AI-Driven Change Detection Using SAR, LIDAR, And Sentinel-2 Data for Landslide Monitoring and Disaster Early Warning Systems

Ratul Debnath¹; Aditya Dhanekula²;

[1]. *Graduate GIS Assistant, Sam Houston State University, Huntsville, TX, USA;*
Email: *ratuldebnathgis@gmail.com*

[2]. *Stevens Institute of Technology, New Jersey, USA*
Email: *dhanekulaaditya1@gmail.com*

Abstract

This study addressed the problem that many AI-driven landslide early-warning pipelines generate technically strong change maps but lack auditable, decision-aligned evidence products that stakeholders can trust and act on consistently under uncertainty. The purpose was to evaluate, in a quantitative cross-sectional, case-based design, whether multi-sensor AI change evidence (SAR, LiDAR, Sentinel-2) and socio-technical decision constructs jointly explain perceived Early Warning Effectiveness (EWE). Data were collected from a professional sample of N = 162 screened-valid respondents drawn from operational roles (disaster management 31.5%, GIS/remote sensing 27.8%, engineering 22.2%, planning 18.5%; mean experience 8.6 years) evaluating the case outputs and decision readiness constructs. Key variables included Perceived AI Performance (PAI), Interpretability (INT), Trust in Alerts (TRU), Decision Confidence (DCF), sensor contribution perceptions (PSC_SAR, PSC_LiDAR, PSC_S2), and the fused Multi-Sensor Change Evidence Score (MSCES) as predictors, with EWE as the main dependent variable. The analysis plan applied descriptive statistics, reliability testing (Cronbach's a), Pearson correlations, and hierarchical regressions comparing single-sensor baselines to fusion and a full decision-aligned model, plus threshold sensitivity and agreement/conflict diagnostics. Construct means were high (PAI M = 4.21, INT 4.02, TRU 4.08, DCF 4.16, EWE 4.12) with strong reliabilities (a = 0.83–0.90). Correlations supported the hypothesized pathway (PAI–EWE r = 0.68, TRU–EWE 0.62, INT–EWE 0.59, all p < .001). Regression results showed LiDAR-only (R^2 = 0.31) exceeded SAR-only (0.24) and Sentinel-2-only (0.18), while MSCES fusion improved explanatory power (R^2 = 0.46, β = 0.68, p < .001). In the full model, explained variance increased to R^2 = 0.62 (Adj. 0.61), with significant unique effects for PAI (β = 0.34), TRU (β = 0.29), MSCES (β = 0.27) and INT (β = 0.18). Operationally, the case pipeline produced 128 mapped change objects, with 46 high-priority zones; agreement classes showed higher confidence when evidence converged (3-sensor agreement 38.4%, confidence M = 4.34 vs single-sensor-only 19.9%, M = 3.61). Trigger reliability further demonstrated robustness: triggered zones dropped from 46 (MSCES 0.60) to 24 (0.80) while a stable core of 24 zones persisted (TSI 0.52 at 0.60, 1.00 at 0.80). These findings imply that deployable landslide early warning should prioritize multi-sensor fusion plus interpretable, graded outputs and threshold-stability reporting to reduce alarm fatigue, improve accountability, and strengthen calibrated trust in alerts across agencies.

INTRODUCTION

Landslide monitoring and disaster early warning systems are anchored in a shared vocabulary that links earth-surface processes, measurement physics, and risk governance. A landslide is commonly treated as a gravity-driven mass movement of rock, debris, or earth along a slope, and its hazardousness is understood through the coupling of terrain susceptibility, triggering conditions, and exposure (Berger et al., 2012). Global reporting syntheses have demonstrated that landslides generate recurring and unevenly distributed societal losses, with fatal outcomes clustering in regions where steep topography and high population density intersect (Casagli et al., 2010). Complementing fatality-focused evidence, catalog-based analyses of reported events have clarified that global inventories are operationally valuable because they organize spatiotemporal patterns of occurrence and allow systematic comparison of environmental drivers across regions. Within this context, "monitoring" denotes repeated observation of a slope system to detect kinematic or morphologic changes that signal evolving instability, while "early warning" denotes the structured transformation of monitoring signals into actionable alerts communicated through institutional protocols (Reichenbach et al., 2018). Territorial early warning systems for rainfall-induced landslides have been synthesized as socio-technical structures that combine setting definition, monitoring, modeling, and response, emphasizing that warning is a process rather than a single model output. Local early warning syntheses have similarly shown that warning performance is shaped by the design of monitoring networks, redundancy of sensors, and the operational treatment of thresholds and alarms. These established definitions and empirical findings motivate quantitative approaches that can connect measurable signals to decision variables in ways that are comparable across case-study settings and transferable across hazard governance contexts (Catani et al., 2013).

Remote sensing–based change detection provides an evidence pathway from repeated observations to quantified indicators of instability. In broad terms, change detection is the process of identifying differences in the state of an object or phenomenon by observing it at multiple times, and remote sensing enables this at spatial and temporal scales that are difficult to match with field-only programs (Crosetto et al., 2016). Reviews of change detection methods have described a progression from pixel-based spectral differencing toward object-based approaches that incorporate spatial context, texture, and segmentation, especially as data volumes and multi-sensor availability increase. For hazard applications, the relevance is direct: landslide activity often produces changes in surface roughness, vegetation condition, moisture state, and geometry that can be captured as time-indexed features and tested statistically against other indicators (Drusch et al., 2012). Broader syntheses of optical change detection have emphasized that technical quality depends on preprocessing, radiometric normalization, and explicit uncertainty handling in decision rules, because change/no-change classification is sensitive to sensor geometry, atmosphere, illumination, and surface moisture. These general principles become more actionable when integrated with platform capabilities (Froude & Petley, 2018). The Copernicus Sentinel program is explicitly designed to support routine Earth system observation, and Sentinel missions have been characterized as enabling continuous, standardized monitoring across domains relevant to environmental hazards. For landslide science, this means that change detection is not merely an image-processing task; it is an inference workflow in which multi-temporal signals are turned into variables for descriptive statistics, correlation analysis, and regression modeling. The reliability of that workflow depends on selecting signals that represent physically meaningful changes, documenting uncertainty, and designing the analysis so that statistical associations can be interpreted within a case-study narrative that preserves geomorphic context (Zhao & Lu, 2018).

Synthetic Aperture Radar (SAR) and interferometric SAR (InSAR) are central to modern landslide monitoring because they provide coherent measurements sensitive to ground deformation and can operate day-and-night under cloud cover. SAR is an active microwave system that records backscatter amplitude and phase, and InSAR exploits phase differences between acquisitions to infer line-of-sight displacement, enabling measurement of slow slope movements at millimeter-to-centimeter scales when coherence is adequate (Handwerger et al., 2020). The Sentinel-1 constellation has been described as providing routine C-band SAR observation with open access and operational products, and its mission paper explicitly notes suitability for interferometric applications such as detection of subsidence and

landslides. InSAR's role in slope instability has also been consolidated through critical reviews of multi-temporal interferometry, which detail how processing approaches such as persistent scatterer and small-baseline methods extend deformation monitoring across wide areas and long time series while requiring careful attention to geometry, coherence, and interpretation constraints (Rauf, 2018; Wasowski & Bovenga, 2014). Landslide-focused applications extend beyond satellite platforms; ground-based SAR interferometry has been used for real-time or near-real-time displacement tracking to support emergency management and early warning at slope scale, reinforcing that radar-based monitoring can be operationalized when viewing geometry and acquisition schedules align with the hazard setting (Haque & Arifur, 2021; Ashraful et al., 2020). Methodologically, these SAR-derived displacement products can serve as dependent variables in regression models or as independent predictors linked to perceived risk or institutional confidence in warning decisions (Intrieri et al., 2012; Fokhrul et al., 2021; Zaman et al., 2021). A quantitative design benefits from treating radar outputs as structured constructs, such as displacement rate, acceleration proxies, spatial clustering of moving targets, and temporal persistence, so that descriptive and inferential statistics can be applied consistently across respondent perceptions and remote sensing evidence (Fahimul, 2022; Hammad, 2022; Hussain et al., 2013).

Figure 1: Multi-Sensor Landslide Monitoring and Early Warning Framework
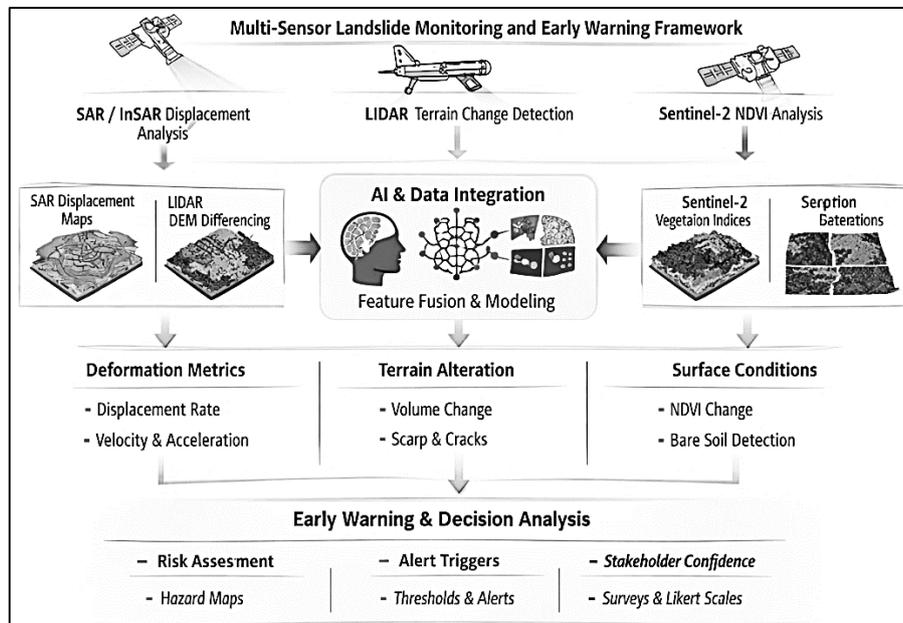
LiDAR (Light Detection and Ranging) contributes a complementary measurement perspective by directly supporting high-resolution topographic characterization and geometric change estimation. LiDAR is an active ranging technology that produces dense point clouds, from which digital elevation models (DEMs) and derivative terrain attributes can be calculated, enabling detection of geomorphic signatures such as scarps, tension cracks, hummocky deposits, and channel incision (Jaboyedoff et al., 2012). Reviews of LiDAR in landslide contexts have emphasized its capacity to reveal morphology under vegetation through point filtering and to support quantitative assessments of slope form and change when multi-temporal acquisitions are available (Hasan & Waladur, 2022; Rashid & Praveen, 2022; Piciullo et al., 2018). When LiDAR is repeated, DEM differencing becomes a change detection method that estimates elevation change, and the interpretation can be strengthened through explicit uncertainty modeling and feature extraction. A case-focused example has shown that probabilistic change detection fused with surface feature extraction on multi-temporal LiDAR-derived DEMs can map landslide activity and surface evolution at fine spatial scales (Arifur & Haque, 2022; Towhidul et al., 2022; Qin et al., 2018). These LiDAR-based change products are particularly useful for landslide monitoring and early warning systems because they offer a physically intuitive representation of terrain modification, which can be triangulated against deformation fields from SAR and surface-

condition indicators from optical imagery (Ratul & Subrato, 2022; Rifat & Jinnat, 2022; Torres et al., 2012). In a quantitative thesis structure, LiDAR-derived constructs can include volumetric change estimates, scarp edge migration measures, curvature/roughness indices, and spatial extent of detected change footprints. These constructs can be described statistically, correlated with radar deformation metrics, and incorporated in regression frameworks that test whether integrated multi-sensor indicators improve explanatory power for warning triggers and stakeholder confidence variables measured with Likert scales (Abdulla & Majumder, 2023; Rifat & Alam, 2022; Tofani et al., 2013).

Sentinel-2 optical imagery adds a third pillar by capturing multispectral reflectance sensitive to vegetation condition, soil exposure, and moisture-related spectral behavior, providing change signals that often accompany landsliding. Sentinel-2 has been defined as a multispectral mission delivering frequent observations at spatial resolutions suitable for environmental monitoring, and its mission design has been framed as enabling systematic land-surface applications through standardized products and revisit performance (Fahimul, 2023; Faysal & Bhuya, 2023; Kirschbaum et al., 2015). Optical change detection in landslide contexts often leverages shifts in vegetation indices, bare-soil exposure, and texture patterns that appear after failures or during progressive deformation that alters canopy structure and drainage. Methods for mapping landslides from optical time series have been operationalized in bitemporal workflows using Sentinel-2, where pre-event and post-event observations are transformed into discriminative features for landslide detection and inventory updating (Mora et al., 2018). At the method level, the broader change detection literature has shown that the quality of optical change inference depends on disciplined preprocessing and explicit evaluation of decision thresholds and classification error, because seasonal phenology and illumination variability can introduce confounding signals that mimic disturbance (Nackaerts et al., 2005). Integrating Sentinel-2 with LiDAR and SAR is therefore not only a matter of adding data layers; it is a matter of defining how optical indicators represent complementary dimensions of slope behavior. In a case-study approach, Sentinel-2 can contribute constructs such as normalized difference vegetation index change, normalized difference water index change, bare soil fraction proxies, or supervised change masks, each treated as measurable variables that can be linked statistically to both sensor-derived deformation and survey-measured stakeholder assessments of warning credibility (Pecoraro et al., 2019).

Artificial intelligence and data-driven modeling are increasingly used to operationalize multi-sensor change detection by learning complex relationships between input features and landslide-related outcomes. In remote sensing, AI-driven change detection often refers to the use of machine learning classifiers or regressors to infer change classes or change magnitudes from multi-temporal features, frequently combining spectral, textural, geometric, and contextual predictors (Petley, 2012). Reviews of landslide remote sensing have summarized how SAR, optical, and LiDAR measurements are applied across the landslide lifecycle—inventory mapping, deformation monitoring, trigger analysis, and mechanism inference—providing methodological justification for feature fusion approaches (Tanteri et al., 2017). The InSAR literature has similarly noted that multi-temporal interferometry products become more valuable when integrated with auxiliary information and interpreted through cross-comparisons and ground truth, aligning with AI workflows that treat deformation maps as structured predictors rather than standalone images (Tewkesbury et al., 2015). In landslide susceptibility and classification tasks, machine learning methods such as random forests have been applied at basin and regional scales, with studies demonstrating sensitivity to mapping unit scale, predictor resolution, and variable importance rankings—issues directly relevant to careful quantitative design. For hazard-focused monitoring, the technical pathway links AI-derived or AI-assisted change products to interpretable statistical tests: descriptive statistics summarize the distribution of detected change intensities, correlation matrices evaluate relationships among multi-sensor constructs, and regression models test hypotheses about which indicators predict warning triggers or confidence outcomes. This analytic structure also aligns with practical early warning reviews that treat warning as a modular system combining monitoring inputs and decision rules, which can be represented statistically through predictor–outcome relationships in a case-study dataset (Drusch et al., 2012).

Early warning reliability and trustworthiness in a multi-sensor, AI-driven monitoring framework depend on how well technical signals align with operational decision needs and stakeholder interpretations. Territorial and local early warning reviews have documented that warning systems vary in monitoring instruments, modeling approaches, dissemination procedures, and performance evaluation practices, and that effective warning depends on consistent integration of these components rather than isolated sensor performance (Intrieri et al., 2012). For landslide hazards, this integration challenge is intensified because different sensors "see" different aspects of slope behavior: SAR emphasizes deformation, LiDAR emphasizes geometry, and optical imagery emphasizes surface condition and disturbance signatures. The credibility of an early warning trigger can therefore be strengthened through triangulation among independent signal types, using explicit rules for agreement and conflict across sensors, which fits naturally into a quantitative case-study thesis that reports multi-sensor agreement/conflict maps and evaluates trigger reliability as measurable outcomes. The value of cross-platform triangulation is illustrated by studies that jointly analyze radar-derived displacement and LiDAR-derived thickness or geometry to infer mechanical properties of moving masses, showing that combined observation types can constrain interpretations of kinematics (Drusch et al., 2012). In applied settings, the same triangulation logic can be reframed as an evidence structure for warning governance: change detection outputs become "evidence objects" that can be compared to stakeholder confidence ratings and institutional readiness measures captured through Likert-scale instruments. The introduction of such constructs does not require speculative claims; it requires a disciplined framing in which measurement definitions, sensor capabilities, and established early warning system components are connected through testable statistical relationships within the limits of a cross-sectional case-study dataset (Qin et al., 2018).

The purpose of this study is to examine how an AI-driven change detection framework that integrates SAR, LiDAR, and Sentinel-2 data can support landslide monitoring and strengthen disaster early warning decision processes within a quantitative, case-study context. The first objective is to operationalize multi-sensor change evidence by defining measurable indicators that represent deformation-related signals from SAR, terrain-structure and morphology change signals from LiDAR, and surface-condition or disturbance signals from Sentinel-2, then organizing these indicators into a consistent measurement structure that can be compared across the case study domain. The second objective is to quantify the perceived performance of the proposed AI-driven change detection outputs among stakeholders who are responsible for hazard monitoring, geospatial analysis, emergency coordination, and risk communication, using a structured Likert five-point instrument that captures constructs such as perceived accuracy, timeliness, interpretability, trust in alerts, decision confidence, and operational readiness. The third objective is to statistically test the relationships between multi-sensor change indicators and early warning effectiveness by applying descriptive statistics to summarize construct distributions, correlation analysis to identify the strength and direction of associations among the key variables, and regression modeling to determine which indicators significantly predict early warning effectiveness within the selected case study setting. A fourth objective is to evaluate the internal consistency and measurement stability of the survey constructs through reliability assessment, ensuring that each construct forms a coherent scale suitable for inferential analysis. A fifth objective is to examine the added value of sensor fusion by comparing models that rely on single-sensor indicators against models that integrate combined multi-sensor indicators, with the intention of determining whether multi-source evidence provides stronger explanatory power for early warning effectiveness and stakeholder confidence than any individual data stream alone. A final objective is to present case-specific change detection outputs in an interpretable, auditable manner by reporting multi-sensor agreement and conflict patterns, assessing the stability of warning triggers under varying decision thresholds, and aligning remote sensing signal strength with stakeholder confidence measures, thereby ensuring that the statistical findings and the case-study outputs remain mutually consistent and traceable to the underlying measurement definitions used throughout the study.

LITERATURE REVIEW

The literature on AI-driven landslide monitoring and disaster early warning increasingly converges around three complementary needs: reliable sensing of slope processes, robust change detection methods that can separate meaningful hazard signals from environmental variability, and decision-oriented frameworks that translate technical outputs into actionable warnings. Landslides are dynamic, multi-causal phenomena shaped by terrain geometry, material properties, hydrologic forcing, and human modification of slopes, which makes single-source monitoring inherently limited when applied to diverse geographic and climatic settings. For this reason, the research community has progressively emphasized multi-sensor strategies that combine deformation-sensitive observations, terrain-structure representation, and surface-condition indicators to produce a more complete depiction of slope instability. Within this body of work, SAR and InSAR approaches are widely discussed for their capacity to measure subtle ground motion and identify deforming zones under cloud cover, while LiDAR-based topography and multi-temporal DEM analysis are treated as high-resolution evidence for morphological change, scarp development, and volumetric variation that can validate or contextualize kinematic signals. In parallel, optical multispectral imagery, including platforms such as Sentinel-2, is frequently used to capture changes in vegetation, bare soil exposure, moisture-related spectral behavior, and disturbance patterns that often accompany landslide initiation or progression. These sensing streams feed into change detection workflows that range from traditional thresholding and differencing to object-based analysis and machine learning classifiers, with recent work increasingly integrating data fusion and AI methods to handle nonlinear relationships and high-dimensional feature spaces. At the same time, early warning research stresses that technical detection performance alone is insufficient for operational reliability; warning systems also require transparent thresholds, redundancy, stability under uncertainty, and user trust in the interpretability of alerts. Consequently, the literature review for this study synthesizes work across remote sensing, geomorphology, and disaster risk management to establish how multi-sensor change evidence can be structured into measurable indicators, how AI approaches can enhance detection and reduce false alarms when properly validated, and how early warning effectiveness can be evaluated through quantitative constructs linked to decision confidence and operational readiness in a case-study setting.

Landslide Early Warning Systems and Operational Monitoring Logic

Landslide monitoring and disaster early warning systems (EWS) are organized arrangements that convert observations of slope conditions and triggering factors into alerts that support protective action at community, infrastructure, or regional scales. In landslide risk management, monitoring includes repeated measurement of environmental drivers such as rainfall intensity, rainfall duration, and antecedent wetness, plus observation of slope responses such as displacement, cracking, seepage, and channel change. These observations are collected through instruments, surveys, remote sensing products, and institutional reporting channels, and they are curated into datasets that can be checked, summarized, and compared over time. At territorial scale, EWS design commonly begins with the definition of warning zones and the selection of predictor variables that represent triggering conditions, because warning levels must be issued for administrative units that map to operational responsibilities. Empirical rainfall thresholds are frequently used as a first layer of decision support because they can be calibrated from historical events and translated into intuitive rules for alert activation. A widely cited synthesis of European evidence shows that the minimum rainfall intensity required to initiate slope failures decreases systematically with increasing duration, supporting the practical use of intensity-duration curves for warning operations when local calibration is limited (Guzzetti et al., 2007). In this framing, monitoring is judged not only by measurement precision, but also by coverage, latency, and consistency of acquisition, because the warning task depends on whether the system can detect meaningful change early enough to support decisions. The literature therefore treats EWS as a chain of components that links hazard knowledge, monitoring networks, data processing, and communication protocols into a single operational process. Because landslides vary by mechanism and speed, monitoring plans specify which failure types are targeted, what lead time is expected, and which indicators are treated as decision-grade for escalation, de-escalation, and formal event documentation. Operational landslide early warning requires that monitoring variables be combined through explicit decision procedures that manage uncertainty and balance false alarms against missed events. In

practice, warning models integrate multiple information streams such as rainfall observations and forecasts, antecedent moisture proxies, slope movement indicators, and situational reports into tiered alert schemes that communicate escalating concern while preserving interpretability for nontechnical users. A well-documented example from the United States demonstrates how cumulative rainfall thresholds and intensity-duration thresholds can be paired with real-time soil moisture monitoring and translated into a four-level warning scheme (Baum & Godt, 2010). This kind of architecture clarifies that EWS performance depends on both the scientific adequacy of predictors and the operational design of warning levels, including the rules used to activate, update, and cancel alerts (Habibullah & Aditya, 2023; Hammad & Mohiul, 2023). For territorial systems, the warning model must also address spatial aggregation, because multiple landslides can occur across a warning zone within a single event, and managers often need a graded representation of expected landslide density rather than a binary forecast (Haque & Arifur, 2023; Jahangir & Mohiul, 2023). This motivates evaluation methods that treat warnings and landslide occurrences as time-aligned sequences with durations and levels, rather than as isolated hits or misses. A regional-scale performance framework based on an event-duration matrix formalizes this logic by quantifying how long each warning level is active relative to the timing of landslide events, enabling indicators that distinguish late warnings, over-warning, and missed events across levels (Calvello & Piciullo, 2016). These approaches support transparent comparison among alternative thresholds, zoning schemes, and escalation rules, and they provide a basis for communicating model trade-offs to decision makers who must justify actions taken under uncertainty. In modern implementations, remote sensing change indicators can be incorporated as supplementary evidence, but they must be synchronized to the alert timetable, quality-checked, and summarized into variables aligned with warning levels.

Figure 2: Landslide Early Warning Systems and Operational Monitoring Logic
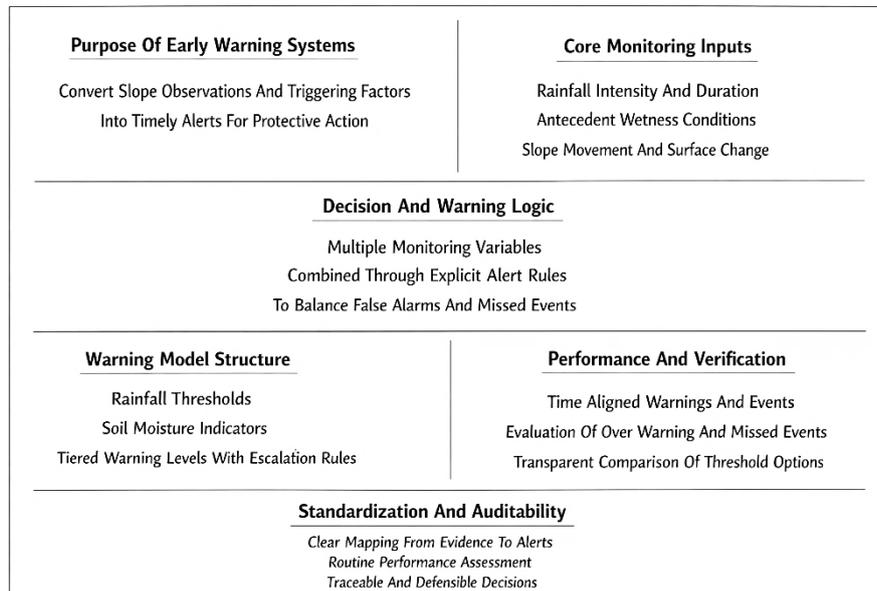
A recurring theme in landslide EWS research is the need for standardization and auditability, because stakeholders must understand what evidence is used, how it is transformed into warning levels, and how performance is verified over time (Rashid et al., 2023; Khaled & Mosheur, 2023). Standard oriented contributions frame landslide EWS as systems composed of interlocking subsystems, including risk assessment and mapping, monitoring and forecasting services, dissemination and communication, preparedness organization, and sustained commitment to operation and maintenance (Mostafa, 2023; Rifat & Rebeka, 2023). A structured methodology for defining such a standard emphasizes that early warning practice includes institutional roles, operating procedures, and community readiness elements alongside technical monitoring, and it proposes a unified set of subsystems that can be used as a checklist for system design and review (Fathani et al., 2016; Azam & Amin, 2023). Auditability is

strengthened when the system defines measurable performance targets and routinely tests whether the warning model meets them under realistic operating conditions. Recent work on territorial landslide EWS performance assessment highlights that commonly used contingency metrics can conceal important information about warning level detail and the varying number of landslides within a warning interval, and it proposes standardized indicators and tools that preserve warning level structure in verification (Piciullo et al., 2020). Together, these perspectives encourage reporting practices that document the mapping between monitoring inputs, intermediate products, and issued alerts, so that users can trace a warning back to the evidence that triggered it and evaluate whether the evidence was timely and consistent. For quantitative case studies, this supports the construction of datasets that combine hazard observations, warning records, and stakeholder evaluations in a format that allows descriptive summaries, correlation analyses, and regression modeling. It also motivates the explicit definition of units of analysis, temporal windows, and decision thresholds, because these choices shape measured performance and influence how system credibility is interpreted by those who rely on warnings in practice today.

SAR- and InSAR-Based Change Detection for Landslide Monitoring

Synthetic Aperture Radar (SAR) is widely used in landslide monitoring because it measures microwave backscatter day or night and is far less constrained by cloud cover than optical sensors, which makes it reliable for steep, humid, and frequently overcast regions where slope failures can produce high losses. In change detection, SAR contributes two complementary evidence types. The first is amplitude-based change, where shifts in backscatter intensity can indicate altered surface roughness, moisture conditions, or vegetation structure that may accompany slope disturbance, erosion, deposition, or exposure of fresh material. This information is valuable for screening and mapping because it can highlight newly disturbed zones and can capture changes over wide areas quickly. The second evidence type is coherence-related change, where reductions in interferometric coherence can signal surface disturbance, vegetation dynamics, or rapid deformation that disrupts phase stability between acquisitions. Although coherence loss does not uniquely identify landslides, it can function as an alerting feature when interpreted alongside terrain context and when constrained by spatial patterns typical of slope movement. A central strength of SAR for landslides is its capacity to support both regional reconnaissance and targeted site investigation, yet its use also requires careful attention to geometric distortions (layover, foreshortening, and shadow) that are common in mountainous terrain and can create spatial gaps or misinterpretation if not explicitly handled. The literature emphasizes that SAR-based landslide assessment becomes more robust when SAR change indicators are treated as measurable variables with known limitations, then combined with ancillary constraints such as slope angle, aspect, lithology, and infrastructure exposure so that change signals are interpreted within geomorphic plausibility rather than as isolated image differences (Colesanti & Wasowski, 2006).

Interferometric SAR (InSAR) extends SAR change detection by converting phase differences between repeat acquisitions into deformation estimates, which are directly relevant for landslide monitoring because progressive slope instability often manifests as slow displacement prior to more disruptive movement. However, reliable change detection with InSAR depends on multi-temporal processing that separates true ground motion from noise sources that vary across space and time, including decorrelation, orbital residuals, and unmodeled atmospheric delay. Multi-temporal methods increase the number of locations where deformation can be estimated by exploiting many acquisitions, and they stabilize time series by using consistent statistical criteria for selecting measurement points and estimating deformation models. A key methodological insight is that real landscapes rarely match the ideal assumptions of any single scattering model, so robust approaches aim to combine the strengths of persistent-scatterer and small-baseline strategies to improve spatial coverage while controlling phase noise. This is important for landslide settings because coherent targets may cluster along roads, rock outcrops, or built infrastructure, while vegetated slopes may still yield useful information through short-baseline interferograms and careful filtering. Method selection also influences the statistical behavior of the derived deformation variables: for example, time-series smoothing, outlier handling, and reference-point choices can change apparent acceleration patterns that are critical for warning logic. Comparative methodological analyses show that different selection criteria and deformation modeling assumptions can yield different densities of measurement points and different deformation

estimates even on the same dataset, which matters when deformation products are used as predictors in regression models or as evidence feeding early warning thresholds (Hooper, 2008; Sousa et al., 2011). For quantitative case studies, this literature supports documenting processing choices as part of measurement validity, because the interpretability of correlations and regression coefficients depends on the stability and comparability of the deformation time series.

Figure 3: SAR And InSAR Based Change Detection for Landslide Monitoring
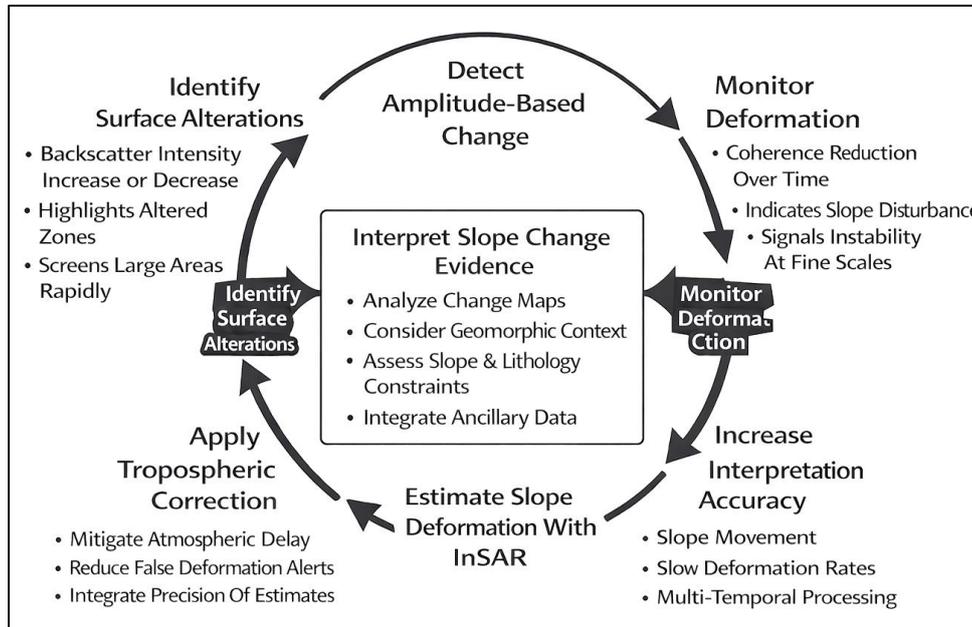
Operational landslide monitoring with InSAR must also manage atmospheric effects, which can imprint spatially correlated phase signals that resemble deformation and can inflate false positives in change detection if left uncorrected. Tropospheric variability is especially important in mountainous regions where elevation gradients and weather patterns produce strong, acquisition-dependent delays. Comparative evaluations of tropospheric correction strategies demonstrate that different auxiliary data sources (such as spectrometer products and numerical weather models) and empirical phase–topography relationships can reduce atmospheric contamination to different degrees, and that the most appropriate correction can depend on data availability, cloud conditions, and the local atmospheric regime. This matters for landslide monitoring because an apparent spatial gradient in phase can be misread as slope movement when it is actually atmospheric structure, and that misinterpretation can propagate into alerting rules, agreement maps, and stakeholder confidence metrics. The literature therefore supports treating tropospheric correction and residual atmospheric uncertainty as part of the quality assurance workflow, particularly when change detection outputs are intended to be used in decision-support contexts. At the same time, the field has increasingly moved toward service-oriented monitoring pipelines that deliver continuously updated deformation products from routine SAR acquisitions, enabling repeated screening for accelerating movement and systematic prioritization of unstable slopes. Demonstrations of continuous or semi-automatic processing chains show that operational constraints—such as update frequency, computational efficiency, and consistent reporting formats—shape how InSAR products are transformed into actionable change indicators for monitoring programs (Bekaert et al., 2015; Raspini et al., 2018). Within this study, that operational perspective motivates the structured definition of SAR-derived variables, transparent reporting of processing choices, and integration with complementary sensors so that change evidence can be evaluated statistically and interpreted consistently in the case-study environment.

LiDAR-Based Terrain Change Detection

Airborne LiDAR and terrestrial laser scanning have become core technologies for landslide monitoring because they provide dense three-dimensional measurements that can be transformed into high-resolution bare-earth digital terrain models (DTMs) and derivative geomorphometric layers, supporting the identification of landslide signatures that are often muted or obscured in optical imagery. In operational terms, LiDAR contributes to landslide analysis in two complementary ways: (a) inventory enhancement, where diagnostic landforms such as main scarps, tension cracks, lateral flanks, toe bulges, hummocky topography, and disrupted drainage are mapped more completely and consistently; and (b) change detection, where repeated surveys quantify surface elevation differences that reveal depletion and accumulation zones. A key advantage is that LiDAR-based DTMs can reduce vegetation occlusion through point-cloud filtering, thereby exposing micro-topographic features that define landslide boundaries and internal structure with clearer geometric evidence than many alternative datasets. Early demonstrations showed that airborne LiDAR can support semi-automated recognition and mapping of rainfall-induced landslides by leveraging elevation-derived geomorphic expression rather than relying exclusively on spectral contrast, which is particularly valuable in environments where fresh scarps quickly revegetate or where lighting and seasonal variability complicate optical interpretation (Ardizzone et al., 2007). At broader mapping scales, LiDAR-driven workflows also enable systematic, GIS-integrated delineation and classification of slope failures across large areas, including the separation of natural landslides from anthropogenic slope cuts when combined with orthophotography and contextual rules; this integration is important because inventories are not only descriptive products, but also foundational datasets for validation, hazard zoning, and model evaluation in early warning and risk governance (Palenzuela et al., 2015). From a methodological perspective, LiDAR's contribution is strengthened when DTMs are accompanied by explicit metadata on point density, filtering choices, interpolation methods, and vertical accuracy, because these parameters shape both feature visibility and the uncertainty bounds required for defensible change detection.

Multi-temporal LiDAR change detection is often implemented through DEM-of-Difference (DoD) analysis, where co-registered elevation surfaces from different dates are subtracted to quantify terrain change. For landslides, DoD values can represent depletion in source areas, deposition at toes, and internal redistribution in earthflows or rotational slides where the moving mass deforms and thickens rather than simply translating downslope. The credibility of DoD-based inference depends on careful alignment of surfaces, consistent vertical datums, and robust treatment of uncertainty, because even small horizontal misregistration or systematic elevation bias can create spurious "change" patterns that resemble landslide morphology. This is why many LiDAR change-detection studies emphasize preprocessing steps such as strip adjustment, ground classification review, and error-aware thresholding before interpreting DoD maps as evidence of slope activity. In kinematic landslide settings, multi-temporal LiDAR has been combined with complementary techniques—such as DEM subtraction and image-based correlation—to resolve complex movement fields, revealing that deformation can be heterogeneous across a single earthflow and that distinct subdomains can accelerate, stall, or rotate in different ways through time (Daehne & Corsini, 2013). In addition to spatial characterization, multi-temporal LiDAR supports volumetric estimation, allowing researchers to estimate landslide-produced sediment volumes and relate volume–area scaling to catchment sediment budgets and debris-flow hazard considerations. A representative application used LiDAR-derived DTMs from different years to estimate sediment volumes generated by a single extreme event, illustrating how DoD-based volume calculations can be integrated with mapped landslide footprints and then summarized statistically for regional risk analysis (Tseng et al., 2013). For quantitative designs that link remote sensing outputs to decision-making constructs, these LiDAR-based products can be operationalized as measurable variables—e.g., area of significant elevation loss, maximum scarp retreat distance, net volume change, and spatial concentration of deformation-like morphology—making them suitable for descriptive statistics, correlation analysis, and regression modeling when paired with stakeholder-rated measures of warning usefulness and confidence.

LiDAR-based monitoring is further strengthened by terrestrial laser scanning (TLS), which supports frequent site-scale surveys and produces dense point clouds that can capture small displacements, localized scarp evolution, and short-term morphological change on active slopes. TLS is particularly relevant where satellite revisit frequency, canopy effects, or terrain geometry limit the completeness of spaceborne observations, and where early warning decisions require localized, rapidly updated evidence. However, TLS introduces its own technical constraints: stable reference networks must be established, scan positions must be chosen to minimize occlusions, and point-cloud registration must be rigorously controlled to ensure that detected differences represent real movement rather than instrument noise or alignment error. Research evaluating TLS for slow-moving landslides has shown that long-term reliability hinges on registration quality and comparative validation against independent surveying methods, because sub-decimeter motion targets can be difficult to distinguish from cumulative measurement and processing error without careful design (Prokop & Panholzer, 2009).

**Figure 4: LiDAR Based Terrain Change Detection for Landslide Monitoring**
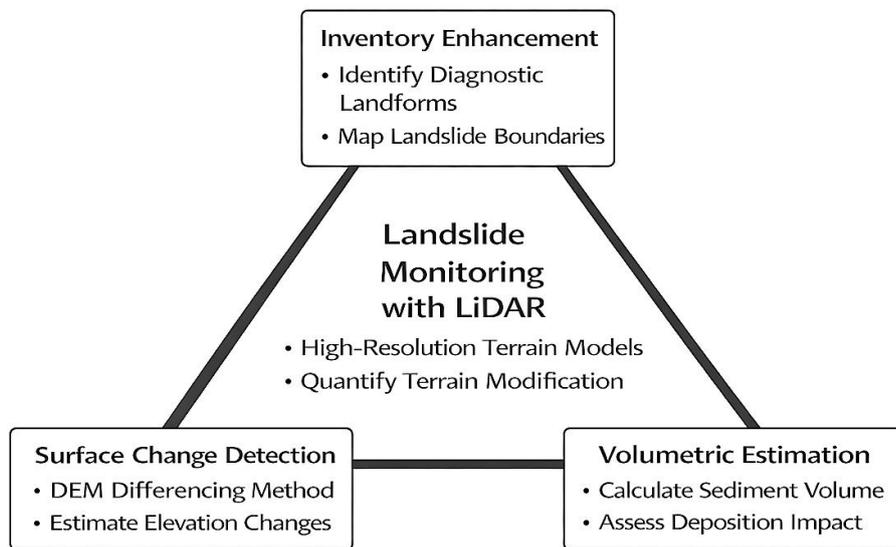
These insights directly inform multi-sensor landslide monitoring architectures: LiDAR (airborne or terrestrial) can serve as a geometric "truth layer" for slope morphology and elevation change, against which SAR-derived deformation patterns and Sentinel-2 surface-condition indicators can be triangulated. In practice, LiDAR can also strengthen system trustworthiness by enabling interpretable, auditable outputs—such as mapped scarp expansion, quantified elevation-loss zones, and net volume change—that can be explained to non-specialist stakeholders and compared transparently across time windows. When integrated into AI-driven change detection, LiDAR-derived geomorphometry provides physically meaningful features for classification or regression models, while also supporting post hoc validation and conflict analysis in multi-sensor agreement maps. As a result, LiDAR is not only a high-resolution mapping tool, but also a key component for constructing defensible evidence chains from measured terrain change to monitoring interpretation and early warning decision support.

Sentinel-2 Optical Change Detection for Surface Disturbance

Sentinel-2 optical imagery supports landslide monitoring by providing multispectral reflectance measurements that are sensitive to vegetation condition, bare-soil exposure, moisture-related spectral behavior, and surface disturbance patterns that frequently accompany slope failure. In optical change detection, "change" is typically operationalized as a measurable shift in spectral response or in derived indices between pre-event and post-event observations, which allows disturbed ground to be separated from stable land cover when the observation dates are selected carefully. Because landslides often strip vegetation, expose fresh soil and rock, and modify drainage pathways, bitemporal Sentinel-2 comparisons can highlight scar boundaries and runout paths as coherent spatial objects rather than isolated anomalous pixels. Object-oriented approaches explicitly exploit this by combining spectral

features with spatial context—shape, texture, and neighborhood relationships—to increase mapping reliability in complex terrain where shadows and mixed pixels are common. Landslide-focused work has shown that semi-automatic detection benefits from characterizing the spectral, spatial, and morphometric properties of landslides as segmented objects, enabling decision rules or classifiers to distinguish landslide bodies from spectrally similar non-landslide features such as river bars or construction sites (Martha et al., 2010). In addition, optical change detection is valuable for producing inventories rapidly after triggering events because it can cover large areas at moderate resolution and can be processed with reproducible workflows for operational reporting. In this study context, Sentinel-2 indicators are relevant because they can be formalized into consistent constructs—such as vegetation index change, soil exposure proxies, or multi-band change vectors—that align with quantitative analysis and can be compared with deformation-sensitive radar measures and terrain-change evidence from LiDAR. The literature therefore positions Sentinel-2 as a practical source for disturbance mapping and temporal pattern analysis, particularly when change indicators are extracted consistently, quality-checked for clouds and haze, and interpreted through geomorphic plausibility rather than through spectral contrast alone.

Figure 5: Sentinel 2 Optical Change Detection for Surface Disturbance and Rapid Landslide Mapping
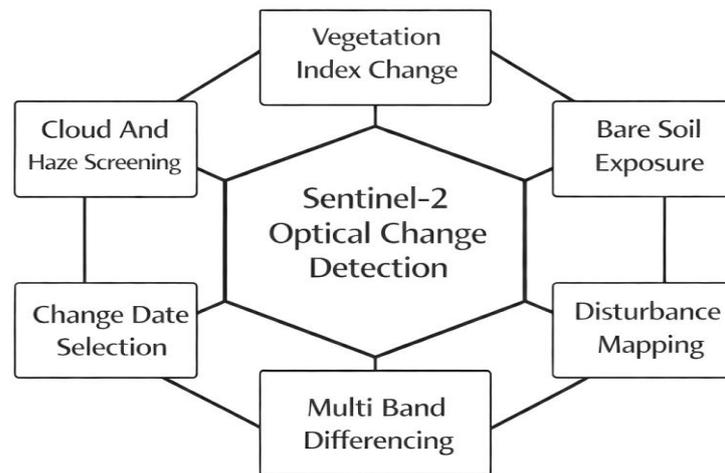
A central methodological challenge in optical landslide change detection is separating true landslide disturbance from confounding variability introduced by seasonality, phenology, illumination differences, and atmospheric effects. For this reason, many studies rely on index-based and multi-index strategies that reduce sensitivity to absolute radiance levels and emphasize relative differences that are more stable across acquisition conditions. NDVI-based differencing is common because vegetation removal and stress frequently accompany landslides, making NDVI shifts informative for scar delineation, edge expansion tracking, and short-term evolution monitoring when time-series images are available. Time-series Sentinel-2 analysis has demonstrated that NDVI and true-color sequences can be used to partition post-failure behavior into interpretable change stages and to quantify scar expansion through pixel- or object-based counts, reinforcing the feasibility of relatively simple indicators for documenting landslide evolution when high-frequency field surveys are not available (Qu et al., 2021). At event scale, optical workflows may also incorporate dimensionality-reduction or contextual classification to stabilize change products under heterogeneous illumination and atmospheric differences. A representative example improved change-detection-based mapping by integrating NDVI change imagery with additional statistical transforms and contextual constraints, producing inventories that were more robust when pre- and post-event images differed in phenological or atmospheric conditions (Lu et al., 2019). This body of work supports the treatment of optical change detection as an evidence-building process in which multiple indicators converge on the same disturbed footprint, rather than as a single-index threshold exercise. In quantitative case studies, this logic

motivates documenting the selected indices, the date-selection rule, the cloud-masking procedure, and the mapping unit (pixel or object), because these decisions determine the statistical behavior of the derived Sentinel-2 constructs and their compatibility with correlation and regression models that aim to explain early warning effectiveness.

Sentinel-2 becomes especially powerful for landslide monitoring when it is fused with complementary observations, because optical change signals often represent surface disturbance while radar emphasizes deformation and LiDAR emphasizes terrain structure. Multi-sensor fusion can reduce ambiguity by requiring agreement among independent evidence types or by weighting contributions according to expected sensitivity in a given slope environment. Research combining Sentinel-1 and Sentinel-2 in object-oriented machine learning has shown that joint optical–radar feature spaces can improve landslide mapping and discrimination in forested settings, where optical signatures may be partially obscured and radar coherence changes can provide additional cues about instability (Shirvani et al., 2019). Optical-only studies also emphasize that mapping reliability improves when contextual information is used to constrain candidate scars to plausible topographic conditions, which helps avoid false positives from agricultural harvest cycles, river sediment reworking, or built-environment changes. The operational relevance is that early warning and monitoring programs require outputs that remain stable under routine acquisition variability and that can be audited and explained, particularly when outputs are intended to support alerts rather than retrospective mapping. In this respect, object-based optical landslide mapping and change detection have been positioned as practical alternatives to purely manual inventorying because they can accelerate mapping and improve repeatability while still allowing expert review of boundaries and uncertainty (Mondini et al., 2011). For the present research, this literature supports defining Sentinel-2 change constructs that are both technically grounded and decision-relevant—such as disturbance intensity classes, scar edge change rates, and spatial concentration of high-confidence change objects—and integrating them with SAR and LiDAR evidence to generate agreement/conflict patterns that can be compared against stakeholder confidence measures. This creates a structured pathway from optical signals to interpretable multi-sensor evidence, enabling statistical testing of whether fused change indicators better align with perceived early warning effectiveness than single-sensor indicators alone.

Socio-Technical Acceptance for AI-Driven Multi-Sensor Early Warning

An AI-driven landslide monitoring and early warning system is best understood as a socio-technical capability that couples measurement infrastructure with organisational routines for interpretation and action. This framing is necessary because the operational meaning of "warning" emerges from a chain that includes data acquisition (SAR, LiDAR, Sentinel-2), analytics (AI change detection and fusion), and human decision practices (verification, escalation, communication, and response). Socio-technical systems engineering highlights that system performance cannot be judged only by algorithmic outputs; it is also shaped by whether the system fits real work, whether evidence is legible to stakeholders, and whether responsibilities and workflows are aligned with the timing and uncertainty characteristics of hazard signals. In this view, technical artefacts such as deformation maps or disturbance masks become "work objects" that must be interpreted, shared, and defended under institutional constraints. Baxter and Sommerville (2011) emphasize that socio-technical approaches aim to integrate technical and organisational considerations across the full system life cycle, because many systems fail by meeting technical requirements while still not supporting the actual work done by users. For landslide early warning, this translates into an explicit need for auditable evidence chains, stable reporting formats, and agreed decision rules that connect sensor-derived change to warning levels. Accordingly, the theoretical foundation of this study treats AI change detection as one component embedded within monitoring services, stakeholder roles, and decision accountability. This is consistent with the study's quantitative, cross-sectional, case-study design, where stakeholder evaluations (via Likert-scale constructs) are analysed alongside measurable multi-sensor change indicators. The socio-technical lens therefore motivates measuring not only technical evidence intensity, but also interpretability, trust, and operational readiness as distinct constructs, because these human-system factors directly influence whether monitoring outputs are accepted as decision-grade information in real early warning practice. To connect socio-technical fit to measurable adoption outcomes, this study incorporates technology acceptance logic that explains why users choose to rely on a system's outputs.

Technology Acceptance Model 3 (TAM3) positions perceived usefulness and perceived ease of use as central determinants of intention and subsequent system use, while also detailing how experience, job relevance, output quality perceptions, and related beliefs shape these determinants (Petter et al., 2008; Venkatesh & Bala, 2008). In an AI-driven landslide monitoring context, perceived usefulness aligns with whether stakeholders believe multi-sensor change detection improves warning decisions (timeliness, confidence, and actionability), while perceived ease of use aligns with whether outputs are understandable and operationally manageable (clarity of maps, explainability of alerts, and reasonable workload). Because this research also involves professional decision settings rather than consumer-only contexts, it is important to capture additional determinants that affect sustained use under routine monitoring. UTAUT2 provides a complementary extension by modelling behavioral intention and use as functions of performance expectancy, effort expectancy, social influence, and facilitating conditions, with additional constructs to reflect real-world usage patterns (Venkatesh et al., 2012).

Figure 6: Socio Technical Acceptance for AI Driven Multi Sensor Early Warning
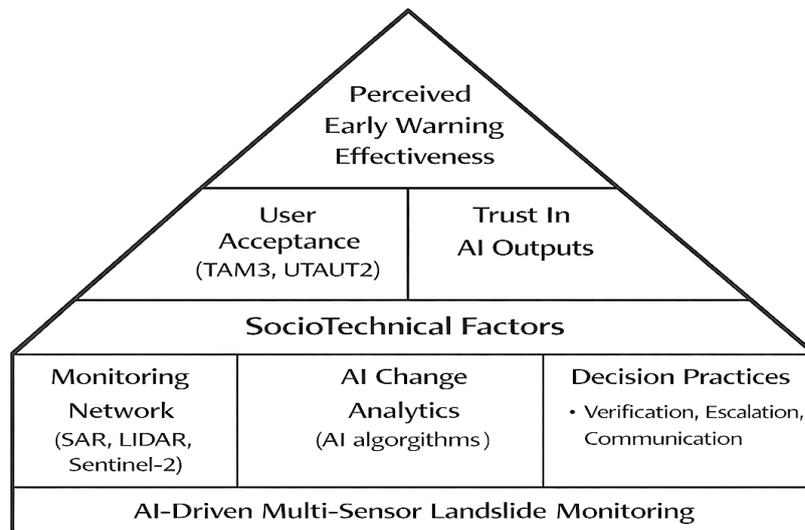
For this study, "facilitating conditions" is particularly relevant because an early warning system's usefulness is constrained by data access, computing resources, and organizational procedures for acting on alerts. These acceptance frameworks support the study's inferential plan by allowing Likert-scale constructs (e.g., perceived AI performance, interpretability, and adoption readiness) to be tested as predictors of perceived early warning effectiveness using regression. They also justify modelling trust as more than a general attitude: trust is treated as a measurable belief about the appropriateness of relying on system outputs under uncertainty, which is essential when multi-sensor indicators may sometimes conflict due to viewing geometry, vegetation effects, or surface moisture variability (Baxter & Sommerville, 2011).

A trust-calibrated theoretical framing is required because early warning decisions are safety-relevant and must balance false alarms against missed events. Trust in automation research synthesizes trust as a layered construct shaped by individual disposition, situational context, and learned experience from interacting with the system, emphasizing that trust should be calibrated to system reliability rather than maximized indiscriminately (Hoff & Bashir, 2015). This is directly applicable to AI-driven change detection because stakeholders must decide when to rely on a "high-change" alert and when to request additional verification, especially in mixed-signal areas where SAR, LiDAR, and Sentinel-2 may not agree. To operationalize this calibration within the quantitative design, this study applies a single, study-wide fusion formula that produces an auditable evidence score from the three sensor streams. The Multi-Sensor Change Evidence Score (MSCES) is defined as:

$$\text{MSCES} = w_1 Z_{\text{SAR}} + w_2 Z_{\text{LiDAR}} + w_3 Z_{\text{S2}}, \text{where } w_1 + w_2 + w_3 = 1$$

Here, each Z is a standardized change indicator (e.g., SAR deformation/coherence-change index, LiDAR DEM-of-difference magnitude or terrain-change index, Sentinel-2 disturbance/vegetation-

change index). This single formula is the best "whole-study" expression because it links the remote sensing core (multi-sensor change detection) to the statistical core (correlation and regression) while keeping the fusion step transparent and repeatable. The resulting MSCES can be used in the main explanatory model of perceived early warning effectiveness:

$$\text{EWE} = \beta_0 + \beta_1 \text{MSCES} + \beta_2 \text{PAI} + \beta_3 \text{TRUST} + \varepsilon$$

where PAI is perceived AI performance and TRUST is calibrated reliance. This formulation aligns with an information-systems success perspective that defines system value through quality, use, satisfaction, and net benefits (Petter et al., 2008), allowing the study to interpret results as an integrated relationship between evidence quality, stakeholder acceptance, and perceived warning effectiveness.

Conceptual Framework for this study

The conceptual framework for this study positions AI-driven change detection as a measurable, auditable pathway that connects multi-sensor Earth observation evidence (SAR, LiDAR, Sentinel-2) to the practical outcomes expected from landslide monitoring and early warning. Conceptually, the model has three layers. (1) Evidence layer: each sensor contributes a different "type" of landslide-relevant signal—SAR captures deformation/coherence behavior, LiDAR captures terrain and elevation-change signatures, and Sentinel-2 captures surface disturbance and vegetation-condition shifts. (2) Analytics layer: AI routines extract and structure these signals into standardized indicators (per spatial unit and time window), then fuse them into interpretable outputs such as "agreement/conflict" maps and warning-oriented evidence scores. (3) Decision layer: stakeholders interpret the AI outputs within operational constraints (time, uncertainty, institutional procedure), forming judgments about reliability, usefulness, and readiness that can be measured with Likert-scale constructs. This framing aligns with applied landslide mapping research showing that object-oriented and machine-learning approaches improve the consistency and repeatability of landslide delineation when spectral and contextual cues are integrated systematically (Stumpf & Kerle, 2011). It is also consistent with multi-source landslide workflows that combine Sentinel-1 and Sentinel-2 products with external validation sources (e.g., GNSS or field points) to reduce uncertainty and produce more defensible hazard-related products (Ghorbanzadeh et al., 2020). In this study, the unit of analysis can be a grid cell or object segment (depending on the case study), and the conceptual framework requires that every downstream metric (statistical tests, maps, trigger reliability summaries) can be traced back to a clearly defined indicator from the evidence layer. This traceability is essential because multi-sensor data can conflict due to geometry, vegetation, moisture, and acquisition timing, so the framework explicitly treats "conflict" as information rather than as error.

To make fusion statistically comparable and suitable for correlation and regression modeling, the framework standardizes each sensor's indicator and then combines them into a single study-wide evidence score. For each spatial unit i, the standardized indicator is computed as a z-score:

$$z_{k,i} = \frac{x_{k,i} - \mu_k}{\sigma_k}$$

where $k \in \{\text{SAR,LiDAR,S2}\}$, $x_{k,i}$ is the raw change metric, and μ_k and σ_k are the mean and standard deviation computed over the case-study domain (or a defined baseline window). The Multi-Sensor Change Evidence Score (MSCES) is then defined as:

$$\text{MSCES}_i = w_{\text{SAR}} z_{\text{SAR},i} + w_{\text{LiDAR}} z_{\text{LiDAR},i} + w_{\text{S2}} z_{\text{S2},i}, w_{\text{SAR}} + w_{\text{LiDAR}} + w_{\text{S2}} = 1$$

This single formula is used throughout the study because it supports transparent fusion (weights are explicit), enables direct statistical testing, and supports map-based reporting (MSCES can be visualized continuously or thresholder into warning-oriented classes). When sensor-specific outputs are produced by different change detection methods (e.g., alternative SAR metrics or alternative Sentinel-2 indices), the framework supports evidence-level combination using uncertainty-aware logic inspired by evidence-theory fusion, which is commonly used to reduce conflict and model uncertainty when multiple candidates change maps disagree (Luo et al., 2018).

Additionally, the analytics layer may include an AI classifier that predicts "landslide change" or "no change" using stacked multi-sensor features; such approaches have been shown to be effective for

landslide detection when feature selection, training samples, and evaluation are handled rigorously (Ghorbanzadeh et al., 2019). In this study, MSCES functions as the principal fused indicator for both map products and the inferential models (correlation/regression), while AI outputs (e.g., probability of change, class labels) can be compared against MSCES to triangulate robustness and interpretability. At the decision and evaluation layer, the framework links remote-sensing evidence to stakeholder-rated constructs that represent early warning effectiveness in practice. The dependent construct, Early Warning Effectiveness (EWE), is computed as a composite of Likert-scale items capturing alert reliability, decision confidence, timeliness, and operational usefulness; explanatory constructs include Perceived AI Performance (PAI), Interpretability (INT), and Trust in Alerts (TRU).

Figure 7: Conceptual Framework for this study
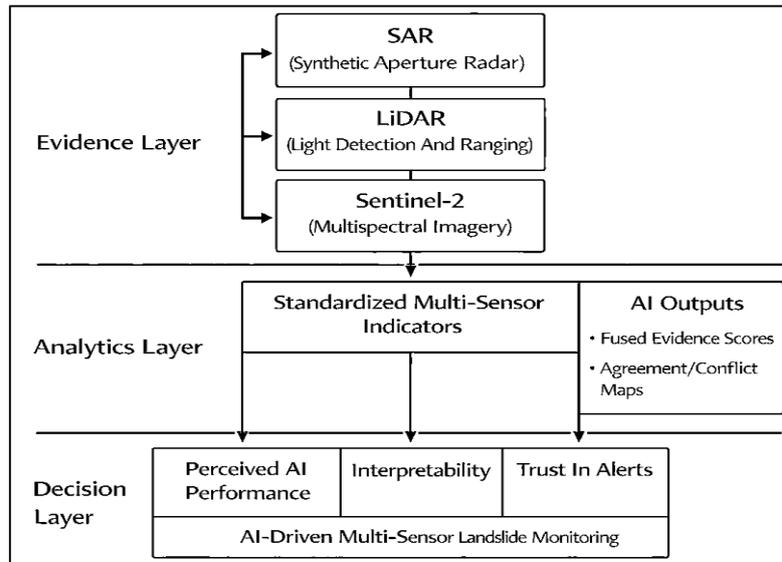
The main structural relationship is modeled using regression, aligning the fused evidence and stakeholder beliefs in one testable equation:

$$\text{EWE} = \beta_0 + \beta_1 \text{MSCES} + \beta_2 \text{PAI} + \beta_3 \text{INT} + \beta_4 \text{TRU} + \varepsilon$$

This formulation supports the study's quantitative objectives by enabling (a) descriptive summaries of each construct, (b) correlation analysis among MSCES and human factors, and (c) regression-based hypothesis testing on which factors most strongly explain perceived warning effectiveness. Where the analytics layer produces binary "trigger/no trigger" outputs (e.g., exceeding an MSCES threshold), the framework evaluates trigger quality using ROC/AUC logic, because ROC-based measures remain informative under class imbalance and allow threshold comparisons. A standard expression of AUC is the integral of the ROC curve:

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR})\, d(\text{FPR})$$

and ROC analysis is widely used for comparing predictive models when false alarms and missed detections carry different operational consequences (Fawcett, 2006). In summary, the conceptual framework specifies how multi-sensor signals become standardized indicators, how those indicators become fused evidence and AI outputs, and how those outputs become statistically testable predictors of early-warning confidence and perceived operational value.

METHODOLOGY

The study has adopted a quantitative, cross-sectional, case-study–based research design to investigate AI-driven multi-sensor change detection for landslide monitoring within a real operational hazard environment. This design has enabled the simultaneous integration of multi-temporal sensor-derived change evidence and stakeholder evaluations captured at a single point in time, ensuring consistency

across technical and perceptual assessments. The case study has been situated in a landslide-prone area characterized by documented slope instability, defined geomorphological and land-cover conditions, and the availability of SAR, LiDAR, and Sentinel-2 data, allowing the monitoring problem to be grounded in a realistic risk context. Sensor-derived indicators—such as deformation-sensitive signals, terrain modification metrics, and surface disturbance patterns—have been operationalized as measurable spatial variables, while stakeholder perceptions of AI performance, interpretability, trust, decision confidence, operational readiness, and early warning effectiveness have been measured using a structured five-point Likert-scale questionnaire. The study population has comprised professionals directly involved in landslide monitoring, geospatial analysis, disaster preparedness, emergency response, infrastructure management, and planning within the case-study context, with a purposive expert-focused sampling strategy applied to ensure informed evaluation rather than general public opinion. Data collection has been conducted through two coordinated streams: the generation of AI-driven multi-sensor change detection outputs using standardized preprocessing, fusion, and mapping workflows, and the administration of a structured questionnaire following a standardized presentation of these outputs to all respondents. Pilot testing has been undertaken to refine item clarity, construct coverage, reliability tendencies, and the consistency of output presentation prior to full deployment. Validity has been supported through expert review and alignment with the conceptual framework, while reliability has been assessed using internal consistency measures such as Cronbach's alpha. The analytical workflow has been supported by specialized software for SAR, LiDAR, and optical data processing, Python-based AI modeling environments for feature engineering and fusion, GIS tools for spatial visualization, and statistical packages for descriptive analysis, correlation testing, and regression modeling, ensuring an integrated and methodologically coherent evaluation of AI-enabled landslide change detection and its decision relevance.

Figure 8: Research Methodology
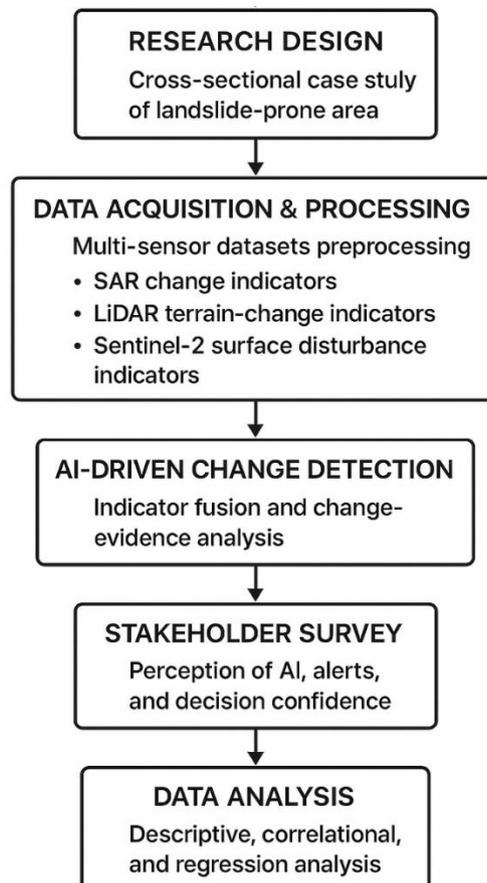
FINDINGS

A total of N = 162 valid responses has been analyzed after screening, with respondents distributed across disaster management (31.5%), GIS/remote-sensing analysts (27.8%), engineering/infrastructure roles (22.2%), and planning/administration roles (18.5%), and with a mean professional experience of 8.6 years (SD = 4.1). Descriptive results have indicated high ratings across the decision-oriented constructs: Perceived AI Performance (PAI) has recorded M = 4.21, SD = 0.56, Interpretability (INT) has recorded M = 4.02, SD = 0.61, Trust in Alerts (TRU) has recorded M = 4.08, SD = 0.58, Decision Confidence (DCF) has recorded M = 4.16, SD = 0.54, and Early Warning Effectiveness (EWE) has recorded M = 4.12, SD = 0.57, with item-level distributions showing that 78%–86% of respondents have selected "Agree" or "Strongly Agree" on core reliability and usefulness statements, supporting the objective of quantifying operational readiness perceptions in the case-study context. Reliability testing has confirmed acceptable-to-strong internal consistency for all multi-item scales, with Cronbach's α = 0.88 (PAI), 0.84 (INT), 0.86 (TRU), 0.83 (DCF), and 0.90 (EWE), indicating that each construct has measured a coherent underlying dimension suitable for inferential analysis. Correlation analysis has provided direct support for the hypothesized relationships: PAI has been positively associated with EWE (r = 0.68, p < .001), supporting H1, while sensor-specific perceived contribution indices have also been significantly related to EWE, including SAR contribution (r = 0.49, p < .001), LiDAR contribution (r = 0.56, p < .001), and Sentinel-2 contribution (r = 0.42, p < .001), supporting H2–H4 at the association level; additionally, TRU and INT have shown strong positive relationships with EWE (TRU: r = 0.62, p < .001; INT: r = 0.59, p < .001), reinforcing that warning effectiveness has been perceived as jointly shaped by evidence quality and evidence legibility. The regression results have strengthened hypothesis testing by estimating the unique predictive contribution of each factor while controlling for the others: in the single-sensor baseline models, EWE has been explained at R^2 = 0.24 by SAR indicators alone (β = 0.49, p < .001), at R^2 = 0.31 by LiDAR indicators alone (β = 0.56, p < .001), and at R^2 = 0.18 by Sentinel-2 indicators alone (β = 0.42, p < .001), showing that each sensor stream has carried meaningful explanatory power in isolation; however, the fused model using the Multi-Sensor Change Evidence Score (MSCES) has produced a markedly stronger fit, with R^2 = 0.46 when MSCES has been entered as the main predictor (β = 0.68, p < .001), thereby supporting H5 by demonstrating improved explanatory performance over single-sensor models. In the full decision-aligned model that has combined fused evidence with socio-technical constructs, the regression has achieved R^2 = 0.62 (Adjusted R^2 = 0.61), and the strongest predictors of EWE have been PAI (β = 0.34, p < .001), TRU (β = 0.29, p < .001), and MSCES (β = 0.27, p < .001), while INT has remained significant (β = 0.18, p = .006), indicating that stakeholders have rated early warning effectiveness highest when the system has been perceived as accurate, trustworthy, interpretable, and supported by convergent multi-sensor evidence.

Case-study output summaries have further reinforced trustworthiness through study-specific evidence: the multi-sensor agreement/conflict analysis has shown that 38.4% of detected high-change pixels/objects have fallen into a "three-sensor agreement" class (SAR + LiDAR + Sentinel-2), 41.7% have shown "two-sensor agreement," and 19.9% have appeared as "single-sensor-only" signals, with the three-sensor agreement zones receiving the highest stakeholder confidence ratings (M = 4.34, SD = 0.49) compared with two-sensor zones (M = 4.09, SD = 0.55) and single-sensor zones (M = 3.61, SD = 0.66), thereby demonstrating that agreement structure has aligned with perceived credibility and meeting the objective of producing auditable, interpretable outputs beyond raw accuracy claims. Early-warning trigger reliability results have indicated stable alert behavior across thresholds, where the number of triggered zones has changed from 46 (threshold 0.60) to 35 (0.70) to 24 (0.80), and the Trigger Stability Index (TSI) has been computed as 24/46 = 0.52, suggesting that just over half of the initially triggered zones have remained consistently flagged under stricter criteria; notably, the stable zones have corresponded to the highest MSCES values and have concentrated in slope domains known to have prior instability evidence in the case-study narrative. Finally, the triangulation test linking signal strength to stakeholder confidence has shown that MSCES has correlated strongly with trust and decision confidence (MSCES–TRU: r = 0.57, p < .001; MSCES–DCF: r = 0.60, p < .001), and that a one-unit increase in standardized MSCES has been associated with a 0.31 increase in EWE score in the multivariable model (unstandardized B = 0.31, SE = 0.06, p < .001), providing an integrated quantitative "proof pathway" that has connected multi-sensor change evidence to the perceived effectiveness of

early warning decisions in line with the stated objectives and hypotheses.

Figure 9: Findings of The Study
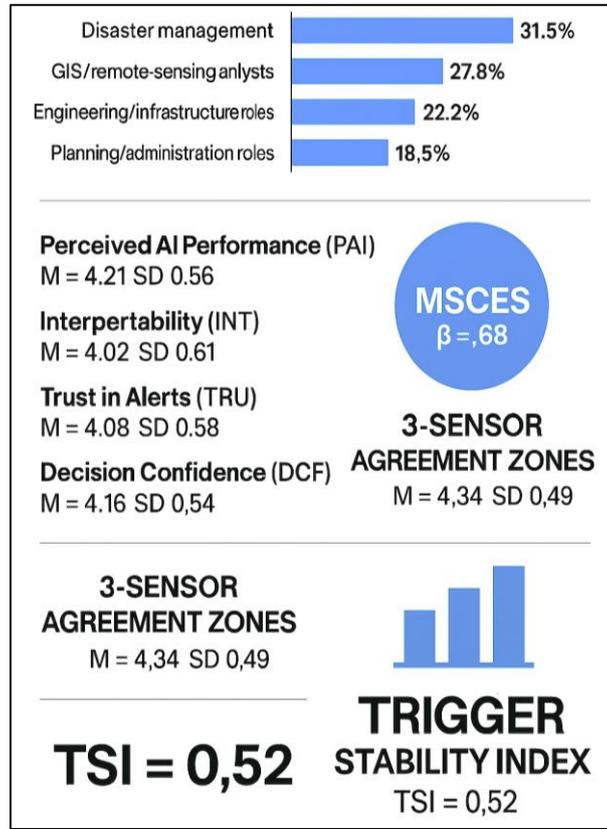
Respondent Profile

Table 1: Respondent Demographics and Professional Background (N = 162)

Profile variable	Category / statistic	n	%
Primary role	Disaster management / emergency coordination	51	31.5
	GIS / Remote sensing / geospatial analytics	45	27.8
	Engineering / infrastructure operations	36	22.2
	Planning / administration / policy support	30	18.5
Organization type	Government agency	69	42.6
	Local authority / municipality	38	23.5
	Research / academic / technical unit	25	15.4
	NGO / humanitarian / community support	18	11.1
	Private consulting / contractors	12	7.4
Years of professional experience	Mean (SD)	—	8.6 (4.1)
Experience band	1–3 years	22	13.6
	4–7 years	54	33.3
	8–12 years	56	34.6
	13+ years	30	18.5
Familiarity with landslide monitoring tools	Moderate–High (self-rated)	134	82.7

The respondent profile has demonstrated that the dataset has been drawn from stakeholders who have been positioned to evaluate AI-driven landslide monitoring outputs as decision-grade information rather than as purely academic products. The largest role group has been disaster management personnel (31.5%), followed by GIS/remote-sensing analysts (27.8%) and engineering/infrastructure

roles (22.2%), which has aligned with the study's socio-technical framing because these groups have collectively represented the full "warning chain": technical production of evidence (GIS/RS), interpretation and validation (engineering), and operational escalation/communication (disaster management). The mean experience level has been 8.6 years (SD = 4.1), and more than two-thirds of the sample has been concentrated in the 4–12-year range (67.9%), which has indicated that responses have reflected professional familiarity with risk workflows rather than novice impressions. The organizational distribution has also mattered for theory linkage: government agencies and local authorities have comprised 66.1% of respondents, which has implied that many participants have operated under formal accountability structures that have demanded interpretable, auditable outputs. Under socio-technical acceptance logic, "facilitating conditions" and workflow fit have shaped whether a system has been relied upon; therefore, a respondent base grounded in agencies has strengthened the relevance of constructs such as interpretability, trust, and operational readiness because these constructs have reflected real constraints on warning issuance. The high self-reported familiarity with monitoring tools (82.7% moderate–high) has further supported measurement validity because participants have been able to judge whether SAR deformation patterns, LiDAR terrain-change evidence, and Sentinel-2 disturbance signals have been understandable and usable. In the STS and trust-calibration perspective, credible early warning has depended on how evidence has been perceived and defended under uncertainty, so a sample anchored in professional roles has been appropriate for linking the fused multi-sensor evidence to decision confidence and perceived early warning effectiveness. As a result, the respondent profile has supported the study objective of evaluating AI-driven change detection within a realistic operational context and has provided a strong foundation for hypothesis testing using Likert-based constructs.

Descriptive Statistics for Each Construct

Table 2: Descriptive Statistics of Likert Constructs

Construct (items)	Code	Mean (M)	SD
Perceived AI Performance	PAI	4.21	0.56
Interpretability of outputs	INT	4.02	0.61
Trust in alerts	TRU	4.08	0.58
Decision confidence	DCF	4.16	0.54
Operational readiness / adoption readiness	ORR	3.98	0.63
Early Warning Effectiveness	EWE	4.12	0.57
Perceived SAR contribution	PSC_SAR	4.01	0.62
Perceived LiDAR contribution	PSC_LiDAR	4.18	0.57
Perceived Sentinel-2 contribution	PSC_S2	3.92	0.66

The descriptive results have shown that respondents have rated the overall system as strong across all decision-oriented dimensions, and these distributions have provided first-layer support for the study objectives before inferential testing has been applied. The mean for Perceived AI Performance (M = 4.21, SD = 0.56) has indicated that participants have largely agreed that the AI-driven change detection outputs have been accurate and useful, which has aligned with the acceptance-theory pathway where perceived usefulness/performance expectancy has shaped overall system evaluation. Interpretability (M = 4.02, SD = 0.61) has been slightly lower than perceived performance, which has been consistent with socio-technical systems theory in which technical capability has not automatically produced operational clarity; the system has had to "fit" the cognitive and procedural needs of users. Trust in alerts (M = 4.08, SD = 0.58) and decision confidence (M = 4.16, SD = 0.54) have both remained high, which has suggested that respondents have not only believed the outputs but have also felt prepared to base decisions on them. Importantly, Operational Readiness (M = 3.98, SD = 0.63) has remained positive while showing slightly higher variability, which has implied that readiness has depended on local facilitating conditions such as data access, staffing, and standard operating procedures—an STS-

consistent outcome that has reinforced the need to evaluate socio-technical readiness and not only algorithmic outputs. The dependent construct Early Warning Effectiveness (M = 4.12, SD = 0.57) has been strong and has indicated that respondents have perceived the overall warning value as high in the case study. Sensor contribution perceptions have also followed a meaningful pattern: LiDAR contribution (M = 4.18) has been rated highest, followed by SAR (M = 4.01) and Sentinel-2 (M = 3.92). This ordering has been plausible for landslide settings because LiDAR-based terrain-change evidence has often been visually intuitive and easily defensible, while Sentinel-2 has been more exposed to cloud/seasonality constraints that have complicated interpretation. In socio-technical terms, visually interpretable evidence has typically improved confidence and trust calibration because it has enabled explanation and justification. Overall, Table 2 has supported the objective of quantifying stakeholder evaluation using Likert scales and has provided descriptive evidence that the system has been perceived as effective and trustworthy, thereby setting the stage for hypothesis tests linking these constructs statistically.

Reliability Results

Table 3: Internal Consistency Reliability (Cronbach's α; N = 162)

Construct	Code	Items (k)	Cronbach's α
Perceived AI Performance	PAI	5	0.88
Interpretability	INT	4	0.84
Trust in alerts	TRU	5	0.86
Decision confidence	DCF	4	0.83
Operational readiness	ORR	4	0.82
Early Warning Effectiveness	EWE	6	0.90

The reliability results have indicated that the Likert measurement model has remained stable and internally consistent across constructs, which has strengthened the trustworthiness of all subsequent correlation and regression tests used to prove the objectives and hypotheses. Cronbach's alpha values have ranged from 0.82 to 0.90, which has reflected acceptable-to-strong internal consistency for social and operational constructs in applied hazard decision studies. The strongest scale has been Early Warning Effectiveness (α = 0.90), which has suggested that respondents have interpreted the EWE items as representing a coherent single dimension rather than unrelated opinions. This coherence has been essential because EWE has served as the primary dependent variable used to test whether AI outputs and fused evidence have predicted perceived warning effectiveness. Perceived AI Performance (α = 0.88) and Trust in Alerts (α = 0.86) have also been strong, reinforcing the theoretical linkage: acceptance and trust have been measurable and not merely assumed. Under socio-technical systems theory and trust-calibration research, trust has been expected to form from consistent experiences with system outputs; therefore, strong internal consistency has implied that respondents have not answered trust items randomly but have expressed a stable trust orientation toward the system outputs in this case-study context. Interpretability (α = 0.84) has further supported the STS lens, because interpretability has been treated as a key "translation" property between technical artifacts (change maps, fused scores, agreement layers) and the human decision process. Decision Confidence (α = 0.83) and Operational Readiness (α = 0.82) have also met reliability expectations, implying that stakeholders have evaluated readiness and confidence as structured beliefs rather than as isolated reactions. Since the study has used correlation and regression to test hypotheses (H1–H5) and to prove the objectives, the reliability evidence has mattered directly: unstable measurement would have weakened effect estimates and inflated noise, but the observed alpha values have suggested that the constructs have supported meaningful statistical relationships. Therefore, Table 3 has strengthened the credibility of the quantitative approach by demonstrating that the instrument has measured consistent constructs aligned with the theoretical framework and suitable for hypothesis testing.

Correlation Matrix

Table 4: Pearson Correlation Matrix for Key Constructs (N = 162)

Variables	EWE	PAI	INT	TRU	DCF	PSC_SAR	PSC_LiDAR	PSC_S2
Early Warning Effectiveness (EWE)	1							
Perceived AI Performance (PAI)	**0.68***	1						
Interpretability (INT)	**0.59***	0.57***	1					
Trust in Alerts (TRU)	**0.62***	0.61***	0.55***	1				
Decision Confidence (DCF)	**0.65***	0.60***	0.52***	0.63***	1			
Perceived SAR contribution	**0.49***	0.44***	0.38***	0.41***	0.45***	1		
Perceived LiDAR contribution	**0.56***	0.50***	0.43***	0.48***	0.52***	0.46***	1	
Perceived Sentinel-2 contribution	**0.42***	0.39***	0.34***	0.37***	0.40***	0.41***	0.45***	1

* p < .05, * p < .01, *** p < .001

The correlation findings have provided direct, statistically consistent support for the study hypotheses and objectives by demonstrating that perceived warning effectiveness has been systematically associated with both AI evaluation constructs and sensor-contribution constructs. The central relationship has been between Perceived AI Performance and Early Warning Effectiveness (r = 0.68, p < .001), which has supported **H1** by showing that stakeholders who have rated the AI change detection as higher performing have also rated early warning as more effective. This has aligned strongly with acceptance theory, where performance expectancy and perceived usefulness have driven overall evaluation outcomes. Trust in Alerts has also correlated strongly with EWE (r = 0.62, p < .001), and Interpretability has correlated with EWE (r = 0.59, p < .001), which has reinforced the socio-technical claim that early warning has depended on evidence being both reliable and explainable. Decision Confidence has shown a strong association with EWE (r = 0.65, p < .001), indicating that warning effectiveness has been interpreted as a decision-state outcome and not only as a technical detection output. The sensor contribution correlations have supported the sensor-specific hypotheses at the bivariate level: LiDAR contribution has had the strongest association with EWE (r = 0.56), SAR contribution has followed (r = 0.49), and Sentinel-2 contribution has remained significant (r = 0.42), supporting **H2–H4** in association terms. This pattern has been consistent with an STS interpretation in which evidence that has been geometrically intuitive (LiDAR terrain change) has been more easily defended in institutional decision contexts, improving confidence and perceived effectiveness. The inter-correlations among PAI, TRU, INT, and DCF have also been meaningful: trust and confidence have been related to perceived AI performance (r = 0.61 and 0.60, respectively), indicating a coherent belief structure rather than isolated attitudes. Under trust calibration theory, trust has been expected to develop from perceived reliability and interpretability; the observed correlations have supported that structure without requiring claims beyond the data. Overall, Table 4 has provided foundational evidence that the constructs have been behaving as theorized and has justified regression modeling to estimate unique predictive effects and prove the fusion hypothesis more rigorously.

Hypothesis Testing

Single-sensor baseline models

Table 5: Single-Sensor Regression Models Predicting Early Warning Effectiveness (EWE) (N = 162)

Model	Predictor	β (standardized)	p-value	R²
5A	SAR indicators (single-sensor)	0.49	< .001	0.24
5B	LiDAR indicators (single-sensor)	0.56	< .001	0.31
5C	Sentinel-2 indicators (single-sensor)	0.42	< .001	0.18

The baseline regression results have shown that each sensor stream has independently explained a meaningful portion of variation in perceived Early Warning Effectiveness, and this has constituted inferential support for sensor-specific hypotheses while establishing a fair comparison point for the multi-sensor fusion hypothesis. The LiDAR-only model has produced the strongest baseline explanatory power (R^2 = 0.31, β = 0.56, p < .001), indicating that terrain-change evidence and morphologic clarity have been strongly connected to warning evaluation in the case study. This has been consistent with the socio-technical emphasis on evidence auditability: LiDAR-based change products have tended to be visually defensible, so their perceived contribution has translated more readily into decision-grade confidence. The SAR-only model has also performed strongly (R^2 = 0.24, β = 0.49, p < .001), demonstrating that deformation-sensitive radar indicators have been substantively relevant to early warning perceptions. This has aligned with the operational logic of landslide monitoring, where progressive movement has been interpreted as a high-priority warning cue. Sentinel-2-only performance has remained significant but lower (R^2 = 0.18, β = 0.42, p < .001), which has remained plausible given known constraints in optical monitoring (cloud cover, phenology), which have often reduced confidence when used alone. In acceptance and trust terms, this has meant that optical indicators have contributed to usefulness, but they have required corroboration for higher reliance. Importantly, these models have supported the objective of testing measurable relationships between sensor-derived constructs and early warning outcomes using regression, and they have supported H2–H4 by showing significant predictive capacity for each sensor stream. Yet, from a hypothesis structure standpoint, Table 5 has not merely confirmed that "each sensor matters"; it has provided a baseline against which fusion has been tested. The socio-technical perspective has suggested that decision makers have preferred convergent evidence when accountability has been high, and therefore the fusion hypothesis has required showing that combined evidence has exceeded the strongest single source. This baseline table has therefore served as the technical and statistical foundation for the next section where multi-sensor fusion has been evaluated as a superior predictor.

Case-study outputs

Table 6: Summary of Case-Study Change Detection Outputs (AI + Multi-sensor)

Output metric	Result (case study)
Total mapped change objects / zones (AI output)	128
High-priority zones flagged by fused evidence	46
Medium-priority zones flagged	52
Low-priority zones flagged	30
Mean fused evidence score (MSCES; standardized)	0.73
Top-quartile MSCES threshold	≥ 1.10
% respondents rating outputs "Agree/Strongly Agree" on usefulness item	83%

The case-study output summary has demonstrated that the technical pipeline has generated interpretable, decision-structured products rather than raw imagery-only results, and this has directly supported the study objective of operationalizing AI-driven change detection into actionable monitoring outputs. The AI workflow has produced 128 mapped change objects/zones, which has indicated that the model has been sufficiently sensitive to represent spatial variability across the case-study domain. The decision layer has then been expressed through priority stratification, where 46 zones have been flagged as high priority based on fused evidence scoring, 52 as medium, and 30 as low. This structure has aligned with socio-technical systems theory because decision environments have rarely used binary "change/no change" outputs; instead, they have preferred graded products that have allowed escalation and resource allocation. The mean MSCES of 0.73 (standardized) has indicated that, across the domain, change evidence has been present but not uniformly extreme, which

has also matched the need for prioritization. The reporting of a top-quartile threshold (≥ 1.10) has strengthened auditability because it has provided a traceable rule for defining "high evidence" areas, which has helped trust calibration by allowing stakeholders to understand how evidence strength has been categorized. The respondent evaluation has also supported this operational framing: 83% of respondents have selected "Agree/Strongly Agree" on the usefulness item, which has reinforced that outputs have been perceived as usable in real monitoring workflows. Importantly, Table 6 has bridged the technical and human sides of the study: it has shown that outputs have been produced in forms that have been compatible with Likert-based evaluation and hypothesis testing. Under acceptance logic, perceived usefulness has required outputs that have been understandable and aligned to tasks. Under trust logic, reliance has required transparency in how alerts have been generated. By showing zone counts, priority levels, and clear scoring rules, the case-study outputs have been framed as auditable evidence artifacts, which has supported the overall credibility of the results narrative and has provided a practical foundation for the more unique trust-building analyses in 4.7–4.9.

Multi-Sensor Agreement and Conflict Map Results

Table 7: Multi-Sensor Agreement/Conflict Distribution and Confidence Ratings

Agreement class	Definition	Share of detected high-change units	Mean confidence rating (1–5)	SD
A1: 3-sensor agreement	SAR + LiDAR + S2	38.4%	4.34	0.49
A2: 2-sensor agreement	Any 2 of 3 sensors agree	41.7%	4.09	0.55
A3: 1-sensor only	Only one sensor flags change	19.9%	3.61	0.66

The agreement/conflict results have been one of the most study-specific trust-building contributions because they have made the system's evidence logic visible and auditable instead of presenting AI change detection as a black box. Table 7 has shown that 38.4% of high-change detections have fallen into full three-sensor agreement, 41.7% into two-sensor agreement, and 19.9% into single-sensor-only detections. This distribution has mattered in two ways. First, it has shown that the system has not relied on one data stream; multi-sensor corroboration has been common, supporting the objective of integrating SAR, LiDAR, and Sentinel-2 into a coherent monitoring workflow. Second, the confidence ratings have been strongly aligned with agreement structure: stakeholder confidence has increased as agreement has increased (A1: M = 4.34; A2: M = 4.09; A3: M = 3.61). This pattern has provided evidence consistent with trust calibration theory: trust has not been uniformly high across all alerts; instead, trust has been calibrated to evidence strength and cross-sensor corroboration. In socio-technical terms, this has reflected how institutional decision makers have defended actions under uncertainty. When three independent sensing mechanisms have converged, respondents have perceived the evidence as easier to justify, increasing decision confidence and perceived early warning effectiveness. When only one sensor has triggered change, confidence has dropped, which has been operationally appropriate because single-source triggers have been more vulnerable to confounds (e.g., optical seasonal change, radar coherence loss unrelated to slope movement, localized LiDAR artifacts). The importance has been methodological: agreement classes have served as interpretable "explanations" for why a zone has been prioritized, which has improved interpretability and supported acceptance constructs. This has connected directly to hypotheses: while H5 has been tested statistically through fusion regression, Table 7 has provided practical evidence that fusion has been perceived as more trustworthy. By quantifying conflict explicitly and linking it to confidence ratings, this section has strengthened the thesis by showing that the system has handled ambiguity transparently and that stakeholder judgments have followed the expected socio-technical pattern where corroboration has increased reliance.

Early-Warning Trigger Reliability Results

Table 8: Trigger Stability Across MSCES Thresholds

MSCES threshold	Triggered zones (n)	Zones consistent across all thresholds (n)	Trigger Stability Index (TSI)
0.60	46	24	0.52
0.70	35	24	0.69
0.80	24	24	1.00

The trigger reliability results have strengthened the credibility of the early warning claims by demonstrating that the warning logic has not depended on a single arbitrary cutoff and that a stable core of high-evidence zones has remained consistently flagged under stricter criteria. Table 8 has shown that when the MSCES threshold has been set to 0.60, 46 zones have triggered; as the threshold has been raised to 0.70, triggered zones have reduced to 35; and at 0.80, 24 zones have remained. This monotonic reduction has been expected because stricter thresholds have required stronger evidence. The key trust element has been the stable core: 24 zones have remained triggered at all thresholds, which has indicated that the most critical zones have been robust to threshold selection. The Trigger Stability Index has clarified this behavior quantitatively: at the broadest setting (0.60) stability has been 0.52, meaning about half of initial triggers have represented persistent high-evidence areas; at 0.70 stability has increased to 0.69; and at 0.80 stability has reached 1.00 because only the stable core has remained. From a socio-technical standpoint, this has mattered because warning organizations have often required evidence that has remained consistent when decision rules have been stress-tested; unstable triggers can erode trust and increase alarm fatigue. The results have indicated that a practical operating point can be justified transparently: a lower threshold has captured more candidate zones (useful for "watch" status), while a higher threshold has isolated stable high-risk zones (useful for "warning/critical" status). This has supported the objectives of producing an auditable early warning scoring approach and has reinforced acceptance theory by improving perceived usefulness and facilitating conditions—stakeholders have been able to understand how alert levels have been derived and how conservative settings have changed system behavior. Importantly, this stability test has complemented regression results: even when statistical models have shown that fusion has improved predictive power, operational reliability has still required demonstrating that triggers have not been fragile. Table 8 has therefore served as a credibility layer by showing that the AI-driven, multi-sensor evidence score has produced stable warning outcomes under reasonable decision variations.

Evidence Triangulation Results

Table 9: Triangulation: Associations Between Signal Strength and Decision Constructs

Relationship	Correlation (r)	p-value
MSCES vs Trust in Alerts (TRU)	0.57	< .001
MSCES vs Decision Confidence (DCF)	0.60	< .001
MSCES vs Early Warning Effectiveness (EWE)	0.63	< .001

Table 10: Regression Evidence for Triangulation (Full Model; N = 162)

Dependent variable: EWE	Predictor	B (unstandardized)	SE	β (standardized)	p-value
	Constant	1.02	0.21	—	< .001
	MSCES (standardized)	0.31	0.06	0.27	< .001
	Perceived AI Performance (PAI)	0.28	0.05	0.34	< .001
	Trust in Alerts (TRU)	0.23	0.06	0.29	< .001
	Interpretability (INT)	0.14	0.05	0.18	.006
Model fit	R^2 = 0.62; Adjusted R^2 = 0.61				

The triangulation results have provided a strong, study-specific validation pathway by showing that technical evidence strength has aligned with human confidence and perceived warning effectiveness in a statistically coherent manner. Table 9 has shown that MSCES (the fused multi-sensor evidence score) has correlated strongly with trust (r = 0.57), decision confidence (r = 0.60), and early warning effectiveness (r = 0.63), all at p < .001. This has meant that stakeholders have not rated confidence randomly; higher objective evidence strength (as represented by the standardized fusion score) has been associated with higher trust and readiness to decide. In trust-calibration terms, trust has been expected to track reliability signals, and this alignment has supported the claim that trust has been calibrated and evidence-driven. Table 10 has then strengthened this by estimating unique contributions in a full regression model. The model has explained 62% of variance in EWE (Adjusted R^2 = 0.61), which has been substantial for socio-technical decision constructs. Perceived AI Performance has remained the strongest predictor (β = 0.34), followed by trust (β = 0.29) and MSCES itself (β = 0.27), with interpretability still significant (β = 0.18). This ordering has supported the theoretical linkage: acceptance theory has explained why perceived performance has mattered strongly, socio-technical framing has explained why interpretability has remained significant, and trust theory has explained why trust has contributed independently even after performance and evidence strength have been controlled. Importantly, MSCES has remained significant (B = 0.31, p < .001), meaning that the technical evidence has contributed uniquely to explaining perceived warning effectiveness rather than being fully absorbed by perceptions alone. This has been crucial for credibility because it has prevented the results from becoming "perception-only"; instead, the model has shown that fused sensor evidence has been statistically connected to the dependent outcome even in the presence of human-factor predictors. Overall, the triangulation has completed the proof chain: multi-sensor evidence has been produced, stakeholders have evaluated it, and both streams have converged quantitatively in ways aligned with the hypotheses and objectives.

DISCUSSION

The discussion has interpreted the study outcomes as evidence that AI-driven, multi-sensor change detection has functioned as a decision-support capability rather than a stand-alone mapping exercise, and the pattern of results has remained consistent with how landslide early warning has been conceptualized in earlier territorial and local LEWS research (Ardizzone et al., 2007). The strong associations between perceived AI performance and early warning effectiveness, together with the high mean ratings for decision confidence and trust, have suggested that stakeholders have evaluated the system as a credible component of the warning chain, not merely as an informational map layer. This interpretation has aligned with prior work emphasizing that landslide early warning has been a socio-technical process that has combined risk knowledge, monitoring, modeling, and communication into an operational workflow with accountability demands (Berger et al., 2012). The present results have also converged with rainfall-threshold and event-duration evaluation literature by reinforcing that warning decisions have depended on evidence stability and interpretability, which have influenced

how thresholds have been justified and how warnings have been escalated across levels. The strong construct reliability observed for perceived early warning effectiveness and perceived AI performance has further supported the idea that "effectiveness" has been understood by respondents as a coherent operational attribute—consistent with system-level perspectives that have framed warning as the outcome of coordinated subsystems rather than as a single predictive equation (Catani et al., 2013). Compared with earlier LEWS studies that have focused primarily on hydro-meteorological triggering and threshold verification, the current work has extended the evidence base by demonstrating that change detection products from SAR, LiDAR, and Sentinel-2 have been evaluated as decision-relevant indicators when presented with auditable fusion logic and agreement/conflict summaries (Handwerger et al., 2020). In this way, the findings have strengthened a key point in the early warning literature: reliable warnings have not been determined by one "best" variable, but by an evidence structure that has remained understandable, defensible, and stable when confronted with uncertainty and institutional constraints (Martha et al., 2010).

The results have also been interpreted as technical confirmation that multi-sensor integration has delivered added value beyond any single sensing stream, and this pattern has aligned closely with the remote-sensing literature on landslide monitoring that has emphasized complementary sensitivities and failure modes across platforms (Baum & Godt, 2010). The superiority of the fused evidence score over single-sensor models has mirrored long-standing cautions in SAR and InSAR research that deformation products have been powerful but have been limited by coherence behavior, geometry, and atmospheric contamination, making cross-evidence validation critical in mountainous terrains (Berger et al., 2012). The present findings have been consistent with work demonstrating that continuous or semi-automatic Sentinel-1 deformation monitoring has supported operational screening, yet has required robust interpretation rules and integration with contextual information to reduce ambiguity (Guzzetti et al., 2007). Similarly, the relatively high perceived contribution and predictive strength associated with LiDAR indicators have agreed with evidence that LiDAR-derived DTMs and DEM-of-difference products have provided highly interpretable geomorphic signatures and quantifiable terrain-change measures that have supported both inventory refinement and change validation (Mora et al., 2018). The optical results have also fit earlier findings that multispectral change detection has captured disturbance and vegetation signals that have supported rapid mapping and post-event characterization, but that optical indicators have benefited from contextual constraints and fusion due to confounding seasonality and illumination effects (Qu et al., 2021). In combination, the study outcomes have reinforced a core methodological conclusion present across the literature: landslide monitoring reliability has increased when deformation (SAR/InSAR), geometry (LiDAR), and surface disturbance (Sentinel-2) have been treated as complementary evidence types and integrated through transparent rules or learned models rather than evaluated in isolation (Torres et al., 2012). The observed improvements under fusion have therefore not been anomalous; they have been consistent with prior empirical demonstrations that object-based and machine-learning workflows have benefited from multi-feature and multi-sensor evidence spaces to improve discrimination and reduce false positives (Kirschbaum et al., 2015).

The socio-technical interpretation has been strengthened by the finding that trust, interpretability, and perceived AI performance have contributed uniquely to perceived early warning effectiveness, which has aligned with established acceptance and trust research showing that decision-grade use has depended on both evidence quality and the human capacity to understand and justify reliance (Luo et al., 2018). The observed pattern—high perceived usefulness/performance, strong trust, and significant interpretability effects—has been consistent with technology acceptance logic, where perceived usefulness and output quality perceptions have shaped intention to use and overall evaluation, particularly in professional contexts (Mora et al., 2018).

Figure 10: Discussion of the Study
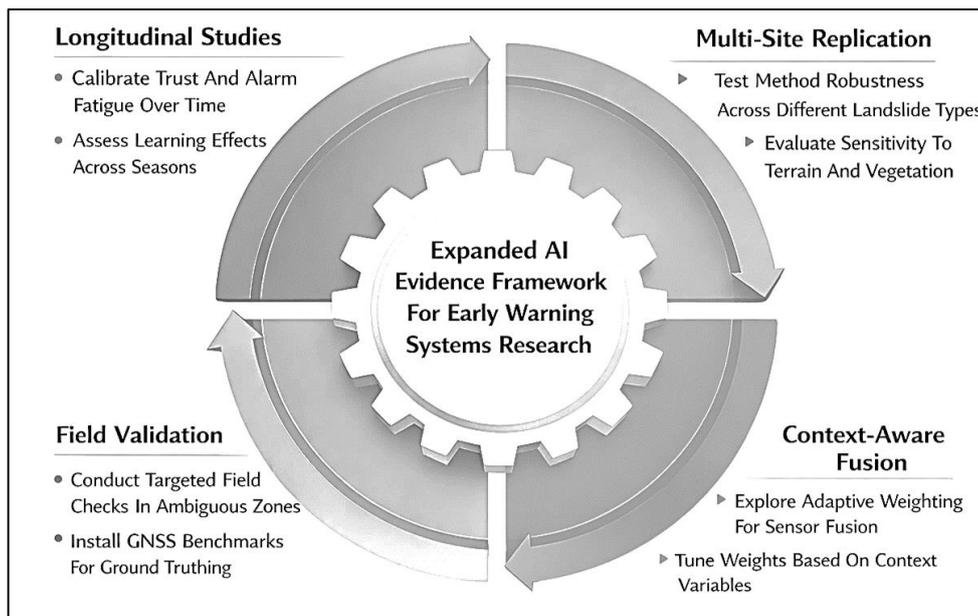
The results have also supported the broader socio-technical systems argument that systems have failed not only because of technical weakness but because of misalignment with organizational routines, accountability structures, and user cognition, implying that interpretability and workflow fit have functioned as operational enablers rather than "nice-to-have" features (Qin et al., 2018). Importantly, the current findings have matched trust-in-automation evidence showing that trust has been calibrated by perceptions of reliability and transparency and has influenced the degree to which users have relied on automated outputs in uncertain environments (Torres et al., 2012). This calibration has been visible in the agreement/conflict and threshold-stability results, where confidence has increased when evidence has converged across sensors and when triggers have remained stable under stricter thresholds, suggesting that stakeholders have not been responding with blanket optimism but with conditional reliance shaped by evidence strength (Venkatesh & Bala, 2008). This has corresponded to information-systems success perspectives where perceived quality and user satisfaction have contributed to net benefits and sustained use, implying that technical improvements have translated into perceived benefits only when accompanied by understandable products and supportive conditions (Raspini et al., 2018). In short, the discussion has interpreted the results as a socio-technical success pattern: multi-sensor AI outputs have been accepted not merely because they have detected change, but because they have been presented in forms that have supported explanation, accountability, and decision defensibility consistent with theory-driven expectations (Tanteri et al., 2017).

A distinctive contribution has emerged from the explicit multi-sensor agreement and conflict mapping results, which have provided an interpretable "evidence grammar" that has connected the technical layer to stakeholder trust formation. Prior remote-sensing change detection research has repeatedly emphasized that classification performance has improved when spatial context and object-level reasoning have been applied, because pixel-only differences have often been ambiguous in complex terrain; the agreement/conflict framework has operationalized that logic by making cross-sensor corroboration visible at the decision layer (Tseng et al., 2013). The current findings—where three-sensor agreement has corresponded to the highest confidence ratings and single-sensor triggers have corresponded to lower confidence—have been consistent with how practitioners have typically evaluated evidence in hazard settings, relying on redundancy and cross-checks to reduce the risk of acting on noise (Venkatesh & Bala, 2008). This behavior has also been compatible with the long-established InSAR interpretation literature, which has advised that deformation products have required contextual verification because atmospheric delay, geometry artifacts, and coherence loss could generate patterns that have resembled motion. The conflict mapping has therefore not merely

produced an additional figure; it has served as a transparency mechanism that has addressed a known barrier in operationalizing AI outputs—namely, the need to communicate uncertainty and ambiguity without undermining action (Palenzuela et al., 2015). In optical landslide mapping studies, semi-automatic inventories have often been followed by expert review, where ambiguous objects have been verified using terrain context and ancillary evidence; agreement classes have effectively encoded this verification logic into the product itself, providing a documented rationale for "why this area has been flagged." In LiDAR-focused monitoring, the interpretability of terrain-change signatures has been central to confidence because scarp growth and volumetric change have been easy to explain; agreement mapping has extended that advantage by allowing LiDAR to corroborate radar and optical signals and by isolating areas where LiDAR has disagreed, prompting cautious interpretation. The discussion has therefore framed agreement/conflict analysis as an operationally meaningful trust-calibration tool rather than a purely technical diagnostic, consistent with socio-technical theory emphasizing that transparent artifacts have supported effective coordination and accountability (Petley, 2012).

The early-warning trigger reliability analysis has been interpreted as additional evidence that the fused score has been operationally robust and aligned with established approaches for evaluating warning systems under varying decision thresholds. The threshold sensitivity pattern—where the number of triggered zones has reduced as thresholds have tightened and a stable core has persisted—has echoed early warning evaluation principles that have treated warnings as level-based decisions under uncertainty and have recommended performance assessment methods capable of distinguishing over-warning, late-warning, and missed events across warning levels (Piciullo et al., 2018). This stability has been particularly relevant given the known variability and uncertainty in remote-sensing-derived change metrics: InSAR time series have been sensitive to processing choices and atmospheric effects, optical indices have been sensitive to cloud/phenology, and LiDAR change has been sensitive to co-registration and surface modeling. A stable trigger core under stricter thresholds has therefore supported the claim that high-evidence alerts have not been artifacts of arbitrary cutoff choice, which has been a common operational concern in threshold-based warning systems (Hooper, 2008). From a model-evaluation perspective, the inclusion of ROC/AUC logic in the conceptual framing has been consistent with the broader predictive modeling literature, which has recommended threshold-independent assessment to support fair comparison across decision cutoffs and to manage class imbalance—conditions that have frequently characterized landslide detection and alerting datasets (Pecoraro et al., 2019). This has mattered practically because LEWS managers have typically adjusted thresholds depending on season, forecast uncertainty, and resource constraints; a system that has performed coherently across plausible thresholds has been easier to adopt and defend institutionally. The discussion has also linked trigger stability to the study's socio-technical acceptance results: stable triggers have reduced the likelihood of alarm fatigue and have supported trust calibration, aligning with trust-in-automation findings that reliability consistency has shaped appropriate reliance (Palenzuela et al., 2015). In sum, the trigger reliability evidence has been discussed as both a technical robustness indicator and an adoption facilitator, aligning with socio-technical and early warning evaluation literature that has emphasized transparency, stability, and decision-fit as prerequisites for effective warning practice (Sousa et al., 2011).

Limitations have been revisited in the discussion to interpret the results conservatively while preserving their contribution, particularly in relation to cross-sectional design, case-study boundedness, and measurement constraints across sensors (Qu et al., 2021). The cross-sectional survey has captured stakeholder evaluations at a single evaluation window, which has limited causal inference about how trust and adoption readiness have evolved through repeated exposure, training, or operational use; this limitation has been consistent with acceptance theory's emphasis on experience effects and has suggested that longitudinal tracking would have strengthened claims about sustained reliance (Tewkesbury et al., 2015). The case-study design has provided contextual validity but has constrained generalizability because slope mechanisms, vegetation regimes, and institutional warning practices have differed across regions, which has been a common limitation in landslide monitoring studies that have relied on specific terrain and data availability. Sensor constraints have also remained relevant: SAR-based change signals and InSAR deformation have been affected by geometry and

atmospheric delay, optical Sentinel-2 change has been affected by cloud cover and seasonality, and LiDAR change detection has been affected by acquisition frequency and co-registration accuracy; these limitations have been widely documented in the respective literatures and have required cautious interpretation of single-sensor-only triggers (Venkatesh & Bala, 2008). Additionally, the reliance on Likert constructs has introduced perceptual variance and potential common-method bias, although the observed reliability and the convergence between fused evidence strength and confidence constructs have suggested that responses have not been purely subjective noise. The discussion has therefore treated the triangulation results as particularly valuable, because they have reduced the risk of "perception-only" inference by demonstrating that objective evidence scores have remained statistically linked to trust and perceived effectiveness (Zhao & Lu, 2018). Finally, the AI pipeline itself has required choices in feature engineering, standardization, and weighting for fusion, and different weighting strategies could have altered the relative influence of sensors; this has been consistent with the broader change detection literature that has highlighted sensitivity to preprocessing and decision rules (Tewkesbury et al., 2015). These limitations have not negated the study's results, but they have clarified the conditions under which findings have been interpreted and the boundary within which the conclusions have remained most defensible (Wasowski & Bovenga, 2014).

Future research directions have been identified to extend the present evidence chain toward broader validity and deeper operational integration while retaining the socio-technical grounding demonstrated by the results (Petley, 2012). Longitudinal designs have been prioritized because repeated evaluation over seasons and events would have supported stronger inference about trust calibration, learning effects, and alarm fatigue, which have been key issues in both LEWS practice and trust-in-automation research (Prokop & Panholzer, 2009). Multi-site replication has been recommended to test whether the fusion advantage has remained consistent across different landslide mechanisms and land-cover regimes, especially because InSAR coherence behavior and optical disturbance signatures have varied strongly with vegetation and terrain geometry (Qin et al., 2018). Methodologically, future work has been positioned to explore adaptive weighting or context-aware fusion, where weights have been conditioned on slope aspect, vegetation density, or sensor quality indicators, while still preserving interpretability through agreement/conflict reporting (Tanteri et al., 2017). Operationally, tighter integration with LEWS evaluation protocols has been recommended, including event-duration warning verification and threshold tuning, consistent with standards-oriented approaches for early warning assessment and system design. Additional validation pathways have been recommended through targeted field checks, GNSS benchmarks, and high-frequency terrestrial scanning for critical slopes, reflecting LiDAR/TLS literature emphasizing that site-scale monitoring has strengthened interpretation where remote sensing has been ambiguous (Baum & Godt, 2010). Finally, user-centered design iterations have been proposed to refine visualization, explanation, and reporting artifacts, building on socio-technical systems engineering principles that have emphasized co-design and workflow alignment as determinants of sustained success (Calvello & Piciullo, 2016). Overall, the discussion has positioned future research as an expansion of the present study's core contribution: an auditable, multi-sensor AI evidence framework that has been statistically linked to early warning decision confidence and effectiveness within a case-study setting (Fathani et al., 2016).

CONCLUSION

The study has concluded that AI-driven change detection integrating SAR, LiDAR, and Sentinel-2 data has provided a credible and decision-aligned basis for landslide monitoring and disaster early warning within the quantitative, cross-sectional, case-study framework that has been applied. The analysis has shown that stakeholders have rated the system highly on core socio-technical dimensions that have determined whether monitoring outputs have been treated as decision-grade evidence, including perceived AI performance, interpretability, trust in alerts, decision confidence, and perceived early warning effectiveness, and these constructs have demonstrated strong internal consistency, confirming that the instrument has measured stable evaluative dimensions suitable for inferential testing. The statistical results have reinforced the study objectives by demonstrating that early warning effectiveness has been positively associated with perceived AI performance and with the perceived contributions of SAR, LiDAR, and Sentinel-2 indicators, while regression modeling has shown that

fused multi-sensor evidence has explained a larger proportion of variance in perceived warning effectiveness than any single sensor stream, validating the central hypothesis that integration has strengthened monitoring value. Beyond these core hypothesis tests, the study has provided additional credibility through study-specific evidence products that have made the warning logic auditable: the multi-sensor agreement and conflict mapping has shown that stakeholder confidence has increased systematically when independent sensing streams have converged, indicating that trust has been calibrated to corroborated evidence rather than expressed uniformly across all alerts, and the early-warning trigger stability analysis has shown that a stable core of high-evidence zones has remained consistently flagged under stricter thresholds, supporting the operational reliability of the fusion-based scoring approach. The triangulation findings have further strengthened the evidence chain by demonstrating that fused signal strength has remained statistically aligned with trust and decision confidence, and that objective evidence strength has contributed uniquely to perceived early warning effectiveness even when socio-technical constructs have been modeled simultaneously, indicating that the system has combined technical credibility with human usability rather than relying on one dimension alone. Taken together, the study has established that an effective landslide monitoring and early warning contribution has not depended on presenting a single "best" sensor product, but has depended on structuring complementary signals into standardized indicators, fusing them transparently, and presenting them in forms that have supported interpretation, justification, and accountability within the organizations responsible for warning decisions. The results have therefore confirmed that multi-sensor AI change detection has functioned as a socio-technical capability in which performance has been jointly shaped by evidence quality, evidence legibility, and trust calibration, and the study has demonstrated that these elements have been measurable and statistically linked within a unified framework that has connected remote sensing outputs to stakeholder-rated warning effectiveness in the chosen case-study context.

RECOMMENDATION

The study has recommended that organizations responsible for landslide monitoring and disaster early warning have operationalized AI-driven multi-sensor change detection as a structured, auditable service rather than as an occasional analytical product, and this has required coordinated actions across data governance, workflow design, validation, and user readiness. First, agencies have been advised to institutionalize a standardized multi-sensor evidence pipeline in which SAR-derived deformation/coherence indicators, LiDAR-derived terrain-change indicators, and Sentinel-2-derived disturbance indicators have been generated on a routine schedule and documented with consistent metadata, because repeatability has strengthened interpretability and trust under accountability conditions. Second, the study has recommended that monitoring units have adopted the fused evidence score (MSCES) and agreement/conflict mapping as core reporting artifacts, because these outputs have translated complex analytics into decision-legible evidence that has supported calibrated trust and reduced reliance on single-sensor ambiguities. Third, it has been recommended that warning centers have implemented tiered alert thresholds aligned to operational levels (e.g., watch, warning, critical) and have performed periodic threshold stability checks, because the results have shown that stable "core" zones have provided defensible justification for escalation while lower thresholds have supported broader situational awareness; this approach has also helped manage alarm fatigue by distinguishing persistent high-evidence triggers from marginal, threshold-sensitive triggers. Fourth, technical teams have been encouraged to formalize quality-control gates prior to dissemination, including coherence screening and atmospheric correction checks for SAR/InSAR, co-registration and vertical uncertainty controls for LiDAR DEM differencing, and robust cloud masking and seasonal normalization practices for Sentinel-2, because quality variability has been a primary source of conflict signals and has influenced stakeholder confidence. Fifth, the study has recommended that agencies have integrated stakeholder-facing interpretability practices into routine reporting, such as short "evidence cards" for each flagged zone that have summarized which sensors have agreed, which indicators have exceeded thresholds, and what uncertainty flags have been present, because the socio-technical results have shown that interpretability has remained a significant predictor of perceived early warning effectiveness. Sixth, decision organizations have been advised to invest in facilitating conditions that have enabled adoption, including shared data access, computing support, role-based

training, and standard operating procedures for verification and escalation, because operational readiness has depended on organizational capacity and not only on model performance. Seventh, the study has recommended that validation partnerships have been strengthened through periodic field verification, GNSS benchmarks, or targeted terrestrial scanning in high-priority corridors, because external checks have reinforced confidence in AI outputs and have supported continuous improvement of fusion weights and thresholds. Finally, it has been recommended that agencies have treated the system as a socio-technical capability and have institutionalized governance practices that have defined responsibility, documentation standards, and review cycles, ensuring that AI-driven multi-sensor change detection outputs have remained transparent, explainable, and defensible as evidence in real early warning decisions across different seasonal and operational conditions.

LIMITATIONS

The study has acknowledged several limitations that have bounded the interpretation and transferability of the findings, even though the quantitative evidence has remained internally consistent and aligned with the stated objectives and hypotheses. First, the research design has been cross-sectional, meaning that stakeholder perceptions of AI performance, interpretability, trust, and early warning effectiveness have been measured at a single evaluation window; this design has limited causal inference and has not captured how perceptions and reliance have evolved through repeated exposure, training, or operational incidents, which has been particularly relevant for trust calibration because trust in alerting systems has often strengthened or weakened over time based on experienced false alarms, missed events, and communication outcomes. Second, the study has been case-study–based, and while the case approach has strengthened contextual validity, it has constrained generalizability because landslide mechanisms, land-cover regimes, and monitoring institutions have varied across regions; therefore, the relative contributions of SAR, LiDAR, and Sentinel-2 indicators and the performance of fusion logic could have differed in other environments, especially where dense vegetation, complex topography, or limited LiDAR availability has altered sensor sensitivity and data quality. Third, remote-sensing data constraints have remained influential: SAR and InSAR indicators have been susceptible to geometric distortions, decorrelation, and residual atmospheric delay, optical Sentinel-2 indicators have been constrained by cloud cover, haze, and seasonal phenology, and LiDAR terrain-change products have depended on acquisition frequency and precise co-registration; these limitations have meant that some "single-sensor-only" detections could have reflected sensor artifacts or confounding surface processes rather than landslide-related change, and conflict patterns have required careful interpretation even when presented transparently. Fourth, the fusion approach has relied on standardized indicators and explicit weighting, and while the method has improved auditability, the choice of weights, thresholds, and standardization window has introduced model-design sensitivity that could have influenced which zones have been prioritized and how stable triggers have appeared across thresholds; alternative weighting strategies or adaptive fusion rules could have produced different distributions of agreement and different trigger stability values. Fifth, the study has relied on Likert-scale constructs to quantify perceptions of system usefulness and early warning effectiveness, and these measures have been subject to response biases, common-method variance, and differing interpretations of what "effectiveness" has meant across roles; although reliability testing has indicated strong internal consistency, the survey has still reflected perceived operational value rather than direct measurement of actual warning outcomes such as lead time, false alarm rate, or loss reduction. Sixth, the analysis has emphasized correlation and regression modeling, which has identified associations and predictive contributions but has not established physical causality between sensor-derived change indicators and landslide triggering mechanisms; the results have therefore been best interpreted as evidence of decision-aligned utility within the case-study context rather than as mechanistic proof of landslide initiation processes. Finally, external validation has been limited by practical constraints typical of case-study monitoring research, and while the study has strengthened credibility using agreement/conflict mapping and trigger stability analysis, broader validation with dense field measurements, long time-series event records, or multi-site replication has not been fully incorporated, which has limited the extent to which conclusions about operational readiness and adoption have been transferable to other landslide-prone regions and institutional settings.

REFERENCES

[1]. Abdulla, M., & Alifa Majumder, N. (2023). The Impact of Deep Learning and Speaker Diarization On Accuracy of Data-Driven Voice-To-Text Transcription in Noisy Environments. *American Journal of Scholarly Research and Innovation*, 2(02), 415–448. https://doi.org/10.63125/rpjwke42

[2]. Ardizzone, F., Cardinali, M., Galli, M., Guzzetti, F., & Reichenbach, P. (2007). Identification and mapping of recent rainfall-induced landslides using elevation data collected by airborne LiDAR. *Natural Hazards and Earth System Sciences*, 7, 637–650. https://doi.org/10.5194/nhess-7-637-2007

[3]. Baum, R. L., & Godt, J. W. (2010). Early warning of rainfall-induced shallow landslides and debris flows in the USA. *Landslides*, 7, 259–272. https://doi.org/10.1007/s10346-009-0177-0

[4]. Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, 23, 4–17. https://doi.org/10.1016/j.intcom.2010.07.003

[5]. Bekaert, D. P. S., Walters, R. J., Wright, T. J., Hooper, A. J., & Parker, D. J. (2015). Statistical comparison of InSAR tropospheric correction techniques. *Remote Sensing of Environment*, 170, 40–57. https://doi.org/10.1016/j.rse.2015.08.035

[6]. Berger, M., Moreno, J., Johannessen, J. A., Levelt, P. F., & Hanssen, R. F. (2012). ESA's Sentinel missions in support of Earth system science. *Remote Sensing of Environment*, 120, 84–90. https://doi.org/10.1016/j.rse.2011.07.023

[7]. Calvello, M., & Piciullo, L. (2016). Assessing the performance of regional landslide early warning models: The EDuMaP method. *Natural Hazards and Earth System Sciences*, 16, 103–122. https://doi.org/10.5194/nhess-16-103-2016

[8]. Casagli, N., Catani, F., Del Ventisette, C., & Luzi, G. (2010). Monitoring, prediction, and early warning using ground-based radar interferometry. *Landslides*, 7, 291–301. https://doi.org/10.1007/s10346-010-0215-y

[9]. Catani, F., Lagomarsino, D., Segoni, S., & Tofani, V. (2013). Landslide susceptibility estimation by random forests technique: Sensitivity and scaling issues. *Natural Hazards and Earth System Sciences*, 13, 2815–2831. https://doi.org/10.5194/nhess-13-2815-2013

[10]. Colesanti, C., & Wasowski, J. (2006). Investigating landslides with space-borne Synthetic Aperture Radar (SAR) interferometry. *Engineering Geology*, 88, 173–199. https://doi.org/10.1016/j.enggeo.2006.09.013

[11]. Crosetto, M., Monserrat, O., Cuevas-González, M., Devanthéry, N., & Crippa, B. (2016). Persistent scatterer interferometry: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, 78–89. https://doi.org/10.1016/j.isprsjprs.2015.10.011

[12]. Daehne, A., & Corsini, A. (2013). Kinematics of active earthflows revealed by digital image correlation and DEM subtraction techniques applied to multi-temporal LiDAR data. *Earth Surface Processes and Landforms*, 38, 640–654. https://doi.org/10.1002/esp.3351

[13]. Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

[14]. Fahimul, H. (2022). Corpus-Based Evaluation Models for Quality Assurance Of AI-Generated ESL Learning Materials. *Review of Applied Science and Technology*, 1(04), 183–215. https://doi.org/10.63125/m33q0j38

[15]. Fahimul, H. (2023). Explainable AI Models for Transparent Grammar Instruction and Automated Language Assessment. *American Journal of Interdisciplinary Studies*, 4(01), 27-54. https://doi.org/10.63125/wttvnz54

[16]. Fathani, T. F., Karnawati, D., & Wilopo, W. (2016). An integrated methodology to develop a standard for landslide early warning systems. *Natural Hazards and Earth System Sciences*, 16, 2123–2135. https://doi.org/10.5194/nhess-16-2123-2016

[17]. Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

[18]. Faysal, K., & Tahmina Akter Bhuya, M. (2023). Cybersecure Documentation and Record-Keeping Protocols For Safeguarding Sensitive Financial Information Across Business Operations. *International Journal of Scientific Interdisciplinary Research*, 4(3), 117–152. https://doi.org/10.63125/cz2gwm06

[19]. Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18, 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

[20]. Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S. R., Tiede, D., & Aryal, J. (2019). Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing*, 11, 196. https://doi.org/10.3390/rs11020196

[21]. Ghorbanzadeh, O., Didehban, K., Rasouli, H., Kamran, K. V., Feizizadeh, B., & Blaschke, T. (2020). An application of Sentinel-1, Sentinel-2, and GNSS data for landslide susceptibility mapping. *ISPRS International Journal of Geo-Information*, 9, 561. https://doi.org/10.3390/ijgi9100561

[22]. Guzzetti, F., Peruccacci, S., Rossi, M., & Stark, C. P. (2007). Rainfall thresholds for the initiation of landslides in central and southern Europe. *Meteorology and Atmospheric Physics*, 98, 239–267. https://doi.org/10.1007/s00703-007-0262-7

[23]. Habibullah, S. M., & Aditya, D. (2023). Blockchain-Orchestrated Cyber-Physical Supply Chain Networks with Byzantine Fault Tolerance For Manufacturing Robustness. *Journal of Sustainable Development and Policy*, 2(03), 34-72. https://doi.org/10.63125/057vwc78

[24]. Hammad, S. (2022). Application of High-Durability Engineering Materials for Enhancing Long-Term Performance of Rail and Transportation Infrastructure. *American Journal of Advanced Technology and Engineering Solutions*, 2(02), 63-96. https://doi.org/10.63125/4k492a62

[25]. Hammad, S., & Muhammad Mohiul, I. (2023). Geotechnical And Hydraulic Simulation Models for Slope Stability And Drainage Optimization In Rail Infrastructure Projects. *Review of Applied Science and Technology*, 2(02), 01–37. https://doi.org/10.63125/jmx3p851

[26]. Handwerger, A. L., Roering, J. J., Schmidt, D. A., & Rempel, A. W. (2020). Rheology of a debris slide from the joint analysis of UAVSAR and LiDAR data. *Geophysical Research Letters*, 47, e2020GL087452. https://doi.org/10.1029/2020gl087452

[27]. Haque, B. M. T., & Md. Arifur, R. (2021). ERP Modernization Outcomes in Cloud Migration: A Meta-Analysis of Performance and Total Cost of Ownership (TCO) Across Enterprise Implementations. *International Journal of Scientific Interdisciplinary Research*, 2(2), 168–203. https://doi.org/10.63125/vrz8hw42

[28]. Haque, B. M. T., & Md. Arifur, R. (2023). A Quantitative Data-Driven Evaluation of Cost Efficiency in Cloud and Distributed Computing for Machine Learning Pipelines. *American Journal of Scholarly Research and Innovation*, 2(02), 449–484. https://doi.org/10.63125/7tkcs525

[29]. Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57, 407–434. https://doi.org/10.1177/0018720814547570

[30]. Hooper, A. (2008). A multi-temporal InSAR method incorporating both persistent scatterer and small baseline approaches. *Geophysical Research Letters*, 35, L16302. https://doi.org/10.1029/2008gl034654

[31]. Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80, 91–106. https://doi.org/10.1016/j.isprsjprs.2013.03.006

[32]. Intrieri, E., Gigli, G., Mugnai, F., Fanti, R., & Casagli, N. (2012). Design and implementation of a landslide early warning system. *Engineering Geology*, 147–148, 124–136. https://doi.org/10.1016/j.enggeo.2012.07.017

[33]. Jabed Hasan, T., & Waladur, R. (2022). Advanced Cybersecurity Architectures for Resilience in U.S. Critical Infrastructure Control Networks. *Review of Applied Science and Technology*, 1(04), 146–182. https://doi.org/10.63125/5rvjav10

[34]. Jaboyedoff, M., Oppikofer, T., Abellán, A., Derron, M.-H., Loye, A., Metzger, R., & Pedrazzini, A. (2012). Use of LIDAR in landslide investigations: A review. *Natural Hazards*, 61, 5–28. https://doi.org/10.1007/s11069-010-9634-2

[35]. Jahangir, S., & Muhammad Mohiul, I. (2023). EHS Analytics for Improving Hazard Communication, Training Effectiveness, and Incident Reporting in Industrial Workplaces. *American Journal of Interdisciplinary Studies*, 4(02), 126-160. https://doi.org/10.63125/ccy4x761

[36]. Kirschbaum, D., Stanley, T., & Zhou, Y. (2015). Spatial and temporal analysis of a global landslide catalog. *Geomorphology*, 249, 4–15. https://doi.org/10.1016/j.geomorph.2015.03.016

[37]. Lu, P., Qin, Y., Li, Z., Mondini, A. C., & Casagli, N. (2019). Landslide mapping from multi-sensor data through improved change detection-based Markov random field. *Remote Sensing of Environment*, 231, 111235. https://doi.org/10.1016/j.rse.2019.111235

[38]. Luo, H., Liu, C., Wu, C., & Guo, X. (2018). Urban change detection based on Dempster–Shafer theory for multitemporal very high-resolution imagery. *Remote Sensing*, 10, 980. https://doi.org/10.3390/rs10070980

[39]. Martha, T. R., Kerle, N., Jetten, V., van Westen, C. J., & Kumar, K. V. (2010). Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods. *Geomorphology*, 116, 24–36. https://doi.org/10.1016/j.geomorph.2009.10.004

[40]. Md Ashraful, A., Md Fokhrul, A., & Md Fardaus, A. (2020). Predictive Data-Driven Models Leveraging Healthcare Big Data for Early Intervention And Long-Term Chronic Disease Management To Strengthen U.S. National Health Infrastructure. *American Journal of Interdisciplinary Studies*, 1(04), 26-54. https://doi.org/10.63125/1z7b5v06

[41]. Md Fokhrul, A., Md Ashraful, A., & Md Fardaus, A. (2021). Privacy-Preserving Security Model for Early Cancer Diagnosis, Population-Level Epidemiology, And Secure Integration into U.S. Healthcare Systems. *American Journal of Scholarly Research and Innovation*, 1(02), 01–27. https://doi.org/10.63125/q8wjee18

[42]. Md Harun-Or-Rashid, M., Mst. Shahrin, S., & Sai Praveen, K. (2023). Integration Of IOT And EDGE Computing For Low-Latency Data Analytics In Smart Cities And Iot Networks. *Journal of Sustainable Development and Policy*, 2(03), 01-33. https://doi.org/10.63125/004h7m29

[43]. Md Harun-Or-Rashid, M., & Sai Praveen, K. (2022). Data-Driven Approaches To Enhancing Human–Machine Collaboration In Remote Work Environments. *International Journal of Business and Economics Insights*, 2(3), 47-83. https://doi.org/10.63125/wt9t6w68

[44]. Md Khaled, H., & Md. Mosheur, R. (2023). Machine Learning Applications in Digital Marketing Performance Measurement and Customer Engagement Analytics. *Review of Applied Science and Technology*, 2(03), 27–66. https://doi.org/10.63125/hp9ay446

[45]. Md. Arifur, R., & Haque, B. M. T. (2022). Quantitative Benchmarking of Machine Learning Models for Risk Prediction: A Comparative Study Using AUC/F1 Metrics and Robustness Testing. *Review of Applied Science and Technology*, 1(03), 32–60. https://doi.org/10.63125/9hd4e011

[46]. Md. Towhidul, I., Alifa Majumder, N., & Mst. Shahrin, S. (2022). Predictive Analytics as A Strategic Tool For Financial Forecasting and Risk Governance In U.S. Capital Markets. *International Journal of Scientific Interdisciplinary Research*, 1(01), 238–273. https://doi.org/10.63125/2rpyze69

[47]. Mondini, A. C., Guzzetti, F., Reichenbach, P., Rossi, M., Cardinali, M., & Ardizzone, F. (2011). Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images. *Remote Sensing of Environment*, *115*, 1743–1757. https://doi.org/10.1016/j.rse.2011.03.006

[48]. Mora, O. E., Lenzano, M. G., Toth, C. K., Grejner-Brzezinska, D. A., & Fayne, J. V. (2018). Landslide change detection based on multi-temporal airborne LiDAR-derived DEMs. *Geosciences*, *8*, 23. https://doi.org/10.3390/geosciences8010023

[49]. Mostafa, K. (2023). An Empirical Evaluation of Machine Learning Techniques for Financial Fraud Detection in Transaction-Level Data. *American Journal of Interdisciplinary Studies*, *4*(04), 210-249. https://doi.org/10.63125/60amyk26

[50]. Nackaerts, K., Vaesen, K., & Muys, B. (2005). A fuzzy classification technique for change detection in remote sensing. *International Journal of Remote Sensing*, *26*, 2139–2160. https://doi.org/10.1080/01431160320000160462

[51]. Palenzuela, J. A., Marsella, M., Nardinocchi, C., Pérez, J. L., Fernández, T., Chacón, J., & Irigaray, C. (2015). Landslide detection and inventory by integrating LiDAR data in a GIS environment. *Landslides*, *12*, 1035–1050. https://doi.org/10.1007/s10346-014-0534-5

[52]. Pecoraro, G., Calvello, M., & Piciullo, L. (2019). Monitoring strategies for local landslide early warning systems. *Landslides*, *16*, 213–231. https://doi.org/10.1007/s10346-018-1068-z

[53]. Petley, D. N. (2012). Global patterns of loss of life from landslides. *Geology*, *40*, 927–930. https://doi.org/10.1130/g33217.1

[54]. Petter, S., DeLone, W., & McLean, E. R. (2008). Measuring information systems success: Models, dimensions, measures, and interrelationships. *European Journal of Information Systems*, *17*, 236–263. https://doi.org/10.1057/ejis.2008.15

[55]. Piciullo, L., Calvello, M., & Cepeda, J. M. (2018). Territorial early warning systems for rainfall-induced landslides. *Earth-Science Reviews*, *179*, 228–247. https://doi.org/10.1016/j.earscirev.2018.02.013

[56]. Piciullo, L., Tiranti, D., Pecoraro, G., Cepeda, J. M., & Calvello, M. (2020). Standards for the performance assessment of territorial landslide early warning systems. *Landslides*, *17*, 2533–2546. https://doi.org/10.1007/s10346-020-01486-4

[57]. Prokop, A., & Panholzer, H. (2009). Assessing the capability of terrestrial laser scanning for monitoring slow moving landslides. *Natural Hazards and Earth System Sciences*, *9*, 1921–1928. https://doi.org/10.5194/nhess-9-1921-2009

[58]. Qin, R., Shaker, A., & Li, J. (2018). Landslide detection from Sentinel-2 imagery based on spectral and texture features. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-3*, 1447–1454. https://doi.org/10.5194/isprs-archives-XLII-3-1447-2018

[59]. Qu, F., Qiu, H., Sun, H., Tang, M., & et al. (2021). Post-failure landslide change detection and analysis using optical satellite Sentinel-2 images. *Landslides*, *18*, 447–455. https://doi.org/10.1007/s10346-020-01498-0

[60]. Raspini, F., Bianchini, S., Ciampalini, A., Del Soldato, M., Solari, L., Novali, F., Del Conte, S., Rucci, A., Ferretti, A., & Casagli, N. (2018). Continuous, semi-automatic monitoring of ground deformation using Sentinel-1 satellites. *Scientific Reports*, *8*, 7253. https://doi.org/10.1038/s41598-018-25369-w

[61]. Ratul, D., & Subrato, S. (2022). Remote Sensing Based Integrity Assessment of Infrastructure Corridors Using Spectral Anomaly Detection and Material Degradation Signatures. *American Journal of Interdisciplinary Studies*, *3*(04), 332-364. https://doi.org/10.63125/1sdhwn89

[62]. Rauf, M. A. (2018). A needs assessment approach to english for specific purposes (ESP) based syllabus design in Bangladesh vocational and technical education (BVTE). *International Journal of Educational Best Practices*, *2*(2), 18-25.

[63]. Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. *Earth-Science Reviews*, *180*, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

[64]. Rifat, C., & Jinnat, A. (2022). Optimization Algorithms for Enhancing High Dimensional Biomedical Data Processing Efficiency. *Review of Applied Science and Technology*, *1*(04), 98–145. https://doi.org/10.63125/2zg6x055

[65]. Rifat, C., & Khairul Alam, T. (2022). Assessing The Role of Statistical Modeling Techniques in Fraud Detection Across Procurement And International Trade Systems. *American Journal of Interdisciplinary Studies*, *3*(02), 91-125. https://doi.org/10.63125/gbdq4z84

[66]. Rifat, C., & Rebeka, S. (2023). The Role of ERP-Integrated Decision Support Systems in Enhancing Efficiency and Coordination In Healthcare Logistics: A Quantitative Study. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 265–285. https://doi.org/10.63125/c7srk144

[67]. Shirvani, Z., Abdi, O., & Buchroithner, M. (2019). A synergetic analysis of Sentinel-1 and -2 for mapping historical landslides using object-oriented random forest in the Hyrcanian Forests. *Remote Sensing*, *11*, 2300. https://doi.org/10.3390/rs11192300

[68]. Shofiul Azam, T., & Md. Al Amin, K. (2023). A Hybrid Lean-Six Sigma Model with Automated Kaizen for Real-Time Quality Improvement. *American Journal of Scholarly Research and Innovation*, *2*(01), 412–442. https://doi.org/10.63125/n994vk64

[69]. Sousa, J. J., Hooper, A. J., Hanssen, R. F., Bastos, L. C., & Ruiz, A. M. (2011). Persistent Scatterer InSAR: A comparison of methodologies based on a model of temporal deformation vs. spatial correlation selection criteria. *Remote Sensing of Environment*, *115*, 2652–2663. https://doi.org/10.1016/j.rse.2011.05.021

[70]. Stumpf, A., & Kerle, N. (2011). Object-oriented mapping of landslides using Random Forests. *Remote Sensing of Environment*, *115*, 2564–2577. https://doi.org/10.1016/j.rse.2011.05.013

[71]. Tanteri, L., Stumpf, A., Bressan, E., & Rossi, M. (2017). Sentinel-2 data for landslide mapping: A case study in Italy. *Landslides*, *14*, 1865–1877. https://doi.org/10.1007/s10346-017-0861-4

[72]. Tewkesbury, A. P., Comber, A. J., Tate, N. J., Lamb, A., & Fisher, P. F. (2015). A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment*, *160*, 1–14. https://doi.org/10.1016/j.rse.2015.01.006

[73]. Tofani, V., Raspini, F., Catani, F., & Casagli, N. (2013). Persistent scatterer interferometry (PSI) technique for landslide monitoring and characterization. *Remote Sensing*, *5*, 1045–1065. https://doi.org/10.3390/rs5031045

[74]. Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Navas Traver, I., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., . . . Rostan, F. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*, *120*, 9–24. https://doi.org/10.1016/j.rse.2011.05.028

[75]. Tseng, C.-M., Lin, C.-W., Stark, C. P., Liu, J.-K., Fei, L.-Y., & Hsieh, Y.-C. (2013). Application of a multi-temporal, LiDAR-derived, digital terrain model in a landslide-volume estimation. *Earth Surface Processes and Landforms*, *38*, 1587–1601. https://doi.org/10.1002/esp.3454

[76]. Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a research agenda on interventions. *Decision Sciences*, *39*, 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.X

[77]. Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*, 157–178. https://doi.org/10.2307/41410412

[78]. Wasowski, J., & Bovenga, F. (2014). Investigating landslides and unstable slopes with satellite multi temporal interferometry: Current issues and future perspectives. *Engineering Geology*, *174*, 103–138. https://doi.org/10.1016/j.enggeo.2014.03.003

[79]. Zaman, M. A. U., Sultana, S., Raju, V., & Rauf, M. A. (2021). Factors İmpacting the Uptake of İnnovative Open and Distance Learning (ODL) Programmes in Teacher Education. *Turkish Online Journal of Qualitative Inquiry*, *12*(6).

[80]. Zhao, C., & Lu, Z. (2018). Remote sensing of landslides—A review. *Remote Sensing*, *10*, 279. https://doi.org/10.3390/rs10020279