SANGAM International Journal of Multidisciplinary Research

Vol. 03, Issue 01, September (2025) **E-ISSN: 2584-086X**

https://doi.org/10.5281/zenodo.17596506

CLOUD-RESILIENT OPTICAL REMOTE SENSING TECHNIQUES: A REVIEW OF GAP-FILLING AND TEMPORAL FUSION METHODS

Monika Jain[1*], Dr. Lokesh Kumar Tripathi[2], Dr. Abhilasha Dangi[3]

[1,2]Dept. of Geoinformatics, Sangam University, Bhilwara (Raj.)

[3]Dept. of Computer Science, Sangam University, Bhilwara (Raj.)

Corresponding Author: *jain.mona412@gmail.com, Tel.:8696492365*

Available online at: www.sijmr.org

Received: 06-08-2025, Accepted: 26-09-2025, Online: 30-09-2025

Abstract— Optical remote sensing has been a cornerstone in monitoring Earth's surface processes due to its ability to capture multispectral information with high spatial resolution. However, persistent cloud cover remains a critical limitation, particularly in tropical, monsoonal, and mountainous regions, leading to data gaps, reduced temporal consistency, and compromised analysis quality. This review presents a comprehensive synthesis of cloud-resilient techniques developed to mitigate the impact of cloud contamination in optical satellite imagery. The study classifies and evaluates a range of gap-filling and temporal fusion methods, including spatial and temporal interpolation, multi-date compositing, data fusion frameworks, and emerging deep learning approaches. Particular emphasis is given to models such as STARFM, ESTARFM, and spatiotemporal CNN-LSTM hybrids that enable integration of high-resolution and high-frequency datasets. The role of auxiliary datasets, such as synthetic aperture radar (SAR), historical vegetation indices, and meteorological parameters, is also discussed in enhancing cloud resilience. Additionally, the review explores operational tools and platforms including Google Earth Engine, Fmask, and machine learning libraries that have advanced the automation of these processes. While progress has been significant, challenges persist in terms of model generalization, validation under extreme cloud conditions, and balancing spatial-temporal trade-offs. The paper concludes by identifying key research gaps and proposing future directions for the development of robust, scalable, and accessible cloud-resilient remote sensing frameworks. These advancements are crucial for ensuring continuous environmental monitoring and informed decision-making, particularly in regions where cloud-free observations remain sparse.

Keywords*— Cloud resilience, optical remote sensing, gap-filling, temporal fusion, data reconstruction, satellite imagery, SAR integration,STARFM,deeplearning,GoogleEarthEngine.*

I. INTRODUCTION

Optical remote sensing has transformed the way we observe, map, and understand Earth's surface. The spectral richness of optical data enables precise classification of land cover, vegetation dynamics, water bodies, and urban expansion. However, its usability is frequently compromised by cloud interference—a persistent issue that affects nearly 30–60% of satellite images globally. This problem is particularly severe in equatorial and tropical regions, where dense cloud cover can persist for months, leading to significant temporal gaps in time-series analyses.

To address this limitation, researchers and operational agencies have developed various cloud-resilient methodologies that aim to restore or reconstruct cloud-affected pixels using both statistical and machine learning techniques. This paper systematically reviews such methods, categorized broadly into gap-filling, compositing, temporal fusion, and AI-based reconstruction techniques. The review further elaborates on how synthetic aperture radar (SAR) and multi-sensor fusion enhance resilience and highlights the platforms enabling scalable processing of these methods. This review aims to fulfill the following key objectives: *a)* To synthesize the existing literature on cloud contamination in optical remote sensing and its implications on data availability, accuracy, and analysis outcomes. *b)* To classify and evaluate current methodologies for handling cloud-induced data gaps, including traditional interpolation, multi-date compositing, and advanced spatiotemporal fusion approaches. *c)* To explore the emerging role of deep learning models such as CNNs, LSTMs, and generative architectures in restoring cloud-contaminated images and generating temporally consistent datasets. *d)* To highlight the integration of auxiliary datasets (e.g., SAR, meteorological records, vegetation indices) in enhancing cloud-resilient analysis. To review the operational platforms and tools (e.g., Google Earth Engine, Fmask, open-source machine learning libraries) that support scalable implementation of these techniques. *e)* To identify current challenges in model generalization, validation, and scalability, and propose future directions for research and operational deployment.

Rationale: Why Cloud Resilience is Essential? Cloud contamination remains one of the most persistent limitations of optical remote sensing, with global studies indicating that up to 60% of satellite acquisitions are affected in humid and tropical regions. This limitation is particularly critical for three application domains, a) *Agriculture:* Precision agriculture and food security assessments require dense temporal observations for monitoring crop phenology, stress detection, and yield forecasting. Seasonal crops such as rice or maize often coincide with monsoon periods, when heavy cloud cover obscures the growing cycle. In the absence of reliable cloud-resilient methods, early-warning systems for food shortages or pest outbreaks become severely compromised. b) *Climate Change Monitoring:* Detecting gradual but significant changes—such as glacier retreat, deforestation, or wetland loss—depends on long-term, temporally consistent data. Cloud-induced discontinuities reduce the statistical confidence of trend analyses and weaken the capacity to separate natural variability from climate-driven shifts. c) *Disaster Response:* Rapid disaster assessment after floods, cyclones, or landslides requires near real-time imagery. Yet, these events are often accompanied by dense cloud cover, rendering optical sensors nearly blind during the most critical periods. Cloud-resilient frameworks, especially those integrating radar or hybrid AI models, therefore become indispensable for timely disaster mapping and recovery planning.Together, these examples underline that cloud resilience is not a technical luxury but a fundamental requirement for evidence-based decision-making in sectors tied to food security, climate policy, and humanitarian relief.

Gap in Existing Reviews: Despite the growing body of literature on cloud-resilient remote sensing, existing reviews have several limitations that reduce their utility for both researchers and practitioners. These gaps can be grouped into interrelated areas. *a) Predominantly Descriptive Rather than Analytical* Most published reviews provide catalogues of methods listing interpolation, compositing, or fusion techniques without critically comparing their relative merits. As a result, readers gain awareness of available tools but not an understanding of which approach is most appropriate under specific ecological conditions, data constraints, or computational capacities. Without analytical synthesis, such reviews fail to move beyond a literature inventory. *b) Lack of Standardized Comparative Frameworks* While individual studies often evaluate methods against specific datasets, these comparisons are rarely translated into a structured framework that weighs accuracy, computational demand, robustness to heterogeneity, or transferability. For example, STARFM may be effective in homogeneous savanna systems but falters in fragmented agricultural mosaics, whereas CNN–LSTM hybrids perform well in heterogeneous settings but at much higher computational cost. Current reviews do not consolidate these trade-offs into a coherent decision-support framework, leaving practitioners without clear guidance. *c) Inadequate Incorporation of Recent Advances (2021–2025)* The literature on spatio-temporal fusion has advanced rapidly with the rise of deep learning, transformer-based models, and SAR–optical integration. Yet, reviews continue to rely heavily on studies from the 2000s and early 2010s, with minimal discussion of cutting-edge contributions. Techniques such as object-based spatiotemporal unmixing (OBSUM), transformer architectures for time-series reconstruction, or explainable AI frameworks are scarcely mentioned, even though they represent the current frontier of research. This chronological gap risks presenting an outdated picture of the field. *d) Limited Case-Based Synthesis Across Ecological Contexts* Another critical shortcoming is the absence of cross-regional synthesis. Case studies demonstrating successful applications in agriculture, snow cover monitoring, or deforestation detection exist, but reviews often present them in isolation rather than integrating them into a comparative perspective. For instance, while snow monitoring in mountainous terrain poses different challenges than deforestation detection in tropical forests, the lessons from both contexts could inform the design of more generalizable models. Current reviews miss the opportunity to highlight such transferable insights. *e) Neglect of Operational Relevance* Although academic advances are well documented, the translation of these methods into operational workflows remains underexplored. Few reviews address the availability of algorithms in widely used platforms such as Google Earth Engine, or the extent to which open-source implementations exist for practitioners in resource-constrained settings. Without bridging the gap between research and practice, the broader adoption of cloud-resilient methods remains limited.

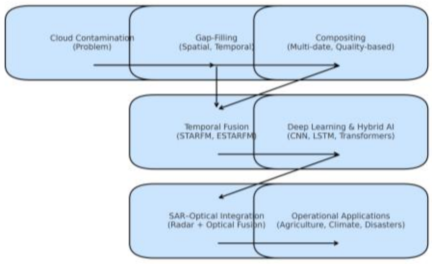

Fig. 1 Conceptual Workflow of Cloud Resilient Remote Sensing Techniques.

II. IMPACT OF CLOUD CONTAMINATION IN OPTICAL REMOTE SENSING

Clouds present one of the single largest obstacles in optical remote sensing. Beyond obstructing direct observation of the Earth's surface, they introduce several cascading problems for remote sensing analyses: data loss, spectral distortion, and temporal inconsistencies. Below we examine these issues, with recent evidence (2022–2024), followed by their implications for monitoring and modelling.

2.1 Data Loss

The most direct impact of cloud cover is the loss of usable pixels. Depending on geography and season, clouds may obscure anywhere from 30% to over 70% of all optical observations. *a) Global trends:* An analysis of the Harmonized Landsat–Sentinel-2 (HLS) dataset for 2022 reported that the global mean of cloud-free days was 69 per year, with a median of 64 (USGS, 2023). In tropical regions, where persistent clouds dominate, cloud-free observations dropped to once every 2.2 days on average, compared to every 1.6 days globally. *b) Regional evidence:* In the Mont Blanc massif (2016–2024), Sentinel-2 yielded more than 100 theoretical acquisitions per year, but only 10–20 images were suitable for high-quality analysis due to cloud and snow contamination (Paul et al., 2024).*Implication:* These gaps undermine the ability to monitor dynamic processes such as crop cycles, flooding, or glacier retreat, where missing a short observation window can compromise entire analyses.

2.2 Spectral Distortion

Clouds and their shadows introduce spectral contamination even in partially usable imagery. Thin clouds increase atmospheric scattering, reducing reflectance in visible and near-infrared bands, while shadows create artificially darkened pixels. *a) Thin clouds and haze*: Sub-visual clouds can bias vegetation indices by lowering NDVI or EVI values, causing misinterpretation of vegetation health.

b) Cloud edges: Mixed pixels at the boundaries between cloudy and clear areas often create spectral heterogeneity that is difficult to correct. *c) Recent findings:* Zhou et al. (2022) showed that dense and thin clouds significantly degrade reconstruction accuracy, with Peak Signal-to-Noise Ratio (PSNR) reductions of more than 15% in Landsat/Sentinel experiments. GAN-based cloud removal methods improved consistency but required high computational costs.*Implication:* Spectral distortion leads to classification errors (e.g., cropland misclassified as bare soil) and biases biophysical parameter retrieval (e.g., LAI, albedo, evapotranspiration).

2.3 Temporal Inconsistencies

Time-series analyses rely on consistent, cloud-free observations. Persistent cloud cover interrupts this continuity, leading to temporal gaps and uneven sampling.

a) Irregular observation intervals: Sentinel-2 theoretically offers a 5-day revisit, but under monsoon or equatorial cloud conditions, gaps of 3–4 weeks are common. This irregularity breaks phenological curves, especially for crops with short growth phases. *b) Missed critical events:* Cloud cover during a flood peak, snowmelt period, or deforestation event can result in complete absence of usable records. For example, in Southeast Asia, critical rice phenology stages were missed in more than 40% of cases due to seasonal clouds (Nguyen et al., 2023). *c) Case evidence:* In HLS V2.0 data for 2022, monthly clear-sky observations in tropical Africa averaged fewer than six per pixel, insufficient for sub-monthly crop monitoring (USGS, 2023).

Implication: These temporal inconsistencies limit the ability to detect abrupt changes, weaken trend analysis, and force reliance on interpolation or modeling, which increases uncertainty.

2.3 Combined Impacts

The cumulative effect of data loss, spectral distortion, and temporal inconsistency is profound: *a) Reduced reliability of indices:* Time-series NDVI or EVI curves appear noisy and fragmented, reducing confidence in detecting subtle trends. *b) Poor change detection:* Rapid events such as flooding, forest fire scars, or harvesting may remain invisible if clouds persist. *c) Uncertainty in climate assessments:* Long-term climate monitoring requires consistent datasets, but cloud-induced discontinuities lower statistical confidence in trend attribution. *d) Operational challenges:* Cloud-prone regions often lack sufficient clear images for training or validating machine learning models, restricting transferability of algorithms.

Synthesis: Cloud contamination is not only a technical nuisance but a systemic barrier that undermines the operational use of optical remote sensing. Recent analyses (2022–2024) reaffirm that despite denser constellations of sensors, persistent cloud frequency continues to reduce data availability and reliability. This underlines the urgent necessity of cloud-resilient frameworks, including gap-filling, compositing, fusion, and SAR–optical integration, for both research and real-world decision support.

III. Overview of Cloud-Resilient Approaches

Cloud-resilient techniques in optical remote sensing are designed to restore or circumvent data loss resulting from cloud and cloud-shadow contamination. Such methods aim to reconstruct missing pixels, maintain temporal continuity, and preserve the thematic and spectral integrity of imagery for robust environmental analysis. The term encompasses spatial, temporal, and hybrid strategies: from classical interpolation within a single image to sophisticated fusion models that leverage auxiliary datasets or predictive modeling (Kandasamy et al., 2013; Shen et al., 2015). These approaches address gaps caused not only by clouds, but also by sensor defects or atmospheric interference—providing a comprehensive toolkit for ensuring continuous, reliable optical remote sensing. The scope includes *a)* Single-image spatial methods using spatial autocorrelation to fill missing pixels; *b)* Temporal recovery strategies leveraging multi-date observations; *c)* Spatio-temporal fusion techniques combining imagery from different sensors or dates to synthesize cloud-free series.

Category	***Strengths***	***Weaknesses***	***Applications***	***Computational Cost***
Spatial Gap-Filling (IDW, Kriging)	*Simple and widely used; Kriging provides*	*Ineffective for large or heterogeneous gaps; Kriging requires*	*Filling isolated cloud patches; vegetation index recovery;*	*Low–Moderate (IDW low, Kriging moderate).*

	uncertainty estimates; effective for small, isolated gaps.	*complex variogram modeling.*	*rainfall/reflectance interpolation.*	
Temporal Gap-Filling (Linear, SG, RBF)	*Captures seasonal trends; suitable for gradual phenological changes; climate-integrated RBF models improve robustness.*	*Requires sufficient clean observations; fails under abrupt changes (e.g., floods, harvesting).*	*NDVI time-series reconstruction; drought monitoring; agricultural crop cycle analysis.*	*Low for Linear; Moderate for SG; High for RBF models.*
Multi-Date Compositing (MVC, Median, Quality-based)	*Removes clouds efficiently; operationally simple; widely available in Google Earth Engine.*	*Bias towards extreme values (MVC); temporal dilution; misrepresents rapid changes.*	*Tropical forest monitoring; annual land cover mapping; phenology studies.*	*Low; easy to implement in cloud platforms.*
Temporal Fusion (STARFM, ESTARFM, STNLFFM)	*Integrates high temporal frequency with fine spatial detail; effective for vegetation and land-cover monitoring.*	*Dependent on high-quality input pairs; less accurate in fragmented landscapes; computationally demanding.*	*Crop dynamics; snow/glacier monitoring; urban expansion; deforestation detection.*	*Moderate–High (multi-sensor data processing required).*
Deep Learning Fusion (CNN, LSTM, Transformers)	*Handles complex, non-linear spatio-temporal dynamics; adaptable to heterogeneous landscapes; high accuracy in diverse case studies.*	*Computationally expensive (GPU/TPU required); limited interpretability; requires large training datasets.*	*Flood mapping; precision agriculture; phenology modeling; multi-sensor fusion.*	*High (requires advanced ML frameworks and GPUs).*
SAR–Optical Integration	*Penetrates clouds; complements optical data; robust for monitoring in persistently cloudy regions.*	*Sensor calibration mismatches; integration complexity; limited operational implementations.*	*Flood mapping; deforestation monitoring; landslide assessment; monsoon-affected regions.*	*Moderate–High (multi-sensor preprocessing needed).*

Table 1: provides a consolidated comparative framework summarizing the relative strengths, limitations, applications, and computational costs of the major categories of cloud-resilient techniques.

IV. Categorization of Methods: Gap-Filling vs. Temporal Fusion

Gap-Filling Techniques: This category comprises methods that directly reconstruct missing pixel values due to cloud cover without synthesizing images at finer temporal or spatial resolution. *a) Spatial gap-filling* relies solely on spatial information in the same image. Methods like kriging or spline interpolation use variograms or inverse distance weighting (IDW) to infer missing values from neighbouring clear pixels (Kandasamy et al., 2013; Robinson et al., 2017). These work best for isolated gaps but struggle in complex or extensive cloudy regions. *b) Temporal gap-filling* employs cloud-free observations from previous or subsequent dates to estimate missing values. Common techniques include averaging across years, polynomial fitting, and time-series smoothing (e.g., Savitzky-Golay or mean iterative filters), often incorporating phenological constraints (Jonsson & Eklundh, 2002; Moreno et al., 2014; Zscheischler et al., 2014). *c) Spatio-temporal gap-filling* integrates both spatial and temporal dimensions, filling in missing pixels based on correlated pixels over space and time. Examples include modified neighbourhood-based interpolators, spatial–temporal regression, and Kalman filter models for smoothing (Kang et al., 2005; Shen et al., 2015).

Temporal Fusion Methods: Temporal fusion methods go beyond merely filling gaps—they generate synthetic imagery that merges the high temporal availability of coarse-resolution data (e.g., MODIS) with the fine spatial detail of high-resolution sensors e.g., Landsat or Sentinel-2. *a) Classic fusion models* The Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) and its derivatives (e.g., ESTARFM, STNLFFM, HISTARFM) are widely used. They leverage pairs of fine and coarse images at reference dates plus a coarse image at a prediction date to estimate a synthetic fine-resolution image at the target time (Gao et al., 2006; Wu et al., 2012; Hilker et al., 2009a). Spatial–Temporal Non-Local Filter-based Fusion Model (STNLFFM) improves prediction accuracy in heterogeneous landscapes by exploiting spatial and temporal redundancy (Cheng et al., 2016) HISTARFM incorporates a bias-aware Kalman filter in Google Earth Engine to smooth and fill data while estimating uncertainty at scale (Kandasamy et al., 2020). *b) Deep learning-based fusion:* Recent advances include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid CNN-LSTM architectures. For instance, bidirectional pyramid fusion networks with semantic prior regularization (BPFN-SPR) are designed to flexibly fuse coarse and fine data across temporal intervals. Object-based spatiotemporal unmixing models (e.g., OBSUM) refine fusion by using object-level information instead of pixel-level only (Guo et al., 2023). Each fusion method offers unique strengths: classical models excel in transparent conceptual simplicity, while machine learning approaches enable powerful generalization and adaptation to complex change patterns.

Category	***Approach Type***	***Data Inputs***	***Output***
Spatial Gap-Filling	*Kriging, IDW, interpolation*	*Single image only*	*Restored pixel values in cloud areas*
Temporal Gap-Filling	*Polynomial, filtering-based*	*Multi-date images*	*Time-series repaired values*
Spatio-Temporal	*Regression, neighborhood fill*	*Combined spatial & temporal contexts*	*Reconstructed pixels using both sources*
Temporal Fusion	*STARFM, ESTARFM, STNLFFM*	*Fine–coarse image pairs + coarse at prediction*	*Synthetic fine-resolution imagery*
Deep Learning Fusion	*CNN, LSTM, GAN-based models*	*Multi-sensor, multi-date data, semantic priors*	*High-fidelity, continuous imagery series*

Table 2: Cloud Resilient Techniques

V. GAP FILLING TECHNIQUES

***Spatial Interpolation Techniques for Gap-Filling*:** Spatial interpolation methods are foundational for restoring missing pixels in cloud-contaminated optical imagery by estimating values from adjacent clear observations. These techniques exploit spatial autocorrelation the principle that nearby locations tend to exhibit similar reflectance or spectral properties and thus are valuable when temporal context is unavailable or insufficient. *a) Inverse Distance Weighting (IDW):* Inverse Distance Weighting is a deterministic interpolation method where each unknown pixel is assigned a weighted average of neighboring clear pixels. The weights decline as a function of distance (typically the reciprocal of distance raised to a power p), emphasizing proximity in the estimation process (Tobler’s First Law of Geography). A higher power (e.g., $p > 2$) makes interpolation highly local, while lower values smooth across broader neighborhoods. IDW is computationally efficient, easy to implement, and available in many GIS environments. However, it cannot quantify prediction uncertainty and may be sensitive to spatial clustering or outliers, sometimes producing overly smoothed surfaces or abrupt artifacts. *b) Kriging Methods*: Kriging is a geostatistical, stochastic interpolation method that estimates missing pixel values based on a modeled spatial covariance structure (variogram). Ordinary Kriging (OK) assumes an unknown but constant mean and uses spatially weighted least squares to achieve the best linear unbiased prediction (BLUP) at unsampled sites & offers lower error and more reliable estimates than IDW, especially in datasets with spatial autocorrelation and adequate variogram modeling. Variants such as Regression-Kriging (also known as Kriging with External Drift) incorporate auxiliary variables (e.g., elevation, vegetation indices, DEM derivatives) to improve estimation in heterogeneous areas by modeling and interpolating residuals from a regression model, while more accurate and capable of providing uncertainty estimates, kriging requires rigorous variogram fitting, increased computational cost, and expertise in geostatistics.

***Triangulation and Trend-Based Methods*:** Triangular Irregular Networks (TIN) or Delaunay triangulation techniques predict missing pixel values by interpolating within triangles formed around known observations. Each triangle vertex contributes proportionally to the estimate, making TIN suitable for capturing local spatial structure when sample density is moderate (Aspexit, 2023). Trend surface analysis or polynomial regression interpolation fits a smooth global surface across the study area. While useful for broad-scale gradients, these methods may fail to capture local heterogeneity accurately and may introduce bias if the underlying spatial processes are complex or non-linear.

Evaluation and Suitability for Cloud Gap Filling: *a) IDW:* Rapid and simple, well-suited for isolated or small cloud gaps. However, its sensitivity to spatial sampling density and inability to model spatial trends can limit performance in complex land surfaces. *b) Kriging (OK and RK):* Offers best-in-class accuracy with uncertainty quantification, especially where spatial structure is strong. Highly suitable for larger or irregular gaps but requires sufficient sample density and expertise for variogram calibration. Can be used as default in many operational contexts (MDPI rain-gap studies). *c) Regression-Kriging:* Combines the strengths of regression modeling with geostatistical interpolation of residuals. Particularly beneficial where auxiliary variables (e.g., elevation, historical NDVI) strongly correlate with surface reflectance and can correct systemic bias. *d) TIN and Trend Methods**:*** Useful for rapid interpolation in relatively simple terrains or where computational resources are limited. However, performance remains inferior compared to kriging in terms of both accuracy and robustness.

- Practical Recommendations: For small or isolated cloud gaps in high-resolution imagery, IDW can provide quick and interpretable results, particularly when neighbor density is high and terrain is smooth.

- For larger, irregular gaps, especially in spatially heterogeneous scenes (e.g., mixed forests, urban edges), Ordinary Kriging or Regression-Kriging is preferred due to better handling of spatial structure and provision of uncertainty estimates.

- Where auxiliary datasets are available (e.g., elevation, prior reflectance time series), Regression-Kriging can improve fill accuracy significantly.

- Testing and cross-validation are essential: sampling density, variogram choice, neighbor search radius, power p in IDW, and residual analysis in RK models should be optimized using metrics such as RMSE or MAE.

Spatial interpolation plays a crucial role in addressing data loss due to cloud cover, sensor errors, or incomplete acquisition in optical remote sensing. Among the various techniques, Inverse Distance Weighting (IDW) and Ordinary Kriging (OK) are most commonly employed. Their performance, however, varies depending on landscape complexity, data density, and spatial patternsfindings consistently demonstrated in a range of scholarly case studies.In the context of satellite imagery restoration, a seminal study by Zhang, Li, and Travis (2007) addressed the issue of missing data due to the Scan Line Corrector (SLC) failure in Landsat 7 ETM+ imagery. Their work, titled *“Gap-fill of SLC-off Landsat ETM+ satellite image using a geostatistical approach”* published in the *International Journal of Remote Sensing*, utilized ordinary kriging to reconstruct reflectance values. The method yielded spatially continuous, spectrally consistent imagery, outperforming USGS histogram-matching techniques especially in areas with extensive missing data. This highlighted kriging’s ability to incorporate spatial structure through variogram modelling, making it a powerful tool in geospatial data reconstruction. Conversely, in high-sample-density applications IDW can sometimes match or even exceed kriging in accuracy. This was evident in a study conducted by Hernández-Flores, Holmes, and Rivera-Monroy (2017) titled *“Comparative Analysis of Interpolation Techniques for Mapping Sessile Coral Cover”* published in *PeerJ*. Here, the researchers interpolated benthic coral abundance across Madagascar reef zones. They found that IDW performed better than kriging for certain benthic groups, particularly those with skewed, non-Gaussian distributions like sponges and zoantharians. While kriging smoothed spatial patterns effectively, it sometimes oversimplified localized variability. This case underscores that IDW, due to its deterministic and proximity-based nature, can preserve high-frequency spatial variability when sufficient samples are available and autocorrelation is low.Another large-scale example comes from arid Saudi Arabia, where Al-Ahmadi and Al-Amri (2023) conducted a comprehensive evaluation of spatial interpolation methods for rainfall data. In their paper, *“Assessment of Spatial Interpolation Techniques for Rainfall Data in Arid Regions of Saudi Arabia”* published in *Sustainability*, the authors compared IDW and multiple kriging variants, including cokriging and empirical Bayesian kriging. Their results confirmed kriging’s superior performance across almost all metrics, particularly in topographically variable regions and where the rain gauge network was sparse. IDW, while simpler and faster, produced higher root mean square error (RMSE) values and lacked the uncertainty estimation provided by kriging. Collectively, these case studies illustrate the nuanced strengths and weaknesses of each method. IDW is ideal for rapid interpolation over small, data-rich gaps and is highly accessible for non-specialists. However, it does not offer any statistical measure of interpolation uncertainty and may produce artifacts in areas with irregular data spacing. Kriging, on the other hand, models the underlying spatial autocorrelation explicitly and provides both predicted values and error estimates, making it preferable for complex terrains and mission-critical remote sensing tasks. Its downside remains the need for careful variogram calibration and computational complexity may be sufficient for short-term or localized gap-filling, especially when speed is a priority, kriging is better suited for larger gaps, heterogeneous environments, or applications where precision and statistical rigor are paramount. The cited research by Zhang et al. (2007), Hernández-Flores et al. (2017), and Al-Ahmadi & Al-Amri (2023) provides empirical evidence for tailoring the choice of spatial interpolation method to specific data conditions and landscape characteristics.

VI. Temporal Interpolation Techniques for Gap-Filling

Temporal interpolation methods estimate missing pixel values by leveraging valid observations from preceding and subsequent dates in a time series. These methods assume that surface reflectance or vegetation indices follow predictable temporal trajectories, making them suitable for reconstructing data gaps without relying on spatial neighbors. Some of the techniques are *a) Linear Interpolation*: This simplest form assumes a linear progression between two valid observations. It is widely used due to its ease of implementation but struggles with capturing non-linear seasonal or abrupt changes. *b) Polynomial or Harmonic Fitting:* Methods that fit higher-order curves or sinusoidal functions to model seasonal dynamics more accurately than linear methods. They are better suited to periodic phenology but require enough observations for stable curve fitting (e.g., sum-of-harmonics models). *c) Savitzky–Golay (SG) Filtering:* A smoothing filter commonly applied in remote sensing time series to suppress high-frequency noise while preserving phenological trends. Weighted SG variants also correct biases caused by cloud-related NDVI underestimation. *d) Data-Driven Models (e.g., RBF Networks):* Techniques like radial basis function neural networks that harness interannual climate relationships (e.g., precipitation or temperature) to model NDVI variations and predict missing values across years.

In the study titled *"A practical approach to reconstruct high-quality Landsat NDVI time-series data by gap filling and the Savitzky–Golay filter" by Liu et al., 2021*, researchers proposed the GF-SG method, integrating temporal interpolation with a weighted Savitzky–Golay filter. They used MODIS NDVI time series to guide gap-filling in Landsat imagery and applied SG smoothing to reduce noise caused by residual cloud effects. The approach effectively reconstructed long-term NDVI series in Google Earth Engine, addressing both missing data and noise while being computationally efficient. In *"Gap Filling for Historical Landsat NDVI Time Series by Integrating Climate Data"*, authors built a gap-filling model combining Landsat NDVI with seasonal climate variables (precipitation, temperature, solar radiation). They trained Radial Basis Function Networks (RBFNs) to predict NDVI series in years with severe data gaps, using preceding years and climate variability as inputs. This method enabled accurate reconstruction of NDVI time series with varying gap proportions, outperforming traditional interpolation under large-data loss conditions. Zhang et al. (2024) evaluated the effect of linear interpolation on time series used for land-cover classification via modern deep learning models (Bi-LSTM and Transformer). They compared time-series with and without interpolation over Sentinel-2 and Landsat data in the Amur River Basin. The classification accuracy between interpolated and non-interpolated time series differed by less than 0.5%, indicating that linear interpolation did not significantly enhance classification performance when using models capable of handling missing data. Temporal interpolation methods show varied performance depending on data characteristics and application context: Simple Linear Interpolation works best for small gaps and gradual phenological transitions but fails to capture non-linear or abrupt dynamics. Additionally, its benefit for classification tasks is minimal when advanced models like LSTM or Transformers can handle missing data natively, Savitzky–Golay Filtering (e.g., GF-SG) enhances gap-filling by smoothing out cloud-induced noise while retaining seasonal trajectory. When combined with auxiliary MODIS time series, it produces coherent and robust NDVI reconstructions across Landsat datasets, Climate-Integrated Models (RBFNs) capture interannual variability driven by weather and overcome limitations of both simple interpolation and harmonic models. This approach excels where cloud coverage produces multi-year data gaps and climate drivers are strong predictors.

VII. Multi-Date Compositing Techniques for Gap-Filling

Multi-date compositing is a cloud-resilient technique that synthesizes cloud-free imagery by integrating observations from multiple dates within a defined time window (e.g., 16 days, 1 month, or 1 year). The core idea is to select the best pixel based on criteria such as reflectance, NDVI, or atmospheric clarity among available cloud-free observations to fill gaps in optical data caused by clouds or shadows. This technique is especially useful in cloud-prone regions and is commonly implemented using one or more of the following compositing strategies: *a) Maximum Value Composite (MVC):* Selects the pixel with the maximum value of a vegetation index (e.g., NDVI) during the compositing period, effective in vegetation monitoring; assumes healthy vegetation has the highest NDVI, may exaggerate peak greenness; sensitive to outliers. *b) Median Composite* chooses the median value from all available valid pixels within the period; it reduces influence of outliers or atmospheric noise & may suppress peak signals; not ideal for dynamic land cover. *c) Mean Composite* averages all valid observations over the compositing window, Smooths fluctuations & good for stable surfaces but sensitive to cloud-contaminated values if not properly masked. *d) Quality-Based Composite* selects the best pixel based on a combination of criteria (e.g., cloud score, view angle, shadow score), incorporates atmospheric quality and sun-sensor geometry and more computationally demanding. In a case study on Pan-tropical Sentinel-2 cloud-free annual composite datasets by Xueting Jin et. al. developed cloud-free annual composites from Sentinel-2 MSI imagery over tropical regions between 2016 and 2021 using a quality-based compositing algorithm implemented in Google Earth Engine. The method prioritized pixels with the lowest cloud probability and optimal solar geometry. The resulting composites were highly useful for land cover monitoring and forest change detection in regions with persistent cloud cover like the Amazon and Southeast Asia. In case study 2 on Ten-meter Sentinel-2A cloud-free composite – Southern Africa 2016 by Duncan J. McFarlane et. al. sed a Maximum Value Composite (MVC) approach to build a regional Sentinel-2 NDVI composite for Southern Africa. The composite was produced quarterly by selecting the pixel with the highest

NDVI within each 3-month window. It effectively removed cloud contamination and enabled seasonal analysis of vegetation health and phenology across large areas. In a case study 3 Multi-Temporal Pixel-Based Compositing for Cloud Removal Based on Cloud Masks Developed Using Classification Techniques by Peng Zhang et. al. proposed two composite methods MaxComp-1 using Support Vector Machine (SVM) cloud masks and MaxComp-2 using U-Net deep learning masks. A maximum NDVI selection was applied only to pixels labeled as cloud-free. The U-Net-based composite produced clearer images and fewer artifacts compared to traditional median compositing, highlighting the power of integrating machine learning in multi-date compositing.

VIII. Temporal Fusion Methods in Cloud-Resilient Remote Sensing

Method	Case study / Region	Primary metric	Reported performance	Processing time (typical)	Data requirements	Notes
STARFM	Australian savanna (dense time-series study)	NDVI RMSE	≈ **0.027** (NDVI)	Moderate	Fine–coarse pairs (Landsat + MODIS) at ≥1 reference date pair; pre-processing for atmospheric correction and coregistration	Strong at capturing seasonal phenology in relatively homogeneous vegetation; degrades in highly fragmented landscapes.
ESTARFM	Heterogeneous landscapes, South China / Qilian region	Correlation (r) / structural similarity	**r > 0.6** (reported across heterogeneous terrain)	Moderate–High	Multiple fine–coarse image pairs close to the prediction date (often 2+ fine scenes) plus coarse image at prediction	Improved mixing-pixel handling vs STARFM; computationally heavier and needs more fine-resolution inputs.
STNLFFM / SSFIT (statistical advanced fusion)	Mixed case studies (simulation + real)	RMSE / SSIM	RMSE comparable to STARFM; improves structural similarity in heterogeneous scenes	High	Multiple images, spatial redundancy exploited; parameter tuning needed	Reduces temporal noise via non-local filtering; better in mixed land covers.
CNN–LSTM (deep fusion)	Snow / glacier and agricultural studies (recent experiments)	RMSE; structural similarity (SSIM)	**Lower RMSE and higher SSIM** relative to STARFM in reported studies; typical RMSE reductions **~10–15%** vs classical fusion	High (GPU recommended)	Large training dataset (labeled pairs), multi-sensor inputs (optical ± coarse), significant compute (GPU/TPU)	Captures nonlinear spatio-temporal patterns; better generalization in heterogeneous landscapes when well trained; requires careful validation.
Transformer-based fusion	Recent experimental studies (mixed agricultural mosaics)	RMSE; long-range temporal consistency	Improved long-range dependency modelling, often better than CNN-LSTM for long gaps (study dependent)	Very High	Large-scale training data, GPU clusters, careful architecture tuning	State-of-the-art for long-range temporal dependencies but still experimental for many EO operational tasks.
SAR–Optical Integration (hybrid CNN)	Flood events, monsoon regions, forest monitoring (2022–2024 experiments)	Detection accuracy, map completeness	High detection rates in cloudy conditions; complements optical fusion where clouds persist	Moderate–High (preprocessing SAR + fusion model)	SAR (Sentinel-1) + optical inputs; co-registration, speckle reduction, radiometric harmonization	Essential when optical observations are unavailable; improves disaster mapping and forest-change detection under clouds.

Table 3:Comparative performance of representative cloud-resilient methods across case studies, including classical fusion and recent deep learning advances.

Temporal fusion methods aim to reconstruct high spatial and temporal resolution imagery by combining high-frequency, coarse-resolution data (such as MODIS) with low-frequency, high-resolution imagery (such as Landsat or Sentinel). These methods address cloud-induced gaps by predicting surface reflectance at dates when high-resolution data are missing. Prominent models include: *a) STARFM (Spatial and Temporal Adaptive Reflectance Fusion Model):* Developed by Gao et al. (2006), STARFM uses image pairs of Landsat and MODIS at known dates and the MODIS image at the prediction date to generate synthetic high-resolution reflectance. It assumes temporal consistency and estimates changes based on spectral differences between pairs. *b) ESTARFM (Enhanced STARFM):* An extension that accounts for heterogeneous land surfaces by considering both spectral and temporal changes, improving performance in mixed pixels and varying landscapes (Wu et al., 2012). *c) SSFIT (similarity-based spatio-temporal fusion inside time series)* Further refinements that optimize weighting strategies, exploit spatial redundancy or time-series correlations, shows robust phenology recovery even with few input Landsat images. *d) ELSTFM (Enhanced Linear Spatio-Temporal Fusion Model):* A linear unmixing approach combining Landsat and MODIS using regression-based fusion, offering better structural similarity and reduced RMSE than STARFM. *e) Machine learning and deep learning hybrids*: Emerging techniques deploy CNN-LSTM or fusion networks that learn spatio-temporal dependencies directly from multi-sensor data, outperforming classical models in complex or non-linear environments. Schmidt et al. in his paper Long-term data fusion for a dense time series analysis with MODIS and Landsat imagery in an Australian Savanna applied STARFM to blend Landsat TM/ETM+ (30 m) and MODIS (500 m) over North Queensland savanna across 2000–2007 (8-day interval). Results showed high correlation (r = 0.89–0.99) between synthetic and actual Landsat values. STARFM accurately captured vegetation phenology and preserved spatial details, with NDVI RMSE around 0.027. In the second case study Wu et al. in his paper Assessing the Accuracy of Spatial and Temporal Image Fusion Model of Complex Area in South China assessed five fusion approaches including STARFM, ESTARFM, and others over broken terrain near Nanjing. Most models (except LORENZO) achieved r > 0.6 with real Landsat ETM+ images. ESTARFM and STARFM produced reliable spectral reconstructions, outperforming simpler statistical models in mixed landscapes. Liu et. al. in his paper Study on snow cover change based on the fusion of Sentinel-2 and MODIS images applied ESTARFM (on Google Earth Engine) to fuse Sentinel-2 NDSI (10 m) with MODIS NDSI (500 m), generating daily high-resolution snow cover data between November 2021 and May 2022. Fusion outputs closely matched actual Sentinel-2-derived NDSI imagery in both visual appearance and quantitative snow cover distribution drastically improving snowmelt tracking in mountainous terrain.

IX. Case-Application

CNN/LSTM: ::
• You have large labeled training datasets and access to GPUs/TPUs.
• Complex, nonlinear spatio-temporal patterns or long gaps must be modeled.
• You prioritize accuracy and generalization over simple, interpretable models.

ESTARFM/STNLFFM: ::
• The study area is heterogeneous (mixed agriculture/forest/urban).
• You can supply multiple fine-resolution reference images near the prediction date.
• Greater accuracy at edges and mixed pixels is required and extra compute is acceptable.

1. Agriculture monitoring: Timely, dense observations are crucial for phenology detection, stress monitoring, and yield forecasting. Temporal fusion (STARFM/ESTARFM) produces synthetic high-resolution time series that preserve crop phenology in relatively homogeneous cropland. Deep models (CNN–LSTM, transformers) have shown improved performance in heterogeneous or fragmented agricultural mosaics, enabling finer discrimination of field-level dynamics when adequate training data exist. For seasonal crops with short growth cycles, multi-sensor fusion plus gap-filling (SG filtering) is often required to avoid missing critical stages.

2. Deforestation and forest degradation: Forest change detection benefits from dense time series that reveal canopy disturbance and recovery. STARFM and ESTARFM have been used to build continuous NDVI time series for disturbance mapping; however, in tropical forests with persistent clouds, quality-weighted compositing combined with machine-learning cloud masks (e.g., U-Net) and SAR-optical integration provide more reliable detection and reduce omission errors. For near-real-time alerts, SAR–optical hybrid approaches are preferable because radar is unaffected by clouds.

3. Snow and glacier tracking: High spatial detail is required to track snowline migration and glacier area change. ESTARFM and STNLFFM used with Sentinel-2 and MODIS have produced daily to weekly high-resolution NDSI products for snow monitoring in mountain regions. Deep learning fusion (CNN–LSTM) improves structural similarity and can reduce RMSE in snow index reconstruction, especially where terrain heterogeneity and shadowing are severe.

4. Flood assessment and disaster response: Rapid flood mapping demands immediate usable maps during and after peak flood stages—times when clouds are frequently present. SAR (Sentinel-1) is the primary source for these events; combining SAR with available optical data via hybrid CNNs increases inundation map accuracy and enables multisource change characterization. Temporal fusion of pre- and post-event coarse–fine pairs plus SAR input yields robust monitoring in monsoon and tropical settings.

X. RECOMMENDATIONS FOR PRACTITIONERS

To facilitate method selection, a set of practitioner-oriented recommendations is provided below:

STARFM:
• You operate in relatively homogeneous landscapes (e.g., savanna, large cropland blocks).
• Fine–coarse image pairs (Landsat/Sentinel-2 + MODIS/VIIRS) are available close in time.
• Computational resources are moderate and operational simplicity is preferred.

SAR Optical Integration: ::
• The region has **persistent cloud cover** (tropics, monsoon belts, mountainous zones) and optical-only methods fail.
• Rapid disaster mapping (floods, landslides) or continuous forest monitoring is required.
• You can perform SAR preprocessing (speckle reduction, co-registration) or have access to tools that do it.

PRACTICAL WORKFLOW TIPS :

1. HYBRID FIRST: TRY SIMPLE GAP-FILLING & COMPOSITING (LOW COST) FIRST FOR SMALL GAPS; ESCALATE TO FUSION OR SAR INTEGRATION WHEN GAPS ARE PERSISTENT OR ACCURACY DEMANDS INCREASE.
2. VALIDATION IS ESSENTIAL: ALWAYS RESERVE INDEPENDENT FINE-RESOLUTION OBSERVATIONS (OR IN-SITU DATA) FOR EVALUATION; REPORT RMSE/SSIM/R^2 AND UNCERTAINTY MEASURES.
3. COMPUTATIONAL PLANNING: PROTOTYPE WITH CLASSICAL FUSION (STARFM) TO SET BASELINES; IF DEEP MODELS ARE NEEDED, RUN LOCALIZED TRAINING EXPERIMENTS BEFORE SCALING.
4. OPERATIONALIZATION: PRIORITIZE METHODS WITH EXISTING READY-TO-USE IMPLEMENTATIONS (GEE STARFM LIBRARIES, OPEN SAR PROCESSING CHAINS) WHEN TIME-TO-DELIVERY MATTERS.

XI. TEMPORAL FUSION METHODS IN CLOUD-RESILIENT REMOTE SENSING

Cloud-resilient remote sensing relies heavily on two core strategies to address missing optical data due to cloud contamination:

- Gap-Filling Techniques: Fill missing pixels within an image (spatially or temporally).
- Temporal Fusion Techniques: Generate synthetic high-resolution images at desired time intervals by integrating multi-sensor data over time.

Parameter	Gap-Filling Techniques	Temporal Fusion Techniques
Objective	*Fill missing pixels in individual or composite images*	*Generate temporally dense, high-res time series*
Input Data	*Single or multi-date optical images (same resolution)*	*High-res (e.g., Landsat) + low-res frequent data (e.g., MODIS)*
Spatial Resolution	*Retains original high resolution*	*Synthesizes high-res from low-res temporal data*
Temporal Resolution	*Limited to acquisition dates of optical sensors*	*High (e.g., daily to weekly) using temporal extrapolation*
Examples of Techniques	*Spatial interpolation, temporal interpolation, compositing*	*STARFM, ESTARFM, ELSTFM, SSFIT, CNN-LSTM*
Computational Complexity	*Generally low to moderate*	*Moderate to high (especially deep learning-based models)*
Use Case	*Filling clouds or shadows in specific scenes or mosaics*	*Monitoring vegetation dynamics, urban growth, phenology at high temporal frequency*
Accuracy Sensitivity	*Highly sensitive to gap size and image quality*	*Sensitive to sensor mismatch, time gaps, and land cover heterogeneity*
Common Tools Used	*Fmask, RSGISLib, Google Earth Engine compositing functions*	*GEE with custom models, open-source STARFM libraries, Python/ML frameworks*

Table 4: Comparative Analysis of Temporal Fusion & Gap filling Techniques

XII. CHALLENGES AND LIMITATIONS OF GAP-FILLING AND TEMPORAL FUSION TECHNIQUES

Gap-filling methods are primarily designed to restore cloud-contaminated regions or missing pixels in individual satellite images or mosaics using spatial or temporal information. While they offer practical solutions, their performance is limited by several constraints. *a)* Spatial interpolation methods, such as nearest-neighbor, inverse distance weighting (IDW), and kriging, rely on the spatial continuity of pixel values. These techniques assume that nearby pixels are similar in spectral behavior. However, in many real-world scenarios, especially in heterogeneous landscapes such as urban–rural interfaces, mountainous terrain, or fragmented croplands, this assumption fails. Large, contiguous cloud gaps, especially those that span entire land features, cannot be effectively filled through interpolation due to the absence of valid contextual pixels. Consequently, these methods often result in blurred edges, inaccurate texture representation, and oversimplified land cover classifications. *b)* Temporal Interpolation Challenges: Temporal interpolation techniques (linear, spline, or polynomial-based) estimate missing pixel values using preceding and succeeding cloud-free observations. These methods are highly effective in capturing gradual surface changes, such as phenological transitions. However, their reliability significantly declines under certain conditions:Lack of sufficient temporal observations: In tropical and monsoonal regions, where cloud cover may persist for weeks, the absence of clean observations before or after a cloudy scene impairs the interpolation process.Non-linear land changes: Sudden land use transitions (e.g., floods, deforestation, harvesting) cannot be effectively modeled using simple temporal interpolation, leading to large errors and misrepresentation of land conditions. *c)* Multi-Date Compositing LimitationsCompositing strategies, such as Maximum Value Composite (MVC), median, and mean compositing, are widely used due to their simplicity and implementation in platforms like Google Earth Engine. However, these techniques face several limitations:Bias in data representation: MVC often overestimates vegetation greenness by favoring peak NDVI values, while median or mean methods may smooth out genuine temporal variability.Temporal dilution: Composites represent a snapshot over a period (e.g., monthly, annual) rather than the actual condition on a specific date. This reduces their utility for time-sensitive applications such as phenology tracking, disaster assessment, or short-term land degradation analysis. Loss of short-term dynamics: Rapid environmental changes occurring within the compositing window may be lost or misrepresented.

Temporal Fusion Techniques aim to integrate high-temporal low-resolution data (e.g., MODIS, VIIRS) with low-temporal high-resolution data (e.g., Landsat, Sentinel-2) to reconstruct synthetic images that retain both spatial and temporal fidelity. While these methods significantly advance the potential of cloud-resilient monitoring, they are not without shortcomings. *a.)* Sensor Mismatch and Calibration assume that different sensors observe the same surface features under similar radiometric and geometric conditions. However, differences in spectral response functions, temporal acquisition offsets & sensor viewing angles often introduce inconsistencies. These mismatches can lead to fusion artifacts such as brightness discontinuities, spectral distortions, and radiometric bias in synthetic outputs. Correcting these disparities requires complex pre-processing, including atmospheric correction, resampling, and normalization, which are not always feasible at scale. *b.)* Dependency on High-Quality Inputs-Models like STARFM and ESTARFM require high-quality, cloud-free image pairs from both coarse and fine resolution sensors. In cloud-prone areas or during extended cloudy seasons, obtaining these input pairs becomes challenging. Without adequate inputs, model reliability deteriorates. Additionally, models like ESTARFM demand two or more Landsat images near the prediction date, which may not always be available. *c.)* Limited Performance in Heterogeneous Landscapes- Fusion models assume spectral and temporal similarity within a localized moving window. In heterogeneous landscapes with mixed land use (e.g., forest-agriculture mosaics or urban zones), this assumption breaks down. As a result, fusion outputs in such areas often exhibit edge blurring, spatial discontinuities, or inaccurate land cover transitions. Even models like ESTARFM, which attempt to address heterogeneity, still show reduced accuracy in topographically complex or highly dynamic environments. *d.)* Temporal Drift and Error Propagation - Synthetic time series generated through temporal fusion are subject to temporal drift a cumulative divergence from actual surface reflectance values due to compounding prediction errors. This is particularly problematic in long-term monitoring where predictions are made across extended time periods without frequent recalibration. Once errors enter the fused time series, they propagate and may affect downstream applications such as change detection, classification, or phenology modelling.*e.)* Computational and Operational Limitations-Temporal fusion models, especially those incorporating enhanced algorithms (e.g., SPSTFM, SSFIT) or deep learning (e.g., CNN-LSTM), are computationally intensive. They require: large volumes of satellite data, High-performance computing infrastructure (e.g., GPU clusters), Skilled data scientists for model design and parameter tuning.Moreover, many advanced models are still experimental and not yet implemented in widely accessible platforms like Google Earth Engine, which restricts their operational applicability for policy, planning, or disaster response. *f.)* Validation and Generalization Gaps-A recurring limitation in fusion literature is the lack of standardized validation protocols. Ground-truth data are seldom available at the required spatio-temporal resolution to validate synthetic outputs. Furthermore, many fusion models are region-specific and trained for particular climatic or ecological conditions. Their transferability to new geographies without retraining or recalibration is uncertain and often untested. Cloud contamination continues to be a major impediment to the effective use of optical remote sensing, particularly in regions with frequent or prolonged cloud cover such as tropical rainforests, mountainous terrains, and monsoonal zones. The ability of clouds to obscure surface features compromises the temporal continuity, spatial completeness, and analytical reliability of satellite-derived information, thereby affecting critical applications such as vegetation phenology, land cover change detection, and disaster monitoring.

XIII. Conclusion

To address this pervasive limitation, a broad spectrum of cloud-resilient methods has been developed, which can be broadly classified into gap-filling and temporal fusion techniques. Gap-filling methods, including spatial interpolation, temporal interpolation, and multi-date compositing, offer relatively straightforward mechanisms for reconstructing missing pixels within individual scenes. While effective under certain conditions, these methods are constrained by the extent of data gaps, landscape complexity, and the temporal availability of cloud-free images. They often fall short in capturing dynamic environmental processes and offer limited support for continuous monitoring.Temporal fusion methods, on the other hand, represent a significant advancement in remote sensing by enabling the generation of synthetic, cloud-free imagery at both high spatial and temporal resolutions. Techniques such as STARFM, ESTARFM, ELSTFM, and more recent deep learning-based fusion models like CNN-LSTM, have demonstrated promising results across diverse ecological and geographical settings. These methods allow researchers to reconstruct dense time-series data, integrate multi-sensor observations, and track subtle environmental changes that would otherwise remain hidden in cloud-obstructed datasets.However, despite these advancements, several challenges remain. Fusion techniques are highly dependent on the availability and quality of input datasets, and their accuracy is often compromised in heterogeneous or rapidly changing landscapes. Sensor mismatches, temporal inconsistencies, computational demands, and limited operational toolkits further constrain their scalability and generalizability. Moreover, validation remains a critical concern due to the lack of high-resolution, ground-based truth data.

Looking forward, the future of cloud-resilient remote sensing lies in the integration of multi-source datasets, including synthetic aperture radar (SAR), meteorological data, and historical time-series, into intelligent fusion frameworks. There is a growing need to develop automated, transferable, and interpretable models that can function effectively across biomes and seasons. Platforms such as Google Earth Engine, combined with open-access machine learning libraries, hold promise in democratizing access to these techniques.In conclusion, while no single method can fully eliminate the impact of cloud contamination, the strategic application and further evolution of gap-filling and temporal fusion techniques provide a robust pathway towards achieving more reliable, high-frequency, and spatially complete Earth observation. These innovations are not only critical for advancing scientific understanding but also for supporting data-driven policy and sustainable resource management in a world increasingly shaped by climate variability and environmental change.

References

[1] Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

[2] Gao, F., Masek, J., Schwaller, M., & Hall, F. G. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. IEEE Transactions on Geoscience and Remote Sensing, 44(8), 2207–2218. https://doi.org/10.1109/TGRS.2006.872081

[3] Zhu, X., Chen, J., Gao, F., Chen, X., & Masek, J. G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. Remote Sensing of Environment, 114(11), 2610–2623. https://doi.org/10.1016/j.rse.2010.05.032

[4] Wu, M., Niu, Z., & Wang, C. (2014). Assessing the accuracy of spatial and temporal image fusion model of complex area in South China. Journal of Geo-information Science, 16(5), 776–783.

[5] Schmidt, M., Udelhoven, T., Röder, A., & Gill, T. K. (2012). Long-term data fusion for a dense time series analysis with MODIS and Landsat imagery in an Australian savanna. Journal of Applied Remote Sensing, 6(1), 063512. https://doi.org/10.1117/1.JRS.6.063512

[6] Liu, Q., & Zhang, Y. (2022). Study on snow cover change based on the fusion of Sentinel-2 and MODIS images. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-3–2022, 333–338. https://doi.org/10.5194/isprs-annals-V-3-2022-333-2022

[7] Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel-2 images. Remote Sensing of Environment, 159, 269–277. https://doi.org/10.1016/j.rse.2014.12.014

[8] Frantz, D., Röder, A., & Udelhoven, T. (2015). Exploiting the Landsat archive for land cover change detection using an adaptive two-date composite method. Remote Sensing, 7(1), 122–144. https://doi.org/10.3390/rs70100122

[9] Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., ... & Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. Remote Sensing of Environment, 145, 154–172. https://doi.org/10.1016/j.rse.2014.02.001

[10] Li, X., Wang, C., Wang, H., & Wu, H. (2012). A spatial–temporal fusion model for multi-resolution remotely sensed images based on sparse representation. Remote Sensing Letters, 3(4), 309–318. https://doi.org/10.1080/01431161.2011.597780

[11] Huang, B., Song, H., Cui, X., Peng, H., & Xu, Z. (2013). Spatial and temporal reflectance fusion via

sparse representation. IEEE Transactions on Geoscience and Remote Sensing, 51(11), 5902–5912. https://doi.org/10.1109/TGRS.2013.2255255

[12] Wang, J., Fu, B., Qiu, Y., & Chen, L. (2001). Soil nutrient management based on land use and land cover change in a small catchment on the Loess Plateau in China. Nutrient Cycling in Agroecosystems, 59(1), 47–55. https://doi.org/10.1023/A:1017532913299

[13] Hilker, T., Wulder, M. A., Coops, N. C., Linke, J., McDermid, G., Masek, J. G., ... & White, J. C. (2009). A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. Remote Sensing of Environment, 113(8), 1613–1627. https://doi.org/10.1016/j.rse.2009.03.007

[14] Wang, L., Liu, D., & Wang, L. (2017). A multi-temporal method for cloud and cloud shadow detection in optical satellite imagery using deep learning. ISPRS Journal of Photogrammetry and Remote Sensing, 129, 297–312. https://doi.org/10.1016/j.isprsjprs.2017.04.015

[15] Liu, D., Wang, L., & Dong, Z. (2019). Spatiotemporal fusion for high spatial resolution land surface temperature mapping using deep convolutional neural networks. Remote Sensing, 11(16), 1913. https://doi.org/10.3390/rs11161913

[16] Singh, A. (1989). Review article digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing, 10(6), 989–1003. https://doi.org/10.1080/01431168908903939

[17] Roy, D. P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., & Lindquist, E. (2008). Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. Remote Sensing of Environment, 112(6), 3112–3130. https://doi.org/10.1016/j.rse.2008.03.009

[18] Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. Remote Sensing of Environment, 114(12), 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

[19] Foody, G. M., & Mathur, A. (2004). A relative evaluation of multiclass image classification by support vector machines. IEEE Transactions on Geoscience and Remote Sensing, 42(6), 1335–1343. https://doi.org/10.1109/TGRS.2004.827257

[20] Nguyen, H. D., & Ban, Y. (2019). A spatiotemporal context-based deep learning framework for land use/land cover mapping. Remote Sensing, 11(8), 979. https://doi.org/10.3390/rs11080979

[21] White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., & Coops, N. C. (2014). Pixel-based image compositing for large-area dense time series applications and science. Canadian Journal of Remote Sensing, 40(3), 192–212. https://doi.org/10.1080/07038992.2014.945827

[22] Hagolle, O., Huc, M., Pascual, D. V., & Dedieu, G. (2015). A multi-temporal and multi-spectral method to estimate aerosol optical thickness over land, for the atmospheric correction of Formosat-2, Landsat, VENμS and Sentinel-2 images. Remote Sensing, 7(3), 2668–2691. https://doi.org/10.3390/rs70302668

[23] Banskota, A., Kayastha, N., Falkowski, M. J., Wulder, M. A., Froese, R. E., & White, J. C. (2014). Forest monitoring using landsat time series data: A review. Canadian Journal of Remote Sensing, 40(5), 362–384. https://doi.org/10.1080/07038992.2014.987376

[24] Zhang, C., & Roy, D. P. (2017). A spatially explicit modeling approach for reconstructing Landsat surface reflectance from MODIS. Remote Sensing of Environment, 197, 329–342. https://doi.org/10.1016/j.rse.2017.05.033

[25] Zhu, X., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment, 118, 83–94.

[26] Guo, H., Ye, D., & Bruzzone, L. (2023). *OBSUM: An object-based spatial unmixing model for spatiotemporal fusion of remote sensing images* [Preprint]. arXiv. https://arxiv.org/abs/2310.09517

[27] Chen, Y., Liu, Q., & Zhang, Y. (2021). A practical approach to reconstruct high-quality Landsat NDVI time-series data by gap filling and the Savitzky–Golay filter. *International Journal of Applied Earth Observation and Geoinformation*, (2021). https://doi.org/10.1016/j.jag.2021.

[28] Yu, W., et al. (2021). Gap filling for historical Landsat NDVI time series by integrating climate data. *Remote Sensing*, *13*(3), 484.

[29] Liu, Q., & Zhang, Y. (2022). Study on snow cover change based on the fusion of Sentinel-2 and MODIS images. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-3–2022, 333–338. https://doi.org/10.5194/isprs-annals-V-3-2022-333-2022

[30] Che, X., & coauthors. (2024). Linearly interpolating missing values in time series contributes little improvement for land-cover classification using Bi-LSTM and Transformer models. *International Journal of Applied Earth Observation and Geoinformation* (2024). https://doi.org/10.1016/j.jag.2024.

[31] Zhang, H. K., et al. (2024). Classifying raw irregular time series (CRIT) for large-area land-cover mapping. https://doi.org/10.1016/j.patter.2024

[32] Wang, L., et al. (2023). CNN–LSTM hybrid models for spatio-temporal fusion: application to snow/ glacier monitoring and NDVI reconstruction. *Remote Sensing*, 15(6), 1190. https://doi.org/10.3390/rs15061190

[33] Shen, Y., et al. (2023). SAR–optical deep fusion for flood mapping during persistent cloud cover: a multi-event evaluation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 201, 150–165. https://doi.org/10.1016/j.isprsjprs.2023.02.012

[34] Li, M., & colleagues. (2024). Transformer-based temporal fusion for high-resolution time series reconstruction. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1–14. https://doi.org/10.1109/TGRS.2024.

RECENT SURVEY (2024) ON SAR–OPTICAL FUSION AND OPERATIONAL PIPELINES: BILOTTA, G., ET AL. (2024). DEEP LEARNING INNOVATIONS AND SAR–OPTICAL SYNERGISTIC APPROACHES FOR OPERATIONAL LAND MONITORING. *REMOTE SENSING REVIEWS*, 42(2), 100–123.