Article

Not peer-reviewed version

GeoAI-Driven Wetland Change Analysis in the Sangamon River Watershed (2000–2025): A Comparative Assessment of Machine Learning and Deep Learning Approaches

Afsheen Sadaf * and Reda Amer

Posted Date: 17 June 2026

doi: 10.20944/preprints202606.1288.v1

Keywords: wetland change analysis; machine learning (ML); deep learning (DL); pixel-based and object-based analysis; GeoAI-driven classification; DNN; U-Net semantic segmentation; Sangamon River watershed

doi:10.20944/preprints202606.1288.v1

Article

GeoAI-Driven Wetland Change Analysis in the Sangamon River Watershed (2000–2025): A Comparative Assessment of Machine Learning and Deep Learning Approaches

Afsheen Sadaf * and Reda Amer

UMSL Geospatial Collaborative, University of Missouri-St. Louis, St. Louis, MO 63121, USA
* Correspondence: afsheensadaf@umsl.edu; Tel.: +13528711170

Highlights

- Sangamon River Watershed, Illinois experienced major landscape transformation over 25 years, including a loss of 54.85% of wetland extent (860.68 km²), 46.5% of forest cover (694.88 km²), and 36.8% of water bodies (66.13 km²), alongside substantial increases in agriculture/grassland/barren land (820.62 km²) and urban/developed areas (799.52 km²).
- The object-based Gradient Tree Boosting (GTB) model achieved the highest performance for Landsat classifications (OA=91%, Kappa=0.88), while the U-Net deep learning model achieved 90% OA and Kappa=0.89 for Sentinel-based classifications.
- This research highlights the importance of combining multi-source Earth observation data, object-based image analysis, machine learning, and deep learning for long-term environmental monitoring and decision-making.
- Results provide valuable geospatial insights for wetland conservation, restoration, and climate-resilient watershed planning.

Abstract

Wetlands monitoring is essential for sustainable watershed and environmental management. This study performs a spatiotemporal wetland change analysis for the Sangamon River Watershed, Illinois between 2000–2025, using Landsat5 TM, Sentinel-2 SR, SAR, GLCM and terrain data through cloud-based processing in GEE, Google Colab and ArcGIS Pro 3.6. We conducted a comparative assessment of DL (DNN, U-Net), and ML models (RF, GTB, and SVM) through pixel-based and object-based methods. NLCD datasets were used for training and validation using stratified random sampling for five categories namely wetlands, forest, agriculture/grassland/barren land, urban/developed and water. 54.85% (860.68 km²) wetlands extent was lost to other land uses particularly agriculture, urban and forest, along with 46.5% (694.88 km²) forest and 36.80% (66.13 km²) water bodies loss. Agriculture/grassland/barren and urban/developed witnessed an increase of 8.56% (820.62 km²) and 72.92% (799.52 km²) respectively. For Landsat-based classification, the object-based GTB model achieved the highest accuracy (OA=91%, Kappa=0.88), with improved wetland delineation (PA=0.99, UA=0.92, F1=0.95). For Sentinel-based classification, U-Net outperformed all ML classifiers with (OA=90%, Kappa =0.89), while object-based GTB, SVM, and RF produced comparable results. Sentinel-based classifications had improved performance than Landsat, while object-based models consistently outperformed pixel-based methods. DEM and slope were the most influential predictors for RF models. The GeoAI framework offers a scalable and transferable approach for watershed-scale wetland monitoring, supporting evidence-based decision-making, restoration prioritization, climate resilience, and sustainable land-use planning in agricultural landscapes.

Keywords: wetland change analysis; machine learning (ML); deep learning (DL); pixel-based and object-based analysis; GeoAI-driven classification; DNN; U-Net semantic segmentation; Sangamon River watershed

---

1. Introduction

Wetlands, often referred as the "kidneys of the landscape", occupy only about 8% of the Earth's surface, yet they provide a diverse range of ecological, hydrological, geological, and socio-economic benefits, including water quality improvement, climate regulation, flood mitigation, soil stabilization, carbon sequestration, and biodiversity support [1–5]. Despite their importance, wetlands have become one of the most threatened ecosystems both globally and in the United States (US). Rapid urbanization, agricultural expansion, accelerated sea-level rise (SLR), and climate-driven stressors, such as prolonged droughts, extreme precipitation events, and storm-induced erosion, have also contributed to wetland loss and degradation at large scale [6–11]. The loss of wetlands is also associated with increased nutrient loading (nitrogen and phosphorus), elevated sediment transport, reduced flood attenuation capacity, and the decline of essential ecosystem services leading to the environmental degradation of overall ecosystem [5,12].

The Sangamon River watershed, located in central Illinois (Figure 1), encompasses approximately 13,915 km² across 18 counties and forms a major sub-basin of the Illinois River system. Illinois, one of the most agriculture-dependent states in the US, has experienced the loss of approximately 85–90% of its historical wetland area [6]. Prior to Euro-American settlement, the watershed had extensive wetland complexes linked with prairie potholes, floodplains, depressional landscapes, and riparian corridors characteristic of the Midwest glacial plains [2,6]. Extensive drainage, hydrologic alterations, and channel modifications to support intensive row-crop agriculture, have historically and largely transformed the Sangamon River Basin wetlands, along with urban expansion, transportation infrastructure development, leading to significant wetland loss in the region over the past century [7,8,12–14]. These cumulative modifications have left remaining wetlands highly fragmented, spatially isolated, and frequently disconnected from natural flow regimes, reducing their ecological integrity and long-term resilience.

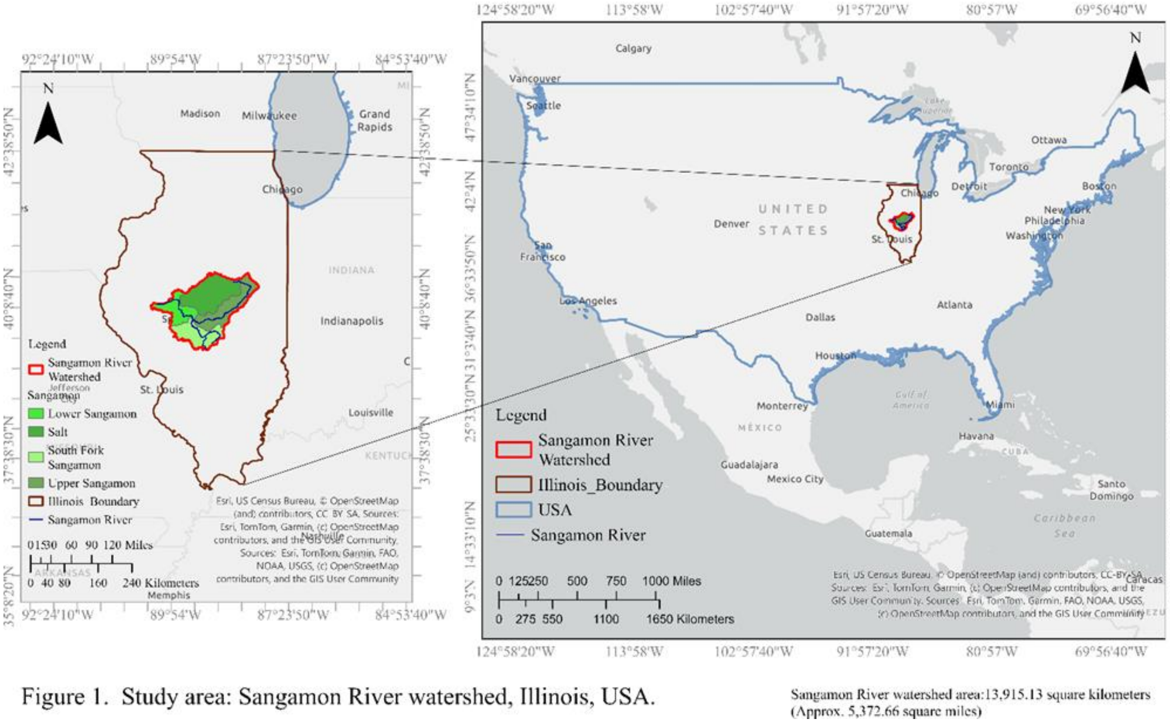

Figure 1. Study area: Sangamon River Watershed, Illinois, USA.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 June 2026 doi:10.20944/preprints202606.1288.v1

Considering the compounding effect of anthropogenic stressors and climate change variability, we can safely anticipate further wetland loss and degradation in the coming decades in this watershed region. These dynamics call for the need for multi-temporal, spatial, remote sensing (RS) and GeoAI-based assessments of wetland change/loss, while forecasting the likelihood for future change/loss, to identify areas of concern, understand scale of wetland change/loss, and support evidence-based restoration, conservation, and protection strategies at the watershed scale.

Monitoring and managing wetlands remain overall a challenging task due to their spatial heterogeneity, dynamic hydrologic regimes, and diverse vegetation structure. Before the emergence of advanced geospatial and RS technologies, traditional wetland analysis methods were often limited by sparse spatial and temporal coverage, high implementation costs, and labor-intensive field surveys, restricting their ability to capture wetland dynamics at watershed to regional scales. Recent advances in GeoAI approaches incorporate machine learning (ML) and deep learning (DL) with RS data, aim to overcome these constraints by enabling scalable, repeatable, data-driven wetland mapping, classification, and change detection across large spatial extents and multi-decadal time periods [15,21,37]. Various studies have examined wetland loss/change across spatial and temporal scales at national and global levels [6,15,20–24], but substantial gaps remain in understanding wetland dynamics at watershed and regional scales, particularly within highly modified, human, and agriculturally dominated landscapes.

Recent studies in wetland classification or change analysis have incorporated multi-temporal RS data, with various ML and DL models along with SAR imagery for improved accuracy in the detection of wetlands/mangroves [20,25,36–39]. Widely applied ML models in related studies including Classification and Regression Trees (CART), Random Forest (RF), Gradient Tree Boosting (GTB), and Support Vector Machines (SVM) have consistently shown high classification accuracy and robust performance [19,22,23,25–32]. While ML models are effective in achieving high accuracy and handle non-linear relationships among predictor variables, DL models have emerged recently as a powerful alternative. DL models such as Dense Neural Networks (DNNs), Convolutional Neural Networks (CNNs), U-Net architectures, Artificial Neural Network (ANN) Recurrent Neural Networks (RNNs), and hybrid CNN-based semantic segmentation have their edge over traditional ML models as they can capture complex spatial patterns, object boundaries, and neighborhood relationships that are particularly important in the detection of fine-scale wetland features through spatial-contextual learning, reducing salt-and-pepper noise, and improving boundary delineation [22,33].

Several studies have adopted a comparative approach of different ML classifiers, particularly in Land Use Land Change (LULC)analysis [19,32,34,35], relatively few have focused specifically on wetland change dynamics [20,36]. Studies using a comparative approach in multiple ML and DL algorithms demonstrate that performance can vary depending on environmental conditions and classification objectives [31,32]. These ML algorithms utilize distinct learning mechanisms to capture non-linear relationships in RS data. RF relies on an ensemble of independently constructed decision trees, GTB develops trees sequentially to iteratively reduce prediction errors, and SVM determines optimal hyperplanes to maximize class separability in high-dimensional feature space. When combined with multi-spectral satellite imagery and spectral indices, RF, GTB and SVM models have been widely shown to outperform other classifiers and in complex landscapes characterized by mixed land covers and subtle wetland signatures [19,20,32,36]. More recent studies have shown that DL models, when integrated with multi-sensor optical and SAR datasets, outperform traditional ML classifiers in wetland mapping applications using multisensory RS imagery [40–42]. CNN, U-Net and other transformer-based semantic segmentation models perform wetland change analysis with higher accuracy while preserving spatial continuity and landscape structure because they allow end-to-end pixel-wise classification. Therefore, evaluating multiple ML and DL classifiers is a comprehensive and integrated approach that provides critical insight into model strengths and limitations, supports the selection of the most appropriate variables and approach for complex wetlands environments, to achieve the overall accuracy of classification outcomes.

Pixel-based or object-based, are two primary image classification methods. After the launch of IKONOS in 1999, a shift was observed from pixel-based approaches, where each pixel is classified independently according to its spectral characteristics, to Object-Based Image Analysis (OBIA or GEOBIA), a method which groups neighboring pixels into meaningful image objects based on spectral, textural, and contextual characteristics [43]. Numerous studies considered OBIA an improvement due to its ability to remove "salt-and-pepper" effect and its within-class spectral variability. In wetland change analysis, OBIA is particularly valuable because wetlands are highly heterogeneous environments characterized by irregular boundaries, mixed vegetation communities, varying hydrologic conditions, and transitional ecotones that are difficult to capture using individual pixels alone. OBIA method sampling can improve the wetland delineation of wetland patches, riparian zones, marsh complexes, and fragmented landscapes by 31% compared with pixel-based methods, particularly when including texture (GLCM) and LiDAR data [33]. Although OBIA produces more visually realistic representations of wetlands, riparian zones, and agricultural landscapes but some comparative studies using RF, SVM, and Decision Tree (DT) classifiers have also found pixel-based methods provide comparable classification accuracy while requiring fewer input variables and lower computational effort, making it more efficient for large-scale wetland mapping [44]. A recent coastal wetland study comparing PBIA, OBIA, and deep learning (DL) reported that DL method provides higher accuracy for wetland classification [45]. Despite these advantages, pixel-based approaches remain valuable, especially for medium-resolution (30m) imagery, where computational efficiency and simpler workflows are important.

Google Earth Engine (GEE), a JavaScript-based programming interface, has become a widely adopted platform for mapping and analyzing LULC change, wetlands, forest and crop dynamics, urban heat islands, and other environmental processes [15,19,46–49]. GEE, an open-source, cloud-based geospatial analysis platform provides free and easy access with multiple spatial and temporal resolutions that allow efficient processing of RS imagery, facilitating reproducible and optimized modeling workflows.

This study employs an integrated and comparative wetland classification and change detection approach using multi-season (spring and summer) Landsat 5 TM and Sentinel-2 SR indices, Sentinel-1 SAR backscatter data, terrain and land use data to quantify wetland change detection in complex, heterogeneous and fragmented landscapes such as the Sangamon River watershed between 2000 and 2025 (Figure 2). GEE's/Colab cloud-based was used for classification and ArcGIS Pro 3.6 for quantifying change, mapping and visualization. We compared ML classifiers, RF, GTB, SVM and DL-DNN, U-Net (FCN) models to quantify and spatialize wetland change through gain, loss, and conversion of wetlands into other land use type such an agriculture, barren, grassland, developed land and water.

The key contributions of this study are:

- integrating multi-season Landsat 5 TM (for years 2000-2008) and Sentinel-2 (for years 2017-2025) optical indices, Sentinel-1 SAR backscatter, and terrain-derived variables within a comparative ML through both pixel-based and object-based approaches and DL-based workflow to evaluate their accuracy performance in wetland mapping and change detection;
- applying a temporally consistent classification approach using identical predictors and training data for 2001 and 2017 to ensure that changes reflect true land cover transitions;
- evaluating DL models performance against ML models for Landsat 5 TM and Sentinel-2 respectively for different spatial resolutions, and temporal periods;
- quantifies and spatially characterizes wetland dynamics, including gains, losses, and conversions to other land use types, providing critical understanding into landscape-level transformations;
- integrating cloud-based GEE and Colab platforms that provide high computational efficiency with ArcGIS-based visualization and quantification,
- determining the "variable importance" of variables used in this study through the RF classifier algorithm.

This study contributes to the growing body of wetland monitoring research through an integrated, comparative, reproducible and scalable framework that is critical for regional planning, conservation and land management policies, future restoration initiatives, and prediction of the future wetland loss based on current study and allows for its application to other watersheds contributing to broader wetland monitoring and management efforts in the US and abroad.

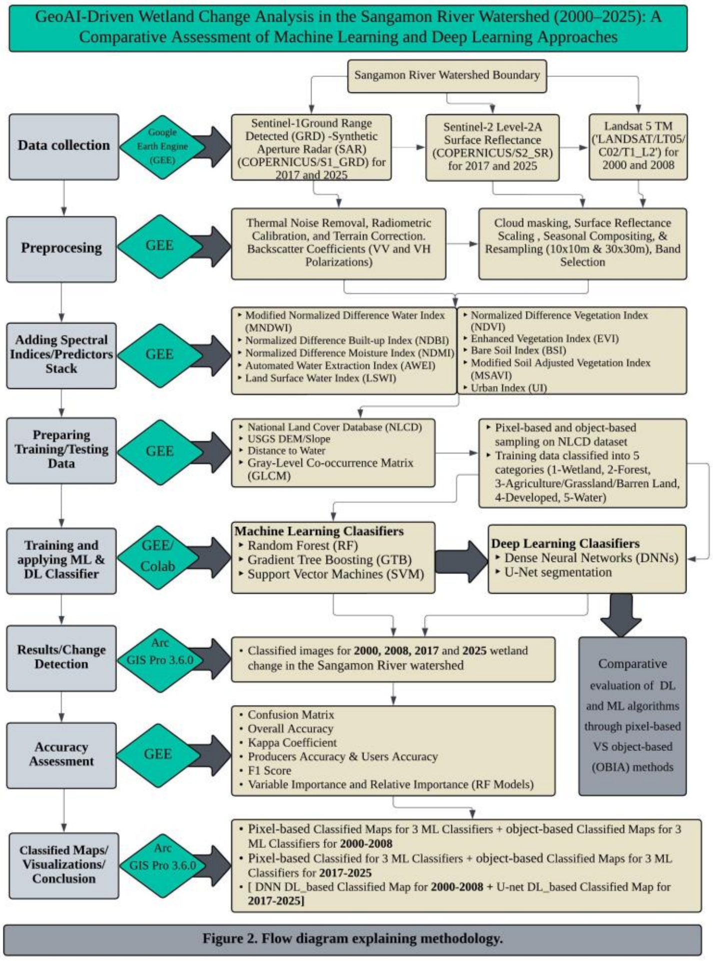

Figure 2. Flow diagram explaining methodology for wetland change analysis in the Sangamon River Watershed.

2. Materials and Methods

2.1. Study Area (Sangamon River Watershed)

The Sangamon River watershed, covers 18 counties and 13,915.13 square kilometers (Approx. 5,372.66 square miles) in central Illinois, lies within coordinates 40°26′43.12″ N, 88°43′57.24″ W and 40°1′21.17″ N, 90°25′58. 45″ W and visualized in ArcGIS Pro 3.6.0 (Figure 1). According to the 2021 National Land Cover Data (NLCD), the current land-use composition in the Sangamon River watershed is approximately consists 77% of row-crop agriculture (corn and soybeans), 12% of developed/urban area, 5% of forested, grassland and wetlands, and 1% of open water, small grains, other land uses, dominating the agricultural land use in this watershed. The Sangamon River watershed was delineated using the USGS Watershed Boundary Dataset by merging four HUC-8 sub-basins: Upper Sangamon (07130006), South Fork Sangamon (07130007), Lower Sangamon (07130008), and Salt Creek (07130009) both in Arc GIS Pro 3.6.0 and GEE.

doi:10.20944/preprints202606.1288.v1

2.2. Data Sources and Data Processing

We used the high-resolution, Sentinel-1(S1) SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected (“COPERNICUS/S1_GRD”) and multispectral Sentinel-2 (S2) Level-2A Surface Reflectance (SR) imagery (“COPERNICUS/S2_SR_HARMONIZED”) for years 2017-2025 and Landsat5 (L5) TM (Thematic Mapper) (“LANDSAT/LT05/C02/T1_L2”) downloaded in GEE for year 2000 and 2008. The other datasets include Digital Elevation Model (DEM) data from the USGS 3DEP 10/30m DEM, and NLCD data (2001 and 2017) to generate training and validation samples for classification.

For the deep learning component, DNNs and U-Net semantic segmentation models were implemented in Python3 (Google Colab) using TensorFlow/Keras, with GEE-exported multi-band image patches derived from the Landsat and Sentinel predictor stacks for all land use classes.

2.3. Cloud Masking, Surface Reflectance Scaling & Band Selection)

All image processing and computational analysis were conducted in JavaScript (GEE) and Python3 (Google Colab). All GIS raster and vector data was projected on “NAD_1983_StatePlane_Illinois_West_FIPS_1202” (in ArcGIS Pro) and “EPSG:5070” (in GEE). L5 imagery was used to conduct analysis between 2000 to 2008, while S1 & S2 imagery for years 2017 to 2025 to calculate wetland change by multiplying respective pixels in ArcGIS Pro.

S1 and S2 optical imagery were preprocessed using standardized approach. For Sentinel-2, we use S2Cloudless DL algorithm to mask cloud, cirrus, cloud shadow, and snow pixels, which has demonstrated improved performance over others cloud masking algorithms including Scene Classification Layer (SCL) in minimizing residual contamination [50]. SAR sensitivity to surface moisture, inundation conditions, and vegetation structure improved class separability and reduced misclassification within moisture-sensitive landscapes. For S1 data, GEE’s built-in SAR processing algorithms, including thermal noise removal, radiometric calibration, and terrain correction, were used. Backscatter coefficients (VV and VH polarizations) were converted to sigma naught (σ^0) values in decibels (dB) and filtered to reduce speckle noise. Temporal compositing and multi-date aggregation were applied to enhance signal stability and reduce noise effects. For L5 Quality Assessment (QA) band (QA_PIXEL) masking was used.

All images were harmonized and resampled to10m (S2) and 30m (L5) spatial resolution. Reflectance values were scaled by a factor of 0.0001 to convert integer digital numbers to physical reflectance units. Detailed sensor characteristics including selected spectral bands, their spatial resolutions, and revisit frequencies are summarized in Table 1.

Table 1. Band Description for Landsat 5 TM, Sentinel-2 SR L2A imagery and Sentinel-1 SAR GRD used in the study.

Landsat 5 TM				**Sentinel-2 SR L2A**				**SAR GRD**			
Band	**Spectral**	**Wavelength (µm)**	**Resolution (m)**	**Band**	**Spectral**	**Wavelength (µm)**	**Resolution (m)**	**Band**	**Polarization**	**Frequency/Wavelength**	**Resolution (m)**
B1	Blue	0.45-0.52	30	B2	Blue	0.45-0.52	10	VV	Vertical transmit / Vertical receive	C-band (~5.6 cm; 5.405 GHz)	10
B2	Green	0.52-0.60	30	B3	Green	0.52-0.60	10				
B3	Red	0.63-0.69	30	B4	Red	0.63-0.69	10				
B4	NIR (Near Infrared)	0.76-0.90	30	B8	NIR	0.76-0.90	10				
B5	SWIR1 (Short wave Infrared)	1.55-1.75	30	B11	SWIR1	1.55-1.75	20	VH	Vertical transmit / Horizontal receive	C-band (~5.6 cm; 5.405 GHz)	10
B7	SWIR2	2.08-2.35	30	B12	SWIR2	2.08-2.35	20				
Temporal resolution 16 day				3-5 days				6-12 days			

2.4. Composite Generation and Normalization

Seasonal L5 composites for 2000 and 2008 and S2 composites for 2017 and 2025 were generated in GEE for the Sangamon River Watershed, Illinois. L5 and S2 imagery was first filtered by seasonal date ranges and clipped to the watershed boundary. For S2, to reduce atmospheric noise, minimize cloud contamination and short-term variability, scaled reflectance images were composited using a median reducer to generate seasonal, cloud-free mosaics for the Spring (March-May) and Summer (June–September) period, which corresponds to peak vegetation growth and maximum wetland inundation in the study area, thereby enhancing the detectability of wetland extent and loss. For L5, a 25th percentile (P25) compositing approach was applied to the image collection for Spring and Summer. The percentile-based composite was selected instead of median compositing as it minimizes the influence of bright atmospheric artifacts and improves the representation of stable land surface conditions for L5. Z-score normalization was applied to standardize spectral variability as shown in Equation 1. Mean and standard deviation statistics for each spectral band were calculated from the seasonal image collection.

$\boldsymbol{z=\frac{\{x-\mu\}}{\sigma}}$ ……… Equation (1), where (x) represents the pixel value, (μ) is the mean reflectance of the band, and (ϭ) is the standard deviation.

A small constant was added during division to avoid instability associated with near-zero standard deviation values. The normalized and denormalized seasonal composites were subsequently renamed according to season-specific band identifiers (e.g., Blue_spring, NIR_summer) and used as predictor variables for ML and DL models. Normalization is used for all classifiers for L5 imagery and SVM for S2 imagery.

2.4.1. Dense Neural Network (DNN) Classification for Landsat 5 TM (2000–2008)

DNN was selected for L5 imagery for years 2000 and 2008, because of its limitation in spatial resolution (30m), where pixels may not be able to distinguish fine land cover details. Sample-based DNN models primarily learn relationships among spectral, environmental, terrain, and texture variables without relying heavily on fine-scale spatial context and enable direct comparison with ML models used in this study.

2.4.2. U-Net Semantic Segmentation for Sentinel-2 and Sentinel-1 Data (2017–2025)

U-Net semantic segmentation model was selected for S2 and S1 datasets for years 2017 and 2025. Predictor stacks and corresponding land-cover labels were exported from GEE as TFRecord image patches using 64×64 patch dimensions. For both DNN and U-Net, spring and summer composites were generated in GEE using similar preprocessing and composite generation as stated in section 3.2 including all the predictor variables used for ML models.

2.5. Spectral Indices

The spectral indices were calculated from both L5 and S2 imagery to enhance differentiation among vegetation, wetlands, built-up areas, open water, and moisture variations [Table 2]. These indices, together with spectral bands, were used as predictor variables for all classifiers for both pixel-based and object-based methods. Among these predictors, MNDWI and the NDMI are considered highly useful for wetland change analysis. While MNDWI provides superior mapping for open water and suppressing built-up noise, NDMI tracks moisture content in vegetation and soil, providing a comprehensive view of wetland degradation, drying, or shifts from aquatic to terrestrial environments.

Table 2. Spectral Indices derived in the study and their interpretation.

	Index	Purpose	Range	Equations	Explanation/References	Interpretation
1	*Normalized Difference Vegetation Index (NDVI)*	Vegetation health	-1 to +1	$\boldsymbol{NDVI} = \frac{(\boldsymbol{NIR}-\boldsymbol{RED})}{(\boldsymbol{NIR}+\boldsymbol{RED})}$ ………. Equation (2)	Equation (2) calculates NDVI, i.e., vegetation greenness to check vegetation density, vigor, and overall health and use to differentiate vegetated areas from other land use types (e.g., soil, water, covered areas) [51]. Healthy vegetation absorbs red and reflects NIR due to chlorophyll and cell structure.	(< 0) are clouds and water bodies. (0) are barren rock, sand, or snow. (0.2 to 0.5) are shrubs, grasslands, or sparse vegetation. (0.6 to 0.9) are dense, healthy vegetation (temperate, and tropical forests).
2	*Modified Normalized Difference Water Index (MNDWI)*	Wetlands & shallow water	-1 to +1	$\boldsymbol{MNDWI} = \frac{(\boldsymbol{GREEN}-\boldsymbol{SWIR})}{(\boldsymbol{GREEN}+\boldsymbol{SWIR})}$ ………. Equation (3)	MNDWI in Equation (3) improves NDWI by using SWIR to suppress built-up land spectral signatures, making it more effective for water mapping in urban regions [52]. It is better at delineating waterlogged areas mixed with vegetation, which are common in the Sangamon River watershed floodplains.	(>0.2) are open water. (-1 to 0) are non-water features (built-up/soil/vegetation)
3	*Normalized Difference Built-up Index (NDBI)*	differentiates built-up areas	-1 to +1	$\boldsymbol{NDBI} = \frac{(\boldsymbol{SWIR}-\boldsymbol{NIR})}{(\boldsymbol{SWIR}+\boldsymbol{NIR})}$ ………. Equation (4)	NDBI in Equation (4) highlights and differentiates built-up areas from vegetation and soil [53].	(-1 to 0) are vegetation, water bodies, or empty, non-urban land. (0 to 1) are built-up, urban areas. (+1) are high-density, dense urban areas. (0) is a mix of built-up, urban areas and sparse, sparse vegetation.
4	*Normalized Difference Moisture Index (NDMI)*	Vegetation moisture / soil moisture	-1 to +1	$\boldsymbol{NDMI} = \frac{(\boldsymbol{NIR}-\boldsymbol{SWIR})}{(\boldsymbol{NIR}+\boldsymbol{SWIR})}$ ………. Equation (5)	NDMI in Equation (5) is useful to estimate vegetation moisture content [54].	(>0.5) indicate higher vegetation water content and moist surfaces (wetlands, irrigated areas). (0.2–0.5) indicate moderate moisture conditions. (-1 to 0) indicate dry vegetation, bare soil, or built-up areas.
5	*Automated Water Extraction Index (AWEI)*	improved water detection from shadows and built-up areas	-1 to +1	$\boldsymbol{AWEI} = \boldsymbol{4}(\boldsymbol{Green} - \boldsymbol{SWIR1}) - \boldsymbol{0.25}(\boldsymbol{NIR}) + \boldsymbol{2.75}(\boldsymbol{SWIR2})$ ………. Equation (6)	AWEI in Equation (6) has improved water mapping accuracy by reducing noise from shadows and dark surfaces than MNDWI [55].	(>0.5) indicate open water and waterlogged areas. (-1 to 0) represent non-water features.
6	*Enhanced Vegetation Index (EVI)*	Enhanced vegetation health	-1 to +1	$\boldsymbol{EVI} = \frac{\boldsymbol{2.5}(}{\{\boldsymbol{NIR} + \boldsymbol{6}\ (\boldsymbol{RE}}$ ………. Equation (7)	EVI in Equation (7) is an improvement upon the NDVI by enhancing sensitivity in high biomass regions and reducing canopy background noise	(< 0) indicate water, clouds, or non-vegetated surfaces. (<0.2) indicate sparse vegetation or bare land.

					and atmospheric interference [56].	(0.2–0.5) represent moderate vegetation. (>0.5) indicate dense vegetation with high biomass.
7	*Bare Soil Index (BSI)*	differentiate bare soil from vegetation, water, and built-up areas	-1 to +1	$\boldsymbol{BSI} = \frac{(\boldsymbol{SWIR1+RED})-(\boldsymbol{NIR}}{(\boldsymbol{SWIR1+RED})+(\boldsymbol{NIR}}$ ………. Equation (8)	BSI in Equation (8) is used to enhance bare soil features that distinguish them from other land uses such as water, vegetation, and built-up areas [57].	(< 0) indicate water bodies or dense vegetation. (≈ -0.2 to 0) represent vegetated surfaces with minimal exposed soil. (0 to 0.3) indicate mixed surfaces (sparse vegetation with exposed soil). (> 0.3) indicate bare soil, dry riverbeds, or exposed land surfaces.
8	*Modified Soil Adjusted Vegetation Index (MSAVI)*	Enhanced early-season crops detection from barren land	-1 to +1	$\boldsymbol{MSAVI} = \frac{\boldsymbol{2} \times \boldsymbol{NIR} + \boldsymbol{1} -}{}$ ………. Equation (9)	Equation (9) illustrates MSAVI, which is useful for distinguishing early-season crops from barren land and reduces soil background impacts [58].	(< 0) indicate water, clouds, or non-vegetated surfaces. (< 0.2) indicate bare soil or very sparse vegetation. (0.2–0.5) represent emerging or moderately dense vegetation, especially useful in early growth stages. (> 0.5) indicate dense, healthy vegetation with minimal soil influence.
9	*Urban Index (UI) (Landsat 5 TM)*	differentiate barren land from other land uses	-1 to +1	$\boldsymbol{UI} = \frac{(\boldsymbol{SWIR-NIR})}{(\boldsymbol{SWIR+NIR})} = \boldsymbol{1.0} \times \boldsymbol{100}$ ………. Equation (10)	Equation (10) illustrates UI which is used specifically for Landsat TM to distinguish built-up areas from bare soil [59].	(< 0) indicate water bodies or vegetated areas. (≈ -0.2 to 0) represents dense vegetation or moist surfaces. (0 to 0.2) indicate mixed land cover (transitional zones). (> 0.2) indicate urban/built-up areas, dry soil, or impervious surfaces.
10	*Land Surface Water Index (LSWI)*	detect the water content in vegetation and soil	-1 to +1	$\boldsymbol{LSWI} = \frac{(\boldsymbol{NIR-SWIR})}{(\boldsymbol{NIR+SWIR})}$ ………. Equation (11)	Equation (11) illustrates LSWI which is used to monitor the water content in vegetation and soil [60].	(< 0) indicate dry soil, impervious surfaces, urban/built-up areas, barren land, or sparse vegetation with low moisture content. (≈ 0 to 0.1) indicate dry vegetation, agricultural fields, grasslands, or mixed land-cover conditions with limited surface moisture. (≈ 0.1 to 0.3) indicate healthy vegetation, moist soils, wetlands, and areas with

moderate water content.
(> 0.3) indicate saturated soils, inundated wetlands, flooded vegetation, and open water bodies with high moisture content.

2.6. DEM, Slope, and Distance-to-Water

DEM, slope [Figure 3 (a) & (b)], and distance-to-water were included as predictor variables to capture topographic gradients, surface hydrological connectivity, and inundation potential. DEM data was downloaded from the USGS 3DEP 10m/30m DEM and clipped to the study area extent and terrain slope was derived from the DEM and resampled using bilinear interpolation and reprojected to consistent 10m/30m for L5 and S2 respectively. Surface water features were extracted from L5 and S2 imagery using the MNDWI, where pixels with values greater than zero were classified as water. A Euclidean distance-to-water layer was then generated using a fast distance transform algorithm, with distances computed from the nearest water pixel and converted to metric units.

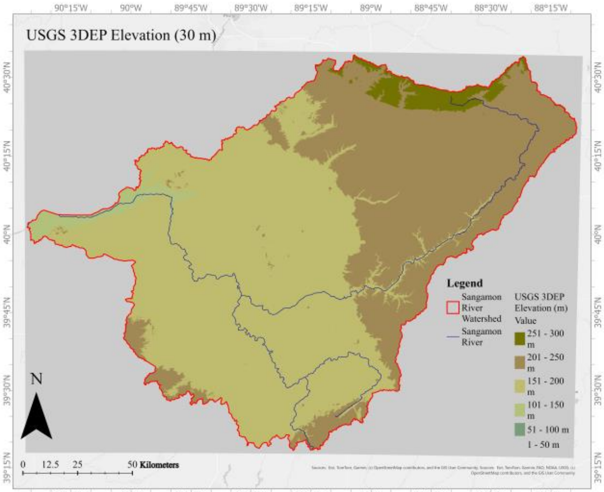

Figure 3 (a). Elevation map - Sangamon River Watershed, Illinois

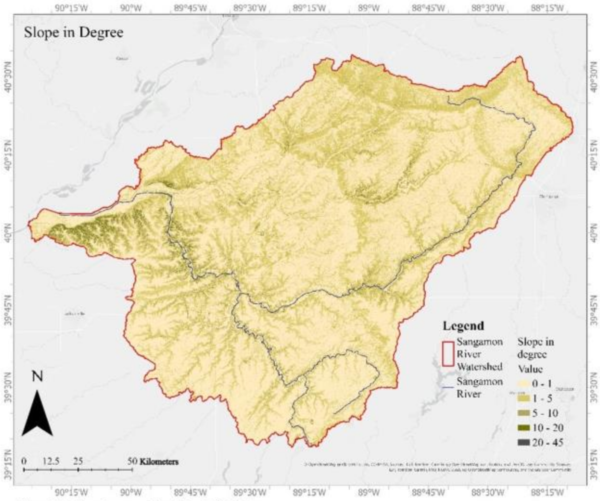

Figure 3(b). Slope-Sangamon River Watershed, Illinois.

doi:10.20944/preprints202606.1288.v1

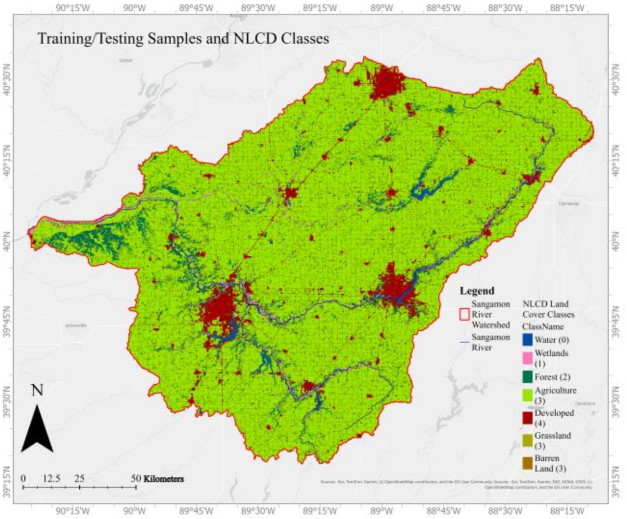

Figure 3(c). Stratified training and testing sample distribution by NLCD land use classes in the Sangamon River Watershed, Illinois

Figure 3. (**a**) USGS 3DEP Elevation (30m), Sangamon River Watershed, Illinois; (**b**) Slope, Sangamon River Watershed, Illinois; (**c**) Stratified training and testing sample distribution by NLCD land use classes in the Sangamon River Watershed, Illinois.

2.7. Gray-Level Co-Occurrence Matrix (GLCM)

Multi-seasonal GLCM features were incorporated for both L5 and S2 imagery to enhance land use/wetland classification accuracy particularly with ML based classifiers like RF and SVM [61,62]. GLCM is a statistical approach for analyzing image texture that explicitly captures the spatial relationships between pixels. It works by quantifying how often pairs of pixels with specific intensity values occur at a defined distance and orientation (e.g., one pixel to the right, above, or diagonally adjacent to another). We computed Contrast, Entropy, Inverse Difference Moment (IDM), using 3x3 window size which is useful for fine details relevant to land used with fragmented wetland composition.

2.8. Training Samples

Wetland training samples were obtained from a rasterized (30m L5 & 10m for S2) NLCD dataset, where wetland classes (90 and 95) woody wetlands (regions that have forest or shrubland vegetation greater than 20% of vegetative cover and the soil or substrate is periodically saturated with or covered with water) and emergent herbaceous wetlands (regions that have perennial herbaceous vegetation greater than 80% of vegetative cover and the soil or substrate is periodically saturated with or covered with water) [63] were classified as Class1, with other Land cover classes were classified as forest (2), agriculture/ grassland/ barren land (3), and urban/developed (4) and water (5) [Figure 3 (c)]. A stratified random sampling approach with same classification scheme and class distribution strategy was used for both L5 and S2 datasets to ensure balanced and representative sampling across all classes. For all pixel-based models, a total of 10,000 sample points were generated at 10/30m resolution, with class-specific sample sizes defined to address class imbalance (2600, 2000, 2500, 2200 and 700 samples per class) [Table 3 (a)].

Sampling was conducted in the upper, middle, and lower reaches of watershed, to capture spatial heterogeneity. To reduce class bias we enforced proportional representation across all land cover classes sampling design, while per-pixel stratified sampling also minimizes label ambiguity compared to vector-based approaches and ensures consistency with the spatial resolution of predictor variables [Table 3 (b)]. Training and validation samples with missing predictor values were removed. The resulting dataset was used to train all the classifiers for wetland change analysis.

For OBIA, image objects were generated using the Simple Non-Iterative Clustering (SNIC) segmentation algorithm, where segment-level mean predictor values and majority land-cover labels were extracted from segmented image objects. These object-level samples (9,000) were subsequently used to train the RF, GTB, and SVM object-based classifiers [Table 3 (a)]. For DNN model (L5), after the stratified random sampling based on reclassified NLCD, predictor values were extracted using sampleRegions() and normalized to reduce differences in variable ranges and improve neural network convergence in GEE and exported as CSV tables for analysis in Google Colab. For U-Net semantic segmentation model (S1/S2) 64×64 image patches exported as TFRecord datasets using a 70:30 split using a reproducible random seed to ensure consistency across all ML and DL models [Table 3 (a)].

Table 3. (a). Class-wise training sample distribution for wetland and non-wetland classes.

		Pixel-based sampling			**Object-based sampling**		
Class Name/ Class ID	**Sub-Classes**	**Samples per class**	**Training distribution**	**Testing distribution**	**Samples per class**	**Training distribution**	**Testing distribution**
Wetland (1)	NLCD 90, 95 (Woody Wetlands/Emergent Herbaceous Wetlands, Freshwater- Forested and Shrub wetland)	2600	1822	778	3500	2,450	1,050
Forest (2)	NLCD 41, 42, 43, 71, 31	2000	1414	586	1500	1,050	450
Agriculture/ Grassland/ Barren land (3)	NLCD 81, 82,	2500	1732	768	2000	1,400	600
Urban/Developed (4)	NLCD 21, 22, 23, 24	2200	1530	670	1500	1,050	450
Water (5)	NLCD 11, 12	700	493	207	500	350	150
Total sample size		10,000				9000	
Training size (70%)		6991				6,300	
Testing size (30%)		3009				2,700	

Table 3. (b). Sampling approach for ML and DL models compared in the study.

Model type	**Sampling approach**
RF (ML)	pixel-based, object-based
GTB (ML)	pixel-based, object-based
SVM (ML)	normalized pixel-based, OBIA
Dense Neural Network (DL)	tabular predictor samples
U-NET (DL)	TFRecord patches

2.9. Training Classifiers [Random Forest (RF), Gradient Tree Boosting (GTB), and Support Vector Machine (SVM)], Deep Learning (DL) Models and Classifying Maps (2000-2008-2017-2025)

All classifiers were trained using predictor variables, including spectral bands and indices, terrain variables, hydrologic proximity layers, SAR backscatter data, and GLCM and the dataset was divided into 70% training/30% validation/testing using a reproducible random seed to ensure consistency across classification. Hyperparameters for each classifier were optimized through iterative testing to improve classification accuracy and reduce overfitting. ML used 500 to 550 trees with 6 variables per split, taking square root of total no of predictors.

Image objects were generated using the Simple Non-Iterative Clustering (SNIC) segmentation algorithm in GEE and segment-level mean predictor variables and majority land-cover labels were extracted and used to train the ML object-based classifiers. Separate classification models were developed for L5 and S2/S1 datasets to account for differences in spatial resolution, spectral characteristics, and sensor properties. Training and validation were performed at 10m/30m resolution for S1/S2 and L5 datasets, respectively.

For DNN, we trained the model using categorical cross-entropy loss and the Adam optimizer with 70:30 split for training and testing subsets. The DNN architecture consisted of multiple fully connected dense layers with batch normalization, dropout regularization, and ReLU activation functions to model non-linear relationships among predictor variables to reduce overfitting. Predictor variables were normalized prior to model training to improve convergence and classification performance.

For U-Net, semantic segmentation multi-band image patches (64×64 pixels) were exported from GEE as TFRecord datasets (patch count=136026) and trained in Google Colab using TensorFlow/Keras. U-Net is efficient in enabling learning of complex spatial patterns and improving segmentation of shoreline boundaries, wetlands, riparian zones, and heterogeneous landscape boundaries. U-Net serves encoder and decoder convolutional blocks with skip connections to preserve fine spatial detail. Convolutional layers were used to extract hierarchical spatial features, while max-pooling operations reduced spatial dimensionality during encoding. Decoder layers reconstructed segmentation outputs using upsampling operations and skip connections from corresponding encoder layers. The final output layer utilized softmax activation for multi-class semantic segmentation.

Final classified maps were generated for years 2000, 2008, 2017, and 2025 (Figure 5), and subsequent wetland change (Figure 6) was calculated using raster calculator in ArcGIS Pro.

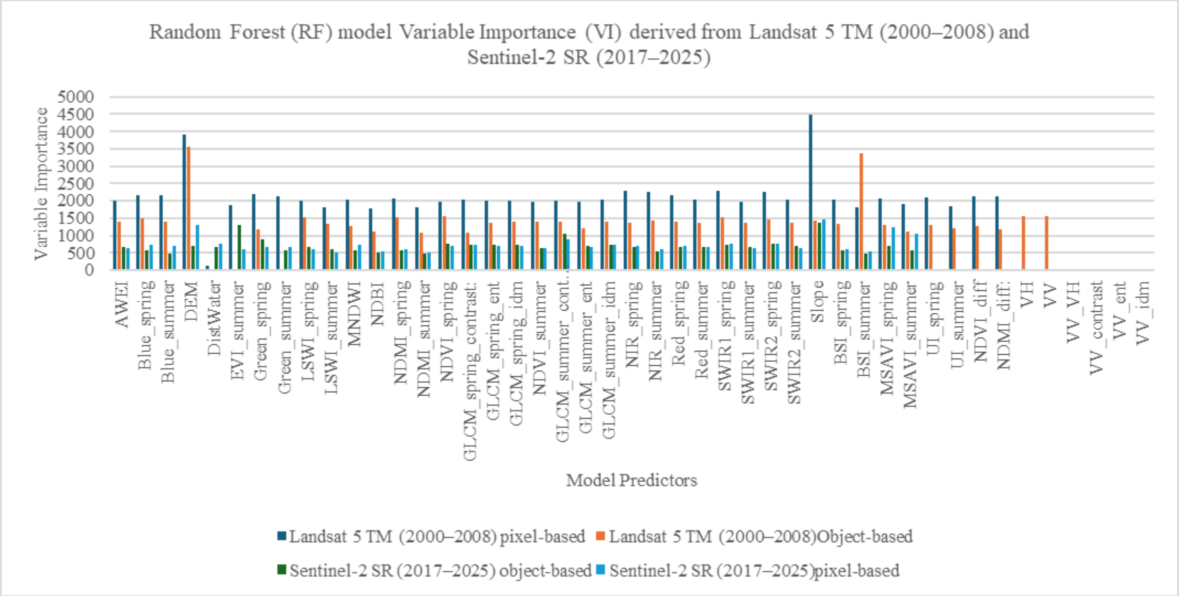

Figure 4. Random Forest (RF) model Variable Importance (VI) derived from Landsat 5 TM (2000–2008) and Sentinel-2 SR (2017–2025) datasets for both pixel-based and object-based methods.

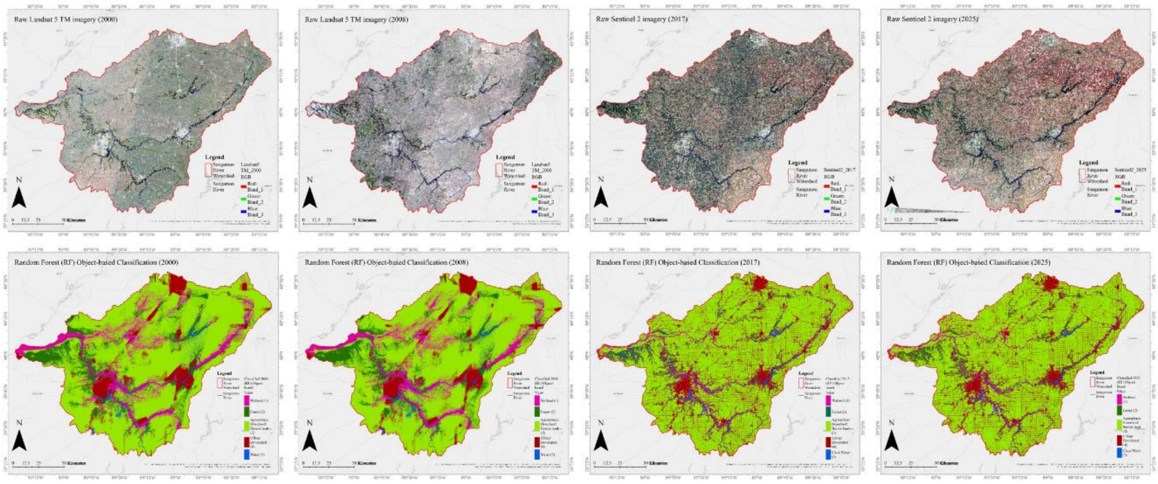

Figure 5. Satellite imagery and object-based classification maps for years 2000, 2008, 2017 and 2025 using Landsat 5 TM, Sentinel 2, and SAR through Random Forest (RF) models in the Sangamon River Watershed, Illinois, USA.

Figure 5. Satellite imagery and object-based classification maps for years 2000, 2008, 2017 and 2025 using Landsat 5 TM, Sentinel 2, and SAR through Random Forest (RF) models in the Sangamon River Watershed, Illinois, USA.

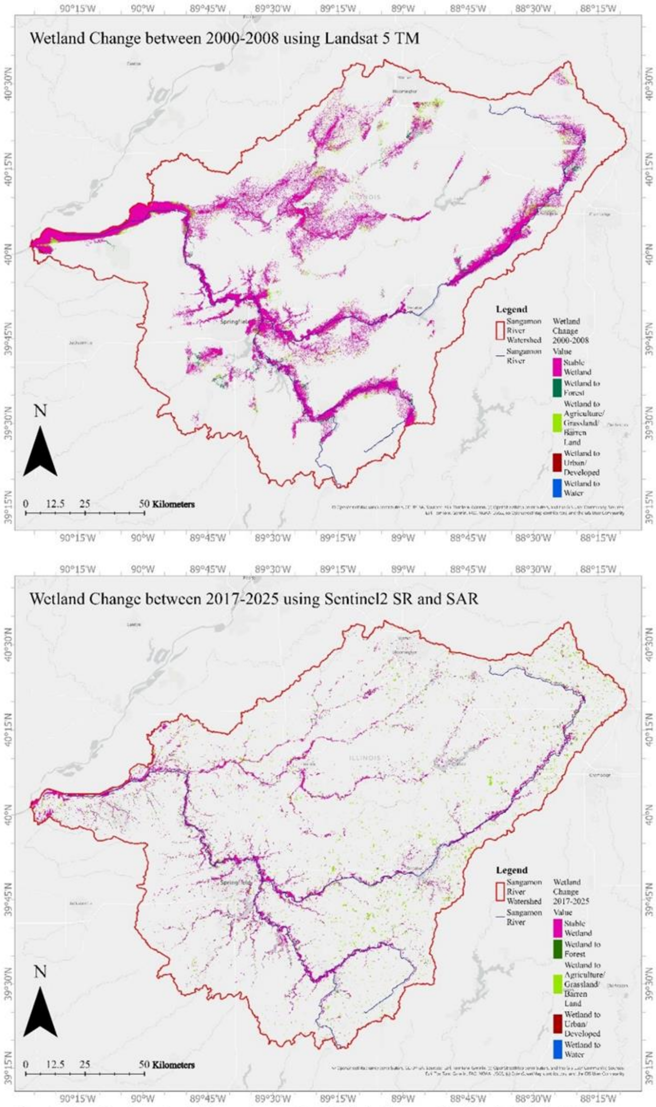

Figure 6. Change detection and wetland loss map between 2000-2025 in the Sangamon River Watershed, Illinois, USA.

Figure 6. Change detection and wetland loss map in the Sangamon River Watershed, Illinois, USA.

3. Results

3.1. Accuracy Assessment

Confusion Matrices (CM), Overall Accuracy (OA), Kappa coefficient (KC), Producer's Accuracy (PA), User's Accuracy (UA), and F1-scores for each class Table 4 (a).

Table 4. (a). Accuracy Assessment Matrices used in the study.

	Accuracy Metric	Equations	Explanation
1	Overall Accuracy (OA)	$\boldsymbol{OA} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$ ………… Equation 12, where TP is true positive, FP is false positive, TN is true negative, and FN is false negative.	OA measures correctly classified pixels in all cases, representing equation 12.
2	Kappa Coefficient (KC)	$\boldsymbol{KC} = \frac{2(TP.TN-FN.FP)}{(TP+FN)(FN+TN)+(TP+FP)(FP+TN)}$ ………… Equation 13, where KC is the Kappa coefficient.	KC measures classification agreement beyond chance, represented with equation 13.
3	Producer Accuracy (PA)	$\boldsymbol{PA} = \frac{TP}{(TP+FN)}$ ………… Equation 14	PA measures the ability of the classifier to correctly identify reference samples for each class (i.e., omission error) and represents equation 14.
4	User Accuracy (UA)	$\mathbf{UA} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$ ………… Equation 15	UA measures the reliability of classified pixels for each class and represents equation 15.
5	Class-wise F1-score	$\boldsymbol{F1i} = 2\frac{(PAi \times UAi)}{(PAi+UAi)}$ …………… Equation 16, where $F1_{i}$ is F1-score for class i.	Class-wise F1-scores calculates the harmonic Mean of PA and UA (precision and recall) for evaluating classifier performance on each class (i=1-5) and represents equation 16.

3.2. Landsat 5 TM for Period 2000-2008

As Table 4 (b) illustrates, for pixel-based classification, RF and GTB classifiers showed strong performance for wetlands, agriculture/grassland/barren land, and water compared to SVM and DNN models with 67% OA , 0.57 KC. RF and GTB showed improved detection of wetland areas with PA between 0.84–0.87 and F1-scores between 0.76–0.78, despite some confusion with forest and agricultural classes. Water showed the highest UA (>0.90) for both RF and GTB models. However, urban/developed areas showed confusion with agriculture and wetland classes due to mixed spectral responses and fragmented urban land cover with lower PA (~0.50–0.51). The pixel-based SVM model compromised in separating spectrally similar and spatially heterogeneous classes, with low OA (53%) and KC (0.38) with only agriculture/grassland/barren land (0.71) performed well, while urban/developed remained low with an F1-score of 0.27. The pixel-based DNN model with 61% OA and 0.50 KC outperformed SVM with wetland and water classes showed comparatively stronger performance compared to urban and forest.

As Table 4 (b) illustrates, object-based classification methods significantly enhance wetland classification compared to pixel-based methods. The object-based RF model with an OA of 88% and a 0.85 KC, achieved high PA (0.98), UA (0.89), and F1-score (0.93) for wetlands. Agriculture/grassland/barren land and urban/developed with F1-scores 0.85 and water with 0.97 UA and 0.88 F1-score showed substantial improvements.

As Table 4 (b) illustrates, the object-based GTB classifier with 91% OA and 0.88 KC achieved the highest performance among all classified models. An improved class separability was observed in wetland classification with F1-score of 0.95, water 0.97, urban/developed and agriculture/grassland/barren land with >0.90 F1-score, highlighting its robustness for wetland mapping in fragmented watershed environments as it captured complex non-linear relationships among all variables. The object-based SVM performed better than pixel-based counterpart, achieving 88% OA and 0.84 KC. Wetland and water with strong F1-scores of 0.91 and 0.93 respectively, indicates segmentation improved SVM classification stability and reduced spectral confusion. Overall RF and GTB consistently outperformed SVM and DNN models in both pixel-based and object-based models.

3.3. Sentinel-2 for Period 2017-2025

As Table 4 (c) illustrates, for pixel-based classification, RF and GTB classifiers showed strong performance with 80–81% OA, 0.77 KC and improved performance for wetlands, with PA between 0.84–0.85 and F1-scores between 0.82–0.83, due to Sentinel multispectral and SAR-derived predictors. F1-scores for agriculture/grassland/barren land and urban/developed were between 0.85–0.86 and 0.80–0.84 respectively, suggest improved class separability compared to Landsat model. Water showed the highest UA (>0.92) and F1-scores (~0.88–0.92), while forest showed some confusion between wetlands and other vegetated land-cover classes with F1-scores (~0.72–0.73). The pixel-based SVM with 61% OA and 0.54 KC, showed reduced PA, UA, and F1-scores for wetlands, forests, and urban respectively compared to RF and GTB. SVM showed substantial confusion in all classes except water with strong F1-score (0.84).

As Table 4 (c) illustrates that the Sentinel-based U-Net model achieved the highest classification performance with 90% OA and 0.89 KC, with other object-based classification also showing improved performance and OA. The encoder–decoder architecture and convolutional feature extraction capabilities of U-Net substantially improve the identification of wetlands riparian corridors, and mixed class boundaries relative to other ML models with high PA (0.97) UA (0.92), and F1score (0.94). Water, urban/developed, agriculture/grassland/barren and forest also produced highest F1-scores (0.97, 0.94, 0.92 and 0.90) respectively.

As Table 4 (c) illustrates, object-based GTB classifier achieved 88% OA and 0.84 KC and produced the highly balanced PA, UA, and F1-scores in all classified models, with F1-score of 0.90 for wetland, 0.89 for water, 0.88 for urban/developed and agriculture/grassland/barren land. The object-based SVM had 88% OA and 0.84 KC. Wetlands and urban/developed indicate strong class separability with reduced omission errors. Water achieved 0.99 UA with and 0.91 F1-score, while forest and agriculture/grassland showed comparatively lower balance between PA and UA, suggesting some confusion between vegetation. The object-based RF model showed strong performance in identifying wetland strands and reducing classification noise with an 87% OA, 0.82 KC, 0.94 PA, 0.85 UA, and 0.89 F1-score. Agriculture/grassland/barren land and urban/developed with F1-scores of 0.87 and 0.86 and UA of 0.85 and 0.82 respectively, while forest with F1-score (0.71) show some confusion with adjacent wetland vegetation.

Overall, object-based RF and GTB showed stronger and more balanced classification performance while object-based SVM also enhanced classification stability. Sentinel-based classification showed improved performance compared to Landsat-based models. The importance of incorporating spatial context, texture, and segment-level information in heterogeneous wetland landscapes is evident through object-based approaches consistently outperformed pixel-based classifications in all ML models included in this study.

Table 4. (b). Comparative accuracy assessment performance of Machine Learning (RF, GTB, SVM) and Deep Learning (DNN) Models derived from Landsat 5 TM for wetland change analysis in Sangamon River watershed (2000–2008).

Confusion Matrix	Wetland (1)	Forest (2)	Agriculture/ Grassland/ Barren land (3)	Urban/Developed (4)	Waer (5)	Producer Accuracy	User Accuracy	Overall Accuracy (%)	F1 Score
Pixel-based									
Random Forest (RF) Kappa Coefficient = 0.57									
Wetland (1)	1358	110	64	25	4	0.87	0.70	67%	0.78
Forest (2)	256	711	102	85	3	0.61	0.62		0.61

Agriculture/ Grassland/ Barren land (3)	132	147	1000	206	1	0.67	0.62		0.64
Urban/Developed (4)	113	135	421	665	1	0.50	0.66		0.57
Water (5)	78	30	21	22	247	0.62	0.96		0.75
Gradient Tree Boosting (GTB) Kappa Coefficient = 0.57									
Wetland (1)	1312	111	83	40	15	0.84	0.70	67%	0.76
Forest (2)	241	670	145	97	4	0.58	0.62		0.60
Agriculture/ Grassland/ Barren land (3)	90	120	1036	239	1	0.69	0.60		0.64
Urban/Developed (4)	100	122	427	685	1	0.51	0.63		0.56
Water (5)	60	22	25	25	266	0.66	0.93		0.77
Support Vector Machine (SVM) Kappa Coefficient = 0.38									
Wetland (1)	1121	168	216	49	8	0.71	0.55	53%	0.62
Forest (2)	336	482	263	105	6	0.40	0.49		0.44
Agriculture/ Grassland/ Barren land (3)	206	90	1092	127	9	0.71	0.48		0.57
Urban/Developed (4)	236	207	635	257	8	0.19	0.46		0.27
Water (5)	113	30	52	17	253	0.54	0.89		0.67
Object-based									
Random Forest (RF) Kappa Coefficient = 0.85									
Wetland (1)	2056	10	17	5	1	0.98	0.89	88%	0.93
Forest (2)	116	703	40	38	3	0.78	0.83		0.80
Agriculture/ Grassland/ Barren land (3)	54	63	1038	33	0	0.87	0.88		0.87
Urban/Developed (4)	48	46	71	753	2	0.82	0.90		0.86
Water (5)	30	19	7	3	231	0.80	0.97		0.88
Gradient Tree Boosting (GTB) Kappa Coefficient = 0.88									
Wetland (1)	2090	14	12	0	0	0.99	0.92	91%	0.95
Forest (2)	48	784	36	25	47	0.87	0.90		0.88
Agriculture/ Grassland/ Barren land (3)	32	21	1147	24	30	0.93	0.91		0.92
Urban/Developed (4)	17	20	35	828	1	0.92	0.92		0.92
Water (5)	8	2	4	2	300	0.95	0.99		0.97
Support Vector Machine (SVM) Kappa Coefficient = 0.84									
Wetland (1)	1973	23	33	21	7	0.96	0.87	88%	0.91
Forest (2)	117	723	40	35	2	0.78	0.89		0.83
Agriculture/ Grassland/ Barren land (3)	89	41	961	39	2	0.85	0.87		0.86
Urban/Developed (4)	82	22	55	731	5	0.81	0.88		0.84
Water (5)	10	4	6	0	265	0.93	0.94		0.93

Dense Neural Networks (DNNs) Kappa Coefficient = 0.50									
Wetland (1)	1294	86	144	27	9	0.83	0.64	61%	0.72
Forest (2)	262	662	170	101	5	0.55	0.59		0.57
Agriculture/ Grassland/ Barren land (3)	219	140	965	173	3	0.64	0.53		0.58
Urban/Developed (4)	146	195	487	486	6	0.36	0.60		0.45
Water (5)	100	34	23	12	251	0.59	0.91		0.72

Table 4. (c). Accuracy Assessment Matrices used in the study.

Confusion Matrix	**Wetland (1)**	**Forest (2)**	**Agriculture/ Grassland/ Barren land (3)**	**Urban/Developed (4)**	**Waer (5)**	**Producer Accuracy**	**User Accuracy**	**Overall Accuracy (%)**	**F1 Score**
Pixel-based									
Random Forest (RF) Kappa Coefficient = 0.77									
Wetland (1)	675	93	13	17	4	0.85	0.82	81%	0.83
Forest (2)	125	415	43	26	1	0.69	0.75		0.72
Agriculture/ Grassland/ Barren land (3)	7	23	591	65	0	0.87	0.85		0.86
Urban/Developed (4)	10	22	46	566	7	0.87	0.82		0.84
Water (5)	9	0	1	10	189	0.90	0.94		0.92
Gradient Tree Boosting (GTB) Kappa Coefficient = 0.77									
Wetland (1)	632	84	12	14	14	0.84	0.80	80%	0.82
Forest (2)	119	457	32	38	1	0.70	0.77		0.73
Agriculture/ Grassland/ Barren land (3)	6	23	629	73	0	0.86	0.84		0.85
Urban/Developed (4)	10	30	74	494	0	0.81	0.79		0.80
Water (5)	29	0	1	4	174	0.84	0.92		0.88
Support Vector Machine (SVM) Kappa Coefficient = 0.54									
Wetland (1)	504	112	81	36	23	0.66	0.62	61%	0.64
Forest (2)	150	339	89	63	6	0.52	0.58		0.55
Agriculture/ Grassland/ Barren land (3)	82	71	411	166	1	0.56	0.57		0.56
Urban/Developed (4)	56	64	150	334	4	0.54	0.55		0.54
Water (5)	18	2	5	8	175	0.84	0.83		0.84
Object-based									
Confusion Matrix	Wetland (1)	Forest (2)	Agriculture/ Grassland/ Barren land (3)	Urban/Developed (4)	Waer (5)	Producer Accuracy	User Accuracy	Overall Accuracy (%)	F1 Score
Random Forest (RF) Kappa Coefficient = 0.82									
Wetland (1)	2022	73	31	17	17	0.94	0.85	87%	0.89
Forest (2)	268	566	32	27	3	0.63	0.81		0.71

Agriculture/ Grassland/ Barren land (3)	26	33	1044	80	0	0.88	0.88		0.88
Urban/Developed (4)	31	24	68	766	10	0.85	0.85		0.85
Water (5)	27	0	0	8	277	0.88	0.90		0.89
Gradient Tree Boosting (GTB) Kappa Coefficient = 0.84									
Wetland (1)	1984	123	24	13	16	0.92	0.88	88%	0.90
Forest (2)	203	641	26	23	3	0.70	0.77		0.73
Agriculture/ Grassland/ Barren land (3)	17	38	1042	86	0	0.87	0.90		0.88
Urban/Developed (4)	22	25	54	788	11	0.87	0.86		0.86
Water (5)	25	0	0	5	282	0.89	0.90		0.89
Support Vector Machine (SVM) Kappa Coefficient = 0.84									
Wetland (1)	2123	7	18	0	0	0.98	0.97	88%	0. 97
Forest (2)	55	810	3	0	0	0.93	0.65		0.77
Agriculture/ Grassland/ Barren land (3)	3	412	789	41	1	0.63	0.96		0.76
Urban/Developed (4)	0	1	10	868	0	0.98	0.90		0.94
Water (5)	0	0	0	48	251	0.83	0.99		0.91
U-Net DL Model Kappa Coefficient = 0.89									
Wetland (1)	2145	42	16	5	2	0.97	0.92	90%	0.94
Forest (2)	96	781	18	4	0	0.87	0.93		0.90
Agriculture/ Grassland/ Barren land (3)	11	72	1112	36	1	0.90	0.95		0.92
Urban/Developed (4)	3	4	44	873	1	0.94	0.95		0.94
Water (5)	1	0	0	12	298	0.96	0.99		0.97

3.4. Variable Importance (VI) and Relative Importance (RI) in RF Models

The comparative assessment of RF model variable importance (VI) in Table 4 (d) between L5 (2000–2008) and S2 (2017–2025) datasets indicate OBIA significantly improved model stability, spatial consistency, and predictor utilization for wetland classification in heterogeneous watershed environments by reducing OOB error from 0.34 to 0.11 in Landsat-based model, 0.19 to 0.13 in Sentinel-based model, respectively.

3.4.1. Landsat 5 TM (2000–2008)

Table 4 (d) and Figure 5 indicate that for both pixel-based and object-based Landsat-based RF models, slope showed the highest RI (~5.60%), followed by DEM (~4.88%). With RI (~2.4–2.9%), SWIR1, SWIR2, NIR, and NDVI are important spectral predictors because of their sensitivity to surface moisture conditions, inundation, and land-surface detection in wetlands. GLCM are also significant in distinguishing fragmented wetland systems of spatial heterogeneity, particularly in OBIA model. AWEI, LSWI, MNDWI, NDMI, MSAVI, and BSI showed moderate but stable VI, whereas DistWater with low VI, showing that terrain and spectral characteristics provided stronger discriminatory capability than simple proximity-based hydrologic variables.

3.4.2. Sentinel-2 SR (2017–2025)

Table 4 (d) and Figure 5 indicate that S2 imagery had an improved and balanced RI of spectral and GLCM variables. With RI ~5%, Slope and DEM were the most influential variables. A comparatively higher VI was observed for GLCM summer contrast and entropy showing improved sensitivity to fine-scale spatial variability and fragmented landscape patterns. SWIR1, SWIR2, NIR, and NDVI-derived variables showed high VI, particularly for OBIA, with improved classification of wetlands, forests, agriculture, and urban categories. MNDWI, NDMI, LSWI, and AWEI signify improved VI delineation of inundated and moisture-rich environments compared to SAR variables and SAR texture metrics, which showed very low VI.

Table 4. (d). Variable Importance (VI) and Relative Importance (RI) (%) of variables used in the Random Forest (RF) model derived from Landsat 5 TM (2000–2008) and Sentinel-2 SR (2017–2025).

		Landsat 5 TM (2000–2008)				Sentinel-2 SR (2017–2025)			
		Pixel-based		Object-based		Pixel-based		Object-based	
No.	Variables	RF VI	RI (w.r.t. Slope) (%)	RF VI	RI (w.r.t. Slope) (%)	RF VI	RI (w.r.t. DEM) (%)	RF VI	RI (w.r.t. Slope)) (%)
1	AWEI	1998.51	2.50	1392.28	2.50	681.78	2.77	637.84	2.57
2	Blue_spring	2171.77	2.53	1483.51	2.67	563.83	2.80	730.06	2.95
3	Blue_summer	2162.47	2.25	1388.20	2.50	480.97	2.75	709.05	2.86
4	DEM	3898.17	4.88	3572.73	6.44	690.84	5.30	1300.62	5.26
5	DistWater	139.03	0.17	43.12	0.07	677.04	3.58	765.77	3.09
6	EVI_summer	1883.86	2.36	1183.89	2.36	1305.76	2.32	590.51	2.38
7	Green_spring	2192.45	2.74	1537.69	2.74	880.79	2.75	669.41	2.70
8	Green_summer	2117.58	2.65	1350.96	2.65	572.64	2.49	655.60	2.65
9	LSWI_spring	2014.94	2.52	1277.33	2.52	677.02	2.35	605.54	2.44
10	LSWI_summer	1813.86	2.27	1108.57	2.27	613.79	1.97	509.25	2.05
11	MNDWI	2048.35	2.56	1531.15	2.56	579.81	2.80	719.04	2.90
12	NDBI	1777.19	2.22	1091.20	2.22	495.99	2.01	526.67	2.13
13	NDMI_spring	2071.63	2.59	1562.58	2.59	580.54	2.36	611.86	2.47
14	NDMI_summer	1817.91	2.28	1091.03	2.28	483.51	1.96	521.09	2.10
15	NDVI_spring	1960.67	2.45	1357.21	2.45	765.15	3.11	714.26	2.88
16	GLCM_spring_contrast:	2033.63	2.55	1411.11	2.55	743.57	3.02	719.85	2.91
17	GLCM_spring_ent	1999.56	2.50	1398.42	2.50	718.72	2.92	695.94	2.81
18	GLCM_spring_idm	1995.57	2.50	1389.26	2.50	743.08	3.02	716.18	2.89
19	NDVI_summer	1980.28	2.48	1224.30	2.48	651.94	2.65	647.28	2.61
20	GLCM_summer_contrast	1999.84	2.50	1411.39	2.50	1052.80	4.27	893.82	3.61
21	GLCM_summer_ent	1985.15	2.48	1369.51	2.48	702.47	2.85	680.04	2.75
22	GLCM_summer_idm	2019.73	2.53	1418.07	2.53	733.13	2.98	732.11	2.96
23	NIR_spring	2288.32	2.87	1409.68	2.87	664.49	2.70	707.84	2.86
24	NIR_summer	2272.23	2.85	1370.25	2.85	556.39	2.26	595.85	2.41
25	Red_spring	2156.80	2.70	1511.61	2.70	664.83	2.70	690.66	2.79
26	Red_summer	2043.31	2.56	1357.56	2.56	675.30	2.74	683.77	2.76
27	SWIR1_spring	2297.34	2.88	1458.11	2.88	739.67	3.00	774.10	3.13
28	SWIR1_summer	1966.76	2.46	1371.00	2.46	668.18	2.71	636.96	2.57
29	SWIR2_spring	2253.85	2.82	1418.44	2.82	768.42	3.12	765.47	3.09

30	SWIR2_summer	2031.24	2.54	1350.84	2.54	710.72	2.88	626.44	2.53
31	Slope	4470.74	5.60	3362.89	5.60	1363.57	5.54	1473.20	5.95
32	BSI_spring	2024.63	2.53	1312.47	2.36	563.83	2.29	611.16	2.47
33	BSI_summer	1798.20	2.25	1113.93	2.00	480.97	1.95	533.10	2.15
34	MSAVI_spring	2062.33	2.58	1307.49	2.58	687.67	2.79	1245.09	2.81
35	MSAVI_summer	1901.76	2.38	1210.56	2.38	567.14	2.30	1049.83	2.31
36	UI_spring	2092.16	2.62	1278.16	2.62				
37	UI_summer	1848.08	2.31	1177.39	2.31				
38	NDVI_diff	2135.98	2.67	1568.33	2.67				
39	NDMI_diff:	2130.13	2.50	1562.58	2.50				
40	VH					0	0.00	0	0.00
41	VV					0.15	0.00	0.15	0.00
42	VV_VH					0	0.00	0	0.00
43	VV_contrast					0.98	0.00	0.98	0.00
44	VV_ent					1.33	0.00	1.33	0.00
45	VV_idm					2.09	0.00	2.09	0.00
	Out of bag error estimate:	0.34		0.11		0.19		0.13	

3.5. Change Detection and Wetland Loss Calculations (2000–2008-2017-2025)

We selected the object-based RF classified maps for both periods (2000–2008) and (2017-2025) for spatiotemporal wetland change analysis due to its comparatively higher and balanced classification performance (strong OA/Kappa, balanced PA and UA, and stable F1-scores) and spatial consistency, although numerically RF and GTB computed similar values. We calculated the wetland gain, wetland loss to all other classes, and stable wetland area for 2000 and 2017 in Km2 as shown in Table 5 (a) and (b). The result showed major wetland loss and fragmentation in Sangamon River Watershed between 2000 and 2025 as approximately 54.85% wetlands (860.68 km^2) area was lost to other land uses. Other land uses that lost their spatial extent includes forest with 46.51% (694.88 km^2), and water bodies with 36.80% (66.13 km^2), whereas agriculture/grassland/barren covered area witnessed an increase of 8.56% (820.62 km^2), and urban/developed land with 72.92% (799.52 km^2) as shown in Table 6 (a) based on their existing land use in year 2000.

Table 5 (b) represents a larger percentage of 87.90% of stable wetlands existed between 2000-2008 than 59.67% between 2017-2025 periods, indicating major changes in wetland occurred during the latter period. Between 2000-2008, 3.18% of wetland converted to forest, 7.57% wetland converted to agriculture/grassland/barren land, 1.25% of wetland converted to urban, 1.07% of wetland lost to other water bodies. While for periods between 2017-2025, 11.54% of wetland converted to forest, 22.91% to agriculture/grassland/barren land, 4.70% to urban, 1.18% to other water bodies. The wetland loss percentage was calculated based on the total wetland area of the base year i.e., 2000 and 2017, respectively.

Overall, the results indicate that Sangamon River Watershed witnessed substantial environmental and landscape changes in the past 25 years in the form of a significant 54.85% wetland loss in addition to a proportional forest cover loss. While agricultural practices intensified continuously throughout this period, as the most dominant land-cover class, an alarming rate of urban expansion caused significant anthropogenic pressure on watershed ecosystems and hydrologically sensitive landscapes. Figure 6 spatialized substantial landscape transformation within the Sangamon River Watershed between 2000 and 2025.

Table 5. (a). Change detection and wetland loss calculations in the Sangamon River Watershed, Illinois.

Class name	**Area (km²) in 2000**	**% of Land use**	**Area (km²) in 2008**	**% of Land use**	**Area (km²) in 2017**	**% of Land use**	**Area (km²) in 2025**	**% of Land use**	**Change in area between 2000-2025**	**% of Land use change based on the year 2000**	**Increase /decrease**
Wetland (1)	1569.2436	11%	1551.91	11.15%	771.9249	5.54%	708.5648	5.09%	860.6788	54.85%	Decrease
Forest (2)	1494.0486	11%	1499.76	10.78%	729.5228	5.24%	799.1661	5.74%	694.8825	46.51%	Decrease
Agriculture/ Grassland/ Barren land (3)	9575.7489	69%	9563.58	68.75%	10246.1084	73.66%	10396.3652	74.74	820.6163	8.56%	Increase
Urban/Developed (4)	1096.3863	8%	1141.87	8.20%	2051.9583	14.75%	1895.9107	13.62	799.5244	72.92%	Increase
Class Water (5)	179.700	1%	157.98	1.13%	113.3917	0.81%	113.5674	0.81	66.1326	36.80%	Decrease
Total watershed area: 13,915.13											

doi:10.20944/preprints202606.1288.v1

Table 5. (b). Wetland loss calculations to other land uses in the Sangamon River Watershed, Illinois.

Class Name	Area (km²) 2000-2008	% of watershed loss to other land uses	Area (km²) 2017-2025	% of watershed loss to other land uses
Stable Wetland between 2000-2008	1,378.7964	87.90%		
Stable Wetland between 2017-2025			460.6786	59.67%
Wetland to Forest (2)	50.4567	3.18%	88.794	11.54%
Wetland to Agriculture/ Grassland/ Barren Land (3)	119.9862	7.57%	176.7847	22.91%
Wetland to Urban (4)	18.3033	1.25%	36.2248	4.70%
Wetland to Water (5)	17.01	1.07%	9.1491	1.18%
Total	1,584.5526	100%	771.6312	100%
Total watershed area: 13,915.13				

4. Discussion

This study represents one of the first comprehensive multi-temporal wetland and LULC change analyses conducted for the Sangamon River Watershed to identify spatial patterns of wetland loss, agricultural expansion, forest cover loss, urban, climate and anthropogenic pressure in the watershed and aims to support evidence-based decision-making for wetland conservation and restoration, climate resilience planning, agricultural best management practices and sustainable land-use policy development. Most existing studies on wetland change analysis have relied on a limited set of classification methods that incorporates comparative framework integrating ML, DL, pixel-based and OBIA classification. This study addresses these gaps by comprehensively evaluating and comparing multiple ML and DL classifiers across both pixel-based and object-based approaches all in one framework to have an accurate wetland classification and change detection in a complex agricultural watershed environment. This study demonstrates that integrating multi-source Earth observation datasets with advanced GeoAI-based (ML & DL) classification models significantly improved wetland delineation and that both DL and ML models can produce significant accuracy in classification models based on spatiotemporal data and variable selection, however, object-based classification performance has consistently outperformed pixel-based methods.

For this study we strategically divide the analysis period into two temporal phases (2000–2008 and 2017–2025) based on the availability of the satellite datasets. For 2000 and 2008 we incorporated L5 imagery since it provides one of the longest continuous Earth observation archives for long-term analysis and S2 imagery is not available years prior to 2017. For 2017 and 2025 we used S2 integrated with S1 SAR data. For wetland change computations, we incorporated classified models in ArcGIS Pro while retaining spatial resolution of two temporal phases. However, as evident in Table 4 (a) and (b), that S2 data enabled improved delineation of fragmented wetlands, narrow riparian corridors, small hydrologically connected depressions, and land-cover transitions, compared to L5 imagery. L5 imagery tends to overlap land-cover classes, leading to underestimation of smaller wetland patches and boundary inaccuracies, while S2 imagery reduces mixed-pixel effects. Consequently, many small or narrow wetlands embedded within agricultural and riparian landscapes may have remained undetected or poorly classified in Landsat-based analyses (Figure 5). The integration of SAR backscatter variables also improved the model accuracy in identification of saturated soils, flooded vegetation, and hydrologically connected wetland environments, particularly in areas where spectral confusion existed between wetlands, agriculture, and riparian vegetation. GLCM (contrast, entropy, and IDM) by incorporating spatial variability, vegetation structure, and surface roughness characteristics also enhanced identification of riparian vegetation gradients and heterogeneous assemblages, wetland edge conditions, distinguishing fragmented and heterogeneous wetland environments that are not fully represented through spectral information of Sentinel imagery alone.

Between 2000–2025, Sangamon River Watershed experienced significant hydrologic alteration and ecological vulnerability with 54.85% wetland decline along with 46.51% forest cover and 36.80% water bodies loss out of total watershed area. Table 6 (a) shows that wetlands remain comparatively stable during 2000-2008, with a significant shift of 50.25% between 2008-2017, while the forest cover is also proportionally declined. During the 25-year period, 72.92% urban expansion in this region added impervious surface expansion, which consequently increased pollutant loading, altered runoff regimes, increasing flood vulnerability, pressure on riparian corridors, along with steady increase in agricultural land: the two key drivers of wetland conversion. This transition is directly associated with hydrologic and habitat connectivity, loss of biodiversity, sediment transport, low nutrient retention, riparian instability, increased erosion, declining overall ecosystem resilience in the watershed.

On the computational analysis side, this study classified and compared a set of 14 models. Each model was independently calibrated based on its algorithm-specific parameters, input predictor variables, sampling strategy, and spatial classification settings, to achieve an optimized output. For ML-based models, these calibrations include numbers of trees adjustment (500-550), learning rates, bag fractions (0.7), kernel functions, and regularization parameters. For DL-based models, we used standard architecture-specific optimization including normalization, dropout regularization, batch normalization, learning-rate scheduling, class weighting, and patch dimensions. Each model was rerun with multiple adjustments until an improved classification accuracy is achieved.

Overall, terrain variables (slope and DEM) with the highest VI for all RF classifications, while Sentinel-1 SAR variables and SAR texture metrics) showed comparatively low VI, nevertheless, SAR variables may still provide complementary information for wetland delineation under cloud-contaminated conditions and for capturing hydrologic variability. Spectral indices such as MNDWI, NDMI, LSWI, AWEI, MSAVI and GLCM also showed strong contributions in Sentinel-based analyses showing improved delineation of inundated and moisture-rich environments.

Object-based RF and GTB produced comparable classification accuracies and wetland change computation for both Landsat and Sentinel-based analyses, the RF model for both periods was selected for wetland change analysis and change detection, because it generated more spatially coherent and visually realistic wetland patterns with reduced classification noise and improved landscape continuity.

As established by previous studies, object-based models consistently produced improved classification robustness and reduced uncertainty through segmentation-based spatial aggregation. The object-based GTB classifier produced the strongest overall performance among the ML models, followed closely by RF and SVM, while U-Net semantic segmentation model with an OA of approximately 90% outperformed all ML models with improved delineation of fragmented wetlands. Among the pixel-based ML models, RF and GTB produced strong and relatively balanced classification results. These findings indicate that integrating object-level spatial information with DL and ML approaches significantly enhances wetland and LULC classification performance within complex watershed environments. The improved Sentinel-based classification results further show the importance of combining multispectral, SAR, terrain, hydrologic, and texture variables for accurate wetland and LULC classification.

Some of the computational and technical challenges of the study encountered include frequent time-out in GEE of the SNIC segmentation algorithm, reduceConnectedComponents() operations and object-level feature extraction using large predictor stacks, high-resolution S2 imagery, SAR variables, terrain data and texture metrics that require high memory usage. Segmentation of S2 imagery generated extremely large numbers of image objects, which increased computational overhead compared to Landsat-based analyses. To resolve the issue, we subdivided the workflow into two segments; one where we trained the data, and second, we perform the classification using assets import (trainingsets) from GEE Assets. While U-Net substantially improved wetland delineation and spatial coherence, the computational resources, processing time, and workflow complexity required for large-scale DL implementation remain considerably higher than those associated with ML models. Each epoch/iteration in TensorFlow training and classification in Google Colab for U-Net took 4 hours. Ideally U-net model must be run for at least 30 iterations, but due to time limitations we stopped after 10 iterations as the models stabilized and achieved 90-92% OA. Also, exporting TFRecord image patches from GEE requires a careful check as it may introduce some inconsistencies in patch generation. Data formatting and exporting of large TFRecord and all other datasets also require large storage capacity of G-drive. Runtime disconnections are a big hurdle for cloud-based platforms. Overall, these issues required careful improvisation strategies to balance computational feasibility and classification performance.

These findings emphasize that although DL and OBIA methods can improve classification accuracy and spatial realism, their operational implementation for regional-scale environmental monitoring remains constrained by computational infrastructure, cloud-processing limitations, storage capacity, scalability, and workflow reproducibility challenges. Future advances in cloud-native geospatial computing, GPU accessibility, distributed processing, and scalable AI frameworks may help reduce these limitations and improve the operational feasibility of large-scale GeoAI-based wetland monitoring applications.

5. Conclusions

This study concludes that integrating Sentinel and Landsat imagery, SAR-derived hydrologic input, object-based methods, ML and DL models substantially improved wetland delineation and classification accuracy. Sentinel-based classification models showed improved performance compared to Landsat-based analyses due to the higher spatial resolution, improved spectral separability, incorporation of SAR backscatter data, and enhanced object-level spatial representation. Object-based ML and DL models enhanced wetland classification accuracy and produced more spatially coherent classification outputs and reduced pixel-level noise than pixel-based models but require considerably greater computational resources and adjustments due to workflow complexity. Therefore, their broader application remains dependent on advances in computational infrastructure, scalable cloud-based processing, and efficient geospatial AI workflows. In conclusion, object-based DL methods are more effective than pixel-based ML methods in classifying complex wetland landscapes.

The wetland change analysis revealed significant 54.85% wetland and 54.85% forest losses in addition to continued expansion of agricultural and urban land uses during the 25-year study period, making this watershed more ecologically vulnerable for water quality, flood regulation, biodiversity conservation, and climate resilience. From a planning, conservation, and environmental management perspective, this study provides one of the first comprehensive evaluations of wetland dynamics in the Sangamon River Watershed using an integrated GeoAI framework. The developed methodology offers a transferable approach for wetland monitoring, watershed-scale environmental assessment and future predictions for further wetland loss and can support evidence-based decision-making, restoration prioritization, and sustainable land-use planning in agricultural watersheds facing increasing environmental and climate-related pressures.

Author Contributions: Conceptualization, A.S.; Methodology, A.S.; Software, A.S.; Validation, A.S.; Formal Analysis, A.S.; Investigation, A.S.; Resources, R.A.; Data Curation, A.S.; Writing—Original Draft Preparation, A.S. Writing—Review & Editing, A.S. and R.A.; Visualization, A.S.; Supervision, R.A.; Project Administration, R.A.; Funding Acquisition, R.A. All authors have read and agreed to the published version of the manuscript.”.

Funding: This research received no external funding.

Data Availability Statement: The links for the GEE/Colab codes and datasets are available through: Sadaf, Afsheen (2026), “GeoAI-Driven Wetland Change Analysis in the Sangamon River Watershed (2000–2025): A

Comparative Assessment of Machine Learning and Deep Learning Approaches.", Mendeley Data, V1, doi: 10.17632/sjdbtm73p4.1.

Acknowledgments: I would like to express my sincere gratitude to my supervisor and co-author, Dr. Reda Amer, for his steadfast support and invaluable guidance throughout every phase of this research. The authors also gratefully acknowledge Zachary Ridings (UMSL Geospatial Collaborative) for his technical assistance and support with software installations, and technology-related aspects of the project. I would also like to thank Paula from Esri Support and Zachary Ridings for their extensive support and long troubleshooting sessions. Their technical guidance was greatly appreciated. We are grateful to Dr. Jason Knouft of Saint Louis University, Dr. Jenna Shelton of the Illinois State Water Survey, and Dr. Karen Johannesson of the University of Massachusetts Boston for their valuable insights, constructive discussions, and expertise.

We further acknowledge the United States Geological Survey (USGS), National Land Cover Database (NLCD), National Aeronautics and Space Administration (NASA), COPERNICUS, Google Earth Engine (GEE) and Google Colab for providing open-access datasets, computational resources, and cloud-based geospatial processing capabilities that made this research possible.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
AWEI	Automated Water Extraction Index
BSI	Bare Soil Index
CART	Classification and Regression Trees
CM	Confusion Matrix
FCN	Fully Convolutional Networks
CNN	Convolutional Neural Networks
DL	Deep learning
DNNs	Dense Neural Networks
DEM	Digital Elevation Model
DEP	Digital Elevation Prodcut
EVI	Enhanced Vegetation Index
GEE	Google Earth Engine
GTB	Gradient Tree Boosting
GLCM	Gray-Level Co-occurrence Matrix
GRD	Ground Range Detected
HUC	Hydrologic Unit Code
IDM	Inverse Difference Moment
KC	Kappa coefficient
LSWI	Land Surface Water Index
LULC	Land Use Land Change
L5	Landsat5
ML	Machine learning
MNDWI	Modified Normalized Difference Water Index
MSAVI	Modified Soil Adjusted Vegetation Index
NLCD	National Land Cover Data
NIR	Near Infrared
NDBI	Normalized Difference Built-up Index
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
NAD	North American Datum
OBIA	Object-Based Image Analysis
OA	Overall Accuracy

PA	Producer's Accuracy
QA	Quality Assessment
RF	Random Forest
RNN	Recurrent Neural Networks
RI	Relative Importance
SLR	Sea-level rise
S2	Sentinel-2
SWIR	Short wave Infrared
SVM	Support Vector Machines
SR	Surface Reflectance
SAR	Synthetic Aperture Radar
TFRecord	TensorFlow Record
TM	Thematic Mapper
US	United States
USGS	United States Geological Survey
UI	Urban Index
UA	User's Accuracy
VI	Variable Importance

References

1. Groot, R. D., Stuip, M., Finlayson, M., & Davidson, N. (2006). Valuing wetlands: guidance for valuing the benefits derived from wetland ecosystem services (pp. ii+-45).
2. Mitsch, W. J., Bernal, B., & Hernandez, M. E. (2015). Ecosystem services of wetlands. International Journal of Biodiversity Science, Ecosystem Services & Management, 11(1), 1-4.
3. Barbier, E. B. (2019). The value of coastal wetland ecosystem services. In Coastal wetlands (pp. 947-964). Elsevier.
4. Xu, X., Chen, M., Yang, G., Jiang, B., Zhang, J., 2020. Wetland ecosystem services research: a critical review. Glob. Ecol. Conserv. 22, e01027. https://doi.org/10.1016/j.gecco.2020.e01027.
5. Ballut-Dajud, G. A., Sandoval Herazo, L. C., Fernández-Lambert, G., Marín-Muñiz, J. L., López Méndez, M. C., & Betanzo-Torres, E. A. (2022). Factors affecting wetland loss: A review. Land, 11(3), 434.
6. Dahl, T. E. (1990). Wetlands losses in the United States, 1780's to 1980's. US Department of the Interior, Fish and Wildlife Service.
7. Gibbs, J. P. (2000). Wetland loss and biodiversity conservation. Conservation biology, 14(1), 314-317.
8. Zedler, J.B., Kercher, S., 2005. Wetland RESOURCES: status, trends, ecosystem services, and restorability. Annu. Rev. Environ. Resour. 30, 39–74. https://doi.org/10.1146/annurev.energy.30.050504.144248. Volume 30, 2005.
9. Hu, S., Niu, Z., Chen, Y., Li, L., & Zhang, H. (2017). Global wetlands: Potential distribution, wetland loss, and status. Science of the total environment, 586, 319-327.
10. Fluet-Chouinard, E., Stocker, B. D., Zhang, Z., Malhotra, A., Melton, J. R., Poulter, B., ... & McIntyre, P. B. (2023). Extensive global wetland loss over the past three centuries. Nature, 614(7947), 281-286. DOI: https://dx.doi.org/10.1080/15481603.2021.1947623.
11. Wilkening, J. L., Esralew, R. A., & Newman, B. F. (2026). From assessment to action: informing water resource management in protected areas amid global change. Frontiers in Water, 7, 1719814.
12. Groschen, G. E., Harris, M. A., King, R. B., Terrio, P. J., & Warner, K. L. (2000). Water Quality in the Lower Illinois River Basin, Illinois, 1995-98 (No. 1209). US Geological Survey.
13. Borah, D. K., Bera, M., & Shaw, S. (2003). Water, sediment, nutrient, and pesticide measurements in an agricultural watershed in Illinois during storm events. Transactions of the ASAE, 46(3), 657.
14. Dahl T.E., Allord G.J. 1996. History of wetlands in the conterminous United States. Pages 19–25 in Fretwell J.D., Williams J.S., Redman P.J. (compilers). National Water Summary on Wetland Resources Water-Supply Paper 2425. Washington, DC: U.S. Geological Survey.

15. Farda, N. M. (2017, December). Multi-temporal land use mapping of coastal wetlands area using machine learning in Google earth engine. In IOP Conference Series: Earth and Environmental Science (Vol. 98, No. 1, p. 012042). IOP Publishing.
16. Islam, H., Abbasi, H., Karam, A., Chughtai, A. H., & Ahmed Jiskani, M. (2021). Geospatial analysis of wetlands based on land use/land cover dynamics using remote sensing and GIS in Sindh, Pakistan. Science Progress, 104(2), 00368504211026143.
17. Christensen, J.R., Golden, H.E., Alexander, L.C., Pickard, B.R., Fritz, K.M., Lane, C.R., Weber, M.H., Kwok, R.M., Keefer, M.N., 2022. Headwater streams and inland wetlands: status and advancements of geospatial datasets and maps across the United States. Earth Sci. Rev. 235, 104230. https://doi.org/10.1016/j.earscirev.2022.104230.
18. Jafarzadeh, H., Mahdianpari, M., Gill, E. W., Brisco, B., & Mohammadimanesh, F. (2022). Remote sensing and machine learning tools to support wetland monitoring: A meta-analysis of three decades of research. Remote Sensing, 14(23), 6104.
19. Sultana, S., & Inayathulla, M. (2022). Precision land use and land cover classification using Google Earth Engine: Integrating random forest and support vector machine algorithms. Geo. Eye, 11(2), 9-14.
20. Salas, E. A. L., Kumaran, S. S., Bennett, R., Willis, L. P., & Mitchell, K. (2024). Machine learning-based classification of small-sized wetlands using Sentinel-2 images. AIMS Geosciences, 10(1), 62.
21. Day, J. W., Britsch, L. D., Hawes, S. R., Shaffer, G. P., Reed, D. J., & Cahoon, D. (2000). Pattern and process of land loss in the Mississippi Delta: a spatial and temporal analysis of wetland habitat change. Estuaries, 23(4), 425-438.
22. Mahdianpari, M., Granger, J.E., Mohammadimanesh, F., Salehi, B., Brisco, B., Homayouni, S., Gill, E., Huberty, B., Lang, M., 2020. Meta-Analysis of wetland classification using remote sensing: a systematic review of a 40-Year trend in North America. Remote Sens. 12 (11), 1882. https://doi.org/10.3390/rs12111882.
23. Montgomery, J., Mahoney, C., Brisco, B., Boychuk, L., Cobbaert, D., & Hopkinson, C. (2021). Remote sensing of wetlands in the prairie pothole region of North America. Remote Sensing, 13(19), 3878.
24. Mattson, M., Sousa, D., Quandt, A., Ganster, P., & Biggs, T. (2024). Mapping multi-decadal wetland loss: Comparative analysis of linear and nonlinear spatiotemporal characterization. Remote Sensing of Environment, 302, 113969.
25. Ghorbanian, A., Zaghian, S., Asiyabi, R. M., Amani, M., Mohammadzadeh, A., & Jamali, S. (2021). Mangrove ecosystem mapping using Sentinel-1 and Sentinel-2 satellite images and random forest algorithm in Google Earth Engine. Remote sensing, 13(13), 2565.
26. Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS journal of photogrammetry and remote sensing, 67, 93-104.
27. Delalay, M., Tiwari, V., Ziegler, A. D., Gopal, V., & Passy, P. (2019). Land-use and land-cover classification using Sentinel-2 data and machine-learning algorithms: operational method and its implementation for a mountainous area of Nepal. Journal of Applied Remote Sensing, 13(1), 014530-014530.
28. Tu, Y., Lang, W., Yu, L., Li, Y., Jiang, J., Qin, Y., ... & Xu, B. (2020). Improved mapping results of 10 m resolution land cover classification in Guangdong, China using multisource remote sensing data with Google Earth Engine. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 5384-5397.
29. Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y. A., & Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations—A review. Remote sensing, 12(7), 1135.
30. Chasmer, L., Mahoney, C., Millard, K., Nelson, K., Peters, D., Merchant, M., Hopkinson, C., Brisco, B., Niemann, O., Montgomery, J., Devito, K., Cobbaert, D., 2020. Remote sensing of boreal wetlands 2: methods for evaluating boreal wetland ecosystem State and drivers of change. Remote Sens. 12 (8), 1321. https://doi.org/10.3390/rs12081321.
31. Sultan, M., Saleous, N., Issa, S., Dahy, B., & Sami, M. (2025). Optimizing land use classification using Google Earth Engine: A comparative analysis of machine learning algorithms. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 863-869.

32. Zhao, Z., Islam, F., Waseem, L. A., Tariq, A., Nawaz, M., Islam, I. U., ... & Hatamleh, W. A. (2024). Comparison of three machine learning algorithms using google earth engine for land use land cover classification. Rangeland ecology & management, 92, 129-137.
33. Dronova, I. (2015). Object-based image analysis in wetland research: A review. Remote Sensing, 7(5), 6380-6413.
34. Delalay, M., Tiwari, V., Ziegler, A. D., Gopal, V., & Passy, P. (2019). Land-use and land-cover classification using Sentinel-2 data and machine-learning algorithms: operational method and its implementation for a mountainous area of Nepal. Journal of Applied Remote Sensing, 13(1), 014530-014530.
35. Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y. A., & Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations—A review. Remote sensing, 12(7), 1135.
36. Mahdianpari, M., Jafarzadeh, H., Granger, J. E., Mohammadimanesh, F., Brisco, B., Salehi, B., ... & Weng, Q. (2020). A large-scale change monitoring of wetlands using time series Landsat imagery on Google Earth Engine: a case study in Newfoundland. GIScience & Remote Sensing, 57(8), 1102-1124. https://doi.org/10.1080/15481603.2020.1846948
37. Ballanti, L., Byrd, K. B., Woo, I., & Ellings, C. (2017). Remote sensing for wetland mapping and historical change detection at the Nisqually River Delta. Sustainability, 9(11), 1919.
38. Chasmer, L., Mahoney, C., Millard, K., Nelson, K., Peters, D., Merchant, M., Hopkinson, C., Brisco, B., Niemann, O., Montgomery, J., Devito, K., Cobbaert, D., 2020. Remote sensing of boreal wetlands 2: methods for evaluating boreal wetland ecosystem State and drivers of change. Remote Sens. 12 (8), 1321. https://doi.org/10.3390/rs12081321.
39. Fickas, K. C., Cohen, W. B., & Yang, Z. (2016). Landsat-based monitoring of annual wetland change in the Willamette Valley of Oregon, USA from 1972 to 2012. Wetlands ecology and management, 24(1), 73-92.
40. Zerrouk, M., Ait El Kadi, K., Sebari, I., & Fellahi, S. (2025). Machine and Deep Learning for Wetland Mapping and Bird-Habitat Monitoring: A Systematic Review of Remote-Sensing Applications (2015–April 2025). Remote Sensing, 17(21), 3605.
41. Pashaei, M., Kamangir, H., Starek, M. J., & Tissot, P. (2020). Review and evaluation of deep learning architectures for efficient land cover mapping with UAS hyper-spatial imagery: A case study over a wetland. Remote sensing, 12(6), 959.
42. Yuan, S., Liang, X., Lin, T., Chen, S., Liu, R., Wang, J., ... & Gong, P. (2025). A comprehensive review of remote sensing in wetland classification and mapping. arXiv preprint arXiv:2504.10842.
43. Hossain, M. D., & Chen, D. (2019). Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. ISPRS journal of photogrammetry and remote sensing, 150, 115-134.
44. Duro, D. C., Franklin, S. E., & Dubé, M. G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote sensing of environment, 118, 259-272.
45. Zheng, J.-Y., Hao, Y.-Y., Wang, Y.-C., Zhou, S.-Q., Wu, W.-B., Yuan, Q., Gao, Y., Guo, H.-Q., Cai, X.-X., & Zhao, B. (2022). Coastal wetland vegetation classification using pixel-based, object-based and deep learning methods based on RGB-UAV. Land, 11(11), 2039. https://doi.org/10.3390/land11112039
46. Badruzzaman, A., Wulandari, P., Sainal, S., Ashley, M., Jobling, S., Austen, M. C., & Praptiwi, R. A. (2025). Satellite imagery pre-processing and feature extraction for the mapping of coastal ecosystems using Google Earth Engine: A workflow for practitioners. MethodsX, 103516.
47. Shafizadeh-Moghadam H, Khazaei M, Alavipanah SK, Weng Q. Google Earth Engine for large-scale land use and land cover mapping: an objectbased classification approach using spectral, textural and topographical factors. GIScience & Remote Sensing. 2021;58(6):914–928.
48. Tu, Y., Lang, W., Yu, L., Li, Y., Jiang, J., Qin, Y., ... & Xu, B. (2020). Improved mapping results of 10 m resolution land cover classification in Guangdong, China using multisource remote sensing data with Google Earth Engine. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 5384-5397.

49. Berra, E. F., Fontana, D. C., Yin, F., & Breunig, F. M. (2024). Harmonized Landsat and Sentinel-2 Data with Google Earth Engine. Remote Sensing, 16(15), 2695.
50. Wright, N., Duncan, J. M., Callow, J. N., Thompson, S. E., & George, R. J. (2024). CloudS2Mask: A novel deep learning approach for improved cloud and cloud shadow masking in Sentinel-2 imagery. Remote Sensing of Environment, 306, 114122.
51. Pettorelli, N. (2013). The normalized difference vegetation index. Oxford University Press, USA.
52. Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International journal of remote sensing, 27(14), 3025-3033.
53. Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International journal of remote sensing, 24(3), 583-594.
54. Wilson, E. H., & Sader, S. A. (2002). Detection of forest harvest type using multiple dates of Landsat TM imagery. Remote Sensing of Environment, 80(3), 385-396.
55. Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. Remote sensing of environment, 140, 23-35.
56. Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote sensing of environment, 83(1-2), 195-213.
57. Nguyen, C. T., Chidthaisong, A., Kieu Diem, P., & Huo, L. Z. (2021). A modified bare soil index to identify bare land features during agricultural fallow-period in southeast Asia using Landsat 8. Land, 10(3), 231.
58. Qi, J., Chehebouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). Modified soil adjusted vegetation index (MSAVI) Remote Sensing of Environment, vol. 48.
59. Kawamura, M., Jayamanna, S., & Tsujiko, Y. (1997). Quantitative evaluation of urbanization in developing countries using satellite data. Doboku Gakkai Ronbunshu, 1997(580), 45-54.
60. Xiao, X., Boles, S., Liu, J., Zhuang, D., & Liu, M. (2002). Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. Remote sensing of environment, 82(2-3), 335-348.
61. Adeli, S., Quackenbush, L. J., Salehi, B., & Mahdianpari, M. (2022). The importance of seasonal textural features for object-based classification of wetlands: New York state case study. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 43, 471-477.
62. Denghao Yang , Nan Zhou , Zhiyu Zhu , Huilin Ge , WenChao Wang , Cheng Xu & Jing Zhang (2025) Coastal wetland classification method based on UAV imagery: integrating hierarchical sample enhancement and multiscale sample selection techniques, Geomatics, Natural Hazards and Risk, 16:1, 2585167, DOI: 10.1080/19475705.2025.2585167
63. Dewitz, J., and U.S. Geological Survey, 2021, National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021): U.S. Geological Survey data release, doi:10.5066/P9KZCM54.