Assessing Factors Associated with Apparent Hotspot-Cluster Spread in Upper Northern Thailand Using a Rothermel-Inspired Model

Piyachet Saengchan[1*], Kritakan Meejan[1]
and Winwan Santipapwiwatana[1]

[1]*Princess Chulabhorn Science High School Chiang Rai, 345 Moo 2, Rob Wiang, Mueang Chiang Rai, Chiang Rai 57000, Thailand*

**E-mail: Piyachet.barnaby@gmail.com*

Abstract

Satellite active-fire data provide broad spatial coverage, but hotspot coordinates represent nominal pixel centres rather than directly observed fire fronts. This study evaluated whether daily centroid displacement of VIIRS hotspot clusters (R_{obs}) could support assessment of fuel moisture, wind speed, and terrain slope in upper northern Thailand. We processed 221,418 detections from Suomi NPP, NOAA-20, and NOAA-21 collected from 1 February to 1 May 2026. After de-duplication and confidence filtering, 186,675 detections remained. Spatial DBSCAN with a 2-km radius, followed by temporal-gap splitting, produced 155 groups with at least two observed days and 1,016 complete R_{obs} records. A validity audit found that a single dominant algorithmic group (Group 0) contained 169,861 detections (92.22% of grouped data), spanned all 90 days, and covered almost the entire 332 × 334 km study box, indicating spatial percolation rather than a single fire. Under physically constrained calibration, the moisture, wind, and slope effects approached zero; the reduced model collapsed to an approximately constant prediction ($R^2 \approx 0$, RMSE = 2.096 m min^{-1}, MAPE = 225.6%). Row-level correlations with R_{obs} also opposed the expected physical directions. For centroid movements of at least 1 km, the median angular difference from downwind was 85.8° and the Rayleigh p-value was 0.829. These results do not support using daily hotspot-centroid displacement to calibrate a physical fire-spread model. The main contribution is a reproducible validity assessment showing that spatiotemporal event construction and a validated fire-front target must precede factor ranking or model calibration.

Keywords: VIIRS active fire, apparent spread rate, Rothermel model, DBSCAN, validity assessment

Introduction

Wildland fire spread emerges from coupled fuel, weather, and topographic processes. The Rothermel surface-fire model represents these processes through fuel loading, fuel-bed depth, surface-area-to-volume ratio, heat content, mineral content, particle density, moisture, wind, and slope terms (Rothermel, 1972; Andrews, 2018). It was designed to predict the advance of a flame front in a continuous fuel bed, so both the target variable and model inputs must have physical meaning.

Satellite active-fire products are valuable for regional surveillance. However, a VIIRS hotspot coordinate is the nominal center of a pixel flagged for thermal anomaly, not the measured location of a fire front; the actual fire may occupy only part of that pixel (NASA Earthdata, 2025; Schroeder et al., 2014). Daily changes in the centroid of many detections may therefore reflect sampling footprint, cloud and overpass effects, multiple ignitions, or an invalid event grouping rather than front propagation.

This project initially aimed to estimate and rank the effects of estimated fuel moisture (*M*), daily 10-m wind speed (*U*), and terrain slope (ϑ) using a five-parameter reduced equation inspired by Rothermel. The analysis was reframed to test a prior question: is daily hotspot-centroid displacement (R_{obs}) a valid response for physical fire-spread calibration? The objectives were to construct a reproducible pipeline, audit event and response validity, calibrate the reduced equation under physical constraints, and test whether centroid movement aligned with downwind direction.

Methodology

Study Area and Satellite Data

The study used a rectangular window of 97.5–100.5°E and 17.5–20.5°N, covering much of upper northern Thailand but not following provincial boundaries. FIRMS VIIRS detections from Suomi NPP, NOAA-20, and NOAA-21 were collected for 90 days from 1 February to 1 May 2026. The raw file contained 221,418 detections. Records were de-duplicated by acquisition day and coordinates rounded to three decimals, then low-confidence detections were removed, leaving 186,675 records. No forest or land-use mask was available, so the sample could include agricultural burning and other thermal sources.

Event Construction and Apparent Rate

Coordinates were projected approximately in kilometers and clustered by spatial DBSCAN with ε = 2.0 km and min_samples = 5 across the full 90-day period (Ester et al., 1996). Each spatial cluster was then split when consecutive detections were separated by more than two days. Groups with fewer than 10 detections or fewer than two observed days were excluded. Daily arithmetic centroids were calculated, and R_{obs} for each consecutive observed-day pair was defined as the Haversine distance between centroids divided by elapsed minutes. Values outside 0.05–10.00 m min^{-1} were removed. The word group is used because no independent fire perimeter or incident record confirmed that each algorithmic cluster represented one fire.

Environmental Variables and Reduced Equation

Terrain elevation came from SRTM, and slope magnitude was estimated from a five-point stencil around each pair midpoint (Farr et al., 2007). Daily mean 10-m wind speed, 2-m relative humidity, and 2-m temperature were obtained from the Open-Meteo Historical Weather API. Estimated equilibrium fuel moisture was calculated from relative humidity and temperature using the Simard formulation (Simard, 1968). These gridded daily covariates are proxies and were not matched to a satellite overpass time or converted to midflame wind.

Because the complete fuel parameters required by the Rothermel model were unavailable, the fitted expression was an empirical reduced equation rather than the full model:

$$R_{pred} = R_{max}\exp(-kM)[1 + cU^{b} + a\tan^{2}(\vartheta)].$$

Calibration used three stages. First, rows with wind speed at most 8 km/h and slope at most 5° were regressed as $\ln(R_{obs}) = \ln(R_{max}) - kM$. Second, c, b, and a were estimated from R_{obs}/R_0. Third, all five parameters were fitted simultaneously by bounded nonlinear least squares, enforcing nonnegative physical effects. Performance before and after calibration was summarized by RMSE, MAPE, and R^2 on the same data; this was an in-sample fit, not independent predictive validation.

Validity and Wind-Direction Tests

Validity checks examined group size, duration, spatial extent, and whether any group percolated across the study area. Pearson and Spearman associations and quintile means were compared with the physical directions expected for M, U, and ϑ. Fire-movement bearing was computed between consecutive centroids. Meteorological wind-from direction was converted to downwind direction, and the absolute angular difference Δ was tested using a one-sided binomial comparison of $\Delta < 45°$ against 25% and a Rayleigh approximation. Subsets included movements of at least 1 km and groups with elongation ratio ≥ 2. The tests treated rows as independent, so p-values are exploratory when many rows belong to the same group.

Results and Discussion

Data Flow and Event Validity

The pipeline produced 162 groups after the point-count filter, 155 groups with at least two observed days, 1,036 R_{obs} pairs after the rate filter, and 1,016 records with complete environmental variables. The decisive finding occurred before model calibration: a single dominant algorithmic group (Group 0) contained 169,861 detections, or 92.22% of all detections assigned to the 155 groups; it persisted for all 90 days and extended approximately 332 × 334 km across almost the entire study box. This is spatial percolation through a chain of neighboring points, not a plausible single fire incident. Daily centroids of this group therefore aggregate many spatially unrelated detections.

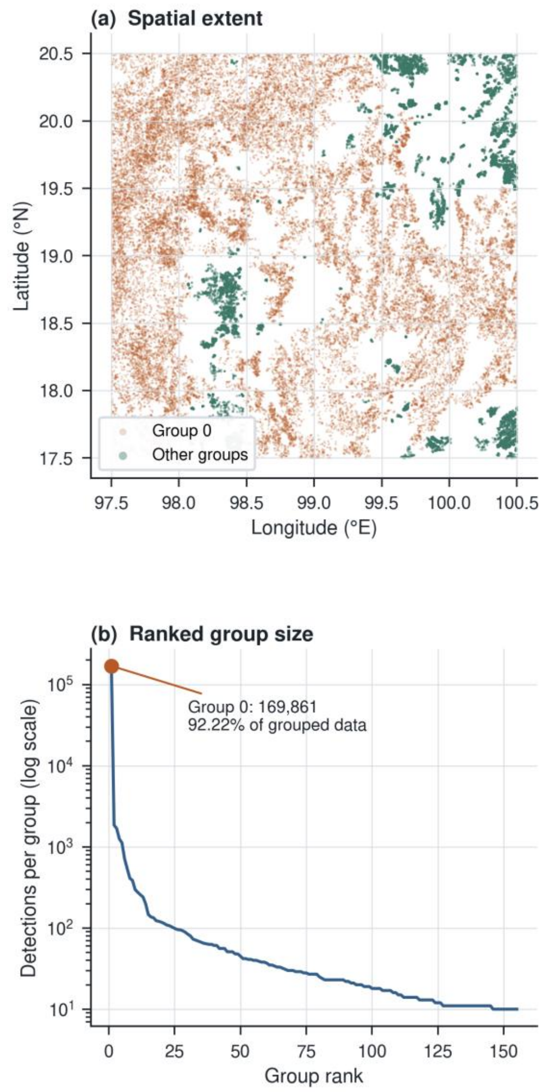

Fig. 1. Spatial percolation diagnosis. Group 0 spans almost the entire study box and is an extreme group-size outlier.

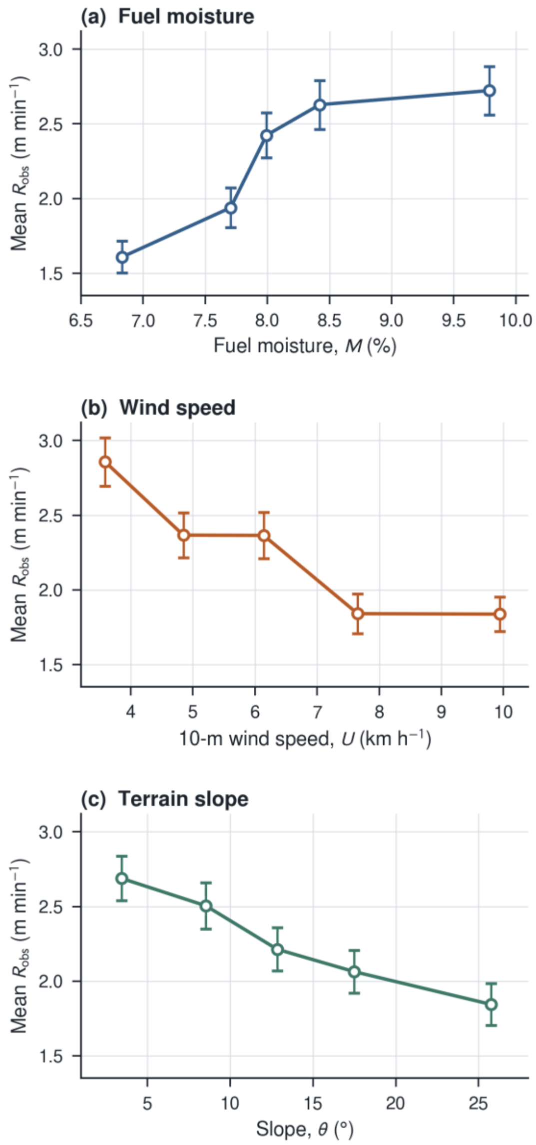

Fig. 2. Mean R_{obs} across predictor quintiles. Points are quintile means and error bars are ± SE; all three observed directions oppose the model's expected physical directions.

Observed Relationships

For the 1,016 complete records, R_{obs} had mean 2.263, standard deviation 2.097, median 1.583, and range 0.063–9.981 m min^{-1}. The row-level Pearson correlations of R_{obs} with M, U, and ϑ were +0.1955, −0.1699, and −0.1455, respectively. Thus, higher estimated moisture coincided with higher R_{obs}, while higher wind and steeper terrain coincided with lower R_{obs}. Quintile means showed the same reversed patterns (Fig. 2). R_{obs} and M both declined through the 90-day period, while U increased, indicating seasonal and spatial composition as plausible confounders. The median R_{obs} fell within the broad range reported from plot-scale experiments in deciduous forest in northern Thailand by Junpen et al. (2013), but agreement in magnitude does not establish construct validity because the measurement definitions are different.

Model Calibration

Stage 1 used 52 rows from 28 groups and estimated R_{max} = 0.407 and k = −0.226, with linear R^2 = 0.065; the negative k already opposed the assumed moisture effect. Under the final nonnegative bounds, R_{max} = 2.2651 ± 0.8218, k = 0.0001 ± 0.0263, c = 0.0000 ± 0.1538, b = 0.3000 but non-identifiable because c = 0, and a = 0.0000 ± 0.2744. The model therefore collapsed to an almost constant 2.27 m min^{-1}. RMSE improved only from 2.250 to 2.096 m min^{-1}; MAPE worsened from 169.8% to 225.6%, and R^2 changed from −0.152 to approximately zero (Fig. 3). Any sensitivity ranking derived from boundary values would be an optimization artifact, not an empirical ranking of physical drivers.

Wind Alignment

Across all 1,036 movements, median Δ was 87.6°, 26.64% had Δ < 45°, and Rayleigh p = 0.833. For the 799 movements of at least 1 km, median Δ was 85.8°, the aligned proportion was 26.91%, and Rayleigh p = 0.829. R_{obs} was nearly uncorrelated with the along-track wind component $U\cos(\Delta)$ (r = +0.0183, p = 0.609). The elongation ≥ 2 and distance ≥ 1 km subset had a small binomial excess (28.72%, p = 0.034) but no circular concentration (Rayleigh p = 0.994; Fig. 4).

This mixed result should not be reported as uniformly random, but neither does it support wind-driven centroid motion without event-level independence and validated incidents.

Table 1. Key diagnostic results

Diagnostic	Result
Group 0	169,861 detections (92.22%); 90 days; 332 × 334 km
Complete records	n = 1,016; median R_{obs} = 1.583 m min^{-1}
Correlations with R_{obs}	$r(M)$ = +0.1955; $r(U)$ = −0.1699; $r(\vartheta)$ = −0.1455
Constrained fit	k = 0.0001; c = 0; a = 0; $R^2 \approx 0$
Movement ≥ 1 km	median Δ = 85.8°; Rayleigh p = 0.829
Elongation ≥ 2; ≥ 1 km	binomial p = 0.034; Rayleigh p = 0.994

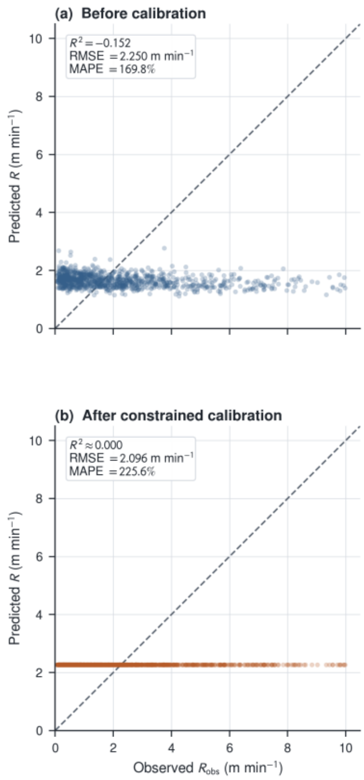

Fig. 3. Observed versus predicted apparent rates before and after physically constrained calibration. The dashed line denotes perfect agreement.

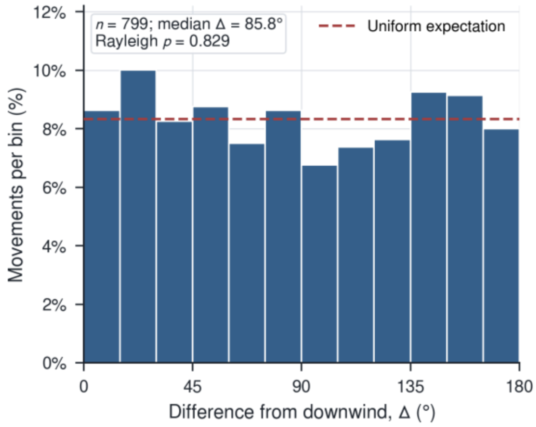

Fig. 4. Distribution of angular differences between centroid movement and downwind direction for movements ≥ 1 km. The dashed line is the uniform expectation.

Interpretation and Limitations

Independent diagnostics converge on target mismatch. The event-building algorithm joined detections across the region, so daily centroid changes need not belong to one fire, and centroid bearings did not consistently concentrate downwind. Calibration failure therefore does not show that moisture, wind, or slope are unimportant to real fire spread; it shows that the current proxy and grouping do not preserve those effects well enough to estimate them.

Sequential human ignitions could move a hotspot centroid without a continuous flame front, but the dataset lacks land-use, ignition, parcel, and road records; human activity remains a hypothesis rather than a demonstrated cause. Other limitations are daily rather than overpass-matched weather, 10-m rather than midflame wind, slope without aspect, repeated rows within groups, in-sample evaluation, and a reduced equation without complete fuel inputs.

Conclusions

The current VIIRS hotspot pipeline cannot validly rank moisture, wind, and slope as drivers of physical fire-front spread. A percolated group dominated the data, factor relationships reversed expected signs, constrained effects collapsed to zero, and centroid movement lacked consistent downwind concentration. The contribution is an explicit validity audit: event definition and response meaning must be verified before calibration. Future work should use speed-constrained spatiotemporal tracking, land-cover masks, event validation against swaths or perimeters, overpass-matched weather, and independent testing by event and area. Validated fire-front displacement—not an unverified daily centroid—should be the response variable.

Acknowledgments

The authors gratefully acknowledge NASA FIRMS for VIIRS active-fire data, the SRTM mission for elevation data, and Open-Meteo for historical weather data. The authors also thank Princess Chulabhorn Science High School Chiang Rai and the project adviser for guidance and support.

References

Andrews, P. L. (2018). The Rothermel surface fire spread model and associated developments: A comprehensive explanation. General Technical Report RMRS-GTR-371. https://doi.org/10.2737/RMRS-GTR-371

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of KDD-96, 226–231.

Farr, T. G., et al. (2007). The Shuttle Radar Topography Mission. Reviews of Geophysics, 45, RG2004. https://doi.org/10.1029/2005RG000183

Junpen, A., Garivait, S., Bonnet, S., & Pongpullponsak, A. (2013). Fire spread prediction for deciduous forest fires in Northern Thailand. ScienceAsia, 39, 535–545. https://doi.org/10.2306/scienceasia1513-1874.2013.39.535

NASA Earthdata. (2025). Collection 2 VIIRS 375 m active fire product user's guide (Version 1.2).

Rothermel, R. C. (1972). A mathematical model for predicting fire spread in wildland fuels. Research Paper INT-115. USDA Forest Service.

Schroeder, W., Oliva, P., Giglio, L., & Csiszar, I. A. (2014). The new VIIRS 375 m active fire detection data product. Remote Sensing of Environment, 143, 85–96. https://doi.org/10.1016/j.rse.2013.12.008

Simard, A. J. (1968). The moisture content of forest fuels—I: A review of the basic concepts. Information Report FF-X-14.