



ARTICLE

Evaluating the Capability of Sentinel-3 as an Alternative to MODIS for Downscaling High Spatiotemporal Resolution LST Data Using ESTARFM and XGBoost Models

Nahid Haghshenas and Ali Shamsoddini*

Department of RS and GIS, Tarbiat Modares University, Tehran, Iran

*Corresponding Author: Ali Shamsoddini. Email: ali.shamsoddini@modares.ac.ir

Received: 14 November 2025; Accepted: 23 February 2026; Published: 21 May 2026

ABSTRACT: This study aimed to evaluate the potential of Sentinel-3 as an alternative to Moderate Resolution Imaging Spectroradiometer (MODIS) for generating high spatiotemporal resolution land surface temperature (LST) data. The Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) and the machine-learning-based Extreme Gradient Boosting (XGBoost) algorithm were independently assessed for fusing MODIS–Landsat and Sentinel-3–Landsat data. This comparison enabled the evaluation of each model’s capability to reconstruct spatiotemporal LST variations and assess the performance of the two sensors in the fusion process. The results showed that XGBoost outperformed ESTARFM in capturing complex and heterogeneous LST patterns, particularly under strong diurnal fluctuations and phenological differences. The mean Root Mean Square Error (RMSE) values for MODIS were 1.87 Kelvin (K) and 2.62 K using XGBoost and ESTARFM, respectively, while for Sentinel-3, they were 1.73 and 2.52 K, confirming XGBoost’s superior accuracy for both sensors. Sentinel-3, owing to its higher spatial resolution, better radiometric quality, earlier overpass time closer to Landsat-8/9 acquisitions, and improved angular effect control, more accurately reconstructed spatial variations and daily temperature dynamics. In contrast, MODIS, with its broader temporal coverage and larger dynamic range, provided more stable Spatiotemporal Fusion (STF) results with slightly higher mean values. The model transferability analysis showed that training models with MODIS data and applying them to Sentinel-3 yielded higher accuracy than the reverse configuration, highlighting the importance of sensor selection and model generalizability. Overall, the findings indicate that Sentinel-3 can serve as a viable alternative to MODIS for STF of LST data under data gaps or reduced-quality time series. Moreover, integrating data from both sensors can ensure the continuity and stability of downscaled LST products and mitigate data limitations in long-term monitoring.

KEYWORDS: Spatiotemporal fusion; MODIS; Sentinel-3; LST; ESTARFM; XGBoost

1 Introduction

LST is a key parameter for understanding and monitoring climatic, hydrological, and biophysical processes. It plays a fundamental role in studies of climate change, agriculture, natural resource management, and environmental monitoring [1]. Access to LST data with high accuracy, appropriate spatial and temporal resolution, and long-term continuity is essential for applied research at both regional and global scales. However, the inherent limitations of satellite sensors typically offer either high spatial or high temporal resolution, but rarely both simultaneously, posing a major challenge for generating consistent and high-quality LST imagery [2].

Although Landsat LST data are widely used as reference datasets in many studies due to their relatively high spatial resolution (100 m), their temporal resolution (16-day revisit cycle), and high sensitivity to cloud contamination limit their independent use for continuous monitoring of land surface dynamic processes, particularly in agricultural and hydrological applications [3]. In contrast, lower spatial resolution satellite data, such as MODIS and Sentinel-3, provide near-daily temporal coverage but, due to their coarse spatial resolution, are unable to capture thermal variations at field and sub-field scales [4]. This inherent trade-off between spatial and temporal resolution underscores the need for STF approaches to generate LST imagery with high resolution in both dimensions [5].

Since 2006, the development of Spatiotemporal Fusion (STF) models aimed at producing datasets with high resolution in both spatial and temporal dimensions has attracted considerable attention [6]. Among various sources, the MODIS sensors onboard the Terra and Aqua satellites have served as the primary global source of LST data for more than two decades, providing near-daily temporal coverage [7]. Many STF models were originally developed using MODIS and Landsat data and have since been widely applied across diverse fields, including agricultural monitoring [8], environmental assessment [9], urban studies [10], and land management [11,12]. Despite the notable success of the MODIS mission, the gradual degradation of its radiometric quality, increased radiometric noise in the thermal bands, orbital drift, and the approaching end of its operational lifetime (expected around 2025–2026) have underscored the urgent need for reliable alternative sensors to ensure the continuity of long-term LST time series [13,14].

In addition to limitations related to the operational lifespan of the sensor, MODIS-derived LST products face well-known technical challenges that constrain their use in applications requiring high spatial resolution, particularly in agricultural studies. These limitations include: (1) cloud contamination and undetected thin clouds, which introduce biases in LST estimates; (2) the relatively coarse spatial resolution of MODIS, which obscures sub-field thermal variability; (3) reduced sensitivity to land cover heterogeneity and management practices such as irrigation; and (4) the direct propagation of errors from the MODIS base imagery into STF model outputs [15]. This highlights the importance of employing alternative sensors and advanced STF approaches to mitigate systematic errors.

NASA's Land Discipline Team has sought to ensure the continuity of global land observations through the Visible Infrared Imaging Radiometer Suite (VIIRS). However, VIIRS exhibits several limitations in the thermal domain, including the discontinuation of the split-window LST product (MxD11) and its primarily afternoon overpass. These limitations pose challenges for applications requiring morning observations, such as energy balance assessment, evapotranspiration estimation, and agricultural monitoring (e.g., crop growth and water stress detection) [16]. Consequently, the use of complementary missions, particularly Sentinel-3 (SLSTR), is essential to preserve the continuity and consistency of LST observations.

The Sentinel-3 mission, equipped with the Ocean and Land Colour Instrument (OLCI) and the Sea and Land Surface Temperature Radiometer (SLSTR), provides a data stream comparable to that of Terra-MODIS. Sentinel-3A (launched in 2016) and Sentinel-3B (launched in 2018) operate in a 10:00 a.m. sun-synchronous orbit, while Sentinel-3C and Sentinel-3D are scheduled for launch in 2026 and 2027, respectively. As a result, Sentinel-3 data will remain available at least until 2031, ensuring long-term continuity of land surface observations [16]. The SLSTR sensor, with its advanced split-window algorithms, view-angle corrections, and surface emissivity parameterization, has demonstrated strong capabilities in producing accurate and stable LST data [17].

Given the differences in overpass time, imaging geometry, and LST retrieval algorithms, evaluating the performance of Sentinel-3 relative to MODIS has become increasingly important in the context of STF modeling. In recent years, numerous STF approaches have been proposed, ranging from weight-based

algorithms such as ESTARFM to machine learning and deep learning methods, which are particularly valued for their ability to model complex and nonlinear relationships [18].

The present study aims to assess the potential of Sentinel-3 as a substitute for MODIS within the framework of spatiotemporal LST fusion modeling. Using Landsat data as a reference, the performance of two representative approaches, ESTARFM and XGBoost, was evaluated and compared in generating high-resolution LST imagery. The innovative aspects of this research include:

1. A systematic evaluation of Sentinel-3’s capability to replace MODIS under conditions involving phenological differences, daily temperature fluctuations, and varying temporal gaps between image pairs;
2. A comparative analysis of weight-based and machine learning models in addressing cross-sensor discrepancies and spatiotemporal inconsistencies; and
3. An assessment of model transferability and stability under real-world conditions, emphasizing generalization and practical applicability for long-term and complex time-series analyses.

Overall, this study provides important empirical and practical evidence in the field of STF of LST, which not only contributes to improving the accuracy of existing models but also paves the way for developing forward-looking strategies in data source selection and cross-sensor transferability.

2 Study Area and Data

2.1 Study Area

The study area is located in the Yanco region, within the Murrumbidgee River catchment in southeastern Australia. Covering approximately 40 km × 40 km (Fig. 1), the region is predominantly composed of irrigated agricultural lands and pastures. It lies between latitudes 34.185° S and 34.980° S and longitudes 145.834° E and 146.769° E. The topography is generally flat, with elevations ranging from 117 to 150 m and gentle slopes [19,20].

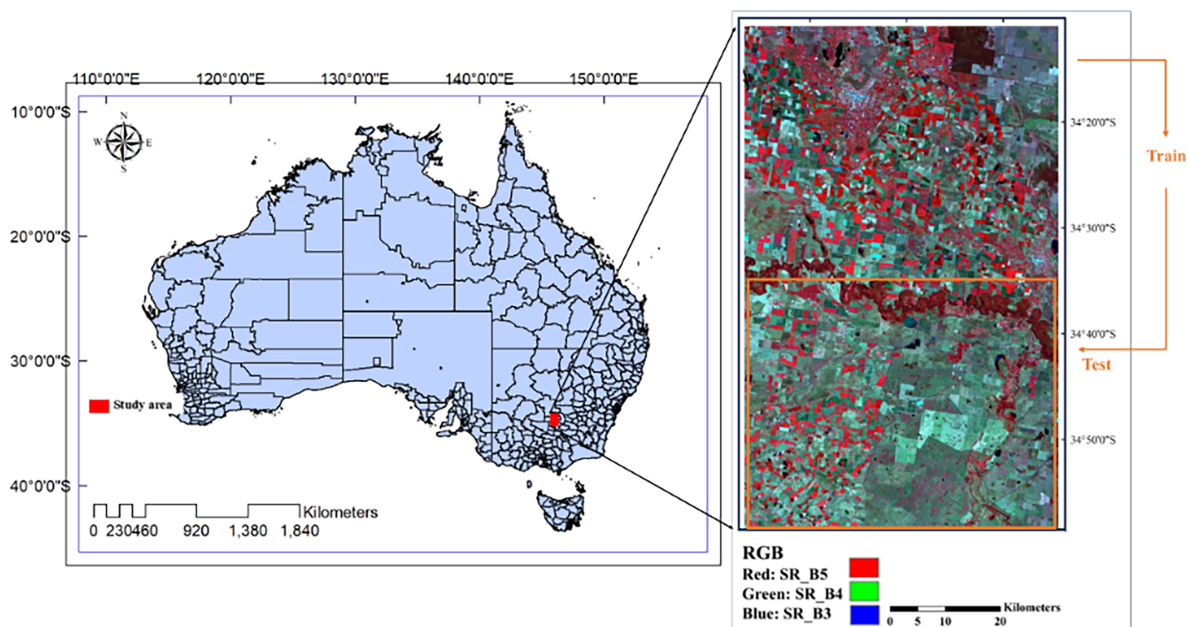


Figure 1: The study area is located in the Yanco region, in New South Wales, Australia, shown in the Universal Transverse Mercator (UTM) coordinate system (WGS84, Zone 55S).

The area experiences a dry continental climate, characterized by an annual average precipitation of 418.5 mm, which predominantly occurs in late autumn and winter. Daily temperatures vary from approximately 13.5°C in July (winter) to 32.1°C in January (summer) [21,22].

2.2 Data

In this study, image triplets from the MODIS, Landsat, and Sentinel-3 sensors were acquired during the summer and early autumn of 2021–2024. Detailed information about this dataset is provided in [Table 1](#).

Table 1: LST datasets and their characteristics.

Date of Acquisition	Satellite	Data Collection	Seasons	Spatial Resolution (m)	LST Retrieval Algorithm	Daily Mean Skin Temperature Variability (K)	Time Period (Day)
17 Jan 2021	LANDSAT_8 Terra/MODIS Sentinel-3B	Landsat 8 LST product MOD11A SLSTR LST Product	Summer	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 21.83 SD = 7.86	48
2 Feb 2021	LANDSAT_8 Terra/MODIS Sentinel-3B	Landsat 8 LST product MOD11A SLSTR LST Product	Autumn	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 18.15 SD = 6.49	
6 Mar 2021	LANDSAT_8 Terra/MODIS Sentinel-3B	Landsat 8 LST product MOD11A SLSTR LST Product	Autumn	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 23.06 SD = 8.50	
12 Jan 2022	LANDSAT_8 Terra/MODIS Sentinel-3A	Landsat 8 LST product MOD11A SLSTR LST Product	Summer	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 15.24 SD = 5.54	32
20 Jan 2022	LANDSAT_9 Terra/MODIS Sentinel-3A	Landsat 9 LST product MOD11A SLSTR LST Product	Summer	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 17.22 SD = 6.10	
13 Feb 2022	LANDSAT_8 Terra/MODIS Sentinel-3A	Landsat 8 LST product MOD11A SLSTR LST Product	Summer	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 20.34 SD = 6.99	
15 Jan 2023	LANDSAT_9 Terra/MODIS Sentinel-3A	Landsat 9 LST product MOD11A SLSTR LST Product	Summer	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 23.18 SD = 8.41	96
31 Jan 2023	LANDSAT_9 Terra/MODIS Sentinel-3B	Landsat 9 LST product MOD11A SLSTR LST Product	Summer	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 24.20 SD = 8.87	
21 Apr 2023	LANDSAT_9 Terra/MODIS Sentinel-3B	Landsat 9 LST product MOD11A SLSTR LST Product	Autumn	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 19.56 SD = 6.95	
3 Feb 2024	LANDSAT_9 Terra/MODIS Sentinel-3A	Landsat 9 LST product MOD11A SLSTR LST Product	Summer	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 26.60 SD = 9.94	51
19 Feb 2024	LANDSAT_9 Terra/MODIS Sentinel-3A	Landsat 9 LST product MOD11A SLSTR LST Product	Summer	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 25.61 SD = 9.32	
22 Mar 2024	LANDSAT_9 Terra/MODIS Sentinel-3A	Landsat 9 LST product MOD11A SLSTR LST Product	Autumn	100 m 1000 m 1000 m	Single-channel Split-Window Split-Window	Range = 20.81 SD = 7.57	

2.2.1 Landsat LST Data

Landsat 8 and Landsat 9 were launched on 12 February 2013 and 27 September 2021, respectively, and both provide LST data at 100 m spatial resolution through thermal infrared (TIR) sensors. The satellites

cross the equator at around 10:00 a.m. local time and offer consistent global coverage with a 185 km swath. The selected LST data are Level-2 Collection 2 products derived using the single-channel algorithm (ST RIT v1.3.0), which incorporate Top-of-Atmosphere (TOA) reflectance, brightness temperature, the Global Emissivity Database (GED), and atmospheric profiles. The data were obtained from the Google Earth Engine (GEE) cloud-based platform.

2.2.2 MODIS LST Data

The MODIS sensors onboard the Terra (launched 1999) and Aqua (launched 2002) satellites have provided global LST data for over two decades. The MOD11A1 (Terra) and MYD11A1 (Aqua) products, with 1 km spatial resolution, are derived from thermal bands 31 (10.78–11.28 μm) and 32 (11.77–12.27 μm) using the split-window algorithm, and have a reported LST accuracy of less than 1.3 K for homogeneous surfaces. Terra crosses the equator at approximately 10:30 a.m. local time, while Aqua crosses at around 1:30 p.m.; both have a swath width of approximately 2330 km, enabling near-daily global coverage. In this study, MOD11A1 and MYD11A1 data were used based on availability within the required time periods. The MODIS LST datasets (MOD11A1 and MYD11A1, Version 6.1) were accessed via the Google Earth Engine (GEE) cloud-based platform.

2.2.3 Sentinel-3 LST Data

The SLSTR instruments onboard Sentinel-3A and Sentinel-3B, part of the European Union (EU) Copernicus program, have been operational since 2016 and 2018, respectively. Using a conical scan with dual view and a swath width of approximately 1420 km, SLSTR provides LST data at 1 km spatial resolution from thermal infrared bands. Advanced split-window and view-angle correction algorithms enable accurate and stable LST retrievals. Key advantages include precise calibration, high thermal stability, and improved atmospheric correction compared to sensors such as MODIS. The daily global Sea and Land Surface Temperature Level-2 product (SL_2_LST) includes LST estimates with quality flags and achieves an accuracy better than 1 K under stable atmospheric conditions. In this study, Sentinel-3 LST data were obtained from the Copernicus Open Access Hub (<https://scihub.copernicus.eu/>). Fig. 2 shows the LST histograms for MODIS, Landsat, and Sentinel-3 datasets used in this study.

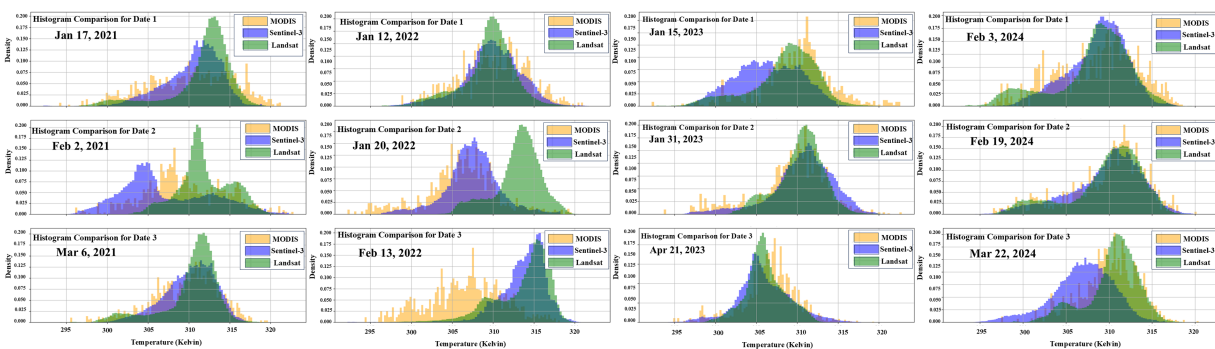


Figure 2: Histograms of landsat, MODIS, and sentinel-3 LST data for the study periods.

2.2.4 Weather Forecasts ReAnalysis 5 Land Portion (ERA5-Land) Hourly Temperature Data

ERA5 data, which represent reanalysis products from the European Centre for Medium-Range Weather Forecasts (ECMWF), are provided at an hourly temporal resolution and a spatial resolution of approximately 0.25° [23]. In this study, hourly temperature data were extracted, and daily mean, standard deviation, and

range (the difference between daily maximum and minimum) were calculated to analyze diurnal thermal variations for diagnostic and descriptive purposes. It should be noted that these data were not directly used for training or prediction in the LST modeling; they were solely employed to examine thermal patterns and support diagnostic analyses. The data were obtained in a gridded Network Common Data Form (NetCDF) format from the Copernicus Climate Data Store (CDS).

3 Methodology

As illustrated in Fig. 3 and the corresponding workflow diagram, this study utilized LST data derived from MODIS, Sentinel-3, and Landsat sensors to predict high-resolution LST maps. In the first step, all satellite images underwent geometric correction and were then resampled to a uniform spatial resolution of 100 m, ensuring spatial consistency among datasets and enabling accurate multi-sensor comparison and modeling.

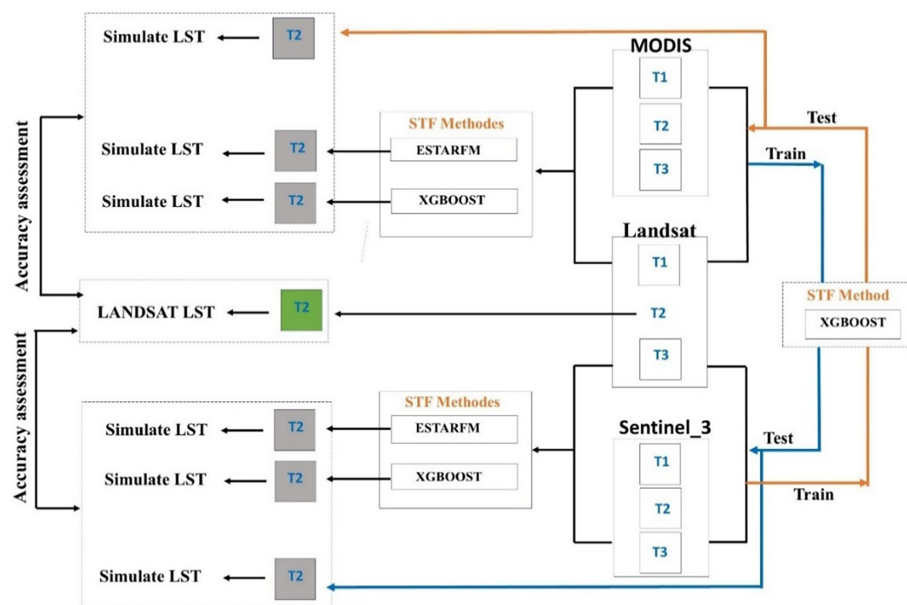


Figure 3: Research workflow diagram (T1 represents the time before prediction, T3 represents the time after prediction, and T2 represents the target date for LST prediction).

In the second step, the data were organized on an annual basis. For each year, three high-resolution Landsat images and six low-resolution images, including three MODIS and three Sentinel-3 images, were considered. In each modeling cycle, two high-resolution Landsat images at times T1 and T3, along with three low-resolution images at times T1, T2, and T3, were used. In this framework, T1 represents the time before prediction, T3 the time after prediction, and T2 the target date for LST prediction.

In the third step, the LST downscaling process was performed using two Spatiotemporal fusion approaches: the ESTARFM algorithm and the XGBoost machine learning model. ESTARFM was applied independently for both MODIS–Landsat and Sentinel-3–Landsat combinations. In each execution, two temporally corresponding pairs of high- and low-resolution images at T1 and T3, along with the low-resolution image at the target date (T2), were used as input data. The generated LST outputs for each sensor combination were then compared with the Landsat-derived reference LST at T2 to evaluate algorithm performance.

Simultaneously, the XGBoost machine learning model was developed for each target date. The training dataset included high-resolution Landsat images at T1 and T3 as the input variables and low-resolution

MODIS or Sentinel-3 images at T1, T2, and T3 as input variables. After training, XGBoost was applied to generate LST maps at 100 m resolution for the target date (T2). This procedure was performed independently for both MODIS–Landsat and Sentinel-3–Landsat combinations, and the predicted results were compared with the reference Landsat images.

In the fourth step, to assess cross-sensor transferability, the trained XGBoost model was evaluated under two scenarios: (1) trained with Sentinel-3–Landsat data and tested with MODIS–Landsat data, and (2) trained with MODIS–Landsat data and tested with Sentinel-3–Landsat data. For each scenario, LST maps were generated for the target dates and compared with the corresponding Landsat observations to assess the model's generalization and robustness across sensors.

Finally, six sets of downscaled LST outputs were obtained for each target date, representing different combinations of sensors and algorithms. This framework allowed for a comprehensive evaluation of algorithm performance, assessment of cross-sensor consistency, and analysis of the potential interoperability and transferability of satellite-derived LST data.

3.1 Applied Algorithms

3.1.1 ESTARFM

ESTARFM is an improved version of the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) algorithm, developed to fuse satellite images with different spatial and temporal resolutions. The algorithm requires two pairs of high- and low-resolution images acquired on two base dates, as well as low-resolution images from the target prediction dates. In this model, land-cover changes are modeled assuming a linear temporal trend. During implementation, the spectral change coefficient for each pixel is calculated and used to transfer information between images of different resolutions. To improve prediction accuracy, the algorithm employs a moving window approach that identifies spectrally similar pixels, assigns appropriate weights, and incorporates them into the estimation of the central pixel value [24].

3.1.2 XGBoost

XGBoost is an advanced machine learning algorithm based on gradient boosted decision trees, developed by Chen and Guestrin (2016). It represents an optimized implementation of the Gradient Boosting Machine (GBM) framework, combining multiple weak decision trees to produce a strong and accurate predictive model. By incorporating regularization techniques, XGBoost effectively prevents overfitting and enhances model generalization. Features such as efficient tree pruning, depth control, and parallelized block processing further improve computational performance, enabling rapid training even on large datasets. Owing to these advantages, XGBoost has become one of the most widely used and robust algorithms for regression, classification, and other complex data analysis tasks [25].

The intrinsic properties of boosting-based algorithms render XGBoost exceptionally well-suited for Spatiotemporal fusion. This algorithm is capable of capturing complex nonlinear interactions among temperature, surface characteristics, and temporal variables, demonstrates strong resilience to multicollinearity, and maintains high generalization performance even under irregular temporal sampling [26,27]. Owing to these strengths, XGBoost is particularly effective for fusing low- and high-resolution sources and producing predictions with high spatial accuracy and fine-grained detail.

In this study, the XGBoost model was specifically applied to satellite-derived LST data to enhance spatial resolution and predict high-resolution LST maps. For each target date (T2), the input features included low-resolution LST observations from MODIS or Sentinel-3 acquired before (T1) and after (T3) the target date, along with low-resolution LST data corresponding to the prediction date (T2). In addition, high-resolution

Landsat LST images at T1 and T3 were used as input high-resolution information. During the model training process, Landsat-derived LST images at the target date (T2) were considered as the reference output (target variable).

The XGBoost model was trained independently for each target date and for each sensor combination (MODIS–Landsat and Sentinel-3–Landsat). To optimize model performance, the main hyperparameters of XGBoost, including the number of trees, tree depth, learning rate, and regularization terms, were tuned using a grid search approach. After training, the model was applied to new low-resolution input data to predict LST maps at a spatial resolution of 100 m for the target date (T2). Finally, the predicted LST results were compared with the corresponding Landsat-derived reference LST images at T2 to evaluate model accuracy and reliability.

3.2 Performance Assessment

To evaluate model accuracy, for each prediction date, two pairs of high- and low-resolution images acquired before and after the target date, together with the low-resolution image on the prediction date, were used as inputs to the algorithms. The high-resolution image from the second date of each year (100 m spatial resolution) was considered as the reference.

For the machine learning model, the dataset was divided into training and testing subsets. The model was trained using the training data and subsequently evaluated on the testing data. In parallel, the classical ESTARFM algorithm was applied to the same image pairs to enable a direct comparison between the rule-based and machine learning approaches.

Model performance was assessed using the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Correlation Coefficient (CC), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM), as defined in Eqs. (1)–(5). RMSE and MAE quantify numerical differences between predicted and reference images, with lower values indicating higher accuracy. CC measures linear correlation, where values close to 1 denote strong spectral and structural similarity. PSNR evaluates the quality of the predicted image relative to the reference, with higher values indicating lower distortion. SSIM incorporates local brightness, contrast, and structure, providing a comprehensive measure of visual similarity. These complementary metrics collectively enable a multidimensional evaluation of STF performance, highlighting the strengths and weaknesses of each model [3,28,29].

$$RMSE = \sqrt{\frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [y(i,j) - x(i,j)]^2} \quad (1)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (2)$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [y(i,j) - x(i,j)]^2 \quad (3)$$

$$CC = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (x(i,j) - \mu_x)(y(i,j) - \mu_y)}{\sqrt{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (x(i,j) - \mu_x)^2 (y(i,j) - \mu_y)^2}} \quad (4)$$

$$PSNR = 20 \times \log_{10} \left(\frac{MAX_y}{\sqrt{MSE}} \right) \quad (5)$$

In these equations, $x(i, j)$ and $y(i, j)$ represent the pixel values of the reference image (Landsat) and the predicted image at the position (i, j) , respectively. Also, m and n denote the number of rows and columns of

the image. MAXy indicates the maximum pixel value in the image, while μ_x and μ_y represent the mean pixel values of the real and predicted images.

4 Results

According to Table 2, the downscaled LST results obtained from the STF models based on machine learning (XGBoost) and the ESTARFM algorithm are presented for four different dates over agricultural areas. In this analysis, low-spatial-resolution data derived from MODIS and Sentinel-3 sensors and high-spatial-resolution data from Landsat were utilized. To examine the consistency and efficiency of different datasets in the STF of LST, a comparison was conducted between MODIS and Sentinel-3 data for the selected dates. Furthermore, to assess cross-sensor transferability, two scenarios were evaluated: the Sentinel-3→MODIS model, trained on Sentinel-3 data and tested on MODIS data; and the MODIS→Sentinel-3 model, trained on MODIS data and validated using Sentinel-3 data. It is important to note that for each date, factors such as the temporal gap between reference image pairs, data acquisition times, crop growth stages, and daily temperature fluctuations varied. Careful consideration of these factors is crucial for accurately interpreting the results and for ensuring the reliability and validity of the model outcomes.

Table 2: Results of Spatiotemporal fusion of LST using ESTARFM and XGBoost models from MODIS and Sentinel-3 Data.

B	Data	Method	RMSE (K)	MAE (K)	CC	PNSR	SSIM
19 February 2024	MODIS→Sentinel-3	XGBOOST	1.3938	1.0647	0.9599	25.56	0.9168
	MODIS	XGBOOST	1.4287	1.0986	0.9562	25.35	0.9146
	MODIS	ESTARFM	1.4708	1.0742	0.9535	25.09	0.9151
	Sentinel-3→MODIS	XGBOOST	1.3836	1.0387	0.9546	25.62	0.9190
	SENTINEL	XGBOOST	1.9137	1.8258	0.9900	22.81	0.9719
	SENTINEL	ESTARFM	1.0297	0.8446	0.9887	28.19	0.9787
31 January 2023	MODIS→Sentinel-3	XGBOOST	1.5596	1.1886	0.8647	23.58	0.7837
	MODIS	XGBOOST	1.7692	1.3670	0.8599	22.49	0.7810
	MODIS	ESTARFM	3.4606	3.1713	0.8547	16.66	0.8336
	Sentinel-3→MODIS	XGBOOST	1.8854	1.4976	0.8703	21.94	0.7699
	SENTINEL	XGBOOST	1.7251	1.3402	0.8611	22.71	0.7934
	SENTINEL	ESTARFM	3.4640	3.1679	0.8501	16.65	0.8319
20 January 2022	MODIS→Sentinel-3	XGBOOST	2.4863	2.004	0.7883	19.58	0.6710
	MODIS	XGBOOST	2.2935	1.7920	0.7821	20.29	0.6853
	MODIS	ESTARFM	2.5967	1.9832	0.7340	19.21	0.7395
	Sentinel-3→MODIS	XGBOOST	2.4938	1.9664	0.7575	19.56	0.6577
	SENTINEL	XGBOOST	2.0265	1.5918	0.8093	21.36	0.6902
	SENTINEL	ESTARFM	2.6486	2.0774	0.7583	19.04	0.7376
2 February 2021	MODIS→Sentinel-3	XGBOOST	2.2417	1.7873	0.8964	22.70	0.7584
	MODIS	XGBOOST	2.002	1.4744	0.8557	23.68	0.7658
	MODIS	ESTARFM	2.9634	2.4566	0.6628	20.28	0.7698
	Sentinel-3→MODIS	XGBOOST	2.2878	1.8289	0.7994	22.53	0.7598
	SENTINEL	XGBOOST	1.2778	0.9299	0.9187	27.59	0.8070
	SENTINEL	ESTARFM	2.9654	2.4626	0.6617	20.27	0.7698

4.1 Comparison of MODIS and Sentinel-3 for STF of LST Using ESTARFM

According to Table 2 and Fig. 4, the comparison of evaluation metrics across different dates indicates that the ESTARFM model achieved its best performance in 2024 for both sensors, with RMSE values of 1.47 K for MODIS and 1.02 K for Sentinel-3. The results also demonstrate consistently high prediction accuracy across all indices for this date. The superior performance observed during this period can be attributed to the short temporal interval between the reference images, combined with moderate daily temperature fluctuations and stable crop growth conditions. These factors enabled the model to more accurately capture and predict the spatiotemporal characteristics of LST, resulting in optimal performance.

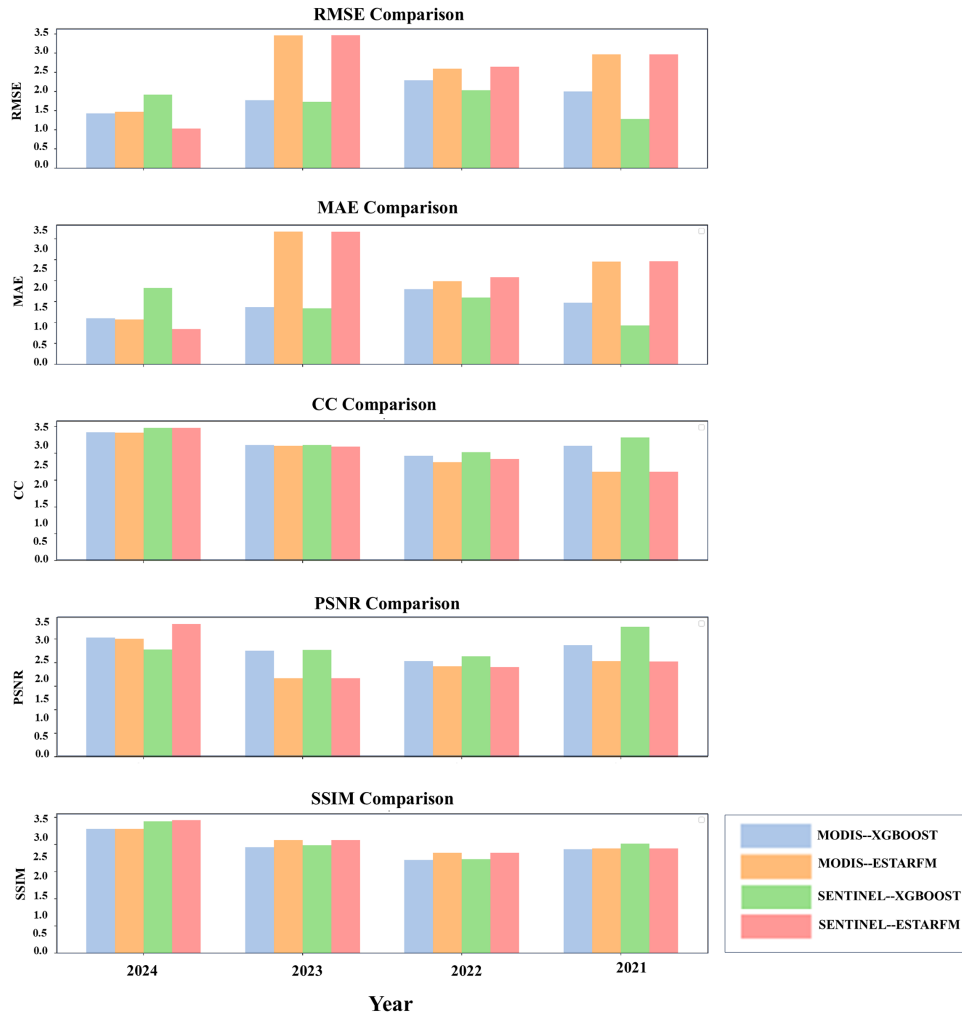


Figure 4: Performance comparison of STF models (ESTARFM and XGBoost) for downscaled LST using MODIS and Sentinel-3 data.

In contrast, the year 2023 exhibited the largest decline in performance, with higher errors and lower quality metrics due to significant variations between the base images. The average RMSE, MAE, and PSNR values for the two sensors were 3.46 K, 3.16 K, and 16.65, respectively. Despite this, the average correlation coefficient (CC = 0.83) remained relatively high, indicating that the model was still able to follow the overall temperature variation trends even under unstable conditions. The year 2022 showed intermediate performance: the average RMSE and MAE values (2.62 and 1.99 K, respectively) were lower than those in

2021 and 2023, but due to spatial heterogeneity in the data, the SSIM index was lower (0.73). This suggests that although the numerical errors were smaller, the model was less effective in accurately predicting the spatial structure.

In 2021, substantial variations between the base images led to a decline in correlation ($CC = 0.66$). Nevertheless, the average PSNR (20) and SSIM (0.76) values were higher than those in 2022, highlighting the model's ability to better capture spatial details, even though it was somewhat less effective in reflecting overall temperature trends. Furthermore, as shown in Table 2 and Fig. 4, the comparison between sensors indicates that, from 2021 to 2023, the performance of STF-derived LST results from MODIS and Sentinel-3 within the ESTARFM model remained nearly identical.

The RMSE values for Sentinel-3 were 2.97, 2.60, and 3.46 K, while those for MODIS were 2.96, 2.65, and 3.46 K, respectively, indicating only minor differences.

However, in 2024, the differences became more pronounced: the RMSE for Sentinel-3 was 1.02 K compared to 1.47 K for MODIS, an approximate difference of 0.45 K, suggesting that Sentinel-3 performed slightly better than MODIS. Spatially, as shown in Fig. 5, the differences between the predicted temperature maps of the two sensors were minimal in the earlier years, but in 2024, a notable divergence in both accuracy and the spatial pattern of downscaled LST became apparent.

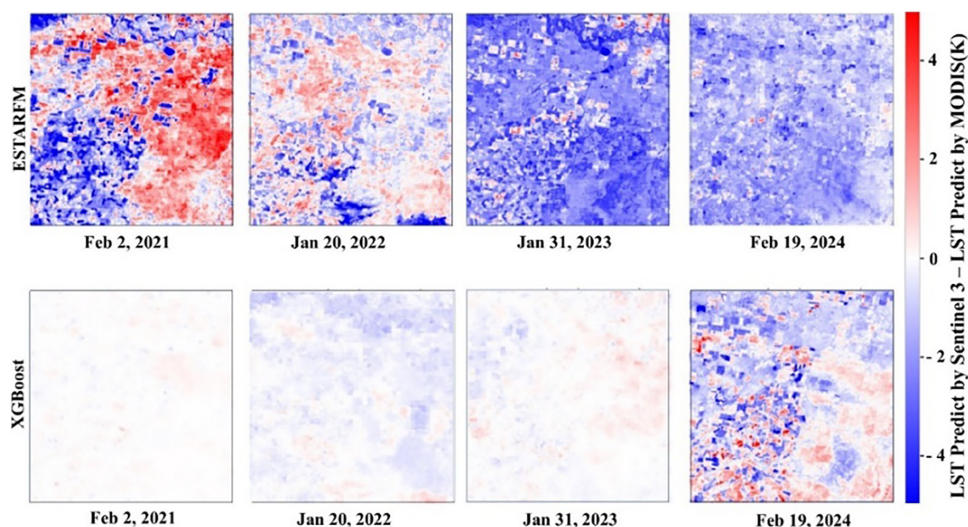


Figure 5: Maps of LST differences derived from MODIS and Sentinel-3 using different models, highlighting warmer MODIS regions (blue), warmer Sentinel-3 regions (red), and areas of high agreement (white).

Fig. 6 illustrates the spatial bias maps of downscaled LST for different dates and for both sensors. In these maps, blue areas represent underestimated values relative to the reference data, while brown areas indicate overestimation. As shown, in 2021, the western regions exhibited slight overestimation for both sensors, whereas the eastern parts displayed noticeable underestimation. In 2022, underestimation occurred mainly in the upper portion of the image, while the remaining areas were mostly overestimated.

The 2023 results revealed that both sensors tended to overestimate the LST across most of the scene. In 2024, spatial discrepancies became more pronounced: the Sentinel-3 prediction map predominantly showed overestimation, whereas MODIS exhibited a mixed pattern of both overestimation and underestimation across different regions. This distinct spatial contrast between the two sensors highlights the model's sensitivity to sensor characteristics and annual environmental variability.

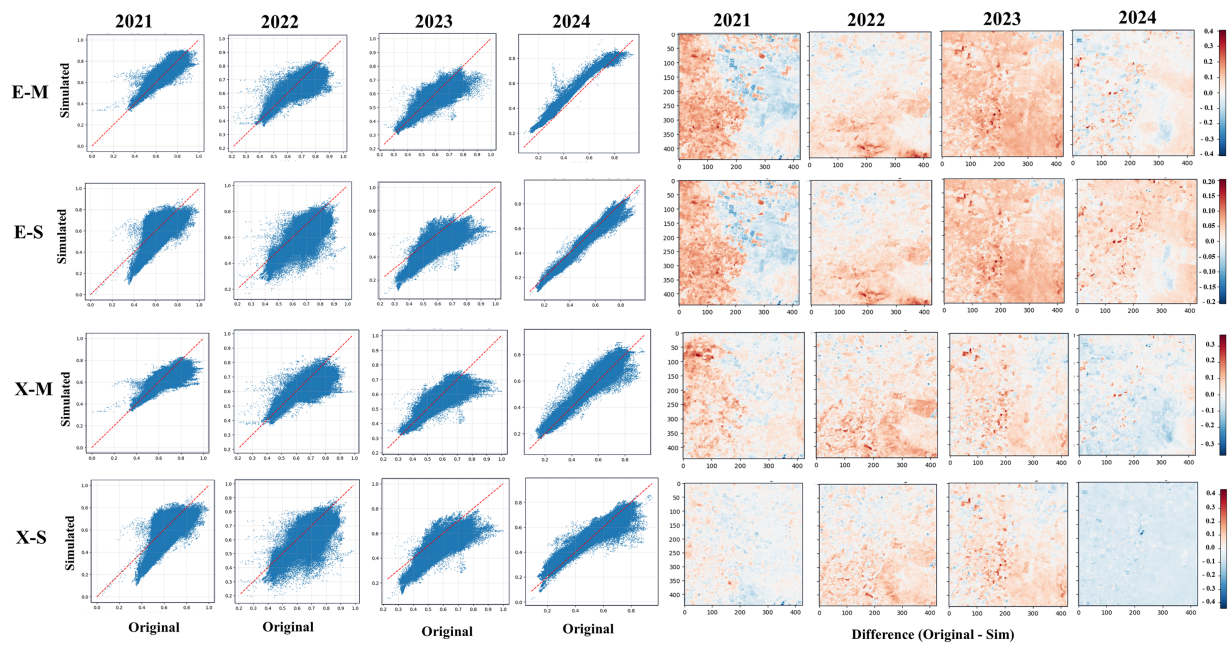


Figure 6: Scatter plots of high-spatiotemporal-resolution LST from MODIS (M) and Sentinel-3 (S), along with the spatial distribution of prediction errors (Observed-Estimated) across years using ESTARFM (E) and XGBoost (X).

4.2 Comparison of MODIS and Sentinel-3 for STF of LST Using XGBoost

According to Table 2 and Fig. 4, the downscaled LST results obtained using the XGBoost machine learning algorithm for the MODIS and Sentinel-3 sensors exhibited a distinct pattern compared with the ESTARFM model. During the period from 2021 to 2023, Sentinel-3 consistently outperformed MODIS, with RMSE values of 1.2778, 2.0265, and 1.7251 K for Sentinel-3, compared with 2.0020, 2.2930, and 1.7692 K for MODIS. However, in 2024, MODIS showed improved performance, achieving an RMSE of 1.4287 K, while Sentinel-3 recorded a slightly higher value of 1.9137 K.

The evaluation metrics (RMSE, MAE, and PSNR) in 2024 indicated better numerical accuracy for MODIS, whereas the structural indices (CC and SSIM) demonstrated Sentinel-3's superior ability to preserve spatiotemporal patterns and spatial details. These findings suggest that although MODIS performed better in minimizing absolute temperature errors, Sentinel-3 was more effective in predicting spatial relationships and maintaining correlation with the reference Landsat data. In other words, error-based indices (RMSE and MAE) reflect only part of the model's accuracy, while structural metrics (CC and SSIM) provide deeper insights into the model's ability to represent spatial structures and temperature patterns.

From a spatial perspective, as shown in Fig. 5, the spatial discrepancies between the two sensors were minor in 2021 and 2022, with MODIS overestimating temperatures in some regions and Sentinel-3 overestimating in others. In 2023 and 2024, Sentinel-3 tended to predict slightly lower LST values than MODIS, indicating a shift in the spatial bias pattern between the two sensors.

According to Fig. 6, the bias relative to the reference image was substantially reduced using the XGBoost model. The magnitude of overestimation errors decreased across all images and gradually shifted toward underestimation. In 2024, the bias was uniformly oriented toward underestimation across the entire image, indicating greater model stability in predicting LST. Overall, these results suggest that although MODIS performed better in predicting average temperature values during 2024 when thermal and phenological variations were relatively small, Sentinel-3 demonstrated superior capability in accurately representing

spatial patterns and fine-scale spatial details. This distinction is particularly important for applications such as agricultural monitoring and spatial temperature variation analysis.

Fig. 7 presents the radar chart of statistical indicators, including SNR, RMSE, MAE, SSIM, and CC for the downscaled LST results from the two STF models. In this chart, higher normalized values correspond to better performance. As shown, the polygon representing Sentinel-3 exhibits a larger area in both the ESTARFM and XGBoost models, indicating higher accuracy, lower error, and stronger correlation with the Landsat reference data. In contrast, MODIS covers a smaller area, reflecting its relatively weaker performance in predicting LST compared with Sentinel-3. A closer examination of the indices reveals that CC, PSNR, and SSIM show the greatest differences between Sentinel-3 and MODIS, clearly indicating that Sentinel-3 not only exhibits lower absolute errors (lower RMSE and MAE) but also performs better in preserving spatial structure, predicting image details, and maintaining spatiotemporal consistency with Landsat. In essence, these indicators demonstrate that Sentinel-3 is more capable of accurately reproducing spatial temperature patterns and temporal trends, whereas MODIS shows lower precision in these aspects.

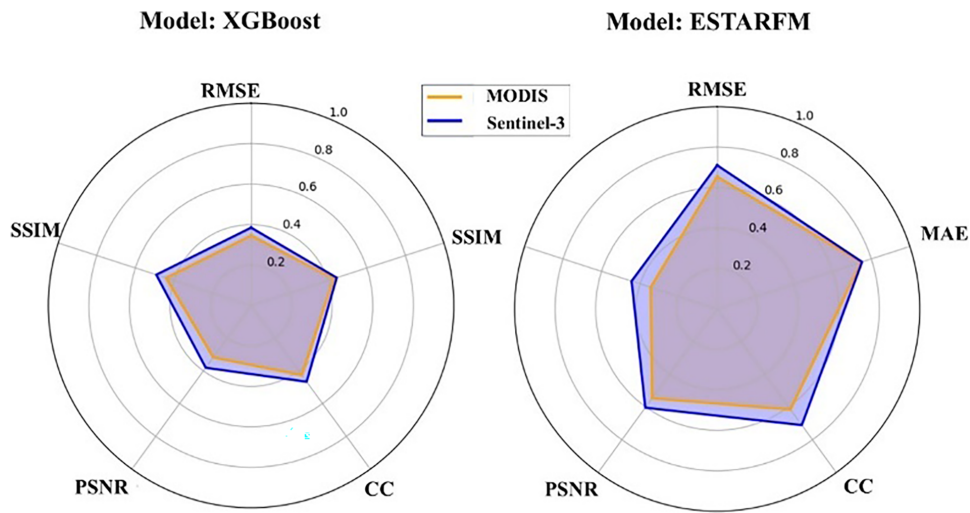


Figure 7: Radar chart of evaluation metrics for STF of LST from MODIS and Sentinel-3.

4.3 Evaluation of MODIS and Sentinel-3 Base Image Pairs Used in STF of LST

Fig. 8 illustrates the spatial distribution of LST differences (Δ LST) between MODIS–Landsat and Sentinel-3–Landsat image pairs. In these maps, lighter areas represent smaller discrepancies and higher agreement with the reference Landsat data, whereas darker areas indicate greater deviations. According to the results, the largest Δ LST values in the MODIS data were observed in pairs 1, 2, 5, and 6, while for Sentinel-3, the highest differences occurred in pairs 2, 5, and 12. Specifically, pairs 2 and 5 exhibited considerable discrepancies for both sensors, whereas pairs 1 and 6 in MODIS and pair 12 in Sentinel-3 showed the greatest mismatch with the Landsat reference imagery. These findings indicate that the base datasets (MODIS and Sentinel-3) inherently contain biases relative to Landsat. In particular, during 2021 and 2022, both sensors tended to overestimate LST, suggesting that such biases could propagate into subsequent fusion results. However, the machine learning models applied in this study effectively mitigated and corrected these deviations. In 2023 and 2024, the overall bias magnitude decreased, and only minor differences were observed except for one Sentinel-3 image in 2024, which still exhibited relatively higher deviations from Landsat. Overall, these results confirm that the machine learning–based fusion models successfully reduced the impact of inherent sensor biases and significantly improved the accuracy of LST downscaling.

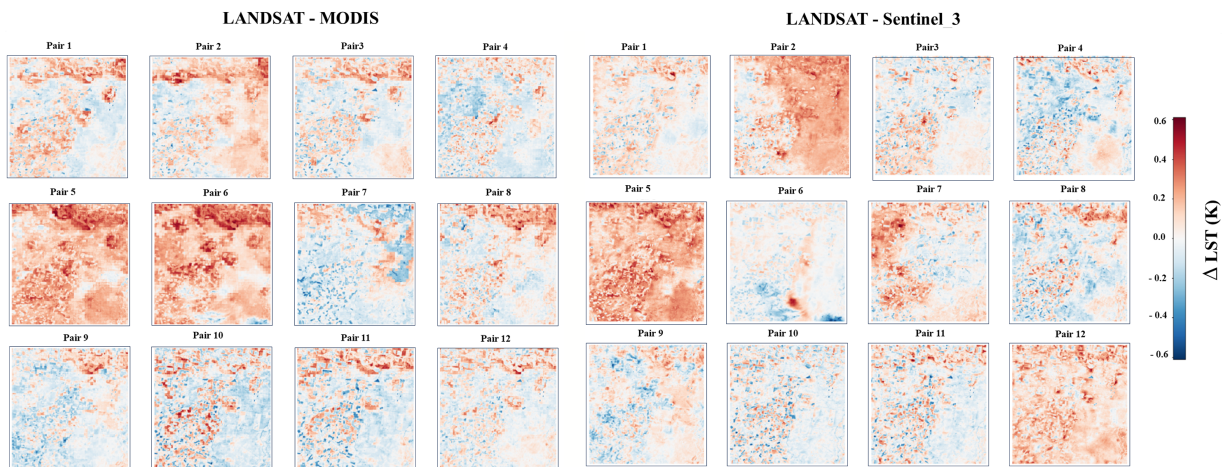


Figure 8: Correlation between MODIS and Sentinel-3 data with reference Landsat observations based on Δ LST (LST Difference) analysis.

Fig. 9 illustrates the Δ LST analysis (mean and standard deviation) between MODIS–Landsat and Sentinel-3–Landsat image pairs, highlighting the spatial and temporal differences between the two sensors. In this analysis, 12 image pairs were used (X-axis of Fig. 9), which correspond to the selected datasets in this study. These include high-resolution Landsat images as the reference and low-resolution images from MODIS or Sentinel-3. The purpose of these pairings was to examine the spatial and temporal discrepancies of each sensor relative to Landsat and to assess their ability to capture LST variations.

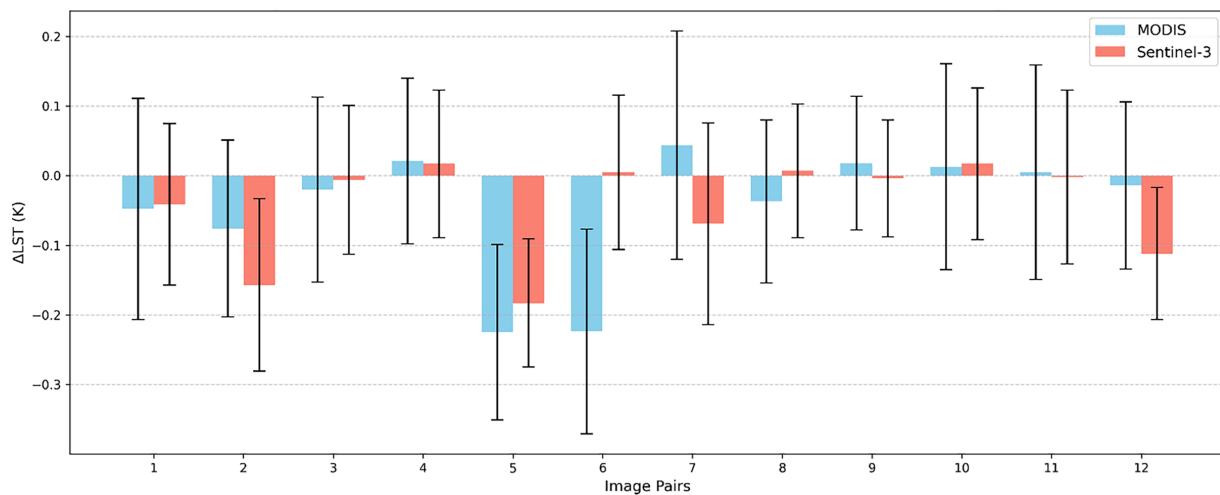


Figure 9: Comparison of Δ LST statistics (Mean and SD) between MODIS and Sentinel-3 relative to Landsat.

It should be noted that these images were used solely for descriptive analysis and were not employed in the training of the XGBoost model or the ESTARFM algorithm. This analysis clearly demonstrates how each sensor records land surface temperature variations and helps reveal the limitations of spatial resolution and sensor sensitivities in representing fine-scale details. In this figure, the blue color represents the differences between MODIS and Landsat, while the red color represents the differences between Sentinel-3 and Landsat for each image pair.

For pairs 1, 2, and 5, Δ LST values for both sensors are predominantly negative, although distinct numerical patterns are observed: in pairs 1 and 3, MODIS exhibits higher Δ LST values, whereas Sentinel-3 shows lower ones. In contrast, for pair 2, Sentinel-3 records a stronger negative Δ LST compared to MODIS. These variations indicate the superior capability of Sentinel-3 in capturing fine-scale spatial variations and detailed LST patterns, while MODIS tends to represent broader-scale fluctuations.

In terms of variability, the standard deviation of Δ LST for MODIS is higher in pairs 1 and 3, reflecting greater spatial fluctuations and wider dispersion of LST values. This may be attributed to MODIS's wider swath width, differences in acquisition timing, and lower spatial sensitivity. In pair 6, the difference between the two sensors becomes more pronounced: Sentinel-3 shows a negative and notable Δ LST (-0.2 K), while MODIS remains nearly neutral (0.001 K), yet still exhibits higher variability. This pattern suggests that Sentinel-3 records temperature changes more uniformly and within a narrower range, whereas MODIS captures broader temperature variability across the scene.

In pair 7, Δ LST is negative for Sentinel-3 and positive for MODIS, with nearly identical standard deviations for both sensors. This behavior reflects the sensitivity of each sensor to environmental conditions and crop growth stages. In the subsequent pairs, both sensors exhibit negative Δ LST values; however, Sentinel-3 records higher magnitudes and greater variability. These findings suggest that Sentinel-3 outperforms MODIS in capturing fine-scale spatial variations and heterogeneous temperature distributions, whereas MODIS represents broader but less spatially detailed fluctuations.

Overall, the combined analysis of Δ LST behavior and sensor performance reveals that the accuracy of LST downscaling is influenced by sensor resolution, image quality, acquisition timing, phenological stage, and diurnal temperature variations. Sentinel-3 demonstrates a stronger ability to capture fine spatial details and daily temperature dynamics, while MODIS, with its broader temporal coverage and higher stability along the scan center, enhances cross-sensor model transferability. These results highlight that integrating multi-sensor data, while accounting for both environmental and temporal conditions, can significantly improve the accuracy of LST downscaling and its applicability in agricultural and climate-related studies.

4.4 Assessing the Cross-Sensor Transferability of MODIS and Sentinel-3 for Downscaled LST Using an XGBoost STF Model

Table 2 and Fig. 10 present the results of downscaled LST using MODIS, Sentinel-3, and their hybrid configurations. The XGBoost model was evaluated under four scenarios: (1) MODIS→Sentinel-3 (trained on MODIS and tested on Sentinel-3), (2) Sentinel-3→MODIS (trained on Sentinel-3 and tested on MODIS), (3) MODIS (trained and tested on MODIS), and (4) Sentinel-3 (trained and tested on Sentinel-3). Several statistical indices were calculated to assess and compare the performance of each configuration.

The results indicate that in 2021 and 2022, the Sentinel-3 configuration achieved the best performance (RMSE = 1.27 and 2.02 K), followed by MODIS (RMSE = 2.00 and 2.29 K), while the hybrid configurations (MODIS→Sentinel-3 and Sentinel-3→MODIS) yielded slightly higher RMSEs (2.24–2.49 K). In 2023, the MODIS→Sentinel-3 configuration achieved the highest accuracy (RMSE = 1.55 K), followed by Sentinel-3 and MODIS (RMSE = 1.72 and 1.76 K, respectively). In 2024, the Sentinel-3 configuration performed slightly worse (RMSE = 1.91 K), although structural indices (CC and SSIM) still indicated better preservation of spatial patterns and fine-scale details.

Overall, the hybrid configurations, particularly MODIS→Sentinel-3, provided superior accuracy compared with the reverse case (Sentinel-3→MODIS). These findings highlight the importance of selecting appropriate training data and demonstrate how the choice of reference sensor can influence model transferability and predictive performance in machine learning-based STF applications.

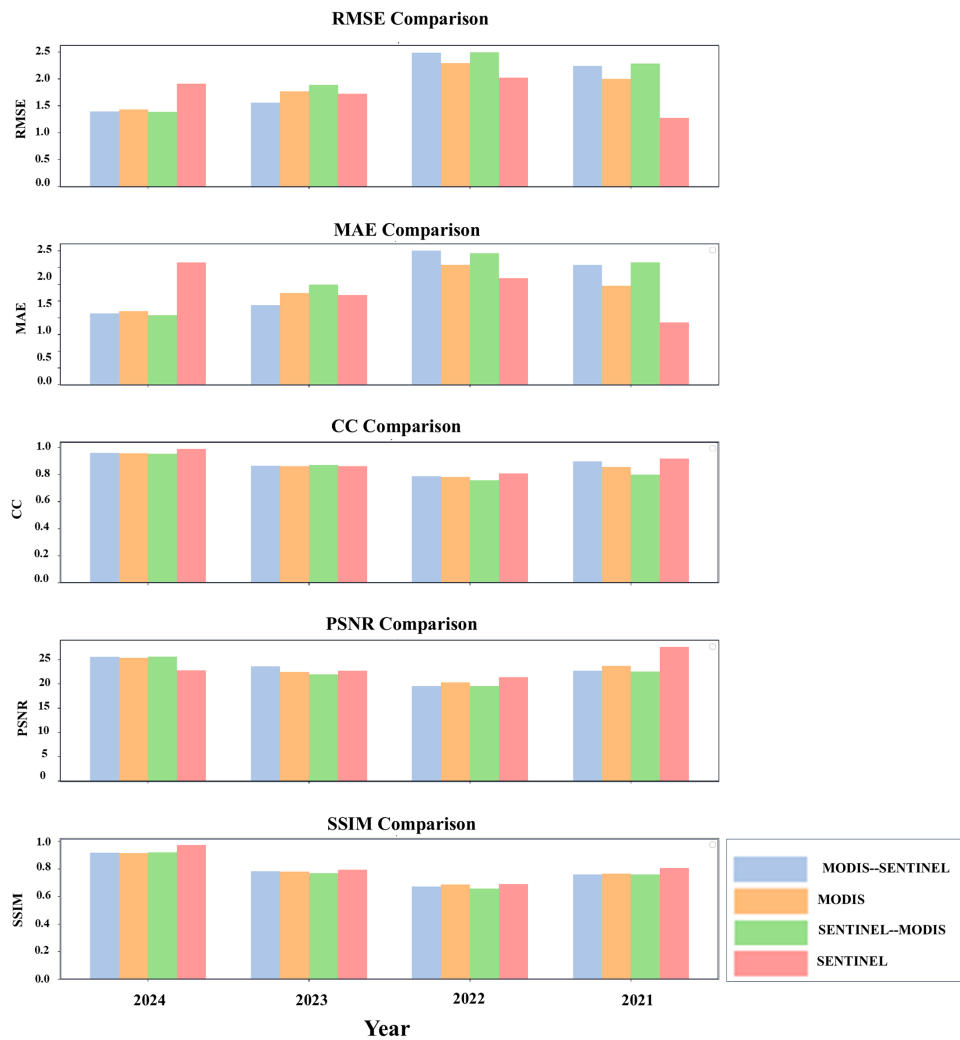


Figure 10: Comparison of high-spatiotemporal-resolution LST predictions using different and hybrid sensor configurations with XGBoost.

5 Discussion

5.1 Comparison of STF Models: ESTARFM vs. XGBoost

Results of the analysis indicate that the machine learning-based XGBoost model generally outperforms the classical ESTARFM model in downscaling LST for both MODIS and Sentinel-3 sensors, particularly under conditions with varying temporal temperature ranges, phenological differences in crops, and spatial heterogeneity. This superiority is mainly attributed to the ability of XGBoost to capture complex and nonlinear relationships among multi-sensor data and spatiotemporal temperature variations. The gradient boosting tree structure and the algorithm's effective feature selection capability enable it to model subtle patterns and intricate variations in LST [30,31].

In line with this advantage, the inherent strengths of boosting algorithms, especially XGBoost, make them highly suitable for Spatio-temporal data fusion, as they effectively capture nonlinear interactions among temperature, land surface features, and temporal variables, while remaining robust to multicollinearity and irregular temporal sampling [26].

Nevertheless, both ESTARFM and XGBoost implicitly assume that coarse-resolution LST fields can be guided by high-resolution Landsat references to reproduce fine-scale thermal structures. The scale mismatch between 1 km MODIS/SLSTR and 100 m Landsat data, especially in irrigated and heterogeneous agricultural landscapes, can substantially affect fusion performance. ESTARFM, which relies on locally linear temporal changes and preserves spatial similarity within a moving window, may violate these assumptions at field boundaries and mixed pixels, leading to systematic underestimation of LST. In contrast, XGBoost, through nonlinear modeling, can partially compensate for such violations; however, it remains unclear whether this correction reflects the learning of true physical relationships between temperature and surface features or merely a statistical adjustment of spatial biases.

Previous studies have confirmed the superior performance of XGBoost in estimating land surface temperature (LST), and recent applications of this algorithm in thermal modeling and energy-related processes such as the study by Kumar et al. (2024) demonstrate that boosting-based approaches offer high stability and robustness in capturing surface temperature variations and can effectively identify complex spatiotemporal relationships [26].

Moreover, Li et al. (2024a) employed XGBoost to generate global instantaneous and daily mean LST products and demonstrated a significant improvement in accuracy (RMSE = 2.787 K for instantaneous and 2.175 K for daily values). In another study conducted in Shiraz, Iran, several algorithms were evaluated for analyzing LST data across different land-cover types (urban, bare soil, and vegetation), and XGBoost provided the most accurate predictions [32]. Moreover, this algorithm exhibits high spatial adaptability and low prediction error when dealing with complex environmental factors [30,32].

The advantage of XGBoost is particularly evident in areas characterized by substantial spatial and diurnal temperature fluctuations [33]. In the present study, the XGBoost model effectively captured the complex interactions between land surface characteristics and multisource datasets. In contrast, the traditional ESTARFM model, due to its methodological limitations and lower temporal resolution, exhibited reduced accuracy in regions with pronounced spatiotemporal variability, particularly in heterogeneous agricultural areas, although it performed reasonably well under relatively stable conditions. Similarly, Filgueiras et al. (2020) demonstrated the superiority of learning-based models such as GBM over linear approaches in spatiotemporal data fusion, especially under heterogeneous vegetation conditions.

5.2 The Performance of MODIS and Sentinel-3 in the STF of LST

The performance of the two sensors varied considerably across the study years. During the 2021–2023 period, Sentinel-3 consistently outperformed MODIS in downscaling LST. Although this superiority was relatively modest for the ESTARFM model, the differences were much more pronounced for the XGBoost model, where indices such as RMSE, MAE, and CC consistently showed better performance for Sentinel-3. These results are in line with the findings of Qi et al. (2023), who also reported slightly lower accuracy in LST retrievals from MODIS compared to SLSTR data [34].

This discrepancy mainly arises from differences in acquisition time and sensor characteristics between the two satellites. The SLSTR sensor onboard Sentinel-3 typically overpasses about one hour earlier than MODIS, which may cause spatial discrepancies in LST distribution, particularly during periods of strong diurnal temperature variation. Moreover, since Landsat data were used as reference imagery in the STF process, datasets acquired closer in time to Landsat typically yield more accurate LST predictions. This temporal discrepancy may vary over time, as the Terra satellite has gradually drifted from its designated orbit following its last inclination-adjustment maneuver in March 2020. By September 2022, its equatorial crossing time had shifted to approximately 10:15 a.m., and it is projected to reach around 9:00 a.m. by 2026. This drift is primarily attributed to orbital decay and the gradual degradation of the satellite's onboard systems [16].

From a technical perspective, certain instrumental limitations should also be taken into account. The thermal infrared bands (S7–S9) of SLSTR experience gradual saturation above approximately 305 K and become fully saturated around 318 K [35], which may produce invalid values in extremely hot regions. However, in typical agricultural environments with moderate surface temperatures, Sentinel-3 provides more accurate LST retrievals owing to its higher radiometric quality and improved calibration. Conversely, the MODIS sensor, with its higher saturation threshold and broader dynamic range, performs more stably in very hot areas. Still, its coarser spatial resolution can obscure fine-scale thermal variations within small agricultural plots. Therefore, the choice of sensor for agricultural studies should consider both the temperature range and the scale of field parcels.

The superior performance of Sentinel-3 can be primarily attributed to its higher spatial resolution and radiometric quality, which enable more precise identification of spatial temperature gradients and diurnal fluctuations. Additionally, angular effects in thermal infrared measurements are a significant issue for polar-orbiting sensors. The observed temperature difference between nadir and off-nadir viewing angles over heterogeneous terrain can reach up to 10 K [36]. The SLSTR instrument mitigates this angular effect through its dual-view conical scanning system and maintains stable in-flight calibration, ensuring accurate thermal and radiometric control. This feature, absent in MODIS, contributes to the higher accuracy of Sentinel-3 in LST prediction [37].

Previous studies have also confirmed these differences. For instance, Shrestha et al. [38] examined the stability and calibration consistency of MODIS thermal bands 31 and 32 and SLSTR bands S8 and S9. They reported that the temperature discrepancy between MODIS band 31 and SLSTR S8 could reach up to 2 K in low-temperature ranges (<200 K), whereas no significant difference was observed between MODIS band 32 and SLSTR S9. Since the temperature range in the present study (290–325 K) falls within a moderate domain, such discrepancies are less pronounced here; however, studies covering broader or different temperature ranges may exhibit distinct results. Furthermore, comparisons of sensor performance across different applications indicate that outcomes are highly dependent on the study region. For example, Varestefanica et al. [39] demonstrated that Terra MODIS imagery achieved higher accuracy than Sentinel-3 in predicting air temperature, with a correlation coefficient of 0.95 and RMSE of 0.51 for MODIS, compared to 0.78 and 0.93 for Sentinel-3, respectively. In another study, Zargari et al. [40], focusing on Tehran and its surroundings, found that MODIS data were more suitable for analyzing nocturnal urban heat island effects. In contrast, Su et al. [41] reported that Sentinel-3 SLSTR data provided highly accurate hourly LST predictions at 30 m spatial resolution, with RMSE values ranging from 0.95 to 1.25 K. This higher precision compared with other satellites, including MODIS, highlights the superior capability of Sentinel-3 in capturing spatial and temporal LST variations across both urban and agricultural areas. In the field of vegetation monitoring, Li et al. [42] showed that Sentinel-3 OLCI achieved higher accuracy than MODIS in estimating vegetation indices . . . such as the Leaf Area Index (LAI), Fractional Vegetation Cover (FVC), and Above-Ground Biomass (AGB) across the eastern Eurasian steppe, with mean RMSE values being 4%–10% lower for Sentinel-3. These findings further confirm the spectral and practical advantages of Sentinel-3 for vegetation studies.

On the other hand, as reported by Colditz et al. [43], data quality represents a critical issue in long-term time series analyses. The MODIS-derived LST data are unavailable for pixels covered by clouds [44]. Moreover, Wan [44] noted that some pixels contaminated by thin clouds remain undetected in the MODIS LST products because the cloud detection algorithms are unable to identify and remove them effectively. Consequently, the accuracy of all applications relying on MODIS LST data can be adversely affected. Similarly, Ackerman et al. [45] reported that approximately 15% of pixels in MODIS LST products are cloud-contaminated. In another study, Williamson et al. [46] evaluated cloud contamination in daily MODIS

Terra LST data under clear-sky conditions over southwestern Yukon using ground-based meteorological observations and found that 13%–17% of MODIS LST data corresponded to undetected clouds. These findings clearly indicate that cloud contamination can significantly influence the outcomes of MODIS LST-based analyses. Furthermore, similar limitations are also present in Sentinel-3 data, as its thermal measurements are acquired from spectral bands within comparable wavelength ranges, making them equally sensitive to cloud cover and atmospheric effects.

5.3 Spatial Analysis of Downscaled LST Results

The spatial analysis of the downscaled LST revealed that the ESTARFM model tended to underestimate LST values across most regions of the study area. This pattern was consistently observed for both the MODIS and Sentinel-3 datasets. In contrast, the XGBoost model effectively reduced this underestimation, producing more accurate results overall. This finding is consistent with previous studies (e.g., [31]), which have shown that machine learning models generally possess a stronger ability to mitigate systematic errors.

However, in 2024, the results exhibited a tendency toward overestimation. This could partly be attributed to the relatively low temporal variability and overall uniformity of temperature patterns during that period. Under such conditions, machine learning models often tend to over-smooth predictions and converge toward the mean, typically resulting in higher overall LST estimates. This behavior was particularly evident in datasets corresponding to dates with low temporal variation.

The observed differences can be interpreted from both technical and data-related perspectives. The XGBoost model, due to its tree-based structure and its ability to model nonlinear relationships between spectral bands and temperature, may occasionally overgeneralize thermal patterns. Consequently, in areas with high surface reflectance (such as dry or sparsely vegetated regions), the model may predict higher LST values than observed. Furthermore, when the training data do not fully capture spatial and temporal variability, the model tends to overestimate mean LST, especially in warmer regions.

In contrast, the ESTARFM model, because of its linear formulation and strong dependence on spatial-temporal weighting between base images, struggles to predict rapid or heterogeneous changes (e.g., at vegetation–bare soil boundaries). As a result, it tends to over-smooth the data, leading to an underestimation of true temperatures.

In fact, the analysis of the mean and standard deviation of Δ LST revealed that part of the discrepancies arises from inherent sensor biases, such as radiometric calibration, satellite overpass time, and viewing angle, while another part stems from the data fusion methods, including window averaging in ESTARFM and regression smoothing in XGBoost. Specifically, ESTARFM tends to propagate base-image bias almost linearly, whereas XGBoost generally reduces the magnitude of bias, although it may occasionally lead to slight overestimation under conditions of low temporal variability.

Previous studies have demonstrated the role of machine learning methods, including boosting-based approaches, in mitigating these limitations. For instance, Zhang et al. (2022) employed a Light Gradient Boosting Machine (LGBM) to learn and correct biases between MODIS and Landsat imagery, significantly improving the accuracy and robustness of spatiotemporal fusion [47]. These findings highlight that machine learning not only reduces sensor bias but also enhances the consistency and quality of input data, ultimately improving the reliability and precision of high-resolution LST products.

The analysis of the mean and standard deviation of Δ LST between MODIS and Sentinel-3 pairs (against Landsat references) was conducted to examine the inherent biases of the sensors. The results indicated that the higher Δ LST variability in MODIS, particularly in pairs 1 and 3, was likely caused by differences in overpass times, sensor viewing angles, and stronger atmospheric scattering effects. In contrast, Sentinel-3

exhibited more stable and homogeneous patterns, especially in pair 6, suggesting better radiometric stability and improved atmospheric correction performance.

The sign reversal of ΔLST in pair 7 (negative for Sentinel-3 and positive for MODIS) was likely due to differences in the thermal responses of vegetation at different growth stages. During mid-growth periods, surface reflectance and thermal emission vary with solar angle, surface moisture, and canopy shading structure. This phenomenon suggests that each sensor exhibits distinct spectral sensitivities to biophysical surface characteristics such as vegetation density, evapotranspiration rate, and soil moisture, which may explain the divergent thermal responses observed between MODIS and Sentinel-3 data.

Such dynamics have been documented in the dryland agricultural systems of Australia, where LST is strongly correlated with root-zone soil moisture and vegetation water status, indicating that LST responds to plant water stress [48]. Moreover, thermal drought indices, such as the Temperature Rise Index (TRI), exhibit more rapid increases in LST under dry conditions, highlighting that water availability substantially controls surface temperature dynamics in agricultural landscapes [49]. Integrated analyses of LST and soil moisture have also shown that LST anomalies often reflect soil moisture deficits and vegetation-related stress [50], supporting our interpretation that phenological changes and water availability can drive the observed differences between sensors.

5.4 The Transferability of STF Models between MODIS and Sentinel-3

The results of the transferability experiments indicated that the model transfer from MODIS to Sentinel-3 (MODIS→Sentinel-3) generally performed better than the reverse configuration. The main reason for this is that MODIS provides a smoother and more temporally consistent representation of LST dynamics, enabling XGBoost to learn robust temporal patterns. In contrast, although Sentinel-3 offers greater spatial richness, its lower temporal continuity limits its effectiveness as a training source for coarser-resolution data.

This finding is consistent with previous studies [51,52], which have demonstrated that datasets with higher temporal continuity and broader spatial coverage enable models to learn more stable and generalizable environmental patterns.

One of the key factors contributing to this superior performance is the extensive temporal and spatial coverage of MODIS data. Despite its coarser spatial resolution, MODIS effectively captures the general and temporally consistent patterns of land surface temperature variations. Models trained on MODIS can therefore transfer these stable temporal patterns to Sentinel-3 data with finer spatial detail, leveraging Sentinel-3's higher spatial resolution to enhance prediction accuracy. In other words, the fine-scale spatial structures provided by Sentinel-3 complement the broader and temporally consistent MODIS patterns, allowing MODIS→Sentinel-3 models to benefit from both sources of information [53,54].

Conversely, models trained on Sentinel-3 data (Sentinel-3→MODIS) often exhibited reduced accuracy when transferred to MODIS. This is primarily because Sentinel-3 data capture highly localized and fine spatial patterns, which are not preserved when applied to the coarser MODIS data. As a result, much of the spatial detail learned by the model is lost, leading to a decline in predictive performance since MODIS lacks the spatial granularity necessary to represent these features effectively [52].

6 Conclusion

The results of this study demonstrated that the machine learning-based XGBoost model outperformed the classical ESTARFM approach in the spatiotemporal fusion of multi-sensor data (MODIS and Sentinel-3) for LST estimation. This superiority was particularly evident in areas exhibiting rapid temporal changes, high spatial heterogeneity, and pronounced phenological variations in crops.

The gradient boosting framework and XGBoost's capacity to capture nonlinear and complex relationships among spatial, temporal, and spectral variables resulted in a significant reduction in prediction errors and improvements in accuracy metrics such as RMSE, MAE, and the correlation coefficient. In contrast, the ESTARFM model performed satisfactorily in relatively stable regions but showed lower accuracy under conditions of rapid temporal dynamics and strong spatial heterogeneity, primarily due to its linear assumptions and dependence on spatial-temporal weighting.

A comparison between the two sensors revealed that Sentinel-3, owing to its higher spatial resolution, superior calibration, and enhanced radiometric quality, achieved higher accuracy in LST prediction than MODIS when cloud-free images were available for the target dates. These findings confirm that Sentinel-3 LST data can serve as a reliable alternative in cases where MODIS data are unavailable or degraded. Nevertheless, MODIS remains indispensable for long-term analyses due to its broader temporal coverage and higher thermal stability under elevated temperature conditions.

The cross-sensor transfer analysis also indicated that transferring models from MODIS to Sentinel-3 (MODIS→Sentinel-3) yielded higher accuracy than the reverse configuration (Sentinel-3→MODIS). This superiority likely stems from MODIS's broader temporal sampling, which captures more stable temperature variation patterns that Sentinel-3 subsequently refines through its finer spatial detail. From a temporal perspective, the year 2024 exhibited the highest model performance, attributed to improved synchronization of satellite acquisitions with key crop phenological stages and reduced thermal fluctuations. Conversely, 2023 showed the weakest performance, primarily due to greater phenological variability and suboptimal temporal alignment between satellite observations. In summary, the findings demonstrate that the use of machine learning approaches, particularly XGBoost, can effectively overcome the limitations of traditional spatiotemporal fusion algorithms, enabling more accurate LST prediction at both local and regional scales. This advancement opens new opportunities for the integration of intelligent methods in agricultural monitoring, water resource management, and climate change studies. Furthermore, the results confirm that, given access to high-quality and cloud-free imagery, Sentinel-3 can serve as an effective replacement for MODIS in generating high-resolution LST maps.

Acknowledgement: We sincerely thank the personnel from the Department of Remote Sensing and GIS, Tarbiat Modares University, Tehran, Iran, for their support and contributions to this research.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Ali Shamsoddini: Conceptualization, Methodology, Supervision, Project administration, Writing—Reviewing and editing. Nahid Haghshenas: Conceptualization, Methodology, Software, Data curation, Investigation, Validation, Formal analysis, Visualization, and Writing—Original draft preparation. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The datasets generated and analyzed in the current study are available from the corresponding author on reasonable request.

Ethics Approval: Not applicable.

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

References

1. Mo Y, Xu Y, Chen H, Zhu S. A review of reconstructing remotely sensed land surface temperature under cloudy conditions. *Remote Sens.* 2021;13(14):2838. doi:10.3390/rs13142838.
2. Zhou D, Xiao J, Bonafoni S, Berger C, Deilami K, Zhou Y, et al. Satellite remote sensing of surface urban heat islands: progress, challenges, and perspectives. *Remote Sens.* 2019;11(1):48. doi:10.3390/rs11010048.
3. Zhu X, Song X, Leng P, Li X, Gao L, Guo D, et al. A framework for generating high spatiotemporal resolution land surface temperature in heterogeneous areas. *Remote Sens.* 2021;13(19):3885. doi:10.3390/rs13193885.
4. Hurduc A, Ermida SL, Trigo IF, DaCamara CC. Importance of temporal dimension and rural land cover when computing surface urban Heat Island intensity. *Urban Clim.* 2024;56(1):102013. doi:10.1016/j.uclim.2024.102013.
5. Wang Q, Tang Y, Ge Y, Xie H, Tong X, Atkinson PM. A comprehensive review of spatial-temporal-spectral information reconstruction techniques. *Sci Remote Sens.* 2023;8(10):100102. doi:10.1016/j.srs.2023.100102.
6. Shi W, Guo D, Zhang H. A reliable and adaptive spatiotemporal data fusion method for blending multi-spatiotemporal-resolution satellite images. *Remote Sens Environ.* 2022;268:112770. doi:10.1016/j.rse.2021.112770.
7. Reiners P, Sobrino J, Kuenzer C. Satellite-derived land surface temperature dynamics in the context of global change—a review. *Remote Sens.* 2023;15(7):1857. doi:10.3390/rs15071857.
8. Dhillon MS, Kübert-Flock C, Dahms T, Rummeler T, Arnault J, Steffan-Dewenter I, et al. Evaluation of MODIS, landsat 8 and sentinel-2 data for accurate crop yield predictions: a case study using STARFM NDVI in Bavaria. *Germany Remote Sens.* 2023;15(7):1830. doi:10.3390/rs15071830.
9. Xu Q, Guo Y, Chen W, Ji G, Shi L, Li Y, et al. Comprehensive assessment of Spatiotemporal fusion methods in inland water monitoring. *GISci Remote Sens.* 2024;61(1):2343200. doi:10.1080/15481603.2024.2343200.
10. Tang Y, Wang Q, Atkinson PM. Filling then spatio-temporal fusion for all-sky MODIS land surface temperature generation. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 2023;16:1350–64. doi:10.1109/jstars.2023.3235940.
11. Negahbani S, Momeni M, Moradzadeh M. Improving the spatiotemporal resolution of soil moisture through a synergistic combination of MODIS and LANDSAT8 data. *Water Resour Manag.* 2022;36(6):1813–32. doi:10.1007/s11269-022-03108-1.
12. Haghshenas N, Shamsoddini A. Evaluation of spatio-temporal downscaling algorithms of MODIS data to sentinel-2 data in different land cover classes. *Iran J Remote Sens GIS.* 2023;15(4):100–83.
13. Lin G, Zhang P, Dellomo J, Tan B. Terra orbit drift and its impacts on MODIS geometric performance. *AGU Fall Meet Abstr.* 2024;2024(338):GC51Z-0338.
14. Xiong X, Butler JJ. MODIS and VIIRS calibration history and future outlook. *Remote Sens.* 2020;12(16):2523. doi:10.3390/rs12162523.
15. Srivastava A, Sahoo B, Raghuvanshi NS, Singh R. Evaluation of variable-infiltration capacity model and MODIS-*Terra satellite*-derived grid-scale evapotranspiration estimates in a river basin with tropical monsoon-type climatology. *J Irrig Drain Eng.* 2017;143(8):04017028. doi:10.1061/(asce)ir.1943-4774.0001199.
16. Román MO, Justice C, Paynter I, Boucher PB, Devadiga S, Endsley A, et al. Continuity between NASA MODIS collection 6.1 and VIIRS collection 2 land products. *Remote Sens Environ.* 2024;302(20):113963. doi:10.1016/j.rse.2023.113963.
17. Ghent D, Anand JS, Veal K, Remedios J. The operational and climate land surface temperature products from the sea and land surface temperature radiometers on sentinel-3A and 3B. *Remote Sens.* 2024;16(18):3403. doi:10.3390/rs16183403.
18. Ao Z, Sun Y, Pan X, Xin Q. Deep learning-based spatiotemporal data fusion using a patch-to-pixel mapping strategy and model comparisons. *IEEE Trans Geosci Remote Sens.* 2022;60:1–18. doi:10.1109/tgrs.2022.3154406.
19. Senanayake IP, Yeo IY, Walker JP, Willgoose GR. Estimating catchment scale soil moisture at a high spatial resolution: integrating remote sensing and machine learning. *Sci Total Environ.* 2021;776(1):145924. doi:10.1016/j.scitotenv.2021.145924.
20. Ye N, Walker JP, Wu X, de Jeu R, Gao Y, Jackson TJ, et al. The soil moisture active passive experiments: validation of the SMAP products in Australia. *IEEE Trans Geosci Remote Sens.* 2021;59(4):2922–39. doi:10.1109/tgrs.2020.3007371.

21. Ye N, Walker JP, Bindlish R, Chaubell J, Das NN, Gevaert AI, et al. Evaluation of SMAP downscaled brightness temperature using SMAPEX-4/5 airborne observations. *Remote Sens Environ.* 2019;221:363–72. doi:10.1016/j.rse.2018.11.033.
22. Merlin O, Walker JP, Kalma JD, Kim EJ, Hacker J, Panciera R, et al. The NAFE'06 data set: towards soil moisture retrieval at intermediate resolution. *Adv Water Resour.* 2008;31(11):1444–55. doi:10.1016/j.advwatres.2008.01.018.
23. Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, et al. The ERA5 global reanalysis. *Quart J Royal Meteorol Soc.* 2020;146(730):1999–2049. doi:10.1002/qj.3803.
24. Zhu X, Chen J, Gao F, Chen X, Masek JG. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens Environ.* 2010;114(11):2610–23. doi:10.1016/j.rse.2010.05.032.
25. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2016 Aug 13–17; San Francisco, CA, USA. doi:10.1145/2939672.2939785.
26. Kumar M, Agrawal Y, Adamala S, Pushpanjali, Subbarao AVM, Singh VK, et al. Generalization ability of bagging and boosting type deep learning models in evapotranspiration estimation. *Water.* 2024;16(16):2233. doi:10.3390/w16162233.
27. Duan SB, Lian Y, Zhao E, Chen H, Han W, Wu Z. A novel approach to all-weather LST estimation using XGBoost model and multisource data. *IEEE Trans Geosci Remote Sens.* 2023;61:1–14. doi:10.1109/tgrs.2023.3324481.
28. Guo S, Li M, Li Y, Chen J, Zhang HK, Sun L, et al. The improved U-STFM: a deep learning-based nonlinear spatial-temporal fusion model for land surface temperature downscaling. *Remote Sens.* 2024;16(2):322.
29. Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM. In: *Proceedings of the 2010 20th International Conference on Pattern Recognition*; 2010 Aug 23–26; Istanbul, Turkey. doi:10.1109/icpr.2010.579.
30. Song S, Shi J, Fan D, Cui L, Yang H. Development of downscaling technology for land surface temperature: a case study of Shanghai, China. *Urban Clim.* 2025;61(9):102412. doi:10.1016/j.uclim.2025.102412.
31. Liu F, Wang X, Sun F, Wang H, Wu L, Zhang X, et al. Correction of overestimation in observed land surface temperatures based on machine learning models. *J Clim.* 2022;35(16):5359–77. doi:10.1175/jcli-d-21-01447.1.
32. Tanoori G, Soltani A, Modiri A. Machine learning for urban heat island (UHI) analysis: predicting land surface temperature (LST) in urban environments. *Urban Clim.* 2024;55(11):101962. doi:10.1016/j.uclim.2024.101962.
33. Bushenkova A, Soares PMM, Johannsen F, Lima DCA. Towards an improved representation of the urban heat island effect: a multi-scale application of XGBoost for Madrid. *Urban Clim.* 2024;55(9):101982. doi:10.1016/j.uclim.2024.101982.
34. Qi Y, Zhong L, Ma Y, Fu Y, Wang X, Li P. Estimation of land surface temperature over the Tibetan Plateau based on sentinel-3 SLSTR data. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 2023;16:4180–94. doi:10.1109/jstars.2023.3268326.
35. Coppo P, Ricciarelli B, Brandani F, Delderfield J, Ferlet M, Mutlow C, et al. SLSTR: a high accuracy dual scan temperature radiometer for sea and land surface monitoring from space. *J Mod Opt.* 2010;57(18):1815–30. doi:10.1080/09500340.2010.503010.
36. Lagouarde JP, Hénon A, Kurz B, Moreau P, Irvine M, Voogt J, et al. Modelling daytime thermal infrared directional anisotropy over Toulouse city centre. *Remote Sens Environ.* 2010;114(1):87–105. doi:10.1016/j.rse.2009.08.012.
37. Smith D, Hunt SE, Etxaluz M, Peters D, Nightingale T, Mittaz J, et al. Traceability of the sentinel-3 SLSTR level-1 infrared radiometric processing. *Remote Sens.* 2021;13(3):374. doi:10.3390/rs13030374.
38. Shrestha A, Angal A, Xiong X. Evaluation of MODIS and Sentinel-3 SLSTR thermal emissive bands calibration consistency using Dome C. In: *Proceedings of the Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIV*; 2018 Apr 15–19; Orlando, FL, USA. doi:10.1117/12.2303987.
39. Varestefanica JA, Yuliara IM, Yuda IWA, Suarabawa KN, Sumadiyah M, Sandi IN, et al. Comparative analysis of sentinel-3 and Terra MODIS satellite images for air temperature observation in Denpasar area. *Asian J Res Rev Phys.* 2023;7(3):15–24. doi:10.9734/ajr2p/2023/v7i3141.
40. Zargari M, Mofidi A, Entezari A, Baaghideh M. Climatic comparison of surface urban heat island using satellite remote sensing in Tehran and suburbs. *Sci Rep.* 2024;14(1):643. doi:10.1038/s41598-023-50757-2.

41. Su Q, Yao Y, Chen C, Chen B. Generating a 30 m hourly land surface temperatures based on spatial fusion model and machine learning algorithm. *Sensors*. 2024;24(23):7424. doi:10.3390/s24237424.
42. Li B, Liang S, Ma H, Dong G, Liu X, He T, et al. Generation of global 1 km all-weather instantaneous and daily mean land surface temperatures from MODIS data. *Earth Syst Sci Data*. 2024;16(8):3795–819. doi:10.5194/essd-16-3795-2024.
43. Colditz RR, Conrad C, Wehrmann T, Schmidt M, Dech S. TiSeG: a flexible software tool for time-series generation of MODIS data utilizing the quality assessment science data set. *IEEE Trans Geosci Remote Sens*. 2008;46(10):3296–308. doi:10.1109/tgrs.2008.921412.
44. Wan Z. New refinements and validation of the MODIS land-surface temperature/emissivity products. *Remote Sens Environ*. 2008;112(1):59–74. doi:10.1016/j.rse.2006.06.026.
45. Ackerman SA, Holz RE, Frey R, Eloranta EW, Maddux BC, McGill M. Cloud detection with MODIS. Part II: validation. *J Atmos Ocean Technol*. 2008;25(7):1073–86. doi:10.1175/2007jtecha1053.1.
46. Williamson SN, Hik DS, Gamon JA, Kavanaugh JL, Koh S. Evaluating cloud contamination in clear-sky MODIS *Terra* daytime land surface temperatures using ground-based meteorology station observations. *J Clim*. 2013;26(5):1551–60. doi:10.1175/jcli-d-12-00250.1.
47. Zhang H, Huang F, Hong X, Wang P. A sensor bias correction method for reducing the uncertainty in the spatiotemporal fusion of remote sensing images. *Remote Sens*. 2022;14(14):3274. doi:10.3390/rs14143274.
48. Holzman M, Srivastava A, Rivas R, Huete A. Evaluating the relationship between vegetation status and soil moisture in semi-arid woodlands, central Australia, using daily thermal, vegetation index, and reflectance data. *Remote Sens*. 2025;17(4):635. doi:10.3390/rs17040635.
49. Hu T, Renzullo LJ, van Dijk AIJM, He J, Tian S, Xu Z, et al. Monitoring agricultural drought in Australia using MTSAT-2 land surface temperature retrievals. *Remote Sens Environ*. 2020;236:111419. doi:10.1016/j.rse.2019.111419.
50. Hu T, van Dijk AIJM, Renzullo LJ, Xu Z, He J, Tian S, et al. On agricultural drought monitoring in Australia using Himawari-8 geostationary thermal infrared observations. *Int J Appl Earth Obs Geoinf*. 2020;91:102153. doi:10.1016/j.jag.2020.102153.
51. da Silva Junior JA, Pacheco AP, Ruiz-Armenteros AM, Henriques RFF. Evaluation of the ability of SLSTR (sentinel-3B) and MODIS (*Terra*) images to detect burned areas using spatial-temporal attributes and SVM classification. *Forests*. 2023;14(1):32. doi:10.3390/f14010032.
52. Xu W, Wooster MJ. Sentinel-3 SLSTR active fire (AF) detection and FRP daytime product-Algorithm description and global intercomparison to MODIS, VIIRS and landsat AF data. *Sci Remote Sens*. 2023;7(1):100087. doi:10.1016/j.srs.2023.100087.
53. Zhang T, Zhou Y, Zhu Z, Li X, Asrar GR. A global seamless 1 km resolution daily land surface temperature dataset (2003–2020). *Earth Syst Sci Data*. 2022;14(2):651–64. doi:10.5194/essd-14-651-2022.
54. Petracca I, De Santis D, Picchiani M, Corradini S, Guerrieri L, Prata F, et al. Volcanic cloud detection using Sentinel-3 satellite data by means of neural networks: the Raikoke 2019 eruption test case. *Atmos Meas Tech*. 2022;15(24):7195–210. doi:10.5194/amt-15-7195-2022.