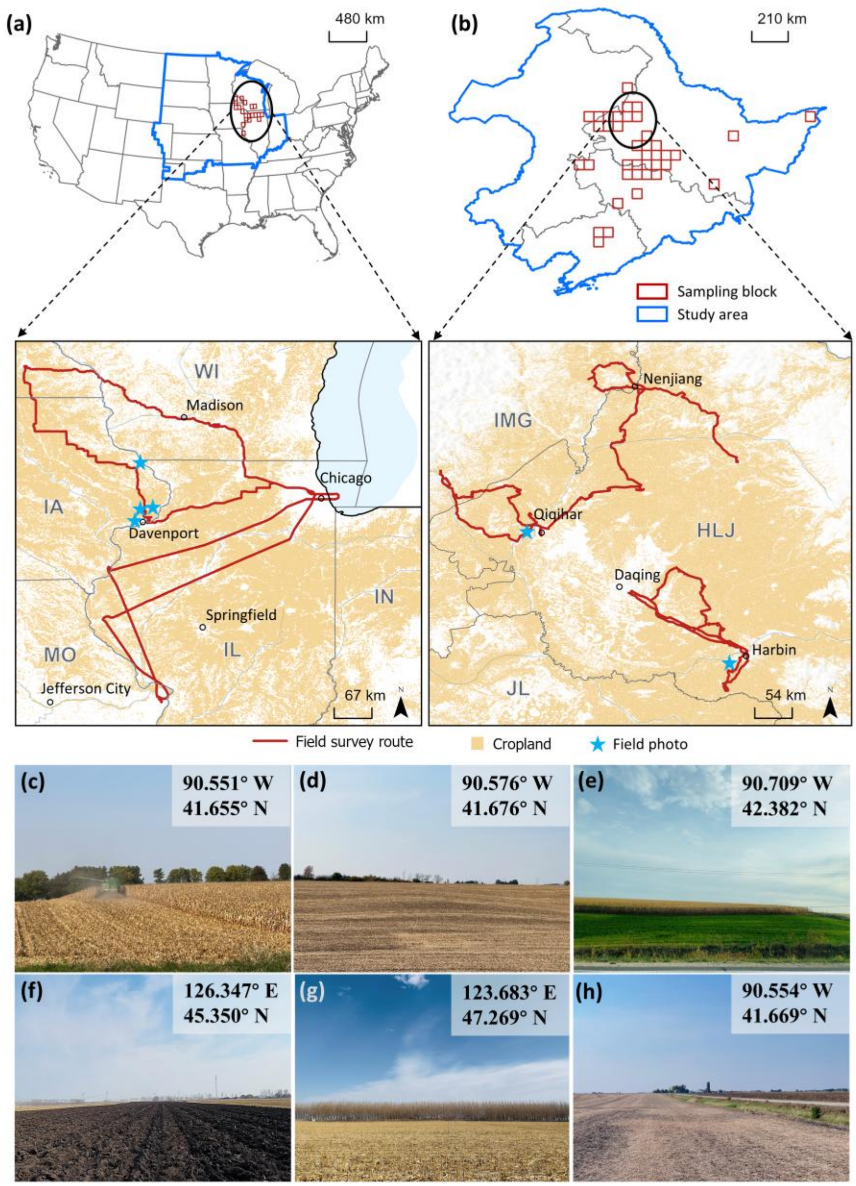


Figure 7 Field survey routes and representative ground photographs of non-growing-season cropland surfaces. Panels (a) and (b) show the field survey routes in the U.S. Midwest and Northeast China, respectively. Panels (c), (d), and (h) show harvested corn and soybean fields in the U.S. Midwest with visible residue cover. Panel (e) shows other crops that remained in the active growing season during the survey window. Panel (f) shows tilled bare soil in Northeast China, and panel (g) shows a corn field with residue cover in Northeast China. These field observations were used as ground reference for visual interpretation and sample augmentation.

The collected samples primarily included corn and soybean residue cover, corresponding to the dominant crop types in the two survey regions (**Figure 7** c, d, g, h). Additional observations were collected for non-residue or exclusion classes, including other crops that remained in the active growing season during the observation window and bare soil exposed after tillage in

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Northeast China (**Figure 7** e, f). These contrasting samples captured the main surface conditions encountered during the non-growing-season mapping window and provided field-based anchors for subsequent visual interpretation and sample augmentation. By combining residue-covered fields with visually similar or seasonally confounding non-residue surfaces, the field survey helped establish an independent reference basis for accuracy assessment.

3.3.2 Visual interpretation and sample augmentation

Although the windshield surveys provided ground truth, field samples were limited in both spatial coverage and temporal representativeness. Ground truth data were difficult to obtain for all Mollisol regions, especially Ukraine and South America, and crop residue cover is a transient, management-driven surface condition that changes rapidly after harvest and tillage. These characteristics together make it difficult to rely on generic LULC sample libraries or Google Earth high-resolution RGB imagery, which often lack the temporal specificity and spectral information needed for residue identification during the non-growing season (Cui et al., 2024b; Stanimirova et al., 2023).

We therefore took the field-survey samples as ground references and expanded them through annual and regional visual interpretation of multi-band Sentinel-2 imageries composites to construct validation samples across the four Mollisol regions and mapping years. Three types of image composites were compared for the representative sites: true-colour Red-Green-Blue composites, commonly used false-colour NIR-Red-Green composites, and SWIR2-NIR-Red false-colour composites (**Figure 8**). The true-colour and NIR-Red-Green composites provided useful contextual information on field boundaries and general land-surface conditions, but residue-covered fields were often difficult to separate visually from bare soil, harvested fields with weak residue signals, or other cropland surfaces. In contrast, the SWIR2-NIR-Red composite enhanced the spectral contrast of crop residue cover during the non-growing season, because residue-covered surfaces showed distinctive colour tones relative to tilled bare soil, actively growing crops, and other non-residue surfaces. Based on this visual separability, the SWIR2-NIR-Red composite was used as the primary image source for sample augmentation. Field-survey points and photographs were first used as anchor samples to identify the image appearance of corn and soybean residue cover, as well as non-residue surfaces such as tilled bare soil and actively growing crops. We then visually interpreted neighbouring fields and parcels with similar spectral appearance, crop context, and spatial consistency to generate additional reference samples. Field parcel boundaries were used to guide sample delineation and to avoid mixed pixels near roads, field edges, settlements, water bodies, and other heterogeneous surfaces. Samples with ambiguous colour signals, cloud or shadow contamination, snow cover, mixed land-cover conditions, or inconsistent appearance across the field parcel were excluded. This procedure produced an augmented validation sample set that retained direct linkage to field observations while increasing the spatial coverage and class diversity needed for regional accuracy assessment.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

The augmented validation samples were organized using Military Grid Reference System (MGRS) tiles as the spatial units for both sample interpretation and map evaluation. In total, 110 MGRS tiles were selected across the global Mollisol regions (**Figure 9**), with their spatial distribution covering the main cropland-concentrated areas in each region to provide a representative sampling frame. For each tile, available non-growing season (NGS) Sentinel-2 imagery was collected and screened using strict quality-control criteria, including (i) cloud cover <20% and (ii) complete spatial coverage of the tile (NODATA_PIXEL_PERCENTAGE = 0). Random samples were then generated within these units, manually labelled, and cross-checked through visual interpretation according to SWIR2-NIR-Red composites (**Figure 8**). Two independent teams evaluated the samples separately. Samples with inconsistent labels were reviewed by a third interpreter, whose decision was used as the final reference label. Across 2019-2024, the validation dataset comprised 158,000 reference samples. The annual validation coverage included 76 tiles and 23,000 samples in 2019, 84 tiles and 26,000 samples in 2020, 82 tiles and 27,000 samples in 2021, 72 tiles and 23,200 samples in 2022, 84 tiles and 29,800 samples in 2023, and 88 tiles and 29,000 samples in 2024. These tile-based samples were used to calculate accuracy metrics for residue-cover maps.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

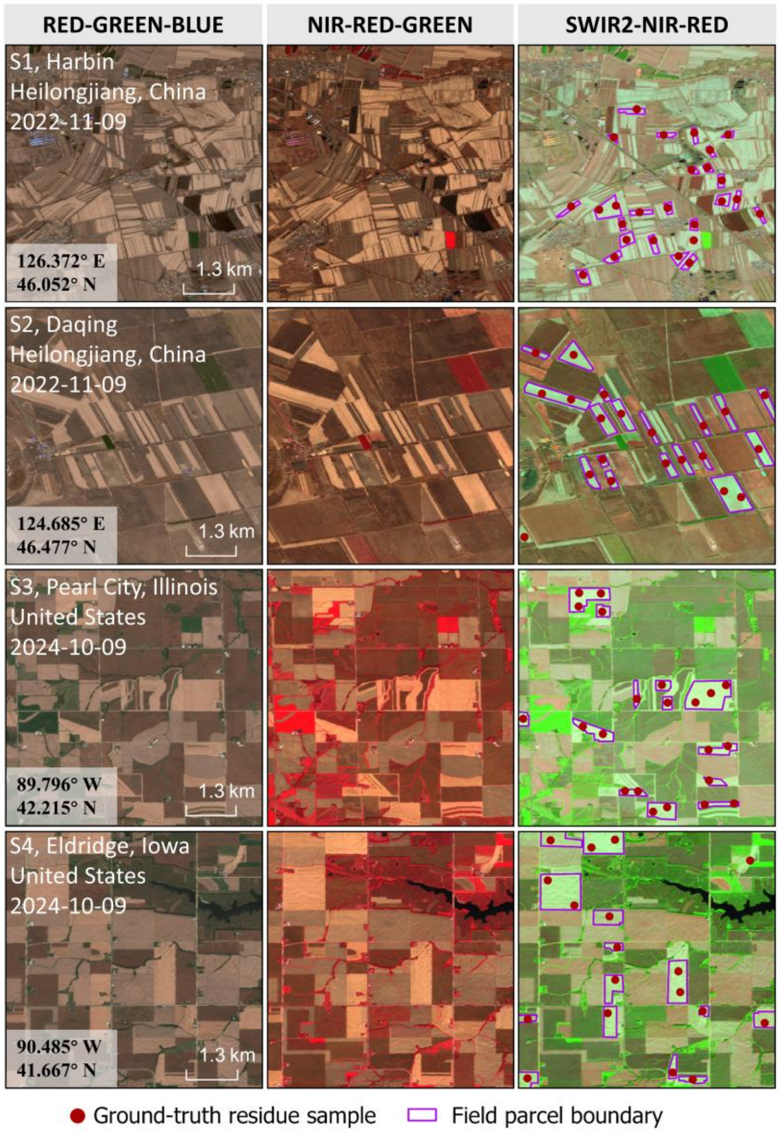


Figure 8 Typical images of the four study sites composed using different spectral band combinations for visual interpretation and sample augmentation. The three columns show Red-Green-Blue, NIR-Red-Green, and SWIR2-NIR-Red composite images, respectively. Red points indicate ground-truth residue samples, and purple outlines indicate field parcel boundaries. Compared with the true-colour and NIR-Red-Green composites, the SWIR2-NIR-Red composite enhances the visual contrast between crop residue-covered fields and other non-growing-season cropland surfaces, providing the primary basis for residue sample augmentation.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

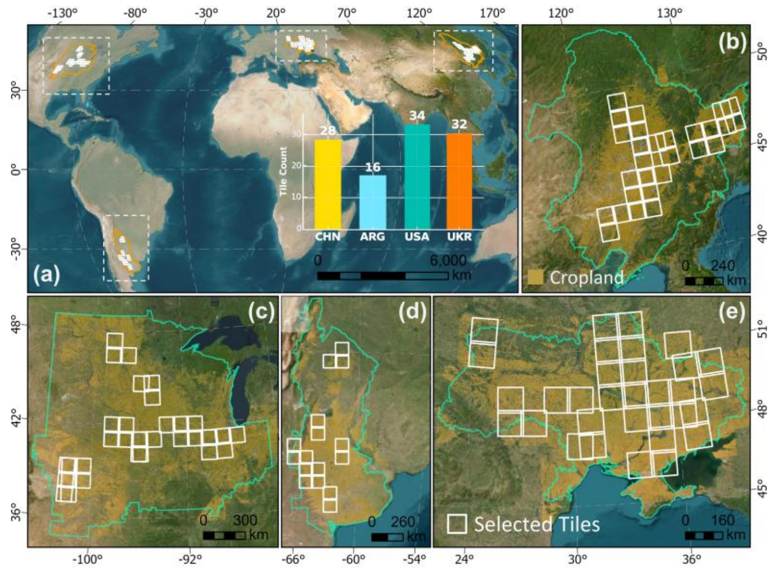


Figure 9 Spatial distribution of validation samples. Panels (a–e) show the geographic distribution of selected tiles used for validation across global Mollisol regions: (a) global overview with tile counts by region, and regional zoom-ins for (b) Northeast China, (c) U.S. Midwest, (d) Argentine Pampas, and (e) Ukraine.

3.3.3 Accuracy metrics

Classification accuracy was assessed using confusion matrices constructed from the validation samples (**Table 3**). Five metrics were calculated to evaluate the performance of CrRUC-M: overall accuracy (OA), omission error (OE), commission error (CE), F1 score, and relative bias (relB). Overall Accuracy (OA) was used to assess the overall agreement between the mapped results and the reference samples. Omission Error (OE) and Commission Error (CE) were calculated to quantify underestimation and overestimation errors for the residue class, respectively. Given the potential class imbalance between residue-covered and non-covered samples, the F1-score was further introduced to provide a balanced measure of classification performance by integrating precision and recall. Additionally, to quantify systematic bias in area estimation, the relative bias (relB) was calculated following Padilla et al. (2015), assessing the degree of overestimation or underestimation of the area of crop residue cover relative to the reference samples. The definitions of these metrics are provided in Eqs. (7)–(11).

$$\mathrm{OA} = \frac{(TP + TN)}{(TP + FN + FP + TN)} \tag{7}$$

$$OE = \frac{FN}{(TP + FN)} \tag{8}$$

$$CE = \frac{FP}{(TP + FP)} \tag{9}$$

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Open Access Earth System Science Data Discussions

$$F1 = \frac{2TP}{2TP + FP + FN} \tag{10}$$

$$\mathrm{relB} = \frac{FP - FN}{TP + FN} \tag{11}$$

Table 3 Confusion matrix of residue datasets

	Validation Class		
Predicted Class	Residue	Non-Residue	Row Total
Residue Cover	TP	FP	TP + FP
Non-Residue	FN	TN	FN + TN
Column Total	TP + FN	FP + TN	TP + FP + FN + TN

4 Results

4.1 Examples of crop residue cover mapping across global Mollisol

The CrRUC-M dataset reveals pronounced spatial heterogeneity in crop residue cover across global Mollisol regions (**Figure 10**). Representative subsets show clear regional contrasts in field shape, parcel size, and residue distribution patterns. In the U.S. Midwest, residue cover occurs in relatively regular and large rectangular parcels (**Figure 10**g-h). In Northeast China, narrower and more fragmented field structures produce elongated and strip-like residue patterns (**Figure 10**d-f). The Ukrainian Mollisol region is characterized by large and consolidated fields with comparatively uniform residue distribution (**Figure 10**i-j). By contrast, the Argentine Pampas shows mixed parcel geometries and more spatially variable residue-cover patterns (**Figure 10**a-c). Across these regions, residue-covered and non-residue cropland pixels are interspersed within intensively cultivated landscapes, indicating substantial within-region variability in post-harvest management. These spatial patterns demonstrate that CrRUC-M can resolve field-scale variability in crop residue cover and provide spatially explicit information across continental Mollisol regions.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

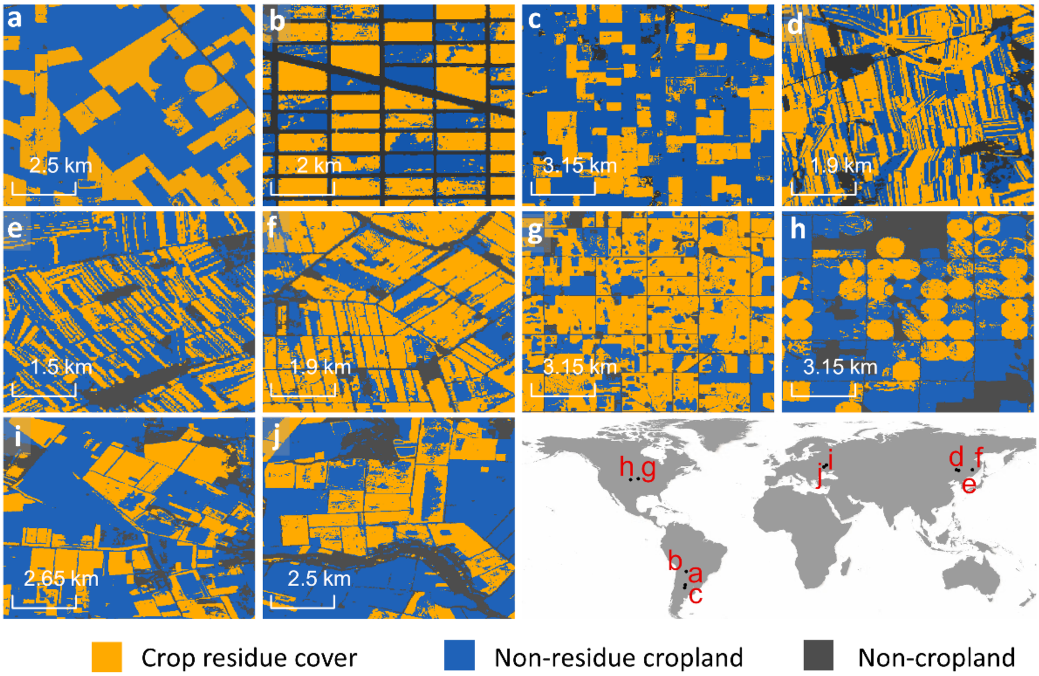


Figure 10 Examples from the CrRUC-M dataset illustrating spatial patterns of crop residue cover across global Mollisol regions. The inset map indicates the geographic distribution of the example locations worldwide.

4.2 Accuracy assessment

CrRUC-M achieved stable classification performance across the global Mollisol validation samples, with a mean OA of 0.830 and F1 score of 0.806 during 2019-2024 (**Figure 11**). Across all region-year estimates, omission error (OE = 0.268) exceeded commission error (CE = 0.082), indicating that classification uncertainty was mainly associated with missed crop residue-covered pixels rather than false detections. The negative relative bias (relB = -0.195) further showed an overall underestimation of residue cover extent. Regional performance varied moderately among the four Mollisol regions. Argentine Pampas showed the highest mean OA (0.841), while the U.S. Midwest had a comparable F1 score (0.820) and the smallest negative bias (relB = -0.123). Northeast China showed slightly lower accuracy, with a mean OA of 0.819 and F1 score of 0.796, together with relatively high OE (0.270). Ukraine had intermediate OA (0.827) but the lowest F1 score (0.788), the highest OE (0.330), and the strongest negative bias (relB = -0.312). In contrast, Ukraine also showed the lowest CE (0.025), indicating that its errors were dominated by omission rather than commission. Overall, CrRUC-M maintained consistent accuracy across major global Mollisol regions, although the remaining uncertainty was primarily driven by omission errors and systematic underestimation of crop residue cover.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

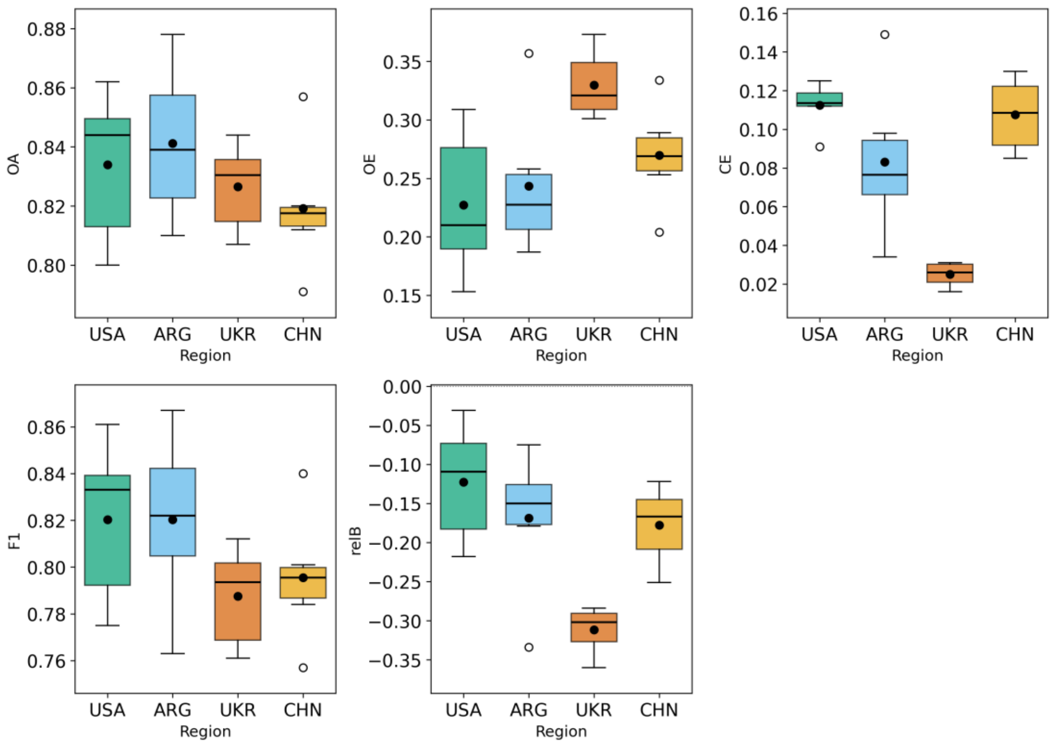


Figure 11 Average classification accuracy across the four Mollisol regions: the U.S. Midwest (USA), Argentine Pampas (ARG), Ukraine (UKR), and Northeast China (CHN). The performance metrics displayed include OA (Overall Accuracy), OE (Omission Error), CE (Commission Error), F1 (F1-score), and relB (Relative Bias).

4.3 Product intercomparison

4.3.1 Comparison with the Corn Residue Coverage dataset in Northeast China

The CrRUC-M datasets were compared with the Corn Residue Coverage (CRC) product for the Northeast China (Dong et al., 2024). Following the CRC definition of conservation tillage, pixels with residue cover greater than 30 % were extracted from the 2021 spring CRC data and denoted as CRC_gt_30. These data were compared with CrRUC-M across six representative ecological sub-regions (**Figure 12**, a1–f1), including the Songnen Plain (SLP), Lesser Khingan Mountains-Changbai Mountains (LK), and Sanjiang Plain (SP).

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

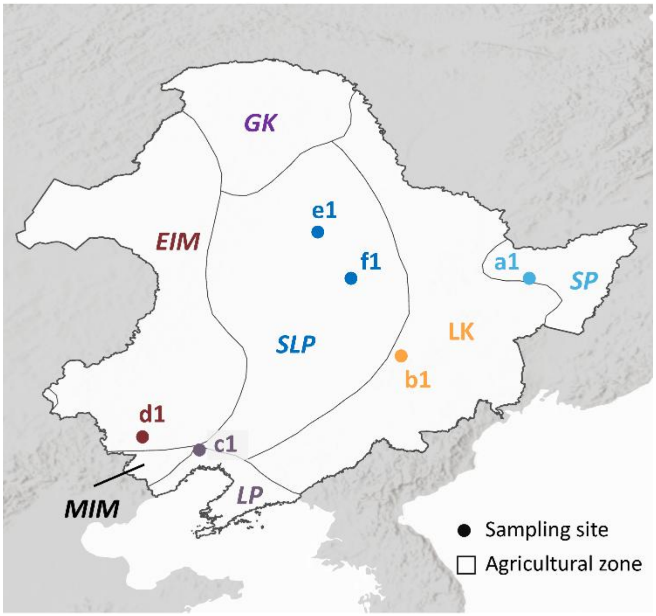


Figure 12 Agricultural zones and sampling site distribution in Northeast China. The study area is divided into several agricultural zones: GK (Northern Greater Khingan Mountains), SP (Sanjiang Plain), LK (Lesser Khingan-Changbai Mountains), SLP (Songliao Plain), EIM (Eastern Inner Mongolia), MIM (Middle Inner Mongolia), and LP (Liaodong-Shandong Peninsula). Colored dots labeled a1–f1 indicate the locations of sampling sites across different zones, specifically within SP (a1), LK (b1), LP (c1), MIM (d1), and SLP (e1, f1).

CrRUC-M and CRC both capture broad regional patterns of crop residue distribution in Northeast China, but they differ in spatial detail, the presence of Landsat 7 ETM+ SLC-off striping artifacts, and consistency between product forms or sensors (**Figure 13**). First, CrRUC-M preserves more detailed field boundaries and within-field variability. For example, the Sentinel-2-derived CrRUC-M product (CrRUC-S2; **Figure 13**e3 and f3) resolves strip-like variations and small residue patches within elongated fields, whereas the corresponding CRC_gt_30 maps (**Figure 13**e4 and f4) classify larger areas as exceeding the 30 % threshold, resulting in more spatially generalized residue patterns. Similar differences are observed in **Figure 13**a3 and c3, where CrRUC-M retains fine-textured features and clearer plot boundaries, while the CRC product appears as smoother clusters. Together, these examples show that CrRUC-M retained finer local-scale spatial variability than CRC in the selected sub-regions. Second, CrRUC-M shows fewer striping and tiling artifacts in the selected examples. In Figure 13a2 and b2, the Landsat-derived CrRUC-M product (CrRUC-LS) presents relatively continuous spatial patterns. By contrast, the corresponding CRC maps (**Figure 13**a5 and b5) show pronounced linear artifacts along orbital tracks, which locally interrupt the mapped residue-cover patterns. CrRUC-LS also maintains stronger continuity across mosaic boundaries, as shown in **Figure 13**d2, whereas the CRC product shows visible block-like discontinuities. Third, CrRUC-M shows consistent spatial patterns between the Sentinel-2 and Landsat products. Despite differences in spatial resolution and spectral response, CrRUC-S2 and CrRUC-LS show similar spatial patterns in the same regions, particularly in **Figure 13**c2-c3 and d2-d3. In contrast, the CRC continuous residue-cover maps and the threshold-derived binary maps differ in spatial structure in some examples. For

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

instance, **Figure 13**c4 and d4 show abrupt transitions and blocky field-scale patterns, whereas the corresponding continuous CRC maps in **Figure 13**c5 and d5 present smoother spatial gradients.

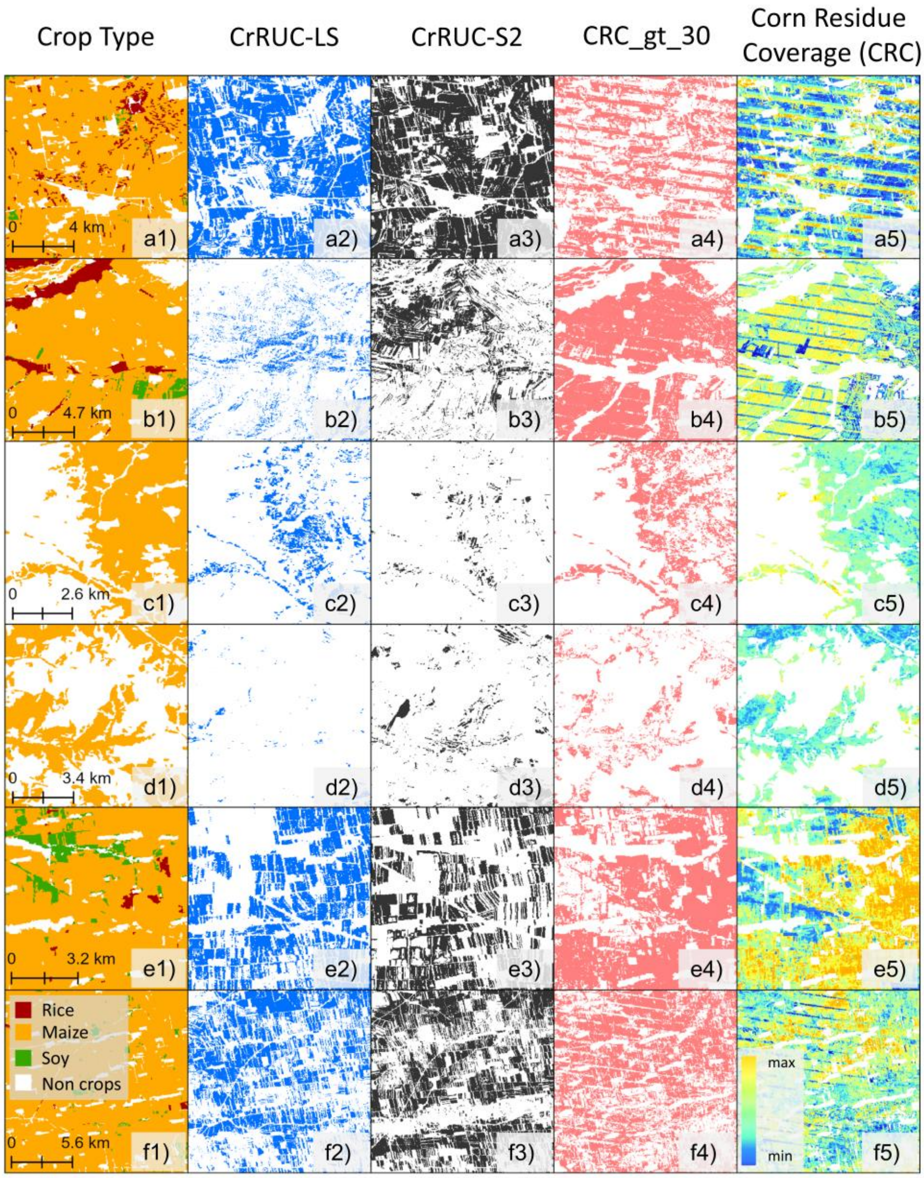


Figure 13 Comparison of our dataset with CRC dataset; the distribution of sample (a1–f1) are shown in Figure 12.

470 **4.3.2 Comparison with USDA-NASS Survey data**

CrRUC-M and the USDA-NASS survey data show broadly consistent state-level patterns of residue-related tillage adoption across the selected states in the U.S. Midwest Mollisol region (**Figure 14**). The two datasets identify a similar spatial hierarchy

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

among states, indicating that CrRUC-M captures the main regional gradients in conservation-practice prevalence despite differences in data source, observation target and reporting unit. For no-till practices, CrRUC-M agrees well with NASS in several major Corn Belt states, including Iowa, Illinois, Missouri and Minnesota, where both datasets indicate relatively comparable adoption levels (**Figure 14**a). Larger discrepancies are observed in Nebraska, Kansas, Colorado, South Dakota and Indiana, suggesting that the agreement between satellite-derived residue cover and survey-reported no-till adoption varies across agroecological and management contexts. For conservation tillage excluding no-till, CrRUC-M also reproduces the broad regional pattern reported by NASS, but the estimated proportions are generally lower than the survey-based values across most states (**Figure 14**b). This consistent offset indicates that CrRUC-M provides a more conservative estimate of residue-related conservation practices. The difference is expected because NASS represents farmer-declared tillage categories, whereas CrRUC-M identifies surface residue cover that is directly observable from satellite imagery during the non-growing season. Thus, fields reported as conservation tillage may not always retain sufficient, spatially continuous or persistent residue cover to be detected by CrRUC-M at the time of observation. Overall, the comparison demonstrates that CrRUC-M can effectively represent the state-level distribution and relative ranking of residue-related conservation practices in the Midwest Mollisol region, while offering an observation-based estimate that is more directly tied to actual surface residue conditions.

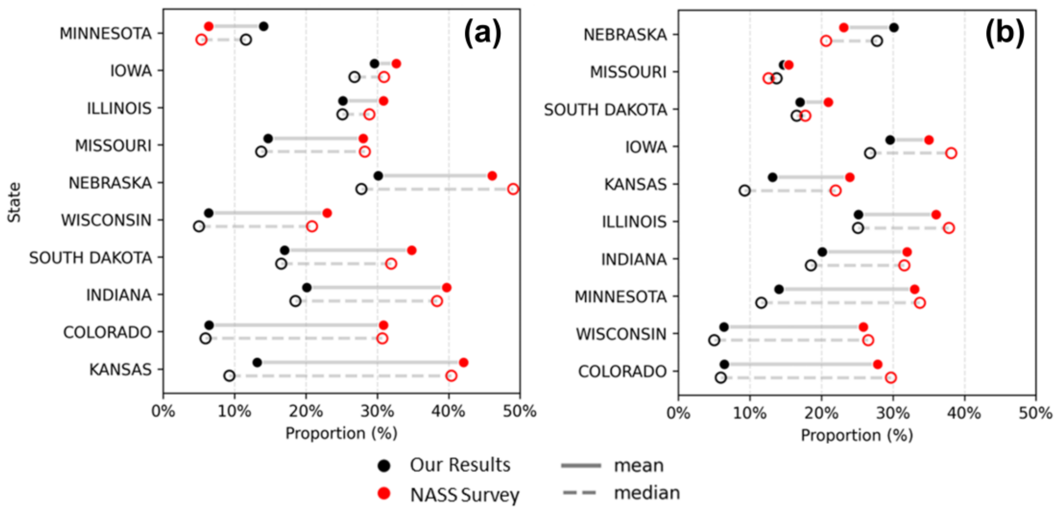


Figure 14 Comparison of crop residue proportions in ten states of U.S. Midwest Corn Belt based on CrRUC-M and NASS survey data. (a) represents no-till area proportions, and (b) represent conservation tillage (excluding no-till).

4.4 Multi-year frequency and crop source composition of residue cover

Spatial distribution of crop residue cover frequency (2019–2024) across the global Mollisol regions reveals a distinct core–periphery gradient, where persistent residue cover (4–6 years) is concentrated in intensive agricultural hubs compared to lower temporal consistency in peripheral zones (**Figure 15**). In Northeast China, high-frequency pixels are mainly concentrated in the central Songnen and Sanjiang Plains, especially in northern Heilongjiang and western Jilin, while southern Liaoning and

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

marginal areas show lower frequencies. The U.S. Midwest shows widespread medium- to high-frequency cover across the Corn Belt (Iowa, Illinois, and Indiana), with values diminishing toward the southern and western margins. Similarly, in the Argentine Pampas, high-frequency occurrences are localized in northern Buenos Aires, Córdoba, and Santa Fe, declining toward the regional boundaries. Conversely, Ukraine displays the lowest overall frequency, with most cropland recorded at three or fewer occurrences and high-frequency pixels appearing only sporadically within isolated high-productivity zones.

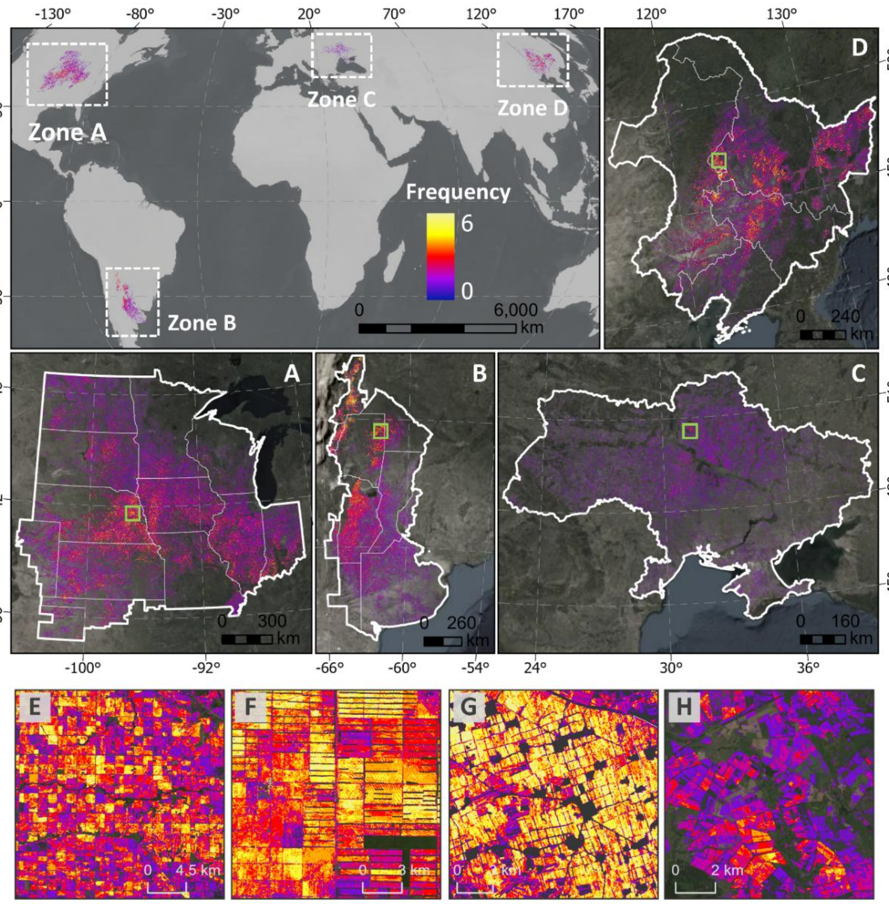


Figure 15 Spatial distribution of crop residue cover frequency (2019–2024) across global Mollisol regions. Panels A–D present regional maps for the U.S. Midwest, Argentine Pampas, Ukraine, and Northeast China, respectively. Green boxes mark subregions enlarged in Panels E–H, where E corresponds to A (U.S.), F to B (Argentine Pampas), G to D (China), and H to C (Ukraine), highlighting field-scale heterogeneity in residue cover patterns.

The source of crop residue across the Mollisol regions aligns closely with regional cropping systems, with corn consistently serving as the primary contributor despite varying secondary crop structures. In the Argentine Pampas (**Figure 16**a), residue production increasingly concentrates in corn and soybean, evidenced by the corn share rising from 35% in 2019 to 62% in 2024 alongside a substantial decline in "other" crops. The U.S. Midwest (**Figure 16**b) exhibits the highest concentration of

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Open Access Earth System Science Data Discussions

residue sources derived almost exclusively from the corn–soybean rotation; here, the soybean contribution expanded from 9% to 30% as the corn share adjusted from 80% to 64%. Northeast China (**Figure 16**d), maintains a more stable and diversified residue structure, where corn consistently accounts for 54%–62%, followed by rice (14%–24%) and a moderate share of soybean. Conversely, Ukraine displays the most complex and variable composition, with corn as the largest but fluctuating contributor, while wheat has emerged as the primary complement to corn after 2021 (**Figure 16**c).

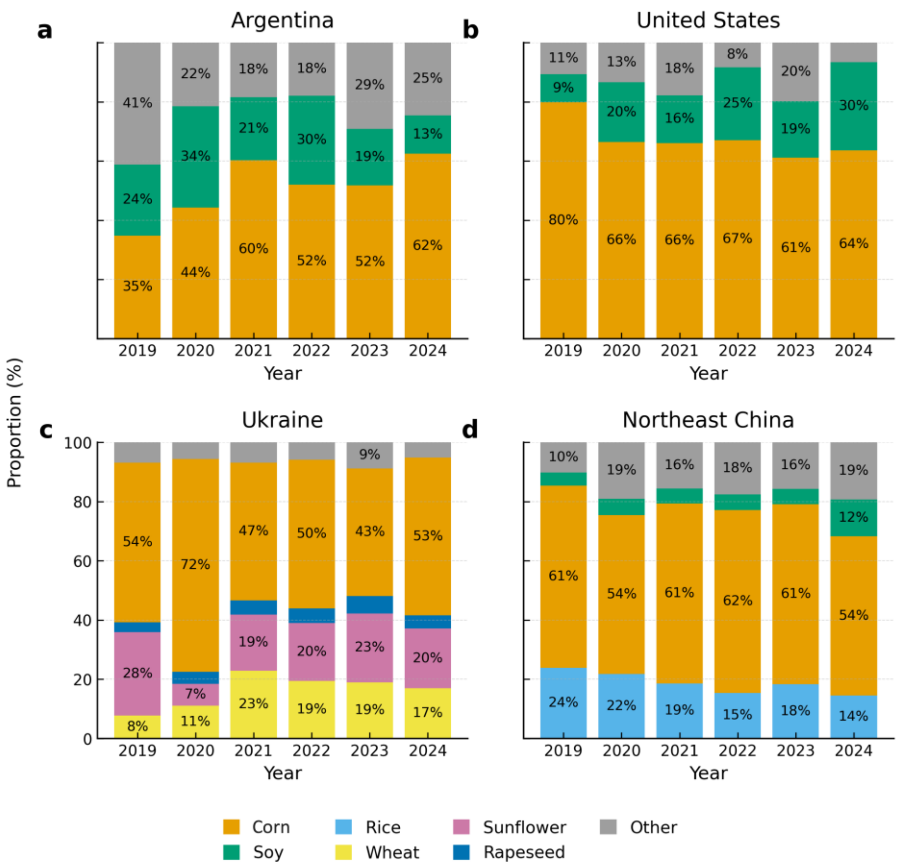


Figure 16 Crop-type composition and proportional contributions of residue sources across global Mollisol regions from 2019 to 2024. Panels show (a) Argentine Pampas (ARG), (b) U.S. Midwest (USA), (c) Ukraine (UKR), and (d) Northeast China (CHN).

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

5 Discussion

5.1 Advances and contributions

Crop residue cover represents a seasonal and management-driven land-surface condition that occurs primarily during the non-growing season, between two crop growth cycles. Unlike static land-cover classes (e.g., cropland or forest), residue cover exhibits high variability dictated by shifting agricultural practices. However, existing land-use and land-cover (LULC) datasets and crop-type maps primarily delineate cropland distribution or crop categories and do not explicitly characterize post-harvest surface conditions (Lesiv et al., 2025; Zhang et al., 2025a). As a result, the spatial and temporal dynamics of residue management remain largely unrepresented in conventional global land datasets. In this context, the CrRUC-M dataset provides a spatially explicit, wall-to-wall crop residue cover product for 2019–2024 across the global Mollisol regions. The wall-to-wall mapping strategy enables consistent, annual monitoring across entire agricultural landscapes, supporting cross-regional comparison and dynamic assessment of management practices. Compared with model-based simulations, statistical data spatialization, or meta-analytical synthesis approaches, remote-sensing-derived wall-to-wall mapping offers direct observational evidence with spatial continuity, pixel-level detail, and year-to-year comparability. This observational consistency is particularly important for capturing heterogeneous residue patterns that vary across fields, regions, and years.

By combining 10 m Sentinel-2 imagery with dense temporal observations, CrRUC-M resolves field-scale residue patterns in fragmented agricultural landscapes, notably in Northeast China (**Figure 17**). The framework is simple and interpretable, using prior knowledge of crop phenology, post-harvest surface dynamics, and residue spectral behaviour to guide image compositing and residue extraction, rather than relying on large training datasets or complex black-box models. This knowledge-guided design improves robustness across regions and years while maintaining scalability and reproducibility. Overall, CrRUC-M provides a high-resolution observational benchmark for crop residue cover and a practical foundation for future transfer to other satellite archives.

5.2 Extension of the framework to Landsat archives

To evaluate the transferability and robustness of the proposed framework, the same mapping strategy was further applied to Landsat series imagery. Because Landsat provides a substantially longer observation record than Sentinel-2, this extension is important for exploring the potential of long-term dynamic monitoring of crop residue cover. At the 1° grid scale, Landsat-derived residue-cover area showed generally strong agreement with Sentinel-2 estimates across all Mollisol regions (**Figure 17**). The strongest consistency was observed in Argentine Pampas ($R^2 = 0.935$), followed by Northeast China ($R^2 = 0.842$), Ukraine ($R^2 = 0.813$), and the USA ($R^2 = 0.810$). Similar relationships were observed across individual years, indicating that the framework retained general spatial and temporal consistency after transferring to Landsat archives (Fig.S9). However, Landsat-derived estimates were systematically lower than Sentinel-2 estimates in all regions, with fitted slopes consistently

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

below the 1:1 line. Independent accuracy assessment further confirmed this tendency. Compared with the Sentinel-2-based product, the Landsat-based product showed lower classification performance across the four Mollisol region (**Table S5**). Averaged across the four Mollisol regions, the Sentinel-2-based product achieved an OA of 0.83 and an F1 score of 0.81, whereas the Landsat-based product achieved an OA of 0.74 and an F1 score of 0.68. The Landsat-based product also had a much higher OE (0.43) and stronger negative bias (relB = -0.35) than the four-region mean of the Sentinel-2-based product (OE = 0.27; relB = -0.20), indicating stronger omission-related underestimation of residue cover extent (Fig.S10). These differences are mainly associated with the coarser spatial resolution and lower revisit frequency of Landsat (Liepa et al., 2024). Sentinel-2 provides denser clear-sky observations during the non-growing season, improving compositing stability and increasing the chance of capturing the short residue-exposure period after harvest (Zhou et al., 2026). In contrast, the longer revisit interval of Landsat increases the risk of temporal mismatch. In some cases, Landsat may capture only the pre-harvest state and the post-tillage state, while missing the intermediate residue-covered stage entirely. This issue is particularly important in regions such as the U.S. Mollisol belt and Northeast China, where fields can be rapidly tilled, ridged, or planted with cover crops soon after harvest. Coarser spatial resolution further increases the omission of small or fragmented residue patches. These factors explain the stronger negative relB and higher OE observed in the Landsat-derived product. Nevertheless, the overall cross-sensor consistency suggests that the proposed framework remains transferable to historical Landsat archives and provides a feasible basis for future long-term residue-cover monitoring.

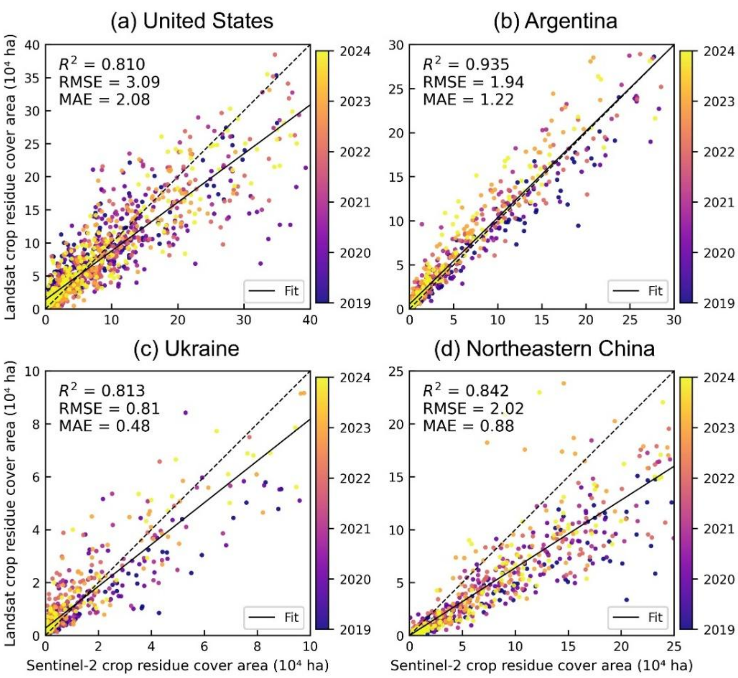


Figure 17 Regional comparison of Sentinel-2-derived crop residue cover area against Landsat estimates from 2019 to 2024 at grid scales. Panels represent (a) the U.S. Midwest, (b) Argentine Pampas, (c) Ukraine, and (d) Northeast China. Year is encoded by a continuous color

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

scale. The dashed line represents the 1:1 reference and the solid line indicates the linear regression fit. Statistical metrics (R^2, RMSE, MAE) quantify cross-sensor consistency. Areas are reported in 10^4 ha.

5.3 Uncertainty and limitations

Several uncertainties remain inherent in satellite-based monitoring of crop residue cover. The reliance on optical imagery introduces constraints related to atmospheric interference and observation frequency. In high-latitude or humid Mollisol regions, particularly the Ukrainian and Northeast China Mollisol region, persistent cloud cover, seasonal precipitation, and early snow events can substantially narrow the effective post-harvest observation window and increase uncertainty in the day-of-year (DOY) timing of residue exposure. Although multi-temporal compositing reduces data gaps, uneven cloud-free acquisitions may still cause localized detection instability. Such limitations are intrinsic to wall-to-wall optical monitoring, where temporal coverage and data quality must be carefully balanced. Harmonized Landsat and Sentinel-2 (HLS) archives (Gao et al., 2023) provide a promising direction for reducing these uncertainties. By integrating observations from multiple sensors within a common spatial and radiometric framework, HLS can increase the density of usable observations during the non-growing season and improve the chance of capturing short-lived residue exposure after harvest. This may be particularly valuable in regions where residue cover changes rapidly due to tillage, ridging, residue removal, or cover-crop establishment. Future applications of HLS could therefore improve temporal continuity, reduce data gaps, and support longer-term monitoring of crop residue cover beyond the Sentinel-2 era.

A conceptual trade-off also arises from adopting a binary classification framework instead of fractional residue coverage mapping. While fractional residue coverage provides more detailed surface characterization, its large-scale implementation is hindered by scale mismatches between sub-meter field measurements, such as line-transects, and the integrated 10 to 30 m reflectance of satellite pixels (Dong et al., 2024). This discrepancy complicates calibration and validation at continental scales. By focusing on the presence or absence of residue cover, CrRUC-M prioritizes methodological robustness and cross-regional comparability. Moreover, residue cover represents a measurable surface condition, whereas agricultural statistics and tillage classifications are typically defined by management intent. A single coverage threshold, for example 30 percent, may be achieved through diverse tillage strategies, meaning spectral detection does not directly correspond to standardized management categories. Representing residue cover as a discrete variable may therefore limit its use in high-precision modelling applications, such as soil carbon or erosion simulations (Anand et al., 2022). Nevertheless, the binary framework provides a scalable and internally consistent basis for large-area monitoring under heterogeneous management conditions.

Finally, continental-scale validation remains constrained by the scarcity of harmonized reference data. Unlike cropland extent or crop type, residue management lacks standardized reporting systems and is rarely recorded in official statistics. Field surveys are labour-intensive, spatially sparse, and often rely on inconsistent definitions across regions(Shuai et al., 2025). Although this study incorporated the public and regional datasets currently available, the absence of a globally consistent ground

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

reference network remains a fundamental limitation for verifying residue cover dynamics across diverse pedo-climatic backgrounds.

5.4 Implications and potential applications

The CrRUC-M dataset provides a harmonized observational basis for assessing crop residue management across the global Mollisol regions. By delivering annual wall-to-wall maps under a unified framework, it enables cross-regional comparison of residue adoption patterns and interannual variability. The dataset can complement survey-based or administrative statistics by providing spatially explicit evidence of surface residue conditions, thereby supporting the evaluation of conservation tillage implementation and examining consistency between reported practices and observed land-surface states. CrRUC-M also provides a useful spatial constraint for process-based agroecosystem models. In particular, uncertainty in biogeochemical simulations often arises from limited information on the fraction of residue retained on the field versus removed, which strongly influences carbon inputs, soil cover, and associated estimates of greenhouse gas emissions (Cao et al., 2026), soil organic carbon dynamics (Zhang et al., 2024), and crop productivity (Song et al., 2026). Incorporating the high-resolution residue cover information from CrRUC-M can therefore reduce uncertainties in model parameterization and improve the representation of management-induced biogeophysical and biogeochemical processes. In addition, crop residues represent an important source of agricultural biomass with potential for bioenergy production (Zhao et al., 2024). The spatially explicit information provided by CrRUC-M can support the quantification of residue availability and its spatiotemporal variability, thereby contributing to improved estimates of biomass energy supply and associated carbon fluxes (Mccarl, 2025). Furthermore, by identifying areas of residue presence, the dataset can be integrated with satellite-derived fire detection products to examine spatial linkages between residue availability and agricultural burning, providing a spatial basis for refining emission-related assessments (Shang et al., 2025). More broadly, CrRUC-M complements conventional land-use and crop-type datasets by representing a dynamic, management-driven surface condition that is typically absent from standard classifications. The demonstrated consistency between Landsat- and Sentinel-2-based products establishes a scalable foundation for future temporal extension and potential adaptation to other soil regions. As continued satellite observations accumulate, the framework presented here can support expanded monitoring of residue management under changing climatic and agricultural conditions.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

6 Conclusions

This study presents CrRUC-M, a 10 m Sentinel-2-based crop residue cover dataset covering major global Mollisol regions from 2019 to 2024. The dataset is generated based on a simple and interpretable knowledge-guided framework that integrates non-growing-season compositing, residue-sensitive spectral information, and adaptive thresholding. Validation using field-anchored reference samples generated through independent visual interpretation showed stable performance across the four global Mollisol regions, including the U.S. Midwest, Northeast China, Ukraine and Argentine Pampas, with mean overall accuracy and F1-score values of 0.83 and 0.81, respectively. The framework also showed transferability to Landsat archives, indicating its potential for extending crop residue cover monitoring to longer historical periods. Spatiotemporal analysis revealed a clear core-periphery gradient in residue-cover frequency, with persistent high-frequency clusters in intensively cultivated areas and lower temporal consistency in marginal regions. Corn was the dominant source of crop residue across the mapped regions, whereas the contribution of secondary crops varied according to regional cropping systems. By providing consistent annual information on crop residue cover in globally important Mollisol regions, CrRUC-M supports the monitoring of conservation agriculture, assessment of land management practices, and parameterization or evaluation of biogeophysical and land-surface models.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Open Access Earth System Science Data Discussions

Data availability

The CrRUC-M dataset, containing annual crop residue cover data for global Mollisol regions from 2019 to 2024, is publicly available from Zenodo at https://doi.org/10.5281/zenodo.18773303 (Cui et al., 2025). The dataset covers the global Mollisol regions considered in this study, including the U.S. Midwest, Northeast China, Ukraine, and the Argentine Pampas. The annual crop residue cover layers are provided as GeoTIFF files in WGS84 geographic coordinates (EPSG:4326). Files are organized by Mollisol region and year and follow the naming convention CrRUC_Mollisol_s2_[REGION]_[YEAR]_v1.tif. Mollisol boundary data are provided as ESRI Shapefiles in WGS84. These boundary files can be opened in standard GIS software and read programmatically using common geospatial libraries such as geopandas.read_file in Python or sf::st_read in R.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

Author contribution

YC: Conceptualization, Methodology, Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. JD: Supervision, Conceptualization, Writing – review & editing. CZ: Supervision, Writing – review & editing. NY: Writing – review & editing. SR: Writing – review & editing. PZ: Writing – review & editing. QD: Methodology. RL: Supervision. CT: Supervision, review & editing.

Competing interests

Editorship CI! The authors have the following competing interests: At least one of the (co-)authors is a member of the editorial board of Earth System Science Data.

Acknowledgements

This study was supported by the National Key Research and Development Program of China (2024YFD1500903), the National Key Research and Development Program of China (2023YFD1500200), the National Natural Science Foundation of China (42271375, 42461144212, 42525108, and 72221002) and the Abdul Latif Jameel Water and Food Systems Lab (J-WAFS) grant at Massachusetts Institute of Technology (MIT).

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

References

Anand, A., Kumar, V., and Kaushal, P.: Biochar and its twin benefits: Crop residue management and climate change mitigation in India, Renewable & Sustainable Energy Reviews, 156, 111959, 10.1016/j.rser.2021.111959, 2022.

Azzari, G., Grassini, P., Edreira, J. I. R., Conley, S., Mourtzinis, S., and Lobell, D. B.: Satellite mapping of tillage practices in the North Central US region from 2005 to 2016, Remote Sens. Environ., 221, 417-429, 10.1016/j.rse.2018.11.010, 2019.

Bao, Y., Meng, X., Liu, H., Wang, M., Yao, F., and Mouazen, A. M.: Adjusting the main cropping types in Mollisol regions could improve the net primary productivity of low-producing areas by 20%–30% under future climate change, Earth's Future, 13, e2025EF006074, 10.1029/2025EF006074, 2025.

Bonney, M. T. and Zhang, Y.: Monitoring snow cover dynamics at 30-m resolution in higher latitude regions using Harmonized Landsat Sentinel-2, ISPRS J. Photogramm. Remote Sens., 233, 360-382, 10.1016/j.isprsjprs.2026.01.032, 2026.

Boryan, C., Yang, Z., Mueller, R., and Craig, M.: Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program, Geocarto Int., 26, 341-358, 10.1080/10106049.2011.562309, 2011.

Cao, P., Bilotto, F., Gonzalez Fischer, C., Mueller, N. D., Carlson, K. M., Driscoll, A. W., Gerber, J. S., Smith, P., Tubiello, F. N., West, P. C., You, L., and Herrero, M.: Spatially explicit global assessment of cropland greenhouse gas emissions circa 2020, Nat. Clim. Change, 16, 1-10, 10.1038/s41558-026-02558-4, 2026.

Chen, B., Tu, Y., An, J., Wu, S., Lin, C., and Gong, P.: Quantification of losses in agriculture production in eastern Ukraine due to the Russia-Ukraine war, Commun. Earth Environ., 5, 336, 10.1038/s43247-024-01488-3, 2024.

Chen, S., Zhuang, Q., Taheripour, F., Yuan, Y., and Benavidez, L.: Assessment of global land cover changes using satellite data: intermittent and long-term land cover changes from 2001 to 2020, Environ. Res. Lett., 20, 034045, 10.1088/1748-9326/adb5a3, 2025.

Cui, Y., Dong, J., Chao, Z., Nanshan, Y., Shuai, R., Peng, Z., Quan, D., Cesar, T., and Ronggao, L.: CrRUC-M: Crop residue cover datasets across global Mollisol regions (2019 to 2024), Zenodo [dataset], 10.5281/zenodo.18773303, 2025.

Cui, Y., Liu, R., Li, Z., Zhang, C., Song, X.-P., Yang, J., Yu, L., Chen, M., and Dong, J.: Decoding the inconsistency of six cropland maps in China, Crop J., 12, 281-294, 10.1016/j.cj.2023.11.011, 2024a.

Cui, Y., Dong, J., Zhang, C., Yang, J., Chen, N., Guo, P., Di, Y., Chen, M., Li, A., and Liu, R.: Validation and refinement of cropland map in southwestern China by harnessing ten contemporary datasets, Sci. Data, 11, 671, 10.1038/s41597-024-03508-5, 2024b.

de Abelleyra, D., Iturralde Elortegui, M. d. R., Zelaya, K., Portillo, J., Melilli, M., Franzoni, A., Volante, J., Goytia, Y., Kurtz, D., Ybarra, D., Barrios, R., Benedetti, P., Morales, C., Berriolo, J., Lanceta Pereyra, M., Scavone, A., Murray, F., Santillan, J., Continelli, M., Gerlero, G., Salas, D., Reinaldi, J., Lopez Juane, P., Gomez, D., Krapovickas, S., Sapino, V., Regonat, A., Cracogna, M., Espíndola, C., Valiente, S., Parodi, M., Colombo, F., Scarel, J., Ayala, J., Martins, L., Basanta, M., Rausch, A., Almada, G., Boero, L., Calcha, J., Chiavassa, A., Calandroni, M., Murillo, M., Pascale, B., Borrachi, S., Erreguerena, J., Besteiro, I., Oyesqui, L., Lazaeta, M., Pezzola, A., Winschel, C., Rodriguez Pérez, J., German, A., and Viatale, J.: Argentina National Map of Crops 2023/2024, Zenodo [dataset], 10.5281/zenodo.13984185, 2024.

Deines, J. M., Wang, S., and Lobell, D. B.: Satellites reveal a small positive yield effect from conservation tillage across the US Corn Belt, Environ. Res. Lett., 14, 124038, 10.1088/1748-9326/ab503b, 2019.

Dong, Y., Xuan, F., Huang, X., Li, Z., Su, W., Huang, J., Li, X., Tao, W., Liu, H., and Chen, J.: A 30-m annual corn residue coverage dataset from 2013 to 2021 in Northeast China, Sci. Data, 11, 216, 10.1038/s41597-024-02998-7, 2024.

Duan, Q., Liu, R., Chen, J., Wei, X., Liu, Y., and Zou, X.: Burned area detection from a single satellite image using an adaptive thresholds algorithm, Int. J. Digital Earth, 17, 2376275, 10.1080/17538947.2024.2376275, 2024.

Fang, Y. R., Hossain, M. D. S., Chen, Y. S., Guo, L. Y., Wu, X. F., Kong, Q. Y., and Yang, P. J.: Optimizing crop straw utilization for enhancing bioenergy production and sustainable development, Renewable Energy, 256, 124212, 10.1016/j.renene.2025.124212, 2026.

Gao, F., Jennewein, J., Hively, W. D., Soroka, A., Thieme, A., Bradley, D., Keppler, J., Mirsky, S., and Akumaga, U.: Near real-time detection of winter cover crop termination using harmonized Landsat and Sentinel-2 (HLS) to support ecosystem assessment, Sci. Remote Sens., 7, 100073, 10.1016/j.srs.2022.100073, 2023.

Gibbs, H. K. and Salmon, J. M.: Mapping the world's degraded lands, Appl. Geogr., 57, 12-21, 10.1016/j.apgeog.2014.11.024, 2015.

Griffiths, P., Nendel, C., and Hostert, P.: Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping, Remote Sens. Environ., 220, 135-151, 10.1016/j.rse.2018.10.031, 2019.

Holben, B. N.: Characteristics of Maximum-Value Composite Images from Temporal Avhrr Data, Int. J. Remote Sens., 7, 1417-1434, Doi 10.1080/01431168608948945, 1986.

Holden, C. E. and Woodcock, C. E.: An analysis of Landsat 7 and Landsat 8 underflight data and the implications for time series investigations, Remote Sens. Environ., 185, 16-36, 10.1016/j.rse.2016.02.052, 2016.

Hoskovec, L., Ogle, S. M., Breidt, F. J., Keyser, A. R., Liu, T., Spencer, S. L., and Williams, S.: Agricultural Land Management Practices in the Conterminous United States from 1980–2023, Earth Syst. Sci. Data Discuss., 2026, 1-45, 10.5194/essd-2025-842, 2026.

Kassam, A., Friedrich, T., and Derpsch, R.: Global spread of conservation agriculture, Int. J. Environ. Stud., 76, 29-51, 10.1080/00207233.2018.1494927, 2019.

Lesiv, M., Fritz, S., Dürauer, M., Georgieva, I., Buchhorn, M., Bertels, L., Tsendbazar, N., Van De Kerchove, R., Zanaga, D., Schepaschenko, D., See, L., Herold, M., Smets, B., Cherlet, M., Brink, A., and McCallum, I.: A global reference data set for land cover mapping at 10 m resolution, Earth Syst. Sci. Data, 17, 6149-6155, 10.5194/essd-17-6149-2025, 2025.

Liao, X., Li, Y., Wang, J., Dong, J., Yao, Q., Li, J., Hou, R., Li, Z., Zhang, Y., Jia, X., and Leng, G.: Governing black soils for food and climate security, Nat. Commun., 17, 1488, 10.1038/s41467-026-69018-7, 2026.

Liepa, A., Thiel, M., Taubenböck, H., Steffan-Dewenter, I., Abu, I.-O., Singh Dhillon, M., Otte, I., Otim, M. H., Lutaakome, M., Meinhof, D., Martin, E. A., and Ullmann, T.: Harmonized NDVI time-series from Landsat and Sentinel-2 reveal phenological patterns of diverse, small-scale cropping systems in East Africa, Remote Sensing Applications-Society and Environment, 35, 101230, 10.1016/j.rsase.2024.101230, 2024.

Liu, H., Sun, J., Wan, W., Liu, S., Li, J., Lv, W., and Liu, Z.: Impact of conservation tillage on maize growth dynamics in black soils of Northeast China through multi-platform remote sensing, Field Crops Res., 341, 110423, 10.1016/j.fcr.2026.110423, 2026.

Liu, J., Fang, L., Qiu, T., Chen, J., Wang, H., Liu, M., Yi, J., Zhang, H., Wang, C., Sardans, J., Chen, L., Huang, M., and Penuelas, J.: Crop residue return achieves environmental mitigation and enhances grain yield: a global meta-analysis, Agron. Sustainable Dev., 43, 78, 10.1007/s13593-023-00928-2, 2023.

Liu, R.: Compositing the minimum NDVI for MODIS data, IEEE Trans. Geosci. Remote Sens., 55, 1396-1406, 10.1109/TGRS.2016.2623746, 2016.

Liu, X., Lee Burras, C., Kravchenko, Y. S., Duran, A., Huffman, T., Morras, H., Studdert, G., Zhang, X., Cruse, R. M., and Yuan, X.: Overview of Mollisols in the world: Distribution, land use and management, Can. J. Soil Sci., 92, 383-402, 10.4141/Cjss2010-058, 2012.

Liu, Y., Diao, C., Yang, Z., Mei, W., and Guo, T.: A novel Normalized Harvest Phenology Index (NHPI) for corn and soybean harvesting date detection using Landsat and Sentinel-2 imagery on Google Earth Engine, Remote Sens. Environ., 331, 115016, 10.1016/j.rse.2025.115016, 2025.

Macfarlane, M., Jia, R., Fricke, E., Vallicrosa, H., Yu, J., Mirzagholi, L., Coleman, E., Olivetti, E., and Terrer, C.: Nature-Based Climate Solutions: Current Uncertainties and Data Gaps in the Assessment of Soil Carbon Sequestration Potentials, 2024.

McCarl, B. A.: Climate change: What do we do about it? Economic issues regarding agricultural adaptation and mitigation, Am. J. Agric. Econ., 107, 368-389, 10.1111/ajae.12517, 2025.

Meng, X., Bao, Y., Luo, C., Zhang, X., and Liu, H.: SOC content of global Mollisols at a 30 m spatial resolution from 1984 to 2021 generated by the novel ML-CNN prediction model, Remote Sens. Environ., 300, 113911, 10.1016/j.rse.2023.113911, 2024.

Nelson, D.: USDA-National Agricultural Statistics Service, Third National Small Farm Conference, 11, 10.22004/ag.econ.350168,

Otsu, N.: A Threshold Selection Method from Gray-Level Histograms, IEEE Trans. Syst. Man Cybern., 9, 62-66, 10.1109/TSMC.1979.4310076, 1979.

Padilla, M., Stehman, S. V., Ramo, R., Corti, D., Hantson, S., Oliva, P., Alonso-Canas, I., Bradley, A. V., Tansey, K., Mota, B., Pereira, J. M., and Chuvieco, E.: Comparing the accuracies of remote sensing global burned area products using stratified random sampling and estimation, Remote Sens. Environ., 160, 114-121, 10.1016/j.rse.2015.01.005, 2015.

Pasquarella, V. J., Brown, C. F., Czerwinski, W., and Rucklidge, W. J.: Comprehensive quality assessment of optical satellite imagery using weakly supervised video learning, 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 17-24 June 2023, 2125-2135, 10.1109/CVPRW59228.2023.00206,

Porwollik, V., Rolinski, S., Heinke, J., and Müller, C.: Generating a rule-based global gridded tillage dataset, Earth Syst. Sci. Data, 11, 823-843, 10.5194/essd-11-823-2019, 2019.

Potapov, P., Turubanova, S., Hansen, M. C., Tyukavina, A., Zalles, V., Khan, A., Song, X.-P., Pickens, A., Shen, Q., and Cortez, J.: Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century, Nat. Food, 3, 19-28, 10.1038/s43016-021-00429-z, 2022.

Prestele, R., Hirsch, A. L., Davin, E. L., Seneviratne, S. I., and Verburg, P. H.: A spatially explicit representation of conservation agriculture for application in global change studies, Glob Chang Biol, 24, 4038-4053, 10.1111/gcb.14307, 2018.

Qiu, S., Zhu, Z., Olofsson, P., Woodcock, C. E., and Jin, S.: Evaluation of Landsat image compositing algorithms, Remote Sens. Environ., 285, 113375, 10.1016/j.rse.2022.113375, 2023.

Roberts, S. D., Harrington, C. A., and Terry, T. A.: Harvest residue and competing vegetation affect soil moisture, soil temperature, N availability, and Douglas-fir seedling growth, For. Ecol. Manage., 205, 333-350, 10.1016/j.foreco.2004.10.036, 2005.

Shang, Y. Q., Pei, Y. Y., Fu, P., Ren, C. T., Li, Z. C., Ren, J. F., Zheng, X. Q., Di, Y. Y., Zhou, Y., and Dong, J. W.: Increased corn cultivation exacerbated crop residue burning in Northeast China in the 21st century, Geogr. Sustainability, 6, 100244, 10.1016/j.geosus.2024.09.008, 2025.

Shuai, G., Sahajpal, R., Whitcraft, A. K., Presley, D., Graber, R., Schlender, R., and Fross, D.: A tale of two datasets: using survey data and optical data to characterize long-term tillage intensity dynamics in Kansas, Int. J. Appl. Earth Obs. Geoinf., 142, 104725, 10.1016/j.jag.2025.104725, 2025.

Silva, T. S., Malone, L. C., Ruark, M. D., Lee, C. D., Jordan, D., Poffenbarger, H. J., Kandel, H. J., Ross, J., Gaska, J. M., Lauer, J. G., Lindsey, L. E., Singh, M. P., Licht, M. A., Plumblee, M., Vann, R. A., Werle, R., Mourtzinis, S., Naeve, S. L., Roberts, T. L., and Conley, S. P.: Impacts of rotation, tillage, cover cropping, and drainage on soil health in soybean-based cropping systems: Evidence from 4–50-year trials across the US, Agric. Ecosyst. Environ., 395, 109950, 10.1016/j.agee.2025.109950, 2026.

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.

Song, Y., Li, Z., He, X., Zhang, J., Fu, J., Zheng, F., and Li, Z.: Unraveling the regional dynamics of straw mulching and incorporation on crop yields in Northeast China, Eur. J. Agron., 174, 127943, 10.1016/j.eja.2025.127943, 2026.

Stanimirova, R., Tarrio, K., Turlej, K., McAvoy, K., Stonebrook, S., Hu, K.-T., Arévalo, P., Bullock, E. L., Zhang, Y., Woodcock, C. E., Olofsson, P., Zhu, Z., Barber, C. P., Souza, C. M., Chen, S., Wang, J. A., Mensah, F., Calderón-Loor, M., Hadjikakou, M., Bryan, B. A., Graesser, J., Beyene, D. L., Mutasha, B., Siame, S., Siampale, A., and Friedl, M. A.: A global land cover training dataset from 1984 to 2020, Sci. Data, 10, 879, 10.1038/s41597-023-02798-5, 2023.

Teluguntla, P., Thenkabail, P. S., Oliphant, A., Xiong, J., Gumma, M. K., Congalton, R. G., Yadav, K., and Huete, A.: A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform, ISPRS J. Photogramm. Remote Sens., 144, 325-340, 10.1016/j.isprsjprs.2018.07.017, 2018.

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.

https://doi.org/10.5194/essd-2026-613
Preprint. Discussion started: 19 August 2026

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, 10.1016/j.rse.2018.12.026, 2019.

Wang, Y., Tao, F., Chen, Y., and Yin, L.: Mapping the spatiotemporal patterns of tillage practices across Chinese croplands with Google Earth Engine, Comput. Electron. Agric., 216, 108509, 10.1016/j.compag.2023.108509, 2024.

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, H., Su, G., Li, C., and Deng, W.: The differences, advantages, and disadvantages of various image compositing methods on the Google Earth Engine platform: an exploration, Geo-spatial Inf. Sci., 28, 1-28, 10.1080/10095020.2025.2542534, 2025.

Yan, J., Wang, S., Feng, J., He, H., Wang, L., Sun, Z., and Zheng, C.: New 30-m resolution dataset reveals declining soil erosion with regional increases across Chinese mainland (1990-2022), Remote Sens. Environ., 323, 114681, 10.1016/j.rse.2025.114681, 2025.

You, N. and Dong, J.: Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth Engine, ISPRS J. Photogramm. Remote Sens., 161, 109-123, 10.1016/j.isprsjprs.2020.01.001, 2020.

You, N., Dong, J., Li, J., Huang, J., and Jin, Z.: Rapid early-season maize mapping without crop labels, Remote Sens. Environ., 290, 113496, 10.1016/j.rse.2023.113496, 2023.

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.

Zhang, C., Kerner, H., Wang, S., Hao, P., Li, Z., Hunt, K. A., Abernethy, J., Zhao, H., Gao, F., Di, L., Guo, C., Liu, Z., Yang, Z., Mueller, R., Boryan, C., Chen, Q., Beeson, P. C., Zhang, H. K., and Shen, Y.: Remote sensing for crop mapping: A perspective on current and future crop-specific land cover data products, Remote Sens. Environ., 330, 114995, 10.1016/j.rse.2025.114995, 2025a.

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, M., Wu, B., Zeng, H., He, G., Liu, C., Tao, S., Zhang, Q., Nabil, M., Tian, F., Bofana, J., Beyene, A. N., Elnashar, A., Yan, N., Wang, Z., and Liu, Y.: GCI30: a global dataset of 30 m cropping intensity using multisource remote sensing imagery, Earth Syst. Sci. Data, 13, 4799-4817, 10.5194/essd-13-4799-2021, 2021.

Zhang, S., Chen, W., Wang, Y., Li, Q., Shi, H., Li, M., Sun, Z., Zhu, B., and Seyoum, G.: Human interventions have enhanced the net ecosystem productivity of farmland in China, Nat. Commun., 15, 10523, 10.1038/s41467-024-54907-6, 2024.

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, 2025b.

Zhang, Y., Yan, H., Jiao, W., Kang, Y., Schmer, M. R., Stewart, R. D., Tracy, B. F., and You, Y.: A 1-km dataset of crop residue production and usage pathways in the conterminous US from 2001 to 2021, Earth Syst. Sci. Data Discuss., 2026, 1-24, 10.5194/essd-2026-237, 2026.

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 to carbon neutrality in China, Resources Conservation and Recycling, 203, 107450, 10.1016/j.resconrec.2024.107450, 2024.

Zheng, B., Campbell, J. B., and de Beurs, K. M.: Remote sensing of crop residue cover using multi-temporal Landsat imagery, Remote Sens. Environ., 117, 177-183, 10.1016/j.rse.2011.09.016, 2012.

Zhou, Q., Guan, K., Wang, S., Wu, X., Stroebel, S., and Hipple, J.: Multi-sensor assessment of phenology-based field-level cover cropping detection using satellite vegetation time series from Harmonized Landsat-8 and Sentinel-2, MODIS, and PlanetScope, Int. J. Appl. Earth Obs. Geoinf., 146, 105004, 10.1016/j.jag.2025.105004, 2026.

Zhou, Q., Guan, K., Wang, S., Jiang, C., Huang, Y., Peng, B., Chen, Z., Wang, S., Hipple, J., Schaefer, D., Qin, Z., Stroebel, S., Coppess, J., Khanna, M., and Cai, Y.: Recent Rapid Increase of Cover Crop Adoption Across the US Midwest Detected by Fusing Multi-Source Satellite Data, Geophys. Res. Lett., 49, e2022GL100249, 10.1029/2022GL100249, 2022.

Zhou, Y., Ferdinand, M. S., van Wesemael, J., Dvorakova, K., Baret, P. V., Van Oost, K., and van Wesemael, B.: A framework for mapping conservation agricultural fields using optical and radar time series imagery, Remote Sens. Environ., 328, 114858, 10.1016/j.rse.2025.114858, 2025.

Zhu, Z. and Woodcock, C. E.: Object-based cloud and cloud shadow detection in Landsat imagery, Remote Sens. Environ., 118, 83-94, 10.1016/j.rse.2011.10.028, 2012.