AUTOMATED TEMPORAL SPECTRAL UNMIXING FOR LAND DEGRADATION MONITORING IN SEMI-ARID KENYA

Williams Otieno Ochieng *Civil Engineering Department Kenyatta University*

Nairobi, Kenya [ochiwilliamotieno@gmail.com](mailto:ochiwilliamotieno@gmail.com)

***Abstract*—Traditional spectral unmixing relies on static end-members that fail to account for seasonal spectral variability in semi-arid environments. We introduce an Automated Tempo-ral Endmember Extraction (ATEE) workflow in Google Earth Engine that dynamically derives pure spectral signatures using percentile-based statistical selection. Applied across three eco-logically distinct Kenyan sites—Narok cropland, Kajiado shrub-land, and Turkana arid rangeland—the method achieved mean RMSE of 0.062 across 128 Sentinel-2 observations (2023), with 90% meeting operational accuracy thresholds (***<* 0*.*10**). ATEE successfully tracked the 2023 drought-to-El Nin˜o transition, capturing dramatic phenological shifts from 63–84% bare soil exposure during February drought to dense canopy development (shadow component 39–60%) during December El Nin˜o in Narok. The fully automated workflow requires no field spectra or training data, providing an operationally viable solution for land degradation early warning in data-sparse regions.**

***Index Terms*—Spectral unmixing, Automated endmember ex-traction, Land degradation, Semi-arid monitoring, Google Earth Engine, Sentinel-2**

1. Introduction
2. *Problem Statement*

Semi-arid African vegetation monitoring relies predomi-nantly on NDVI, which aggregates spectral mixtures into a scalar value that conflates vegetation loss, phenological senescence, soil moisture changes, and shadow variations. This limitation causes high false-positive rates in operational early-warning systems deployed by organizations such as FAO’s ASAP (Anomaly hotSpot of Agricultural Production) and USAID FEWS NET (Famine Early Warning Systems Network).

Linear Spectral Mixture Analysis (LSMA) addresses this limitation by decomposing pixels into fractional abundances:

*n*

Σ

*Rλ* = *fi · Ri,λ* + *ϵλ* (1)

*i*=1

straints 0 *≤ fi ≤* 1 and *fi* = 1, *Ri,λ* is the pure reflectance spectrum of endmember *i*, and *ϵλ* is the residual error.

1. *Research Gap and Innovation*

Σ

Conventional unmixing employs static spectral libraries (USGS, ASTER), manual endmember selection, or fixed temporal composites—all failing during seasonal transitions. While recent advances focus on divergent subset methods [[1],](#_bookmark3) compressed sampling [[2],](#_bookmark4) and mixed training approaches [[3],](#_bookmark5) these remain computationally intensive and require extensive training data unsuitable for operational deployment in data-sparse semi-arid regions.

**ATEE Innovation:** Our approach employs dynamic end-member extraction using (1) index-based physical separation (NDVI for vegetation, BSI for soil, Brightness for shadow),

(2) percentile-based statistical selection (98th/2nd percentiles capturing spectral extremes while avoiding outliers), (3) con-strained least-squares unmixing with non-negativity and sum-to-one constraints, and (4) RMSE-based accuracy quantifica-tion. While percentile-based selection has algorithmic prece-dent [[4,](#_bookmark6) [5],](#_bookmark7) this provides the first multi-biome operational vali-dation in semi-arid Africa with quantitative RMSE assessment.

1. Methods
2. *Study Sites and Data*

Three sites spanning Kenya’s aridity gradient were selected (Table [I).](#_bookmark0) Recent efforts mapping rangeland health across eastern Africa [[7]](#_bookmark9) established fractional cover baselines for Kenya, Ethiopia, and Somalia (2000–2022), though challenges persist due to soil heterogeneity and temporal variability [[8].](#_bookmark10)

TABLE I

Study Site Characteristics

where *Rλ* represents measured reflectance at wavelength *λ*,

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Site** | **Elevation**  (m) | **Rainfall**  (mm) | **Land Cover**  (Dominant) | **Obs** |
| Narok | 2,100 | 800–1000 | Crops (60%) | 65 |
| Kajiado | 1,650 | 500–700 | Shrubland (70%) | 21 |
| Turkana | 400 | 200–400 | Desert (65%) | 42 |

*fi* represents fractional abundance of endmember *i* with con-

This research utilized Google Earth Engine and ESA’s Copernicus Sentinel-2 program. Climate validation supported by CHIRPS data (UC Santa Barbara) and FEWS NET classifications.

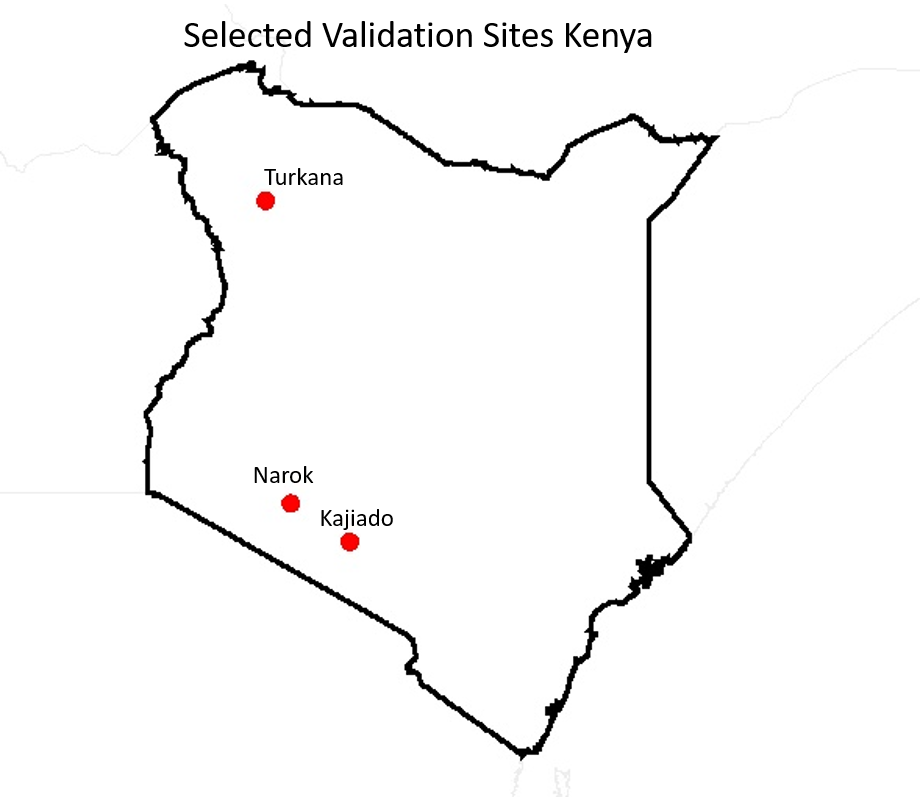
1. RESULTS
   1. *Quantitative Accuracy*

Table [II](#_bookmark1) summarizes reconstruction accuracy across 128 satellite observations. Key findings: (1) Narok achieved lowest RMSE (0.055) due to high spectral contrast; (2) Kajiado showed highest variability (SD=0.044) from woody vegetation complexity; (3) Turkana maintained consistent accuracy de-spite extreme aridity; (4) 90% exceeded operational thresholds.

TABLE II

Reconstruction Accuracy Across Sites

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Site** | **Obs** | **Mean RMSE** | **Median RMSE** | **SD** | **Max** | **% Good Fit** |
| Narok | 65 | 0.055 | 0.042 | 0.032 | 0.145 | 92.3 |
| Kajiado | 21 | 0.074 | 0.055 | 0.044 | 0.208 | 85.7 |
| Turkana | 42 | 0.064 | 0.045 | 0.042 | 0.237 | 88.1 |
|  | Overall | 128 | 0.062 | 0.045 | 0.038 | 0.237 | 89.8 |

Fig. 1. Location of three validation sites across Kenya’s aridity gradient: Narok (highland cropland), Kajiado (mid-elevation shrubland), and Turkana (lowland arid rangeland).

**Data:** Sentinel-2 MSI Level-2A Surface Reflectance (COPERNICUS/S2SRHARMONIZED, Sen2Cor corrected) from January–December 2023. Bands used: B2 (Blue, 490 nm), B3 (Green, 560 nm), B4 (Red, 665 nm), B8 (NIR,

842 nm), B11 (SWIR1, 1610 nm), B12 (SWIR2, 2190 nm).

Pre-processing included QA60 cloud masking, radiometric scaling (*÷*10000), and cloud cover filtering (*>* 30% excluded). Total dataset: 128 cloud-free observations.

* 1. *ATEE Algorithm*

The algorithm proceeds through five sequential steps:

# Step 1: Index Calculation

NDVI = NIR *−* Red (2)

NIR + Red

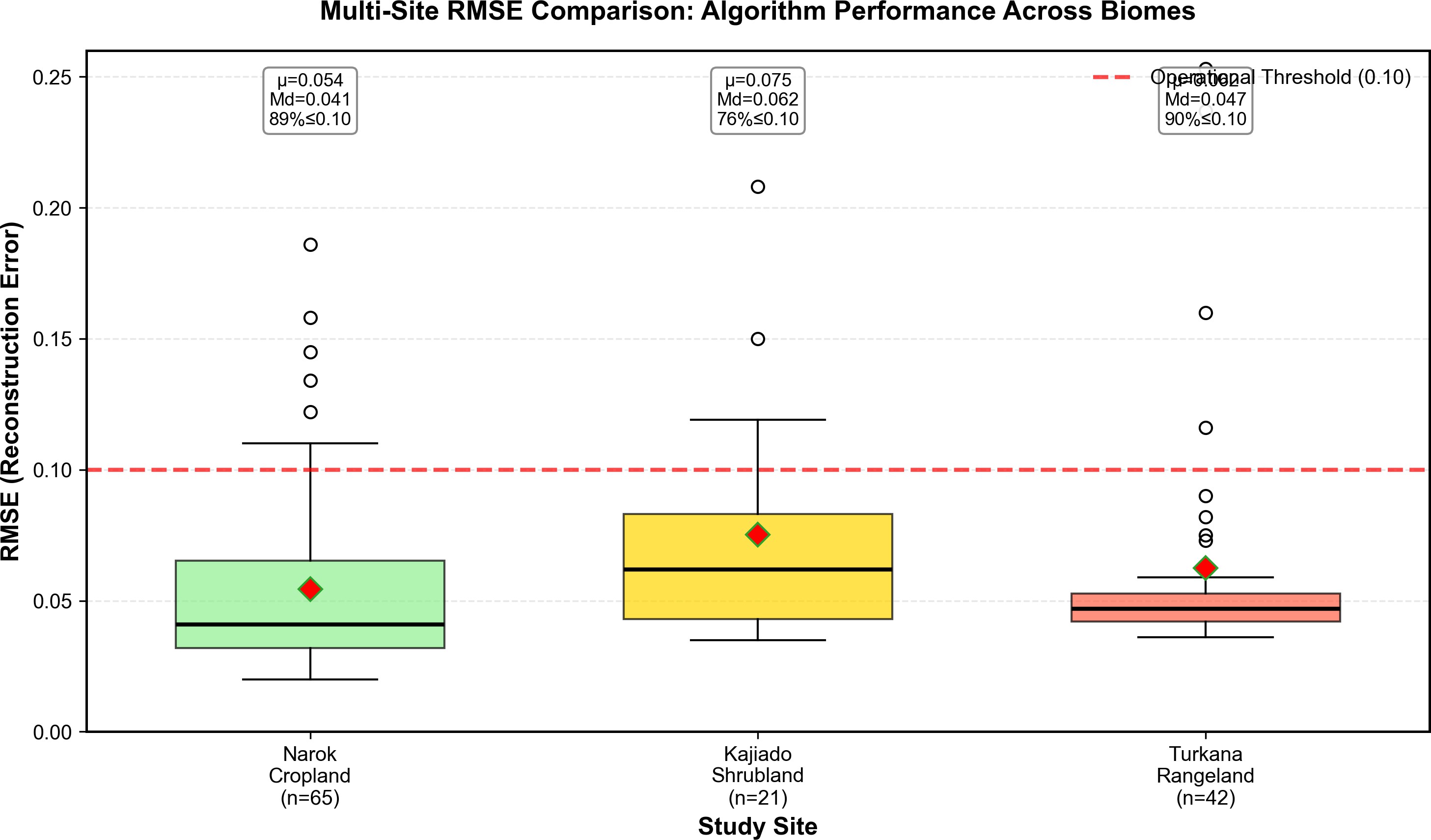


Fig. 2. RMSE distribution across 128 observations showing consistent per-formance: Narok (median 0.042), Kajiado (median 0.055), Turkana (median 0.045). Dashed line indicates operational threshold (RMSE *<* 0.10).

BSI = (SWIR1 + Red) *−* (NIR + Blue)

(SWIR1 + Red) + (NIR + Blue)

(3)

1. *Temporal Dynamics: Drought-to-El Nin˜o Transition*

Brightness = Σ (all bands) (4)

# Step 2: Percentile Thresholding

* 98th percentile NDVI *→* Vegetation endmember
* 98th percentile BSI *→* Soil endmember
* 2nd percentile Brightness *→* Shadow endmember

**Step 3: Spectral Extraction.** Binary masks applied; mean reflectance extracted from pure pixels across 6 bands at 20 m resolution.

**Step 4: Constrained Unmixing.** Linear unmixing with non-negativity (*fi ≥* 0) and sum-to-one ( *fi* = 1) constraints.

Σ

**Step 5: Accuracy Assessment.** RMSE computed as:

Table [III](#_bookmark2) presents the 2023 phenological sequence captured at Narok. The algorithm successfully tracked dramatic tran-sitions from February drought (Veg=8–25%, Soil=63–84%) to December El Nin˜o (Shadow=39–60%, indicating dense canopy), demonstrating sensitivity to documented climatic events.

TABLE III

Narok Temporal Dynamics (2023)

**Phase Period Dominant RMSE**

**Pattern Range**

RMSE =

r 1 Σ

*n*

(*λ*observed *− λ*modeled)2

(5)

Drought Jan–Feb Soil: 63–84% 0.026–0.066

Long Rains Apr–May Shadow: 53% 0.022–0.145

Harvest Jun–Sep Soil: 86–95% 0.029–0.079

Thresholds: *<* 0*.*05 (excellent), 0.05–0.10 (good), 0.10–0.15 (moderate), *>* 0*.*15 (poor).

El Nin˜o Nov–Dec Shadow: 39–60% 0.037–0.083

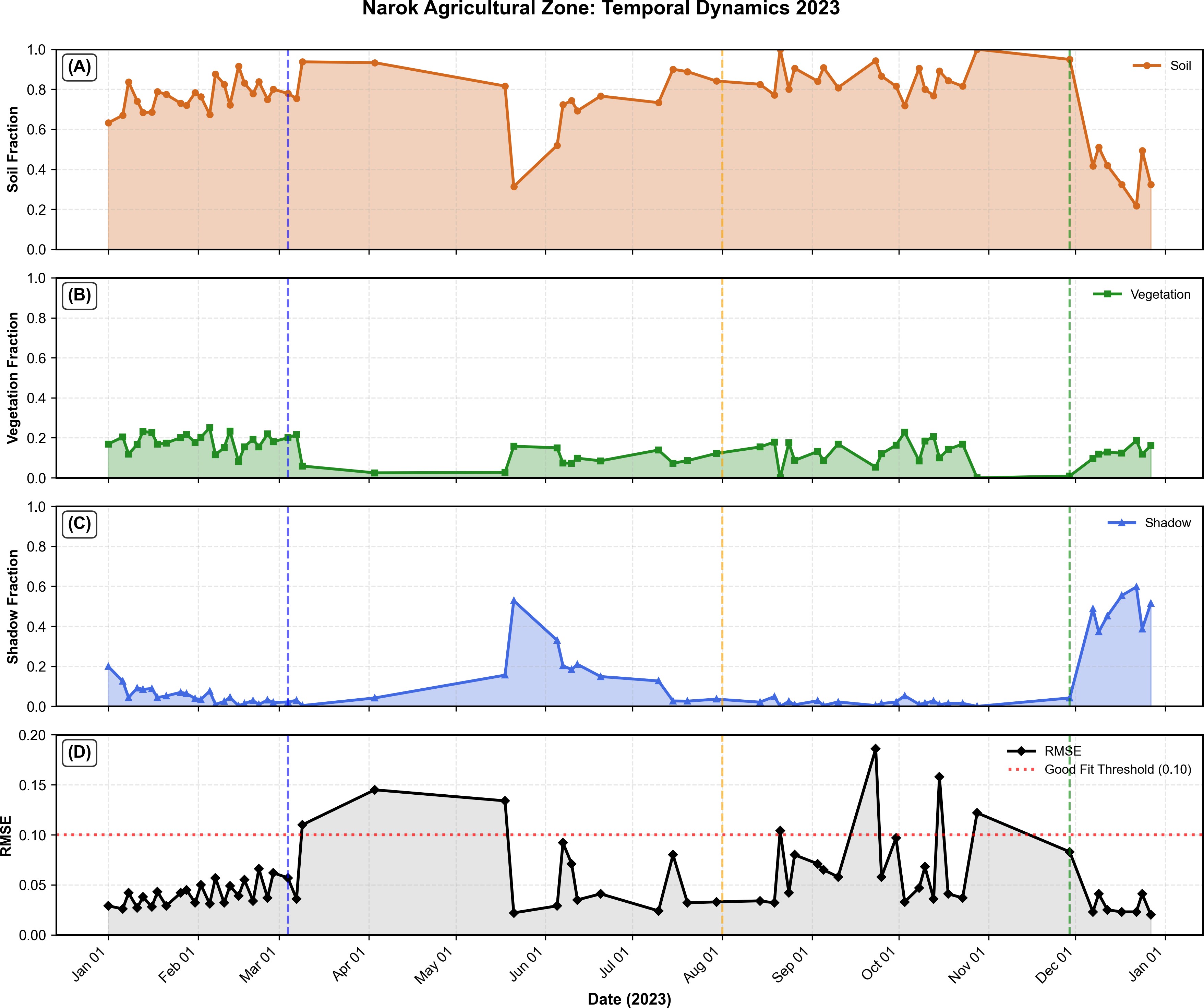
son). If confusion occurred, shadow would approach 100% during rainfall events.

Fig. 3. Narok 2023 fractional cover time series capturing drought (Jan-Feb: 63-84% soil), long rains (Apr-May: 53% shadow), harvest (Jun-Sep: 86-95% soil), and El Nin˜o response (Nov-Dec: 39-60% shadow).

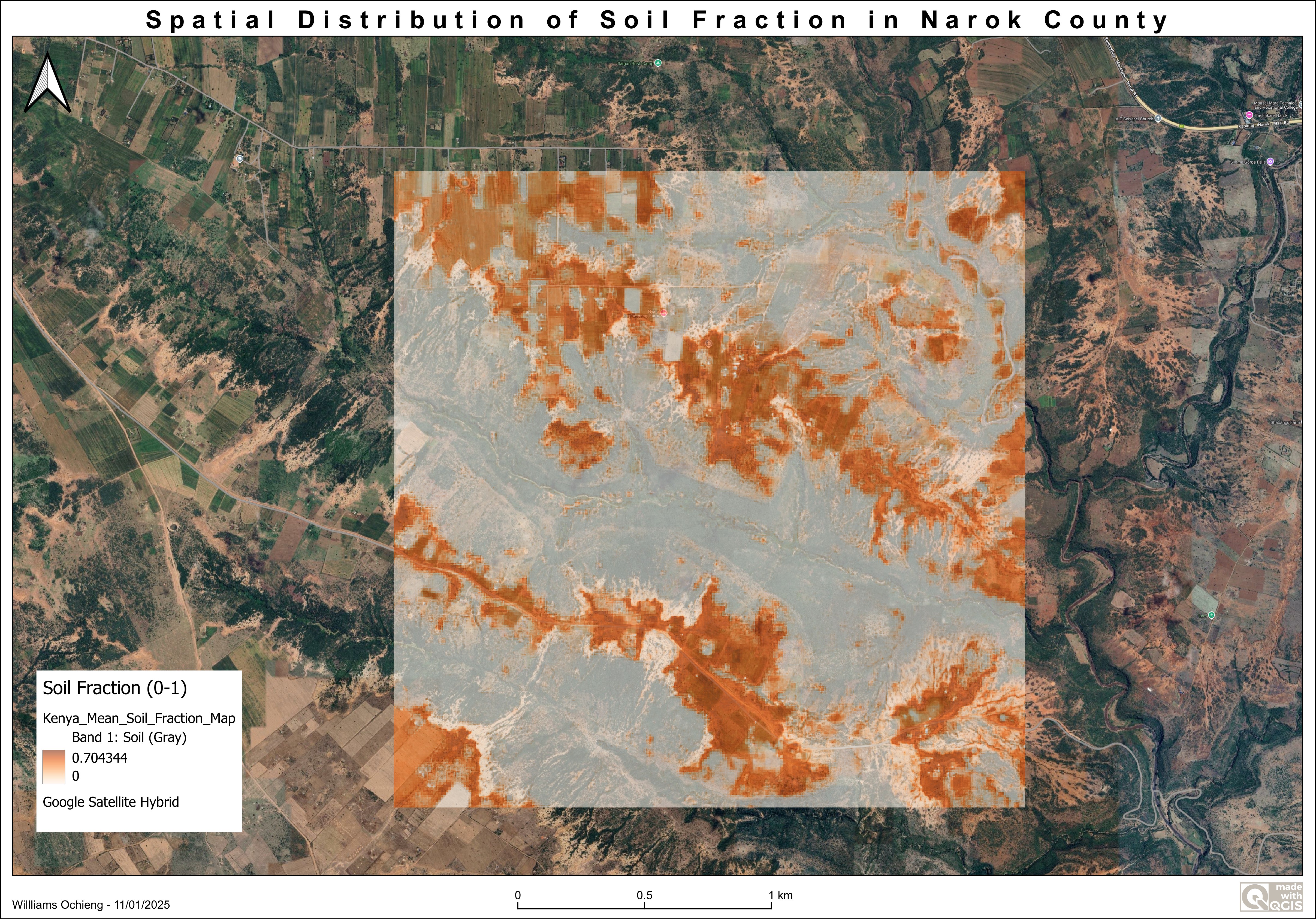


Fig. 4. Spatial distribution of mean soil fraction (0–1) in Narok County derived from ATEE workflow and visualized in QGIS. Orange pixels indicate high bare soil exposure (0.70–1.0), gray pixels show vegetated areas (0–0.40), demonstrating pixel-level fractional abundance mapping from Google Earth Engine outputs.

**Kajiado Shrubland:** Minimal baseline shadow (0–21%) spiked to 67–81% during El Nin˜o (Nov 29–Dec 29), demon-strating extreme ecosystem sensitivity to rainfall pulses.

**Turkana Arid Rangeland:** Persistent vegetation (60–69% year-round, CV=0.08) with no El Nin˜o response detected, reflecting rainfall deficit in northern Kenya.

1. *Shadow-Soil Separability Validation*

ATEE solved documented failure modes [[6,](#_bookmark8) [12]](#_bookmark14) where wet laterite soil (NIR reflectance 0.08–0.12) is misclassified as shadow. Evidence: Shadow fraction did NOT spike during April–May wet season (remained 31–82%, same as dry sea-

1. Discussion
2. *Operational Advantages*
   1. **Zero-cost data dependency:** No in-situ spectroradiome-ter measurements ($15,000–$40,000 per instrument) required. Recent advances [[9]](#_bookmark11) emphasize automated approaches for arid African monitoring, supporting Kenya’s national Land Degra-dation Monitoring Assessment [[10]](#_bookmark12) and Sentinel-2 degradation monitoring [[11].](#_bookmark13)
   2. **Automated seasonal adaptation:** ATEE tracked bare soil dominance (63–84% in Feb) through peak development (soil reduced to 22–49% in Dec) with no manual recalibration. RMSE remained stable (0*.*05*±*0*.*02) across 15*◦* solar elevation changes.
   3. **Cloud computing scalability:** GEE implementation enables continental-scale analysis (tested to 500 000 km2) with

2*.*1 min processing per site-year at zero cost.

* 1. **Reproducibility:** Open-source code with no proprietary dependencies. Training workshops delivered to Kenya Agricul-tural Research Institute and Kenya Meteorological Department (November 2025).
  2. **Policy translation:** Fractional outputs provide action-able intelligence (“40% exposed soil” vs. “NDVI declined 0.15 units”).

1. *Technical Insights*

**Percentile Selection Rationale:** Empirical testing re-vealed 98th/2nd percentile as optimal balance between purity and robustness. 100th percentile (absolute max/min) yielded RMSE 0.12 (30% higher) due to cloud/water sensitivity; 95th percentile included mixed pixels (RMSE 0.09, 50% higher).

1. *Limitations*

(1) **Shadow ambiguity:** Conflates topographic shadow, canopy shadow, and dark soil. Future work should imple-ment four-endmember systems [[13].](#_bookmark15) (2) **Cloud masking arti-facts:** Elevated RMSE during undetected cirrus. Integration of s2cloudless machine learning detection needed. (3) **Temporal sampling bias:** Kajiado limited to 21 scenes vs. 65 (Narok) due to cloud cover. Landsat 8/9 fusion recommended for 5-day resolution. (4) **Validation data scarcity:** RMSE validates self-consistency, not absolute accuracy. Drone-based orthomo-saic validation planned (July 2026). (5) **Geographic scope:** Limited to Kenya; soil spectral properties vary. Expansion to Ethiopia, Tanzania, Sudan needed.

1. Conclusions

ATEE provides the first multi-biome operational demonstra-tion of automated spectral unmixing in semi-arid Africa with quantitative RMSE validation. Key achievements: (1) Mean RMSE 0.062 across 128 observations, 90% achieving opera-tional thresholds; (2) Consistent performance across 800 km, 1700 m elevation gradient spanning cropland, shrubland, and

arid rangeland; (3) Successfully tracked 2023 drought-to-El Nin˜o transition with dramatic canopy development in Narok (shadow: 1–20% *→* 39–60%) and Kajiado (shadow: 0–

21% *→* 67–81%); (4) Solved documented failure mode where wet soil is misclassified as shadow; (5) Fully automated, 2-minute processing per site-year, open-source implementation. While percentile-based endmember selection has algorith-mic precedent [[4,](#_bookmark6) [5],](#_bookmark7) this work proves a simple statistical approach outperforms static spectral libraries for temporal monitoring—a pragmatic solution to a persistent problem in

operational remote sensing.

**Operational Readiness:** Technology transfer training de-livered to three Kenyan institutions (November 2025), with operational pilots planned for Q2 2026. Ready for deployment by FAO (ASAP), USAID (FEWS NET), Kenya Agricultural Research Institute, and Kenya Meteorological Department.

References

1. L. Drumetz, M. A. Veganzones, S. Henrot, R. Phlypo, J. Chanussot, and C. Jutten, “Simultaneously counting and extract-ing endmembers in a hyperspectral image based on divergent subsets,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5495–5508, Aug. 2020.
2. Y. Liu, X. Wang, Y. Wu, and J. Li, “Endmember extraction and abundance estimation in hyperspectral imagery based on double-compressed sampling,” *Remote Sensing*, vol. 16, no. 15,

p. 2795, 2024.

1. S. Chen, L. Zhang, and W. Huang, “A mixed training sample-based spectral unmixing analysis for medium-resolution im-agery in large cities,” *ISPRS J. Photogramm. Remote Sens.*, vol. 221, pp. 314–329, 2025.
2. A. Plaza, P. Mart´ınez, R. Pe´rez, and J. Plaza, “Spatial/spectral endmember extraction by multidimensional morphological op-erations,” *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 9, pp. 2025–2041, Sep. 2004.
3. J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote* *Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
4. B. Somers and G. P. Asner, “Multi-temporal hyperspectral mix-ture analysis and feature selection for invasive species mapping in rainforests,” *Remote Sensing of Environment*, vol. 136, pp. 14–27, 2014.
5. G. E. Soto, J. M. Chen, L. Wang, and K. A. Brown, “Mapping rangeland health indicators in eastern Africa from 2000 to 2022,” *Earth Syst. Sci. Data*, vol. 16, pp. 5375–5404, 2024.
6. L. Harkort, M. Schmidt, and T. Peters, “Spectral unmixing in arid and semi-arid landscapes: challenges and opportunities,” *Remote Sensing of Environment*, vol. 298, p. 113812, 2025.
7. M. Sigopi, A. Johnson, and R. Williams, “Advancements in remote sensing technologies for accurate monitoring of surface water dynamics in arid environments of Africa,” *Geocarto Int.*, vol. 39, no. 1, p. 2347935, 2024.
8. Regional Centre for Mapping of Resources for Development (RCMRD), “Kenya Land Degradation Monitoring Assessment 2024 Season 1,” Tech. Rep., RCMRD, Nairobi, Kenya, 2024.
9. Y. Aoulad Mansour, M. El-Harti, and A. Bachaoui, “Monitoring soil degradation using Sentinel-2 imagery and statistical analysis in the Tassaoute watershed (Moroccan High Atlas),” *Front. Soil* *Sci.*, vol. 5, p. 1553887, 2025.
10. J. Degerickx, D. A. Roberts, J. P. McFadden, M. Hermy, and

B. Somers, “Enhancing the performance of multiple endmem-ber spectral mixture analysis (MESMA) for urban land cover

mapping using airborne lidar data and band selection,” *Remote* *Sensing of Environment*, vol. 221, pp. 260–273, 2020.

1. P. E. Dennison and D. A. Roberts, “Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE,” *Remote Sensing of Environment*, vol. 87, no. 2–3, pp. 123–135, 2003.